III. SAFIR Science Rationale

In this section of the report we develop selected elements of the SAFIR science case, expanding on many of the key science objectives summarized in Section II. We note that as of the date of completion of this Study Report, the results from the Spitzer Space Telescope are just beginning to be digested. It can be assumed that for at least the short wavelength parts of the SAFIR science case, Spitzer will not only help focus our strategic science goals further, but develop entirely new lines of research that SAFIR can attack. As a result, our team would like to emphasize the importance to NASA of near-term updates to the science case for SAFIR based on discoveries coming out of Spitzer.

Distant Universe

Background

The search for the origins of galaxies is a pursuit that is fundamental to our quest to understand the Universe around us. The search requires the most comprehensive range of observations, at all accessible wavelengths. In particular, it is essential to make observations at wavelengths where only relatively shallow coverage has been obtained at present. One wavelength range that has been lightly covered to date is in the mid- and far-infrared. At these wavelengths galaxies emit powerful thermal continuum radiation from their warm and cool interstellar dust grains, the small solid particles that absorb optical and ultraviolet light from stars and active galactic nuclei. There are also a wide range of transition lines from atomic and molecular gas that also appear at these wavelengths, including emission from large molecules/small dust grains - the polycyclic aromatic hydrocarbons (PAHs). These lines can trace the physical conditions in optically thick regions of galaxies that are impossible to probe in other ways, and which can carry information about the way in which stars form directly in their nurseries. Similarly, continuum and line emission from the central regions of active galactic nuclei (AGNs) appear at mid-IR wavelengths, carrying information about the fueling and growth of their black holes.

Dust influences the appearance of galaxies at optical wavelengths both by preferentially absorbing blue light to redden the galaxies, and by making some of the most active regions difficult to identify at all at these wavelengths. Even where the effects of dust can be recognized in optical and near-IR images, it is almost impossible to quantify the total amount of energy absorbed and re-emitted by the dust based on this short-wavelength data, without making direct observations in the far-IR.

The importance of dust emission as a process for revealing and controlling the evolution of galaxies is clear when viewed galaxy by galaxy, including the star-forming regions within the Milky Way, and from the beautiful images that Spitzer has provided of nearby galaxies. These images reveal regions of the galaxies that are radically different from those highlighted in optical images. The same disparity appears to be reflected in the properties of more distant galaxies with less significant detections and less impressive spatial resolution in the images. Additionally, the intensity of the background radiation at far-infrared wavelengths emitted by galaxies across the whole sky and at all distances was revealed by the COBE satellite (see Figure III-1). This background spectrum, due to the sum of the emission from all galaxies over cosmic history, consists of a broad spectral peak from dust emission that lies at 200 µm and is matched in energy content with the background radiation intensity that has been measured very accurately at optical and near-infrared wavelengths. This reveals directly that dust remains an important energetic factor in the Universe out to the highest redshifts, and has reprocessed about half of all the energy released over the history of the Universe.

In contrast with the shorter-wavelength component of the background radiation intensity, at least half of which has been identified with known faint galaxies detected in the deepest HST images, much less is known about the details of which galaxies generate the background radiation seen at far-IR wavelengths. A large-aperture very-sensitive telescope operating at far-infrared wavelengths is essential to provide deeper and more detailed images of the galaxies that are responsible. Without this capability we will remain largely ignorant of the origin of about half of the energy that has been emitted by galaxies, as they formed their stars, and especially about the growth of black holes deep in their nuclei.

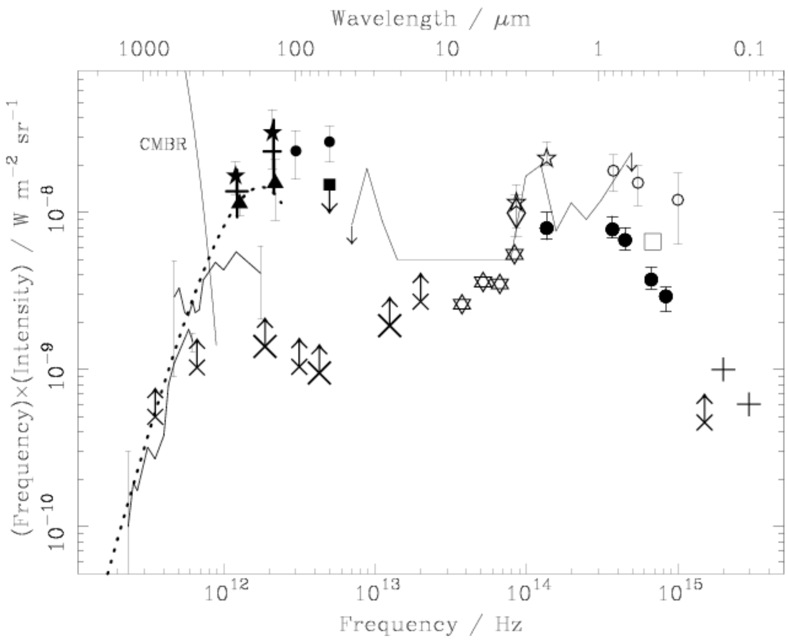


Figure III-1: A composite background radiation spectrum from radio to UV wavelengths. Lower limits are derived from existing far-IR and submillimeter observations from Spitzer and ground-based telescopes (Blain et al. 2002; Physics Reports, 369, 111), updated for recent Spitzer results. The realm of SAFIR is the least-well constrained region of this diagram from 30-500 μm, where the background is strong, but uncertain. The units correspond to equal radiation intensity in equal logarithmic frequency intervals.

SAFIR will provide a whole new class of capabilities to probe the evolution of the Universe at these wavelengths, penetrating deep into the most active regions of the galaxies that are currently known, and identifying galaxies that are far too faint to have been detected by any previous infrared facility.

Space IR telescopes and the Distant Universe

Atmospheric absorption and emission prevent mid- and far-IR observations from the ground. Until the advent of sensitive space-based telescopes there were only hints of the appearance of the dust-enshrouded Universe. The InfraRed Astronomy Satellite (IRAS) in 1984 revealed the first key details of the infrared Universe, providing an all-sky image that remains the baseline for infrared astrophysics. IRAS revealed that the relative energy outputs of galaxies at far-infrared and optical wavelengths were comparable in size, with far-infrared emission being an important contributor in most situations. Moreover, IRAS identified a new class of Ultraluminous Infrared Galaxies (ULIGs) with luminosities that extended up to the most luminous known objects, and which emitted up to 99% of their energy in the far-infrared. Interacting galaxies in the low-redshift Universe were found to often be the hosts of ULIGs, and rare but important examples were found out to redshifts as high as z=3.8. The study of the counterparts of these ULIGs out to high redshifts shows that they represent an important part of the picture of galaxy formation, and are probably associated with the formation of the bulge and elliptical components of the stellar content of galaxies. It was clear from IRAS that the population of far-infrared-selected galaxies undergoes very strong evolution, and that the first half of the Universe's history was much more intense at far-infrared wavelengths than at present.

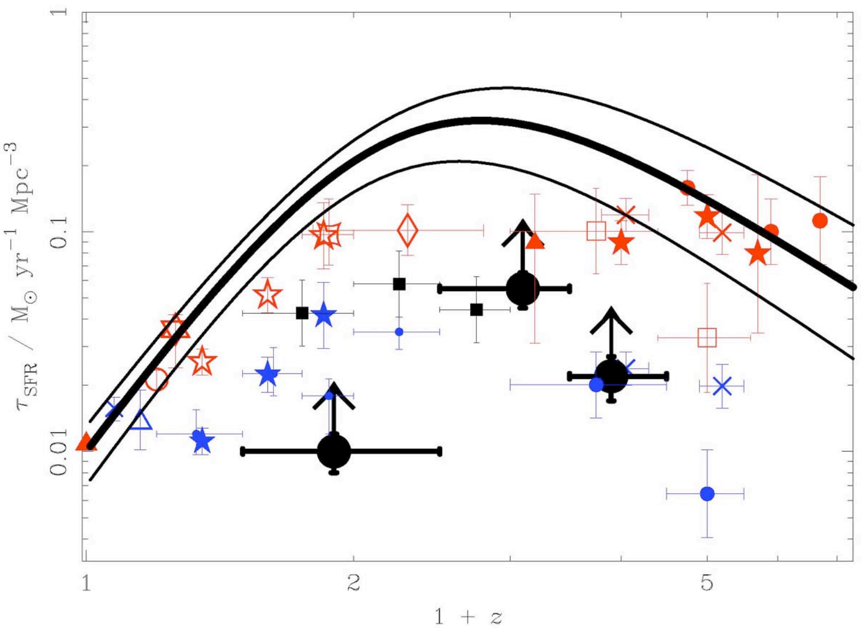


Figure III-2: A compilation of currently known estimates of the rate at which stars formed in the Universe. Small points reflect results from unobscured activity visible at ultraviolet, optical and near-IR wavelengths, large lower limits show the amount of activity so far associated with dust-enshrouded galaxies with redshifts found in submillimeter-wave ground-based surveys with the resolution of SAFIR (Chapman et al. 2005; ApJ, 622, 772). The thick lines represent the inferred best-fitting form of evolution of dust-enshrouded galaxies, incorporating the whole population, including galaxies without known redshifts. SAFIR will image much less luminous galaxies, both to establish the connections between the far-IR and optically visible galaxy populations, including a direct measurement of the reprocessed energy appearing from optical galaxies at far-IR wavelengths, and to identify the sources of a much larger fraction of the inferred total luminosity, providing an excellent census of the energy emitted by galaxies over the Universe's history.

Subsequent space missions have established more details of the strong evolution of galaxies at SAFIR wavelengths, and added the ability to probe details of their spectroscopic properties, providing ways to distinguish between powerful sources of radiation generated by stellar nuclear burning or the gravitational energy released by accreting material onto a black hole in a galaxy's nucleus. The Spitzer Space Telescope, in orbit since 2003 is the current revolutionary facility, providing spectroscopic and imaging capabilities that are far in excess of the earlier missions, including the Infrared Space Observatory (ISO) that enabled the first space-borne mid-IR and far-IR spectroscopy.

Forthcoming surveys at far-IR wavelengths from the Japanese ASTRO-F/IRIS mission, and at mid-IR wavelengths from the NASA MIDEX mission WISE should yield an atlas of galaxies across the whole sky that is complete down to levels between 100 and 1000 times fainter than IRAS reached. Amongst the tens of millions of galaxies that these missions detect, SAFIR will be able to reveal the internal structures, and detailed spectral properties of a sample of these galaxies. This will allow us to understand both their full range of properties and power sources, and provide the means to identify the most extreme and unusual examples that are likely to illuminate the properties of the wider population.

Very Deep Imaging in the Far Infrared and Submillimeter

By making the deepest possible images of the sky at mid-IR wavelengths, subject to its modest aperture, Spitzer is revealing the properties of dust-enshrouded galaxies, and elucidating their relationship to the existing faint samples of optically selected galaxies out to redshifts z~1-2. The ultradeep GOODS project is leading the way. Note however that Spitzer's capabilities to probe the evolution of galaxies is strictly limited at these redshifts by its 0.85-m aperture that imposes a best-possible resolution of several arcseconds. Not only does this prevent Spitzer from revealing the internal structure of distant galaxies that it can readily detected, but moreover, the images of fainter individually undetected galaxies in its field of view merge together to provide the dominant source of noise in the observations. This effect is "confusion noise", which dogged the development of radio and X-ray astronomy, and which prevents the detectors flying on Spitzer from achieving their full capability. Reasonable estimates of the limits to observations using Spitzer are now established based on its in-orbit performance: the maximum depths of its images at observing wavelengths of 24, 70 and 160 μm are about 0.1, 4 and 50 mJy respectively (Dole et al. 2004 ApJ Supp, 154, 93). The luminosities associated with these flux densities mean that Spitzer can detect the less luminous ULIGs only out to z=1, and so a much larger aperture is required to detect typical high-redshift galaxies, and thus to identify the origin of the majority of the diffuse background radiation. Confusion also limits the range of properties of the discrete galaxies that can be studied in detail. In order to identify galaxies in the far-infrared that have already been studied at optical wavelengths, with luminosities of order $10^{11}L_o$ and flux densities of order a few 100 μJy at 200 μm that are expected to lie at redshifts z~2-5, the confusion-beating resolution of a 10 m-class aperture is required at mid- and far-infrared wavelengths.

SAFIR will combine excellent spatial resolution with great sensitivity. At imaging sensitivities of several 10s of μJy, SAFIR should be able to detect a $5\times10^{12}L_o$ ultraluminous infrared galaxy (ULIG) out to redshift z~7, and a $5\times10^{10}L_o$ galaxy similar to the Milky Way out to z~3-4. This will enable the most extreme galaxies to be detected back to the epoch of re-ionization, and probe typical galaxies out beyond the most intense epoch of galaxy formation.

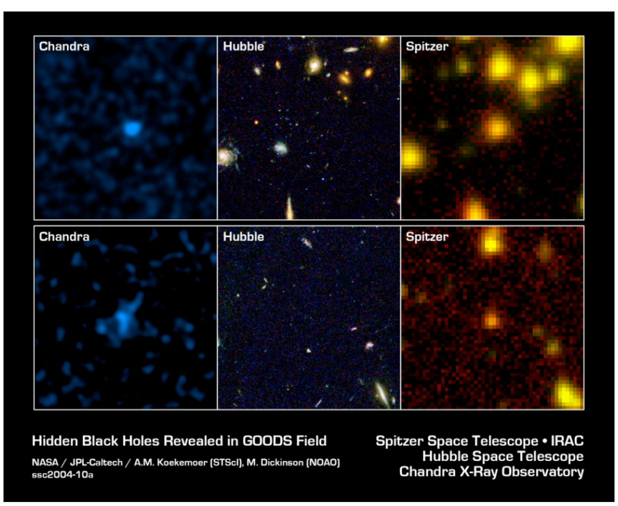


Figure III-3: Some of the deepest multiwavelength observations of the sky ever taken, which combine X-ray (Chandra), optical (Hubble) and infrared (Spitzer) data in 30" wide fields. The X-ray data highlights the presence of two AGN, impossible to detect or at least very difficult to identify in the ultradeep optical image. The infrared image from the Spitzer IRAC camera reveals a readily detectable, red counterpart. Note that the resolution of SAFIR will be comparable to Spitzer, at wavelengths that are several times longer, allowing Spitzer to define the shape of the spectral energy distribution (SED) at the peak.

The resolving power of SAFIR will also make it a very valuable observatory for studying the far-infrared properties of all the galaxies identified using existing methods, either at X-ray, optical, radio or near-infrared wavelengths. This includes reaching the sensitivity levels necessary to detect the typical galaxies that are found at high redshifts in the large samples compiled from optical multicolor surveys, and the reddest and most extreme evolved galaxies and AGN in near-infrared surveys. By measuring colors for the detected galaxies, SAFIR should be able to readily determine the relative amounts of hot and cold dust present, which if hot dust is prevalent, is likely a sign of AGN activity (see Figure III-3). For galaxies out to moderate redshifts (z~0.2-0.5), the resolution of SAFIR can also be used to identify hot emission associated with the nucleus of a galaxy in contrast with cooler emission from the surrounding disk and bulge. This has never been possible for all but the most nearby galaxies at far-infrared wavelengths, and no planned facilities will be able to make these observations.

The detection of hard X-ray sources with Chandra, XMM-Newton and future X-ray facilities can highlight distant AGN, deeply obscured by gas and dust. SAFIR can provide a powerful probe of the astrophysics of these galaxies, in which the supermassive black holes found in galaxies today are growing by both detecting hot dust continuum emission, both by detecting high-excitation spectral lines, and absorption features (see below), and by tracing their SED accurately to assess the total luminosity of the black hole. By comparing spectral features typical of star-formation and AGN activity, the relative contribution of these important processes in distant galaxies can be quantified.

Both to probe inside the sub-arcsec scale structures of high-redshift galaxies, and to ensure that we can continue to probe to ever fainter galaxies in which the bulk of the star-formation and black hole growth in the Universe took place, immune from the effects of confusion noise, a much larger aperture far-IR facility is required than available at present. The specifications of SAFIR are set to achieve these twin goals.

Spectroscopic Probes of the Early Universe
Confusion noise is less of a concern for spectroscopic observations, as the spectral dispersion adds additional effective resolving power. However, in order to probe the distribution of spectral properties from place to place in a distant galaxy, resolution of order 1 arcsec is required. SAFIR can provide this resolution at wavelengths as long as 50 μm.

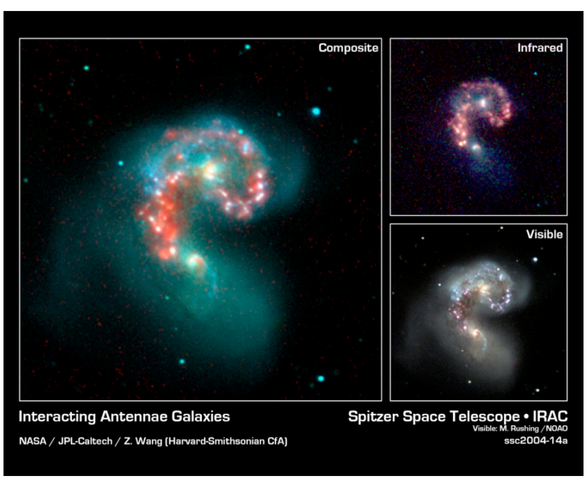


Figure III-4: The power of infrared observations to highlight active regions of galaxies hard to pinpoint in restframe optical images due to dust obscuration. From high redshifts, the 3-8 μm infrared pictures from Spitzer correspond to short wavelengths probed by SAFIR. The infrared image highlights both the emission from the central AGN present, and dust-enshrouded star formation in the Western region of the image.

Less affected by confusion noise, but benefiting from excellent spatial resolution, spectroscopy of PAH emission features, Si-grain absorption features and highly excited Ar and Ne lines in the cores of AGN, cooling lines from rotational-vibrational excitations of molecular hydrogen, fine-structure lines in C, O, N and S from photodissociation regions in molecular clouds are all open to investigation using SAFIR. Spitzer is revealing some of the details of these sources at modest resolutions and shorter wavelengths (see Figure III-5), but for distant objects integrations of many hours are required to detect even the most luminous galaxies. A much larger facility, with a cold aperture, would allow much more rapid progress to be made.

The presence of detectable spectral features in Spitzer observations does confirm that far-IR spectroscopy of galaxies at greater redshifts could provide both redshift determinations, and astrophysical diagnostics. In particular, redshifts could be determined for strongly obscured, very red objects which do not lend themselves to spectroscopy at near-IR or optical wavelengths. For almost all classes of galaxies at all redshifts, the availability of sensitive mid-IR spectra to reveal the conditions in the most-heavily obscured regions, where star-formation or nuclear activity is taking place would provide an unparalleled direct probe of the process of galaxy formation and evolution.

Although the possible power of far-IR spectroscopy is appreciated for studies of galaxy evolution, its execution has not yet become routine. With a warm aperture, the Herschel Space Observatory (HSO) is better suited to high-resolution spectroscopy of Galactic targets, while the proposed Japanese mission SPICA is a useful stepping stone between the performance of Spitzer and the promise of SAFIR. The 10^{-21} Wm^{-2} sensitivity of SAFIR is sufficient to rapidly detect many fine structure lines from a typical galaxies at z=2, and will allow SAFIR to detect [CII] emission from galaxies with luminosities as small as $0.01L_*$. As a result, SAFIR can probe deep into the luminosity function of distant galaxies in the [CII] line, providing an excellent sampling of the physical state of the interstellar gas at specific cosmic times. Probing a series of fine-structure lines in carefully chosen galaxies with similar luminosities but different redshifts could allow changes in the metallicity and state of the interstellar gas to be studied in detail, galaxy by galaxy over the greatest part of the Universe's history from z=1 to beyond 3. The unique access to the complete far-IR wavelength range provided by a space-borne telescope makes these investigations unique for SAFIR.

JWST will have the capability of viewing rest-near-IR features of the galaxies that SAFIR will aim to probe at longer wavelengths. By extending its coverage to longer wavelengths, SAFIR gains access to a new suite of unique and complimentary spectral features, providing great leverage to understand the interplay between gravity and feedback from young stars that regulates the development of stellar populations in galaxies.

SAFIR has uniquely high resolution and sensitivity for making observations of high-redshift spectral lines. Its capabilities provide a natural bridge between those of JWST and ground-based ALMA, spanning the most luminous spectral regions of galaxies.

Transient sources, supernovae and gamma-ray bursts may generate detectable features as they shock and disturb the ISM of their host galaxies. The resulting line emission could be detected by SAFIR, perhaps out to the highest redshifts at which the most massive stars form from almost primordial gas. The absorption spectrum from high-redshift gamma-ray bursts could provide the best way to trace the evolution of the cool interstellar medium in galaxies, highlighted against the bright emission of the burst. With the capability to highlight absorption features in molecular hydrogen gas out to very high redshifts, SAFIR can probe the epoch of the first star formation.

Seeking large-scale structure

In order to probe the way in which galaxies evolved, it is essential to put their growth and development in the context of their environment: proximity to neighbors, and the overall density of nearby galaxies. By enabling surveys to extend in distance well into the high-redshift Universe to

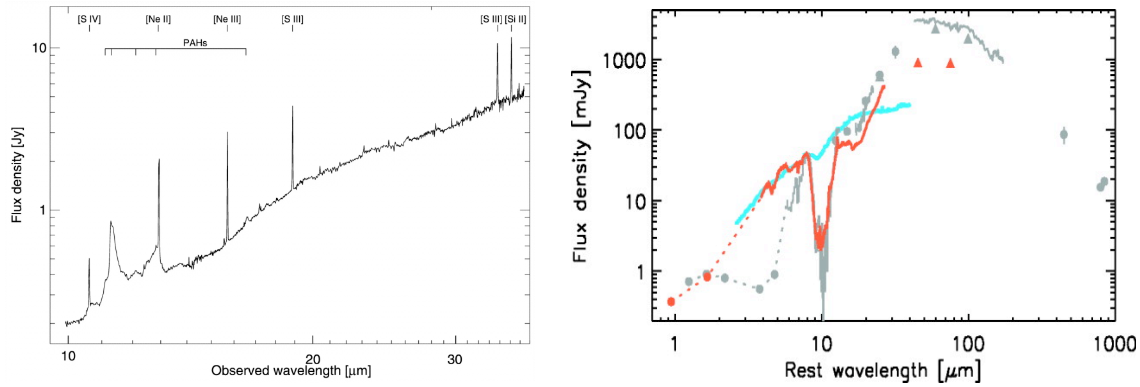


Figure III-5: 2 Spitzer Spectra: The Spitzer IRS spectrograph has started to reveal details of the mid-IR spectra that SAFIR will be able to probe from high-redshift galaxies. This figure shows examples of current spectra, for the typical star-forming galaxy NGC7714 (left: Brandl et al. 2004, ApJS, 154, 188), and for three AGN with different degrees of obscuration (right; Spoon et al. 2004, ApJS, 154, 184): the gray curve shows the deeply enshrouded local AGN NGC4418, the red curve shows the ULIG IRAS F00183-7111 at z=0.327, and the blue curve shows the featureless spectrum of the nucleus of the lightly-obscured source NGC1068. The deep silicate absorption feature that SAFIR can detect out to extremely high redshifts, and the presence of emission lines both provide information about the AGN properties.

reveal populations of galaxies when their typical age was much less than at the present, and by covering wide fields that fully sample the 100 Mpc (comoving) or about 1º scales in the distant Universe on which the largest detected structures have developed, SAFIR can provide a rich sampling of the full range of galaxy properties from the earliest times until they reach maturity.

SAFIR will provide a map of the Universe that traces out to the farthest objects, and which covers a large enough breadth of field to keep track of the full range of cosmic structures that have formed.

The earliest Universe

Observations of the cosmic microwave background radiation imply that the intergalactic medium (IGM), the gas between galaxies was re-ionized by the energy released by stars and galaxies at very early times. However, our view of the possible sources of this energy at optical and near-infrared wavelengths is unavoidably clouded by the small fraction of neutral gas that is know to remain at $z>6.4$ from the spectra of distant QSOs. In the intermediate redshift range the first stars, galaxies and AGN form, and generate the energy required to ionize the IGM, thus ending the “dark ages” of the Universe's history. By probing the first objects to pollute their environment with metals, and perhaps to cool out of molecular hydrogen clouds, SAFIR will directly address the birth of the first generation of cosmic structures. In concert with JWST, but potentially probing even further out into the early Universe, SAFIR will access the starlight now redshifted longwards of 24 μm, and probe the dust-enshrouded regions surrounding the ashes of the first supernovae. SAFIR can search freely within the wide range of possible space to find the first phases of galaxy formation at $6<z<20$.

Active Nuclei in Dusty Primordial Galaxies

By virtue of its large, cooled mirror and high-resolution spectroscopic capabilities, SAFIR promises to deliver huge advances in our understanding of the dynamics and energetics of dusty galaxies both at low and high redshift. These advances will be driven by the great increases in spatial resolution, sensitivity, and long-wavelength coverage afforded by SAFIR over existing (Spitzer) or planned (Herschel) infrared, space-based observatories.

The highest luminosity, dustiest galaxies are often powered by a mix of thermal (star-formation) and non-thermal (AGN) sources. Determining the fractional contribution of each source to the bolometric energy budget in a particular galaxy is extremely difficult, even at low redshift where we have the best S/N and highest possible spatial resolution. It is nearly impossible for sources with z > 2, even with Spitzer. Spatially, the AGN is always unresolved while the starburst is usually extended over many Kpc. Spectrally, the AGN and starburst have very different signatures – strong high-ionization fine structure lines dominate the former, while low-ionization fine structure lines and broad dust emission features (e.g. PAH) dominate the latter (see Figure III-6) Silicate absorption can be present in both. SAFIR, by combining high spatial resolution with greatly increased spectroscopic sensitivity (both being 1-2 orders of magnitude better than Spitzer) will allow us for the first time to separate nuclear from star-forming, disk emission in z < 1 galaxies in the mid and far-infrared, and determine the relative importance of AGN and starbursts for the total energy budgets. Understanding these low-redshift templates is critical for cosmological studies.

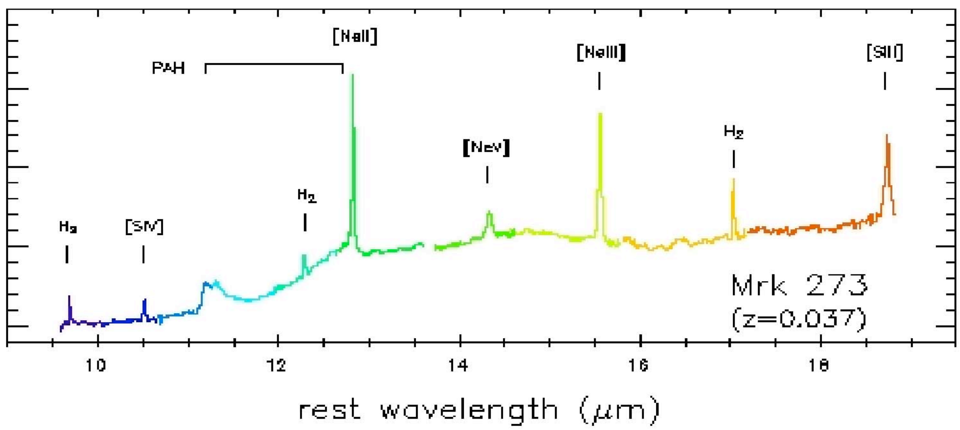


Figure III-6: The Spitzer IRS spectrograph shows the richness of atomic and molecular lines in a typical nearby galactic nucleus that SAFIR will see redshifted in ULIGS.

The sensitivity and long-wavelength coverage of SAFIR will allow us to apply what we learn at low-redshift, to studies of ULIGs at redshifts of z=3 and beyond. SAFIR will be the first observatory capable of obtaining useful infrared spectra of galaxies at these redshifts (well beyond even extremely long integrations with the IRS on Spitzer), separating AGN from starbursts, and tallying up the contributions of these sources to the far-infrared background light. For z > 3, the critical mid-infrared spectral diagnostics ([NeV], [NeVI], [OIV], [SIV]) all move out of range of the IRS on Spitzer (λ > 35 μm) and into the far-infrared. The sesnsitivity of the IRS also limits high-resolution (R=700) studies of these features essentially to z < 2 for all but the brightest sources. Therefore, even after we complete the Spitzer mission, we will have little or no knowledge of the ionization state of the ISM in high-redshift galaxies. Our understanding of the power sources and energetics of high-redshift ULIGs will rest entirely on our ability to fit low-resolution (R=80), low S/N IRS spectra, with local templates. SAFIR would provide a true breakthrough in the study of these distant galaxies, since we would finally be able to measure the state of the gas and the dust and piece-together the evolutionary state of ULIGs at z > 3.

There is one important, yet often overlooked, way in which SAFIR's sensitivity and long-wavelength coverage will provide a breakthrough in our understanding of the true nature of the rapidly evolving population of dusty galaxies at z > 1. The IRAC and MIPS cameras on Spitzer will (and are) generating large samples of luminous galaxies with z > 1. We will only be able to obtain IRS spectra of significant numbers of the brightest of these galaxies, i.e. those with 24 μm flux densities greater than about 0.3 mJy, since each will take many hours of integration with Spitzer. Since the IRS wavelength range is limited at the long end to about 35μm, this means that any flux-limited spectroscopic sample will be biased towards galaxies with strong mid-infrared continua, namely AGN with large quantities of hot dust. Therefore, many of the spectroscopic samples that are built up with the IRS will give us a skewed picture of the relative importance of AGN to the evolution of dusty galaxies at the critical epochs when this evolution is strongest and the contribution to the far-infrared background is largest. We will be able to circumvent some of this built-in bias by carefully examining certain redshift intervals where the PAH features fall, and thus where starburst galaxies shine brightly and dominate the counts. However, with the limited spectral coverage and sensitivity of the IRS, we will find a great deal of AGN-like ULIGs at high redshift. With SAFIR, we will be able to completely remove this bias, and alter our perception of these dusty galaxies. Not only we will measure ULIGs, but we will obtain diagnostic spectra of less luminous, dust-enshrouded galaxies that dominate the number counts, and whose SEDs peak at longer wavelengths because the dust is cooler than around an AGN. Although great strides are being made with Spitzer, we will not be able to put together an accurate picture of the relative numbers of AGN and starburst galaxies at high redshift without SAFIR, and our detailed knowledge of the dusty Universe will be limited.

Complementarity with Other Extragalactic Facilities

From the ground, the long-wavelength tail of thermal emission from dust-enshrouded high-redshift galaxies can be detected, and imaged, while molecular line emission can be probed. Space facilities are however essential in order to trace the properties of these galaxies' emission at the most intense wavelengths, and to survey spectra that cover significant expanses of wavelength, not just restricted to those wavelengths that can penetrate even the upper atmosphere.

A key opportunity from the ground, based on current technologies, is better spatial resolution (see Figure III- 7), owing to much larger apertures than have been flown. Mountain-top sites with 10, 15 and 30 m telescopes have identified a population of very luminous galaxies that are analogous to the ULIGs, and found mostly at z>2. These galaxies can be detected by the 0.85 m aperture Spitzer at 24 μm, but their internal details cannot be resolved, and so it is difficult to trace out the details of their power sources and define their true luminosities well. They are also not detected around the peak of their SEDs at wavelengths ~200 μm, and so the range of their properties are not well sampled. There is sure to be a wide and interesting range of high-redshift galaxies to detect using SAFIR.

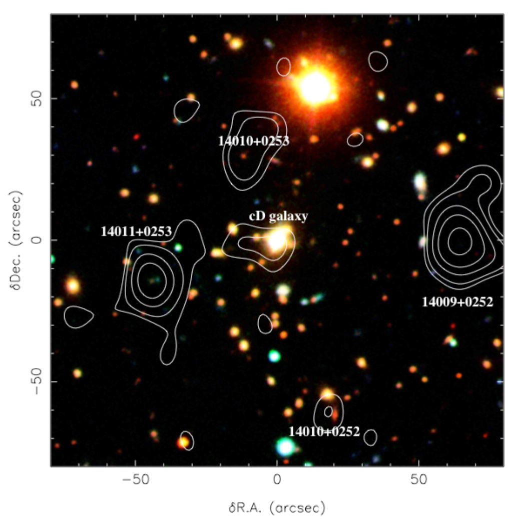


Figure III-7: A multicolor picture combining a Palomar 200-inch image of the core of the cluster A1835, with white contours showing detections of background galaxies using the SCUBA instrument at 850 μm. The brightest background sources are at redshifts z=2.55 and z=2.9. The resolution of the 15 m JCMT is coarser than expected for SAFIR at shorter wavelengths. The lack of overlap between optical and submillimeter sources shows the importance of combining far-IR and optical/near-IR views (Ivison et al. 2000; MNRAS; 315, 209).

Future wide-field mapping instruments imaging the sky at submillimeter wavelengths from ground-based telescopes with arcsec resolution and arcminute fields of view, and the ALMA interferometer with its exquisite 0.01 arcsec resolution will also face these problems. A complementary space-borne capability to probe all available wavelengths, and to reach the peak of the emission spectrum at moderate redshifts is essential for understanding the place of far-infrared emitting galaxies in the full scheme of galaxy evolution. SAFIR will provide several measurements of the SED of the galaxy, allowing an accurate distribution of dust temperatures to be derived, and thus perhaps a reasonably accurate photometric redshift to allow the galaxies to be prioritized for future study.

However, there is excellent synergy with ALMA and future powerful ground-based facilities. Identified as the highest priority for ground-based astronomy during the last decadal review, ALMA will be capable of tracing out the detailed mass distributions within galaxies identified by SAFIR, from molecular line spectroscopy, while SAFIR has unique access to line tracers of AGN activity and metallicity in the galaxy.

A uniquely large, cold aperture allows SAFIR to probe deeper than any other facility. Its wavelength range complements that of JWST, and SAFIR is able to trace the invisible and enshrouded fraction of the energy escaping from distant galaxies that JWST can probe from the emission of starlight, and from the hottest fraction of dust particles. By fully accounting for the fraction of energy absorbed by dust SAFIR will put the cosmological discoveries of HST and JWST in context. In the meantime, Spitzer, Herschel, ASTRO-F and SPICA will add to our knowledge of the dust-enshrouded Universe; however, in order to probe the faintest, most distant, most typical galaxies, the resolving power and imaging speed of SAFIR, equipped with wide-field imaging and spectroscopy detector arrays will be required. The raw power of the 10 m aperture is robust against future discoveries: the most exciting results from Spizter and Herschel demand SAFIR in order to understand the details and consequences of their view.

Activity in Nearby Galaxies

Unification Schemes for Active Galaxies

The vast majority of galaxies in the observable universe are “normal” in that they generate most of their energy by nuclear fusion in the cores of stars. Those galaxies for which a significant fraction of their energy is not generated by stars are called “active galaxies”. These galaxies can have luminosities in excess of $10^{13.5}L_o$, coming from engines of physical size less than 1 pc – the luminosity of a thousand galaxies generated in a space typical of the distance between stars. These active galactic nuclei (AGN) have a large number of subclasses, but there exists a single “Unified Model” for what makes up the broad Seyfert 1 and 2 classes of AGN. Seyfert 1 galaxies have very broad (up to 10,000 km s^{-1}) permitted optical emission lines, and narrow forbidden lines, while Seyfert 2 galaxies have both sets of lines, but they are both relatively narrow (< 1000 km s^{-1}). These two major AGN classes are unified by positing that energy is derived in both cases by accretion onto a supermassive black hole, surrounded by an accretion disk which is fed by a circumnuclear torus, or the tidal disruption of a nuclear star cluster. (Figures III-8 and III-9) The broad lines are thought to arise from modest amounts of ionized gas within 1 pc of the black hole, and heated by the very hot accretion disk. The narrow lines are thought to come from more massive clouds at 10 to 1000 pc from the nucleus. The difference between Seyfert 1 and Seyfert 2 galaxies is thus the result of viewing geometry. Seyfert 1 galaxies are viewed face on so that the broad line region is exposed, while Seyfert 2 galaxies are viewed edge-on so that the broad line region is obscured by the torus.

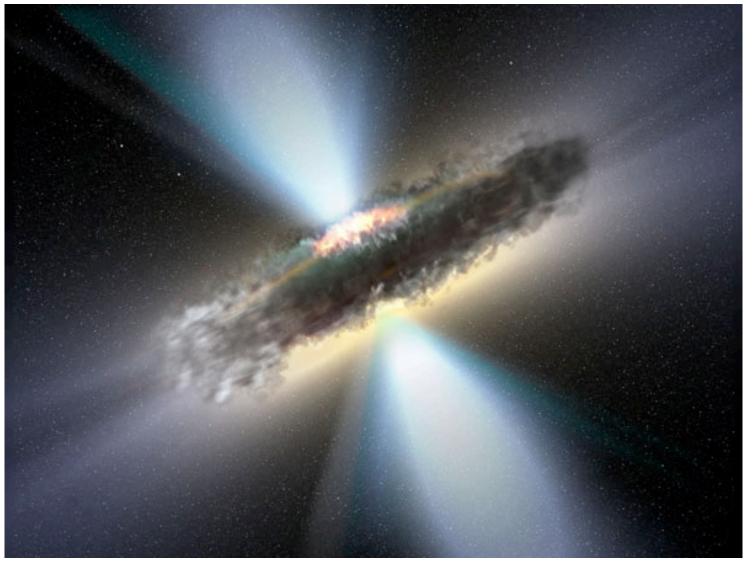


Figure III-: Artist's conception of the doughnut-shaped torus that confines the emission from an active nucleus. (Credit ESA.)

SAFIR provides a unique opportunity to investigate active galaxies and the unification models using the wide variety of far-IR/submillimeter spectral lines as probes. These investigations target specific aspects of the AGN.

The Engine: There is a variety of lines in the far-infrared (e.g. neon sequence, [OIII] and [OIV]) that directly sample the hardness of the radiation fields from the central engine. Particularly useful are lines from different ionization states within an element, e.g. [NeII] (12.8 μm), [NeIII] (15.6 and 36.0 μm), [NeV] (24.3 μm), or [OIII] (51.8 and 88.4 μm) and [OIV] (25.9 μm), as within a species gas phase abundance effects divide out. Intra-element ionization tests are also effect with moderate caveats on assumed elemental abundance ratios.

Unification: Detection of the high ionization lines in the nuclei of Seyfert II galaxies would confirm that the broad line region of Seyfert II's is obscured by foreground material (the confining torus). It is particularly important that the resolving power ($\lambda/\Delta\lambda > 1000$) be sufficient to discern the high velocity wings expected in these lines. Kinematic structure in the centers of galaxies is an important driver for the spectral resolution capabilities of SAFIR instrumentation.

The Torus: The confining torus is also detectable in both the dust continuum and in its submillimeter and far-IR line emission, particularly in the rotational lines of CO, H_2, and H_2O and the fine structure lines of [OI]. Particularly exciting are predictions that very high J CO line emission is detectable from the highly excited molecular torus (Krolik and Lepp 1989 ApJ 347, 179). Such highly excited line emission will be a clear signal for the presence of a confining torus.

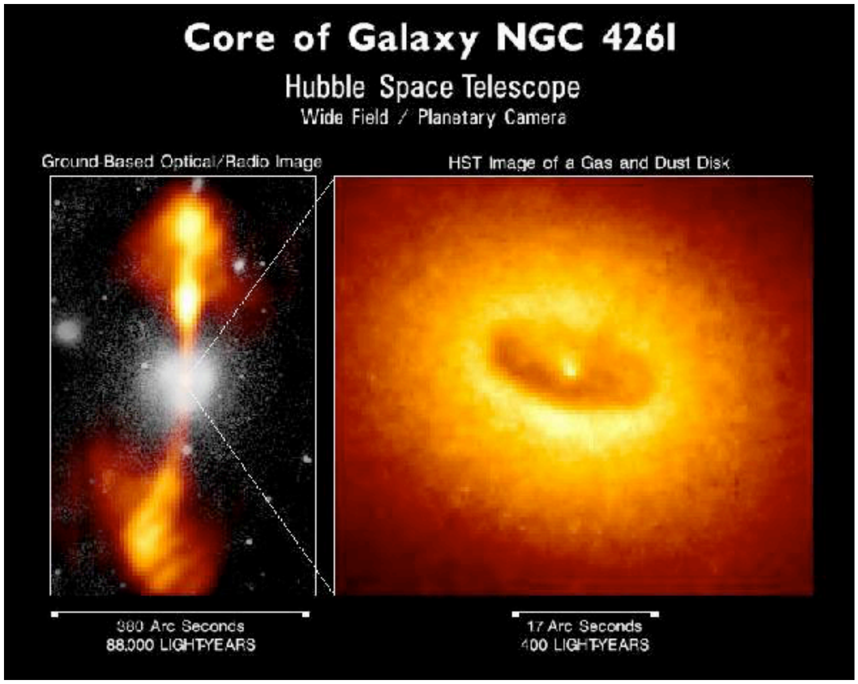


Figure III-9: HST/WFPC2 mage of the active nucleus NGC 4261. NGC 4261 is an elliptical galaxy with radio jets extending at least 25 kpc from the nucleus (left image). These jets are launched from regions near a supermassive black hole (at the bright spot in the center of the right image), which is likely fed by a 50 pc radius ring of gas and dust (the doughnut). SAFIR would be able to image this ring, discerning its physical properties and the ionization parameter of the AGN at 30 pc spatial resolution. Credit: NASA/HST /WFPC2

Detecting CO Cooling Lines from the Torus

Krolic and Lepp have modeled the heating and cooling within a pc scale molecular torus enveloping an AGN. The torus is found to be both very warm (1000 K), and very dense (~ 10^7 cm^{-3}) so that the dominant gas cooling lines are the high – J rotational lines of CO. Remarkably, the emission is predicted to peak in the J ~ 58 → 57 (48 μm) line. Quantitatively, if the ring is exposed to an X-ray source who's ionizing luminosity in units of 10^{44} ergs s^{-1} is L_{X44}, and a fraction, f_{abs} of this flux is absorbed, then the expected line luminosity is: $L_{58\rightarrow57} \sim 7 \times 10^{40}$ f_{abs} L_{X44} ergs s^{-1}, $L_{17\rightarrow16} \sim 2 \times 10^{39}$ f_{abs} $L_{X44,}$ and $L_{7\rightarrow6} \sim 1 \times 10^{38}$ f_{abs} L_{X44} ergs s^{-1}, in the CO J=58→57, 17→16, and 7→6 lines respectively. A standard source at 100 Mpc has line fluxes of 6×10^{-18}, 2×10^{-19}, and 8×10^{-21} W m^{-2} in the three lines respectively assuming f_{abs} ~ 10%. The lines are detectable with SNR ~ 3000, 100, and 4, respectively in 5 minutes of integration time with an R ~ 2000 grating spectrometer on SAFIR.

Comparing active galaxies with with superstarbursts

No high-J CO lines have been observed from external galaxies. However, the mid-J lines have been observed from several starburst nuclei, and there brightness demonstrates that while the molecular gas is very warm and dense, it is far lower excitation than that predicted in the Krolik and Lepp model. Therefore, when comparing emission from a starburst to the torus, the higher the J, the more the torus stands out. As a benchmark, we use the CO rotational model for the nearby starburst nucleus of NGC 253 ($L_{far-IR} \sim 2 \times 10^{10}$ L_o), which is constrained by its observed mid-J (7→6) line emission, and scale this to a 10^{12} L_o starburst (Bradford et al. 2003, ApJ 586, 89 Figure III-10). For this case, $L_{58\rightarrow57,SB} << 10^{32}$, $L_{17\rightarrow16,SB} \sim 1 \times 10^{39}$ ergs s^{-1}, $L_{7\rightarrow6,SB} \sim 1 \times 10^{41}$ ergs s^{-1}. The emission in the mid-J lines is totally dominated by the starburst. Even as high as J = 17→16, the starburst is likely 5 times as bright as the torus emission. At the highest J, however, the starburst is not detectable, and all of the line emission arises from the torus.

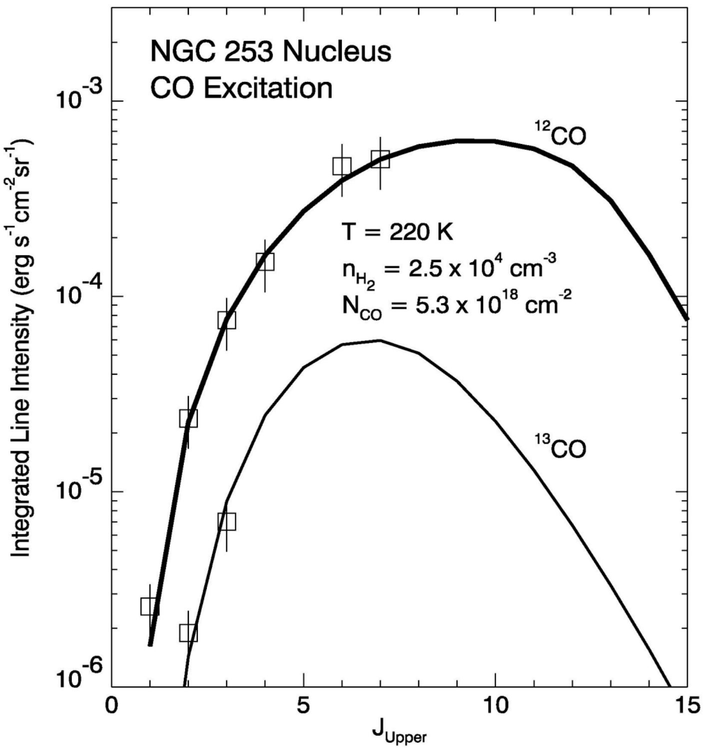


Figure III-10: Best fit LVG model for the CO excitation in the inner 180 pc of NGC 253 plotted on top of the observed ^{12}CO and ^{13}CO lines (Bradford et al. 2003)

Synergy with ALMA

ALMA can just resolve a pc sized torus in the most nearby AGN (10 Mpc), which is quite exciting. However, only SAFIR can detect the higher-J molecular lines that constrain the physical conditions and cooling of the torus. It is the highest J lines that are the "smoking gun" for the confining molecular torus. Furthermore, with its high sensitivity, SAFIR can unambiguously reveal the presence of a warm, dense molecular torus in a statistically significant number of sources through its unsurpassed sensitivity and access to the unambiguous high-J lines. At 100 Mpc, the CO(58→57) and (17→16) lines can be detected with SNR > 100 in just 5 minutes of integration time from a standard source.

SAFIR advantage for unification studies

These ideas are not new, have been pursued with ISO, and are being pursued with Spitzer. However, even for nearby systems, where its intrinsic sensitivity advantage is less important, SAFIR will do better than these facilities. This is because both ISO and SIRTF have relatively large beams, so that much of the host galaxy is included in addition to the AGN. Therefore the AGN lines are swamped by emission from the host galaxy resulting in a low line to continuum ratio, making it difficult to securely identify lines and their intensities.

SAFIR observations are complementary to those planned with the ALMA submillimeter array. While ALMA will offer superb spatial resolution, the mm-wave lines (CO as high as J=7-6) to which

ALMA is sensitive do not strongly distinguish between a warm, dense torus and the more quiescent molecular gas in the host galaxy, so that these lower-J observations do not make a strong test of the Seyfert galaxy hypotheses.

We plan to survey the local Universe with SAFIR for toroids enveloping massive black holes, investigating about 1000 nearby systems. The Milky Way galaxy has a form of torus in the 1.5 pc radius circumnuclear disk circulating about the central black hole. This torus is easily detectable in the 40 μm dust continuum by SAFIR to redshifts z ~ 0.1. The torus in the Seyfert 2 nucleus of NGC 1068 is detectable to high redshifts (z ~ 5) in the continuum. However, at high z the torus would be blended into the emission from other (starforming) regions of the galaxy.

We also plan a survey of the nearest 1000 Seyfert galaxies in the diagnostic lines listed above. With the superb sensitivity of SAFIR, this survey will go quickly. For example, the observed [NeV] 24.31 μm line emission observed from the IR luminous ($L \sim 3 \times 10^{12}$ L_o) Seyfert 1 nucleus of Mrk 1014 (5×10^{-17} W/m^2, Armus et al 2004 ApJ Supp 154,178) at z = 0.1631 would easily be detectable with SAFIR at redshifts in excess of 5, or from systems with $L \sim 3 \times 10^{8}$ L_o in the local Universe.

The Energy Source in ULIG Galaxies

One of the most exciting astronomical discoveries in the last two decades has been the class of Ultraluminous Infrared Galaxies (ULIGs). These are among the most powerful objects in the local Universe with far IR luminosities exceeding 10^{12} L_o. They are often colliding or merging galaxies, and it is likely that the mergers caused the inordinate far-IR luminosities, by either compressing the natal ISM triggering a global starburst, or by triggering accretion onto a central massive black hole forming an AGN. Mid-IR ISO observations suggest that ~ 70% of ULIGs are powered primarily by starbursts (Genzel et al. 1998 ApJ 498, 599), and the remaining ~ 30% are powered by AGN. The first Spitzer data appear to confirm this result (Armus et al. 2004). However, other data appear to contradict these findings. For example, the 158 μm [CII] line luminosity, which is typically quite strong relative to the far-IR luminosity in local starforming galaxies, is substantially weaker in many ULIG galaxies (Luhman et al. 2003 ApJ 594, 758). This suggests that much of the observed far-IR luminosity arises from a buried AGN, or, if the far-IR energy arises from star formation, the starburst is much more intense, and confined than in lower luminosity systems (Stacey et al. 1991 ApJ 373,423), or much of the far-IR continuum arises from dust bounded HII regions (Luhman et al. 2003). For these cases, starbursts in ULIGs are quite different from those in lower luminosity systems.

What powers ULIGs? If they are powered by starbursts, in what way are they different from lower luminosity systems? What fraction of ULIGs are powered by nucleosynthesis (starbursts), and what fraction are powered by gravity (AGN)? These questions are important to understand the local systems, but also vital for the study of star formation in the early Universe. ULIGs may well be the local "slimmed down" version of the recently discovered bright submillimeter galaxies. The far-IR/ submillimeter continuum emission from the submillimeter galaxies comprises a large fraction of the energy budget of the early Universe. Therefore, it is important to fully understand the energy budgets of ULIGs so as to model the relative contribution of gravity and nucleosynthesis to the energy budgets in the early Universe. If powered by star formation, in what way do starbursts differ in ULIGs from lower luminosity systems.

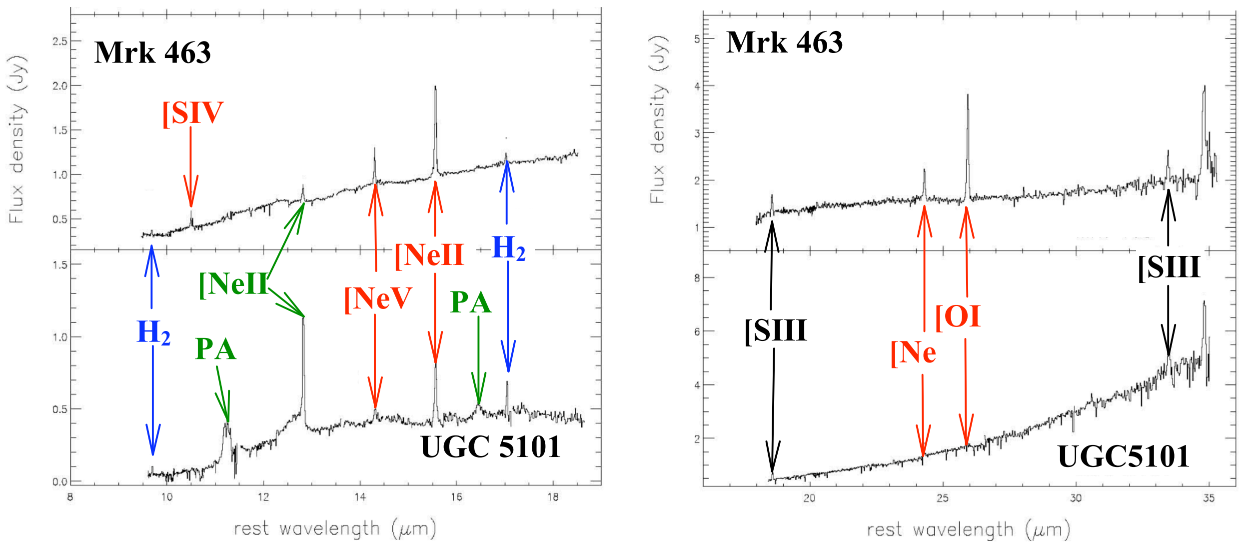


Figure III-11: Spitzer IRS spectra of the AGN Mrk 463, and the starburst galaxy UGC 5101. Prominent diagnostic lines are indicated. Note the dramatic shift from low ionization features (green) to high ionization features (red) between the starburst and AGN spectra (especially the shift from [NeII] to [NeV] – from Armus et al. 2004).

We will use SAFIR to investigate the energy sources in ULIGS using the many atomic and ionic fine structure lines and molecular rotational lines that are available in the far-IR/submillimeter regime. These lines have the distinct advantage over the optical or near-IR lines in that they suffer relatively little extinction, so that they can probe the dust enshrouded emission regions, and that the energy levels are typically low-lying, with modest critical densities, so that they are the dominant coolants and physical probes of much of the interstellar medium in galaxies. SAFIR can:

Determine the energy source in ULIGs: The hardness of the local radiation fields is given by the relative strengths of fine-structure lines from the various ionization states of elements, such as [NeII] (12.8 μm), [NeIII] (15.6 and 36.0 μm), and [NeV] (24.3 μm), [NII] 122 and 205 μm, and [NIII] 57 μm, or [OIII] (51.8 and 88.4 μm) and [OIV] (25.9 μm). Strong [NeV] and/or [OIV] emission is only consistent with the hard radiation fields from AGN (Figure III-11), while strong [NeIII] or [OIII] emission is consistent with a very young starburst with a mass function headed by early type stars.

Determine the Age of the Starburst: The most massive main sequence star determines the hardness of the UV radiation field in galaxies. By observing the relative strengths of the lower ionization neon, oxygen, or sulfur ([SIII] 18.7 and 33 μm) lines, one determines the hardness of the UV radiation field, hence most massive main sequence star, hence age of the starburst. To understand what triggers starbursts in galaxies, it is vitally important to understand the chronology.

Measure the Concentration of Starburst: The strength of the UV radiation field, and the physical parameters of the interstellar medium (temperature, density, clumping, mass) are well characterized by the brightness of the lines arising from photodissociation regions (PDRs), e.g. [OI] (63 and 146 μm), [CII] (158 μm), [CI] (370 and 610 μm), mid-J CO lines (e.g. (7-6) 372 μm) lines, rotational lines of H_2 (e.g. 17 and 28 μm), and the far-IR continuum (cf. Wolfire et al. 1990 ApJ 358, 115), Kaufman et al. 1999 ApJ 527, 795). The beam filling factor for PDRs is derived by comparing the derived strength of the UV radiation field to the observed far-IR continuum flux. If there are two lines from a single ionized gas species (e.g. [OIII]), their ratio yields ionized gas density. One can therefore establish the size of starburst regions without fully spatially resolving them.

SAFIR advantage for understanding energy sources
The spatial resolution of SAFIR is sufficient (~ 1" at 40 μm) to resolve the cores of local ULIGs in the far-IR continuum and lines. This is important, as the lines from a high excitation AGN will be confined to the inner regions of the nucleus – at 50 Mpc, 1" corresponds to about 250 pc. Many systems will be AGN/starburst composites, so that without sufficient spatial resolution, the emission from nearby starburst regions may significantly lower the line to continuum ratio for AGN lines, making their detection problematic. The resolution is also sufficient to enable mapping and statistical studies of merger sequences to investigate the relationship between starburst intensity and age, or the infall process towards an active nucleus.

Armus et al. 2004 present the first Spitzer IRS spectra of the active ULIG galaxies Mrk 1014, Mrk 463, and UGC 5101. SAFIR will be able to detect the observed [NeII], [NeIII], [NeV], [SIII], H_2 (17 μm), and [OIV] line emission from these systems well beyond redshifts of 5 – well beyond the peak of star formation history, and into the regime where galaxies are first assembling. The situation is similar in the longer wavelength lines (scaling from observed [CII] and [OI] flux from ULIGS (Dale et al. 2004 ApJ 604, 555). SAFIR is therefore a platform to probe the evolution of the brightest dust-enshrouded galaxies from their first formation to the present. Equivalently, SAFIR can detect systems with $L \sim 5 \times 10^8 L_o$ in these lines in the local ($z < 0.2$) Universe. We therefore will conduct a survey of the nearest 1000 systems in these lines. This survey will have significant overlap with the Seyfert survey outlined above.

Star Formation in Normal Galaxies

To understand the star formation process in the early Universe, it is important to investigate the star formation process in nearby systems where we can resolve star formation sites and put them into context with the rest of the galaxy. The resolution that SAFIR will achieve is illustrated for nearby galaxies in Figure III-12 below. We will therefore use SAFIR to survey and map about a hundred nearby starforming galaxies. The group will include grand design spirals such as NGC 6946, barred and interacting spirals such as M83 and M51, Milky way analogues such as NGC 7731 and the edge-on system NGC 891, low metallicity dwarf galaxies such as IC 10, and nearby, well resolved interacting/starburst systems such as Arp 244 (the Antennae). Our goal is to fully characterize star formation as a function of galaxy mass, morphology, metallicity, and environment, so that the lessens learned can be applied to the more distant, unresolved sources in our deep surveys. SAFIR will do many things.

Investigation of Star Formation in Low Metallicity Systems
The investigation of the star formation properties of low Metallicity systems is of particular interest, as these systems are the local analogue of the first starforming galaxies in a bottom up scenario. We plan to observe these systems in the variety of far-IR fine structure lines and molecular rotational transitions. We will observe lines that: (1) trace the hardness of the interstellar radiation fields and density of the ionized gas (e.g. [NeII], [NeIII], [OIII], [SIII], [NII], [NIII]); (2) reveal interstellar abundances thereby tracing the degree of processing (age) of the ISM (e.g. [OIII] 52 μm/[NIII] 57 μm); (3) trace the strength of the UV radiation fields, physical conditions of the photodissociated gas, and beam filling factors (e.g. [OI], [CII], mid-J CO, [CI], H_2); and (4) trace the physical conditions of the cool molecular clouds that will form the next generation of stars (e.g. mid, and low-J CO, [CI] lines). We are particularly interested in searching for weak H_2 rotational line emission from the

molecular ISM. There is evidence that much of the molecular ISM is hidden to CO observers (cf. Stacey et al. 1991, Poglitsch et al. 1995 ApJ 454, 293, Madden et al. 1997 ApJ 483, 200).

Search for reservoirs of CO free H_2 in normal spiral galaxies

Investigate the warm molecular ISM in normal spiral galaxies by searching for the rotational line of H_2 at 17 and 28 μm. ISO observations of the edge-on Sb galaxy NGC 891 and the face-on Sc galaxy NGC 6946 revealed surprisingly strong H_2 line emission from these "normal" spirals in regions that are weak in CO line emission. Valentijn & van der Werf (1999 ApJLett 522, L29) suggest that this warm molecular gas may make up a significant amount of the total mass of these galaxies. Early Spitzer results confirm strong H_2 emission from many galaxies (e.g. Smith et al. 2004 ApJ 154, 199).

Investigation of roles of bars, spiral arms, and galactic environment on star formation

SAFIR will investigate the rolls of bars potentials, spiral density waves, and galactic environment (collisions) on star formation in nearby galaxies in the lines that provide the cooling of the natal ISM. As a bar potential or spiral density wave sweeps through quiescent molecular clouds, they are compressed and stimulated to collapse and form stars. The cooling lines for collapse of neutral clouds include the rotational transitions of H_2O, CO, OH, and H_2, as well as the fine structure lines of [CI]. Nearly all of the important cooling lines lie in the SAFIR spectral regime. As stars form, they illuminate their local environments giving rise to photodissociation regions, cooled by the far-IR fine structure lines of [CI], [CII], [OI], and [SiII] and the mid-J rotational lines of CO. Early type stars will form HII regions, cooled by the ionized gas fine-structure lines (e.g. [SIII], [OIII], [NII], [NIII]). The physical conditions (density, temperature, mass, abundances, etc.) of the ISM are traced by the ratios of these lines, and comparison with the far-IR continuum. The line ratios also reveal the hardness of the ambient interstellar radiation field (most massive star on the main sequence), and the strength of the far-UV radiation fields (concentration of the newly formed star clusters). We will investigate the roles of supernovae in energizing the ISM by observing the "galactic chimneys" found in edge-on spiral galaxies such as NGC 891. The spatial resolution available with SAFIR is impressive: at M51, the 1.5" [OI] 63 μm beam corresponds to about 40 pc – or about the scale of large GMCs (cf. Figure 5).

An example of this is mapping M83 with an R ~ 2000 spectrometer. For this example, we assume either an image slicer, or a long slit grating spectrometer with an instantaneous field of view of 256 positions on the sky (as in Figure III-13). Based on ISO results, and scaling to the 1.5" beam of SAFIR, the [OI] line flux in the interarm regions of M83 is ~ 2.3 x 10^{-18} W/m^2 per beam. The [OI] line is detected by SAFIR with SNR > 100 in one second of integration time. The time required to map the entire galaxy in this line (a 6' x 8' region) would take about 300 pointings, or assuming 50% observing efficiency, about 10 minutes of SAFIR time. To map in the 10 other bright far-IR fine-structure lines ([OI] (63 and 146 μm), [OIII] (52 and 88 μm), [SIII] (19 and 33 μm), [NII] (122 and 205 μm), [CII] (158 μm), [NIII] (57 μm), [SiII] (35 μm) requires ~ 1.5 hours of integration time.

With SAFIR's 1.5" beam at 63 μm, we will have imaged nearby starforming galaxies at a linear resolution of ~ 40 pc – the size scale of giant molecular clouds. We therefore investigate the interstellar medium (heating and cooling, density, mass and abundances) and young stars (strength, and hardness of radiation fields) in star formatting galaxies at the size scale relevant for star formation. How do spiral arms, galactic bars, and intergalactic collisions stimulate star formation in galaxies. What is the effect of environment on the initial mass function and size of the starburst?

Figure III-12: Composite visible and IRAC band image of the nearby spiral galaxy, M51 (from the Spitzer public web site). SAFIR will achieve similar spatial resolution in the far-IR spectral lines.

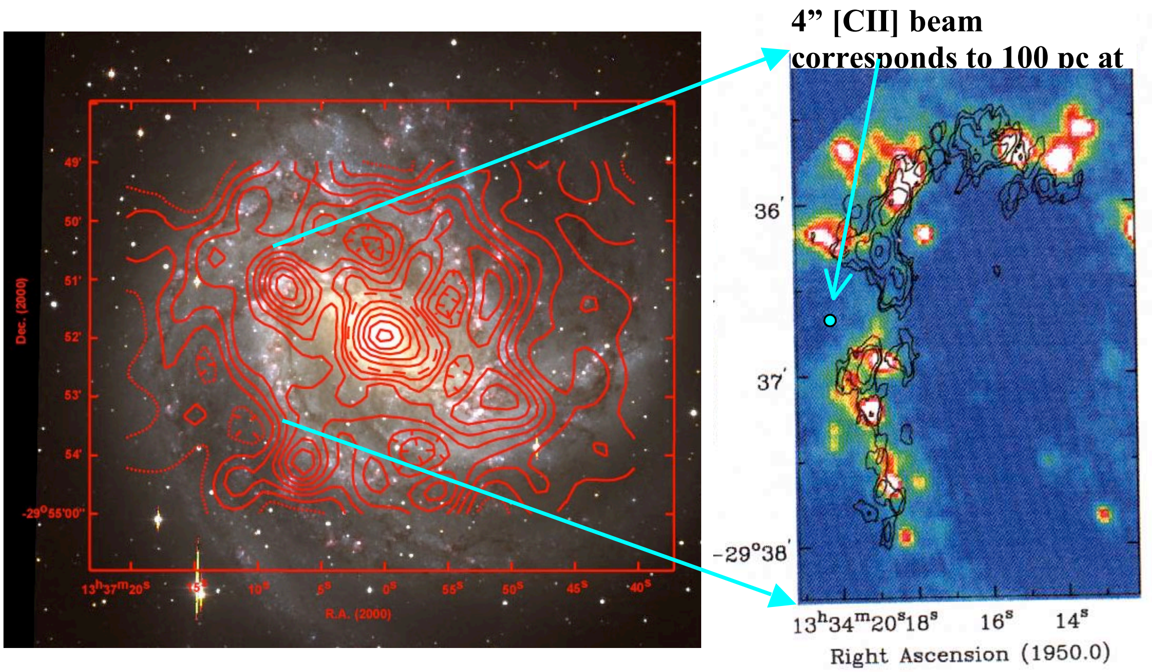


Figure III-13: (left) 158 μm [CII] line map of the barred spiral galaxy M83 obtained with the KAO at 55” resolution (Geis et al. 2005) superposed on an optical image. The [CII] line is a signpost of OB star formation (Stacey et al. 1991) and is strongly enhanced at the bar-spiral arm interfaces to the north-east and south-west. (right) Size of the SAFIR beam at 158 μm superposed on a composite interferometric CO(1→0) line (countours) and optical Hα line (false color) images (Rand, Lord and Higdon, 1999)

We plan to obtain complete maps of about 50 key nearby galaxies, (e.g. IC 10, II Zw40, NGC 891, M51, M83, NGC 6946) in the far-IR spectral lines and continuum, and to extend similar observations to a statistically meaningful (~ 1000) sample of normal galaxies within redshift of 0.5. The sensitivity of SAFIR is such that mapping speed will likely be dominated by pointing overheads, and the overall map size will be determined by the format of the spectrometer involved (spatial and spectral multiplexing). Assuming a spectrometer with a 256 spatial multiplex advantage, each nearby galaxy will take about an hour to map in the brightest 8 lines (50 hours total). The more distant galaxy set will go more quickly since they subtend a much smaller region on the sky (~1'). We expect that mapping 1000 galaxies will require an additional 100 hours of SAFIR time if we integrate 3 seconds per pointing.

SAFIR advantage for studies of star formation in normal galaxies
SAFIR will have unprecedented surface brightness sensitivity, very high spatial resolution (only eclipsed by submillimeter interferometry missions) and excellent mapping capabilities in the most important cooling lines and probes of the interstellar medium. Nearly all of these lines are inaccessible from the ground. The critically important [OI] is blocked by telluric absorption at modest redshifts from aircraft and balloon altitudes, and the very important H_2O cooling lines are totally inaccessible to these platforms. SAFIR observations are complementary to those planned with the ALMA submillimeter array. While ALMA will offer superb spatial resolution, SAFIR accesses unique lines (especially at high frequencies) not available to ALMA observers.

Birth of Stars, Planets and Complex Molecules

Background

Except at the earliest epochs, stars form in the densest parts of dusty molecular clouds in the interstellar medium and the formation of their associated planetesimals, planets, moons, comets, KBOs, etc. is intimately related to the physical conditions in the immediate vicinity of these forming stars. Stars form because the force of gravity ultimately overcomes gas and radiation pressure, magnetic fields, rotational forces, and turbulence. Stars almost universally form in groups, resulting in strong mutual interactions. Solar system objects form through complex processes of dust grain coagulation, solid particle compactization, gravitational attraction, and gas accretion.

Dust plays a prominent role in the formation history of stars and planets. It is a principal player in the thermal evolution and structure of the dense molecular clumps from which they form. During early phases of gravitational collapse, dust obscures the emission at short wavelengths but allows the thermal energy generated during compression to be radiated away, enabling continued contraction. Grain surfaces act as a chemical catalyst for the production of complex molecules, including organics, both prior to and during the star formation process. The result of chemical processing in star forming regions is a richness of molecular species. By studying line shapes, intensities, and excitation, information on the physical conditions can be extracted, because many ions and molecules do not co-exist but rather trace regions with differing conditions.

Because of the high degree of extinction due to dust, the cradles of star and planet formation are completely obscured at wavelengths shorter than about 30 μm. Future missions such as JWST, TPF-C, and TPF-I, which operate at mid-infrared through optical and the ultraviolet spectral regimes will therefore not be able to probe the earliest phases of this formation process. By contrast, the ground-based radio observatory ALMA, the currently operating Spitzer Space Telescope and the future far-infrared/submillimeter Herschel Observatory and SOFIA will provide important observational constraints for our understanding the star and planet formation process. This combination of operating and planned facilities, however, will not offer the wavelength coverage combined with both the sensitivity and spatial/spectral resolution necessary for answering some of the key scientific questions associated with the sequence of events leading to the formation of stars, planets, and life on Earth.

Figure III-14: The massive star-forming region W49 (Alves and Homeier 2003) with 100 O7V "equivalents" to account for the total ionizing flux (blue = H; green = K; red = 3.6 cm). Approximately 40 highly obscured UCHIIs are shown in pure red in this composite.

The unique contributions to be made by SAFIR were summarized in the McKee-Taylor Decadal Survey: "The SAFIR Observatory will …study the relatively unexplored region of the spectrum between 30 and 300 μm. It will enable the study of galaxy formation and the earliest stage of star formation by revealing regions too enshrouded by dust to be studied by [JWST] and too warm to be studied effectively with ALMA." Several of the major star and planet formation science objectives to be addressed by SAFIR are:

- Formation of first stars in the early universe and their evolution
    - How, when, and with which IMF did the first stars (with zero metallicity) form?
    - What is the subsequent star formation history?
    - Is there an unknown population of high-z IR galaxies?

- Formation of stars and the physics of the interstellar medium
    - How do stars form out of the interstellar medium?
    - Circulation / enrichment / chemical processes of the interstellar medium
    - Detailed studies of nearby (resolvable) protostars, star forming regions, "mini-starbursts", starbursts, thus providing templates for studying star formation in distant galaxies

- Formation of new solar systems in protostellar disks
    - How do the disks and their outflows form and evolve?
    - How are planetesimals and larger bodies built up out of interstellar dust?
    - How do terrestrial and Jupiter-type planets form and interact with the disk?
    - What is the chemical state (pre-biotic?) of material that enters into planetary atmospheres?

- Debris disks (including our Solar System) containing cometary, planetary, and satellite bodies and atmospheres
    - History of our and neighboring solar systems
    - Finding and analyzing pristine material in comets
    - Secondary dust production

SAFIR can be expected to make key unique contributions in understanding star and planet formation by enabling a combination of the high spatial resolution with high sensitivity, large area photometry and high-resolution spectroscopy. The instrument and telescope requirements can be derived from the needs of star/planet formation science, which we outline below.

The dust continuum emission of star forming regions peaks at 60 – 120 μm. Thus, we require that SAFIR's imaging photometry cover the wavelength range from at least 40 - 200 μm to characterize the energy output of local star formation regions and to 1mm in order to characterize the energy output of distant star formation regions out to $z < 5$ (thus covering the bulk of star formation history). Studies of local star formation at spectral resolution R~3 is not a principal driver of photometer sensitivity requirements, but similar studies at epoch z=5 will be enabled by background-limited sensitivity. Typical scale sizes, distances, and separation of local sources of star formation require 2" spatial resolution (comparable to non-AO ground-based optical telescopes) at the wavelengths of interest. For example, a 200 AU diameter protoplanetary disk at 100 pc subtends a 2" angle, the diffraction limit attainable at 100 μm with a 10 m telescope. Such a disk would be resolved at 40 μm.

R~1000 dust spectral features in the far infrared/submillimeter wavelength regime enable dust crystallography during the various phases of star and planet formation. Of particular interest are the so-called PAH (polycyclic aromatic hydrocarbon) features. PAH dust spectral features appear in the 3-12+ μm regime but are difficult to uniquely interpret because they arise from e.g. C-H and C-C bending and stretching modes. The "large scale" floppy and torsional modes, which result in far infrared/submillimeter emission of large complex molecules and PAH-dust, will allow their unique identification. Even the lack of such features provides us with important constraints on the physical make-up of the dust. Because these dust features will be associated with a strong continuum emission, the sensitivity of an R~1000 spectrograph is an enabling constraint. ISO was unable to identify far infrared dust features, but the Herschel PACS and SPIRE spectrographs may be able to observe the strongest dust features. A sensitive background-limited R~1000 spectrograph such as that envisioned for SAFIR will enable the study of dust evolution during the early phases of star formation and planetesimal build-up.

The very strong 3-12 μm PAH features shift into the SAFIR window beyond z=7. The Spitzer Space Telescope has demonstrated that the 8 μm camera in particular is a sensitive tracer of star formation activity. Thus, background-limited sensitivity, which allows identification of these PAH features, will enable quantitative studies of star formation rate and metallicity history out to the reionization epoch.

Detailed line studies of many important diagnostic spectral lines are not possible from ground-based or airborne observatories (e.g. HD, H_2O/HDO, OH, fine-structure lines, bending modes of hot complex molecules, high-J CO), and submillimeter atmospheric "windows" do not permit access to all needed diagnostic lines at a given redshift. Detailed line profile shapes and centroid velocities of several key lines will enable the study of the structure and dynamics of circumstellar disks before, during and after planet-forming stages as well as the detection and mapping of large (pre-biotic?) molecules. Typically, one would require sub-km/s velocity resolution for such studies which translates into a spectral resolution requirement of R > 300,000, easily attainable with heterodyne receivers.

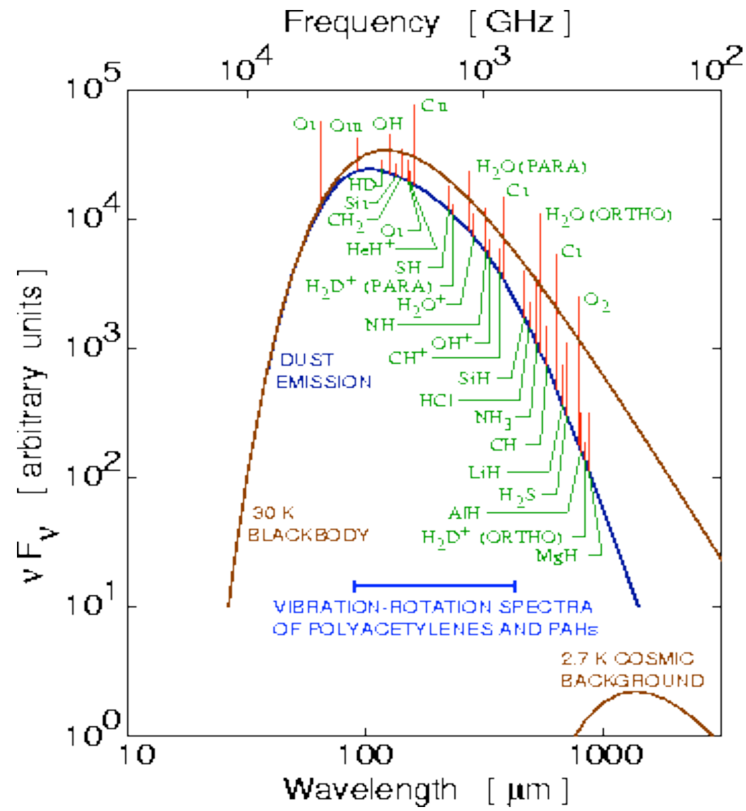


Figure III-15: Emission from a star-forming region (~70K) with spectral lines imposed on the dust continuum. Large organic (pre-biotic?) molecules can be expected to have torsional and floppy modes with bands in the far infrared, which – unlike the near infrared "PAH" features, allow their unambiguous identification.

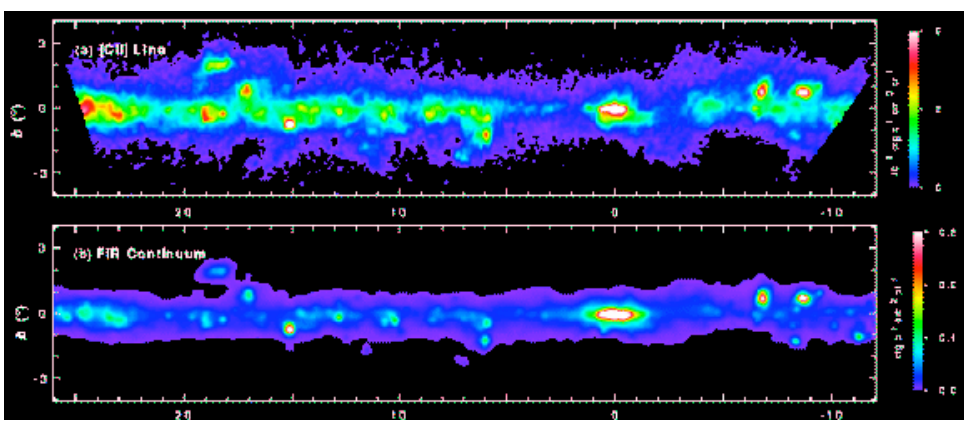


Figure III-16: CII at 158 µm, the strongest cooling line in the ISM, is visible everywhere in the galactic plane. BICE Galactic maps of CII at very low spectral resolution (top) and continuum dust emission (bottom). With the addition of velocity information from a high resolution SAFIR instrument, pieces along the line of sight can be separated, yielding a 3D picture of our galaxy.

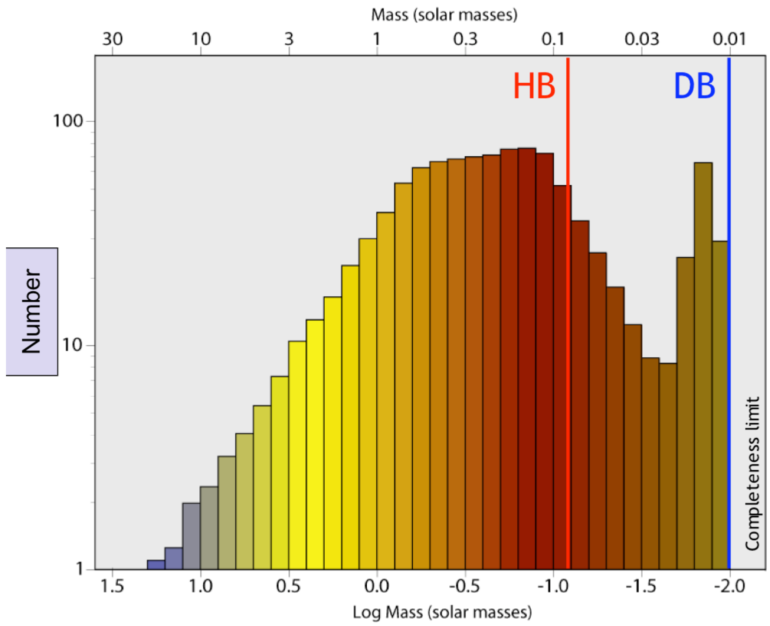


Figure III-17. The initial mass function of the Trapezium cluster (Muench et al. 2001). Stars more massive than 30 M_{sun} are absent simply because the cluster does not contain enough stars to have significant numbers at the high end of the mass spectrum. Observing the formation of high mass stars requires one to consider clusters at much greater distances than the Orion Trapezium. HB and DB denote minimum masses for hydrogen and deuterium burning.

From a collapsing cloud to a solar system?

The sequence of events beginning from the collapse of a gravitationally unstable clump within a molecular cloud and ultimately resulting in several planets, their satellites, comets, asteroids, etc. are still poorly known – primarily due to our inability to observe many of these events during the earliest stages of evolution. Surely this evolution strongly depends on the local environment and the individual parameters of the clump, in particular the clump's mass, magnetic field, rotation, and turbulent motions. Specifying the clump's Jeans mass, $M_J = 1\ M_{sun}\ (T/10\ K)^{3/2}\ (n_{H2}/10^4\ cm^{-3})^{-1/2}$ from the temperature and density of the clump gives only part of the picture. The initial mass spectrum resulting from the collapse of a massive molecular cloud easily spans some 4 orders of magnitude, from 0.01 M_{sun} to 100 M_{sun} (e.g. Figure III-17 above). What determines the final mass of the star cannot only be a function of the temperature and density. The collapse of massive molecular clumps can produce high mass stars, but the evolution is not simply a scaled-up version of low mass star formation. Outflows and radiation strongly influence the star formation process.

Figure III-18 depicts some of the relevant evolutionary time scales during the formation and evolution of stars. Stars more massive than about 5-8 M_{sun} are not optically visible as pre-main sequence stars, because they evolve quickly to core hydrogen burning while the star is obscured optically and is still growing in mass. Massive stars destroy their accretion disks on a time scale comparable to the accretion time scale and it is thus highly unlikely that planets can form around such stars.

Perhaps massive stars are formed through an entirely different process than that envisioned for their low mass counterparts: accretion onto a disk with subsequent central core growth via mass flow through an accretion disk. Do massive stars grow instead by coalescence of lower mass bodies?

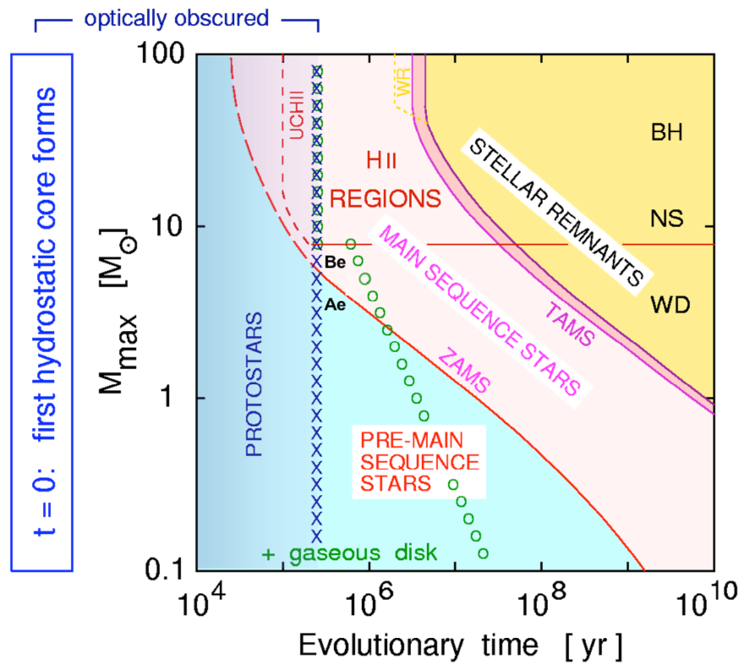


Figure III-18. Times scales for the evolution of single stars of given mass from the birth of the first embryonic hydrostatic core of about 1 Jupiter mass (t=0). Because stars both gain and lose mass during their evolution, the star's maximum mass is used for the "given mass". The time for accretion and rapid mass growth (and high obscuration) is given by the symbol "X". The mean lifetime of gaseous disks is given by the symbol "O". ZAMS = Zero Age Main Sequence; TAMS = Terminal Age Main Sequence (end of core hydrogen burning); UCHII = Ultracompact HII Region; WR = Wolf—Rayet star; BH = Black Hole; NS = Neutron Star; WD = White Dwarf.

$M > M_*/M_{sun}$	N_* ($\alpha = 2.35$)	N_* ($\alpha = 2.3$)
10	5400	6300
20	2000	2400
30	950	1200
50	390	480

Table III-1: Expected number N_* of stellar progenitors in the Galaxy currently in the accretion phase which eventually attain a final mass M_* above the given value. Two values for the IMF power law exponent α are given.

Table III-1 gives the number of stars in the entire Galaxy, which we can reasonably expect to be in the accretion phase, assuming a power law distribution at the high mass end. If we restrict ourselves to the local (< 1 kpc) disk, these numbers have to be decreased by a factor of about 300. Thus, we can expect to find only a single star more massive than 50 M_{sun} in our immediate vicinity. The relative scarcity and the short evolutionary time scales affect our ability to observe massive stars during the earliest phases of formation. Small number statistics are partially offset by higher luminosities – we can observe massive stars at greater distances, especially if we observe at longer wavelengths which are unaffected by dust extinction within the galactic disk. However, insufficient spatial resolution and confusion may be an issue; a large aperture telescope is imperative. SAFIR will have sub-arcsec resolution below 50 µm.

Due to high dust obscuration in the star-producing molecular clumps, observing star formation process during the earliest evolutionary phases is primarily a long wavelength (far infrared – submillimeter – radio) endeavor. Space-based far infrared and submillimeter studies complement and supplement ground-based radio and submillimeter studies through their unique ability to directly establish the total luminosity, certainly one of the most important parameters of the star forming system. Line studies of diagnostic spectral lines unique to the far infrared and submillimeter (e.g. HD, H_2O/HDO, OH, fine structure lines of atoms and ions, bending modes of hot complex molecules, high-J CO) permit precise characterization of the immediate environment of the forming massive star (e.g Figure III-19). Detailed line profile shapes and centroid velocities of several key lines will enable the study of the structure and dynamics of circumstellar disks, material infall, and outflow before, during and after planet-forming stages as well as the detection and mapping of chemical species, such as large (pre-biotic?) molecules. The ability to distinguish various velocity components along the line-of-sight is necessary to avoid confusion when studying complex processes in the galactic plane.

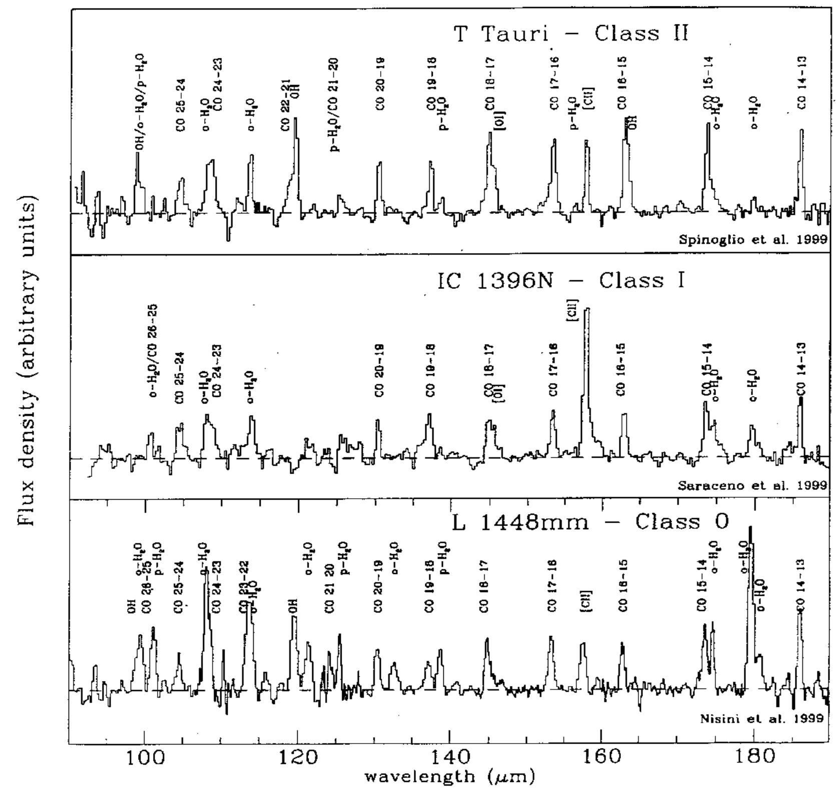


Figure III-19: ISO-LWS spectra of line emission from low mass protostars in different phases of development (oldest at top). These lines are all unresolved; many are highly blended

Exploiting a line shape requires a spectrometer with a minimum of 3 spectral resolution elements across the line (10-20 is optimal). The trade-off is between information and sensitivity. Spectral line widths are determined by thermal broadening and by line-of-sight motions of the emitting material. Typical temperatures range from less than 50 K to 10,000 K in star-forming regions, corresponding to thermal lines widths (for, say H_2O) of less than 0.2 km/s to ~3 km/s. Large molecules can have thermal widths that are much narrower. Line-of-sight velocities resulting from bulk motions and turbulence can range from sub-km/s (quiescent disks) to several 10 km/s (jets and outflows).
For the far infrared one has four choices of spectrometer types:

- Dispersive spectrometer (grating/prism or equivalent)
  - R = ~10^3 limited by grating size (R = length/wavelength)

- Fabry-Perot Spectrometer
  - R = ~10^4 limited by finesse of cavity

- Fourier Transform Spectrometer
    - $R = \sim 10^5$ limited by length of moving arm

- Heterodyne Spectrometer
    - R determined by LO spectral purity and backend spectrometer $\sim 10^{6\text{-}8}$ in far infrared

In order to attain sub-km/s resolution SAFIR requires either a superb FTS or a heterodyne spectrometer system.

Planet Formation

According to our current understanding, planets begin to form as soon as their parent disks are formed, which occurs while the protostar-disk system is highly obscured. The time scales and structure for this process are illustrated in Figures III-20 and III-21. Significant grain reprocessing can occur within a few thousand years, which should be detectable by SAFIR. Whereas dust can quickly sediment and coagulate into larger particles in gaseous disks – within several hundred thousand years, the growth of planetesimals into larger bodies requires a relatively long-lived gaseous disk – of the order of several million years. Thus, stars more massive than 3-5 M_{sun} may have difficulty producing Jupiters through accretion onto a solid core. The importance of the spatial resolution of SAFIR in understanding the structure of these disks is illustrated in Figure III-22.

It is conceivable, however, that Jupiter-type planets form through a process different than the core-accretion mechanism described above. Gravitational instabilities in the disks may produce Jupiter-like planets, ranging in mass from Saturn mass to values typical for brown dwarfs. Indeed, there may be no clear-cut distinction between the formation mechanism of brown dwarfs and Jupiters.

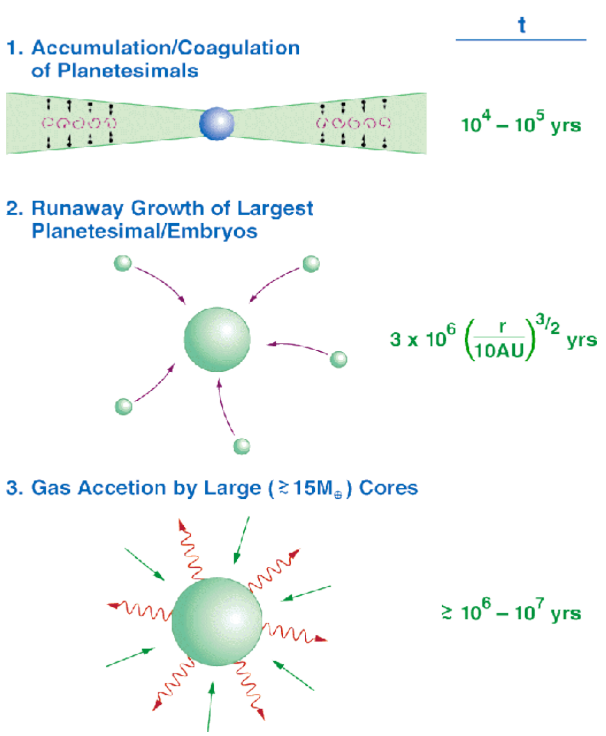
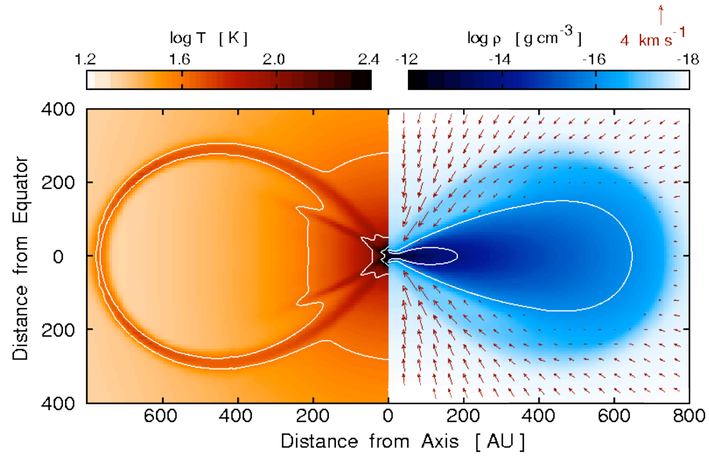


Figure III-20: Time scales for the main phases of planet formation: 1. Dust settles towards the disk midplane and coagulates into larger bodies as it is pushed around by internal gas motions; 2. Runaway growth of the largest planetesimals into planet-sized objects; 3. Accretion of gas onto a solid (Earth-like) core (according to the core-accretion model for Jovian planet formation)

Figure III-21: Numerical simulations of the collapse of a slowly rotating clump which produces a disk several thousand years after the first hydrostatic core is formed. The temperature structure is displayed to the left, velocity and density structure to the right. Note the occurrence of multiple accretion shocks.



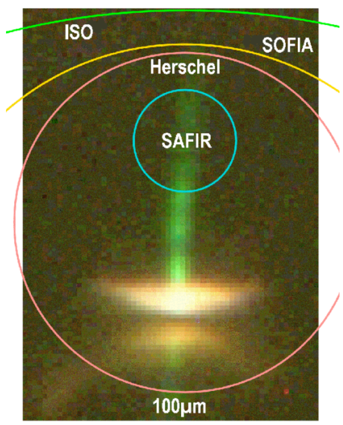


Figure III-22: The spatial resolution of SAFIR is compared to other future facilities at 100 µm. SAFIR's spatial resolution will allow detailed studies of nearby young star-disk systems, spatially resolving for the first time the line and continuum contributions of jets from the gaseous disks. The source in the picture is HH30 IRS. From analysis of the scattered light, the (obscured) illuminating source is estimated to be a M0 T Tau star with 0.2 L_o. HH30 is at a distance of 140pc.

SAFIR and the chemistry of the interstellar medium

A question central to our understanding of origins is how much of the organics in comets, asteroids and meteorites is pristine interstellar material and how much of it was formed by processing in the presolar nebula. This also has implications for the origin of the prebiotic material that initiated the development of life on Earth. In order to understand the chemical history of our own Solar System and that of extrasolar systems we need to understand the composition and history of the organic component of the interstellar medium and young stellar systems.

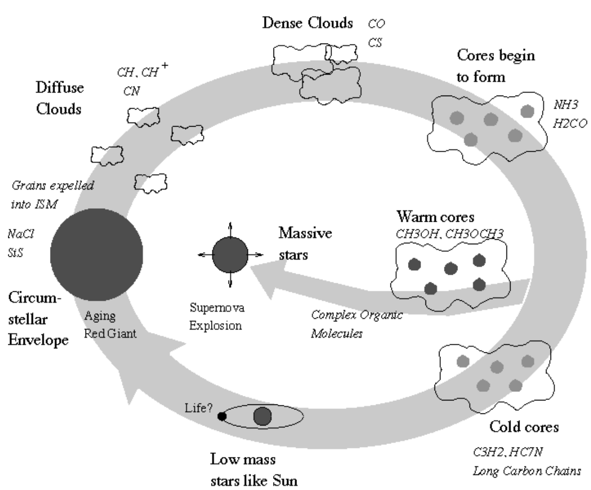


Figure III-23: Chemical processing in the interstellar medium.

The interstellar medium (ISM) consists of regions that cover a wide range of temperatures (10 – 10^4 K) and densities (100 – 10^8 cm^{-3}). It consists of mainly of gas with about 1% of the mass being in dust grains. The gas is mostly hydrogen with 10% helium and 0.1% elements such as carbon, oxygen and nitrogen. Other elements e.g. Si, S and Fe are even less abundant. Within these regions with diverse physical conditions complex chemistry can develop. There is evidence for a link between molecules formed in the dense prestellar cores and those observed in comets suggesting that chemistry in the ISM plays a role in determining the composition of solar system bodies (e.g. Irvine et al. 2000). Models have shown that PAHs, fullerenes and other complex organics can form in molecular clouds and are important for the chemistry and laboratory work suggests that they can be also be formed by UV photolysis of ices, but observationally it is not yet well known what degree of chemical complexity it is possible to achieve in cold interstellar clouds. This is something that SAFIR is well suited to explore. Understanding the chemistry in all regions of the ISM, from the formation of dust grains in the envelopes around late-type stars, to the formation of molecules in gas and ices contained in cold dark cores and the evolution of cores through collapse to form a young stellar object and protostellar disk and eventually planets (Figures III-23 and III-25) is crucial to understanding the chemistry of our solar system and other planetary systems. Figure III-24 illustrated examples of the carbon complexity in these chemical processes.

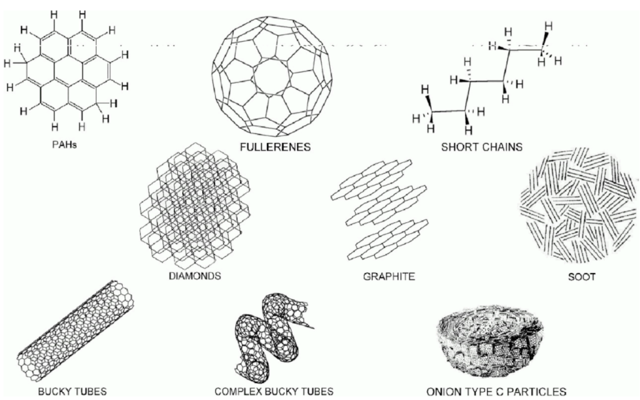


Figure III-24: Examples of carbon molecules likely to be present in the ISM and in young stellar systems (Ehrenfreund & Charnley 2000, Ann Rev. Astron. Astrophys. 38, 427).

Table III-2 shows the molecules that have been observed in the ISM to date. Most are organic, with carbon chains being more common than ring molecules. In addition, there is observational evidence for the presence of large molecules such as PAHs and fullerenes. There is also a tentative detection of glycine, the simplest amino acid, but this is still in dispute.

SAFIR will be able to make a crucial contribution to the understanding of interstellar chemistry by providing access to the wavelength region from 20 μm to 1 mm. Not only does this region contain transitions of important molecules such as HD, H_2O, HDO, OH and atomic fine structure lines which can trace the physical properties of the ISM and determine the contribution of molecules to the cooling of the gas, but it also provides the opportunity to detect many transitions of more complex molecules such as PAHs, fullerenes and amino acids (see Figure III-24 and Table III-2 for examples of organic molecules expected to exist in space).

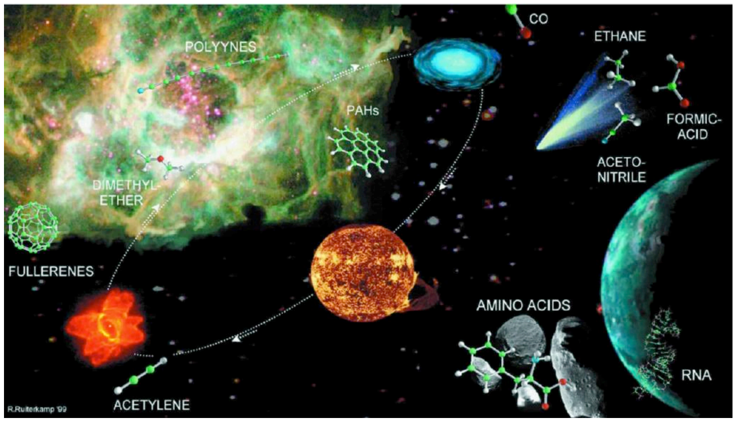


Figure III-25: The cycle of organic molecules in the Universe. Interstellar organics are formed in the interstellar gas, in stellar outflows and on dust grains. This organic material is integrated into solar systems and is partly chemically processed or destroyed. In the final stages of stars, dust and elements are returned to the ISM.

Number of Atoms							
2	**3**	**4**	**5**	**6**	**7**	**8**	**9**
H_2	H_2O	NH_3	SiH_4	CH_3OH	CH_3COH	CH_3CO_2H	CH_3CH_2OH
OH	H_2S	H_3O^+	CH_4	NH_2CHO	CH_3NH_2	HCO_2CH_3	$(CH_3)_2O$
SO	SO_2	H_2CO	CHOOH	CH_3CN	CH_3CCH	CH_3C_2CN	CH_3CH_2CN
SO^+	HN_2^+	H_2CS	HCCCN	CH_3NC	CH_2CHCN	C_7H	$H(CC)_3CN$
SiO	HNO	HNCO	CH_2NH	CH_3SH	HC_4CN	H_2C_6	$H(CC)_2CH_3$
SiS	SiH_2 ?	HNCS	NH_2CN	C_5H	C_6H	C_8^- ?	C_8H
NO	NH_2	CCCN	H_2CCO	HC_2CHO	c-CH_2OCH_2	CH_2OHCHO ?	C_9^- ?
NS	H_3^+	HCO_2^+	C_4H	CH_2=CH_2	C_7^-	l-HC_6H	
HCl	NNO	CCCH	c-C_3H_2	H_2CCCC			**10**
NaCl	HCO	c-CCCH	CH_2CN	HC_3NH^+			CH_3COCH_3
KCl	HCO^+	CCCO	C_5	C_5N			CH_3C_5N ?
AlCl	OCS	CCCS	SiC_4	C_6^- ?			$NH_2CH_2CO_2H$?
AlF	CCH	HCCH	H_2CCC	C_5S ?			
PN	HCS^+	$HCNH^+$	HCCNC				**11**
SiN	c-SiCC	HCCN	HNCCC				HC_9N
NH	CCO	H_2CN	H_3CO^+				
CH	CCS	c-SiC_3					**12**
CH^+	C_3	CH_2D^+ ?					c-C_6H_6 ?
CN	MgNC	CH_3					
CO	NaCN						**13**
CS	CH_2						$HC_{11}N$
C_2	MgCN						
SiC	HOC^+						**15**
CP	HCN						$(C_2H_5)_2O$?
CO^+	HNC						
HD	H2D+						
HS	CO_2						
FeO	SiCN						
O_2?	AlNC						
HF	KCN ?						

Table III-2: Observed interstellar and circumstellar molecules. ? indicates a tentative detection, c- indicates a ring molecule.

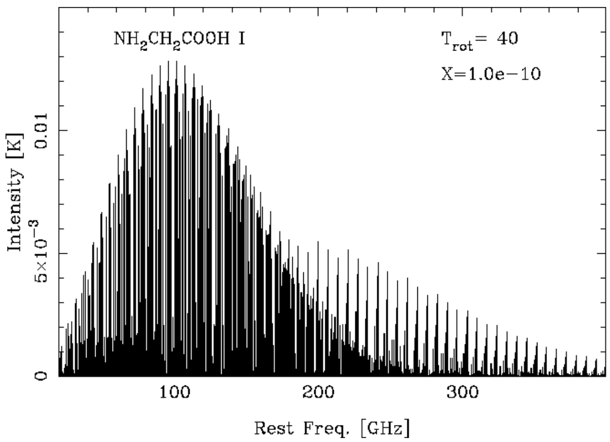


Figure III-26: Predicted millimeter spectrum of glycine demonstrating the forest of lines that are produced as a consequence of the large partition function. (J. Pearson, JPL)

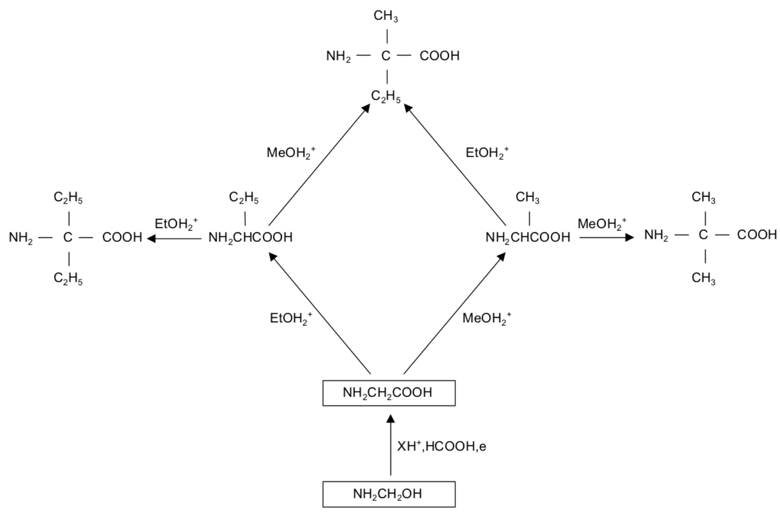


Figure III-27: Possible formation routes to amino acids starting from aminomethanol. Me represents a CH_2 group and Et a C_2H_5 group**.**

Several mid-IR lines between 3 and 12 µm have been attributed to the C-C or C-H bending and stretching modes of PAHs. However, these lines are not sufficient to allow PAHs to be uniquely identified or to establish the size distribution of such molecules. Complex molecules have complex structures and large partition functions which means that there are many ways for them to lose energy. Consequently a transition of a complex molecule is much weaker than those for simple molecules e.g. assuming the same abundance of glycine as for CH_3OH the strongest glycine transition is predicted to be 10 mK as opposed to 400 mK for methanol. The large number of transitions also means that many of them need to be observed in order for a positive identification to be made. PAHs and other complex molecules are expected to have large-scale flopping or torsional modes in the FIR and submillimeter – regions of the spectrum which will be accessible to SAFIR. A combination of the data from SAFIR with laboratory spectra will allow the unique identification of many of the organic molecules in the ISM.

Amino acids are important biological molecules which might have an interstellar origin. Laboratory data is available for glycine ($NH_2CH_2CO_2H$), the simplest amino acid, over the SAFIR wavelength range (Figure III-26). Amino acids have been shown to form in the laboratory when interstellar ices are irradiated by UV photons. Possible chemical pathways to their formation by ice chemistry are shown in Figure III-28. Complex molecules could also form in the gas phase. Figure III-27 shows a reaction network that starts from alcohols. Alcohols can form easily (and have been observed) in interstellar ices. In warm regions such as hot cores they can be desorbed and returned to the gas where they can drive a complex chemistry. Observations of complex molecules and their identification will provide us with clues as to their formation mechanisms and their ability to survive in the interstellar medium and through the star formation process (e.g amino acids are likely to be quickly destroyed by UV photons in highly irradiated regions, but may be able to survive in low UV fields or if embedded in ices). SAFIR is uniquely placed with an ideal wavelength coverage and high sensitivity to make a significant contribution to our understanding of the organic chemistry of the interstellar medium and to the question as to the origin of prebiotic molecules that led to the development of life on Earth.

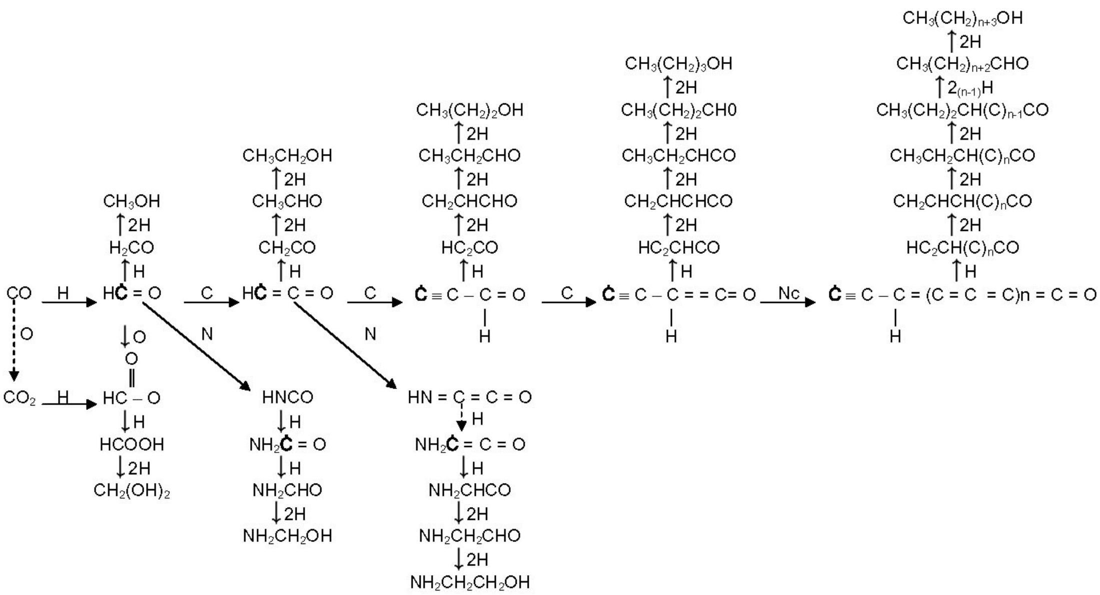


Figure III-28: Schematic illustrating the types of ices expected to be present in molecular clouds, and the chemical pathways that might form them.

Debris Disks and Signatures of Planets

Background

Long before extra-solar planets were discovered by the Doppler technique, there was indirect evidence of planet building from observations of debris disks around several main sequence stars. These debris disks, for which our own counterpart is the zodiiacal cloud and Kuiper Belt dust disk, were first detected in their thermal emission by IRAS (see Figure III-29). Stars harboring debris disks are too old to have remnants of the primordial disk from which the star formed. The timescale of dust grain removal due to Poynting-Robertson (P-R) drag for a solar type star is $\sim 10^5$ years, much shorter than the age of main sequence stars, $>10^7$ years, indicating that these dust disks are not primordial, but the result of mutual collisions between asteroid-like bodies or the evaporation of comets close to the star. The observation of debris disks indicates that planetesimal formation is indeed a common by-product of the star formation process. One would expect, therefore, that some stars harbor both giant planets and extended emission, as is the case of the Sun. Indeed, a recent study with the Spitzer MIPS instrument confirms that out of 26 FGK field stars known to have radial velocity planets, 6 show 70 μm excess at 3σ confidence level, implying the presence of cool material (<100 K) located beyond 10

AU. These stars, with a median age of 4 Gyr, are the first identified as having both well-confirmed planetary systems and well-confirmed IR excesses.

A sensitive far infrared observatory that can resolve such structures on these scales such as SAFIR would contribute strongly to our understanding of the environments in such solar systems. For the nearby stars that will be the targets of TPF, the number of detectable debris disks rises rapidly with sensitivity. SAFIR will provide, for TPF-identified planets, a strong assessment of the late bombardment rate for these planets, which is indicative of habitability. While evidence for such disks has also come from scattered light at optical and near IR wavelengths, such detections are biased to regions close to the central star, and do not provide particularly strong information about composition or mass. Reaching out into the thermal infrared, JWST will probe the hottest regions of these clouds. To the extent that these clouds are in thermal equilibrium with the radiation field from the star, long wavelength measurements are more revealing. With the spectrum of these excesses peaking at about 100μm, observations in the far infrared offer special advantages. Such measurements give access to the bulk of the heated dust, and also offer sensitivity to color temperature variations, and grain-size distribution. In particular, the mid- and far-infrared offer spectral probes of disk composition, including detection of entrained gas.

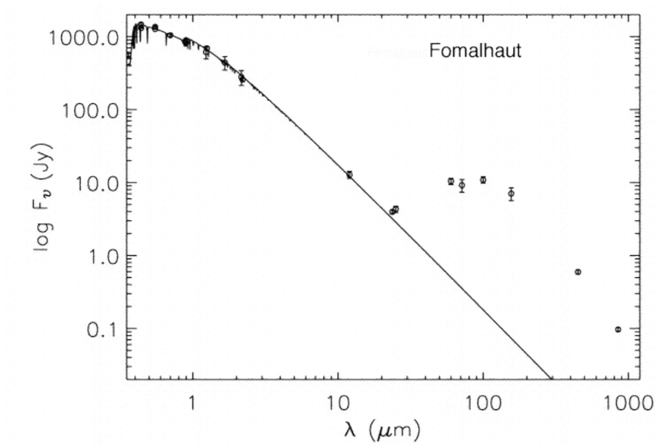


Figure III-29: The spectrum of Fomalhaut shows the component due to the debris disk, in excess over the photospheric light indicated by the solid line. This excess peaks at wavelengths of ~100 μm. Observations in the far infrared are thus especially useful for exploring these debris clouds. While ground-based submillimeter observations are possible, flux levels are much lower. (From Stapelfeldt et al. 2004 ApJS, 154, 458.)

Target Statistics

The spatial resolution requirements for useful debris disk studies are determined by the importance of having a useful sample size for characterization. It is clear from early Spitzer results, as well as ground based submillimeter results, on a handful of resolved debris disks that their morphologies are very different. Whether the variety that is represented is due to differences in solar systems around these stars or other factors is not at all clear. It is understood that for our own solar system the P-R clearing time is of order the replenishment time, which is the collision timescale for 10 km asteroids. As a result, it can be understood that the density and distribution of material in disks around solar systems like ours is likely to be highly stochastic. For those few disks that have been resolved by Spitzer at 70 μm, the morphology is remarkably different than seen at submillimeter wavelengths Figure III-30), so longer wavelength studies from the ground do not appear to provide the same information as a far infrared telescope in space. This is one of the surprises of the Spitzer mission. SAFIR will offer more than an order of magnitude higher spatial resolution than Spitzer.

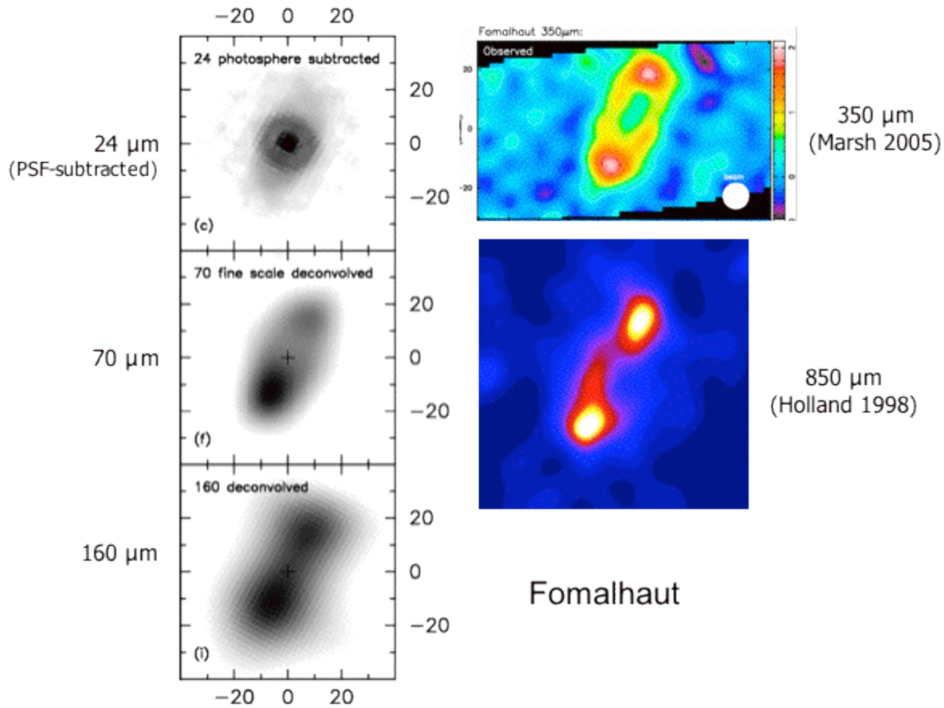


Figure III-30: Spitzer far infrared observations of Fomalhaut at left compared with ground-based observations at right show the dramatic differences at different wavelengths. Ground-based studies of debris disks cannot, by themselves, fully represent the thermal and structural complexity of these sources. The spatial resolution of SAFIR at 100 μm would exceed that in the 850 μm map and in the Spitzer 70 and 160 μm maps by a factor of approximately ten. (Breyer et al. 2004)

A 10 m aperture of SAFIR allows to have a spatial resolution of ~1" at 30 μm. Disk inner holes created by planets more massive than Neptune and located at 20–30 AU will be resolved out to a distance of 20–30 pc, large enough to contain a significant number of debris disks systems to allow a statistical approach to the study of the frequency of long period planets.

Assuming thermal equilibrium of grains of emissivity ε with albedo a with a central star of luminosity L_* at a distance D, the characteristic scale size for thermal emission from the grains is given approximately by

$$\theta_{char} = (\lambda^2/Db^2)(L_*(1-a)/4\pi\sigma\varepsilon)^{1/2}$$

Where b is the constant in the Wien law. Using this equation, one can predict a characteristic angular scale for the disk at this wavelength as a function of stellar luminosity. It is noteworthy that the angular scale increases as the square of the wavelength, while the spatial resolution of a given aperture size decreases as the first power of the wavelength, so there is value in making these measurements at long wavelength. Comparing these calculations with the solar neighborhood population, it is found that SAFIR's 2" resolution at 100 μm should provide a reservoir of about five hundred main sequence stars that, if they have debris clouds, would be resolvable. This compares with about one hundred for Herschel, and a few dozen for Spitzer. It is furthermore understood that a substantial fraction of stars in the solar neighborhood do, in fact, have well developed debris disks. While Spitzer results show a strong correlation of debris disk emission with age, distance-limited surveys such as the TPF-target list MIPs survey of Beichman et al (2004 ApJ 622, 1160), and more general ISO work show that of order 10% have debris disks that are significantly more massive than that of our solar system. This kind of spatial resolution in the far infrared should thus provide access to of order fifty resolvable disks more massive than our own. Those very few disks that have already been reasonably well resolved at 350 μm with 5" resolution, such as Vega, Beta Pic, Fomalhaut, and Epsilon Eri, will reveal dramatically more detail when viewed with arcsecond-scale resolution closer to the spectral peak.

A holy grail of debris disk studies is to understand how unusual our solar system is. At a nominal distance of 10 pc our solar system would have a 100 μm integrated emission of about 100 μJy, corresponding to $L_{dust}/L_{*} \sim 10^{-7}$. Assuming a "Fabulous Four" debris disk scale size of order 300 AU diameter, this would correspond to ~2 μJy per 2" pixel. This surface brightness is, in fact, detectable by the baseline background-limited 10 m SAFIR with even modestly cold telescope temperatures and an observatory venue within the inner ecliptic where our own zodiacal cloud contributes strongly to the background. Since the zodi in our solar system is fairly uniform in the vicinity of the Earth, it is a trivial result that the surface brightness of an extrasolar debris cloud like our own has the same brightness per pixel as our own zodi background at high ecliptic latitude.

With recognized debris disks having angular extents up to 1-2 arcminutes, a SAFIR imager should have a format size of at least 200 x 200 in order to fully sample the debris cloud around a single system. Short wavelength leaks in the bandpass filters will need careful management. Multicolor imaging will give important constraints on grain temperature and grain size distribution. The former will test grain heating mechanisms, and the latter will pertain strongly to the kind of bombardment that a planet in that system would endure, and help define the survivability of life on such a planet.

Coexistence of Planets

Another hint of the co-existence of debris disks and giant planets come from high-resolution images of debris disks both in scattered light and in thermal. Some of these images show the presence of density structure, such as gaps, arcs, rings, warps and clumps of dust. These features are likely to be the result of gravitational perturbations by one or more massive planets on the dust disk. There are several mechanisms by which giant planets can sculpt the disks: (a) capture of dust in mean motion resonances with the planet, as the dust particles drift toward the central star due to Poynting-Robertson drag; (b) resonance capture of dust producing planetesimal due to planet migration; (c) secular planetary perturbations; and (d) gravitational scattering of dust particles by the planet. It is important to notice that debris disk structure is sensitive to the presence of long period planets, complementing a parameter space not covered by radial velocity or transit studies, and understanding of the orbits of long period planets is fundamental for the study of the stability of orbits in the habitable zone, where terrestrial planets could form and survive. Therefore, even if the planets are not directly visible, we can learn about the diversity of planetary systems by studying their dusty "fingerprints". It is noteworthy that recent results from Spitzer show that debris disks are somewhat what more common around stars known to have radial velocity-detected planets.

The dust density structure carved by giant planets affects the shape of the disk Spectral Energy Distribution (SED), that depends on the grain properties (chemical composition, density and size distribution) and the mass and location of the perturbing planet. The SED of a debris disk with embedded giant planets is fundamentally different from that of a disk without planets, the former showing a significant decrease of the near/mid-IR flux due to the clearing of dust inside the planet's orbit. The SED is particularly sensitive to the location of the planet, i.e. to the area interior to the planet's orbit that is depleted in dust.. However, there are some degeneracies that can complicate the interpretation of the SED in terms of planet location. For example, the SED of a dust disk dominated

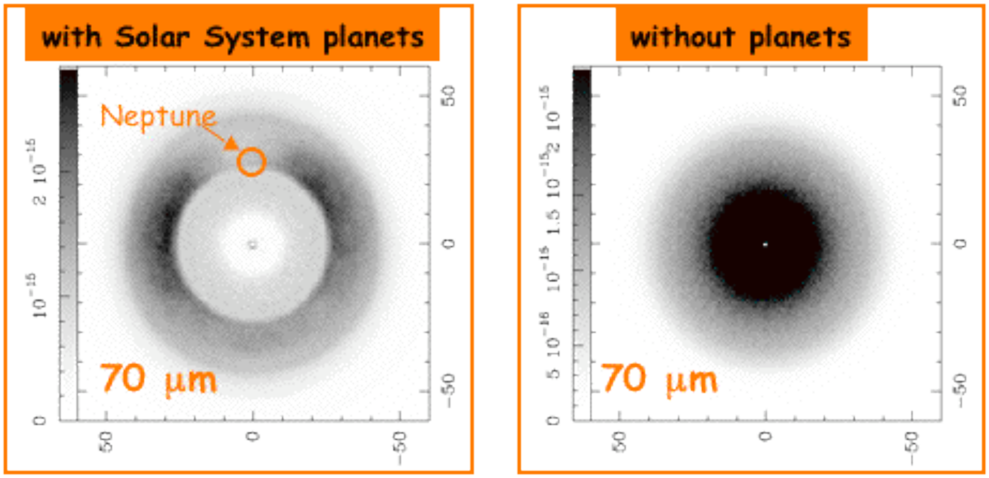


Figure III-31: The figure illustrates the clearing of debris in a solar system with massive planets around a solar type star. A Neptune-like body at 30 AU (Jupiter-like body also enclosed) around a solar neighborhood star would clear a region easily resolved by SAFIR. Note resonant clumps at the L3 and L4 regions. Compare with an azimuthally and radially uniform debris cloud at right. (Moro-Martin et al. 2005 in press)

by weakly absorbing grains (e.g. Fe-poor silicates) has its minimum at wavelengths longer than those of a disk dominated by strongly absorbing grains (e.g. carbonaceous and Fe-rich silicate). Because the SED minimum also shifts to longer wavelengths when the gap radius increases (owing to a decrease in the mean temperature of the disk), there might be a degeneracy between the dust grain chemical composition and the semimajor axis of the planet clearing the gap. The conclusion is clear: if we want to constrain the planet location, high resolution images that spatially resolve the disk are optimal, as indicated by simulations in Figure III-31.

One of the main features predicted by the dynamical models is the depletion of dust inside the planet orbit due to gravitational scattering by the planet. Dynamical models show that for a planet in a circular orbit with semi major a_{pl}, the radius of the depleted region is between $0.8a_{pl}$ and $1.25a_{pl}$, allowing us to constrain the semimajor axis of the planet from the sizes of the observed inner gaps. The models also show that the dust depletion factor (i.e. the ratio between the dust density inside and outside the depleted region) depends largely on the planet mass when $1M_{Nep} < M < 3M_{Jup}$, so in principle one should also be able to roughly constrain masses of planets at large astrocentric distances from observations of the density discontinuity at the inner edge of the disk.

While the disk SED can only yield information about the radial density distribution of dust particles (because the grain temperature depends only on the distance to the central star), the high resolution images will allow us to study the disk's azimuthal structure. Because this resonant structure depends on the planet mass and the location of the planet along its orbit, its study can also serve to constrain these planetary parameters. A constraint of the planet location can help the planning of observations aimed for direct detection. Observations at different epochs can detect proper motions in the dust features, that together with dynamical models, can serve as a test of the perturbing planet hypothesis (i.e., is the disk structure carved by a planet?). Low resolution spectroscopy with SAFIR of even distant and spatially unresolved debris disks will thus provide clues about the debris disk structure and provide clues about massive planets that are hosted within.

While clean grains show weak, though distinctive features (Figure III-33), biogenic and water ices have strong bands that will be accessible to SAFIR (Figure III-32). For example, intermolecular lattice translational modes in water ice at 44, 62, and 154 µm will be observable in such disks. SAFIR will probe chemical fractionation in these ices, most likely from cometary material, and by tracing both ices and their gaseous counterparts, will provide information about the evaporation of those ices.

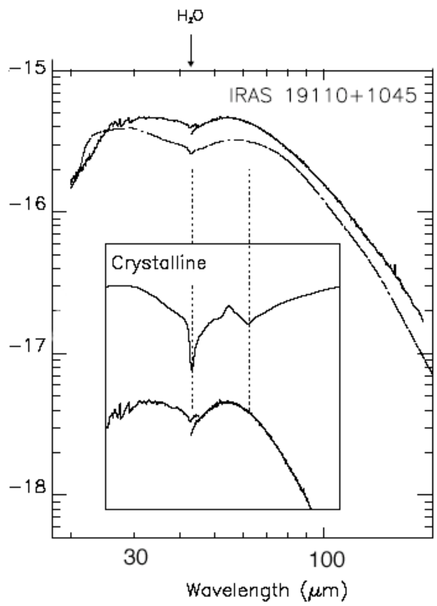


Figure III-32: ISO spectrum of the protostar IRAS19110+1045 (Dartois et al. 1998, A&A, 331, 651) shows the clear signature of crystalline water ice. Such ice mantles on grain cores will offer important probes of grain characteristics in solar systems around other stars.

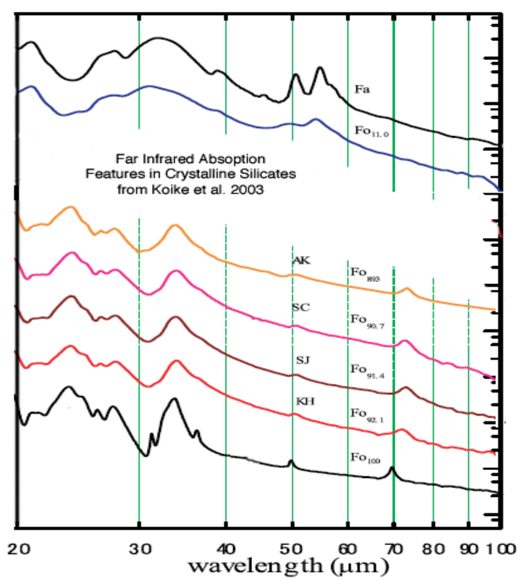


Figure III-33: Laboratory spectra of Mg and Fe-rich crystalline silicates (Forsterite and Fayalite) show distinctive bands in the far infrared, the details of which are sensitive to composition and annealing. These spectra are taken from the compilation by Koike et al. 2003. The 20-30 µm bands are being seen in Spitzer spectra of asteroids. These bands will be useful as a probe of refractory chemistry in circumstellar disks.

Operational Considerations

Debris disk science does not put any difficult operational constraints on acquisition and pointing. In fact, the association with a bright visible star makes fine tracking at short wavelengths more easily achieved than for most sources. The target stars are spread more or less uniformly over the sky, and while repeated measurements for confirmation are desirable, there is no particular constraint on time separation. Debris disk science does derive value from a long-lived observatory, however. Orbital resonance models predict structural changes (such as position angle of a resonance-driven clump) on a time scale of order that of the orbital period of the planet that determines the resonance, and we might hope to look for changes on this decadal time scale. One strong justification for observational repeats of debris disk imaging on a time scale of years is that it allows removal of confusion by background galaxies which, at the μJy level, will be significant at 100 μm. Ideally, SAFIR would allow confirmation imaging on a time scale that would spatially resolve the proper motion of the system. For Vega and Fomalhaut, with motions of order 0.4”/yr, a several year baseline would suffice.

SAFIR and the Role of Other Telescopes for Debris Disk Work

Ground based submillimeter telescopes are now in the planning stages (e.g. AT-25, LST) that will offer 2” resolution at 200-350 μm, and a far infrared telescope that can achieve that performance would be well matched to these scientifically. The spatial dynamic range provided by this resolution on at least these stars offers a good match to the detail in debris disk modeling that is now being done – detail that would directly test current theories of planet stirring and resonant clumping. This resolution corresponds to a 10m-class telescope. As it turns out, this resolution is similar in angular extent to the exclusion zone of TPF-C, such that a 10 m SAFIR will be investigating debris disk structure in the same radial regions (if not at the same resolution) as TPF-C.

ALMA will eventually allow us to image debris disks with an order of magnitude higher spatial resolution (10 milliarcseconds) than VLA and HST, in systems which are more than an order of magnitude fainter; i.e. it will be able to search for analogs of our own Kuiper Belt dust disk. The reason why (sub)millimeter observations of debris disks with ALMA are not sufficient is because spatially resolved images of the Vega disk taken by Spitzer have taught us that mm and mid/far-IR observations can be very different from each other, and both need to be considered in the interpretation of the system. Different wavelengths trace different particles sizes, giving complementary information about the dynamics of dust in planetary systems. Large particles dominate the emission at longer wavelengths, and their location might resemble that of the dust producing planetesimals; dynamical models show that the disk structure is more pronounced for these larger grain because the trapping in resonances is more efficient when the drag forces are small. The small grains dominate at shorter wavelengths, they interact with the stellar radiation field more strongly so that their lifetime in the disk is shorter, and therefore their presence may signal a recent dust-producing event (like a planetesimal collision). As noted above, flux levels near 1 mm wavelength for debris disks are much lower than in the far infrared, so even ALMA will be sensitivity challenged at these long wavelengths.

About 15% of the Spitzer GTO programs, and two out of six Legacy programs are focused on the detection and characterization of circumstellar disks around nearby stars. Imaging and spectroscopy is used to study the spatial structure and composition of the disks, and the frequency and duration of the protoplanetary disk phase (constraining the probabilities and timescales for the formation of the major planetary bodies). The SEDs obtained with Spitzer are capable of diagnosing the radial distribution of dust, and in some cases they show inner gaps that may be the imprints of embedded giant planets. SAFIR, with its superior spatial resolution, will be able to build up on the Legacy of Spitzer by studying these potential planet-harboring systems in unprecedented detail. As mentioned above, debris disk structure is sensitive to the presence of long-period planets, complementing a parameter space not covered by other methods. It can therefore allow us constrain the orbital parameters of long period planets, which in turn determines the stability of orbits in the habitable zone of the star, where terrestrial planets could form and survive. In other words, the study of debris disk structure with SAFIR can help us identify the stars that could be potential targets for TPF/Darwin.

KBOs in Our Own Solar System

Individual KBOs in our solar system, with spectral peaks at 70-100μm, will be well studied with SAFIR. The prime task on the path to exploring this new frontier of planetary science according to the COMPLEX study is to document fully the chemical and physical makeup of the objects that compose these trans-Neptunian region objects. In fact, the baseline SAFIR at 50-100 μm will be sensitive to populations of KBOs (~1 km at 35 AU, ~70 km at 100 AU) that are far smaller than can be detected by reflected visible light (~50 km at 35 AU). Surveys with SAFIR, or even just attention to the presence of moving objects at low ecliptic latitudes in deep SAFIR images, would dramatically enhance our understanding of the mass distribution and radial extent of KBOs. The baseline 10 m size for SAFIR is critical to these studies, as models of KBO evolution predict a break in population characteristics (e.g. Kenyon & Luu 1998 AJ 115, 2136) at diameters of order 10 km. While smaller infrared telescopes could in principle see such KBOs source confusion would dramatically compromise our ability to distinguish them. For this reason, the 10 m baseline aperture size of SAFIR, diffraction limited in a 2" beam at 100 μm, is needed. The small diffraction spot is also of value in mitigating the effects of strong zodi continuum emission at the low ecliptic latitudes where these sources are concentrated. It is worth noting that assuming that the size distribution can be extrapolated to 10km, the expectation value for such KBOs is about $0.3/\text{arcmin}^2$. As a result, most low ecliptic latitude SAFIR fields will contain a detectable KBO!

The motion of these KBOs across the sky is small, and a circular orbit at this distance would correspond to degradation of the diffraction spot in about five minutes of integration. The baseline SAFIR would achieve S/N ~ 100 broad-band on a 10 km target in this time, providing excellent color temperatures if observed in two bands. While a special moving source tracking mode is thus not particularly important nor challenging, this kind of motion allows for recognition of a KBO with a convenient observational cadence. Two consecutive five minute exposures with fixed source tracking would allow straightforward identification of a KBO, and distinction from background stars and galaxies.