

THE JAMES WEBB SPACE TELESCOPE

JONATHAN P. GARDNER^{1,*}, JOHN C. MATHER¹, MARK CLAMPIN²,
RENE DOYON³, MATTHEW A. GREENHOUSE¹, HEIDI B. HAMMEL⁴,
JOHN B. HUTCHINGS⁵, PETER JAKOBSEN⁶, SIMON J. LILLY⁷, KNOX S. LONG⁸,
JONATHAN I. LUNINE⁹, MARK J. MCCAUGHREAN^{10,11}, MATT MOUNTAIN⁸,
JOHN NELLA¹², GEORGE H. RIEKE¹³, MARCIA J. RIEKE¹³,
HANS-WALTER RIX¹⁴, ERIC P. SMITH¹⁵, GEORGE SONNEBORN¹,
MASSIMO STIAVELLI⁸, H. S. STOCKMAN⁸, ROGIER A. WINDHORST¹⁶ and
GILLIAN S. WRIGHT¹⁷

¹Laboratory for Observational Cosmology, Code 665, Goddard Space Flight Center, Greenbelt, MD 20771, U.S.A.

²Laboratory for Exoplanet and Stellar Astrophysics, Code 667, Goddard Space Flight Center, Greenbelt, MD 20771, U.S.A.

³Departement de Physique, Université de Montreal, C.P. 6128 Succ. Centre-ville, Montreal, Quebec, Canada H3C 3J7

⁴Space Science Institute, 4750 Walnut Avenue, Suite 205, Boulder CO 80301, U.S.A.

⁵Herzberg Institute of Astrophysics, 5071 West Saanich Road, Victoria, British Columbia, Canada V9E 2E7

⁶Astrophysics Division, RSSD, European Space Agency, ESTEC, 2200 AG Noordwijk, The Netherlands

⁷Department of Physics, Swiss Federal Institute of Technology (ETH-Zurich), ETH Hönggerberg, CH-8093 Zurich, Switzerland

⁸Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, U.S.A.

⁹Lunar and Planetary Laboratory, The University of Arizona, Tucson, AZ 85721, U.S.A.

¹⁰Astrophysikalisches Institut Potsdam, An der Sternwarte 16, 14482 Potsdam, Germany

¹¹School of Physics, University of Exeter, Stocker Road, Exeter EX4 4QL, U.K.

¹²Northrop Grumman Space Technology, 1 Space Park, Redondo Beach, CA 90278, U.S.A.

¹³Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, U.S.A.

¹⁴Max-Planck-Institut für Astronomie, Königstuhl 17, Heidelberg D-69117, Germany

¹⁵NASA Headquarters, 300 E Street Southwest, Washington, DC 20546, U.S.A.

¹⁶Department of Physics and Astronomy, Arizona State University, Box 871504, Tempe, AZ 85287, U.S.A.

¹⁷Astronomy Technology Centre, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, U.K.

(*Author for correspondence, E-mail: jonathan.p.gardner@nasa.gov)

(Received: 8 March 2006; Accepted in final form: 15 May 2006)

Abstract. The James Webb Space Telescope (JWST) is a large (6.6 m), cold (<50 K), infrared (IR)-optimized space observatory that will be launched early in the next decade into orbit around the second Earth–Sun Lagrange point. The observatory will have four instruments: a near-IR camera, a near-IR multiobject spectrograph, and a tunable filter imager will cover the wavelength range, $0.6 < \lambda < 5.0 \mu\text{m}$, while the mid-IR instrument will do both imaging and spectroscopy from $5.0 < \lambda < 29 \mu\text{m}$.

The JWST science goals are divided into four themes. 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. 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. 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. 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. Within these themes and objectives, we have derived representative astronomical observations.

To enable these observations, JWST consists of a telescope, an instrument package, a spacecraft, and a sunshield. The telescope consists of 18 beryllium segments, some of which are deployed. The segments will be brought into optical alignment on-orbit through a process of periodic wavefront sensing and control. The instrument package contains the four science instruments and a fine guidance sensor. The spacecraft provides pointing, orbit maintenance, and communications. The sunshield provides passive thermal control. The JWST operations plan is based on that used for previous space observatories, and the majority of JWST observing time will be allocated to the international astronomical community through annual peer-reviewed proposal opportunities.

Keywords: galaxies: formation, infrared: general, planetary systems, space vehicles: instruments, stars: formation

1. Introduction

The James Webb Space Telescope (JWST; Table I) will be a large, cold, infrared(IR)-optimized space telescope designed to enable fundamental breakthroughs in our understanding of the formation and evolution of galaxies, stars, and planetary systems. It is a project led by the United States National Aeronautics and Space Administration (NASA), with major contributions from the European and Canadian Space Agencies (ESA and CSA). It will have an approximately 6.6 m diameter aperture, will be passively cooled to below 50 K, and will carry four scientific instruments: a Near-IR Camera (NIRCam), a Near-IR Spectrograph (NIRSpec), a near-IR Tunable Filter Imager (TFI), and a Mid-IR Instrument (MIRI). It is planned for launch early in the next decade on an Ariane 5 rocket to a deep space orbit around the Sun–Earth Lagrange point L2, about 1.5×10^6 km from Earth. The spacecraft will carry enough fuel for a 10 yr mission.

In this paper, we describe the scientific capabilities and planned implementation of JWST. The scientific planning is based on current theoretical understanding, interpretation of Hubble Space Telescope (HST) and Spitzer Space Telescope observations (e.g., Werner *et al.*, 2004 and additional papers in that volume), 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 plans are necessarily derived from theoretical predictions and extrapolations from known objects. In these cases, we outline the basis of the predictions and describe the observatory capabilities needed to verify them.

Additional results from HST and Spitzer, and other advances in theory and observation will continue to refine the observational plans for JWST. The scientific

TABLE I
Table of acronyms

Acronym	Definition
AB	Absolute Bolometric
AGN	Active Galactic Nuclei
AI&T	Assembly, Integration and Test
ALMA	Atacama Large Millimeter Array
AMSD	Advanced Mirror System Demonstrator
AU	Astronomical Unit
AURA	Association of Universities for Research in Astronomy
CDM	Cold Dark Matter
CMB	Cosmic Microwave Background
CSA	Canadian Space Agency
DHS	Dispersed Hartmann Sensor
DWS	Deep-Wide Survey
DSN	Deep-Space Network
EGG	Evaporating Gaseous Globule
ESA	European Space Agency
ESO	European Southern Observatory
EW	Equivalent Width
FGS	Fine Guidance Sensor
FOR	Field of Regard
FORS1	Focal Reducer and Low Dispersion Spectrograph 1
FOV	Field of View
FPA	Focal-Plane Array
FSM	Fine Steering Mirror
FUSE	Far-Ultraviolet Spectroscopic Explorer
FWHM	Full-Width Half-Maximum
GOODS	Great Observatories Origins Deep Survey
GSC	Guide-Star Catalog
GSFC	Goddard Space Flight Center
HDF	Hubble Deep Field
HgCdTe	Mercury Cadmium Telluride
HST	Hubble Space Telescope
IFU	Integral Field Unit
IGM	Intergalactic Medium
IMF	Initial Mass Function
IR	Infrared
IRAS	Infrared Astronomical Satellite
ISIM	Integrated Science Instrument Module

(Continued on next page)

TABLE I
(Continued)

Acronym	Definition
ISM	Interstellar Medium
ISOCAM	Infrared Space Observatory Camera
ISO LWS	Infrared Space Observatory Long Wavelength Spectrograph
ISO SWS	Infrared Space Observatory Short Wavelength Spectrograph
JPL	Jet Propulsion Laboratory
JWST	James Webb Space Telescope
KBO	Kuiper Belt Object
L2	Second Lagrange Point
LF	Luminosity Function
LMC	Large Magellanic Cloud
LRS	Low-Resolution Spectrograph
MIRI	Mid-Infrared Instrument
MSA	Micro-shutter Assembly
MJy/sr	Mega-Jansky per steradian
NASA	National Aeronautics and Space Agency
NGST	Northrop Grumman Space Technology
NIRCam	Near-Infrared Camera
NIRSpec	Near-Infrared Spectrograph
nJy	nano-Jansky
NTT	New Technology Telescope
OPR	Ortho-Para Ratio
OS	Operating System
OTA	Optical Telescope Assembly
OTE	Optical Telescope Element
PAH	Polycyclic Aromatic Hydrocarbon
PSF	Point-Spread Function
PM	Primary Mirror
PMSA	Primary Mirror Segment Assembly
QA	Quality Assurance
QPM	Quarter Phase Mask
QSO	Quasi-Stellar Object
rms	Root Mean Squared
RoC	Radius of Curvature
SDSS	Sloan Digital Sky Survey
SED	Spectral Energy Distribution
SFR	Star-Formation Rate
SI	Science Instrument

(Continued on next page)

TABLE I
(Continued)

Acronym	Definition
SM	Secondary Mirror
SMC	Small Magellanic Cloud
SN	Supernova
SNe	Supernovae
SOFI	Son of ISAAC (Infrared Spectrograph and Imaging Camera)
SSM	Space Support Module
STIS	Space Telescope Imaging Spectrograph
STScI	Space Telescope Science Institute
SWG	Science Working Group
S&OC	Science and Operations Center
TAC	Time Allocation Committee
TBD	To Be Determined
TFI	Tunable Filter Imager
TMA	Three Mirror Anastigmat
UDF	Ultra-Deep Field
UDS	Ultra-Deep Survey
ULIRG	Ultra-Luminous Infrared Galaxy
UV	Ultraviolet
VLA	Very Large Array
VLT	Very Large Telescope
WFPC2	Wide-Field Planetary Camera 2
WFE	Wavefront Error
WFS&C	Wavefront Sensing and Control
WMAP	Wilkinson Microwave Anisotropy Probe
YSO	Young Stellar Object

investigations we describe here define the measurement capabilities of the telescope, but they do not imply that those particular observations will be made. In this paper, we do not list all potential applications of JWST. Instead, the scientific programs we discuss here are used to determine the key parameters of the mission such as collecting area, spatial and spectral resolution, wavelength coverage, etc. A mission which provides these capabilities will support a wide variety of astrophysical investigations. JWST is a facility-class mission, so most of the observing time will be allocated to investigators from the international astronomical community through competitively-selected proposals.

The plans for JWST reflect scientific and engineering discussions, studies, and development over the last 16 yr. At a workshop held in 1989, the astronomical community began discussions of a scientific successor to HST (Bély *et al.*,

1989). In the mid-1990s, the “HST and Beyond” committee recommended that NASA build an IR-optimized telescope to extend HST discoveries to higher redshift and longer wavelength (Dressler, 1996). Initial studies of the mission, then called the Next Generation Space Telescope, were reported by Stockman *et al.* (1997). Its scientific program was given top priority by the National Academy of Sciences survey “Astronomy and Astrophysics in the New Millennium” (McKee and Taylor, 2001). In 2002, the Next Generation Space Telescope was renamed the James Webb Space Telescope in honor of the Administrator of NASA during the Apollo era (Lambright, 1995). This paper provides an update to these previous studies.

The capabilities and performance specifications for JWST given here are preliminary. The mission is currently in its detailed design phase and has not yet been given official permission to proceed beyond that stage. NASA missions receive authority, and the accompanying budget, to proceed to launch only after they pass a nonadvocate review and transition from detailed design into development phases. Hence, the specifications given in this paper are subject to change based upon technical progress of individual elements of the observatory and available budgets. JWST is expected to receive final approval and make the transition to the development phase in 2008.

The scientific objectives of JWST fall into four broad themes: The End of the Dark Ages; First Light and Reionization; The Assembly of Galaxies; The Birth of Stars and Protoplanetary Systems; and Planetary Systems and the Origins of Life.

We organize this paper as follows. The remainder of this section provides a scientific introduction to these themes. In Sections 2–5, we expand on the themes, and describe the scientific capabilities JWST will use to address them. In Section 6 we describe the planned implementation of JWST.

1.1. 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, at about 180 Myr after the Big Bang (Barkana and Loeb, 2001). According to theory, the first stars were 30–300 times as massive as the Sun and millions of times as bright, burning for only a few million years before meeting a violent end (Bromm and Larson, 2004). Each one produced either a core-collapse supernova (SN) (type II) or a black hole. The supernovae (SNe) 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, which

grew and merged to become the huge black holes now found at the centers of nearly all massive galaxies (Magorrian *et al.*, 1998). The SNe and the mini-quasars could be individually observable by JWST.

Some time after the appearance of the first sources of light, hydrogen in the intergalactic medium (IGM) was reionized. Results from the Wilkinson Microwave Anisotropy Probe (WMAP; Kogut *et al.*, 2003; Page *et al.*, 2006; Spergel *et al.*, 2006) combined with data on quasars at $z \sim 6$ from the Sloan Digital Sky Survey (SDSS; Fan *et al.*, 2002) show that this reionization had a complex history (Cen, 2003b). Although there are indications that galaxies produced the majority of the ultraviolet (UV) radiation which caused the reionization, the contribution of quasars could be significant.

JWST will address several key questions in this theme: What are the first galaxies? When and how did reionization occur? What sources caused reionization? JWST will conduct ultra-deep near-IR surveys with spectroscopic and mid-IR follow-up to find and identify the first galaxies to form in the early universe. It will determine the processes that caused reionization through spectroscopy of high-redshift quasars or galaxies, studies of the properties of galaxies during that epoch.

1.2. THE ASSEMBLY OF GALAXIES

Theory predicts that galaxies are assembled through a process of the hierarchical merging of dark matter concentrations (e.g., White and Frenk, 1991; Cole *et al.*, 1994). Small objects formed first, and were drawn together to form larger ones. This dynamical buildup 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 diverse properties of present-day galaxies, organized into the Hubble Sequence of galaxies. This galaxy assembly process is still occurring today, as the Magellanic Clouds fall into the Milky Way, and as the Andromeda Galaxy heads toward the Milky Way for a possible future collision. To date, 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.

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, and what makes them form stars. We do not know how the chemical elements are generated and redistributed through the galaxies, and whether the central black holes exert great influence over the galaxies. We do not know the global effects of violent events as small and large parts join together in collisions.

JWST will address several key questions in this theme: When and how did the Hubble Sequence form? How did the heavy elements form? What physical

processes determine galaxy properties? What are the roles of starbursts and black holes in galaxy evolution? To answer these questions, JWST will observe galaxies back to their earliest precursors, so that we can understand their growth and evolution. JWST will conduct deep-wide imaging and spectroscopic surveys of thousands of galaxies to study morphology, composition, and the effects of environment. It will conduct detailed studies of individual ultraluminous IR galaxies (ULIRGs) and active galactic nuclei (AGN) to investigate what powers these energetic sources.

1.3. 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 stars are powered by nuclear fusion, and 50 yr ago we did not know that stars are continually being formed. We still do not know the details of how stars are formed from clouds of gas and dust, why most stars form in groups, or how planetary systems form. Young stars within a star-forming region interact with each other chemically, dynamically, and radiatively 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.

We also know that a substantial fraction of stars, solar-type and later, have gas-giant planets, although the discovery of large numbers of these planets in very close orbits around their stars was a surprise. The development of a full theory of planet formation requires substantially more observational input, including observations of young circumstellar disks and older debris disks in which the presence of planets can be traced.

JWST will address several key questions in this theme: How do protostellar clouds collapse? How does environment affect star formation and vice versa? What is the initial mass function (IMF) of stars at substellar masses? How do protoplanetary systems form? How do gas and dust coalesce to form planetary systems? JWST will observe stars at all phases of their evolution, from infall onto dust-enshrouded protostars, through the formation of planetary systems, penetrating the dust to determine the physical processes that produce stars, planets, and debris disks.

1.4. PLANETARY SYSTEMS AND THE ORIGINS OF LIFE

Understanding the origin of the Earth and its ability to support life is an important objective for astronomy. Key parts of the story include understanding the formation of planetesimals, and how they combine to form larger objects. We do not know how planets reach their present orbits, and how the large planets affect the smaller ones in solar systems like our own. We want to learn 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 will address several key questions in this theme: How do planets form? How are circumstellar disks like our Solar System? How are habitable zones established? JWST will 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 will use coronagraphy to investigate extrasolar planets and debris disks, and will compare these observations to objects within our own Solar System.

2. 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.

The emergence of the first individual sources of light in the universe marks the end of the “Dark Ages” in cosmic history, a period characterized by the absence of discrete sources of light (Rees, 1997). Understanding these first sources is critical, since they greatly influenced 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 by studying these objects 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 elemental 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 (CMB) 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 that is as interesting as the emergence of first light. The epoch of reionization is the most recent global phase transition undergone by the universe after the Big Bang.

Before the epoch of 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 the stars that formed

from this process to be very massive and very hot. Historically, this primordial stellar population has been given the name of population III stars. The most overdense peaks in the perturbations emerging from recombination will collapse first. Hierarchical-clustering models predict that these overdense peaks sit in larger overdense regions that have larger mass but lower contrast. We may expect the first stars to be markers of, and possibly reside in, star clusters or even small galaxies.

At the end of their short lives, some 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 instead of from stellar remnants. Thus, in these models, the first sources of light would be powered by gravitational accretion, rather than by nuclear fusion.

Soon after the first light sources appear, both high-mass 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 IGM. AGN produce a highly energetic synchrotron power spectrum, and would reionize helium as well as hydrogen. Because observational evidence reveals that helium is reionized at a much later time, hydrogen was probably reionized by starlight at earlier epochs. 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 Wilkinson Microwave Anisotropy Probe (WMAP) of the CMB polarization (Kogut *et al.*, 2003; Page *et al.*, 2006; Spergel *et al.*, 2006) with spectra of $z > 6$ quasars found by the SDSS (Fan *et al.*, 2001, 2002) has revealed the possibility that there were two reionization epochs for hydrogen (Cen, 2003a,b). In these models, the completion of the reionization epoch that is seen at $z \sim 6$ 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 redshifts. Although the observations allow for other possibilities, in general, there is evidence that the reionization history of the universe was complex (e.g., Gnedin, 2004).

2.1. WHAT ARE THE FIRST GALAXIES?

When did the first luminous sources arise and what was their nature? What were their clustering properties?

In 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 $z \gtrsim 15$ (see Figure 1; Couchman and Rees, 1986; Haiman *et al.*, 1996; Ostriker and Gnedin, 1996; Haiman and Loeb, 1997; Abel *et al.*, 1998, 2000; Barkana and Loeb, 2001).

While we do not know whether the first sources of light are powered by nuclear energy from fusion reactions in stars, or by gravitational accretion (Haiman and Loeb, 1999), it is possible that population III stars are responsible for the reionization of hydrogen at $z > 6$ (Madau and Shull, 1996; see also Gnedin and

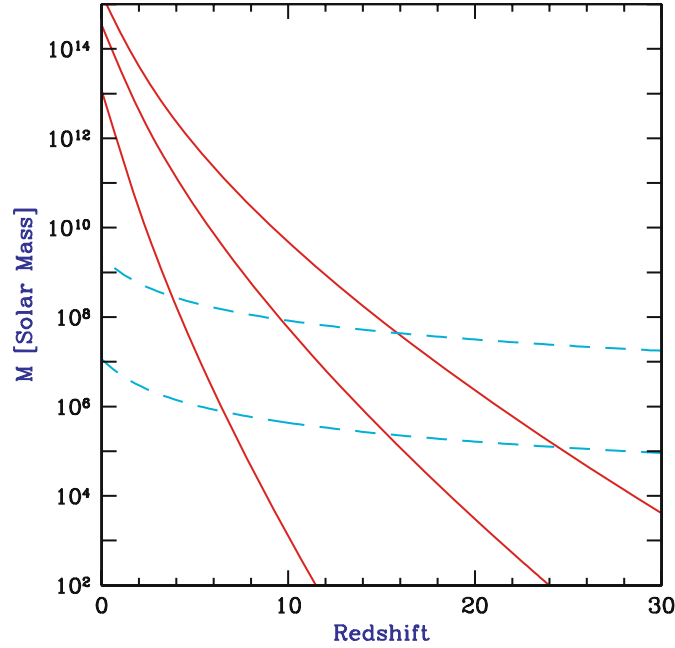


Figure 1. The mass of collapsing dark matter halos in the early universe. The *red solid* curves show the mass of collapsing halos corresponding to 1, 2, and 3σ fluctuations (in order from *bottom* to *top*.) The *blue dashed* curves show the mass corresponding to the minimum temperature required for efficient cooling with primordial atomic species only (*upper curve*) or with the addition of molecular hydrogen (*lower curve*). The intersection of these lines indicates that the epoch of formation for the first galaxies is likely to be $10 < z < 20$ (From Barkana and Loeb, 2001).

Ostriker, 1997; Haiman and Loeb, 1999; Chiu and 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 $z \sim 20$ (Cen, 2003b; Haiman and Holder, 2003).

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 process, which leads to the formation of very massive objects, with masses exceeding $100M_\odot$ (Bromm *et al.*, 1999, 2002) and possibly going as high as $500M_\odot$. The spectral energy distribution (SED) of these massive stars resembles a blackbody 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 analogs (Tumlinson and Shull, 2000).

Two 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 through the Lyman α line, resulting in a Lyman α equivalent width (EW) of 3000 Å (Bromm *et al.*, 2001). The helium lines are also strong, with the intensity of HeII λ 1640 comparable to that of H β (Tumlinson *et al.*, 2001; Panagia *et al.*, 2003).

An interesting feature of these models is that the HII region emission longward of Lyman α is dominated by a strong two-photon nebular continuum. The H α /H β ratio for these models is 3.2. Both the red continuum and the high H α /H β ratio could be 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 $10^{-2}Z_{\odot}$ (Miralda-Escudé and Rees, 1998), one finds that a combined surface brightness of about $AB = 32 \text{ mag arcsec}^{-2}$ is needed. (The AB magnitude system is defined to be $AB = 31.4 - 2.5 \log(f_{\nu})$, where f_{ν} is in nJy, Oke, 1974.) This surface brightness is about two orders of magnitude brighter than the surface brightness derived later for reionization (see Section 2.2). For reasonable luminosity functions (LFs), these sources would be either 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 seen at increasing redshifts (after properly accounting for the effects of sample completeness.)

The deepest images of the universe include the Hubble Ultra-Deep Field (UDF) in the optical (Beckwith *et al.*, 2003), which reaches $AB = 29.0 \text{ mag}$ in the I band, HST near-IR images of the UDF, which reach $AB = 28.5$ in the J and H bands (Bouwens *et al.*, 2005a), and the Spitzer Great Observatories Origins Deep Survey (Dickinson, 2004), which reaches $AB = 26.6 \text{ mag}$ at $3.6 \mu\text{m}$. Galaxies are detected in these observations at $6 < z < 7$ (e.g., Yan *et al.*, 2005) with potential candidates at even higher redshift. The rest-frame UV LF of $z \sim 6$ galaxies is intrinsically fainter than that at $z \sim 3$ (Dickinson *et al.*, 2004; Bouwens *et al.*, 2005b), showing that the global star-formation rate (SFR) is climbing. However, the detection of galaxies with stellar populations as old as 400–500 Myr at $z \sim 6.5$ (Egami *et al.*, 2005; Eyles *et al.*, 2005; Mobasher *et al.*, 2005) indicate that the first galaxies formed much earlier, perhaps in the range $7.5 < z < 13.5$. The Spitzer detections point to the importance of using mid-IR observations for galaxy age determinations through stellar population model fitting.

The number of SNe expected before reionization also strongly depends on the assumptions made about the nature of the ionizing sources. Based on relatively normal stellar populations and a metallicity of $5 \times 10^{-3} Z_{\odot}$ at the end of reionization, one arrives at an estimate of about 1 SN arcmin^{-2} at any give time (Miralda-Escudé and Rees, 1997). However, if the ionizers are very massive population III stars and the metallicity at the end of reionization is lower than $5 \times 10^{-3} Z_{\odot}$, the SN rate would be one hundred, or even one thousand times smaller. SNe with very

massive population III progenitors could be much brighter than regular type II SNe.

2.1.1. *Observations*

The direct detection of an individual population III star is not feasible even for JWST, as a $1000M_{\odot}$ star at $z = 30$ would have an AB magnitude of ~ 36 . However, JWST can detect super star clusters or dwarf galaxies made of population III stars, as well as SNe with population III progenitors. In order to directly detect these first luminous objects and to identify the redshift when they appear, we need to 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 candidate first light objects are identified, JWST will study them in detail to confirm their nature.

Number Evolution: 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 of the first sources. A strong upper limit on the number density of objects at redshifts greater than that of the most distant object observed is a likely indication that first light objects were detected.

Evolution of the SFR: In addition to using UV emission as an indicator of star formation in galaxies, one can determine the star formation rate as a function of redshift by measuring the SN rate.

Metallicity Evolution: The metallicity of first light objects should be zero, while nonzero metallicity would indicate that the object formed from gas that had already been enriched. For the brightest objects, JWST will be able to obtain spectra. At low metallicity, the ratios of oxygen lines to Balmer lines, such as $[OIII]/H\beta$, are a linear measure of metallicity.

Confirmation: 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 SED. Alternatively, identifying the age of an older stellar population sets a lower limit to the redshift of the first star formation.

JWST will need two observing programs to make these measurements: an ultra-deep imaging survey, and in-depth follow-up of candidate high-redshift sources with low-resolution spectroscopy and mid-IR photometry. The depth of the ultra-deep survey will be built up in several epochs, so that SNe can be identified for subsequent observations. A wider survey, described in the next section, could also use multiple epochs to search for SNe.

2.1.2. *Ultra-Deep Imaging Survey*

To identify a sample of high redshift galaxies, JWST will make an ultra-deep imaging survey using several broadband filters (Figure 2). The Lyman break technique will identify objects at increasing redshifts up to $z = 20$ or higher. For dwarf galaxies

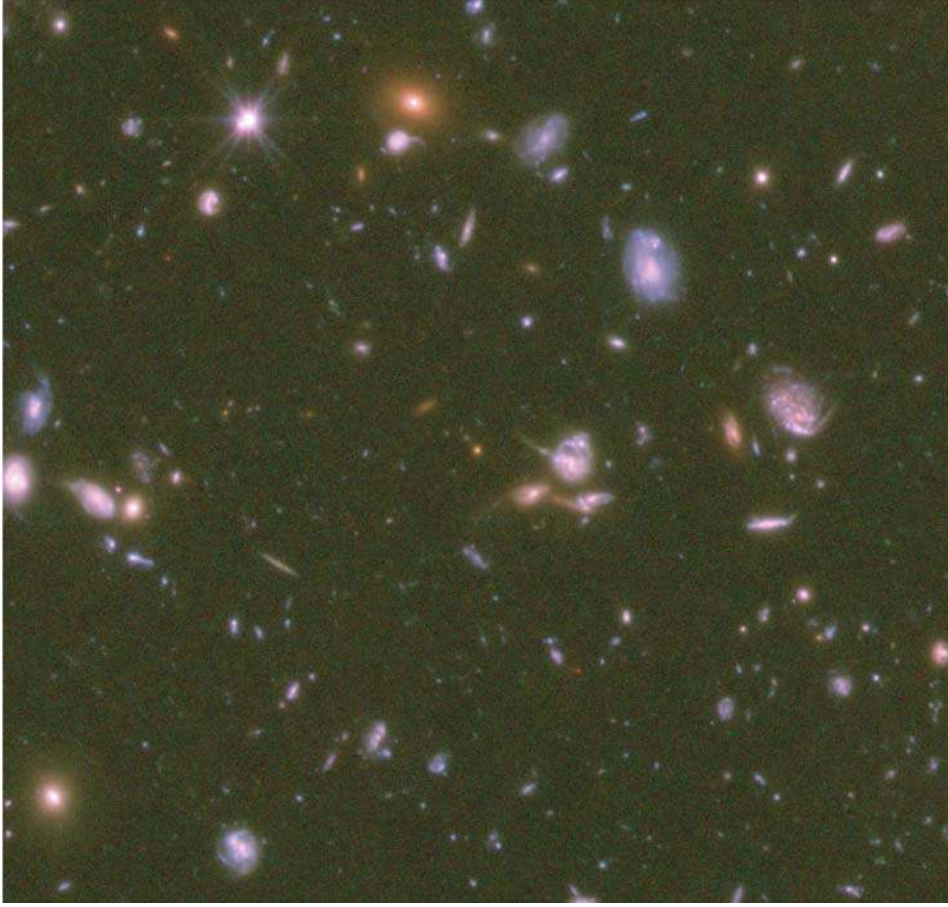


Figure 2. A simulated JWST galaxy field. The three *colors* correspond to 0.7, 0.9, and 2.0 μm . The Hubble UDF was imaged with the Advanced Camera for Surveys onboard HST. The UDF will probably be the deepest survey before JWST. For this simulation, we have taken the HST UDF and convolved it with the JWST point-spread function and scaled it to a 20 h exposure. (From Cohen *et al.*, in preparation).

with $10^6 M_{\odot}$ of zero-metallicity massive stars at $z \sim 20$, the expected AB magnitude at emitted wavelengths just longward of Lyman α is ~ 31 mag. A similar value (AB = 31 mag) is obtained by redshifting the brightest local super star clusters to $z = 20$. To enable this survey, JWST will have the sensitivity to reach AB = 31 mag in a feasible (although long) exposure time, about 100–200 hr exposure per filter, depending on the signal-to-noise ratio needed. The expected number densities are about 1 object arcmin^{-2} , thus a significant sample requires deep observations over several 10 arcmin^2 . Another driver for a large area is that it is necessary to cover a volume of at least 50 Mpc on the side in order to average over cosmic structures. Such a volume is obtained at $z = 15$ over a $\Delta z = 3$ for an area of 35 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 the derivation of $N(z)$. 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 nonionizing continuum can be calibrated to star formation rate (SFR) to yield $\text{SFR}(z)$. The required observations are deep broadband imaging in the near-IR, with mid-IR follow-up observations. 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. This survey could be done with a total of $2\text{--}4 \times 10^6$ s exposure time, depending on the number of filters required.

2.1.3. *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-IR 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 $8.4 \mu\text{m}$ for $z < 20$) will establish the absence of an older generation of stars, confirming the nature of the sources as first generation objects. 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.

This program combines deep near-IR spectroscopy and deep mid-IR imaging, using total integration times of up to $\sim 10^6$ s. It is possible that in order to achieve the required signal-to-noise ratio it will be necessary to exploit the gravitational lensing effect of a cluster of galaxies.

2.1.4. *Supernova Search in Galaxy Surveys*

Individual population III stars are too faint to be detected, but SNe can be identified up to very high redshift, since they could peak at levels brighter than $AB = 27$ mag (Figure 3; Weinmann and Lilly, 2005). Although predictions are model dependent, the brightest known SNe would be visible to $z > 30$, and it is possible that population III stars will produce bright type II SNe. Detection of a number of SNe at high redshift will require multiple visits and will yield a SN-based star formation rate. The redshift of each SN will in general be determined photometrically, although spectroscopy may be possible on the brightest sources. The expected number of population III SNe is very uncertain; predictions range between 2500 (Wise and Abel, 2003) and 50 (Mackey *et al.*, 2003) $\text{yr}^{-1} \text{deg}^{-2}$, while Weinmann and Lilly (2005) argue that the rate is $4 \text{yr}^{-1} \text{deg}^{-2}$. Large areas need to be surveyed in order to obtain a significant sample. This program requires broadband near-IR imaging. Because of time dilation, the time between visits of the search field will need to be up

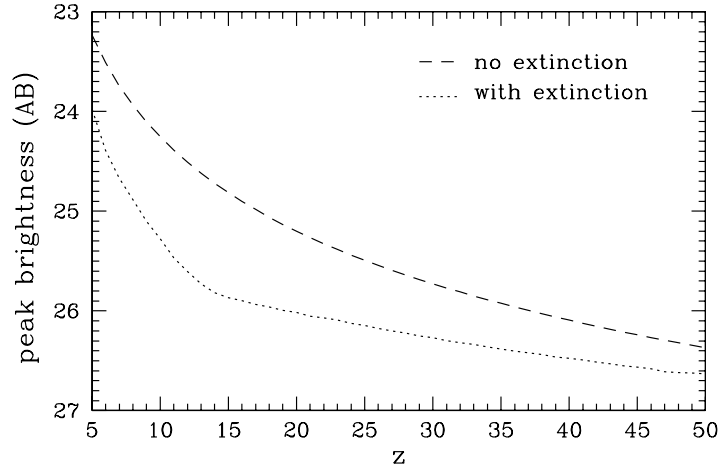


Figure 3. Predicted peak brightness of population III SNe as a function of redshift. The observed peak brightness of a $250M_{\odot}$ SN in the spectral region just longward of Lyman α is plotted assuming no extinction and with a worst case extinction. Although the number of SN expected in JWST surveys could be very low, they are bright enough and can be easily seen at any redshift. A $200M_{\odot}$ SN would be 1.7 mag fainter and a $175M_{\odot}$ SN would be 3.5 mag fainter (From Weinmann and Lilly, 2005).

to 6 months or more. 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, or with the wider surveys described in the next section. JWST will have a continuous viewing zone around each of the ecliptic poles that will enable repeated observations throughout the year.

2.2. WHEN AND HOW DID REIONIZATION OCCUR?

Was reionization a single event? What is the ionization history of the universe prior to the final reionization?

The most direct observational evidence of reionization is the detection of a Gunn–Peterson trough (Gunn and Peterson, 1965) in the spectrum of high redshift quasars. Neutral hydrogen clouds along the line of sight (the Lyman α forest) produce increasing absorption as the redshift increases. At $z \sim 5$, some signal is detected shortward of the Lyman α line, suggesting that the universe is fully ionized at $z = 5$ and that reionization was completed at still higher redshifts.

Fan *et al.* (2001, 2003, 2004) detected high redshift quasars using the SDSS, including some at $z > 6$. One quasi-stellar object (QSO) at $z = 6.28$ shows a drop in continuum flux just shortward of Lyman α by a factor 150 (see Figure 4). This is evidence that a Gunn–Peterson trough has been detected in this object (Becker *et al.*, 2001; Fan *et al.*, 2002). Other QSOs at slightly lower redshift show a much smaller

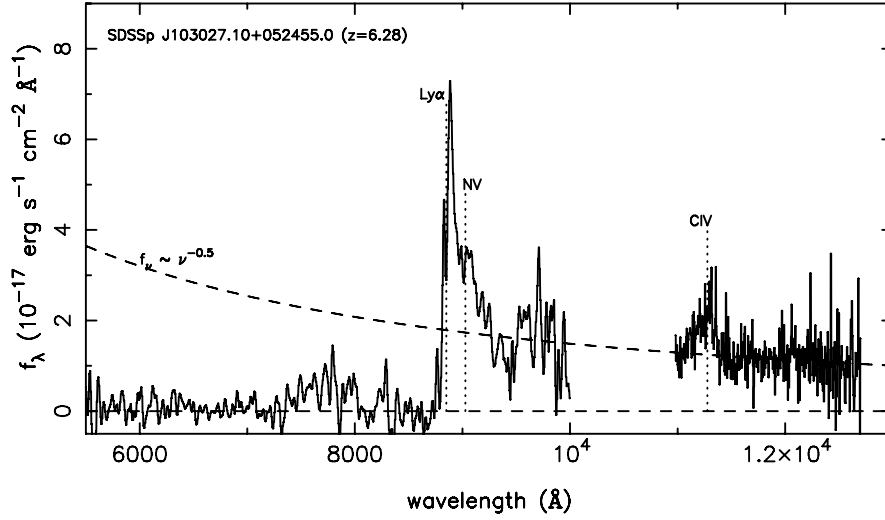


Figure 4. Spectrum of quasar SDSSp J103027.10 0552455.0 at $z = 6.28$. The absence of flux over the 300 Å region shortward of the Lyman α line is a possible indication of a Gunn–Peterson trough, indicating that the fraction of neutral hydrogen has increased substantially between $z = 5.7$ and $z = 6$, and that the universe is approaching the epoch of complete reionization at $z = 5.7$ (From Fan *et al.*, 2001).

continuum drop. Variation in the QSO properties indicates that the reionization did not occur abruptly at the same time throughout the universe. Haiman and Holder (2003) argue for an extended “percolation” period of reionization.

We cannot conclude that the reionization epoch has been determined on the basis of these few objects, particularly since even a very modest local neutral hydrogen column density could produce the observed Gunn–Peterson troughs. However, these detections open up the possibility that reionization was completed at the relatively low redshift of $z \sim 6$. There are few constraints on the number density of galaxies at redshift greater than 6. By extrapolating the LF 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 5; Yan and Windhorst, 2004b), which are at the level of a few AB = 28 galaxies per arcmin square.

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 LF of galaxies at $z > 6$. However, these facilities are unlikely to push to $z > 8$, measure the internal properties of these objects, or characterize the population of galaxies.

The correlations between the CMB temperature and polarization, as measured by WMAP, support an earlier reionization of hydrogen, giving $z_{\text{reion}} = 10.9^{+2.7}_{-2.3}$ under the assumption of a single epoch of full reionization (Figure 6; Spergel *et al.*, 2006). This may be an indication that hydrogen at least partially recombined after the first epoch of reionization, only to be reionized again at a lower redshift. In contrast to the reionization of hydrogen, the epoch of helium reionization has been

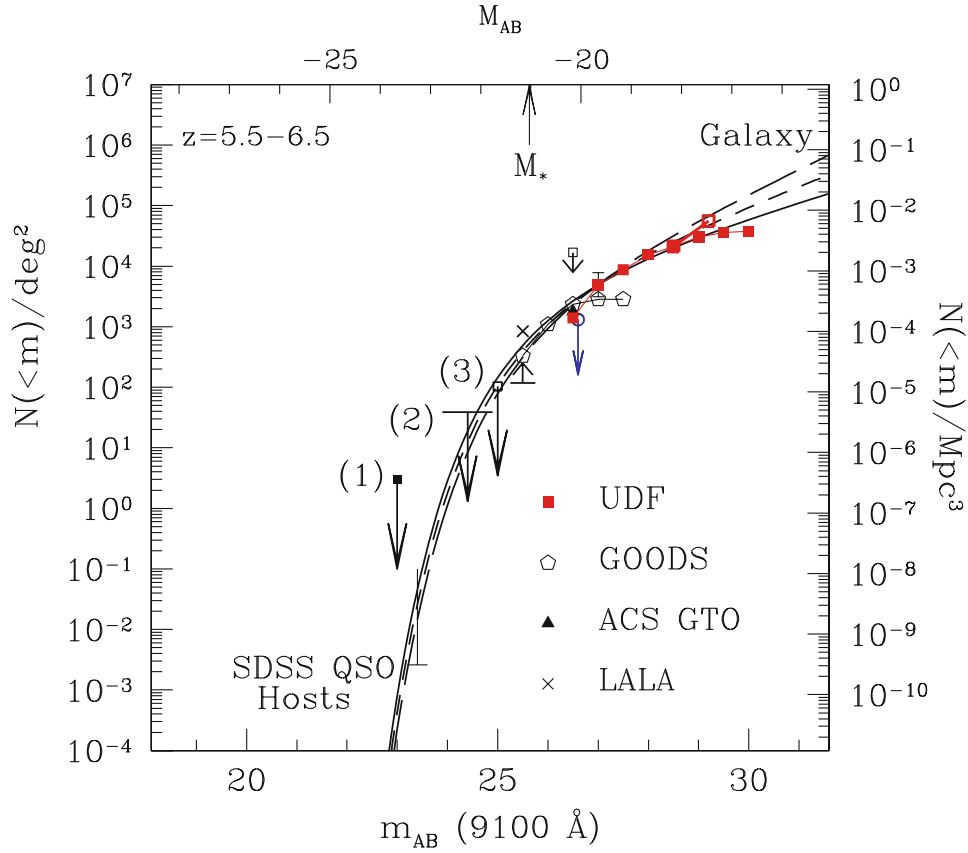


Figure 5. Cumulative galaxy counts for $z \sim 6$. Galaxy counts for $z \sim 6$ are predicted on the basis of lower-redshift measurements. In this figure, the AB magnitudes refer to a band from 9100 to 9800 Å in the observed frame. Different curves refer to different cosmologies and different normalizations of the LF. At AB = 28, one expects a few $z \sim 6$ galaxies per arcmin² but this number is uncertain by about an order of magnitude. JWST will go fainter by three magnitudes and reach completely uncharted territory (From Yan and Windhorst, 2004b).

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).

Even though one often refers to the epoch of 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. The WMAP detection of a significant Compton opacity is evidence of either an extended reionization process, or of two distinct reionization epochs (Cen, 2003a,b; Haiman and Holder, 2003; Holder *et al.*, 2003; Stiavelli *et al.*, 2004; Page *et al.*, 2006; Spergel *et al.*, 2006). Regardless of the specifics of the reionization process, inhomogeneities along the line of sight may create dispersion in optical depth shortwards of Lyman α . Moreover, only a very low residual fraction of neutral hydrogen is needed to

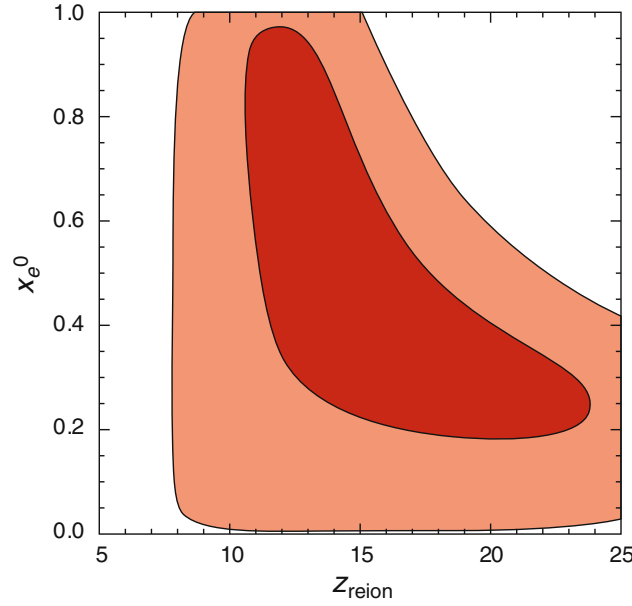


Figure 6. WMAP constraints on the reionization history. The plot shows the 68 and 95% joint two-dimensional marginalized confidence level contours for a model in which the Universe is partially reionized with an ionization fraction x_e^0 at z_{reion} , and then fully reionized at $z = 7$. The WMAP data are inconsistent with a single epoch of reionization at $z \sim 6$, and argue for a complex reionization history. (From Spergel *et al.*, 2006).

produce a Gunn–Peterson trough in the spectra of high redshift quasars. In addition, the opacity near Lyman α would be modified in the neighborhood of ionizing sources (Miralda-Escudé and Rees, 1994), in analogy to the proximity effect in QSOs (Møller and Kjaergaard, 1992).

It is possible to compute the minimum surface brightness required to reionize the universe, under the assumptions that the universe was reionized by hot population III stars, and that all UV photons can escape the system. This minimum surface brightness of ionizing sources at $z > 6$ is $AB \cong 29 \text{ mag arcmin}^{-2}$ 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 LF similar in shape to that of $z = 3$ Lyman break galaxies and with M^* not fainter than -15 mag , this implies a few sources per square arcmin with $AB = 31$ or brighter.

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 of population II stars in the mini halos of dwarf galaxies at $6 < z < 10$. This 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 nonzero 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 the 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 $6 < z < 11$. For example, the Cen (2003a) models predict an increase in the cosmic star-formation history 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 population II stars that we see today were born in dwarf galaxies, but most were not born until about $z = 8$ (consistent with the oldest ages of population II measured today of 12.8 Gyr), and it was likely the high-mass end of those population II stars that completed the epoch of reionization by $z = 6$. In WMAP cosmology (Spergel *et al.*, 2003), there was only 300 Myr 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 population II stars seen today.

2.2.1. Observations

The epoch of reionization is revealed through signatures in the Lyman α forest: a black Lyman α Gunn–Peterson trough, islands in the Lyman α forest, and appearance of a Lyman α 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 α emitters. However, a sharp transition in the Lyman α LF 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 not only increasingly more absorbed Lyman α 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, use the evolution of the ratio between Lyman α and Balmer lines.

Three observing programs are needed to firmly establish the epoch of reionization and to probe the possibility that a first reionization took place at very high redshifts. A starting sample of Lyman-break selected galaxies will be obtained from the ultra-deep observations required to identify the first light sources.

2.2.2. Lyman α Forest Diagnostics

JWST will make deep spectroscopic observations of QSOs or bright galaxies to study the Lyman α forest. High signal-to-noise, $R \sim 1000$, near-IR spectra of the brightest high-redshift QSOs or galaxies will reveal the presence of a Gunn–Peterson trough or of a Lyman α damping wing. The targets will be the brightest known high-redshift objects, perhaps from JWST surveys, or perhaps found by other means. High signal-to-noise is needed to discriminate between optical depths τ 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 Strömgren 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 β “island.” This is relevant if reionization occurs relatively abruptly. In this case, objects at redshifts between the redshift of reionization, z_{reion} , and $z = (\lambda_{\alpha}/\lambda_{\beta})(1 + z_{\text{reion}}) - 1$, will show an island of normal, finite, forest absorption between the Lyman α and the Lyman β forests. Here, λ_{α} and λ_{β} are the rest-frame wavelengths of Lyman α and Lyman β , respectively.

If there are indeed two distinct reionization epochs, the (possibly partial) recombination following the first reionization may be detectable in continuum spectra of high-redshift objects as an absorption signature in the region shortward of Lyman α .

2.2.3. *Survey for Lyman α Sources*

When the universe was still neutral, Lyman α was efficiently scattered over a large volume or absorbed by dust. The faintest Lyman α sources and those with narrow Lyman α emission will therefore not be visible before reionization. Thus, at reionization, one expects a fast evolution of the faint end of the Lyman α LF of star-forming objects (Malhotra and Rhoads, 2004). To detect a transition in the properties of Lyman α sources at the epoch of reionization, JWST will select Lyman α emitters at a variety of increasing redshifts by using the narrow-band excess technique in near-IR 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 α LF at one or two specific redshifts. Such an 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 α requires one to attempt the detection of a second line, e.g., $\text{H}\beta$. This will in general be 30 times fainter than Lyman α . An alternative method for finding Lyman α emitters would be to search in the spectral domain with spectroscopy of blank areas.

By following the properties of Lyman α emitters to the highest redshifts, we will be able to identify a period of partial recombination that would appear as a statistical brightening followed by dimming of Lyman α sources in the intermediate non-fully ionized period. This might be more sensitive than the equivalent test 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.

2.2.4. *The Ratio Between Lyman α and Balmer Lines*

If neither the metallicity nor the dust content of the universe changes abruptly at reionization, then detection of a rapid change in the Lyman α to $\text{H}\alpha$ (or $\text{H}\beta$) ratio can be used to identify the reionization epoch. By measuring the hydrogen Balmer lines in addition to Lyman α , it is possible to determine the amount by which Lyman α 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. This measurement requires $R = 1000$ spectroscopy of Lyman α , and H α or H β , and requires measurements of line intensities down to $2 \times 10^{-19} \text{ erg cm}^{-2} \text{ s}^{-1}$ at $\lambda > 3 \mu\text{m}$. This flux limit corresponds to a Lyman α intensity of $6 \times 10^{-18} \text{ erg cm}^{-2} \text{ s}^{-1}$ (detected at $1 \mu\text{m}$) and a flux ratio of 30 between Lyman α and H β .

2.3. WHAT SOURCES CAUSED REIONIZATION?

What were the sources responsible for reionization? Were they powered by nuclear fusion or gravitational accretion? How is the evolution of galaxies and black holes affected by the possibly extended period of reionization?

It is often assumed that the population III stars were responsible for the reionization of hydrogen, mainly because it is not clear how seed black holes could form in the absence of stars. This is supported by the measured LF of $z \sim 6$ quasars, which does not produce a sufficient number of ionizing photons to keep the universe ionized (Fan *et al.*, 2004; Yan and Windhorst, 2004a), and by observations of the soft X-ray background, which set limits on accretion by black holes at high redshift (Dijkstra *et al.*, 2004). However, the observed local black-hole mass – host galaxy bulge velocity dispersion relation (Ferrarese and Merritt, 2000; Gebhardt *et al.*, 2000) clearly indicates that the evolution of AGN and their galaxy hosts are closely related. Although Walter *et al.* (2004) conclude that this $M_{\text{BH}} - \sigma_{\text{bulge}}$ relation is unlikely to hold at high redshift, this result is controversial (Shields *et al.*, 2003), so determining the relative contributions of fusion and accretion to reionization and investigating the relationship between galaxies and black holes during this epoch will connect the first light sources to the processes that assembled galaxies after reionization.

Nuclear processing of only $\sim 10^{-6}$ of the baryons would be sufficient to reionize the universe (Loeb and Barkana, 2001), leading to a minimum average metallicity of the universe at reionization of $\sim 10^{-3} Z_{\odot}$. 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 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, as seen in the nearly-Solar metallicities in $z \sim 6$ QSOs (Freudling *et al.*, 2003).

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 (Madau and Rees, 2001; Heger and Woosley, 2002). Both scenarios would lead to the presence of seed black holes at the end of reionization, with implications for the formation of AGN and galaxies (Silk and Rees, 1998; Stiavelli, 1998).

Barkana and Loeb (2000) predict a distinct drop in the cosmic SFR around the reionization redshift. As the IGM is photoionized, the temperature increases, which suppresses the formation of low-mass galaxies. The LF of galaxies should show a much steeper faint-end slope before reionization than after. This may have already been seen in the Hubble UDF (Yan and Windhorst, 2004b; Bouwens *et al.*, 2004).

2.3.1. *Observations*

When 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. A combination of spectroscopic diagnostics (line shapes, line widths, and line intensity ratios) and photometry can be used to distinguish between stellar and nonstellar photoionization. The ionizing continuum can be derived indirectly by estimating its slope and intensity. This slope can be derived from the ratio between hydrogen and helium lines. The hydrogen Balmer lines can provide the intensity.

2.3.2. *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-AGN. Mid-IR photometry can help distinguish the flat UV-optical continuum of a star-bursting galaxy from the redder quasars. This program requires a combination of deep near-IR $R = 1000$ spectroscopy and mid-IR imaging.

2.3.3. *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-IR spectroscopy of very faint objects. The expected observed surface brightness of the sources responsible for reionization ranges between $AB = 27$ and $29 \text{ mag arcmin}^{-2}$, 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% escape fraction for a more extended reionization period. This program requires near-IR $R = 1000$ spectroscopy of high-redshift galaxies. It is likely to be satisfied by the same data set that was used to determine the nature of the reionizing sources.

2.3.4. *Luminosity Function of Dwarf Galaxies*

The LF of dwarf galaxies over the redshift range $6 < z < 10$ will reveal the completion of reionization and the birth of population II stars. High-mass population II stars likely completed the reionization at $6 < z < 8$, and low-mass population II stars are still visible today (Yan and Windhorst, 2004b).

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 $0.8 < \lambda < 1.3 \mu\text{m}$. JWST will measure any structural properties of these objects at wavelengths $> 2.0 \mu\text{m}$, where it will be diffraction limited. While JWST will not be diffraction limited at shorter wavelengths, 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 signal-to-noise. This program requires near-IR ultra-deep imaging as for the ultra-deep survey.

2.4. SUMMARY

Table II summarizes the measurements needed for the End of the Dark Ages theme. They include:

- *Ultra-Deep Survey (UDS)*: The UDS will be the deepest NIRCам survey, probably done in Treasury or Legacy mode. The survey will use a full set of broadband NIRCам filters, with exposure times optimized to find high-redshift objects using the drop-out technique. If done in the continuous viewing zone, the observations could be scheduled in several epochs to find high-redshift SNe.
- *In-Depth Study*: Follow-up observations of very high-redshift objects found in the ultra-deep survey will be used to investigate their nature. NIRSpec low-resolution spectroscopy will be used to search for continuum breaks and emission lines. MIRI photometry of high-redshift objects will give age estimates, relying on upper limits for very young populations.
- *Lyman α Forest Diagnostics*: Spectra of the brightest high-redshift objects will be used to look for Gunn–Peterson troughs and determine the epoch and nature of reionization.
- *Survey for Lyman- α Sources*: A narrow-band imaging program will search for Lyman α -emitting companions to known high- z objects. The properties of these objects are expected to be different before and after reionization.

TABLE II
JWST measurements for the end of the dark ages theme

Observation	Instrument	Depth, Mode	Target
Ultra-deep survey (UDS)	NIRCam	1.4 nJy at $2\ \mu\text{m}$	$10\ \text{arcmin}^2$
In-depth study	NIRSpec	23 nJy, $R \sim 100$	Galaxies in UDS area
	MIRI	23 nJy at $5.6\ \mu\text{m}$	Galaxies in UDS area
Lyman α forest diagnostics	NIRSpec	$2 \times 10^{-19}\ \text{erg cm}^{-2}\ \text{s}^{-1}$, $R \sim 1000$	Bright $z > 7$ quasar or galaxy
Survey for Lyman α sources	TFI	$2 \times 10^{-19}\ \text{erg cm}^{-2}\ \text{s}^{-1}$, $R \sim 100$	$4\ \text{arcmin}^2$ containing known high- z object
Transition in Lyman α /Balmer	NIRSpec	$2 \times 10^{-19}\ \text{erg cm}^{-2}\ \text{s}^{-1}$, $R \sim 1000$	UDS or wider survey area
Measure ionizing continuum	NIRSpec	$2 \times 10^{-19}\ \text{erg cm}^{-2}\ \text{s}^{-1}$, $R \sim 1000$	Same data as above
Ionization source nature	NIRSpec	$2 \times 10^{-19}\ \text{erg cm}^{-2}\ \text{s}^{-1}$, $R \sim 1000$	Same data as above
LF of dwarf galaxies	MIRI	23 nJy at $5.6\ \mu\text{m}$	UDS data
	NIRCam	1.4 nJy at $2\ \mu\text{m}$	

- *Transition in Lyman α /Balmer*: This program will determine the epoch of reionization through the effect on the galaxy population by measuring spectral lines of galaxies before and after reionization. It needs to see both Lyman α and $\text{H}\alpha$, in galaxies at a range of redshifts.
- *Measure Ionizing Continuum*: Ratios of the hydrogen and helium Balmer lines will reveal the hardness of the ionizing continuum.
- *Ionization Source Nature*: Near-IR line widths and mid-IR photometry will separate star-formation from AGN as source of ionizing continuum.
- *LF of Dwarf Galaxies*: The number of dwarf galaxies as a function of redshift changes as the universe is reionized.

3. 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.

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 manifest within galaxies, and the formation of galaxies represents a key link between this complexity and the relative simplicity of the early universe. On the one hand, the most basic properties of galaxies reflect the distribution of dark matter in the universe, which is believed to 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 halos produces the morphological symmetries 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.

The CDM cosmological model provides a conceptual framework for understanding the formation of galaxies through the hierarchical assembly of progressively more massive objects. However, many of the most basic questions about this process remain unanswered due to the difficulty of observing faint objects at high redshifts. The origins of the most fundamental scaling relations for galaxies are not well understood, and the CDM paradigm has not yet been tested on galactic scales. On the theoretical side, the “semianalytic” models for galaxy formation and evolution include many free parameters, while numerical gravito-hydrodynamic simulations do not yet have the resolution and dynamic range needed to simultaneously model individual star-formation events and 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 an understanding of galaxy assembly 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 reionization of the IGM 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 point-source and surface-brightness sensitivity.

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 will address the most pressing of these 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 other 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.

3.1. PREVIOUS INVESTIGATIONS

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

At $z \sim 1$, the universe appears roughly similar at optical and near-IR wavelengths 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 LF 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.*, 1996, 1997), and so on. The metallicities of the star-forming gas are close to solar. Some clear evolutionary effects are apparent, as luminous galaxies at $z \sim 1$ have signatures of vigorous star-formation activity, such as blue colors, strong emission lines, irregular morphologies. These indications are usually seen locally only in smaller galaxies, the so-called “downsizing” effect (Cowie *et al.*, 1995). In addition, the overall luminosity density in the UV, and in emission lines, is about a factor of 5 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 UV, and it is not yet clear whether there is a real absence of well-developed spiral or quiescent elliptical galaxies at this redshift; nor is it clear when such galaxies first appear (see e.g., Giavalisco *et al.*, 1996; Zepf, 1997; Abraham *et al.*, 1999; Dickinson, 2000; Franx *et al.*, 2003). Recent Spitzer results have begun to address this question by examining the population at $z \sim 2$ (Yan *et al.*, 2004; Labbé *et al.*, 2005). There are indications in these data that some galaxies have substantial old stellar populations by $z \sim 3$, but that there is not a large, previously hidden population of old galaxies (Barmby *et al.*, 2004). Spitzer 24 μm detections of extremely

red galaxies at $z \sim 2$ show two populations, merger-induced dusty starbursts and galaxies with old stellar populations (Chary *et al.*, 2004; Yan *et al.*, 2004).

The first samples selected through deep K-band imaging appear to show large numbers of red galaxies at redshifts approaching $z \sim 2$ (Abraham *et al.*, 2004; McCarthy *et al.*, 2004), although their stellar masses are sufficiently uncertain that it is not yet clear what fraction of the $z \sim 1$ population these represent. The handful of UV-selected galaxies studied in detail at $z \sim 3$ show evidence for significantly sub-solar metallicities ($Z \sim 0.3Z_{\odot}$) and for galactic winds of several hundred km s^{-1} , indicating substantial ejection of enriched material into the IGM.

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 ($z \sim 6-7$). Small samples of galaxies are known, generally found through their strong Lyman α emission (Hu *et al.*, 2002, 2004; Rhoads *et al.*, 2003) or by extensions of the Lyman break “drop-out” technique to longer wavelengths (Bouwens *et al.*, 2003; Dickinson *et al.*, 2004; Yan and Windhorst, 2004b), but the systematic and detailed study of these exceedingly faint objects is difficult, and relies on the brightest end of the LF (e.g., Bouwens *et al.*, 2005).

Results from COBE showed that the extragalactic background light has equal energy in the far-IR as in the optical and near-IR, and that the absorption and reradiation 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., Barger *et al.*, 1998; Hughes *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 to or greater than the local ULIRGs discovered by Infrared Astronomical Satellite (IRAS). Although little is known reliably about their redshifts, it is clear that they 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 Infrared Space Observatory Camera (ISOCAM) surveys, and Chary and Elbaz (2001) show that the rapid evolution required to account for these must flatten at $z = 1$, so as not to overproduce the background seen beyond $100 \mu\text{m}$ (see also Lagache *et al.*, 2003). This is broadly similar to the behavior seen in the UV, 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.

3.2. WHEN AND HOW DID THE HUBBLE SEQUENCE FORM?

Where were stars in the Hubble Sequence galaxies formed? When did luminous quiescent galaxies appear? How does this process depend on the environment?

To answer these questions, we need observations of the morphologies, stellar populations, and SFRs in a very large sample of galaxies observed in deep imaging and spectroscopic surveys. This investigation has substantial overlap with the chemical enrichment of galaxies, the measurement of masses, and the nature of the highly obscured luminous galaxies.

JWST will characterize the SFRs in individual galaxies, ideally as a function of their mass, environment, and cosmic epoch. JWST will 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 to 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 starbursts, as compared with the steady-state star formation in stable galactic disks, will reveal the modes of star formation that dominate different phases of galactic evolution, and that 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.

3.2.1. Observations

Except in objects with very high levels of dust extinction, the SFR of massive stars in a galaxy can best be estimated from measurements of the H α emission line, complemented by those of other emission lines, the UV continuum, and the bolometric luminosity at longer wavelengths. JWST should have the capability to spectroscopically measure the H α emission (5×10^{-19} erg s $^{-1}$ cm $^{-2}$) that would be produced by a SFR of only $1 M_{\odot}$ /yr at $z \sim 5$ (Kennicutt, 1999).

The existence of older stellar populations is best revealed by continuum imaging at rest wavelengths $\lambda > 0.5 \mu\text{m}$, or even at $\lambda > 1 \mu\text{m}$, if possible. Based on the Local Group and Milky Way Galaxy, the deepest near-IR imaging should be able to detect the Small Magellanic Cloud (SMC) (with $M_V = -16.2$ mag) if placed at $z \sim 5$, where it would be unresolved and have AB ~ 30.3 mag at $3.5 \mu\text{m}$.

With imaging data that span the rest-frame UV and optical with at least five filters, redshifts for essentially all galaxies above a faint flux threshold (typically $\geq 10\sigma$) can be estimated using photometric redshift techniques (e.g., Hogg *et al.*, 1998). These techniques have a typical accuracy $\delta_z/(1+z) < 0.1$, and with only a few percent of the estimates falling far from the actual redshift. Confirmation of these will be possible using either $R \sim 100$ or $R \sim 1000$ near-IR spectra, as required.

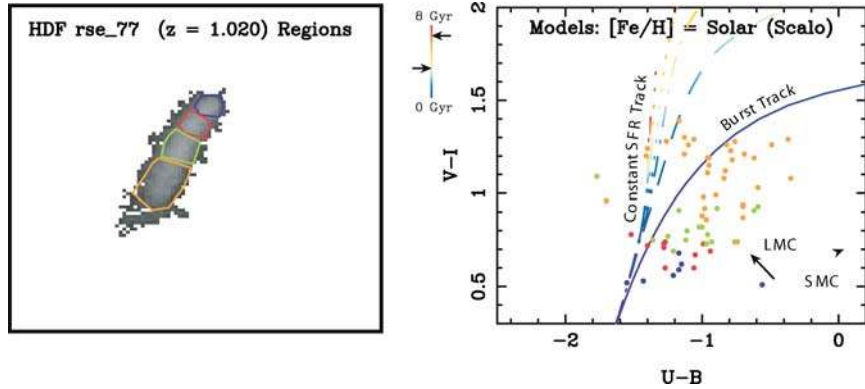


Figure 7. Galaxies in deep HST images are separated on the basis of *color* into regions with different star-formation histories. On the *left*, we show four different color regions in the galaxy image. On the *right*, these regions are placed, pixel by pixel on a *color-color* diagram and compared to model predictions to determine the ages of the regions. The *arrows* labeled LMC and SMC are the *reddening* curve from the LMCs and SMCs, respectively. (From Abraham *et al.*, 1997).

Broadband colors on their own can reveal information on the ages and reddening of individual components within a galaxy (Figure 7; Abraham *et al.*, 1997, 1999), possibly revealing the physical causes of episodes of star-formation, such as sequential star-bursts.

If the JWST galaxy surveys are conducted in the same regions as existing HST observations such as the Hubble Deep Field (HDF), Great Observatories Origins Deep Survey (GOODS), and the UDF, the data will allow a full representation of the SED of each galaxy, and of the distinct morphological components within it. Full SEDs will be obtained from the rest-frame Lyman limit at 912 Å out to a solid anchor in any older population in the rest-frame 0.6 μm 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 fully, 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 completely 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.

3.3. HOW DID THE HEAVY ELEMENTS FORM?

Where and when are the heavy elements produced? To what extent do galaxies exchange material with the IGM?

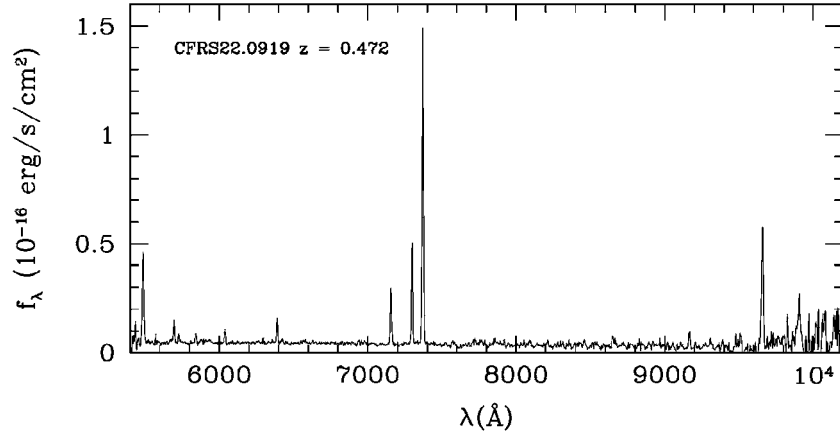


Figure 8. Spectrum of a galaxy at $z \sim 0.5$, taken as part of the Canada–France Redshift Survey, shows O[II] to H α , the lines that make up the R_{23} index (From Lilly *et al.*, 2003).

The average metallicity of the universe and of the objects in it as a function of epoch 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 (e.g., Hamann and Ferland, 1999). The lines of sight to quasars probe random regions of the universe. The study of the metallicities of material in 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 planetary and astrobiological studies, since it is likely to be 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 the R_{23} index (Pagel *et al.*, 1979), which is based on strong emission lines such as [OII] $\lambda 3727$, H β , [OIII] $\lambda\lambda 4959, 5007$, H α , [NII] $\lambda 6583$, [SII] $\lambda\lambda 6717, 6731$ (Figure 8). Such measurements require $R \sim 1000$ to separate H α and [NII].

3.3.1. Observations

JWST will be able to measure the H α , H β , 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 $3M_{\odot}/\text{yr}$. This SFR is comparable to that of the Milky Way today and requires a line sensitivity of $5 \times 10^{-19} \text{ erg s}^{-1} \text{ cm}^{-2}$ in the $2\text{--}4 \mu\text{m}$ range at 10σ . In order to assemble sufficient samples for statistical determinations, JWST will have a

multiobject spectrograph, and will be able to make these measurements with high multiplexing gain.

Extensive studies of the local universe have recently been extended to $z > 1$ 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 IR. Emission-line measurements have only been made for a fraction of galaxies at redshifts much greater than $z \sim 1$. With JWST, all of these lines will be observable in the near-IR over the redshift range $1.7 < z < 6$, enabling metallicities of individual star-forming galaxies to be measured to a precision of about 0.2 dex.

Metallicities will be determined over this range of redshifts to tie in with the observations of “first light” and the very first enrichment, and to trace 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 JWST to address the origin of the enriched IGM, the enrichment histories of different types of galaxies, and the degree to which merging or accretion of galaxies alters the metallicity of growing galaxies.

3.4. WHAT PHYSICAL PROCESSES DETERMINE GALAXY PROPERTIES?

When and how are the global scaling relations for galaxies established? Do luminous galaxies form through the hierarchical assembly of dark matter halos?

Global Scaling Relations: Despite the variety of galaxy properties observed today, galaxies obey a number of remarkably tight scaling relations between basic properties of luminosity, size, kinematics and metal enrichment. These include the Tully-Fisher relation for disk galaxies (Tully and Fisher, 1977) and the “fundamental plane,” and projections thereof, for spheroids (Faber and Jackson, 1976; Kormendy, 1977; Bender *et al.*, 1992). More 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 (Ferrarese and Merritt, 2000; Gebhardt *et al.*, 2000; Tremaine *et al.*, 2002). 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 evolution of the stellar population).

Simulations of galaxy formation have managed to reproduce the slopes, but not the normalizations of these 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.

Hierarchical Assembly: In the standard CDM paradigm, 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 (Press and Schechter, 1974; Schechter, 1976; Percival, 2001). 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.

3.4.1. Observations

Global Scaling Relations: Construction of the global scaling 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-IR. This is much superior to the lines that are accessible in the observed optical band, such as Lyman α , 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 bandheads 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 will have sufficient mid-IR 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 $AB = 24.5$ mag in the R band, and the SED of a present-day Scd galaxy, this observation needs a continuum sensitivity at $9\ \mu\text{m}$ of $AB = 21.3$ mag at 10σ per resolution element.

Hierarchical Assembly: The dark matter mass function of bound objects at very high redshifts can be uniquely measured in two ways with JWST. First, the dynamics of groups of galaxies or sub-galactic fragments can be used to determine the typical masses of halos (Zaritsky and White, 1994).

These measurements require observations of emission lines in the rest-frame optical, such as [OII] 3727, [OIII] 5007, and $H\alpha$. These are very difficult to measure from the ground when redshifted into the near-IR.

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 of galaxies 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 $z \sim 1$ (e.g., Rhodes *et al.*, 2004; Rhodes, 2004). 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 expect that HST will establish the statistical mass functions for spiral and elliptical galaxies at $z \sim 1$, but not much beyond that, because of its limited sensitivity and sampling at $\lambda > 1.6 \mu\text{m}$.

JWST will 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-IR imaging with high spatial resolution and sensitivity to achieve this greater depth. Background galaxies with a size comparable to the resolution of JWST will be measured at $\sim 20\sigma$.

The same near-IR 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 arcmin 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 beginning of dark energy dominance at $z \sim 1$.

3.5. WHAT ROLES DO STARBURSTS AND BLACK HOLES PLAY IN GALAXY EVOLUTION?

What are the redshifts and power sources of the high-redshift ULIRGs? What is the relation between the evolution of galaxies and the growth and development of black holes in their nuclei?

ULIRGs: The optical identification of high-redshift ULIRGs, found at sub-mm wavelengths, is extremely difficult with ground-based 8–10 m telescopes. The objects are very faint, and the detected images are at the confusion limits of the sub-mm telescopes. At present, none of the deepest field samples are securely identified at a level greater than 50%. Intensive efforts with ground-based telescopes will improve this before JWST's launch, but it is almost certain that many currently-known sub-mm sources will still be unidentified by the time JWST is launched. Spitzer observations have revealed the power of the mid-IR in ULIRG and AGN identification (Egami *et al.*, 2004; Frayer *et al.*, 2004; Ivison *et al.*, 2004). Analogs of known $z \sim 2$ ULIRGs, if they exist at $z > 5$, will have remained unidentified from the ground until JWST, even though they may well already be present in today's sub-mm samples. The Atacama Large Millimeter Array (ALMA) will resolve the confusion in the sub-mm, but deep imaging with JWST at $\lambda > 2 \mu\text{m}$ is needed to identify these sub-mm sources.

AGN: One of the most surprising discoveries in the study of galaxies in the last 10 yr has been that the masses of central black holes are tightly correlated with the bulge stellar population in present-day galaxies (e.g., Tremaine *et al.*, 2002). These estimates have been extended using proxy indicators to redshifts $z \sim 2$ in QSOs, and

there are indications that this correlation still holds at high-redshift (Shields *et al.*, 2003; but see also Walter *et al.*, 2004). 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 (Fan *et al.*, 2001; Freudling *et al.*, 2003). The close connection between central black holes and spheroid 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 emission-line widths in AGN broad-line regions. They show a good level of agreement and are probably correct to within a factor of 2 or 3. Many of these methods will be applicable at high redshifts with JWST. Bulge stellar populations are characterized by the bulge luminosity profiles, velocity dispersion, and overall flux, with appropriate mass-to-light ratios according to the stellar populations.

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

3.5.1. Observations

Mid-IR 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-IR spectroscopy with JWST will 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 $\text{H}\alpha$ line, which would be expected to be the strongest line in these highly obscured but vigorously star-forming galaxies, redshifts out of the ground-based K-band window at $z > 2.6$, but will be readily observable with JWST near-IR 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 $R = 1000$ spectroscopy 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-IR, high resolution spectra will allow the 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 lines and PAH features are good diagnostics of the energy sources in the center of these systems (Armus *et al.*, 2004; Soifer *et al.*, 2004). Star-bursts have strong PAH features, while AGN 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 will allow application of these same diagnostics which have proven most useful in the low-redshift ULIRG systems.

JWST will 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 (Figure 9).

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

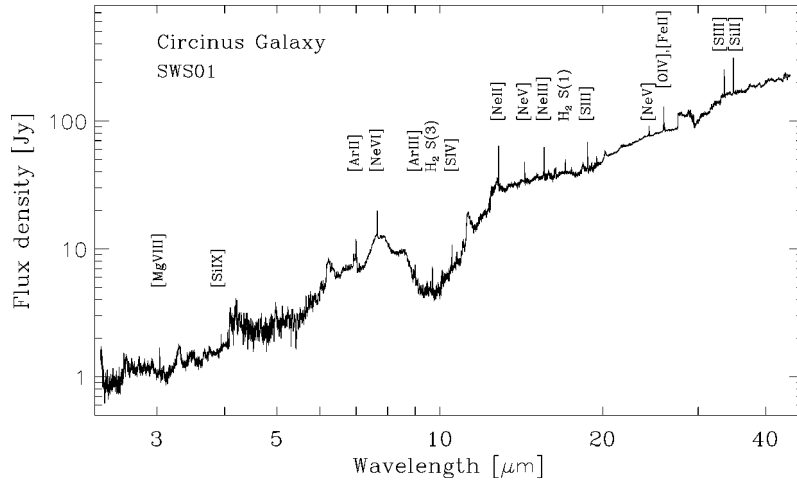


Figure 9. Mid-IR spectrum of the Circinus galaxy taken with ISO shows an abundance of emission lines useful for diagnosing the energy sources which power ULIRGs (From Moorwood *et al.*, 1996).

TABLE III
JWST measurements for the assembly of galaxies theme

Observation	Instrument	Depth, Mode	Target
Deep-wide survey (DWS)	NIRCam	3 nJy at $3.5 \mu\text{m}$	100 arcmin^2
Metallicity determination	NIRSpec	$5 \times 10^{-19} \text{ erg s}^{-1} \text{ cm}^{-2}$, $R \sim 1000$	Galaxies in DWS
Scaling relations	MIRI	$11 \mu\text{Jy}$ at $9 \mu\text{m}$, $R \sim 3000$	Lyman Break galaxies at $z \sim 3$
Obscured galaxies	NIRCam	3 nJy at $3.5 \mu\text{m}$	DWS data
	MIRI	23 nJy at $5.6 \mu\text{m}$	ULIRGs
	NIRSpec	$5 \times 10^{-19} \text{ erg s}^{-1} \text{ cm}^{-2}$, $R \sim 1000$	ULIRGs and AGN
	MIRI	$1.4 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$ at $24 \mu\text{m}$, $R \sim 2000$	ULIRGs and AGN

3.6. SUMMARY

Table III summarizes the measurements needed for the Assembly of Galaxies theme. They include:

- *Deep-Wide Survey (DWS)*: A deep-wide multi-filter NIRCam survey will be used for faint galaxy identification and morphology. Galaxies would be assigned to approximate redshift bins using photometric redshifts over the range $1 < z < 6$. The stellar populations that make up the morphological features in the galaxies would be identified on the basis of their broad-band colors. This program is designed to detect all galaxies brighter than the SMC at $z = 5$.
- *Metallicity Determination*: Follow-up multi-object spectroscopy of hundreds or thousands of galaxies in the DWS will reveal the buildup of heavy elements as galaxies are assembled. The depth is sufficient to determine R_{23} from emission line ratios for a galaxy with $\text{SFR} = 3 M_{\odot}/\text{yr}$ at $z = 5$.
- *Scaling Relations*: MIRI spectroscopy of the CO bandhead at rest wavelength $2.2 \mu\text{m}$ will measure the velocity dispersion and put the galaxy on the fundamental plane or Tully-Fisher relation. The depth is sufficient to measure the stellar velocity dispersion for an $R = 24.5 \text{ mag}$ Lyman-Break galaxy at $z = 3$. In addition, a weak lensing analysis of the DWS data will reveal the relationship between the masses of galactic halos and their star light out to $z \sim 2.5$.
- *Obscured Galaxies*: Imaging of ULIRGs will penetrate the obscuring dust to reveal the presence of merger-induced starbursts. Redshift identification of highly obscured systems can be done with $\text{H}\alpha$ out to $z \sim 6.5$. $R \sim 1000$ spectroscopy will also reveal the kinematics of merging systems. MIRI

spectroscopy will determine the energy sources that power these objects. The depth is sufficient to measure [NeVI] in a $z \sim 2$ ULIRG with Arp220 bolometric luminosity, assuming a Circinus spectrum.

4. 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 onto dust-enshrouded protostars, to the genesis of planetary systems.

The formation of stars and planets is a complex process, even in the well-developed paradigm for a single, isolated low-mass star (see, e.g., Shu *et al.*, 1987; Figure 10). We now know, however, that things are even more complicated, as stars very rarely form in isolation. The current picture of star formation starts on large scales, as molecular cloud cores cool and fragment to form highly dynamic clusters

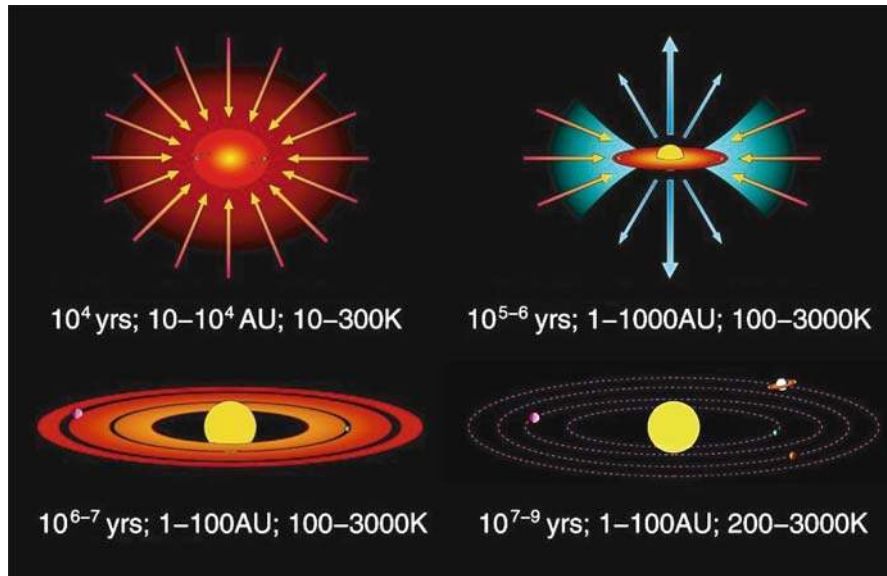


Figure 10. The formation of a single, isolated low-mass star and its planetary system. Following a deeply embedded protostellar collapse phase (Class 0 YSO; *top-left*), a circumstellar disk and collimated outflow are established, which renders the central star visible at most orientations (Class I/II; *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 (10–3000 K) involved and the associated circumstellar dust extinction implies that the bulk of the radiation seen in the early phases comes out at near-IR through millimeter wavelengths. Typical size scales (1–1000 AU, or 0.002–2 arcsec at 500 pc) imply that high spatial resolution is required for such studies (After Shu *et al.*, 1987).

of protostars, spanning the mass spectrum from O stars to planetary-mass brown dwarfs. Within those clusters, individual young sources are often encircled by disks of warm gas and dust, where material aggregates to form protoplanetary systems. These disks are the source of highly-collimated jets and outflows, which transfer energy and angular momentum from the infalling material into the surrounding medium, and clear away the remainder of the birth core. On larger scales, the intense UV 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-IR wavelengths, and at the earliest stages, the shorter wavelength emission is absorbed by the dust surrounding them. To probe these obscured regions and detect emission from gas and dust at temperatures ranging from 3000 to 100 K, imaging and spectroscopic observations from roughly 1 to 30 μm are required. High sensitivity, high spatial resolution, and a large dynamic range are needed to study the physical properties, composition, and structure of faint stellar 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.

4.1. HOW DO PROTOSTELLAR CLOUDS COLLAPSE?

How do clouds of gas and dust collapse down to the dense cores that form stars? What is the early evolution of protostars?

Clouds and Cores: Stars form in small (~ 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), in which the center forms first and the outer envelope rains down upon it. The collapse propagates at an effective sound speed of about 0.3 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, rather than differential, rotation of the cloud core (Mouschovias and Paleologou, 1981; Crutcher *et al.*, 1994). Furthermore, cores may be externally pressure-confined (Alves *et al.*, 2001), or may be altogether more chaotic and dynamic structures formed in the intersections of fractal clouds (Bate *et al.*, 2003; MacLow and Klessen, 2004). These different models predict different density distributions for star-forming cores. 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-mm telescopes (about 10 arcsec) is also a problem that will be partly alleviated by new sub-mm interferometers such as the Sub-Millimeter Array and ALMA.

An alternative technique involves mapping the extinction seen along various lines of sight through a cloud core, by measuring the near-IR colors of discrete background field stars shining through it. The extinction can be directly related to the dust column density, so a two-dimensional projection of the core density profile can be deduced, assuming a fixed gas-to-dust ratio. Used on ground-based telescopes, a typical maximum depth of $A_B = 22$ mag in the K-band can be reached with seeing-limited resolution, providing a resolution in the resulting extinction map 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 Figures 11 and 12). The much greater sensitivity and substantially improved spatial resolution of the JWST will yield much more detailed profiles through greater column density.

Another approach maps the attenuation of the diffuse mid-IR 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 μm 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 values of up to $A_V \sim 50$ mag. Again, JWST's mid-IR spatial resolution and high sensitivity will enable mapping through much higher extinctions and with greater fidelity.

Protostars: 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 *et al.*, 1993), is deeply embedded in, and obscured by, the massive envelope from which it is accreting, and its SED is dominated by this cold (~ 20 K) material. As a result, these young ($\sim 10^4$ yr) sources emit the bulk of their flux at millimeter and sub-mm wavelengths, and are generally undetected at shorter wavelengths to date (Figure 13).

Detecting and studying the 10–20 μm emission from protostars is important. Radiative transfer models (Wolfire and Cassinelli, 1986, 1987; André *et al.*, 1993; Boss and Yorke, 1995) predict that there should be a warm ‘shoulder’ in the mid-IR in the otherwise single ~ 20 K blackbody SED, and that protostars should be

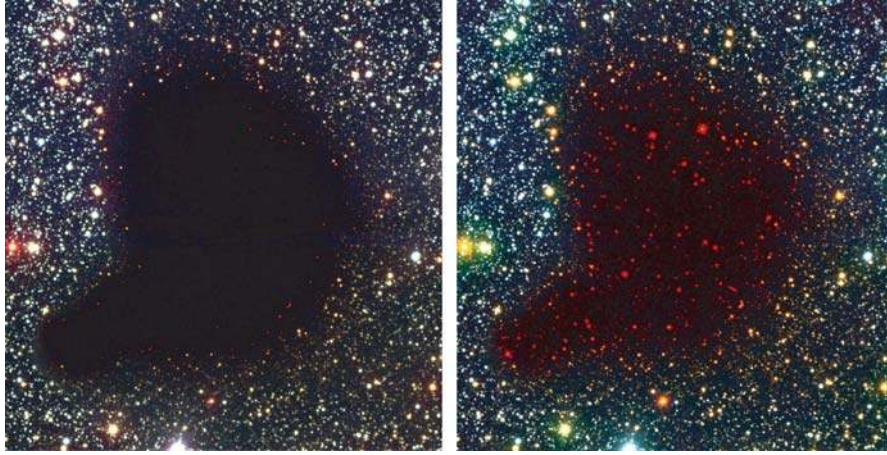


Figure 11. The low-mass dark cloud Barnard 68 imaged at optical and near-IR wavelengths using FORS1 on the ESO VLT and SOFI on the ESO NTT. The *left panel* shows a color composite of optical B, V, and I images, while the *right panel* shows a composite of the B and I images with a near-IR K_S image. The images cover 4.9×4.9 arcmin, or 0.18×0.18 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 IR wavelengths, they appear *red*. Using other IR images and measuring the extinction on a star-by-star basis, the dust column density profile of Barnard 68 and similar dark clouds could be measured (From Alves *et al.*, 2001).

roughly 1000 times brighter than the blackbody at some wavelengths, since 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-IR 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, $7.5 \mu\text{m}$) between ice and silicate absorption features, warm material (~ 700 K) was observed through effective extinctions of $A_V \sim 80\text{--}100$ mag, and with flux coming from within 4 AU of the accreting protostars. More detailed observations of this kind are required for a much 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.

The dynamics of the protostellar collapse can be diagnosed through imaging and spectroscopy of shocks, which form as material accretes onto the inner envelope and disk, and as the vertical velocity component is dissipated. The models of Yorke and Bodenheimer (1999) predict at least two shock fronts at 500–1000 AU from the protostar, with positions changing as a function of evolution in the system,

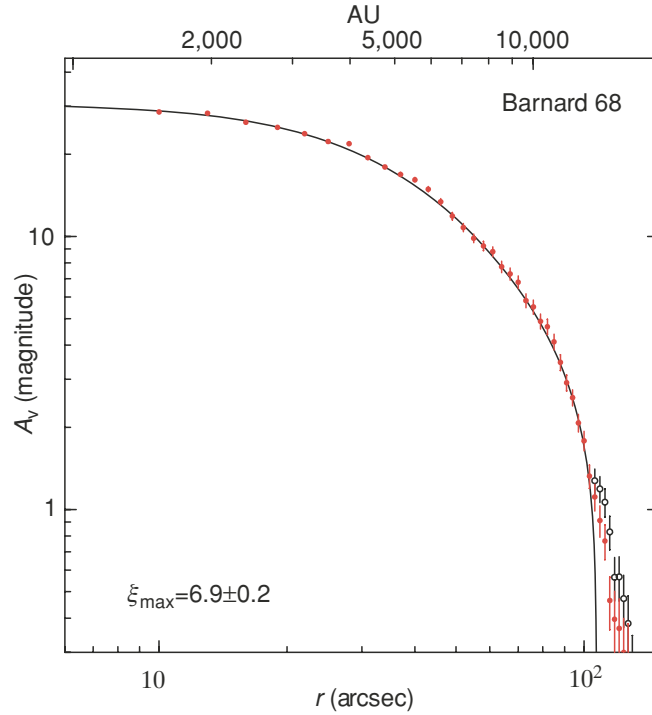


Figure 12. Azimuthally averaged radial dust column density profile for Barnard 68. The *red circles* show the averaged profile of a subsample of data that does not include the southeast prominence of the cloud (see Figure 11), while the *open circles* include the prominence. The *solid line* represents 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).

i.e., moving further from the source as the disk grows, but disappearing once the accretion terminates.

Finally, it is now clear that the majority of stars form in binaries or high-order multiples, but their origin is not well understood. While some theoretical predictions of fragmentation models are supported indirectly by statistical studies of evolved binary systems at optical and near-IR 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 therefore needed 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

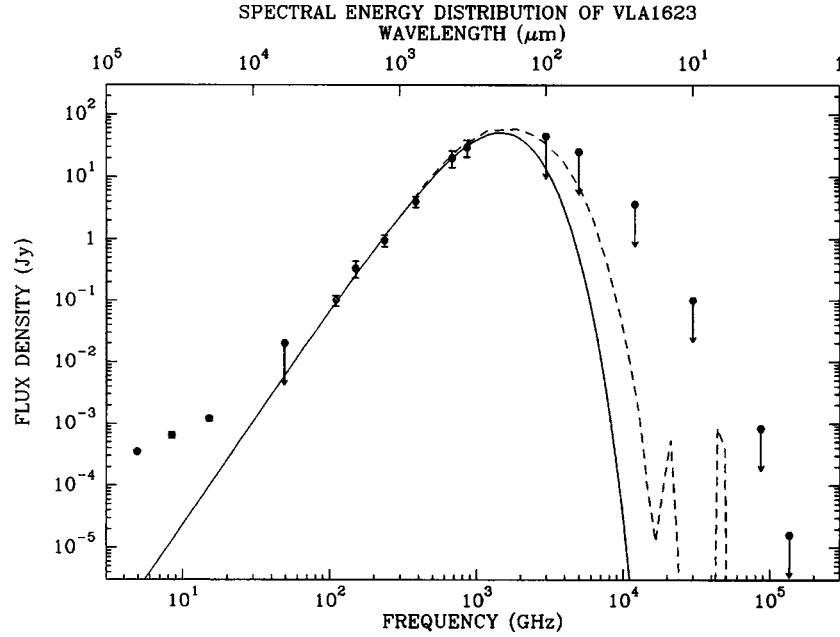


Figure 13. The SED of the Class 0 protostar, VLA 1623, in Ophiuchus. Despite a nominal black-body temperature of only ~ 20 K, radiative transfer models predict significant mid-IR flux as emission from the warm core is scattered off dust grains in the inner envelope (From André *et al.*, 1993).

binary fractions and separations, and how these properties change as stars evolve through the pre-main-sequence phases.

4.1.1. Observations

Clouds and Cores: 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 carry out these observations, at least a factor of 2 increase in extinction penetration is required relative to present ground-based observations, i.e., up to $A_V \sim 120$ mag, which requires an additional 7 mag of sensitivity at $2 \mu\text{m}$.

To achieve this, JWST will reach a 10σ 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 ~ 1 – 2 arcsec resolution, i.e., 200–500 AU for clouds at a few hundred parsecs distance. 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 2–4 arcmin is required.

The centers of cores and Class 0 objects have even more extinction, and thus mid-IR observations using the extended background emission technique are required.

A typical mid-IR background flux of 10 MJy sr^{-1} yields an unattenuated surface brightness of $240 \mu\text{Jy arcsec}^{-2}$.

To observe this surface brightness through $A_V \sim 300 \text{ mag}$ of extinction (or $A_{7\mu\text{m}} \sim 7 \text{ mag}$, assuming a standard extinction law) requires a 10σ surface-brightness sensitivity of $1 \mu\text{Jy arcsec}^{-2}$ over the $6.7\text{--}7.7 \mu\text{m}$ region where the background is bright and the dust extinction low; binning of the pixels to 1 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-IR, JWST will have a mid-IR field-of-view larger than 1 arcmin (i.e., 0.15 pc at 500 pc).

Protostars: 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 $5\text{--}7 \mu\text{m}$ extinction windows are required. To detect such young protostars and protostellar cores at a distance of $\sim 150 \text{ pc}$ (the distance of the nearest star-forming regions like Taurus and Chamaeleon), JWST will reach sensitivities of $1 \mu\text{Jy}$ at $6 \mu\text{m}$ and $10 \mu\text{Jy}$ at $15 \mu\text{m}$ (Figure 13). Going out to a distance of 500 pc , and thus encompassing a much wider range of star-forming environments, 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 ($< 1 \text{ arcsec}$ Full-Width Half-Maximum (FWHM)) is required at mid-IR wavelengths, as Class 0 protostars should be slightly extended due to dust scattering, with diameters of order 1000 AU , i.e., 6 to 2 arcsec in star-forming regions at 150 to 500 pc . In addition, the sources need to be resolved out against the larger ($2000\text{--}5000 \text{ AU}$) envelope in which they are embedded (e.g., Boss and Yorke, 1995).

Resolving binary protostars at the peak of the separation distribution for pre-main-sequence stars ($\sim 30 \text{ AU}$) 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 dominates. The preferred wavelength range to detect embedded protostars directly is 20 to $25 \mu\text{m}$; diffraction-limited imaging ($\sim 0.8 \text{ arcsec}$ 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, JWST will use mid-IR integral-field spectroscopy of a range of narrow-band features (extinction windows, shock tracers, and PAHs) and the intervening continuum simultaneously, thus yielding a large number 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 $7 \times 10^{-19} \text{ erg cm}^{-2} \text{ s}^{-1}$ per resolution element is needed to detect a $10 \mu\text{Jy}$ source, such as a typical Class 0 protostar in Taurus-Auriga, at a distance of about 140 pc .

4.2. HOW DOES ENVIRONMENT AFFECT STAR FORMATION AND VICE VERSA?

How do very massive stars form? How do stellar winds and ionizing radiation from massive stars affect nearby star formation?

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. The mechanism by which massive stars form is not yet known. The standard disk-accretion scenario appropriate for low-mass stars cannot be simply scaled up, since the radiation pressure from the growing central source would build up so quickly that no more material could accrete, limiting the mass of the source (Yorke and Sonnhalter, 2002).

Various alternatives have been proposed to solve this problem. McKee and Tan (2002) have suggested that very high accretion rates can overcome the radiation pressure, allowing a massive star to form “normally” via infall and a disk. However, to date, no definitive evidence for circumstellar disks around massive stars exists, leaving this hypothesis unconfirmed. Alternatively, Bonnell *et al.* (1998) proposed that massive stars might form by the agglomeration of colliding low-mass stars in the very deep potential well at the center of a dense young cluster. In this model, stellar densities of 10^6 to 10^8 stars per cubic parsec are required, on the order of 100–1000 times the density seen in the 0.1 pc cores of present-day young embedded clusters.

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 0.01 pc, or ~ 2 arcsec at 1 kpc: such densities are predicted in some models of early cluster evolution (see, e.g., Kroupa *et al.*, 1999; Stahler *et al.*, 2000). There are a number of observational consequences of massive star formation via this collisional route (Bally and Zinnecker, 2005), including high-luminosity IR flaring and impulsive wide-angle outflows, in addition to the simple fact of a very high density of low-mass stars crammed into a small region, all of which should be traceable with the JWST.

Once born, massive stars can be disturbing neighbors. On small scales, ionizing photons and strong winds from O and B stars can destroy disks around young low-mass stars (Johnstone *et al.*, 1998; Bally *et al.*, 2000). 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 and Lazareff, 1994; Hester *et al.*, 1996).

An example of these environmental impacts is seen in M16, 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 M16) 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-IR observations show that about 10–20% contained embedded low-mass stellar and

brown dwarf candidates (McCaughrean and Andersen, 2002; see Figure 14), in addition to the higher-mass young stellar objects (YSOs) in the tips of the ablating columns (see also Sugitani *et al.*, 2002; Thompson *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 surrounding O and B stars, or whether pre-existing YSOs are revealed as the parent cores are eroded away (see also Smith *et al.*, 2005). Do O and B stars have an 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?

4.2.1. Observations

In order to solve the massive star formation paradox, deep, high spatial resolution imaging is required in the mid-IR to see through the high extinction into the very dense, very young cluster cores. These stars are obscured by the considerable density of gas and dust that accompany them: a prototypical hyper-compact HII region might contain a volume density of 10^7 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 μm would be preferred, but some tradeoff with spatial resolution will be sought, given the potentially very high density of point sources that would be expected in the agglomeration hypothesis where, for example, as many as 50–100 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 μm , $A_V \sim 150 \text{ mag}$ would be reduced to 6 and 4 mag, respectively. The aim would be to measure the full stellar IMF and thus in order to detect a fiducial $0.1 M_\odot$, 1 Myr old star at 1 kpc, seen through this dust column, JWST will reach point source limits of ~ 2 and 10 μJy , 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, ~ 0.15 arcsec resolution should yield about 30 fully-sampled resolution elements across 4 arcsec, equivalent to 4000 AU at 1 kpc distance.

To determine the impact of massive stars on their environments, comprehensive surveys will 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 (see, e.g., Smith *et al.*, 2005).

Sensitive mid-IR observations are needed to penetrate the dust in such trunks, and reveal pre-main-sequence stars and brown dwarfs, as well as very young protostellar



Figure 14. Near-IR ($1\text{--}2.5\ \mu\text{m}$) image of the M16 Elephant Trunks made using the ESO VLT. The J_S -band data are shown as *blue*, H-band as *green*, and K_S -band as *red*. The image covers $158 \times 214\ \text{arcsec}^2$ ($1.5 \times 2.0\ \text{pc}^2$ at $1.9\ \text{kpc}$); north is *up*, east *left*. Subimages show more detail, including evaporating gaseous globules from Hester *et al.* (1996) found to be associated with point sources; E23, a globule with no near-IR 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 and Andersen, 2002).

sources. Broadband imaging at 3–20 μm is needed 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 models. In addition, narrow-band line imaging in a variety of ionized, atomic, and molecular tracers is needed 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 will be conducted at high spatial resolution to permit unambiguous identification of a given source with a given globule: typical sizes are 200–2000 AU or 0.1–1 arcsec at 2 kpc distance.

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, $0.02M_{\odot}$, low-mass brown dwarf (Baraffe *et al.*, 1998) with typical IR excess emission due to a circumstellar disk (Kenyon and Hartmann, 1995) would be ~ 55 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 $A_V = 100$ mag. Under similar conditions, a $0.075M_{\odot}$ 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. Broadband imaging photometry is needed in the near- and mid-IR down to the $0.02M_{\odot}$ limit, with multiobject spectroscopy at $R \sim 100$ for the shorter near-IR wavelengths, and $R \sim 1000$ at the longer. In the mid-IR, integral field spectroscopy at $R \sim 2000$ of selected individual sources down to the star-brown dwarf boundary is required.

4.3. WHAT IS THE INITIAL MASS FUNCTION AT SUB-STELLAR MASSES?

Does cloud fragmentation explain low-mass star formation, and is there a lower limit to the mass? How does the sub-stellar IMF depend on metallicity or environment?

The IMF is a key product of star formation (Salpeter, 1955; Miller and Scalo, 1979). Remarkably, the IMF is almost entirely feature-free, all the way from the most massive stars down to $\sim 0.3M_{\odot}$. Below $0.3M_{\odot}$, studies show that the IMF flattens somewhat, but continues to increase. Below $0.1M_{\odot}$ ($100M_{\text{JUP}}$), 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 $0.1M_{\odot}$, the physics of star formation produces a turnover and decline in the IMF, perhaps with some lower-limit boundary condition.

The classical theory of opacity-limited fragmentation has long predicted a significant boundary around $0.003\text{--}0.01M_{\odot}$, equivalent to $3\text{--}10M_{\text{JUP}}$, below which it is believed that cores become opaque to their own radiation, and therefore cannot cool and fragment any further (Hoyle, 1953; Low and 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 $1M_{\text{JUP}}$. However, 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 and Nordlund, 2002), dynamical interactions between protostars (Bate *et al.*, 2002), feedback due to strong bipolar outflows (Adams and Fatuzzo, 1996), and ionizing radiation from massive stars (Palla and 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 in the $10\text{--}100M_{\text{JUP}}$ regime yield surprisingly different results. On one hand, a rising substellar IMF has been found in the 100 Myr Pleiades (Bouvier *et al.*, 1998; 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 $5\text{--}10M_{\text{JUP}}$ (Hillenbrand and Carpenter, 2000; Lucas and Roche, 2000; Luhman *et al.*, 2000; McCaughrean *et al.*, 2002; Muench *et al.*, 2002).

Searching for and characterizing sources at $1M_{\text{JUP}}$ and below in nearby star-forming regions would allow us to constrain the physics of any lower-limit boundary (Figure 15). In addition, free-floating objects less massive than $10M_{\text{JUP}}$ can serve as important proxies for true planets, yielding crucial insights into their early evolution

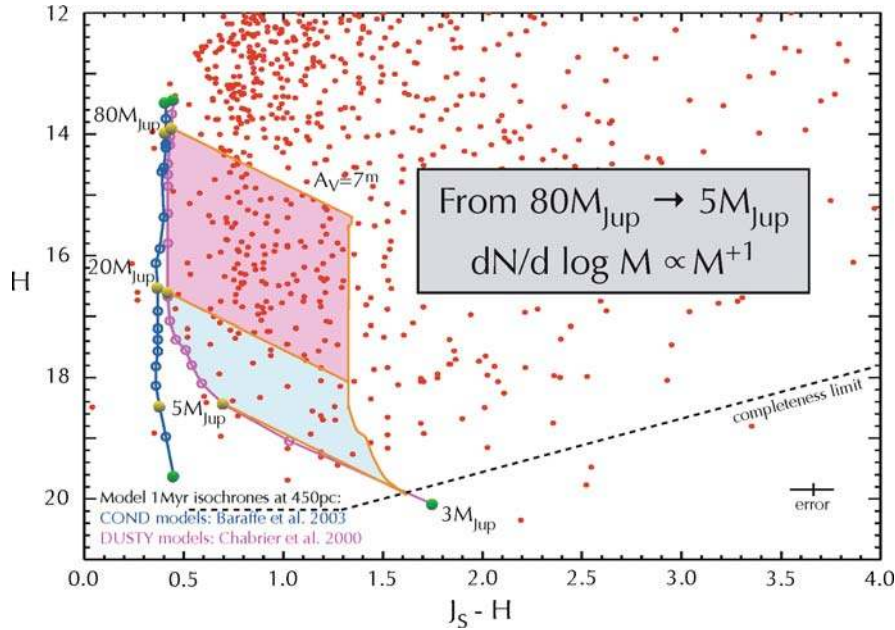


Figure 15. H vs. $(J_S - H)$ color-magnitude diagram for the Orion Trapezium Cluster. An extinction-limited sample of sources in the range $0.005\text{--}0.08M_{\odot}$ ($5\text{--}80M_{\text{JUP}}$), assumed to be 1 Myr old, shows that the brown dwarf end of the IMF is falling, as roughly characterized by the form $dN/d\log M \sim M^{+1}$ (From McCaughrean *et al.*, 2002).

and complementing JWST studies that will be made of mature giant planets in orbit around nearby stars.

Furthermore, the near-IR sensitivity and high spatial resolution of 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.

4.3.1. *Observations*

In order to clearly identify the bottom of the IMF, it is important that we probe to well below $1M_{\text{JUP}}$. However, sources with this mass at 1 Myr have effective temperatures just below 1000 K, and thus the bulk of the flux from objects at this mass and lower is emitted in the thermal IR longward of $3\mu\text{m}$, making ground-based observations prohibitive. JWST will be able to detect such sources readily, permitting surveys at $3\text{--}10\mu\text{m}$ in nearby young clusters to $1M_{\text{JUP}}$ and below.

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

Multiobject spectroscopy at $2\text{--}5\mu\text{m}$ will 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, making it possible to accurately deredden the sources and locate them in theoretical pre-main-sequence Hertzsprung-Russell diagrams. Typically, within a given cluster, there will be 100–1000 sources with masses in the range $1\text{--}100M_{\text{JUP}}$ (see Figure 15).

$R \sim 1000$ multiobject spectra across the $2\text{--}5\mu\text{m}$ regime will allow spectral typing with the H_2O and methane absorption bands, atomic Na, K and Ca lines, and H_2 collisionally-induced absorption. Surface gravity indicators will be used to eliminate older, high-gravity non-cluster members and spectral-type based effective temperatures will be assigned to the true cluster members.

In extremely crowded regions with bright nebulosity, contrast and light leakage in the spectrograph becomes issues, 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 $2\text{--}5\mu\text{m}$ 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

TABLE IV
Predicted fluxes for substellar objects

Mass (M_{JUP})	$T_{\text{eff}}(K)$	$2.2 \mu\text{m}$	$3.8 \mu\text{m}$	$4.8 \mu\text{m}$
0.5 kpc distance				
1	941	290	270	1,300
10	2,251	34,000	33,000	28,000
100	2,856	950,000	820,000	760,000
5 kpc distance				
1	941	2.9	2.7	13
10	2,251	340	330	280
100	2,856	9,500	8,200	7,600
50 kpc distance				
1	941	0.03	0.03	0.13
10	2,251	3.4	3.3	2.8
100	2,856	95	82	76

Note. Predicted fluxes in nJy at near-IR wavelengths for $1M_{\text{JUP}}$, $10M_{\text{JUP}}$, and $100M_{\text{JUP}}$ sources are modeled at 1 Myr age, and 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; Baraffe *et al.*, 1998; Chabrier *et al.*, 2000; Burrows *et al.*, 2001; Marley *et al.*, 2002; Baraffe *et al.*, 2003).

vary considerably. Table IV shows the sensitivities required to reach mass limits of 1, 10 and $100M_{\text{JUP}}$, 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.

For the nearby clusters at ~ 500 pc, JWST will carry out accurate broadband near- and mid-IR photometry, as well as classification spectroscopy at $2\text{--}4 \mu\text{m}$ down to $1M_{\text{JUP}}$, reaching a limiting continuum sensitivity of 300 nJy. This classification will be done with either multiobject spectroscopy or narrow-band $R \sim 100$ imaging in (approximately) 10 bands.

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

In all young clusters, $3\text{--}10 \mu\text{m}$ 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 and Testi, 2001; Liu *et al.*, 2003).

High spatial resolution images of clusters are needed 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 will 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 centroid accuracy for point sources equal to the Gaussian sigma of the point spread function divided by the signal-to-noise ratio, a 50σ measurement will yield 0.8 milliarcsec precision at $2\mu\text{m}$ or 0.4 AU at 500 pc. Measurements at 3 epochs over 3 yr will permit individual proper motions to be measured for sources down to $1M_{\text{JUP}}$ to an accuracy of about 1 km s^{-1} in nearby clusters, resolving the typical cluster velocities of 3 to 4 km s^{-1} (Jones and Walker, 1988). Uncertainties in the geometrical distortion correction may limit this technique.

4.4. HOW DO PROTOPLANETARY SYSTEMS FORM?

How do circumstellar disks form and evolve? What determines their physical sizes? How do dust grains within the disks form planets?

The existence of disks around young low-mass stars was firmly established in the 1990s, initially based on indirect measurements, such as SEDs, asymmetric wind profiles, polarization mapping (Beckwith and Sargent, 1993; Strom *et al.*, 1993), and more recently through direct imaging (McCaughrean *et al.*, 2000; Wilner and 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 through mid-IR 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 (Jayawardhana *et al.*, 1998; Koerner *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 and Wen, 1994; Bally *et al.*, 1995, 2000; O'Dell and Wong, 1996) 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 HII region or within the proplyds (O'Dell and Wen, 1994; McCaughrean and O'Dell, 1996; Figure 16).

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 (50–1000 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 and 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 (nonionized) 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-IR wavelengths, an estimate can be made of the dominant particle size in its outer reaches. HST

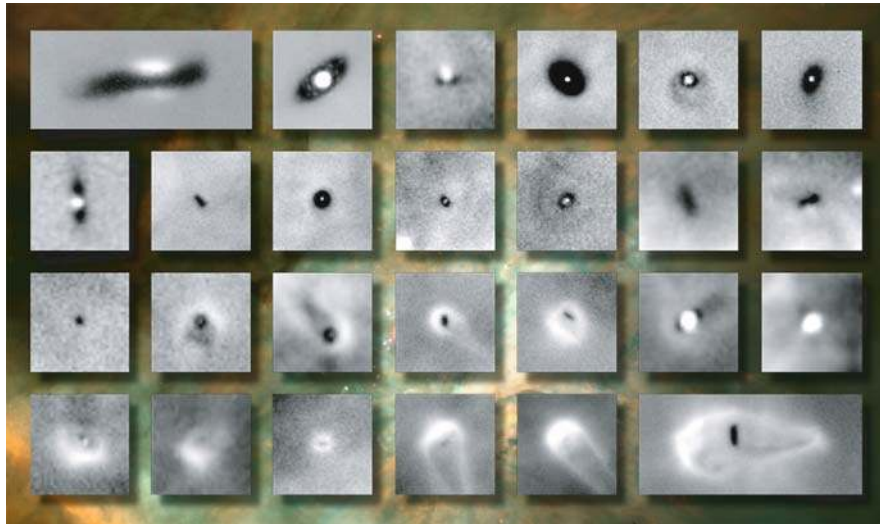


Figure 16. A collection of young circumstellar disks seen as silhouettes against the bright background emission of the Orion Nebula HII region, imaged using HST. All of these sources are associated with young (0.5–2 Myr) low-mass ($0.3\text{--}1.5M_{\odot}$) members of the Trapezium Cluster. In several 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 $1.67 \times 1.67 \text{ arcsec}^2$ or $750 \times 750 \text{ AU}^2$ 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 [OIII] filter of STIS. Data from McCaughrean and O'Dell (1996); Bally *et al.* (2000).

and ground-based adaptive-optics studies to date have proven ambiguous, with contradicting suggestions of ISM-like grains and much larger $5\ \mu\text{m}$ grains (cf. McCaughrean *et al.*, 1998; Throop *et al.*, 2001; Shuping *et al.*, 2003). However, due to the relatively limited spatial resolution of HST at the critical near-IR wavelengths, it has not been possible to attempt these studies for any but the single largest disk.

4.4.1. Observations

JWST will provide high spatial resolution imaging of a substantially larger sample of silhouette disks in selected near-IR wavelengths where the background HII region is particularly bright. These observations will 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 α line, inaccessible from the ground, and the $4.05\ \mu\text{m}$ Br α line, where adequate sensitivity is hard to achieve from the ground.

Narrowband ($R \sim 100$) diffraction-limited JWST imaging will yield angular resolutions of 0.07 and 0.15 arcsec in the Pa α and Br α lines, respectively, corresponding to 30 and 70 AU at the Orion distance of 500 pc. These resolutions imply that for the more than 10 silhouette disks with diameters of 200 AU (0.4 arcsec) or greater, it will be possible to measure the outer-disk radial profiles. In combination with similar resolution HST images in the H α , [OIII], and [OII] 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 α fluxes around the Orion silhouette disks and the known foreground dust extinction, predicted Pa α line fluxes in the outer parts of the nebula are $\sim 10^{-13}\ \text{erg s}^{-1}\ \text{cm}^{-2}\ \text{arcsec}^{-2}$ or $10^{-16}\ \text{erg s}^{-1}\ \text{cm}^{-2}\ \text{pixel}^{-1}$, assuming 0.03 arcsec pixels. In the disk centers, the fluxes will be roughly 10 times fainter, about $10^{-17}\ \text{erg s}^{-1}\ \text{cm}^{-2}\ \text{pixel}^{-1}$, and accurate mapping of the disk profiles requires imaging at 100σ to this flux level through a 1% narrow-band filter in the near-IR.

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

4.5. WHAT ARE THE LIFE CYCLES OF GAS AND DUST?

How do gas-phase molecules interact with dust grains in quiescent cloud cores? How does the formation of a star and planetary system affect the astrochemical evolution of the gas and dust? What is the origin of water and organic materials in a planetary system?

Generations of both low- and high-mass stars have converted primordial hydrogen and helium into successively heavier elements, including carbon, oxygen, and

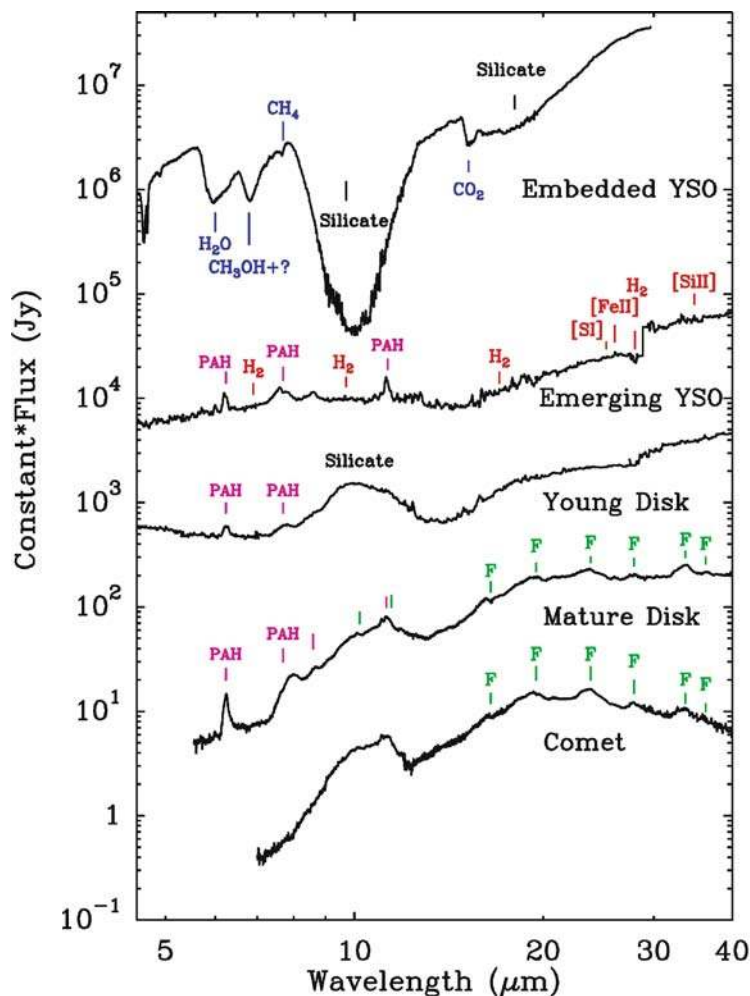


Figure 17. A series of ISO SWS mid-IR spectra of young stars and circumstellar disks at different stages in their evolution (Malfait *et al.*, 1998; Gibb *et al.*, 2000; van den Ancker *et al.*, 2000a,b). From *top to bottom*, in a rough evolutionary sequence, the spectra change from being dominated by solid-state absorption features and shocked gas emission lines, to PAH features and photodissociation region lines, to amorphous and crystalline silicates with HI recombination lines. An ISO SWS spectrum of Comet Hale-Bopp is shown for comparison (Crovisier *et al.*, 1997).

nitrogen, elements that make up life on Earth. Returned to the interstellar medium via winds and SNe, these elements 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 cycles of gas and dust from prestellar cores to planetary systems (van Dishoeck and Blake, 1998; Ehrenfreund and Charnley, 2000; Waters, 2000; see Figure 17).

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 IR and millimeter observations yield indirect evidence that there is substantial depletion, with up to 90% of the heavy elements frozen out onto grains (Kramer *et al.* 1999; Lada *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 CO₂ bending mode at 15 μm , the OCN⁻ band at 4.62 μm , and the unidentified 6.85 μm feature (Schutte and Greenberg, 1997; Ehrenfreund *et al.*, 1998; Gerakines *et al.*, 1999; Schutte and 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 H₂, CO, CH₄, and C₂H₂ 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 1–50 AU from the central star in nearby debris disks can be traced using the pure rotational lines of H₂: $J = 5-3$ 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 H₂ results with dust measurements. These H₂ 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 determine 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-IR spectroscopy of individual sources in star-forming regions spanning a range of ages and environments. Spectra from the Infrared Space Observatory Short Wavelength Spectrograph (ISO SWS) have shown signs of material evolution in the disks around more massive stars (Malfait *et al.*, 1998; Bouwman *et al.*, 2001; Meeus *et al.*, 2001) and the H₂ 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 larger sample, including proto-solar-type stars.

4.5.1. Observations

Spectroscopy of individual sources covering 3–28.3 μm is needed at spectral resolutions of $R = 1000\text{--}3000$, depending on the source and the tracers under study, with smoothing to $R \sim 100$ (with an 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 5σ limiting sensitivity of $10\ \mu\text{Jy}$ is needed 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 pc.

To study the evolution of gas in disks, a line flux sensitivity of $2.6 \times 10^{-17}\ \text{erg s}^{-1}\ \text{cm}^{-2}\ \text{arcsec}^{-2}$ at $17\ \mu\text{m}$ would enable the detection of about half an Earth mass of 100 K molecular gas in a debris disk at 30 pc. It is also desirable that JWST spectroscopy be sensitive to the lowest H_2 transition line at $28.22\ \mu\text{m}$, which will be the brightest line and sensitive to the coldest gas, although the sensitivity at this wavelength will be severely limited by the detector technology. High spectral resolution is essential at this wavelength to obtain adequate contrast of the emission line relative to the continuum.

Finally, integral field spectroscopic capability in the mid-IR with JWST will allow accurate mapping of the emission and absorption components in a given system. A field of at least 3 arcsec will 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 for near-IR spectroscopy will allow the faint disk material to be well separated from the potentially bright central source, and to map out spatial structures in the disks.

4.6. SUMMARY

Table V summarizes the measurements needed for the Birth of Stars and Protoplanetary Systems theme. They include:

TABLE V
JWST measurements for the birth of stars and protoplanetary systems theme

Observation	Instrument	Depth, Mode	Target
Cloud collapse	NIRCam	9 nJy at $2\ \mu\text{m}$	e.g., Barnard 68
	MIRI (imaging)	1 μJy arcsec at $7\ \mu\text{m}$	
Evolution of protostars	MIRI (imaging)	0.1 μJy at $6\ \mu\text{m}$, 1 μJy at $15\ \mu\text{m}$	e.g., Taurus–Auriga
	MIRI (spectra)	$7 \times 10^{-19}\ \text{erg cm}^{-2}\ \text{s}^{-1}$ at $15\ \mu\text{m}$	Class 0 protostars in Taurus–Auriga
Massive stars	NIRCam	2 μJy at $3.8\ \mu\text{m}$, 10 μJy at $4.8\ \mu\text{m}$	e.g., Eagle Nebula
	NIRCam	55 nJy at $3.5\ \mu\text{m}$, 2.3 μJy at $4.8\ \mu\text{m}$	
	NIRSpec	fixed slits; IFU	e.g., Sources in the Eagle Nebula
	MIRI (imaging)	34 μJy at $20\ \mu\text{m}$	
IMF	NIRCam	2.9 nJy at $2.2\ \mu\text{m}$	e.g., Orion Nebula, 30 Doradus
	NIRSpec or TFI	290 nJy at $2.2\ \mu\text{m}$	
	MIRI		
Protoplanetary systems	TFI	$10^{-13}\ \text{erg s}^{-1}\ \text{cm}^{-2}\ \text{arcsec}^{-2}$ at $1.87\ \mu\text{m}$; coronagraph	e.g., Orion Nebula
Astrochemistry	MIRI (spectra)	$2.6 \times 10^{-17}\ \text{erg s}^{-1}\ \text{cm}^{-2}\ \text{arcsec}^{-2}$ at $17\ \mu\text{m}$	Protostars and Debris Disks

- *Cloud Collapse*: NIRCam imaging of background stars through dark dust clouds will map the extinction and thus the density of the clouds. MIRI imaging will reveal the densest inner cores, which are not reachable even with NIRCam.
- *Evolution of Protostars*: MIRI imaging of “Class 0” protostars will show their density and structure. These stars are still embedded in the dense dust clouds, and are not visible at shorter wavelengths. MIRI spectroscopy of protostars will provide diagnostics of extinction windows, shock tracers, and PAHs.
- *Massive Stars*: High spatial resolution imaging of crowded star-forming regions will test the agglomeration hypothesis for massive star formation. Deep imaging of “elephant trunks” will reveal the low-mass stars forming within and test models of the effects of massive stars on their environment. Near-IR spectroscopy will reveal masses, age and evolutionary status of sources forming in the “elephant trunks.” MIRI imaging will see through the densest dust columns to reveal the pre-main-sequence stars and brown dwarfs forming within.

- *IMF*: This program will search for the low-mass end of the IMF in nearby young star-forming clusters, by detecting $1M_{\text{JUP}}$ objects in the Orion Nebula. Scattered light from the bright stars within the nebula are probably the limiting factor. It will also photometrically detect $1M_{\text{JUP}}$ objects in clusters in the inner galaxy, and $10M_{\text{JUP}}$ object in 30 Doradus in the LMC. Spectroscopic follow-up of the more massive objects will reach the sub-brown dwarf stage in the inner galaxy and the faintest M dwarfs in the LMC. MIRI observations of the low-mass end of the IMF will be used to determine the frequency of disks around brown dwarfs, by searching for excess thermal emission. Can planets form around brown dwarfs?
- *Protoplanetary Systems*: TFI imaging will be used to resolve dust disks in silhouette against a brighter background. HST has done this for some disks in Orion. JWST measurements of these same disks will constrain the grain sizes, compositions, and profiles.
- *Astrochemistry*: MIRI spectroscopy of organic molecules (plus H_2 and H_2O) in protostars will be used to study the origins of the “stuff of life.”

5. 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 to 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 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, and the potential for stable habitable regions around other stars.

5.1. HOW DO PLANETS FORM?

What are the physical processes that lead to planets? How common are giant planets and what is the distribution of their orbits? How do giant planets affect the formation of terrestrial planets?

Formation: The formation of multiple objects is a common outcome of star formation, including binary or higher-order star systems, a central star orbited by brown dwarfs and/or planets, or a star with 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 (Chabrier and Baraffe, 2000), while giant planets could

form from a two-stage process in which growth of a rock-ice core triggers the rapid accretion of gas (Lunine *et al.*, 2004). 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 Earth and Neptune, and the two processes may produce distinct IMFs. Even if there is an overlap in the masses generated by the two processes, they would 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, but direct collapse does not need to do so.

The 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 mass, angular momentum, and subsequent disk evolution.

Frequency and Orbits: Planets of Uranian mass or larger exist in orbits that are detectable by the current radial velocity surveys around 8% of F, G, and K-type stars in the solar neighborhood. Extrapolating to larger semimajor axes, and guided by our own dynamically crowded outer solar system (the region from 5 to 30 AU), one can infer that about 15% of mature F, G, and K-type stars should possess at least one giant planet (Marcy *et al.*, 2005). Theoretical studies of the inward migration of giant planets through interactions with the gaseous disk, or with massive remnant particulate disks, suggest that many more stars generate giant planets during their pre-main-sequence phase (Trilling *et al.*, 2002). Many or most of these giant planets are lost through inward migration and merging with the central star, or by ejection. We do not know how many nearby stars 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 semimajor axes, even in the warm inner parts of disks.

Ground-based surveys have already uncovered eight planets that pass in front of their star as seen from Earth (transiting planets) (e.g., Charbonneau *et al.*, 2000; Henry *et al.*, 2000; Alonso *et al.*, 2004; Konacki *et al.*, 2005; Sato *et al.*, 2005; for a review of surveys see Gillon *et al.*, 2005). Transiting planets are the only solar-system-aged extrasolar planets that can be physically characterized until direct imaging is possible. HST and Spitzer have detected the atmospheres of two of these planets, including sodium, an extended atmosphere of hydrogen, and thermal emission (Charbonneau *et al.*, 2002, 2005; Vidal-Madjar *et al.*, 2003; Deming *et al.*, 2005). These atmosphere detections are made possible by a differential measurement of the planet and starlight combined (during primary transit or outside of secondary eclipse) compared with only the starlight (outside of primary transit or during secondary eclipse). In addition to these atmosphere measurements, radii of all transiting planets can be determined, and the wide range of planet densities

has uncovered two of the biggest puzzles in extrasolar planet science so far: an anomalously low density for one of the planets that cannot be explained by the standard planet evolution theory and an anomalously high density for another planet, implying a core several times more massive than those of the giant planets in our own solar system (e.g., Gaudi, 2005).

Effect of Giant Planets on Terrestrial Planets: Given a particular distribution of orbital semimajor 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 Myr 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 (Morbidelli *et al.*, 2000). 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 (Mardling and Lin, 2004). It is even 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.

5.1.1. *Observations*

JWST will provide broad and narrow-band photometry of giant planets and brown dwarfs in bound systems as well as spectra of isolated giant planets and brown dwarfs. JWST will detect giant planets in large orbits that are unreachable by other techniques, provide spectroscopic and photometric information on their thermal properties, and constrain their 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 are incorporated during formation.

JWST will make coronagraphic images at $2.7\ \mu\text{m}$ and $4.44\ \mu\text{m}$ of nearby stars to find mature Jovian companions and of more distant stars to find young Jovian companions. The $2.7\text{--}4.44\ \mu\text{m}$ flux ratio is extremely sensitive to the temperature of planetary companions and provides a rapid diagnosis of an object's age and mass, along with a rough estimate of the orbital parameters (Sudarsky *et al.*, 2003). 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 of clouds on the observations, which in general will be to render the flux ratio less sensitive to temperature.

Following-up the discoveries, JWST will make additional 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 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 of the planet and the age of the companion star, we can then infer the planet's 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.

Isolated or widely-separated giant planets and sub-brown dwarfs can be observed without coronagraphy, making their detection and study much easier, and JWST will do $R \sim 1000$ spectroscopy at $1\text{--}5\ \mu\text{m}$ of these objects (Figure 18). 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 the methane absorption feature is poorly

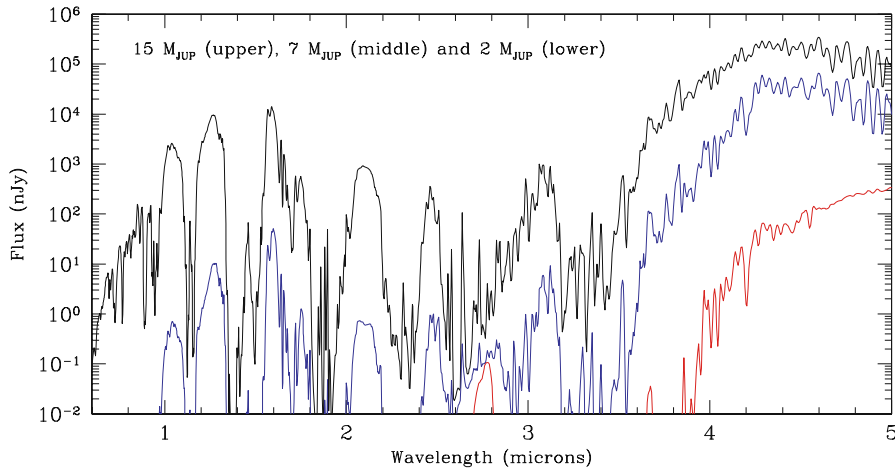


Figure 18. Detectability of extrasolar giant planets and brown dwarfs. We plot model spectra (Sudarsky *et al.*, 2003) of a 15, 7, and $2M_{\text{JUP}}$ extrasolar giant planet, free-floating at 10 parsec from Earth, with an age of 5 billion years. The $5\ \mu\text{m}$ window makes the smallest mass planet detectable with JWST, while the $15M_{\text{JUP}}$ object (transitional to the T-dwarfs) is bright enough to be studied spectroscopically with narrow-band imaging or spectroscopy.

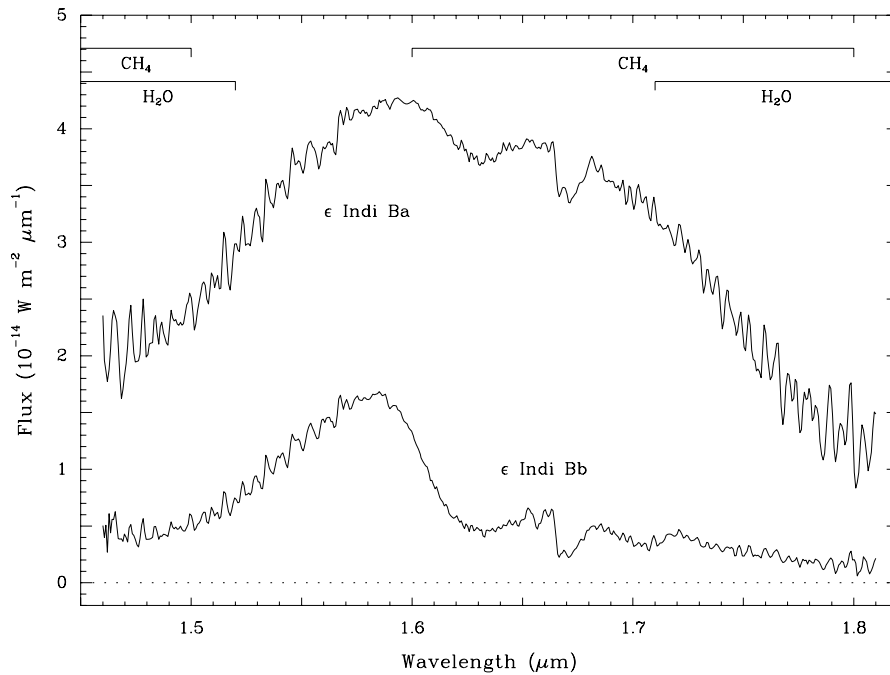


Figure 19. Spectra of the closest T-dwarfs. The closest brown dwarfs to Earth, ϵ Indi B a and b (originally thought to be one object), show distinctly different spectra (McCaughrean *et al.*, 2004). The differences between the two low-resolution spectra relate primarily to effective temperature (1250 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.

understood, since telluric absorption renders these studies difficult with ground-based data. JWST will provide much higher signal-to-noise spectra of T-dwarfs down through warm Jupiters with full coverage through the regions obscured by telluric absorption (Figure 19).

JWST will have the coronagraphic sensitivity to detect a Jupiter analog around a Solar-type star out to ~ 30 pc. As shown in Figure 20, 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 longer wavelengths.

JWST will make spectra at $5\text{--}29\ \mu\text{m}$ of $1M_{\text{JUP}}$ and larger objects within 10 pc. This broad wavelength range is near to, but generally longward, of the Planck-function peak (Figure 20). 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 21 shows that many spectral features are available for filter and spectroscopic study in the near- and mid-IR of objects with masses between that of Jupiter and brown dwarfs.

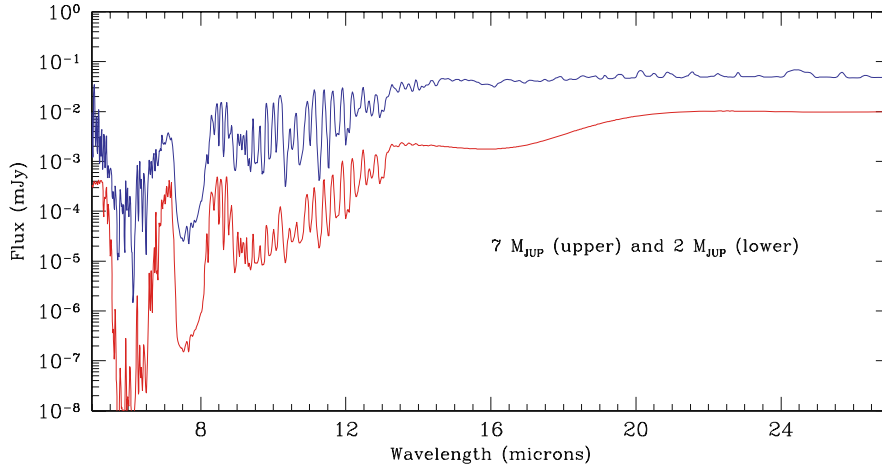


Figure 20. Detectability of Jovian-sized exoplanets in the mid-IR. We plot model spectra of isolated giant planets of 7 and $2M_{\text{JUP}}$, 10 parsec from Earth, around a 5 billion years old star by Sudarsky *et al.* (2003). The $2M_{\text{JUP}}$ object is detectable in JWST broadband coronagraphic imaging; the $7M_{\text{JUP}}$ object is bright enough that spectra at $R \sim 3000$ can be collected, enabling atmospheric structure and composition to be inferred.

JWST will be able to measure density and atmospheric properties of transiting extrasolar planets. These measurements are only possible with the large aperture and stable space environment, together with the excellent calibration due to the precisely predictable on/off nature of both the primary and secondary eclipses. Short-period giant transiting planets will come from the numerous ground-based surveys, while the Kepler space mission will detect Earth-sized to giant planets with semi-major axes out to 1 AU. JWST will use near-IR $R \sim 1000$ spectroscopy to disperse the light from bright stars at a high cadence. JWST will detect atomic and molecular absorption features O_2 , CO_2 , CO , CH_4 , H_2O , Na and K . Extremely precise photometry will be obtained by binning the spectral data together, as was done with HST (Charbonneau *et al.*, 2002). The measurement of a planet's radius and atmosphere will tell us about its composition and evolution. These measurements will help us understand the exotic planets not found in our own solar system, including hot giant planets close to the star, super-Earth-mass rocky planets, and giant planets within 1 AU of the parent star. JWST will also be capable of detecting planetary moons and rings in transit.

5.2. HOW ARE CIRCUMSTELLAR DISKS LIKE OUR SOLAR SYSTEM?

What comparisons, direct or indirect, can be made between our Solar System, circumstellar disks (forming solar systems) and remnant disks?

Detailed observations of disks around other stars, both during and after planet formation and disk clearing, provide a global view of the distribution of major

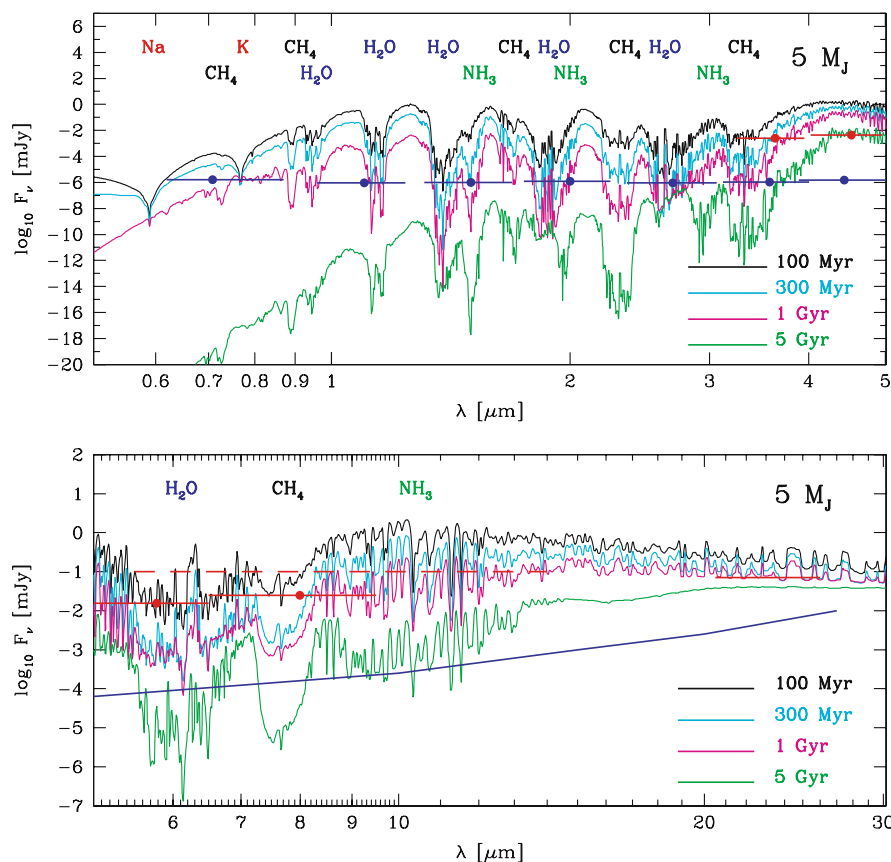


Figure 21. Guide to spectral features in extrasolar giant planets. A $5M_{\text{JUP}}$ planet of varying ages is shown, with major species responsible for the spectral features labeled. Approximate Spitzer (red) and JWST (blue) sensitivities are shown. Detection of these features in the near- and mid-IR samples of different regions of the atmosphere, allowing abundance to be separated from temperature profile and the effects of clouds (From Burrows *et al.*, 2003).

condensables and the effect of planets on the distribution of the dusty, small body material. 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 zodiacal light, interplanetary gas, and dust. 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 mantle and crust, suggesting

that much of the “primitive inventory” is missing or sequestered in the core (Drake and Richter, 2002).

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 H₂O ice, while CO and CO₂ are usually the next most abundant. The spectra also show CH₃OH, H₂CO, and CH₄. In addition, C₂H₆ and C₂H₂ 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-1980s 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 illustrated with ISO observations of Hale–Bopp, which showed that the mid-IR reflectivity of the comet’s dust is strikingly similar to that in a protoplanetary disk surrounding the young star HD 100546 (Figure 22; Malfait *et al.*, 1998). In particular, strong emission features of carbon- and oxygen-rich dust

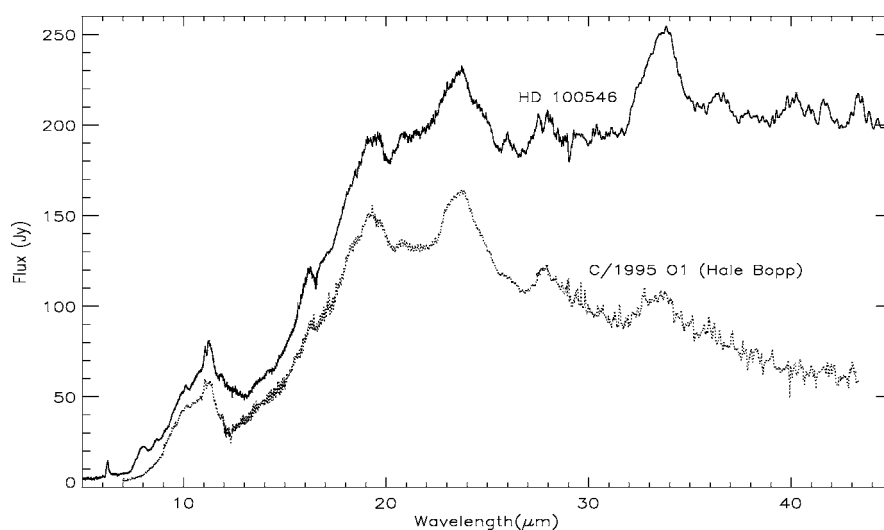


Figure 22. Comets and circumstellar disks. An ISO spectrum of the Herbig Ae/Be star HD 100546 (*upper curve*) shows emission features in the star’s circumstellar disk as predicted if the disk is heated by radiation from the central source. The particle composition in the circumstellar disk appears to be remarkably similar to that of comet Hale–Bopp (*lower curve*). JWST will only reach the shorter wavelengths depicted here, but the similarities are most prominent at those wavelengths (From Malfait *et al.*, 1998).

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.

5.2.1. Observations

Circumstellar Disks: JWST will resolve the details of nearby debris 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 in the interiors of the disks.

JWST will test models of debris-disk evolution in the presence of planets out to distances of 40–50 pc. At these distances, it will provide the same spatial resolution available with Spitzer on the Vega disk (8 parsec away), which has been sufficient to determine the first-order effects of a giant planet on disk morphology (Su *et al.*, 2005). Around Vega, Fomalhaut, and other nearby disk systems, details of the perturbations in disk structure on 10 AU scales are likely to be discernable (Figure 23), so that the nature of debris disks can be constrained.

Spectroscopic resolution of 1000 in the near and mid-IR will 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. The grain temperatures in the disks are generally low enough so

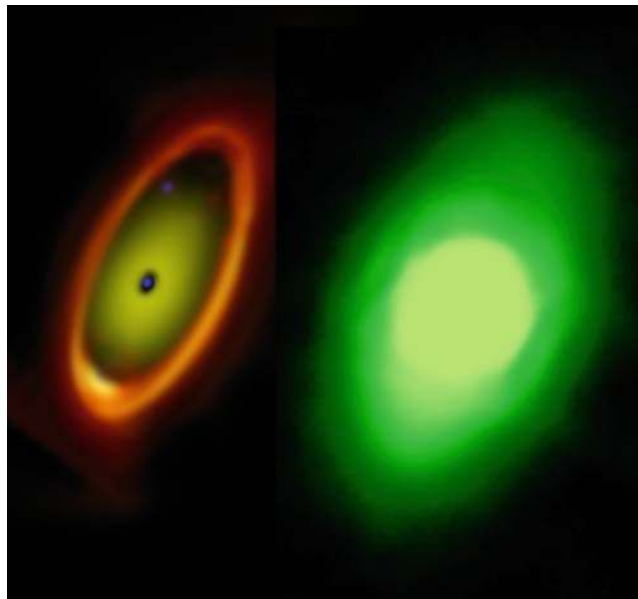


Figure 23. Dusty debris disks. We show the Fomalhaut debris disk as it appears to Spitzer at $24\,\mu\text{m}$ (right; Stapelfeldt *et al.*, 2004), and in a simulated JWST image. Structure within the disk is clearly resolvable by JWST.

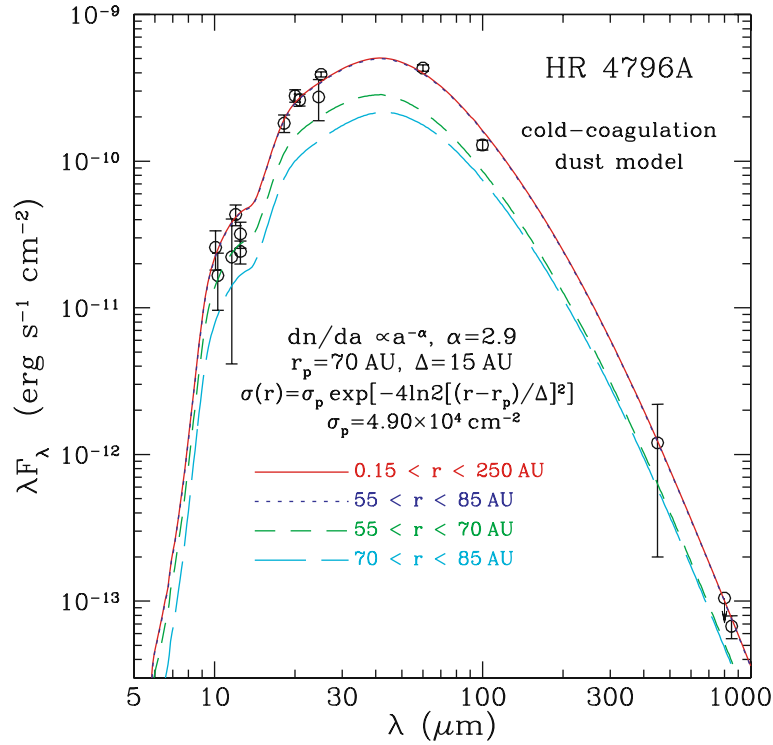


Figure 24. Models and observations of a dusty disk. Models of the SED of the dusty disk around HR 4796 are compared with observations. Emissions from different ranges of semi-major axis are plotted. Most of the action in terms of peak emission and interesting spectral features lies beyond $5\ \mu\text{m}$, as is the case for most dust disks (Li and Lunine, 2003).

that peak emission will be in the mid-IR (Figure 24), and there is little dust at much warmer temperatures.

Comets: Comets are remnants of Solar System formation, and their 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 IR 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_2O , CO , CO_2 , and CH_3OH in even relatively faint comets. Near-IR spectrometry with $R \sim 1000$ resolution will be used to measure the ratio of ortho and para H_2O separately (Crovisier *et al.*, 1997), possibly providing an indication of the comet's formation temperature (Mumma *et al.*, 1987). Likewise, mid-IR spectroscopy can

determine 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-IR 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 the short- and long-period dynamical classes. 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 CO₂ abundance in comets that come within ~ 3 AU of the Earth and Sun. Depending on the circumstances, JWST can measure CO₂ emission in either the ν_3 band near $4.3\ \mu\text{m}$ or the ν_2 band near $15\ \mu\text{m}$, both of which are exceptionally strong (Goody and Yung, 1989). Only JWST can detect CO₂; strong absorption in the terrestrial atmosphere prevents ground-based IR observations. The molecule does not have a permanent electric dipole moment so it does not emit in the radio.

Although we now have several means of measuring H₂O and CO directly in comets from ground-based observatories (Mumma *et al.*, 2003), JWST will do so with much higher sensitivity. From the ground, H₂O can only be observed in intrinsically faint intercombination bands; atmospheric absorption prevents observations in the much stronger fundamental bands. By observing H₂O 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 (Mumma *et al.*, 1987). However, cosmic rays may alter the OPR in the outer layers of comets in the Oort cloud (Mumma *et al.*, 1993). In that case the OPR may provide an unambiguous method for identifying the comet as dynamically new.

Measuring the HDO/H₂O ratio 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 (Meier *et al.*, 1998). This is the key to understanding the relationship between water in comets and that on the Earth and Mars (Lunine *et al.*, 2003). JWST's high-sensitivity spectral observations in the near- and mid-IR will enable this measurement in certain comets and the largest Kuiper Belt Objects (KBOs).

Similarly, the CO abundance can be measured 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 IR facilities. Radio observations of CO are much less sensitive than IR observations, because of CO's small electric dipole moment.

Cometary molecules have spectral lines across the entire JWST spectral range. Primary vibration-rotation lines are concentrated in the $1\text{--}5\ \mu\text{m}$ band. We also need observations at $15\ \mu\text{m}$ to study CO₂.

Spitzer has 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.

Moving Targets: 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/s. 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), 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 s (for very faint and typically slow objects).

Moving target tracking is currently under study by the JWST Project to determine if this capability is affordable.

5.3. HOW ARE HABITABLE ZONES ESTABLISHED?

What are the sources of water and organics for planets in habitable zones? How are systems cleared of small bodies? What are the planetary evolutionary pathways by which habitability is established or lost? Does our Solar System harbor evidence for steps on these pathways?

Water and Organics: Some geochemical evidence suggests that Earth's water did not come from locally formed planetesimals at 1 AU (Morbidelli *et al.*, 2000). 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 the section of the primordial asteroid belt that supplied water to the Earth (Robert, 2001). However, other geochemical evidence severely limits the amount of chondritic material the Earth might have acquired (Drake and Righter, 2002). Cometary HDO/H₂O values measured in three long-period comets are twice that of Earth's ocean water (Meier *et al.*, 1998). 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 (Lunine *et al.*, 2003). 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 (Pierazzo and Chyba, 1999) and the inventory of organics derived from high-sensitivity IR spectra will be of value in constraining the theories. 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.

Small Bodies: Studies of active gas-dust disks and remnant disks from JWST, with Spitzer disk studies as a foundation (Meyer *et al.*, 2004), 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 the early history of planetary systems. 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 be particularly important for mid-IR measurements of these objects.

Pathways to Habitability: 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 coevolution 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?

Evidence in our Solar System: 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 these carbonates today? 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, has a dense atmosphere, and possesses 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 4 yr tour?

Prior to the origin of life, chemical processes may have led to a substantial level of complexity (Lunine, 2005). 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) (Botta, 2004). Remote detection of stereochemical orientation is almost impossible, although amino acid detection might be possible if sufficient concentrations exist on optically exposed surfaces.

5.3.1. Observations

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.

KBOs: JWST will obtain near-IR and mid-IR measurements of the brightness of KBOs to separate effects due to size and albedo and constrain their physical size. Spectra will be obtained for 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 KBO, as indicated by Triton's retrograde orbit and Neptune's dilapidated remnant regular satellite system (McKinnon *et al.*, 1995).

JWST will obtain high-resolution, near-IR 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 in Figure 25. Ratios of isotopes, including $^{13}\text{CO}/^{12}\text{CO}$,

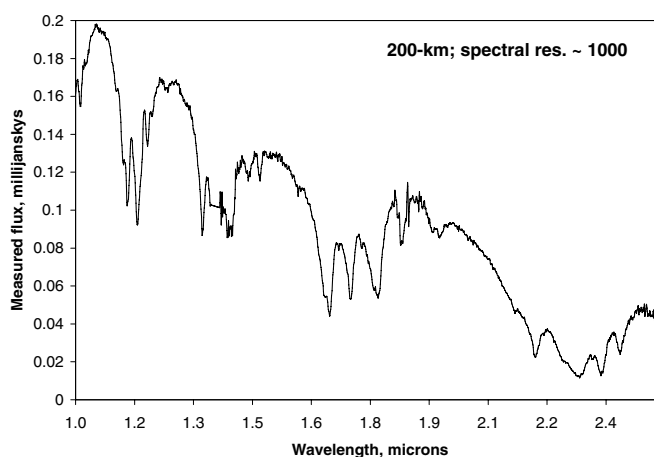


Figure 25. Simulated KBO spectrum. The simulated spectrum of a KBO with a 200 km radius using a spectrum borrowed from that of Pluto (Cruikshank *et al.*, 1997). This spectrum is well above JWST's sensitivity limit. Many or all of the major ices can be identified with absorption features out to 2 μm (Schmitt *et al.*, 1998).

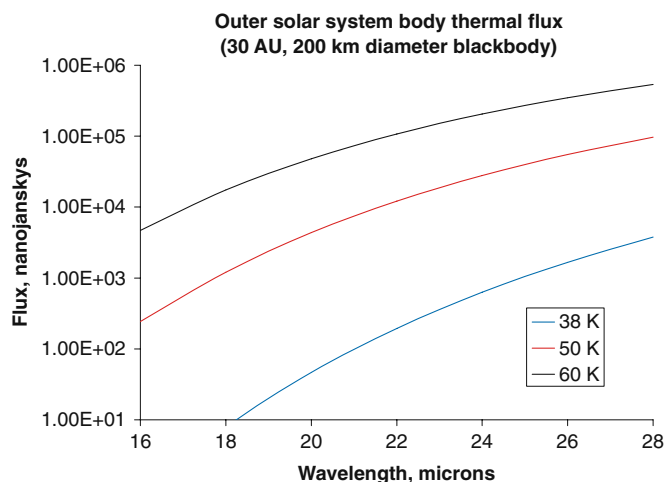


Figure 26. Blackbody flux vs. wavelength in the mid-IR for a 200 km radius KBO at three different possible temperatures, located 30 AU from JWST.

$C^{18}O/C^{16}O$, $^{13}CH_4/^{12}CH_4$, $^{15}NH_3/^{14}NH_3$, CH_3D/CH_4 , and HDO/H_2O , may be detectable on Pluto 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 KBOs.

JWST will be able to detect KBOs in the mid-IR down to typical discovery sizes of hundreds of kilometers, in order to provide a separate determination of size and brightness, when combined with the optical brightness (Figure 26). 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 KBOs (e.g., volatile-rich vs. quiescent).

Target tracking requirements for KBOs are typically less than those of comets (discussed above), although more accurate tracking rates are needed because of the long exposure times. Current ultra-deep surveys with 8 m ground-based telescopes achieve R -magnitudes above 26 (Sheppard *et al.*, 2005). JWST will enable deeper searches for new objects and higher sensitivity characterization of known ones. Photometry of near-IR reflected light and of thermal emission at $20\ \mu m$ will constrain simultaneously the albedo and the radius. Spectroscopic studies at resolutions of ~ 1000 , comparable to the best ground-based studies of Pluto and Triton, will be extended to smaller and more distant bodies. $R = 3000$ spectroscopy of Triton and Pluto will constrain isotopic ratios in the water ice and other components, as well as providing surface temperature monitoring through the nitrogen overtone band. While the near-IR is the most familiar territory in this regard, because of the ground-based spectra of Triton and Pluto, rotational features will appear in the mid-IR and will provide compositional and isotopic data not attainable in the

near-IR overtone region. Mid-IR detection leading to the separation of the size and albedo of such objects is of fundamental importance, and requires the high mid-IR sensitivity that JWST provides.

Titan: 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 10 yr or longer baseline of space-borne near-IR observations of Titan’s surface and atmosphere. JWST will use near-IR spectrometry with spectral resolution a factor of 6 higher 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 midlatitude regions of Titan. These data will reveal whether surface changes or secular atmospheric changes are in evidence over a decadal timescale.

5.4. SUMMARY

Table VI summarizes the measurements needed for the Planetary Systems and Origins of Life theme. They include:

- *Isolated Extrasolar Giant Planets:* Deep NIRCам imaging (done as part of IMF studies) will detect isolated giant planets and brown dwarfs, or planets in large orbits.

TABLE VI
JWST measurements for the planetary systems and origins of life theme

Observation	Instrument	Depth, Mode	Target
Isolated extrasolar giant planets	NIRCам	4 nJy at 2 μ m	Star-forming regions
Bound planets	TFI or NIRCам	200 μ Jy at 4 μ m	Nearby stars
In-depth study	TFI	60 nJy, $R = 100$	Nearby stars
	NIRSpec	1 μ Jy, $R = 3000$	
	MIRI	1 μ Jy at 15 μ m	
Transiting planets	NIRSpec	$R \sim 1000$	Kepler discoveries
Circumstellar disks	MIRI	1 μ Jy at 15 μ m	Debris Disks
Comets	NIRSpec	1 μ Jy, $R = 3000$ at 2 μ m	10/yr
	MIRI	1 μ Jy, $R = 3000$ at 15 μ m.	
KBOs	NIRCам	4 nJy at 2 μ m	10 arcmin ²
	NIRSpec	1 μ Jy, $R = 3000$ at 2 μ m	
	MIRI	1 μ Jy at 25 μ m	
Satellites	NIRSpec	1 μ Jy, $R = 3000$ at 2 μ m	Titan
	MIRI	1 μ Jy at 15 μ m	

- *Bound Planets*: Using TFI and NIRCам coronagraphy, JWST can observe bound giant planets, and characterize their ages and masses through spectral features.
- *In-Depth Study*: Subject to the discovery of planets appropriate for follow-up (i.e., at the right distance from their star, not too faint, etc.), JWST will get spectra with NIRSpec, narrow-band imaging with TFI, and mid-IR (coronagraphic) photometry with MIRI.
- *Transiting Planets*: Using rapid readout of the NIRSpec detectors, JWST will measure differential light curves of transiting planets discovered by Kepler. These spectra will contain atmospheric signatures of O₂, CO₂, CO, CH₄, H₂O, Na, and K.
- *Circumstellar Disks*: Resolved mid-IR spectra of circumstellar disks will reveal the signature of key ices and silicates, constraining the constituents of terrestrial planets. Resolved, coronagraphic imaging of disks may reveal structures due to shepherding planets.
- *Comets*: NIRSpec spectroscopy of comets will be used to measure the abundances of H₂O, CO, CO₂, and more complex molecules like CH₃OH. Ortho and para H₂O can be separated, indicating the comets' formation temperature. HDO/H₂O will constrain models of the source of water on the Earth and Mars. MIRI spectroscopy will reveal the mineralogy of the comets.
- *KBOs*. NIRCам imaging will be used to find and identify KBOs, characterizing the population's dynamics and subclasses. NIRSpec spectroscopy and MIRI photometry and spectroscopy will be used to study individual objects, revealing molecules and isotopic ratios, size and albedo, and temperature.
- *Satellites*. Using NIRSpec and MIRI, JWST will observe weather on Titan to establish a 10 yr baseline follow-on to the Cassini mission, monitoring seasonal changes as the Saturnian system transitions from summer to autumn. NIRSpec's spectral resolution is a factor of 6 X higher than Cassini's, and will reveal the organic species present on the surface.

6. JWST Implementation

To make the scientific measurements described in Section 2 through 5, JWST will 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. It will be launched early in the next decade to an orbit around the Earth–Sun second Lagrange point (L2). 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. In this section we describe the design of the JWST observatory and mission.

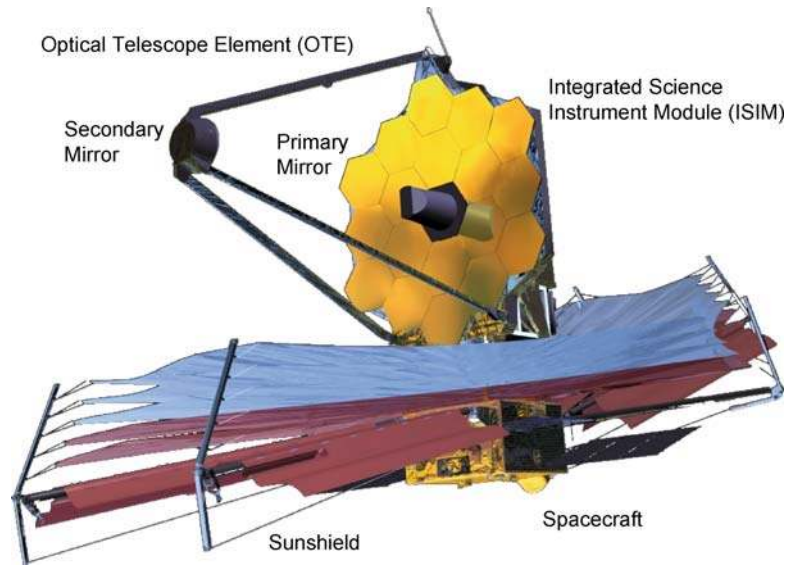


Figure 27. The JWST Observatory. The optical telescope element contains the primary and secondary mirrors, the ISIM element contains the instrumentation, and the spacecraft element consists of the spacecraft and the sunshield.

The JWST Project is organized into three segments: observatory, ground, and launch. The observatory is composed of an optical telescope element, an integrated science instrument module (ISIM) containing the scientific instruments, a spacecraft, and a sunshield (Figure 27). The telescope (Feinberg, 2004) is a deployable optical system that provides diffraction-limited performance at $2\text{ }\mu\text{m}$ using active wavefront sensing and control (WFS&C). The ISIM (Davila *et al.*, 2004; Greenhouse *et al.*, 2004) contains four science instruments (SIs): a NIRCам (Horner and Rieke, 2004), a NIRSpec (Zamkatzian and Dohlen, 2004), a MIRI (Wright *et al.*, 2004), and a near-IR TFI (Rowlands *et al.*, 2004b). The ISIM also contains a fine guidance sensor (FGS) (Rowlands *et al.*, 2004a) to provide active control of pointing. The spacecraft provides the pointing platform and housekeeping functions for the observatory. The wavelength range of JWST and the SIs spans $0.6\text{--}29\text{ }\mu\text{m}$, limited at the short end by the gold coatings on the primary mirror and at the long end by the detector technology. The sunshield shades the telescope and ISIM from solar illumination to allow zodiacal-light-limited performance at $\lambda < 10\text{ }\mu\text{m}$ and high sensitivity out to $29\text{ }\mu\text{m}$, and provides a stable thermal environment for the telescope and ISIM.

6.1. OBSERVATORY

Table VII summarizes the key features of the observatory architecture (Nella *et al.*, 2004). The primary mirror uses semirigid primary mirror segments mounted

TABLE VII
Observatory key features

Element	Feature
Optical telescope element	Three-mirror anastigmat f/20 design Fine steering mirror (FSM) to provide line-of-sight stabilization <7.3 milliarcsec Four separate deployments Semirigid, adjustable hexagonal mirror segments and graphite composite backplane structure
Primary mirror	25 m ² collecting area Primary mirror deploys in two steps (2-chord fold) 18 semirigid hexagonal segments with set-and-monitor wavefront control actuators Mirror segment material is Beryllium
Secondary mirror	Tripod configuration for support structure Deployment using a single redundant actuator Semirigid optic with 6 degrees of freedom alignment
Aft optics	Fixed baffle contains tertiary mirror and FSM
ISIM	Simple semikinematic mount; 8 m ² of thermal radiator area, and 19.9 m ³ volume.
Tower	Contains all science instruments and fine guidance sensor Integral 1 Hz passive vibration isolators
Sunshield	Isolates the telescope from spacecraft both thermally and vibrationally Five-layer “V” groove radiator design reduces solar energy to tens of milliwatt Folded about telescope during launch Sized (~19.5 m × 11.4 m) and shaped to limit solar radiation induced momentum buildup.
Spacecraft bus	Chandra-based attitude control subsystem Two-axis gimballed high-gain Earth-pointing antenna with omnis, Ka and S band 471 Gb solid-state recorder to store 2 days of science and engineering data Propellant for >10 yr

on a stable and rigid backplane composite structure. The architecture is termed “semirigid”, because it has a modest amount of flexibility that allows for on-orbit compensation of segment-to-segment radius of curvature variations. Semirigid segments have a high degree of inherent optical quality and stability, allowing verification of optical performance by 1-g end-to-end ground testing before launch.

The telescope optics are made of beryllium. Spitzer and IRAS had beryllium mirrors and its material properties are known at temperatures as low as 10 K. Beryllium has an extremely small variation in its coefficient of thermal expansion over temperatures of 30–80 K, making the telescope optics intrinsically stable to small temperature variations. Beryllium fabrication and figuring procedures were designed using the results from the Advanced Mirror System Demonstrator (AMSD) program (Feinberg, 2004; Stahl *et al.*, 2004). An engineering demonstration unit will be used as a pathfinder for the flight mirror processing.

The sunshield enables passive cooling and provides a stable cryogenic environment by minimizing the amount of solar energy incident onto the telescope and ISIM. The observatory will not use active wavefront control during observations.

6.2. OBSERVATORY PERFORMANCE

Table VIII summarizes the predicted performance for the JWST Observatory.

6.2.1. Image Quality

The imaging performance of the telescope will be diffraction limited at $2\ \mu\text{m}$, defined as having a Strehl ratio >0.80 (e.g., Bély, 2003). JWST will achieve this image quality using image-based WFS&C of the primary mirror. There will also be a fine guidance sensor in the focal plane and a fine steering mirror to maintain pointing during observations.

TABLE VIII
Performance of the JWST observatory

Parameter	Capability
Wavelength	0.6–29 μm
Image quality	Strehl ratio of 0.8 at $2\ \mu\text{m}$
Telescope FOV	Instruments share $\sim 166\ \text{arcmin}^2$ FOV
Orbit	Lissajous orbit about L2
Celestial sphere coverage	100% annually
	39.7% at any given time
	100% of sphere has at least 51 contiguous days visibility
	30% for >197 days
Observing efficiency	Continuous viewing zone <5 degrees from each ecliptic pole
	Observatory $\sim 80\%$. Overall efficiency $>70\%$
Mission life	Commissioning in less than 6 months
	5 yr minimum lifetime after commissioning
	10 yr fuel carried for station keeping

The Strehl ratio specification is used to determine the allowed optical wave-front error (WFE) and its allocation to low (0–5 cycles/aperture), mid (5–30 cycles/aperture), and high frequencies (> 30 cycles/aperture). Point spread function (PSF) stability is needed to reliably separate the optical PSFs from different targets or for the same target at different observations, and to ensure radiometric stability. A Monte Carlo analysis was performed for the optical design by varying the spatial characteristics of the errors, and worst-case analyses were also performed showing the design and build tolerances are suitable for the science. The allocated top-level WFE is 150 nm root-mean-squared (rms) through to the NIRCam focal plane, and includes both the effect of 7.0 milliarcsec image motion, most of which is line-of-sight jitter, and 51 nm of drift instability.

6.2.2. Sky Coverage and Continuous Visibility

Field of regard (FOR) refers to the fraction of the celestial sphere that the telescope may point toward at any given time. A large FOR increases the number of days per year of target visibility, provides the ability to visit targets repeatedly for time variability studies, flexibility to schedule observations, to revisit failed observations, and to respond to targets of opportunity. JWST's FOR is limited by the size of the sunshield.

Sky coverage performance is shown in Figure 28. A continuous viewing zone within 5° of both the north and south ecliptic poles is available throughout the year. Thirty percent of the sky can be viewed continuously for at least 197 continuous days. All regions of the sky have at least 51 days of continuous visibility per year. The architecture provides an instantaneous FOR at any epoch of approximately 40% of the sky (Figure 29) This FOR extends 5° past the ecliptic pole, and provides 100%

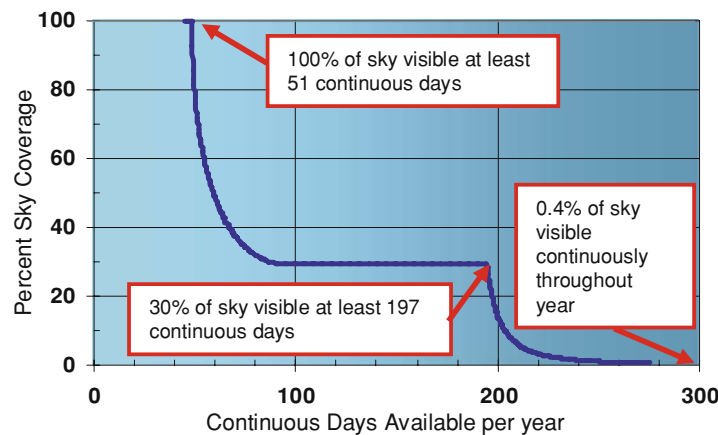


Figure 28. Sky coverage and continuous visibility. There is a continuous viewing zone within 5° of each ecliptic pole. Thirty percent of the sky is viewable for at least 197 days per year, and all of the sky will have at least 51 days of continuous visibility each year.

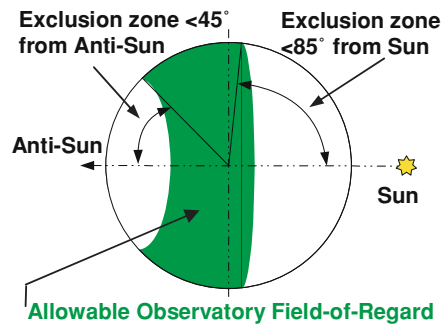


Figure 29. Observatory FOR. JWST will be able to see about 40% of the sky at any given time, in an annulus around the anti-Sun direction.

accessibility of the sky during a 1 yr period. In addition, a nominal 5° angular safety margin will be maintained when determining the allowable Observatory pointing relative to the Sun.

In order to take full advantage of the FOR, JWST will be able to observe any point within 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 without special scheduling or other procedures.

6.2.3. Sensitivity and Stray Light Limitations

Many JWST observations will be background limited. The background is a combination of in-field zodiacal light, scattered thermal emission from the sunshield and telescope, scattered starlight, and scattered zodiacal light. Over most of the sky, the zodiacal light dominates at wavelengths $\lambda < 10 \mu\text{m}$. The flux sensitivities of the JWST instruments are given in Section 6.5. Figure 30 gives the wavelength dependence of the maximum effective background radiance for JWST architecture (the sum of all individual thermal emissions and scattered sources). At $\lambda < 10 \mu\text{m}$, broadband imaging sensitivity will be zodiacal-light limited.

6.2.4. Observational Efficiency

The observational efficiency determines the scientific productivity of the mission. Efficiency is the fraction of the total mission time spent actually “counting photons,” rather than occupied by overhead activities such as slewing, stabilizing, calibrating, and adjusting parts of the observatory. The observatory is designed to be stable and simple to operate, and 19% overhead is allocated to the spacecraft and observatory, including WFS&C operations. An additional 6% overhead is allocated to the instrumentation, including standard-star calibrations, and 5% is allocated to scheduling, including failed guide-star acquisitions. The overall efficiency is expected to be 70%, although this depends on the specific observing programs that are

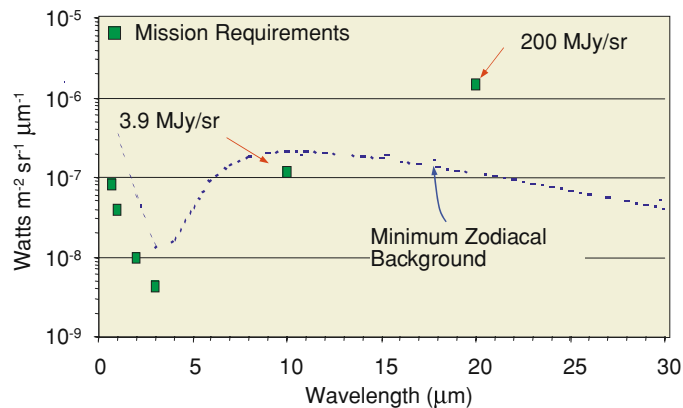


Figure 30. Scattered and self-emission background. JWST will be limited by thermal emission from the telescope at $\lambda > 10 \mu\text{m}$.

implemented. In comparison to missions in low-Earth orbit, there are significant efficiency advantages of the L2 orbit. JWST will use parallel instrument calibration (i.e., darks) to achieve this overall efficiency.

The observatory will be fully commissioned 180 days after launch. The semirigid mirror segment design is expected to be stable, so that wavefront control adjustments will not be needed more than once per week, minimizing the overhead for this activity. Wavefront quality is monitored by using the science target imagery in addition to dedicated observations of bright stars, reducing the overhead. Time intervals used for momentum control are minimized by the sunshield design, which approximately balances the observatory torque from solar radiation pressure over most of the field of regard.

The thermal stability of the observatory is predicted to be good enough that optical adjustments will not be necessary even when large changes in the Sun angle are commanded. Included in the efficiency overhead is an estimate for safe mode based on the design and on-orbit performance for the Chandra Observatory. The largest contributors to the observatory inefficiency are slewing to new targets and the dither of the locations of a single target on the focal plane. This indicates that the JWST architecture is optimized in the sense that the overheads associated with pointing to targets are the dominant contributors to inefficiency.

6.3. OBSERVATORY DESIGN DESCRIPTION

6.3.1. *Optical Telescope Element (OTE)*

The telescope was designed to provide a well-corrected image to the ISIM. It has 132 degrees of freedom of adjustment mechanisms and a composite structure so that passive figure control and passive disturbance attenuation is possible. Figure 31 shows an isometric view of the telescope's primary and secondary mirrors.

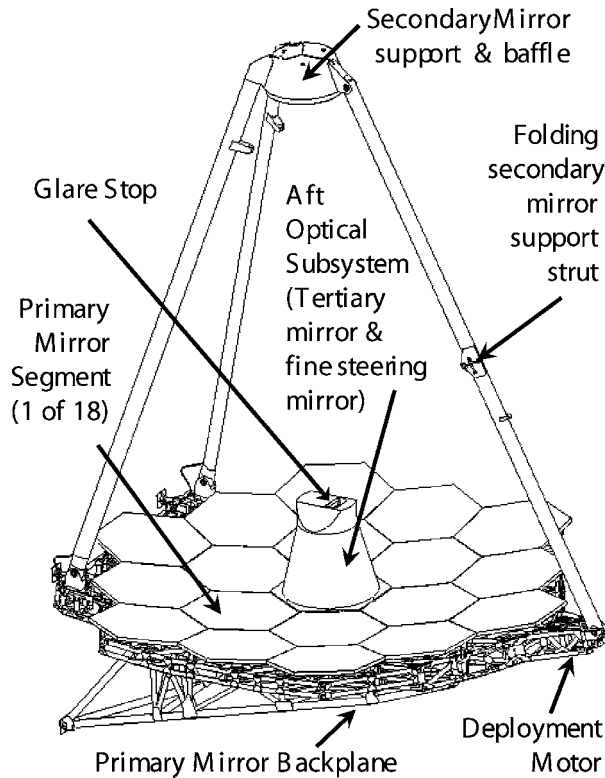


Figure 31. Isometric view of the JWST telescope. The JWST telescope is a three-mirror anastigmat made up of 18 hexagonal segments.

The optical telescope has a 25 m^2 collecting area, three-mirror anastigmat (Figure 32), made up of 18 hexagonal segments. It has a Strehl ratio of approximately 0.8 at $\lambda = 2 \mu\text{m}$. This configuration provides excellent image quality over a large field of view (FOV), and accommodates the science instruments and the guider. The telescope has an effective f/number of 20, and an effective focal length of 131.4 m. Each ISIM instrument reimages the telescope focal plane onto its detectors, allowing for independent selection of detector plate scale for sampling of the optical PSF. A fine steering mirror (FSM) is used for image stabilization. The FSM is located at the image of the pupil, after the tertiary mirror but forward of the focal plane interface to the ISIM. The FSM, coupled with the low structural noise spacecraft, suppresses line-of-sight jitter to allow diffraction-limited performance at $2 \mu\text{m}$. The secondary mirror is supported by a deployable tripod support structure, which latches into position following deployment. This structure provides the necessary jitter stability with minimum beam blockage.

Figure 33 shows the placement of the individual ISIM instrument detectors on the telescope's field of view. The telescope design and control is optimized to

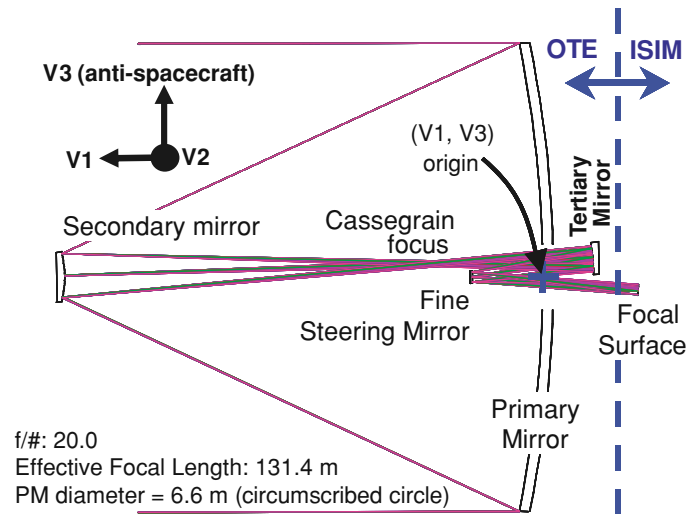


Figure 32. The telescope optical layout. The JWST optical telescope has an effective f/number of 20 and an effective focal length of 131.4 m.

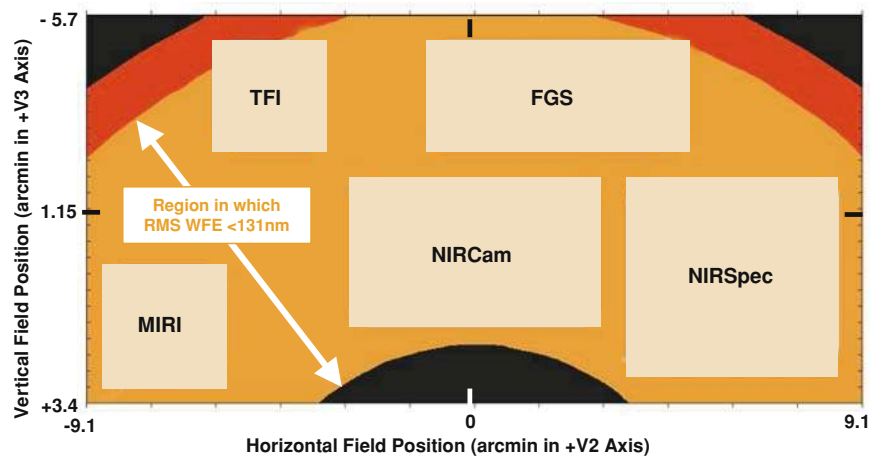


Figure 33. Placement of the ISIM instruments in the telescope FOV.

equalize performance over the total FOV. The WFE that results from residuals in the optical design is a small portion of the total telescope WFE. This allows significant WFE allocation to the manufacturing processes and on-orbit environmental degradation, while still delivering 131 nm rms WFE over the NIRCams field, as required for diffraction-limited performance at $2\ \mu\text{m}$. The outer black area in the figure represents portions of the FOV that have some amount of vignetting.

The NIRCams detectors are placed in a spatial region with the lowest residual WFE to take full advantage of imaging performance. MIRI and NIRSpect are

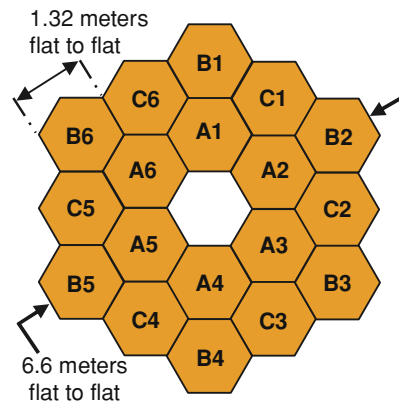


Figure 34. Primary mirror dimension and prescription. The primary mirror is made up of 18 hexagonal segments, each 1.32 m flat to flat. There are three different segment aspheric prescriptions, designated A, B, and C, with six segments of each prescription. The segments with the same prescription are interchangeable.

positioned in an area with slightly larger WFE, but still well within the requirements for instrument performance. Similar considerations exist for placement of the FGS and TFI focal planes within the telescope FOV. The telescope mirrors are gold coated, providing a broad spectral bandpass, from 0.6 to 29 μm .

6.3.2. Mirror Segments

The primary mirror is composed of 18 individual beryllium mirror segments (Figure 34). When properly phased relative to each other, these segments act as a single mirror. Primary mirror phasing is achieved via six degree of freedom rigid body motion of the individual segments, and an additional control for the mirror segment radius of curvature. The six degrees of freedom are decenter (x , y), tip, tilt, piston (z), and clocking. Each mirror segment is hexagonal with a 1.32 m flat-to-flat dimension. There are three separate segment types with slightly different aspheric prescriptions depending on placement. The three prescriptions are identified in Figure 34 by letters A, B, and C. Numbers 1 through 6 represent the sixfold symmetry of the hexagonal packing of the primary mirror. All segments within a type (letter) are completely interchangeable. The mid- and high-spatial-frequency figure errors are caused by manufacturing errors within the segments. The low frequency errors are corrected by the primary mirror adjustments on orbit.

Figure 35 shows the rear portion of the mirror segments and the seven actuators. The six actuators providing rigid body motion are arranged in three bipods (a hexapod) to form a kinematic attachment to the backplane. Each bipod attaches to a triangular structure shown in Figure 35, which is attached to the isogrid structure of the mirror segment. This structure spreads the loads over the surface of the mirror.

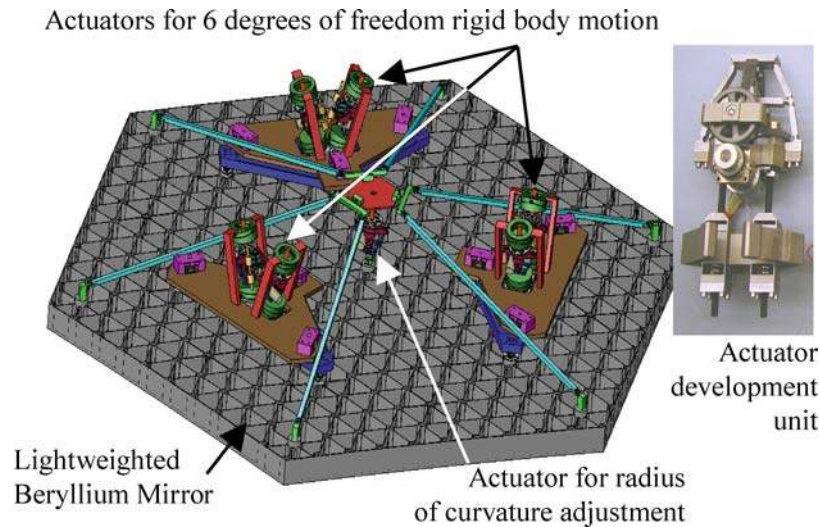


Figure 35. Rear view of a primary mirror segment. Each segment is supported by six actuators, providing rigid body motion with six degrees of freedom. In addition, radius-of-curvature adjustment is provided by a seventh actuator in the center of the segment.

The other end of each actuator attaches through a secondary structure and flexure to the backplane. The seventh actuator controls the segment radius of curvature and is independent of the rigid body actuators and the backplane structure. The actuators operate at cryogenic and ambient temperatures, and have both coarse- and fine-positioning capability. The semirigid segments have very low mirror figure changes due to 1-g effects, thus supporting high-fidelity ground testing and a minimum number of actuators for each segment. The secondary mirror has six actuators which provide six degree of freedom rigid body control, although only five degree of freedom control is actually needed.

6.3.3. Integrated Science Instrument Module

The ISIM contains the science instruments for the observatory, and the support electronics for the science instruments: NIRC*am*, NIRS*pec*, TFI, and MIRI (described in Section 6.5). A cryocooler will be used for cooling MIRI and its Si:As detectors. The near-IR detector arrays in the other instruments are passively cooled HgCdTe. In addition to the science instruments, the ISIM contains the FGS and the computer that directs the daily science observations based on plans received from the ground. The science instruments and FGS have nonoverlapping FOVs as shown in Figure 33. Simultaneous operation of all science instruments is possible; this capability will be used for parallel calibration, including darks and possibly sky flats. FGS is used for guide star acquisition and fine pointing. Its FOV and sensitivity are sufficient to provide a greater than 95% probability of acquiring a guide star for any valid pointing direction and roll angle.

Optical Interface: The optical interface to the ISIM is a curved focal surface located approximately 196 cm behind the primary mirror vertex. The ISIM is supported off the rear of the telescope backplane by four kinematic mounts (two bipods and two monopods.) This mounting configuration cleanly partitions optical and mechanical interfaces to simplify integration and test. During normal observations, the FGS provides the pointing error signal used by the telescope and spacecraft to maintain a stable and accurate optical line of sight.

Packaging Interface: The Observatory architecture provides a 19.9 m³ volume for the ISIM instruments, with 8 m² of radiator surface area for cooling the NIR detectors. The focal plane electronic boxes are mounted near the ISIM in a nominal room temperature environment to keep the cable line length below 6 m to ensure cable noise is minimized. The ISIM computers are mounted within the warm spacecraft.

Thermal Interface: In order to achieve a 70% observational efficiency, the science instruments that are not making science observations will make calibration exposures such as darks in parallel. The supporting structure of the instrument payload has very tight optical alignment requirements which necessitate stable constant power dissipation. Therefore, the baseline operations concept calls for all of the instruments to be on all of the time (similar to HST). The ISIM thermal, power, and data handling systems are designed to support simultaneous operation of the four science instruments and the FGS with appropriate margin.

WFS&C interface: The alignment and phasing of the telescope optics require data gathered by the NIRCcam. For WFS&C procedures, there will be six lenses, filters and grisms in the NIRCcam filter wheels. The physical positioning of these optical elements is no more exacting than for science imaging through filters. During the WFS&C process, the NIRCcam will be operated in the same manner as that used for science (i.e., no special modes are necessary). In addition, NIRCcam has a pupil imaging lens which can be used to image the primary mirror.

6.3.4. Wavefront Sensing and Control Subsystem

The ability to perform on-orbit alignment of the telescope is one of the enabling technologies that allows JWST to be built. The large telescope area is made up of 18 smaller segments that can be manufactured and tested more easily than larger ones. The WFS&C subsystem aligns these segments so that their wavefronts match properly, creating a diffraction-limited 6.6-m telescope, rather than overlapping images from 18 individual 1.3 m telescopes.

The 1.3-m semirigid segments limit segment-level wavefront errors to those controllable through fabrication, radius-of-curvature adjustment, and rigid-body positioning. The WFS&C does not require nor use any other deformation of optical-structural surfaces. Such deformation would crosscouple surface corrections and limit the testability of the optics in the ground 1-g environment. The semirigid mirror architecture supports the use of a set of deterministic WFS&C algorithms that control the fundamental and critical initial telescope alignment operation (Acton







<i>First light NIRCam</i>		After Step 1	Initial Capture	Final Condition
	1. Segment Image Capture		18 individual 1.6-m diameter aberrated sub-telescope images PM segments: < 1 mm, < 2 arcmin tilt SM: < 3 mm, < 5 arcmin tilt	PM segments: < 100 μm, < 2 arcsec tilt SM: < 3 mm, < 5 arcmin tilt
2. Coarse Alignment Secondary mirror aligned Primary RoC adjusted	After Step 2		Primary Mirror segments: < 1 mm, < 10 arcsec tilt Secondary Mirror : < 3 mm, < 5 arcmin tilt	WFE < 200 μm (rms)
3. Coarse Phasing - Fine Guiding (PMSA piston)	After Step 3		WFE: < 250 μm rms	WFE < 1 μm (rms)
4. Fine Phasing	After Step 4		WFE: < 5 μm (rms)	WFE < 110 nm (rms)
5. Image-Based Wavefront Monitoring	After Step 5		WFE: < 150 nm (rms)	WFE < 110 nm (rms)

Figure 36. A summary of the WFS&C commissioning and maintenance phases.

et al., 2004). Another benefit is the use of a closed-form reconstructor algorithm to measure and correct segment-to-segment piston errors, a crucial initial operation in creating a phased primary mirror. Determination of the wavefront error and the necessary telescope mirror commands is done on the ground using the downlinked image data. The algorithm uses least-squares fits to images taken in and out of focus, at different wavelengths, and at multiple field points. The focus is adjusted using weak lenses in the NIRCam filter wheel. This algorithm was used to diagnose and measure the spherical aberration of the HST primary mirror. Mirror adjustment commands are uplinked to the observatory. Figure 36 summarizes the WFS&C process.

The initial telescope commissioning process occurs in four phases: (1) image capture and identification, (2) coarse alignment, (3) coarse phasing, and (4) fine phasing. Shi *et al.* (2004) analyzed the WFS&C architecture and implemented hardware and software demonstrations of the more critical aspects of this architecture. Their analyses and simulations show that the residual errors from one step (final condition) are well within the capture range of the next step (initial condition.) The WFS&C process is straightforward, has a singular deterministic solution, and will be fully tested on the ground.

To accomplish commissioning, the telescope is pointed at a celestial region with specific characteristics at the 2 μm NIRCam operation wavelength. A bright (magnitude 10 or brighter) isolated commissioning star is located on the NIRCam detector. A second star (guide star), with magnitude and position restrictions relative to the commissioning star, is used with the fine guiding sensor. There are many star fields that meet these criteria to ensure availability regardless of the JWST launch date (Green *et al.*, 2004).

6.3.5. Sunshield

The sunshield enables the passively-cooled thermal design. It reduces the ~ 200 kW incident solar radiation that impinges on the sunshield to milliwatts incident on the telescope and ISIM. This solar attenuation is a result of the five-layer configuration of the sunshield. Its physical size and shape determine the celestial field of regard for the observatory. By reducing the solar radiation to the milliwatt level, the observatory has an intrinsically stable point spread function as it is pointed over its FOR. The orientation and angular separation of the individual layers can be seen in the artist's rendition of the JWST Observatory in Figure 27.

A sunshield with such a large surface area will result in angular momentum build up from solar radiation. The angle between the forward and aft sunshield was selected to minimize this momentum buildup. The momentum will be dumped by firing the thrusters that are used for orbit maintenance.

6.3.6. Spacecraft Bus

The spacecraft (Figure 37) provides the housekeeping function of the Observatory. It has a 471 Gb solid state recorder to hold the science and engineering data collected between and during the daily contacts with the ground station. The recorder can hold two days of collected data, providing redundancy against a missed contact. Three star trackers (including one for redundancy) are used to point the observatory toward the science target prior to guide star acquisition and to provide roll stability about the telescope line of sight. Six reaction wheels (including two for redundancy) are mounted on isolators near the center of gravity of the bus to reduce disturbances to the observatory. These reaction wheels offload the fine steering control (operating with a 16 Hz update rate from the FGS), maintain the fine steering mirror near a central position, and limit blurring from differential field distortion. Power is provided by two solar arrays that are canted toward the sun when the observatory is pointed in the middle of its FOR. The downlink operates at Ka band and has a selectable rate of 7, 14, or 28 Mbps. A pair of omnidirectional antennas (at S band) provide nearly complete spherical coverage for emergency communications.

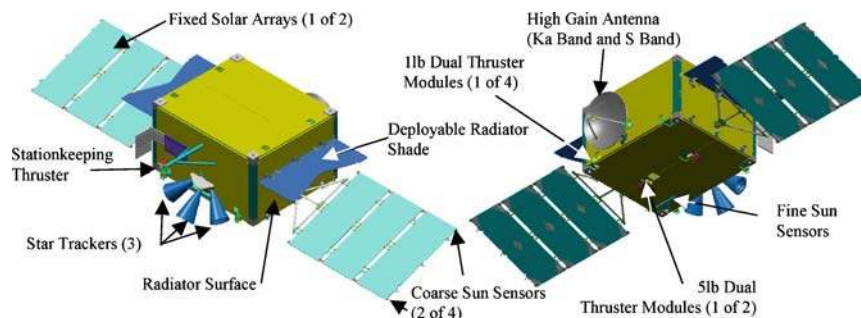


Figure 37. JWST spacecraft bus external configuration.

6.4. INTEGRATION AND TEST

Incremental verification will be used to integrate and test JWST in a way similar to that used for the Chandra X-ray Observatory. Performance parameters are verified incrementally as the observatory is integrated, stopping at the highest level of integration that adequately verifies the system performance. This will culminate in a final end-to-end cryogenic optical test of the observatory.

Once each mirror element, with its actuators installed, successfully completes cryogenic tests, the mirror elements will be attached to the backplane support structure and the telescope aligned (Wells *et al.*, 2004). The telescope and ISIM combination will be cryogenically tested at Chamber A at NASA's Johnson Space Center and verified prior to integration to the spacecraft and sunshield. This telescope/ISIM cryogenic test will be the end-to-end optical performance verification for the Observatory. The test chamber is 16.8 m in diameter and 35.7 m high, giving adequate room to configure cost-effective assembly, test, and verification equipment.

The telescope will be tested with the optical axis vertical (see Figure 38) to minimize gravity effects. Testing the optics in this orientation minimizes the moments induced in the mirrors at the actuator attachment interface. The completed primary mirror will be tested interferometrically at its center of curvature. In addition, a sampled full-aperture test is performed where each segment is sampled in a single wavefront test. The semirigid hexagonal primary mirror architecture allows use of six 1.5 m flats rather than an expensive and high risk 7 m flat or collimator. Another benefit of the sparsely sampled approach is that the autocollimating flats need not be phased to each other. The test is operated using individual triads of three primary mirror segments, allowing the mirror rigid body actuation to correct for phasing of the flats. Only stability of the flats is required.

The sunshield deployment will be tested at room temperature. A scale version of the sunshield will be tested in the deployed configuration with a helium-cooled shroud. By performing a test of the sunshield with a helium-cooled shroud background, we will validate the thermal model of the sunshield at operational temperatures.

6.5. INSTRUMENTATION

Table IX lists the main characteristics of JWST's science instruments. The sensitivities of the instruments are listed in Table X. Although stated as the sensitivity achievable for 10,000 s exposures, it is expected that cosmic ray hits will limit the maximum exposure time for an individual integration to about 1,000 s, and that longer total exposure times will be achieved through coadding. Based on experience with Hubble data, for example, in the Hubble Deep Field, we expect the errors to scale as the square root of the exposure time in coadds as long as 10^5 or even 10^6 s. The absolute photometric accuracy is expected to be 5% for imaging and 10–15%

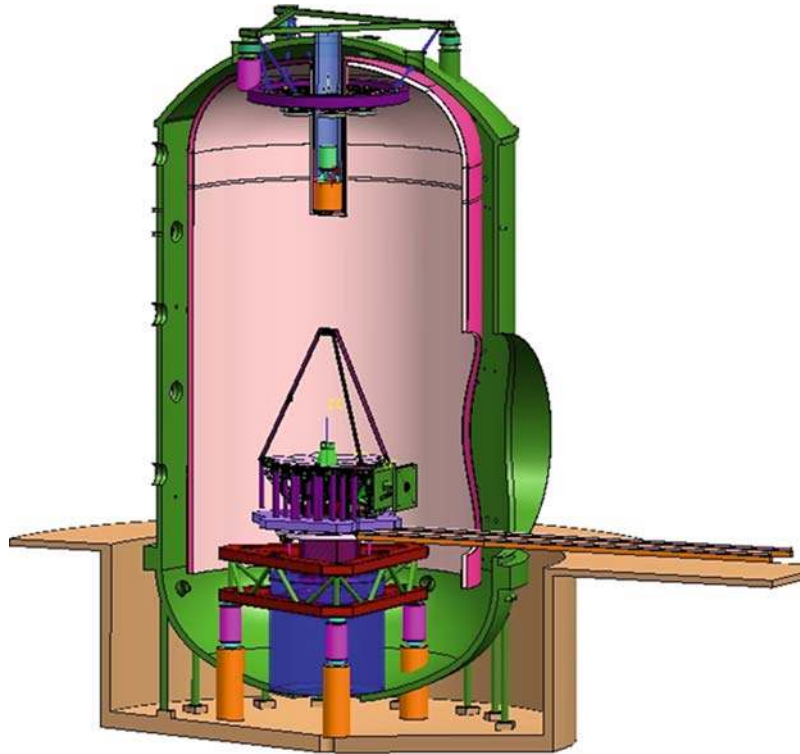


Figure 38. The JWST observatory pictured in the thermal vacuum test chamber at NASA's Johnson Space Center. The test chamber is 16.8 m in diameter and 35.7 m high. The full telescope will be tested at its center of curvature, and a sampled full-aperture test will test the wavefront from each segment.

for coronagraphy and spectroscopy, based on calibration observations of standard stars.

6.5.1. Near-Infrared Camera

NIRCam provides filter imaging in the $0.6\text{--}5.0\,\mu\text{m}$ range with wavelength multiplexing. It includes the ability to sense the wavefront errors of the observatory. NIRCam consists of an imaging assembly within an enclosure that is mounted in the ISIM. The imaging assembly consists of two fully redundant, identical optical trains mounted on two beryllium benches, one of which is shown in Figure 39. The incoming light initially reflects off the pick-off mirror. Subsequently, it passes through the collimator and the dichroic, which is used to split the light into the short ($0.6\text{--}2.3\,\mu\text{m}$) and long ($2.4\text{--}5.0\,\mu\text{m}$) wavelength light paths. Each of these two beams then passes through a pupil wheel and filter wheel combination, each beam having its own pair of pupil and filter wheel. After this, the light passes

TABLE IX
Science instrument characteristics

Instrument	Wavelength(μm)	Detector	Plate scale (milliarcsec/pixel)	Field of view
NIRCam			32	2.2×4.4 arcmin
Short	0.6–2.3	Eight 2048 \times 2048		
Long ^a	2.4–5.0 2048 \times 2048	Two	65	2.2×4.4 arcmin
NIRSpec	0.6–5.0	Two	100	
MSA ^b			2048 \times 2048	3.4×3.1 arcmin
Slits ^c				$\sim 0.2 \times 4$ arcsec
IFU				3.0×3.0 arcsec
MIRI	5.0–29.0	1024 \times 1024	110	
Imaging				1.4×1.9 arcmin
Coronagraphy				26×26 arcsec
Spectra ^d	5.0–10.0			0.2×5 arcsec
IFU	5.0–29.0	Two 1024 \times 1024	200 to 470	3.6×3.6 to 7.5×7.5 arcsec
TFI	1.6–4.9 ^e	2048 \times 2048	65	2.2×2.2 arcmin

Note. ^aUse of a dichroic renders the NIRCam long-wavelength FOV cospatial with the short-wavelength channel, and the two channels acquire data simultaneously.

^bNIRSpec includes a microshutter assembly (MSA) with four 365×171 microshutter arrays. The individual shutters are each 203 (spectral) \times 463 (spatial) milliarcsec clear aperture on a 267×528 milliarcsec pitch.

^cNIRSpec also includes several fixed slits which provide redundancy and high contrast spectroscopy on individual targets, and an IFU.

^dMIRI includes a fixed slit for low-resolution ($R \sim 100$) spectroscopy over the 5–10 μm range, and an integral field unit for $R \sim 3000$ spectroscopy over the full 5–29.0 μm range. The long wavelength cutoff for MIRI spectroscopy is set by the detector performance, which drops longward of 28.0 μm .

^eThe wavelength range for the TFI is 1.6–2.6 μm and 3.1–4.9 μm . There is no sensitivity between 2.6 and 3.1 μm .

through the camera corrector optics and is imaged (after reflecting off a fold flat in the short-wavelength beam) onto the detectors.

The instrument contains a total of ten $2\text{k} \times 2\text{k}$ detector chips, including those in the identical redundant optical trains. The short-wavelength arm in each optical train contains a 2×2 mosaic of these detectors, optimized for the 0.6–2.3 μm wavelength range, with a small gap ($\sim 3 \text{ mm} = \sim 5 \text{ arcsec}$) between adjacent detectors. The detectors arrays are HgCdTe of HAWAII II heritage built by Rockwell Science Center. The detectors will all have thinned substrates to avoid cosmic ray scintillation issues, as well as to extend their sensitivity below 0.85 μm .

TABLE X
Instrument sensitivities

Instrument/mode	λ (μm)	Bandwidth	Sensitivity
NIRCam	2.0	$R = 4$	11.4 nJy, AB = 28.8
TFI	3.5	$R = 100$	126 nJy, AB = 26.1
NIRSpec/Low Res.	3.0	$R = 100$	132 nJy, AB = 26.1
NIRSpec/Med. Res.	2.0	$R = 1000$	$1.64 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$
MIRI/Broadband	10.0	$R = 5$	700 nJy, AB = 24.3
MIRI/Broadband	21.0	$R = 4.2$	$8.7 \mu\text{Jy}$, AB = 21.6
MIRI/Spect.	9.2	$R = 2400$	$1.0 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$
MIRI/Spect.	22.5	$R = 1200$	$5.6 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$

Note. Sensitivity is defined to be the brightness of a point source detected at 10σ in 10,000 s. Longer or shorter exposures are expected to scale approximately as the square root of the exposure time. Targets at the North Ecliptic Pole are assumed. The sensitivities in this table represent the best estimate at the time of submission and are subject to change.

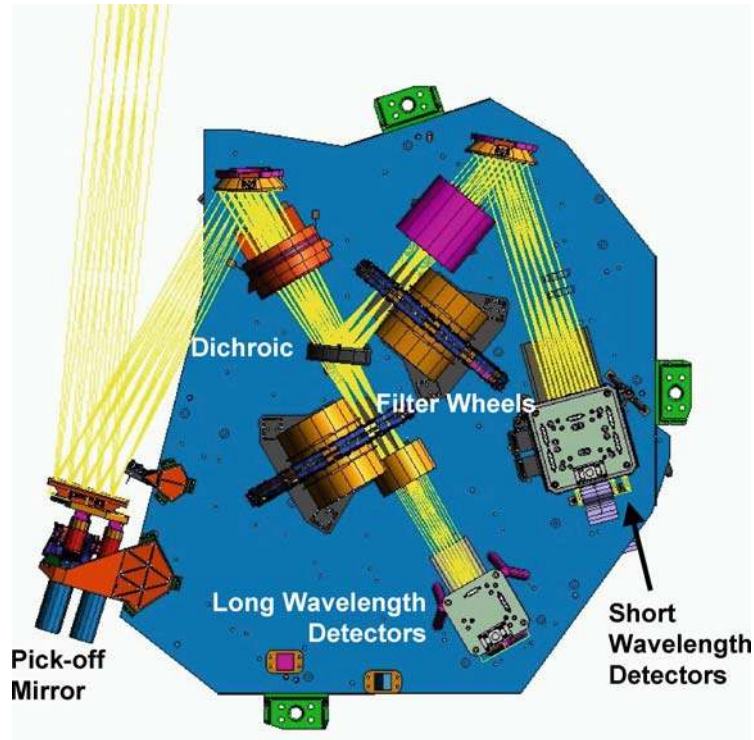


Figure 39. Optical layout of one of two NIRCam imaging modules. A dichroic directs the light into short- and long-wavelength detectors. There are two identical imaging modules, providing redundancy and a large FOV.

TABLE XI
Preliminary list of NIRCcam filters and pupils

Filter wheel	Pupil wheel
Short-wavelength channel	
Broadband 0.7 μm	Imaging pupil
Broadband 0.9 μm	TBD
Broadband 1.15 μm	Flat field pinholes
Broadband 1.5 μm	10% H ₂ O 1.62 μm
Broadband 2.0 μm	10% H ₂ O 1.40 μm
10% CH ₄ 2.10 μm	WFS DHS 1
1% H ₂ 2.12 μm	WFS weak lens 1
WFS 1.15–2.3 μm	WFS weak lens 2
WFS weak lens 3	Outward pinholes
1% H ₂ 2.25 μm /1% [FeII] 1.644 μm ^a	Coronagraphic pupil 1 with wedge
10% H ₂ O 1.82 μm	Coronagraphic pupil 2 with wedge
1% P α 1.875 μm	WFS DHS 2
Long-wavelength channel	
Broadband 2.7 μm	Imaging pupil
Broadband 3.6 μm	Flat field pinholes
Broadband 4.4 μm	1% H ₂ 3.23 μm
10% H ₂ O 3.00 μm	1% H ₂ 4.18 μm
10% PAH 3.35 μm	1% CO 4.60 μm
10% 3.60 μm	1% H ₂ 4.69 μm
10% 3.90 μm	TBD
10% CO ₂ 4.30 μm	Outward pinholes
10% CO 4.60 μm	WFS grism 1
10% 4.80 μm	WFS grism 2
1% Br α 4.05 μm	Coronagraphic pupil 1
10% CH ₄ 2.50 μm	Coronagraphic pupil 2

Note. WFS: Wavefront sensing optics; DHS: Dispersed Hartmann sensing optics; PAH: Polycyclic aromatic hydrocarbon; TBD: To be determined.

^aThis position will have two different filters in the two channels of NIRCcam, resulting in half the FOV in each filter.

Each optical train contains a dual filter and pupil wheel, containing a range of wide-, medium-, and narrow-band filters and the WFS&C optics. Each wheel has 12 slots. The preliminary filter selection is given in Table XI.

- *Coronagraphy:* To enable the coronagraphic imaging, each of the two identical optical trains in the instrument also contains a traditional focal plane coronagraphic mask plate held at a fixed distance from the detectors, so that

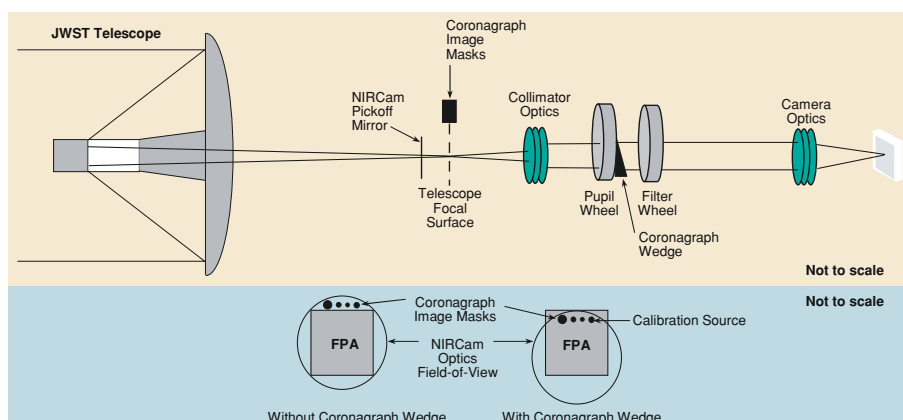


Figure 40. Schematic of NIRCcam coronagraph design. An optical wedge in the pupil wheel brings the coronagraphic spots into the FOV. The spots are matched with Lyot stops.

the coronagraph spots are always in focus at the detector plane. Each coronagraphic plate is transmissive, and contains a series of spots of different sizes, including linear and radial-sinc occulters, to block the light from a bright object. Each coronagraphic plate also includes a neutral density spot to enable centroiding on bright stars, as well as calibration sources at each end that can send light through the optical train of the imager to enable internal alignment checks. Normally these coronagraphic plates are not in the optical path for the instrument, but they are selected by rotating into the beam a mild optical wedge that is mounted in the pupil wheel (see Figure 40), which translates the image plane so that the coronagraphic masks are shifted onto the active detector area. Diffraction can also be reduced by apodization at the pupil mask, thus the pupil wheels will be equipped with both a classical and an apodized pupil with integral wedges in each case. Current models predict a contrast of $\sim 10^4$ at 0.5 arcsec, at a wavelength of $4.6 \mu\text{m}$.

6.5.2. Near-Infrared Spectrograph

NIRSpec is a near-IR multiobject dispersive spectrograph capable of simultaneously observing more than 100 sources over a FOV larger than 3×3 arcmin. In addition to the multiobject capability, it includes fixed slits and an integral field unit for imaging spectroscopy. Six gratings will yield resolving powers of $R \sim 1000$ and ~ 2700 in three spectral bands, spanning the range $1.0\text{--}5.0 \mu\text{m}$. A single prism will yield $R \sim 100$ over $0.6\text{--}5.0 \mu\text{m}$. Figure 41 shows a layout of the instrument.

The region of sky to be observed is transferred from the JWST telescope to the spectrograph aperture focal plane by a pick-off mirror and a system of fore-optics that includes a filter wheel for selecting bandpasses and introducing internal calibration sources.

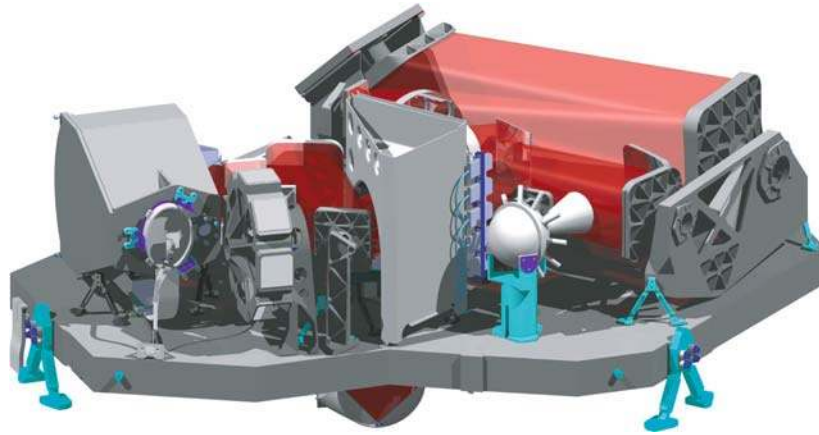


Figure 41. The NIRSpec instrument.

Targets in the FOV are normally selected by opening groups of shutters in a microshutter assembly (MSA) to form multiple apertures. The microshutter assembly itself consists of a mosaic of four subunits producing a final array of 730 (spectral) by 342 (spatial) individually addressable shutters with 203×463 milliarcsec openings and 267×528 milliarcsec pitch. Sweeping a magnet across the surface of the microshutters opens all of the shutters. Individual shutters may then be addressed and released electronically, and the return path of the magnet closes the released shutters. The minimum aperture size is 1 shutter (spectral) by 1 shutter (spatial) at all wavelengths. Multiple pointings may be required to avoid placing targets near the edge of a shutter and to observe targets with spectra that would overlap if observed simultaneously at the requested roll angle. The nominal slit length is 3 shutters in all wavebands. In the open configuration, a shutter passes light from the fore-optics to the collimator. A slitless mode can be configured by opening all of the microshutters. As the shutters are individually addressable, long slits, diagonal slits, Hadamard-transform masks (Riesenberger and Dillner, 1999), and other patterns can also be configured with them.

In addition to the slits defined by the microshutter assembly, NIRSpec also includes five fixed slits that can be used for high-contrast spectroscopy. They are placed in a central strip of the aperture focal plane between subunits of the microshutter assembly and also provide redundancy in case the microshutters fail. Three fixed slits are 3.5 arcsec long and 200 milliarcsec wide. One fixed slit is 4 arcsec long and 400 milliarcsec wide for increased throughput at the expense of spectral resolution. One fixed slit is 2 arcsec long and 100 milliarcsec wide for brighter targets.

The strip between microshutter subunits also contains the 3×3 arcsec entrance aperture for an integral field unit (IFU). The IFU has 30 slices, each 100 milliarcsec wide. The IFU relay optics introduce a 2:1 anamorphic magnification of each slicer

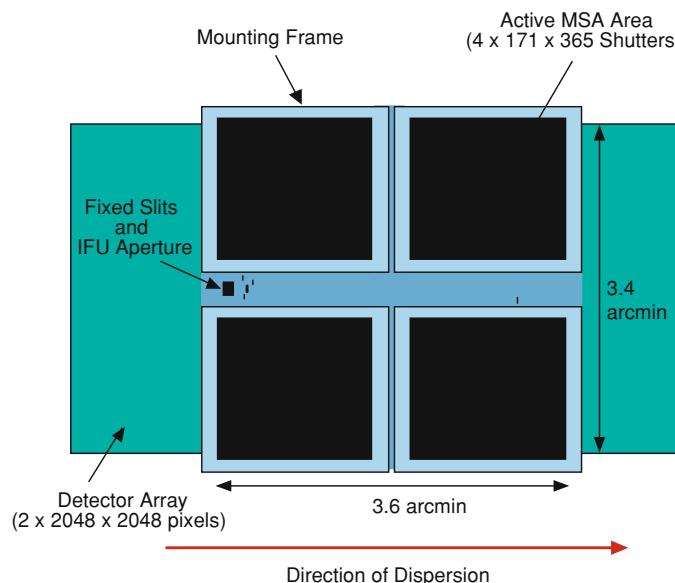


Figure 42. Schematic layout of the NIRSpec slit mask overlaid on the detector array and projected to the same angular scale.

such that the matching projected virtual slits are properly sampled on the detector by two 50 milliarcsec pixels in the dispersion direction and at the (nominal) 100 milliarcsec per pixel in the spatial direction.

The aperture focal plane is reimaged onto a mosaic of two Near-IR detectors by a collimator, a dispersing element (gratings or a double-pass prism) or an imaging mirror, and a camera. Three gratings are used for first-order coverage of the three NIRSpec wavebands at $R \sim 1000$ ($1.0\text{--}1.8\ \mu\text{m}$; $1.7\text{--}3.0\ \mu\text{m}$; $2.9\text{--}5.0\ \mu\text{m}$). The same three wavebands are also covered by first-order $R \sim 2700$ gratings for objects in a fixed slit or in the IFU. The prism gives $R \sim 100$ resolution over the entire NIRSpec bandpass ($0.6\text{--}5\ \mu\text{m}$) but can optionally be blocked below $1\ \mu\text{m}$ with one of the filters. Any of the aperture selection devices (microshutter assembly, fixed slits, or IFU) can be used at any spectral resolution.

The focal plane array is a mosaic of two detectors (see Figure 42), each $2\text{k} \times 2\text{k}$, forming an array of $2\text{k} \times 4\text{k}$ 100 milliarcsec pixels. The detectors will be thinned HgCdTe arrays built by Rockwell Science Center. NIRSpec contains a calibration unit with a number of continuum and line sources.

6.5.3. Mid-infrared Instrument

The MIRI on JWST provides imaging and spectroscopic measurements over the wavelength range $5\text{--}29\ \mu\text{m}$. MIRI consists of an optical bench assembly (Figure 43) with associated instrument control electronics, actively cooled detector modules with associated focal plane electronics, and a cryocooler with associated control

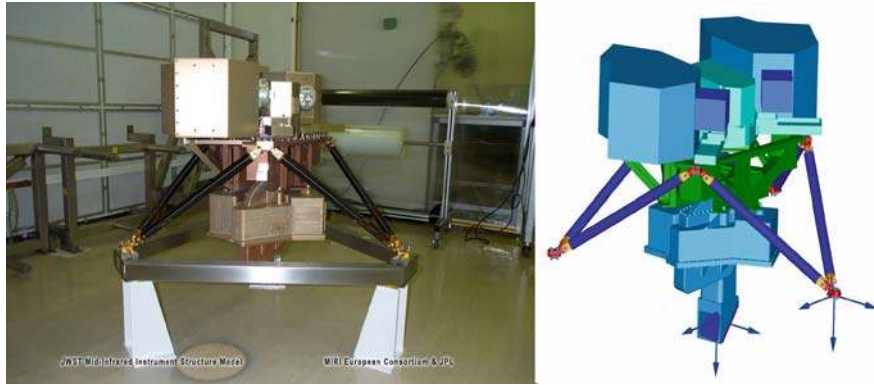


Figure 43. The MIRI structural and thermal model (left) compared to a computer design of the instrument (right).

electronics. The cryocooler electronics interface with the spacecraft command and telemetry processor, while the instrument control electronics interface with the ISIM command and data handling. The optical bench assembly contains two actively cooled subcomponents, an imager and an Integral Field Unit (IFU) spectrograph, plus an on-board calibration unit.

Imaging: The imager module provides broadband imaging (see Table XII for preliminary filter selection), coronagraphy, and low-resolution ($R \sim 100$, $5\text{--}10\ \mu\text{m}$) slit spectroscopy using a single 1024×1024 pixel Raytheon Si:As detector with $25\ \mu\text{m}$ pixels. Figure 44 shows the focal plane arrangement of the elements of the

TABLE XII
MIRI filters

	$\lambda_0\ (\mu\text{m})$	$\Delta\lambda\ (\mu\text{m})$	Comment
B1	5.6	1.2	Broadband
B2	7.7	2.2	PAH, broadband
B3	10	2	Silicate, broadband
I1	11.3	0.7	PAH, broadband
I2	12.8	2.4	Broadband
B4	15	3	Broadband
I3	18	3	Silicate, broadband
B5	21	5	Broadband
B6	25.5	~ 4	Broadband
B6'	25.5	~ 4	Redundant filter, risk reduction
ND#	Neutral density		Coronagraphic acquisition
NIR	Near-IR, TBD		Testing
	Blackened blank	N/A	Darks

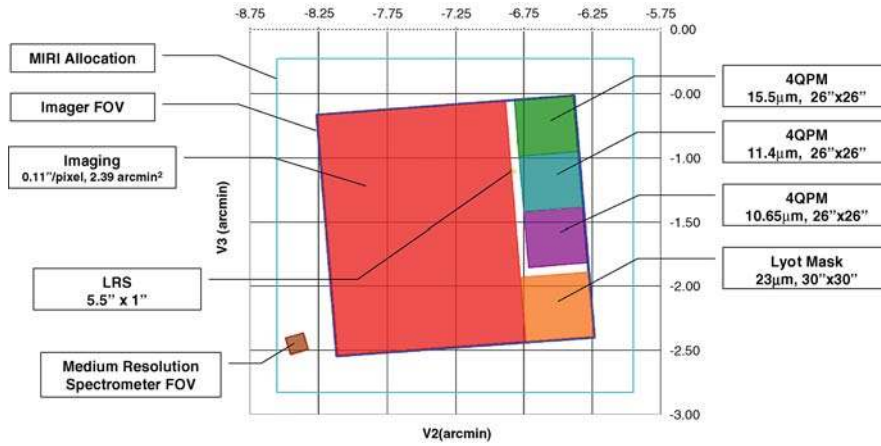


Figure 44. MIRI's imaging focal plane. One quarter of the FOV is devoted to coronagraphy, there is a 5 arcsec wide low-resolution spectroscopy strip, and the remaining $1.4 \times 1.9 \text{ arcmin}^2$ is available for broadband imaging.

MIRI imager. The region on the left is the clear aperture available for imaging. The gray region on the right marks the mechanical frame that supports the coronagraphic masks and the slit for the low-resolution spectrometer. The coronagraphic masks include three phase masks for a quadrant-phase coronagraph and one opaque spot for a Lyot coronagraph. The coronagraphic masks each have a square FOV of $26 \times 26 \text{ arcsec}$ and are optimized for particular wavelengths. The imager's only moving part is an 18-position filter wheel. Filter positions are allocated as follows: 12 filters for imaging, four filter and diaphragm combinations for coronagraphy, one ZnS-Ge double prism for the low-resolution spectroscopic mode, and one dark position. The imager will have a pixel scale of $0.11 \text{ arcsec/pixel}$ and a total FOV of $113 \times 113 \text{ arcsec}$; however, the FOV of its clear aperture is $84 \times 113 \text{ arcsec}$ because the coronagraph masks and the low-resolution spectrograph are fixed on one side of the focal plane.

Integral-Field Spectroscopy: The integral-field spectrograph obtains simultaneous spectral and spatial data on a small region of sky. The spectrograph FOV is next to that of the imager so that accurate positioning of targets will be possible by locating the image with the imager channel and off-setting to the spectrograph. The light is divided into four spectral ranges by dichroics, and two of these ranges are imaged onto each of two detector arrays. A full spectrum is obtained by taking exposures at each of three settings of the grating wheel. The spectrograph uses four image slicers to produce dispersed images of the sky on two 1024×1024 detectors, providing $R \sim 3000$ integral-field spectroscopy over the full $5\text{--}29 \mu\text{m}$ wavelength range, although the sensitivity of the detectors drops longward of $28 \mu\text{m}$. As shown in Figure 45 and Table XIII, the IFUs provide four simultaneous and concentric

TABLE XIII
MIRI integral field spectroscopy parameters

Wavelength (μm)	Pixel size (arcsec)	Slice width (arcsec)	Slices #	FOV (arcsec^2)	Resolving power
5.0–7.7	0.196	0.176	22	3.00×3.87	2,400–3,700
7.7–11.9	0.196	0.277	16	3.50×4.42	2,400–3,600
11.9–18.3	0.245	0.387	16	5.20×6.19	2,400–3,600
18.3–28.3	0.273	0.645	12	6.70×7.73	2,000–2,400

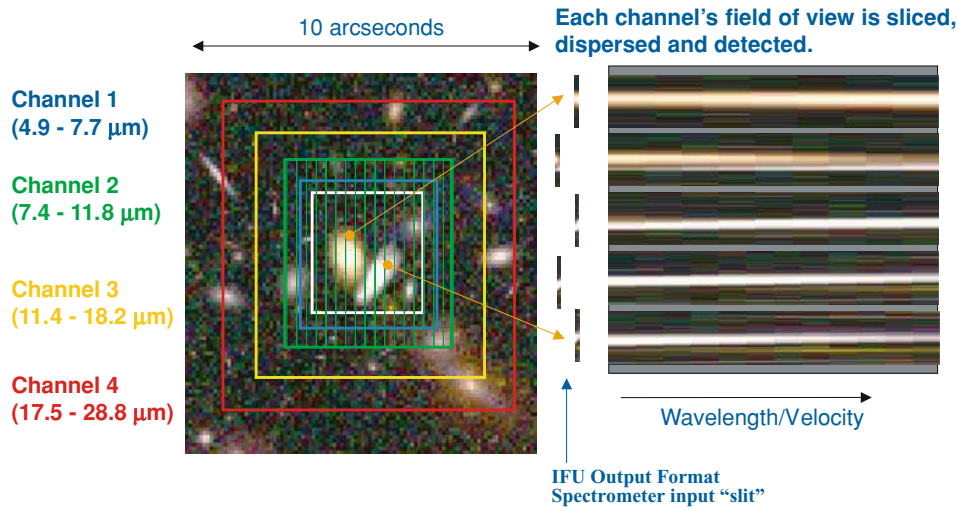


Figure 45. Schematic illustration of the MIRI IFU image slicer format (*left*) and dispersed spectra on detector (*right*).

fields of view. The slice widths set the angular resolution perpendicular to the slice. The pixel sizes set the angular resolution along the slice.

The spectral window of each IFU channel is covered using three separate gratings (i.e., 12 gratings to cover the four channels). Each grating is fixed in orientation and can be rotated into the optical path using a wheel mechanism (there are two wheel mechanisms which each hold three pairs of gratings). The optics system for the four IFUs is split into two identical sections (in terms of optical layout). One section is dedicated to the two short-wavelength IFUs and the other handles the two longer wavelength IFUs, with one detector for each section. The two sections share the wheel mechanisms (each mechanism incorporates three gratings for one of the channels in the short-wavelength section and three gratings for one of the channels in the long-wavelength section). As shown in Figure 45, the image slicers in the MIRI IFU dissect the input image plane. The dispersed spectra from two

IFU inputs are placed on one detector side-by-side. The spatial information from the IFU is spread out into two adjacent rows with the information from each slice separated by a small gap.

Coronagraphy: MIRI features a coronagraph designed for high-contrast imaging in selected mid-IR bandpasses. Three quadrant phase masks (Boccaletti *et al.*, 2004; Gratadour *et al.*, 2005) provide high-contrast imaging to an inner working angle of λ/d , with bandpass of $\lambda/20$, centered at $10.65\ \mu\text{m}$, $11.4\ \mu\text{m}$, and $15.5\ \mu\text{m}$ respectively. A fourth, traditional Lyot mask of radius $0.9\ \text{arcsec}$, will provide $R \sim 5$ imaging at a central wavelength of $23\ \mu\text{m}$. Simulations predict that the quadrant phase masks will achieve a contrast of $\sim 10^4$ at $3\lambda/D$. The Lyot stop is predicted to deliver a contrast of 2×10^3 at $3\lambda/D$.

6.5.4. Tunable Filter Imager

The TFI provides narrow-band near-IR imaging over a FOV of $2.2 \times 2.2\ \text{arcmin}^2$ with a spectral resolution $R \sim 100$. The etalon design allows observations at wavelengths of $1.6\text{--}2.6\ \mu\text{m}$ and $3.1\text{--}4.9\ \mu\text{m}$, although this design is still preliminary. The gap in wavelength coverage allows one channel to reach more than one octave in wavelength. Characteristics of the TFI are listed in Table XIV.

The TFI uses dielectric-coated Fabry–Perot etalon plates with a small air (vacuum) gap. The finesse is about 30 and the filters are used in third order. The finesse

TABLE XIV
Tunable filter imager characteristics

Parameter	Value	Comments
Wavelength range	$1.6\text{--}4.9\ \mu\text{m}$	Gap in coverage at $2.6\text{--}3.1\ \mu\text{m}$
Field of view	$2.2 \times 2.2\ \text{arcmin}$	Half the FOV of NIRCcam
Pixel size	65 milliarcsec	Nyquist sampled at $4.2\ \mu\text{m}$
Detector array	2048×2048	Same as NIRCcam long-wavelength detectors
Spectral resolution	$R > 80$	Etalon intrinsic resolution higher
Clear aperture	56 mm	Pupil size $\sim 40\ \text{mm}$. Clear aperture set by etalon location
Finesse	~ 30	Minimizes # of blocking filters
Surface figure (P–V)	$< 30\ \text{nm}$	Coated etalon surface figure must support reflectance finesse
Transmittance	$> 75\%$	Set by achieved surface figure
Contrast	> 100	Peak spectral transmittance divided by minimum between peaks
Passband shift	$< 1\%$	Typical designs have $< 0.5\%$
Blocking filters	< 8	Minimize filter wheel size and simplify operations

Note. From Rowlands *et al.*, 2004a,b.

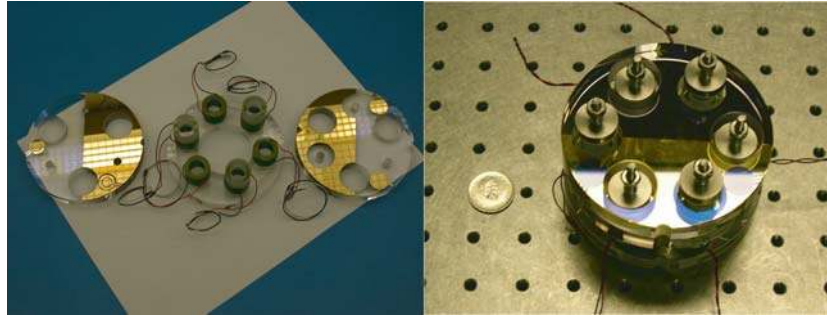


Figure 46. Prototype etalon structure for TFI. Disassembled (*left*) and assembled (*right*). The disassembled view shows the gold coatings for diagnostic measurements (*half circles*) and the gold coatings forming the capacitive displacement pads (*small circles*) (From Rowlands *et al.*, 2004b).

was chosen to be a compromise between the surface figure requirements and the need to minimize the number of blocking filters, while providing a contrast ratio of about 100. The filters are scanned using piezoelectric actuators, consisting of lead–zirconia–titanate ceramic transducers. The air (vacuum) gap ranges from 2.0 to $8.0\ \mu\text{m}$ plate spacings. The Fabry–Perot operates at a nominal temperature of $\sim 35\ \text{K}$.

To demonstrate the feasibility of a cryogenic Fabry–Perot etalon that meets the requirements, Rowlands *et al.* (2004b) fabricated a prototype. The etalon surface figure is the most critical requirement, and is influenced by the coatings, as well as structural and cryogenic issues (Figure 46).

The TFI incorporates four coronagraphic occulting spots permanently to one side of the FOV, and occupying a region $20 \times 80\ \text{arcsec}$. A set of selectable apodization masks is located at the internal pupil images of each channel by the filter wheels. The coronagraph will deliver a contrast ratio of $\sim 10^4$ (10σ) at 1 arcsec separation. The sensitivity is limited by speckle noise. Contrast ratios of 10^5 may be achievable at sub-arcsec scales using roll or spectral deconvolution techniques (Sparks and Ford, 2002; Doyon *et al.*, 2004).

6.5.5. Fine Guidance Sensor

The FGS instrument uses two FOVs in the JWST focal plane to provide fine guidance for the telescope. The FOV locations are chosen to provide optimum lever arms to all instruments for roll about a single guide star. Roll is sensed by star trackers on the spacecraft bus.

The FGS consists of an optical assembly and a set of focal plane and instrument control electronics. The optical assembly of the FGS instrument consists of two channels, imaging separate regions of the sky onto independent $2\text{k} \times 2\text{k}$ detectors. The detectors will be HgCdTe, similar to those in NIRC*am*. The plate scale is 67 milliarcsec/pixel and the field of view is $2.3 \times 2.3\ \text{arcmin}^2$.

The FGS will provide continuous pointing information to the observatory that is used to stabilize the line of sight, allowing JWST to obtain the required image quality. Each of the independent FGS channels provides $>90\%$ probability of obtaining a useable guide star for any observatory pointing and roll angle. With both channels, the probability is $>95\%$. The wavelength region and pixel size have been optimized so that in fine-guidance mode the FGS will provide pointing information to a precision of <5 milliarcsec updated at 16 Hz. The guide stars will be chosen from existing catalogs with $AB < 19.5$ mag in the J band (see Section 6.7.1). Fainter stars may be used for location identification. In the event that a suitable guide star is not available for a particular desired pointing and roll angle combination, alternate choices of roll angles can be considered during scheduling.

JWST will be capable of relative pointing offsets with an accuracy of 5 milliarcsec rms, which will enable subpixel dithering and coronagraphic acquisition. Absolute astrometric accuracy will be limited to 1 arcsec rms by the accuracy of the guide star catalog.

JWST is expected to be capable of moving object tracking, at rates up to $30 \text{ milliarcsec s}^{-1}$ relative to the fixed guide stars, although some degradation of the image quality may occur. The JWST Project is currently doing a cost-benefit analysis of this capability. There is no expectation of the ability to follow curved trajectories, to track objects continuously as the guide stars cross sensor chip boundaries, or to observe moving targets in special orientations.

The FGS is designed to be completely redundant in terms of the guiding function. The loss of any single component would at most result in the loss of one FGS channel. This would reduce the probability of guide star acquisition to $\sim 90\%$ if using current catalogs.

6.6. LAUNCH, ORBIT, AND COMMISSIONING

6.6.1. *Launch and Orbit*

JWST will be launched on an Arianespace Ariane 5 Enhanced Capability-A rocket into orbit about the second Lagrange (L2) point in the Earth-Sun system, approximately 1.5×10^6 km from the Earth. The orbit (shown in Figure 47) lies in a plane out of, but inclined slightly with respect to, the ecliptic plane. This orbit avoids Earth and Moon eclipses of the Sun, ensuring continuous electrical power. JWST will have a 6 month orbit period about the L2 point in the rotating coordinate system moving with the Earth around the Sun. Station-keeping maneuvers are performed after the end of each orbit determination period of 22 days.

The observatory fits into launch vehicles with a 5 m fairing (Figure 47), such as the planned Ariane 5. In the figure, the left view shows the sunshield stowed about the folded telescope. The right view eliminates the sunshield to show the optical telescope in its folded configuration. Because the secondary mirror faces down during launch, it is shielded from particulate redistribution during launch. This

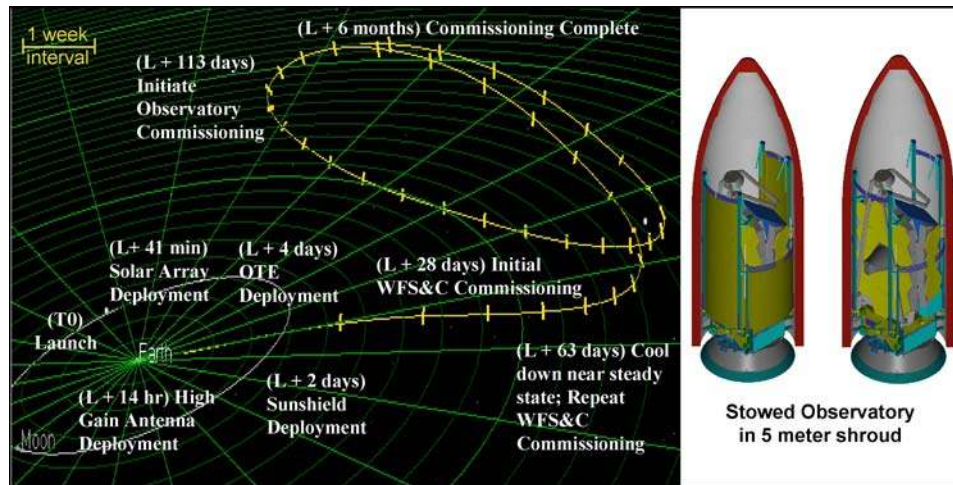


Figure 47. JWST orbit and trajectory to L2, and stowed view in 5 m shroud. The deployments will be complete 4 days after launch, the spacecraft will arrive in the vicinity of L2 in about a month, and commissioning will be complete by 6 months after launch. JWST will carry propellant sized for a 10 yr mission.

configuration with the stowed sunshield protecting the folded telescope reduces contamination which would degrade sensitivity and increase stray light. The total observatory mass is 6500 kg, including station-keeping propellant sized for a 10 yr lifetime.

6.6.2. Deployment

The observatory has the following five deployments: (1) deploy spacecraft bus appendages, including solar arrays and the high-gain antenna, (2) deploy sunshield, (3) extend telescope tower, (4) deploy secondary mirror support structure, and (5) deploy primary mirror wings. The deployment mechanism design includes heaters and other protections that eliminate the need for time-critical events and allow for unlatching and relatching to relieve any residual long-term stress in the structure. The secondary-mirror deployment and primary-mirror deployment sequences are shown in Figure 48.

6.6.3. Commissioning

There will be precommissioning activities during the transfer to L2. About 28 days after launch (Figure 47), the observatory will cool down to a temperature that permits precommissioning activities to begin. Almost continuous ISIM availability for preliminary science observations is provided from this time until 106 days after launch, when a final trajectory burn is required to achieve orbit about L2. The intrinsic passive-thermal stability of the semirigid mirror segment and the two-chord-fold primary allows early operation of science instruments, by providing

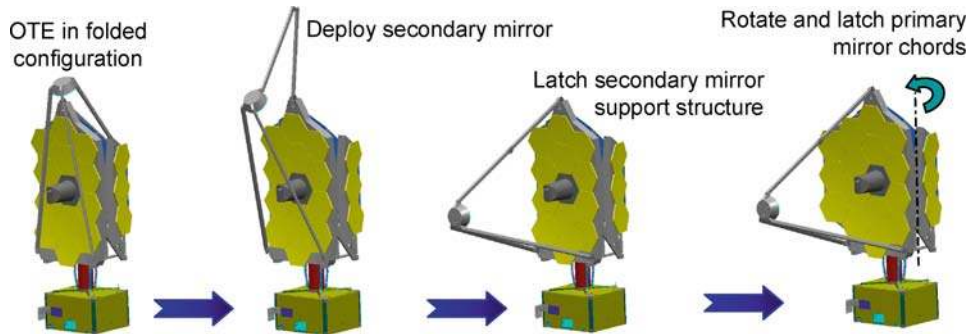


Figure 48. Telescope Deployment Sequence. In deployment steps 4 and 5, as described in the text, the secondary and primary mirrors are deployed. These deployments will be followed by WFS&C operations to align the primary mirror segments.

a stable optical image to the ISIM. These precommissioning activities will develop an operational experience database that allows formal commissioning and science operations to be conducted efficiently.

A final checkout of all systems is initiated after the L2 orbit is achieved. Commissioning (complete 6 months after launch) includes ISIM, telescope, sunshield, and spacecraft operations that were not feasible during the transfer to L2. It also repeats selected operations performed during the transfer, to ensure adequate knowledge of system performance in the final orbital–thermal conditions and allow comparisons with previous measurements. Examples of repeated operations include final optical distortion mappings for the observatory and the characterization of the WFS&C actuators’ transfer function. Allotting 76 days for ISIM precommissioning activities allows for operations rehearsal and increases familiarity with observatory operations without reducing operational availability.

6.6.4. Mission Lifetime

JWST will operate with all science instruments for at least 5 yr after completion of commissioning. In order to exploit the full scientific potential of the mission, a lifetime of 10 yr or longer is desired. Although we will not require mission assurance to guarantee a lifetime greater than 5 yr, JWST will maintain the possibility of a longer mission lifetime, and will carry propellant sized for at least 10 yr of operation after launch. There are no other consumables which would limit lifetime.

6.7. OPERATIONS

JWST will be operated from a Science and Operations Center (S&OC) located at the Space Telescope Science Institute (STScI¹), the organization that operates HST for NASA. From experience with HST, Spitzer, and other astronomical missions,

¹STScI is operated for NASA by the Association of Universities for Research in Astronomy, Inc. (AURA).

NASA has learned that operations can be expensive and inefficient if the scientists and engineers who will operate an observatory for NASA are not involved in the design and development of the mission, and so STScI is supporting JWST during the design and development phases.

Although the capabilities of JWST are being developed to address science themes discussed in the previous sections, almost all of the observing time on JWST will be competitively selected. Approximately 10% of the observing time for the first 5 yr of the mission has already been awarded to the science instrument teams and to other members of the science working group. An additional 5% will be director's discretionary time allocated by the director of the S&OC. The remaining $\sim 85\%$ of the observing time will be awarded through a series of proposal solicitations, which will be open to any astronomer in the world. The proposal solicitations will begin a year or two before the anticipated launch of JWST. This will take advantage of the scientific progress that has been made in the intervening time both from ground-based and spaced-based observatories. The scope of JWST's competitively-selected investigations will range from large legacy-style projects, which last months and address a range of science goals simultaneously, to small programs that target important but very specific science objectives. To maximize JWST's science productivity, the operations concept has to make JWST easily accessible to those who will use it.

6.7.1. *Operations Concept*

JWST will be located at L2 in order to allow effective cooling of the telescope and ISIM. On HST and other low-Earth orbiting satellites about half of the heat load is reflected light and thermal radiation from the Earth. Other orbits for JWST were considered during the design phases of the mission, including a drift-away orbit similar to the one utilized by Spitzer, and highly elliptical Solar orbits that could have taken a smaller telescope out to the orbit of Jupiter. The advantage of L2 is that the distance to the observatory is approximately the same throughout the mission, so the data rate to and from the observatory can be maintained with the same large ground antennae and 1 m class downlink antennae.

The L2 orbit creates a number of simplifications for operations of a general-purpose astronomical observatory. Unlike HST's low-Earth orbit where target visibilities are interrupted every 95 min by earth occultation, astronomical targets from an L2 orbit are visible for long periods of time once or twice a year. Targets near the ecliptic plane are visible for 2 months twice a year; targets within 5 degrees of the ecliptic pole are visible continuously throughout the year. (In low-Earth orbit, observatories such as HST and FUSE have continuous viewing zones but these zones actually move on the time scale of the orbital precession period of ~ 57 days.) The continuous viewing zone of JWST will enable long-term monitoring of SNe and other time-variable objects, and will allow us to establish relative calibration standards that can be observed anytime.

L2 is a saddle point in the gravitational potential of the Earth–Sun system, and as a result propellant is required to maintain JWST at this position. In addition, solar pressure on the sunshield causes torques on the spacecraft that cause the momentum wheels that control the spacecraft pointing to spin up. In low-Earth orbit, the angular momentum that builds up (primarily from the residual atmospheric drag) is usually dumped using devices that interact with the Earth’s geomagnetic field. At L2, this momentum buildup requires thruster firings. Since propellant is limited, its usage must be carefully monitored to meet the design goal of a 10 yr operational lifetime.

From an operations perspective, the instruments on JWST are comparable in complexity to those on HST, but more complex than those on Spitzer. This complexity is needed to accomplish the science mission of JWST, but has the potential to make the Observatory costly to operate if it translates into large numbers of observing modes. Therefore, the S&OC, the instrument developers, and NASA expect to provide a relatively small number of standard modes for each instrument for use by observers, with a limited selection of readout and dither patterns. Limiting the number of observing modes also allows better calibrations of the instruments within the limited resources of the JWST lifetime. Limiting the number of modes is also important for assuring the quality of the archive, and the ease with which archived observations can be used for science other than that intended by the original observer. From this perspective, it is fortunate that the basic readout architecture of all of the detectors on JWST is very similar, if not identical.

The acquisition of targets with JWST will differ in detail from that of HST. The spacecraft will be able to position the observatory to an accuracy of about 5–7 arcsec using its star trackers, but targets must be positioned to accuracies of (in some cases) 5 milliarcsec relative to the instruments’ apertures. The JWST optical system cannot support large-field bore-sight trackers, so the guide stars are relatively faint. The new version of the guide star catalog (GSC) originally developed for Hubble acquisitions, the so-called GSC II (e.g., Spagna *et al.*, 2004), will be used for JWST. GSC II extends to $AB \sim 19$ mag in the J band, at which magnitude the areal density of guide stars is high enough for the two 2.3×2.3 arcmin² fields of the FGS to reliably contain at least one guide star with over 95% confidence anywhere on the sky. To reduce the possibility of acquisition failures, several candidate guide stars will be included in command loads to the FGS. Once a suitable guide star has been located, acquisition of the science targets with the NIRCам, MIRI, and TFI will be very similar to acquisition procedures used for HST’s imaging, long-slit, and coronagraphic applications. For fine-target positioning, this involves obtaining one or more acquisition images with the science instrument, finding the centroid of the desired target, and using the centroid position to offset the science target to a specific position, such as the center of a slit in the science instrument. NIRSpec acquisitions (and observation planning) will likely be the most complex since the programmable microshutter array is intended to allow an astronomer to obtain spectra of up to 100 discrete objects simultaneously. This means that one must accurately determine both the desired offset and the desired roll from acquisition images obtained through the

microshutter array and multistar centroiding algorithms on-board the spacecraft. In contrast, the imaging modes will not require an acquisition sequence, as the 1 arcsec positioning accuracy of the FGS will be sufficient.

The observatory is designed to require very little real-time commanding. This is an important factor in reducing mission costs because continuous communications with the spacecraft will not be required. Fewer people will be needed to staff the S&OC, and after commissioning, S&OC operations will be sufficiently automated so that staffing only is needed during a normal workweek.

6.7.2. *Event-Driven Architecture*

Commanding HST and most other low-Earth-orbiting satellites is based upon absolute time. Command loads are uplinked and the spacecraft's computer executes each command at a specific absolute time. In some cases, a sequence of relative time commands is issued, but the underlying principle of this approach is that one can determine exactly when commands start and stop. This commanding concept was well suited to a time when on-board computers were extremely primitive by modern standards, and is still appropriate in situations where external factors, such as regular Earth occultation in low-Earth orbit of an inertially-pointed satellite, require that many commands be issued at specific times.

JWST's observing program will have few absolute time requirements. Earth blockage is not a problem at L2. Data downlinks and command uplinks will not interrupt science observations, as is the case for Spitzer, since JWST will have a gimbaled antenna. Real-time commanding will be used only for critical operations, such as station-keeping maneuvers or recovery from anomalies. JWST will use a different commanding approach that will be event driven rather than absolute-time driven. The commanding concept for JWST is like a command queue; commands in the queue are executed sequentially and the next command starts when an indication is received that the previous command is completed (successfully or unsuccessfully). This approach sacrifices certainty, but simplifies the software system that is needed for operations planning and lowers operations cost, both in the mission's development and operational phases. It avoids the need to accurately model execution times for every command and then to update that model as the flight software and hardware evolve. It also improves observational efficiency since a basic tenet of absolute-time commanding is that if a command does not complete normally, the spacecraft must either ignore the failure or wait until the next observation is scheduled. For JWST, that wait could be as long as 24 h. By contrast, if an observation fails in an event-driven approach (for example, by failing to acquire a guide star), the observatory can move on to the next observation or the next target. Similarly, with an event-driven approach, if a single instrument goes into a failure or "safe" mode, observations with that instrument are skipped and observations with other instruments are brought forward in time. This allows JWST to continue to conduct high-quality science without loss of overall efficiency.

6.7.3. *Visits and Observation Plan*

The event-driven architecture for JWST will be implemented through a construct of visits and an observation plan. A visit is simply a logically grouped series of activities, along with a set of conditions that can be identified at the beginning of the series and which must be met for the visit to be executed. Typically, a visit would comprise a slew to a new target, the steps required to finely point at the target, the setup of the instrument for the observation itself, and the acquisition of all of the science data at that pointing position. The constraints for executing the visit would include the availability of the prime science instrument for the observation and a time window during which the visit had to start. The observation plan constitutes the single queue that orders the visits. Software in the on-board computer executes the observation plan. Although the visit activities can be relatively complex, the role of the software is to initiate the next visit after the previous visit is complete, to wait until the start time for the next visit, or to skip to the next visit in the queue. If a visit ends early because of a failure to identify the appropriate guide stars, for example, the on-board software starts the next visit in the sequence. The only time the observatory waits is when the earliest start of the next visit is still in the future.

The operational approach is science driven and intended to be simple and understandable to the user. Astronomers will not need to visit the S&OC for their observations. The interaction of astronomers with JWST will be very similar to those of other space-based missions designed for use by large numbers of astronomers.

The S&OC will solicit proposals for JWST observations annually, on behalf of NASA, ESA, and CSA. Following the Spitzer model, it is expected that the first year's observations will be dominated by a small number of large Legacy-style programs and the guaranteed-time observations submitted by the guaranteed-time observers (the instrument teams and other science working group members). In that case, a transition to a larger number of investigations (probably 100–200 programs annually) will occur after the first year. Those proposing observations will use an integrated planning tool, consisting of a graphical user interface and associated widgets, including exposure-time calculators and tools for importing sky maps and accurately positioning the JWST apertures on targets. When writing a proposal, the astronomer will complete those portions required for scientific and technical assessment, including any special requirements for timing and/or fixed orientation of the observatory (Phase I). As with Spitzer, time will effectively be allocated in “wall-clock time,” including estimates for slew and setup times. Once the proposal is selected, the scientist will fill in the remaining details that would be required to execute the approved observations on JWST (Phase II). This two-phase approach was largely developed by STScI for HST to allow astronomers to spend most of their time before selection creating scientific justifications for their ideas. Only those whose proposals are selected must develop the additional details needed to execute the observations on the spacecraft.

Following program selection, the S&OC will construct a long-range plan for observing with JWST. The long-range plan will consist of the approved science programs and the calibration and maintenance activities for the observatory. The same planning tools used by general observers for their science programs will be used for the calibration and maintenance programs. This will reduce cost and complexity within the ground system by limiting the total number of planning systems. It will also provide better service to users since calibration and operations scientists and operators will use the same tools. The long-range plan will specify possible scheduling windows for all the observations and calibration activities for the coming year. The scheduling windows are not simply the times during the year when a target is in the FOR of the observatory, but are limited to assure that there is a high probability that most (90%) of the approved observations will be carried out during the year. This involves assuring that there is a good mix of “easy” and “hard” observations that are available for scheduling when short-term scheduling takes place. Examples of “hard” observations include sets of observations that are extremely long or that have tight timing or orientation constraints. The long-range plan will be updated as observations are completed and modified to reflect new proposals as they are approved and made ready for observation with JWST.

The long-range plan provides a pool of observations that must be turned into a sequence of observations to be executed on the observatory. Development of this sequence, which is known as the observational plan, will normally begin about 1 month before observations are executed on the observatory. Scheduling is carried out as close to the actual observing time as practical so that the latest information about the observatory and the potential observations can be incorporated into the short-term scheduling process. A typical observation plan will likely last about 22 days, the timescale on which orbit maintenance would be carried out. Shorter observation plans will be used during commissioning, for contingency operations, and for some time-critical activities. Automated tools will help planners create the observation plan from the available targets using a priority-based system that ranges from targets that must be scheduled during the 22 days, to those that can be scheduled in a number of scheduling intervals. Imaging observations at $\lambda < 10 \mu\text{m}$ are limited in sensitivity by the zodiacal light, and so sensitivity is maximized when the observations are done at high solar elongation.

Since JWST does not have to contend with Earth eclipses or South Atlantic Anomaly passages as HST does, the optimal schedule is one that maximizes the number of higher priority observations while minimizing slew lengths, and managing the amount of momentum stored in the momentum wheels and the peak data volume on the recorder. JWST is intended to be efficient and science exposures are currently expected to use 70% of wall-clock time, although this depends on the mix of short and long observations. To achieve this high efficiency, JWST has been designed so that certain calibration data can be obtained from one instrument while another instrument, the prime science instrument, is used for science. During

creation of the observation plan, parallel calibration observations will be integrated with the science-observing plan in a way that maximizes their utility, but does not adversely affect the science observation. No “science parallels” are currently planned, although there are no restrictions in the observatory that would prevent them.

The observation plan, along with all of the associated information describing each individual pointing, will be broken into segments and sent to JWST at weekly intervals, during one of the daily contacts with the observatory. If for some reason observations are missed, it will be possible to uplink a new plan whenever JWST is in ground contact. Usually the S&OC will let JWST continue with the same operations plan even if a few observations have been missed. Communications with JWST will be through NASA’s Deep Space Network. Real-time communications and commanding of the spacecraft will be conducted via S band, while high-speed data downlink will use Ka band. The solid-state recorder and downlink will be sized for a typical daily load of 232 Gb (compressed) for science and engineering data. This data volume is required because of the 80 Mpix detector arrays used for science observations with readouts every 20–200 s. These nondestructive rapid readouts are required to compensate for the effects of cosmic rays on the detectors at L2. Typically, daily ground contacts lasting about 3 h in total duration will occur to uplink new commands and downlink up to 232 Gb of data. These contacts will also provide Doppler tracking and ranging data for orbit determination.

On rare occasions, it will be necessary to interrupt the JWST observing plan to make target of opportunity observations of time-critical events, such as SNe or gamma-ray burst sources. By constructing and uploading a new observation plan, JWST will be able to make observations of targets of opportunity on timescales as short as 2 days from the decision to do so.

HST is protected from cosmic rays and high-energy particles from the Sun by the Van Allen belts. At L2, however, the quiescent flux is expected to be 5–10 particles $\text{cm}^{-2} \text{s}^{-1}$. The charge that is deposited as these particles pass through the detectors corrupts about 5–10% of the pixels in 1000 s. To limit the effect of this on sensitivity, images must be collected for transmission to the ground over a shorter (20–200 s) time period, although this uses nondestructive readouts and the detectors are not reset this often. The portions of each image that are affected by cosmic rays can be removed as part of data processing on the ground. An alternative would have been to attempt to remove the bad pixels on the spacecraft; however, this is computationally expensive and risks the possibility that the detectors will not perform exactly as predicted on orbit. Consequently, the raw images will be sent to the ground after lossless compression.

During the commissioning phase of JWST, the telescope’s flight operations system will be staffed 24 h a day, 7 days a week, and real-time contact will be maintained with the observatory for all critical operations. Engineers and scientists from Northrop Grumman, the instrument teams, and NASA, ESA, and CSA will participate fully in a joint mission operations team in the commissioning of the observatory and its instruments. Much of the deployment and initial turn-on

will require real-time commanding, but a gradual transition to semiautonomous operations will commence as soon as feasible.

The ground operations software will be constructed around the largely commercial-off-the-shelf command and telemetry system that will have been used in JWST's integration and testing. Although real-time contact with the spacecraft will typically be 3 h a day in normal operations, much of the data will be cached at ground stations and transmitted within 24 h back to the S&OC. Data received by the flight operations data management system will normally have been processed to create error-corrected, compressed data packets but the S&OC will have the capability to complete the initial level of processing in the event of contingencies. The operations system will automatically monitor the performance and state of health of the observatory, integrate mission scheduling with and oversee execution of the observation plan, and handle the initial receipt of data. It will include capabilities for automatic notification of staff in the case of anomalies, since the experts required to deal with specific problems may not be physically present at the S&OC when an anomaly is discovered. The flight operations team within the S&OC will be responsible for uplink and verification of flight software tables and loads required to update and maintain the observatory. The uplinks will be carried out in real time, but are not expected to interrupt normal operations.

The S&OC is responsible for preserving the image quality of the telescope through a sequence of WFS&C visits. During the commissioning phase, most of the initial telescope adjustments will be carried out via real-time commanding, but in normal operations the wavefront-sensing visits, and other JWST calibration observations, will be executed in the same way as science observations, and carried out under the supervision of the flight software. The visits will usually consist of NIRCам observations of a field with one or more bright stars through a set of special-purpose filters and lenses. The wavefront-sensing visits will occur weekly, or on a time scale that is short compared with the expected changes in the mirror. Data from the wavefront-sensing visits will be retrieved from the observatory and through the Deep Space Network on a priority basis. The S&OC will analyze the images. When the quality of the images degrades to the degree that warrants an update, mirror-actuator corrections will be generated and sent to the spacecraft. At the next WFS&C visit, the corrections will be carried out and data before and after the update recorded.

Maintaining the observatory's L2 orbit is the only regular activity that will require real-time commanding. Because orbital maintenance is so critical to the safety of the mission, it will be carried out during ground contact when operation engineers are able to monitor the performance of the observatory.

An archive of all of the data obtained from the observatory will be maintained in the S&OC. The JWST data management system will use the capabilities of the Multimission Archive at Space Telescope. Originally developed for HST, this archive now houses data from a variety of NASA space astronomy missions. The JWST archive will contain the data stored there in its raw form, the calibration files

necessary to calibrate the data, and databases that describe the data. As with HST, a safe-store archive will be maintained at a separate site. Scientific data received from the spacecraft will be processed within 48 h to assess quality using the calibration files available at the time of the observations and to make an initial version of the science data available to the principal investigator.

When the HST archive was initially built in the early 1990s, users received science data that were calibrated at the time of the observation. A user who retrieved data a year later did not benefit from ongoing improvements in the calibration pipeline or the results of calibrations that took place at the time of the observations, unless he or she recalibrated the data. Since then, however, the price of processing has dropped considerably. Now, HST archive data are normally reprocessed “on-the-fly” each time a request arrives. The same approach will be used for JWST. Users will request data from the archive using Web-based tools similar to those used for HST data. Although a few users may want their data on a physical medium, most are expected to retrieve data directly to their home institution via the Internet.

6.8. MANAGEMENT

JWST is a partnership of NASA, ESA, and CSA. NASA’s Goddard Space Flight Center (GSFC) provides overall project management, systems engineering, and scientific leadership for the project. The prime contractor is Northrop Grumman Space Technologies (NGST), with major subcontracts to Ball Aerospace, ITT, and Alliant Techsystems. The STScI is the science and operations center. The responsibilities of the major JWST partners are given in Table XV (Sabelhaus and Decker, 2004).

7. Summary

The JWST science requirements are divided into four themes. 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. 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. 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. 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. Within these themes and objectives, we have derived representative astronomical observations.

TABLE XV
Responsibilities of the major JWST partners

Organization	Responsibilities
Goddard Space Flight Center	Overall project management Overall systems engineering Scientific leadership Integrated science instrument module (ISIM) ISIM assembly, integration, and test (AI&T) Microshutter assembly and detectors for NIRSpec
Northrop Grumman Space Technology	Observatory Systems engineering and interfaces Observatory AI&T Telescope, spacecraft bus, and sunshield design, manufacturing, AI&T Launch site processing, observatory launch, and commissioning Observatory performance and programmatics Spacecraft, sunshield, and deployables Support ground segment and operations
Ball Aerospace	Telescope optical design and optics Beryllium mirror segment development and cryogenic testing Wavefront sensing & control (WFS&C) design and algorithms
ITT	Telescope and observatory AI&T support Telescope ground AI&T Thermal vacuum test configuration and interfaces
Alliant Techsystems	Telescope backplane and secondary mirror support structure
Space Telescope Science Institute	Ground systems development Flight and science operations Optics and instrument support Science program peer review and selection
European Space Agency/ European Consortium	MIRI optical bench assembly
Arianespace	Ariane 5 launch vehicle
EADS Astrium	NIRSpec instrument
Canadian Space Agency/COM DEV	Fine guidance sensor Tunable filter imager
University of Arizona/Lockheed Martin ATC	NIRCam instrument

(Continued on next page)

TABLE XV
(Continued)

Organization	Responsibilities
Jet Propulsion Laboratory	MIRI management MIRI detectors, cooler, software, and end-to-end verification Wavefront sensing and control technology
Rockwell Science Center	Near-IR (NIR) detectors
Raytheon Vision Systems	MIR detectors
Marshall Space Flight Center	Primary mirror technology development and testing Environmental analysis
Ames Research Center	Detector technology development
Johnson Space Center	Thermal vacuum test facility

To enable these observations, JWST consists of a telescope, an instrument package, a spacecraft, and a sunshield. The telescope consists of 18 beryllium segments, some of which are deployed after launch. The segments will be brought into optical alignment on-orbit through a process of periodic WFS&C. The instrument package contains the four science instruments and an FGS. The spacecraft provides pointing, orbit maintenance, and communications. The sunshield provides passive thermal control. The JWST operations plan is based on that used for previous space observatories, and the majority of JWST observing time will be allocated to the international astronomical community through annual peer-reviewed proposal opportunities.

In this paper we have described the astronomical observations that JWST is designed to make and the implementation that will enable those observations. We have provided scientific justifications for these observations within four science themes. The themes are chosen to span the range of science that we expect JWST will do, and the observations are chosen to be representative of that science. We wish to emphasize, however, that these observations will not necessarily be done. The majority of observing time on JWST will be allocated to the worldwide astronomical community through competitive selection of peer-reviewed proposals. Opportunities for proposals will begin 1 or 2 yr before launch, and will continue annually for the lifetime of the mission. JWST will be a highly capable general-purpose observatory able to address a very wide range of scientific investigations. It represents a major contribution to scientific progress by the governments of the United States of America, of the European nations, and of Canada. Regular, competitive peer-reviewed proposal selection will ensure that this international resource will address the most relevant and strongly justified scientific questions, and will leave a legacy of knowledge and discovery for future generations.

References

- Abel, T., Anninos, P., Norman, M. L., and Zhang, Y.: 1998, *ApJ* **508**, 518.
- Abel, T., Bryan, G. L., and Norman, M. L.: 2000, *ApJ* **540**, 39.
- Abraham, R. G.: 1997, Waller, W. H., Fanelli, M. N., Hollis, J. E., and Danks, A. C. (eds.), *AIP Conference Proceedings 408, The Ultraviolet Universe at Low and High Redshifts: Probing the Progress of Galaxy Evolution*, AIP, Woodbury, NY, p. 195.
- Abraham, R. G., Tanvir, N. R., Santiago, B. X., Ellis, R. S., Glazebrook, K., and van den Bergh, S.: 1996, *MNRAS* **279**, L47.
- Abraham, R. G., Ellis, R. S., Fabian, A. C., Tanvir, N. R., and Glazebrook, K.: 1999, *MNRAS* **303**, 641.
- Abraham, R. G., *et al.*: 2004, *AJ* **127**, 2455.
- Acton, D. S., Atcheson, P. D., Cermak, M., Kingsbury, L. K., Shi, F., and Redding, D. C.: 2004, in J. C. Mather (ed.), *Optical, Infrared, and Millimeter Space Telescopes, Proceedings of SPIE*, No. 5487, SPIE, Bellingham, WA, p. 887.
- Adams, F. C., and Fatuzzo, M.: 1996, *ApJ* **464**, 256.
- Alcock, C., *et al.*: 1998, *ApJ* **499**, L9.
- Alonso, R., *et al.*: 2004, *ApJ* **613**, L153.
- Alves, J., Lada, C. J., Lada, E. A., Kenyon, S. J., and Phelps, R.: 1998, *ApJ* **506**, 292.
- Alves, J., Lada, C. J., and Lada, E. A.: 2001, *Nature* **409**, 159.
- André, P., Ward-Thompson, D., and Barsony, M.: 1993, *ApJ* **406**, 122.
- Armus, L., *et al.*: 2004, *ApJS* **154**, 178.
- Bacmann, A., André, P., Puget, J.-L., Abergel, A., Bontemps, S., and Ward-Thompson, D.: 2000, *A&A* **361**, 555.
- Bally, J., and Zinnecker, H.: 2005, *AJ*, **129**, 2281.
- Bally, J., Devine, D., and Sutherland, R.: 1995, *Rev. Mex. A. & A.* **1**, 19.
- Bally, J., O'Dell, C. R., and McCaughrean, M. J.: 2000, *AJ* **119**, 2919.
- Baraffe, I., Chabrier, G., Allard, F., and Hauschildt, P.: 1998, *ApJ* **337**, 403.
- Baraffe, I., Chabrier, G., Barman, T. S., Allard, F., and Hauschildt, P. H.: 2003, *A&A* **402**, 701.
- Barger, A. J., *et al.*: 1998, *Nature* **394**, 248.
- Barkana, R., and Loeb, A.: 2000, *ApJ* **539**, 20.
- Barkana, R., and Loeb, A.: 2001, *Phys. Rep.* **349**, 125.
- Barnby, *et al.*: 2004, *ApJS* **154**, 97.
- Bate, M. R., Bonnell, I. A., and Bromm, V.: 2002, *MNRAS* **336**, 705.
- Bate, M. R., Bonnell, I. A., and Bromm, V.: 2003, *MNRAS* **339**, 577.
- Becker, R. H., *et al.*: 2001, *AJ* **122**, 2850.
- Beckwith, S. V. W. and Sargent, A. I.: 1993, in Levy, E. H. and Lunine, J. I. (eds.), *Protostars and Planets*, Vol. III, University of Arizona Press, Tucson, p. 521.
- Beckwith, S. V. W., Henning, T., and Nakagawa, Y.: 2000, in Mannings, V., Boss, A. P., and Russell, S. S. (eds.), *Protostars and Planets*, Vol. IV, University of Arizona Press, Tucson, p. 533.
- Beckwith, S. V. W., *et al.*: 2003, *BAAS* **202**, 1705.
- Bély, P. Y.: 2003, *The Design and Construction of Large Optical Telescopes*, Springer, New York.
- Bély, P.-Y., Burrows, C. J., and Illingworth, G. D. (eds.): 1989, *The Next Generation Space Telescope*, STScI, Baltimore.
- Bender, R., Burstein, D., and Faber, S. M.: 1992, *ApJ* **399**, 462.
- Bertoldi, F.: 1989, *ApJ* **346**, 735.
- Boccaletti, A., *et al.*: 2004, *PASP* **116**, 1061.
- Bonnell, I. A., Bate, M. R., and Zinnecker, H.: 1998, *MNRAS* **298**, 93.
- Boss, A. P.: 1988, *ApJ* **331**, 370.
- Boss, A. P.: 2001, *ApJ* **551**, L167.

- Boss, A. P., and Yorke, H. W.: 1995, *ApJL* **439**, 55.
- Botta, O.: 2004, in Ehrenfreund, P., Irvine, W., Owen, T., Becker, L., Blank, J., Brucato, J., Colangeli, L., Derenne, S., Dutry, A., Despois, D., Lazcano, A., and Robert, F. (eds.), *Astrobiology: Future Perspectives*, Kluwer Academic, Dordrecht, The Netherlands, p. 359.
- Bouvier, J., Stauffer, J. R., Martín, E. L., Barradoyy Navascues, D., Wallace, B., and Bejar, V. J. S.: 1998, *A&A* **336**, 490.
- Bouwens, R., Broadhurst, T., and Illingworth, G.: 2003, *ApJ* **593**, 640.
- Bouwens, R., *et al.*: 2004, *ApJ* **616**, L79.
- Bouwens, R. J., Illingworth, G. D., Thompson, R. I., and Franx, M.: 2005a, *ApJ* **624**, L5.
- Bouwens, R. J., Illingworth, G. D., Blakeslee, J. P., and Franx, M.: 2005b, astro-ph/0509641.
- Bouwman, J., Meeus, G., de Koter, A., Hony, S., Dominik, C., and Waters, L. B. F. M.: 2001, *A&A* **375**, 950.
- Brinchmann, J., *et al.*: 1998, *ApJ* **499**, 112.
- Bromm, V., and Larson, R. B.: 2004, *ARA&A* **42**, 79.
- Bromm, V., Coppi, P. S., and Larson, R. B.: 1999, *ApJ* **527**, L5.
- Bromm, V., Kudritzki, R. P., and Loeb, A.: 2001, *ApJ* **552**, 464.
- Bromm, V., Coppi, P. S., and Larson, R. B.: 2002, *ApJ* **564**, 23.
- Burrows, C. J., *et al.*: 1996, *ApJ* **473**, 437.
- Burrows, A., *et al.*: 1997, *ApJ* **591**, 856.
- Burrows, A., Hubbard, W. B., Lunine, J. I., and Leibert, J.: 2001, *Rev. Mod. Phys.* **73**, 719.
- Burrows, A., Sudarsky, D., and Lunine, J. I.: 2003, *ApJ* **596**, 587.
- Cen, R.: 2003a, *ApJ* **591**, 12.
- Cen, R.: 2003b, *ApJ* **591**, L5.
- Cernicharo, J., Noriega-Crespo, A., Cesarsky, D., Lefloch, B., González-Alfonso, E., Najarro, F., Darois, E., and Cabrit, S.: 2000, *Science* **288**, 649.
- Chabrier, G., and Baraffe, I.: 2000, *ARA&A* **38**, 337.
- Chabrier, G., Baraffe, I., Allard, F., and Hauschildt, P.: 2000, *ApJ* **542**, L119.
- Charbonneau, D., Brown, T. M., Latham, D. W., and Mayor, M.: 2000, *ApJ* **529**, L45.
- Charbonneau, D., Brown, T. M., Noyes, R. W., and Gilliland, R. L.: 2002, *ApJ* **568**, 377.
- Charbonneau, D., *et al.*: 2005, astro-ph/0503457.
- Chary, R., and Elbaz, D.: 2001, *ApJ* **556**, 562.
- Chary, R., *et al.*: 2004, *ApJS* **154**, 80.
- Chiu, W. A., and Ostriker, J. P.: 2000, *ApJ* **534**, 507.
- Cole, S., Aragon-Salamanca, A., Frenk, C. S., Navarro, J. F., and Zepf, S. E.: 1994, *MNRAS* **271**, 781.
- Couchman, H. M. P., and Rees, M. J.: 1986, *MNRAS* **221**, 53.
- Cowie, L. L., Hu, E. M., and Songaila, A.: 1995, *Nature* **377**, 603.
- Crovisier, J., *et al.*: 1997, *Science* **275**, 1904.
- Cruikshank, D. P., Roush, T. L., Moore, J. M., Sykes, M. V., Owen, T. C., and Bartholomew, M. J., Brown, R. H., and Tryka, K. A.: 1997, Stern, S. A., and Tholen, D. J. (eds.), *Pluto and Charon*, University of Arizona Press, Tucson, p. 221.
- Crutcher, R. M., Mouschovias, T. C., Troland, T. H., and Ciolek, G. E.: 1994, *ApJ* **427**, 839.
- Davidson, A. F., Kriss, G. A., and Wei, Z.: 1996, *Nature* **380**, 47.
- Davila, P. S., *et al.*: 2004, Mather, J. C. (ed.), *Optical, Infrared, and Millimeter Space Telescopes, Proceedings of SPIE*, No. 5487, SPIE, Bellingham, WA, p. 611.
- Deming, D., Seager, S., Richardson, L. J., and Harrington, J.: 2005, *Nature* **434**, 740.
- Dickinson, M. E.: 2000, *Phil. Trans. R. Soc. Lond. A* **358**, 2001.
- Dickinson, M. E.: 2004, *BAAS* **205**, 163.01.
- Dickinson, M. E., *et al.*: 2004, *ApJ* **600**, L99.
- Dijkstra, M., Haiman, Z., and Loeb, A.: 2004, *ApJ* **613**, 646.

- Doyon, R., Lafrenière, D., Rowlands, N., Evans, C., Murowinski, R., and Hutchings, J., and Alexander, R.: 2004, in Mather, J. C. (ed.), *Optical, Infrared, and Millimeter Space Telescopes, Proceedings of SPIE*, No. 5487, SPIE, Bellingham, WA, p. 746.
- Drake, M., and Righter, K.: 2002, *Nature* **416**, 39.
- Dressler, A.: 1984, *ARA&A* **22**, 185.
- Dressler, A.: 1996, *HST and Beyond: Exploration and the Search for Origins: A Vision for Ultraviolet–Optical–Infrared Space Astronomy*, AURA, Washington, DC.
- Driver, S. P., Windhorst, R. A., Ostander, E. J., Keel, W. C., Griffiths, R. E., and Ratnatunga, K. U.: 1995, *ApJ* **449**, L23.
- Eales, S., *et al.*: 1999, *ApJ* **515**, 518.
- Eales, S., *et al.*: 2000, *AJ* **120**, 2244.
- Egami, E., *et al.*: 2004, *ApJS* **154**, 130.
- Egami, E., *et al.*: 2005, *ApJ* **618**, L5.
- Ehrenfreund, P., and Charnley, S. B.: 2000, *ARA&A* **38**, 427.
- Ehrenfreund, P., Dartois, E., Demyk, K., and D’Hendecourt, L.: 1998, *A&A* **339**, L17.
- Evans, N. J., II, Rawlings, J. M. C., Shirley, Y. L., and Mundy, L. G.: 2001, *ApJ* **557**, 193.
- Eyles, L., Bunker, A., Stanway, E., Lacy, M., Ellis, R., and Doherty, M.: 2005, astro-ph/0502385.
- Faber, S. M., and Jackson, R. E.: 1976, *ApJ* **204**, 668.
- Fan, X., *et al.*: 2001, *AJ* **122**, 2833.
- Fan, X., *et al.*: 2002, *AJ* **123**, 1247.
- Fan, X., *et al.*: 2003, *AJ* **125**, 1649.
- Fan, X., *et al.*: 2004, *AJ* **128**, 515.
- Feinberg, L.: 2004, in Mather, J. C. (ed.), *Optical, Infrared, and Millimeter Space Telescopes, Proceedings of SPIE*, No. 5487, SPIE, Bellingham, WA, p. 814.
- Ferrarese, L., and Merritt, D.: 2000, *ApJ* **539**, L9.
- Fixsen, D. J., Dwek, E., Mather, J. C., Bennett, C. L., and Shafer, R. A.: 1998, *ApJ* **508**, 106.
- Franx, M., *et al.*: 2003, *ApJ* **587**, L79.
- Frayser, D. T., *et al.*: 2004, *ApJS* **154**, 137.
- Freudling, W., Corbin, M. R., and Korista, K. T.: 2003, *ApJ* **587**, L67.
- Gaudi, B.: 2005, *ApJ* **628**, L73.
- Gebhardt, K., *et al.*: 2000, *ApJ* **539**, L13.
- Gerakines, P. A., *et al.*: 1999, *ApJ* **522**, 357.
- Giavalisco, M., Steidel, C. C., and Macchetto, F. D.: 1996, *ApJ* **470**, 189.
- Gibb, E. L., *et al.*: 2000, *ApJ* **536**, 347.
- Gillon, M., Courbin, F., Magain, P., and Borguet, B.: 2005, astro-ph/0508109.
- Gnedin, N. Y.: 2004, *ApJ* **610**, 9.
- Gnedin, N. Y., and Ostriker, J. P.: 1997, *ApJ* **486**, 581.
- Goody, R. M., and Yung, Y. L.: 1989, *Atmospheric Radiation. Vol. I. Theoretical Basis*, Oxford University Press, Oxford, p. 205.
- Gratadour, D., Rouan, D., Boccaletti, A., Riaud, P., and Clénet, Y.: 2005, *A&A* **429**, 433.
- Green, J. J., Dean, B. H., Ohara, C. M., Redding, D. C., and Zhang, Y.: 2004, in Mather, J. C. (ed.), *Optical, Infrared, and Millimeter Space Telescopes, Proceedings of SPIE*, No. 5487, SPIE, Bellingham, WA, p. 944.
- Greenhouse, M. A., Sullivan, P. C., Boyce, L. A., Glazer, S. D., Johnson, E. L., McCloskey, J. C., and Voyton, M.F.: 2004, in Mather, J. C. (ed.), *Optical, Infrared, and Millimeter Space Telescopes, Proceedings of SPIE*, No. 5487, SPIE, Bellingham, WA, p. 754.
- Gunn, J. E., and Peterson, B. A.: 1965, *ApJ* **142**, 1633.
- Haiman, Z., and Loeb, A.: 1997, *ApJ* **483**, 21.
- Haiman, Z., and Loeb, A.: 1999, *ApJ* **521**, L9.
- Haiman, Z., and Holder, G. P.: 2003, *ApJ* **595**, 1.

- Haiman, Z., Thoul, A. A., and Loeb, A.: 1996, *ApJ* **464**, 523.
- Hamann, F., and Ferland, G.: 1999, *ARA&A* **37**, 487.
- Heap, S. R., *et al.*: 2000, *ApJ* **534**, 69.
- Heger, A., and Woosley, S. E.: 2002, *ApJ* **567**, 532.
- Henry, G. W., Marcy, G. W., Butler, R. P., and Vogt, S. S.: 2000, *ApJ* **529**, L41.
- Hester, J. J., *et al.*: 1996, *AJ* **111**, 2349.
- Hillenbrand, L. A., and Carpenter, J. M.: 2000, *ApJ* **540**, 236.
- Hillenbrand, L. A., Massey, P., Strom, S. E., and Merrill, K. M.: 1993, *AJ* **106**, 1906.
- Hogg, D. W., *et al.*: 1998, *AJ* **115**, 1418.
- Holder, G. P., Haiman, Z., Kaplinghat, M., and Knox, L.: 2003 *ApJ* **595**, 13.
- Horner, S. D., and Rieke, M. J.: 2004, in Mather, J. C. (ed.), *Optical, Infrared, and Millimeter Space Telescopes, Proceedings of SPIE*, No. 5487, SPIE, Bellingham, WA, p. 628.
- Hoyle, F.: 1953, *ApJ* **118**, 513.
- Hu, E. M., *et al.*: 2002, *ApJ* **568**, L75.
- Hu, E. M., Cowie, L. L., Capak, P., McMahon, R. G., Hayashino, T., and Komiyama, Y.: 2004, *AJ* **127**, 563.
- Hughes, D. H., *et al.*: 1998, *Nature* **394**, 241.
- Irvine, W. M., Schloerb, F. P., Crovisier, J., Fegley, B., and Mumma, M. J.: 2000, Mannings, V., Boss, A. P., and Russell, S. S. (eds.), *Protostars and Planets*, Vol. IV, University of Arizona Press, Tucson, p. 1159.
- Iverson, R. J., *et al.*: 2004, *ApJS* **154**, 124.
- Jakobsen, P., *et al.*: 1994, *Nature* **370**, 35.
- Jayawardhana, R., Fisher, S., Hartmann, L., Telesco, C., Pina, R., and Fazio, G.: 1998, *ApJL* **503**, L79.
- Johnstone, D., Hollenbach, D., and Bally, J.: 1998, *ApJ* **499**, 758.
- Jones, B. F., and Walker, M. F.: 1988, *AJ* **95**, 1755.
- Kennicutt, R. C.: 1999, *ApJ* **525**, 1165.
- Kenyon, S. J., and Hartmann, L.: 1995, *ApJS* **101**, 117.
- Koerner, D. W., Ressler, M. E., Werner, M. W., and Backman, D. E.: 1998, *ApJL* **503**, L83.
- Kogut, A., *et al.*: 2003, *ApJS* **148**, 161.
- Konacki, M., Torres, G., Sasselov, D. D., and Jha, S.: 2005, *ApJ* **597**, 1076.
- Koresko, C. D.: 1998, *ApJL* **507**, L145.
- Kormendy, J.: 1977, *ApJ* **218**, 333.
- Kramer, C., Alves, J., Lada, C. J., Lada, E. A., Sievers, A., Ungerechts, H., and Walmsley, C.M.: 1999, *A&A* **342**, 257.
- Krist, J. E., *et al.*: 1999, *ApJ* **515**, 35.
- Kroupa, P.: 1998, *MNRAS* **298**, 231.
- Kroupa, P., Petr, M. G., and McCaughrean, M. J.: 1999, *New Astron.* **4**, 495.
- Labbé, I., *et al.*: 2005, *ApJ* **624**, L81.
- Lada, C. J., Lada, E. A., Clemens, D. P., and Bally, J.: 1994, *ApJ* **429**, 694.
- Lada, C. J., Alves, J., and Lada, E. A.: 1999, *ApJ* **512**, 250.
- Lagache, G., Dole, H., and Puget, J.-L.: 2003, *MNRAS* **338**, 555.
- Lambright, W. H.: 1995, *Powering Apollo: James E. Webb of NASA*, Johns Hopkins University Press, Baltimore.
- Larosa, T. N.: 1983, *ApJ* **274**, 815.
- Lefloch, B., and Lazareff, B.: 1994, *A&A* **289**, 559.
- Li, A., and Lunine, J. I.: 2003, *ApJ* **590**, 368.
- Lilly, S. J., Tresse, L., Hammer, F., Crampton, D., and Le Fevre, O.: 1995, *ApJ* **455**, 108.
- Lilly, S. J., Carollo, C. M., and Stockton, A. N.: 2003, *ApJ* **597**, 730.
- Liu, M. C., Najita, J., and Tokunaga, A. T.: 2003, *ApJ* **585**, 372.

- Loeb, A., and Barkana, R.: 2001, *ARAA* **39**, 19.
- Low, C., and Lynden-Bell, D.: 1976, *MNRAS* **176**, 367.
- Lucas, P. W., and Roche, P. F.: 2000, *MNRAS* **314**, 858.
- Luhman, K., Rieke, G. H., Young, E. T., Cotera, A. S., Chen, H., Rieke, M. J., Schneider, G., and Thompson, R. I.: 2000, *ApJ* **540**, 1016.
- Lunine, J. I.: 2005, *Astrobiology: A Multidisciplinary Approach*, Addison Wesley Longman, San Francisco.
- Lunine, J. I., Chambers, J., Morbidelli, A., and Leshin, L. A.: 2003, *Icarus* **165**, 1.
- Lunine, J. I., Coradini, A., Gautier, D., Owen, T.C., and Wuchterl, G.: 2004, in Bagenal, F., Dowling, T., and McKinnon, W. (eds.), *Jupiter*, Cambridge University Press, Cambridge, UK, p. 19.
- MacLow, M.-M., and Klessen, R. S.: 2004, *Rev. Mod. Phys.* **76**, 125.
- Madau, P., and Shull, J. M.: 1996, *ApJ* **457**, 551.
- Madau, P., and Rees, M. J.: 2001, *ApJ* **551**, L27.
- Madau, P., Ferrara, A., and Rees, M. J.: 2001, *ApJ* **555**, 92.
- Magorrian, J., *et al.*: 1998, *AJ* **115**, 2285.
- Malfait, K., Waelkens, C., Waters, L. B. F. M., Vandenbussche, B., Huygen, E., and de Graauw, M. S.: 1998, *A&A* **332**, L25.
- Malhotra, S., and Rhoads, J. E.: 2004, *ApJ* **617**, L5.
- Marcy, G. W., Butler, R. P., Vogt, S. S., Fischer, D. A., Henry, G. W., Laughlin, G., Wright, J. T., and Johnson, J.A.: 2005, *ApJ* **619**, 570.
- Mardling, R. A., and Lin, D. N. C.: 2004, *ApJ* **614**, 955.
- Marley, M. S., *et al.*: 2002, *ApJ* **568**, 335.
- McCarthy, P. J., *et al.*: 2004, *ApJ* **614**, L9.
- McCaughrean, M. J., and O'Dell, C. R.: 1996, *AJ* **111**, 1977.
- McCaughrean, M. J., and Andersen, M.: 2002, *A&A* **389**, 513.
- McCaughrean, M. J., Rayner, J. T., Zinnecker, H., and Stauffer, J. R.: 1996, in Beckwith, S., *et al.* (eds.), *Disks and Outflows Around Young Stars*, Springer, Heidelberg, Germany, p. 33.
- McCaughrean, M. J., Chen, H., Bally, J., Erickson, E., Thompson, R. I., Rieke, M. J., Schneider, G., Stolovy, S., and Young, E.: 1998, *ApJ* **492**, L157.
- McCaughrean, M. J., Stapelfeldt, K. R., and Close, L. M.: 2000, Mannings, V., Boss, A. P., and Russell, S. S. (eds.), *Protostars and Planets*, Vol. IV, University of Arizona Press, Tucson, p. 485.
- McCaughrean, M. J., Zinnecker, H., Andersen, M., Meeus, G., and Lodieu, N.: 2002, *ESO Messenger* **109**, 28.
- McCaughrean, M. J., Close, L. M., Scholz, R.-D., Lenzen, R., Biller, B., Brandner, W., Hartung, M., and Lodieu, N.: 2004, *A&A* **413**, 1029.
- McKay, T. A., *et al.*: 2002, *ApJ* **571**, L85.
- McKee, C. F., and Taylor, J. H. Jr.: 2001, *Astronomy and Astrophysics in the New Millenium*, National Academy Press, Washington, DC.
- McKee, C., and Tan, J.: 2002, *Nature* **416**, 59.
- McKinnon, W. B., Lunine, J. I., and Bandfield, D.: 1995, in D. P. Cruikshank (ed.), *Neptune and Triton*, University of Arizona Press, Tucson, p. 807.
- Meeus, G., Waters, L. B. F. M., Bouwman, J., van den Ancker, M. E., Waelkens, C., and Malfait, K.: 2001, *A&A* **365**, 476.
- Meier, R., and Owen, T. C.: 1999, *Space Sci. Rev.* **90**, 33.
- Meier, R., Owen, T. C., Matthews, H. E., Jewitt, D.C., Bockelée-Morvan, D., Biver, N., Crovisier, J., and Gautier, D.: 1998, *Science* **279**, 842.
- Meyer, M. R., *et al.*: 2004, *ApJS* **154**, 422.
- Miller, G. E., and Scalo, J. M.: 1979, *ApJS* **41**, 513.
- Miralda-Escude, J., and Rees, M. J.: 1994, *MNRAS* **266**, 343.
- Miralda-Escude, J., and Rees, M. J.: 1997, *ApJ* **478**, L57.

- Miralda-Escude, J., and Rees, M. J.: 1998, *ApJ* **497**, 21.
- Mobasher, B., *et al.*: 2005, astro-ph/0509768.
- Møller, P., and Kjaergaard, P.: 1992, *A&A* **258**, 234.
- Moorwood, A. F. M., *et al.*: 1996, *A&A* **315**, L109.
- Morbidelli, A., Chambers, J., Lunine, J. I., Petit, J. M., Robert, F., Valsecchi, G. B., and Cyr, K. E.: 2000, *Meteorit. Planet. Sci.* **35**, 1309.
- Mori, M., Ferrara, A., and Madau P.: 2002, *ApJ* **571**, 40.
- Mouschovias, T. C., and Paleologou, E. V.: 1981, *ApJ* **246**, 48.
- Muench, A. A., Alves, J., Lada, C. J., and Lada, E. A.: 2001, *ApJ* **558**, L51.
- Muench, A. A., Lada, C. J., Lada, E. A., and Alves, J.: 2002, *ApJ* **573**, 366.
- Mumma, M. J., Weaver, H. A., and Larson, H. P.: 1987, *A&A* **187**, 419.
- Mumma, M. J., Weissman, P. R., and Stern, S. A.: 1993, in Levy, E., and Lunine, J. I. (eds.), *Protostars and Planets*, Vol. III, University of Arizona Press, Tucson, p. 1177.
- Mumma, M. J., Disanti, M. A., dello Russo, N., Magee-Sauer, K., Gibb, E., and Novak, R.: 2003, *Adv. Space Res.* **31**, 2563.
- Natta, A., and Testi, L.: 2001, *A&A* **376**, L22.
- Nella, J., *et al.*: 2004, in Mather, J. C. (ed.), *Optical, Infrared, and Millimeter Space Telescopes, Proceedings of SPIE*, No. 5487, SPIE, Bellingham, WA, p. 576.
- O'Dell, C. R., and Wen, Z.: 1994, *ApJ* **436**, 194.
- O'Dell, C. R., and Wong, S.-K.: 1996, *AJ* **111**, 846.
- O'Dell, C. R., Wen, Z., and Hu, X.: 1993, *ApJ* **410**, 696.
- Oke, J. B.: 1974, *ApJS* **236**, 27.
- Ostriker, J. P., and Gnedin, N. Y.: 1996, *ApJ* **472**, L63.
- Padgett, D. L., Brandner, W., Stapelfeldt, K. R., Strom, S. E., Terebey, S., and Koerner, D.: 1999, *AJ* **117**, 1490.
- Padoan, P., and Nordlund, Å.: 2002, *ApJ* **576**, 870.
- Page, L., *et al.*: 2006, *ApJ* submitted, astro-ph/0603450.
- Pagel, B. E. J., Edmunds, M. G., Blackwell, D. E., Chun, M. S., and Smith, G.: 1979, *MNRAS* **189**, 95.
- Palla, F., and Stahler, S. W.: 2000, *ApJ* **540**, 255.
- Panagia, N., Stiaivelli, M., Ferguson, H., and Stockman, H. S.: 2003, *Rev. Mex. A. & A.* **17**, 230.
- Percival, W. J.: 2001, *MNRAS* **327**, 1313.
- Pierazzo, E., and Chyba, C. F.: 1999, *Meteorit. Planet. Sci.* **34**, 909.
- Press, W. H., and Schechter, P.: 1974, *ApJ* **187**, 425.
- Puget, J.-L., Abergel, A., Bernard, J.-P., Boulanger, F., Burton, W. B., Desert, F.-Z., and Hartmann, D.: 1996, *A&A* **308**, L5.
- Rees, M. J.: 1976, *MNRAS* **176**, 483.
- Rees, M. J.: 1997, in Tanvir, N. R., Aragon-Salamanca, A., and Wall, J. V. (eds.), *The Hubble Space Telescope and the High Redshift Universe*, World Scientific, Singapore, p. 115.
- Rhoads, J. E., Dey, A., Malhotra, S., Stern, D., Spinrad, H., Jannuzi, B. T., Dawson, S., Brown, M. J. I., and Landes, E.: 2003, *AJ* **125**, 1006.
- Rhodes, J. D.: 2004, *BAAS* **205**, 72.05.
- Rhodes, J., Refregier, A., Collins, N. R., Gardner, J. P., Groth, E. J., and Hill, R. S.: 2004, *ApJ* **605**, 29.
- Riesenberg, R., and Dillner, U.: 1999, Shen, S. S. (ed.), *Imaging Spectrometry, Vol. V, Proceedings of SPIE*, No. 3753, SPIE, Bellingham, WA, p. 203.
- Robert, F.: 2001, *Science* **293**, 1056.
- Rodmann, J.: 2002, Masters Thesis, University of Potsdam, NY.
- Rowlands, N., *et al.*: 2004a, Mather, J. C. (ed.), *Optical, Infrared, and Millimeter Space Telescopes, Proceedings of SPIE*, No. 5487, SPIE, Bellingham, WA, p. 664.

- Rowlands, N., *et al.*: 2004b, Mather, J. C. (ed.), *Optical, Infrared, and Millimeter Space Telescopes, Proceedings of SPIE*, No. 5487, SPIE, Bellingham, WA, p. 676.
- Sabelhaus, P. A., and Decker, J.: 2004, Mather, J. C. (ed.), *Optical, Infrared, and Millimeter Space Telescopes, Proceedings of SPIE*, No. 5487, SPIE, Bellingham, WA, p. 550.
- Salpeter, E. E.: 1955, *ApJ* **121**, 161.
- Sato, B., *et al.*: 2005, astro-ph/0507009.
- Scally, A., Clarke, C. J., and McCaughrean, M. J.: 1999, *MNRAS* **306**, 253.
- Schade, D., Lilly, S. J., Crampton, D., Hammer, F., Le Fevre, O., and Tresse, L.: 1995, *ApJ* **451**, L1.
- Schechter, P.: 1976, *ApJ* **203**, 297.
- Schmitt, B., Quirico, E., Trotta, F., and Grundy, W.: 1998, Schmitt, B., De Bergh, C., and Festou, M. (eds.), *Solar System Ices*, Kluwer Academic, Dordrecht, The Netherlands, p. 199.
- Schneider, G., Smith, B. A., Becklin, E. E., Koerner, D. W., Meier, R., Hines, D. C., Lowrance, P. J., Terrile, R. J., Thompson, R. I., and Rieke, M.: 1999, *ApJ* **513**, L127.
- Schutte, W. A., and Greenberg, J. M.: 1997, *A&A* **317**, L43.
- Schutte, W. A., and Khanna, R. K.: 2003, *A&A* **398**, 1049.
- Sheppard, S. S., Jewitt, D., and Kleyna, J.: 2005, *AJ* **129**, 518.
- Shi, F., Redding, D. C., Green, J. J., and Ohara, C. M.: 2004, Mather, J. C. (ed.), *Optical, Infrared, and Millimeter Space Telescopes, Proceedings of SPIE*, No. 5487, SPIE, Bellingham, WA, p. 897.
- Shields, G. A., *et al.*: 2003, *ApJ* **583**, 124.
- Shu, F. H.: 1977, *ApJ* **214**, 488.
- Shu, F. H., Adams, F. C., and Lizano, S.: 1987, *ARA&A* **25**, 23.
- Shuping, R. Y., Bally, J., Morris, M., and Throop, H.: 2003, *ApJL* **587**, 109.
- Silk, J.: 1977, *ApJ*: **214**, 152.
- Silk, J., and Rees, M. J.: 1998, *A&A* **331**, L1.
- Smith, N., Stassun, K., and Bally, J.: 2005, *AJ* **129**, 888.
- Soifer, B. T., *et al.*: 2004, *ApJS* **154**, 151.
- Spagna, A., *et al.*: 2004, *Mem. S. A. It. Suppl.* **5**, 97.
- Sparks, W. B., and Ford, H. C.: 2002, *ApJ* **578**, 543.
- Spergel, D. N., *et al.*: 2003, *ApJS* **148**, 175.
- Spergel, D. N., *et al.*: 2006, *ApJ* submitted, astro-ph/0603449.
- Spinrad, H.: 1986, *ARA&A* **25**, 231.
- Stahl, H. P., Feinberg, L., and Texter, S.: 2004, Mather, J. C. (ed.), *Optical, Infrared, and Millimeter Space Telescopes, Proceedings of SPIE*, No. 5487, SPIE, Bellingham, WA, p. 818.
- Stahler, S. W., Palla, F., and Ho. P. T. P.: 2000, Mannings, V., Boss, A. P., and Russell, S. S. (eds.), *Protostars and Planets*, Vol. IV, University of Arizona Press, Tucson, p. 327.
- Stapelfeldt, K. R., *et al.*: 2004, *ApJS* **154**, 458.
- Stapelfeldt, K. R., Krist, J. E., Ménard, F., Bouvier, J., Padgett, D. L., and Burrows, C. J.: 1998, *ApJ* **502**, L65.
- Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., and Pettini, M.: 1999, *ApJ* **519**, 1.
- Stiavelli, M.: 1998, *ApJ* **495**, L91.
- Stiavelli, M., Fall, S. M., and Panagia, N.: 2004, *ApJ* **600**, 508.
- Stockman, H. S., *et al.*: 1997, *Next Generation Space Telescope: Visiting a Time When Galaxies Were Young*, AURA, Baltimore.
- Strom, K. M., Strom, S. E., and Merrill, K. M.: 1993, *ApJ* **412**, 233.
- Su, K., *et al.*: 2005, *ApJ* **628**, 487.
- Sudarsky, D., Burrows, A., and Hubeny, I.: 2003, *ApJ* **588**, 1121.
- Sugitani, K., Tamura, M., Nakajima, Y., Nagashima, C., Nagayama, T., Nakaya, H., Pickles, A. J., Nagata, T., Sato, S., Fukuda, N., and Ogura, K.: 2002, *ApJ* **565**, L25.
- Terebey, S., Shu, F. H., and Cassen, P.: 1984, *ApJ* **286**, 529.

- Thi, W. F., Blake, G. A., van Dishoeck, E. F., van Zadelhoff, G. J., Horn, J. M. M., Becklin, E. E., Mannings, V., Sargent, A. I., van den Ancker, M. E., and Natta, A.: 2001, *Nature*, **409**, 60.
- Thompson, R. I., Smith, B. A., and Hester, J. J.: 2002, *ApJ* **570**, 749.
- Throop, H. B., Bally, J., Esposito, L. W., and McCaughrean, M. J.: 2001, *Science* **292**, 1686.
- Tremaine, S., *et al.*: 2002, *ApJ* **574**, 740.
- Trilling, D. E., Lunine, J. I., and Benz, W.: 2002, *A&A* **394**, 241.
- Tully, R. B., and Fisher, J. R.: 1977, *A&A* **54**, 661.
- Tumlinson, J., and Shull, J. M.: 2000, *ApJ* **528**, L65.
- Tumlinson, J., Giroux, M. L., and Shull, J. M.: 2001, *ApJ* **550**, L1.
- van den Ancker, M. E., Bouwman, J., Wesselius, P. R., Waters, L. B. F. M., Dougherty, S. M., and van Dishoeck, E. F.: 2000a, *A&A* **357**, 325.
- van den Ancker, M. E., Tielens, A. G. G. M., and Wesselius, P. R.: 2000b, *A&A* **358**, 1035.
- van Dishoeck, E. F., and Blake, G. A.: 1998, *ARA&A* **36**, 317.
- Vidal-Madjar, A., Lecavelier Des Etangs, A., Désert, J.-M., Ballester, G. E., Ferlet, R., Hébrard, G., and Mayor, M.: 2003, *Nature* **422**, 143.
- Vogt, N. P., *et al.*: 1996, *ApJ* **465**, L15.
- Vogt, N. P., *et al.*: 1997, *ApJ* **479**, L121.
- Walter, F., *et al.*: 2004, *ApJ* **615**, L17.
- Ward-Thompson, D., Scott, P. F., Hills, R. E., and André, P.: 1994, *MNRAS* **268**, 276.
- Ward-Thompson, D., Motte, F., and André, P.: 1999, *MNRAS* **305**, 143.
- Waters, L. B. F. M.: 2000, in Salama, A., *et al.* (eds.), *ISO Beyond the Peaks: The 2nd ISO Workshop on Analytic Spectroscopy*, ESA-SP 456, p. 39.
- Weinmann, S. M., and Lilly, S. J.: 2005, *ApJ* **624**, 526.
- Wells, C., Whitman, T., Hannon, J., and Jensen, A.: 2004, in Mather, J. C. (ed.), *Optical, Infrared, and Millimeter Space Telescopes, Proceedings of SPIE*, No. 5487, SPIE, Bellingham, WA, p. 859.
- Werner, M. W., *et al.*: 2004, *ApJS* **154**, 1.
- White, S. D. M., and Frenk, C. S.: 1991, *ApJ* **379**, 52.
- Wilner, D. J., and Lay, O. P.: 2000, in Mannings, V., Boss, A. P., and Russell, S. S. (eds.), *Protostars and Planets*, Vol. IV, University of Arizona Press, Tucson, p. 509.
- Wilson, G., Kaiser, N., Luppino, G. A., and Cowie, L. L.: 2001, *ApJ* **555**, 572.
- Wise, J. H., and Abel, T.: 2005, *ApJ* **629**, 615.
- Wolfire, M. G., and Cassinelli, J. P.: 1986, *ApJ* **310**, 207.
- Wolfire, M. G., and Cassinelli, J. P.: 1987, *ApJ* **319**, 850.
- Wright, G. S., *et al.*: 2004, in Mather, J. C. (ed.), *Optical, Infrared, and Millimeter Space Telescopes, Proceedings of SPIE*, No. 5487, SPIE, Bellingham, WA, p. 653.
- Yan, H., and Windhorst, R. A.: 2004a, *ApJ* **600**, L1.
- Yan, H., and Windhorst, R. A.: 2004b, *ApJ* **612**, L93.
- Yan, L., *et al.*: 2004, *ApJS* **154**, 75.
- Yan, H., *et al.*: 2005, astro-ph/0507673.
- Yorke, H. W., and Bodenheimer, P.: 1999, *ApJ* **525**, 330.
- Yorke, H. W., and Sonnhalter, C.: 2002, *ApJ* **569**, 846.
- Zamkotsian, F., and Dohlen, K.: 2004, in Mather, J. C. (ed.), *Optical, Infrared, and Millimeter Space Telescopes, Proceedings of SPIE*, No. 5487, SPIE, Bellingham, WA, p. 635.
- Zapatero-Osorio, M. R., Rebolo, R., Martín, E. L., Hodgkin, S. T., Cossburn, M. R., Magazzu, A., Steele, I. A., and Jameson, R. F.: 1999, *A&AS*, **134**, 537.
- Zaritsky, D., and White, S. D. M.: 1994, *ApJ* **435**, 599.
- Zepf, S. E.: 1997, *Nature* **390**, 377.
- Zucconi, A., Walmsley, C. M., and Galli, D.: 2001, *A&A* **376**, 650.