Journal of Cosmology, 2010, Vol 12, 3904-3911.
JournalofCosmology.com, October-November, 2010

Mars Base First: A Program-level Optimization for Human Mars Exploration

Douglas W. Gage, Ph.D., XPM Technologies, Arlington, VA 22201-4637

Abstract

Much effort has been spent over the past two decades on the development of various plans for the human exploration of Mars, and, while the allocation of functions to specific vehicles has differed, consensus has formed around a conjunction mission (nominally 18 months on the surface) with pre-emplacement of some assets and some level of in-situ resource utilization (ISRU). The exploration program as a whole is envisioned as a sequence of several independent missions, each sent to a different area of interest on Mars, analogous to the Apollo approach to lunar exploration. Creation of a quasi-permanent Mars base is envisioned as following later, at a site to be identified during the initial exploration campaign. This paper argues that a better initial human exploration campaign – i.e., one which will ultimately maximize scientific payoffs and minimize costs and risks, both to crew safety and mission success – is one in which multiple (nominally two, perhaps three) crews land at a single site and live and work in a small underground base habitat constructed in advance by robots. They stay on Mars for an extended period – nominally 96 months, returning not on the first, but on the fourth Hohmann opportunity. They explore an extended area using pressurized rovers, perhaps driving to additional robotically-constructed underground habitats. Such a program-level architecture affords a continuous human presence on Mars, provides better shielding from radiation, reduces the number of crew transits from and to Earth, greatly reduces the maximum mass requirement for EDL (Entry, Descent, and Landing – we would land a crew capsule, not a full habitat), permits deferred development of the return vehicle (which could otherwise be a schedule-limiting element), and allows an initial unmanned return vehicle test supported by "ground crew" to return samples to Earth. Adoption of a "base-first" exploration program will require us to acknowledge and engage the real challenges to the human exploration and colonization of Mars – maintaining the safety, health, productivity, and happiness of a very small population of humans on the surface of Mars for an extended period of time. Apollo/Saturn proved that powerful rocket systems can be developed in less than a decade – but the Mars surface stay presents many specific technical and non-technical challenges that have nothing to do with "rocket science." Now is the time to start thinking seriously about these issues.

Key Words: Mars base, human spaceflight

1. INTRODUCTION

In the early 1990s, in response to the demise of the first President Bush's proposed Space Exploration Initiative, or SEI (Synthesis Group, 1991), Martin Marietta engineers Bob Zubrin and David Baker developed a mission concept that came to be called "Mars Direct" (Zubrin, 1996; Gill, 2005). Several of its key features have since been incorporated into NASA and other Design Reference Missions, or DRMs (Rapp & Andringa, 2005; NASA, 1997, 1998, 2001), including the recently released Design Reference Architecture (DRA) 5.0 (NASA, 2009): (1) a conjunction mission comprised of a 6+ month transit to Mars, approximately 18 months spent on the surface of Mars, and a 6+ month return; (2) pre-emplacement on Mars of unmanned assets, including an Earth Return Vehicle (ERV) or Mars Ascent Vehicle (MAV); and (3) in-situ resource utilization (ISRU) to generate fuel (methane) and/or oxidizer (liquid oxygen) from Martian atmospheric carbon dioxide and hydrogen possibly extracted from Martian water. While the partitioning of functional elements into specific vehicles differs among the various proposals, this nominally 30+ month conjunction mission profile leveraging pre-emplacement of resources and ISRU now represents the consensus of thinking both inside and outside NASA.

Since expected high costs make it unlikely that we will invest in developing a system capable of taking humans to Mars and then use it for just one 18 month surface-stay mission – NASA's DRA 5.0 assumes three consecutive missions (NASA, 2009, p. 20) – we should consider how a sequence of manned Mars flights should be configured to create a rational program – one which maximizes scientific payoff consistent with minimizing costs and risks (maximizing crew safety). The Apollo program provides a baseline for comparison: a series of six independent sorties were made to different locations on the lunar surface in the three and a half years between July 1969 and December 1972, with surface stays ranging from 22 to 75 hours and total mission durations from launch to splashdown of between 8 and 13 days. Launch windows occurred every month when the lighting at the planned landing site was correct, so the factor limiting the pace of launches was the assembly and checkout of the required Saturn launch vehicle stages and Apollo command, service, and lunar modules.

We cannot fly our Mars missions this way, however, since the possible launch windows occur only every 26 months. Moreover, in order to maintain effective and efficient operations of the hardware production lines and launch and mission control facilities, it will be necessary to launch during every conjunction launch window. This means (1) that the second mission crew will launch from Earth two months after the first has begun its return trip from Mars, (2) that for a period of four months we will have two crews in space, and (3) that the second crew will not arrive on Mars until eight months after the first crew has departed. This suggests a fairly obvious opportunity: if we want to place a crew at point B on the surface of Mars, and already have a crew at point A, why should we fly the first crew hundreds of millions of km back to Earth, even as we send a second crew all the way from Earth? Why not send a capable ground transport vehicle and have the first crew drive it from point A to a robotically prepared site at point B? This, of course, would require that the two points be within driving range of each other, perhaps 500-1,000 km.

Manned spaceflight is inherently dangerous, especially in its launch and reentry phases – witness the fates of Challenger, Columbia, Soyuz 1, and Soyuz 11 (NASA, 1997, p. 3-13). While proper operation of a life support system is critical on the surface of Mars as well as in space, a spacecraft requires many more active systems than a ground habitat (e.g., vehicle stabilization). Moreover, radiation and meteoroids pose additional risks in space, and zero-gravity is a major complicating factor. On a per-hour basis, humans inside a habitat on the surface of Mars are likely to be far safer than humans in space – and it should be relatively easy to make them safer still. Since even the basic conjunction mission calls for more time on the surface of Mars (18 months) than in transit (12-13 months), it will pay to invest heavily in making the surface stay as safe as possible. In fact, we have a "virtuous cycle": (1) the longer we are planning to stay on the surface of Mars, the safer we can and should make it, and (2) the safer it is on the surface of Mars, the longer we can and should plan to stay.

2. THE BASE FIRST PROGRAM

Consider the following program plan: an initial crew (of 4 to 6 people) is launched to a carefully prepared site, and remains on the surface of Mars for a full 96 months, returning on the fourth minimum-energy Hohmann opportunity. A second crew follows to the same site 26 months later, and also stays for 96 months. Thus, we have 8-12 people at the base for a period of 70 months, we have no gaps in crew presence on Mars, and the operations team never has to deal with two crews in transit at the same time. We continue to build assets at the initial landing site, expanding to a second site only when the first base has achieved true critical mass. Instead of launching an ERV or MAV to Mars 26 months before the first crew, we send it 26 months after the second crew. Moreover, the first return launch of the ERV can be an unmanned test flight, carrying Mars samples back to Earth, and we will have the advantage of having "ground crew" to service the launch. Finally, because we are delaying the launch of the first ERV, we can also defer its development, thus employing a smaller team of rocket developers for a longer period of time (Gage, 2009).

	Mars Base First Program (2-96 month surface stays)	5 Conjunction missions (18 month surface stays)
Time from first crew arrival to last departure	122 mo	122 mo
Time 2 crews on Mars	70 mo (1@70 mo)	--
Time 1 crew on Mars	52 mo (2@26 mo)	90 mo (5@ 18 mo)
Gap: 0 crews on Mars	--	32 mo (4@8 mo)
# crews launched	2	5
EDL mass requirement	6-10 tonnes	40 tonnes
First ERV/MAV launch	52 months AFTER first crew launch	26 months BEFORE first crew launch

Figure 1. A capsule comparison of the "Base First" plan proposed here with a program consisting of a sequence of five independent conjunction missions.

2.1 Extended missions dictate early base establishment. It is clear that supporting a crew of 8-12 people on the surface of Mars for 10 full years is a very different proposition from supporting 4-6 people for 18 months. NASA's DRA 5.0 and other DRMs all the way back to the Mars Direct plan envision the crew living inside a habitat sitting on the surface. EDL to get this nominally 10-meter diameter 40-tonne "tuna can" unit to the surface of Mars in one piece represents a major technological challenge. The alternative approach proposed here for the "Base First" plan would have the crew live and work in underground tunnels constructed by robots before the crew arrives, creating a true Mars Base. Living underground would provide much better protection than a surface habitat against radiation, which, while much less intense on the Martian surface than in interplanetary space, is much more intense than on Earth or in low earth orbit (LEO) (Rapp, 2006). This approach also carries the additional advantage that a crew-landing vehicle could be much smaller than the 40 tonne tuna can habitat, greatly reducing the structural mass that would have to be brought from Earth and radically simplifying the EDL problem (Gage, 2009). For comparison, the Apollo Command Module (CM) weighed 6 tonnes, while the Apollo Lunar Module (LM, nee LEM), including both descent and ascent stages, was 15 tonnes.

The underground space would be quickly and continually expanded to eventually include living, sleeping, and dining areas, galley, pantry, and garden; medical/dental clinic with mini-intensive care unit, exercise facilities (gym and track), spa, and swimming pool; medical, biological, chemical, and geological laboratories; manufacturing and repair shops; storage for food, other supplies, spare parts, and collections of samples. Many critical systems will be required to support human life and mission operations, including thermal control, air, water, waste, computing, and communications, but these various subsystems can be installed in the constructed underground base in a much more loosely coupled manner than would be possible in a tightly integrated habitat transported from Earth, thus simplifying component repair and replacement.

3. THE CHALLENGES OF THE BASE FIRST PROGRAM

The concept of a base on Mars is obviously not a new one: Google searches for "Mars Base" or "Mars Colonization" yield a plethora of diverse links. The "Mars Homestead" project (Mars Foundation, 2010) seems to be pursuing (at the "enthusiastic volunteer" level) a fairly reasonable broad-based exploration of many of the issues involved in establishing a permanent human presence on Mars. However, a common thread among mainstream efforts is the (usually implicit) assumption that the establishment of a base will occur only after a sequence of manned sorties to identify the best location for the base or colony. The key arguments in this paper are that **(1)** we should plan for a Mars base beginning with the first humans we send to Mars: we should plan to send fewer people to Mars, but have each of them stay much longer, and **(2)** we should invest heavily up-front in developing and refining the surface segment of the human Mars mission, because travelers to Mars will spend (much) more time on the surface of Mars than in space, and this is where the critical challenges and payoffs lie. This will require either that NASA move well beyond its traditional focus on space transportation systems, or that one or more other entities assume a leadership role in the Mars exploration enterprise.

3.1 Can humans survive and succeed on a ten-year mission? Some may object that a mission profile calling for an eight-year stay on the surface of Mars (and ten years away from Earth) is unreasonable – that the psychological stresses of living in such a small isolated group for so long would put the success of the mission, if not the crew's survival, at unacceptable risk. However, the history (and especially the prehistory) of humanity is one of many small groups of people migrating into the unknown with no intention of returning, and, in fact, informal surveys suggest that many people would be willing to sign up for a one-way trip to Mars (Krauss, 2009). We find many examples of small groups that have successfully lived in nearly constant isolation, including bands of hunter-gatherers, Inuit family groups, pre-20th century ship crews, castaways, and some soldiers and prisoners.

However, while humans on Mars will be physically isolated from Earth, they will have high bandwidth connectivity to the rest of the humanity (albeit with a 6-44 minute round trip latency). They need not be lonely; the World Wide Web will grow into the Solar System Wide Web. But we must thoroughly explore the full range of issues associated with long-term connected-but-physically-isolated living, including understanding how and how well high-bandwidth network communications can compensate for the lack of physical contact, and develop an experience base on Earth before we dare send people on such a mission. Since it is likely that the success of the mission may depend on the "chemistry" of the specific personalities involved, it may be that a crew should begin living together as a coherent group (if not in full isolation) well before their launch. The psychological and psychiatric issues associated with spaceflight have been studied since the beginning of the space age; see, for example, Kanas & Manzey, 2003; Kanas & Ritsher, 2005; and Johns, 2004.

Since living beneath five meters of regolith will mitigate the radiation hazard on the surface, the principal physiological challenge posed by the base-first mission (beyond those posed by a 30-month conjunction mission) is the loss of bone density and strength associated with the outward and return 6+ month zero gravity transits and eight years of 0.38 g Mars gravity. A focused exercise regimen, possibly combined with dietary modification, should at least partially mitigate these effects (Keyak et al. 2009), and at some point it might be possible to install a one-g centrifuge in the base. Long-term exposure to a low-pressure high-oxygen atmosphere in the base habitat, which could be adopted in order to reduce EVA prebreathe time (Gage, 2006; NASA, 2001, p. 20), would constitute a second physiological risk factor – but this is a risk which can be evaluated by experimentation on Earth.

An advantage of the base-first exploration strategy is that it will allow people to extend their stay on Mars, which would be absolutely necessary if the ERV or MAV could not be made ready during the return launch window, and might be desirable in other cases – imagine that the crew exobiologist on the first conjunction mission were to discover living Martian life just a few weeks before she is scheduled to return to Earth. And, of course, one of the classical planetary exploration science fiction tropes (e.g., Landis, 2000; Varley, 2005) – is that, when it is time to return to Earth, one or two characters (usually a couple) announce "we're going to stay."

3.2 Required technologies and tools Viewed from an engineering perspective, it is clear that a Mars base will constitute a complex "system of systems", one whose development will involve a large number of technical disciplines, and this fact must be explicitly acknowledged if we are to succeed. Here is a listing of some of the technologies and tools we will need to develop in order to create a human base on Mars:

- Surface nuclear power plant (nominally 150 kW electrical, plus thermal energy)
- Cryogenic storage and handling tools/systems

- Thermal control systems (including insulation) – different on Mars than in space
- Methane (and/or propane?)-oxygen power sources (electrical, thermal, motive; very small to very large)
- Vehicles (manned and unmanned, ground and air, pressurized and unpressurized, all sizes)
- Construction technologies and equipment (including robots, autonomous or supervised)
- Communications and navigation systems (intra-base, off-base, and off-planet; supporting systems, vehicles, and people)
- Ultra-reliable computing and other IT support (redundant, radiation-hard; wearable systems, etc)
- Medical strategies/tools: auto-medicine (taking care of yourself), para-medicine (taking care of each other), and tele-medicine (accessing medical resources back on Earth)

Not only is this not "rocket science", it's not even just technology. We need to think about construction, physiology, and robotics; and psychology and sociology; and nutrition, gardening, and medicine; and architecture, history, insulation, and HVAC; and power distribution, IT, sensors, and AI; and biology, chemistry, geology, and seismology; and…

In fact, the successful development of an effective base on Mars will require more than a solid systems-centric engineering perspective; it will also require a human-centric perspective, involving numerous social as well as technical disciplines. In essence, we are attempting to design the smallest-scale possible viable human economy and supporting ecology, and we don't know in advance what this "nano-society" will look like. But NASA as an organization is focused on the "rocket science." To understate the case considerably, "studies of surface activities and related systems have not always been carried out to the same breadth or depth as those focused on the space transportation and entry or ascent systems needed for a Mars mission" (NASA, 2001, p. 1). Perhaps the National Science Foundation (NSF), with its broad scientific purview and experience managing U.S. Antarctic bases, might effectively participate in the development of the Mars base.

3.3 Base development process and technology context Developing all the pieces for an effective base on Mars will be a complex undertaking, one quite distinct from the development of the system that will be required to transport humans to and from Mars. What is required is the development of an overall plan, starting from the physiological and psychological needs of a human crew, defining their task-oriented and other activities, leading to system and subsystem models, assessing and adopting/adapting technologies to implement them, and eventually validating the various subsystems through extensive testing and simulations here on Earth (Gage, 2006). This should be a "spiral" process that will be iterated until it is time to go, with multiple agencies/entities involved (e.g., development of a surface-sited nuclear power plant by DOE, healthcare systems by NIH). The initiation of this activity need not and should not wait for a specific commitment to build the Mars transportation system. Perhaps the most difficult challenge will be to manage a complicated program with a relatively small budget (as compared to rocket development), across multiple agencies, over a period of many years.

Many technologies and systems developed for Earth will be carried unchanged to Mars; others will have to be adapted to the particular situation of our Mars base. Rapidly changing technology complicates the development process: at what point do we decide to adopt or adapt a given product or system for inclusion in our long-term Earth-based Mars base prototyping/simulation enterprise? We can freely experiment with COTS elements, but the decision to embark on a costly program to modify existing products for use on Mars must not be taken too early, or we will, like the U.S. military with its communications systems, be trapped in an expensive web of obsolete proprietary systems even as the rest of the world adopts technologies with much higher performance and much lower cost.

The rapid evolution of technology also carries short-term challenges with respect to what we actually send to Mars: given the 26 month synodic period between launch windows, an assembly-test-launch (ATL) time that is not much shorter, and the 12-18 month COTS electronics product innovation cycle, we will have to decide whether to introduce a new generation of IT for each successive mission, and it will clearly be impossible to perform a full-mission duration test of new subsystems as they are deployed. Fortunately, the loose coupling of subsystems in the Mars base environment will allow easy module upgrades and the use of redundant units to ensure system-level reliability.

4. CONCLUSIONS

Because of the limitations placed by orbital mechanics on energy-affordable transits between Earth and Mars (transits that last 6+ months, and are possible only every 26 months), it would be suboptimal to execute the initial human exploration of Mars as a sequence of independent sorties, analogous to the Apollo program. Costs and risks can be significantly reduced by pursuing a program in which the first humans we send to Mars remain there for many (nominally eight) years, living and working in a safe and productive underground base, constructed in advance of their arrival by robots (Gage, 2010). Twenty-six months after the first crew's arrival, a second crew will land at the same base, and other sites of interest can be visited using ground vehicles. Human presence on Mars will be continuous for the (nominally ten year) extent of the program, but fewer crews will be exposed to the risks of spaceflight than in a sequence of independent conjunction missions.

The enthusiasm of wannabe Martians (Gill, 2005) notwithstanding, and independent of the cleverness of NASA's rocket scientists in developing the vehicles and systems needed to transport humans to Mars, many years of preparation involving many other technologies and disciplines will be required before humans will actually set foot on the Martian surface. A major challenge will be the management of this extended process of preparation, with a relatively small budget (as compared to rocket development), across multiple agencies, and over a period of many years.

References

Gage, D.W. (2006). Begin High Fidelity Mars Simulations Now. Ninth International Mars Society Convention, Washington, DC, 3-6 August.

Gage, D.W. (2009). Prepare Now for the Long Stay on Mars. Twelfth International Mars Society Convention, College Park, MD, 30 July - 2 August.

Gage, D.W. (2010). Unmanned systems to support the human exploration of Mars. Proc. SPIE Vol 7692, 7692M.

Gill, S. (2005). The Mars Underground. DVD available online at http://store.oculefilms.com/buy-dvds, Ocule Films, Venice, CA.

Johns, W. (2004). Psychological Factors for Mars Settlement. Online at http://www.marshome.org/files2/Psychological%20Factors%20for%20Mars%20Settlement.ppt.

Krauss, L. (2009). A One-way Ticket to Mars. New York Times Op-Ed. August 31. Online at http://www.nytimes.com/2009/09/01/opinion/01krauss.html.

Kanas, N., Ritsher, J. (2005). Psychosocial issues during a Mars mission. AIAA 1st Space Exploration Conference, Orlando, FL, January 30-February 1.

Kanas, N., Manzey, D. (2003). Space Psychology and Psychiatry. Microcosm Press, El Segundo CA.

Keyak, J.H., Koyama, A.K., LeBlanc, A., Lu, Y., Lang, T.F. (2009). Reduction in proximal femoral strength due to long-duration spaceflight. Bone, vol 44, issue 3, pp 449-453.

Landis, G.A. (2000). Mars Crossing. Tom Doherty Associates, NewYork.

Mars Foundation. (2010). The Mars Homestead Project. http://www.marshome.org/

NASA. (1997). Human Exploration of Mars: The Reference Mission of the NASA Mars Exploration Study Team, NASA SP-6107, NASA Johnson Space Center, Houston TX.

NASA. (1998). Reference Mission 3.0; Addendum to the Human Exploration of Mars: The Reference Mission of the NASA Mars Exploration Study Team, NASA SP-6107-ADD, also EX13-98-036, NASA Johnson Space Center, Houston TX.

NASA. (2001). The Mars Surface Reference Mission: A Description of Human and Robotic Surface Activities, NASA TP-2001-209371, NASA Johnson Space Center, Houston TX.

NASA. (2002). Analysis of Synthesis Group Architectures. ExPO Document XE-92-004.

NASA. (2009). Human Exploration of Mars Design Reference Architecture 5.0. Online at http://www.nasa.gov/pdf/373665main_NASA-SP-2009-566.pdf.

Rapp, D., Andringa, J. (2005). Design Reference Missions for Human Exploration of Mars. JPL Report D- 31340, also presented at ISDC, Arlington, VA, May, 2005.

Rapp, D. (2006). Radiation effects and shielding requirements in human missions to the moon and Mars, Mars 2, 46-71 (2006), online at http://marsjournal.org/contents/2006/0004/files/rapp_mars_2006_0004.pdf.

Synthesis Group. (1991). America at the Threshold: America's Space Exploration Initiative. Online at http://history.nasa.gov/staffordrep/main_toc.pdf.

Varley, J. (2005). In the Hall of the Mountain Kings, in the anthology Fourth Planet from the Sun. Thunder's Mouth Press, New York.

Zubrin, R. (1996). The Case for Mars. Simon & Schuster, New York.

Zubrin, R. (2003). Mars on Earth. Tarcher/Penguin, New York.