

Supernova Remnant OH Masers: Signposts of Cosmic Collision

Mark Wardle^{1*} and Farhad Yusef-Zadeh²

A supernova explosion, the final death throes of a massive star, creates an expanding bubble of hot gas that overruns up the surrounding medium. When a supernova remnant encounters a dense interstellar cloud, the compression may trigger gravitational collapse and the formation of a new generation of stars. This event can be detected through intense stimulated emission in the 1720-megahertz transition of the hydroxyl radical, OH, which yields unique insights into the physical processes and conditions occurring during the interaction.

Shock waves from supernova explosions have played a key role in the evolution of the Milky Way Galaxy by depositing energy and material enriched in heavy elements into interstellar space (1–4). The passage of a shock wave past an interstellar molecular cloud contributes to the cloud's eventual dispersal, but the compression of the cloud's gas during the interaction may lead to star formation. Interactions between supernovae and clouds are common: About half of all supernova explosions occur when a massive star ($>8M_{\odot}$, where M_{\odot} is the mass of the Sun) exhausts its supply of nuclear fuel (5); and during its short life, the star does not drift very far from its parent cloud.

The evolution and appearance of a supernova remnant (SNR) is also affected by an adjacent cloud in several ways. The expansion is slowed at the cloud's location and the remnant is distorted. A reflected shock is driven into the remnant's interior as it runs into the obstacle, and the incorporation and heating of cloud material may fill the interior with gas, substantially modifying its x-ray emission (6–9). Enhanced particle acceleration in the shock front produces synchrotron radiation at radio frequencies, and, if the particle energy is high enough (as in the SNR RCW 86), at x-ray frequencies (10).

Interaction between a SNR and a molecular cloud is signaled by bright, compact spots—masers (see below)—emitted by OH molecules at a frequency of 1720 MHz (11). About 10% of SNRs have associated masers and are probably interacting with molecular clouds (12–15). Here we review what we have learned about the interaction from these

masers, and we describe how they have provided a crucial window into the physical processes induced within the cloud by the interaction.

Supernova Remnant Masers

The term “maser” is derived from microwave amplification of stimulated emission radiation, the microwave-frequency analog of a laser (16, 17). Amplification occurs in a par-

state, a condition referred to as population inversion. When this occurs, seed photons originating from spontaneous emission or background radio sources will be exponentially amplified by stimulated emission, leading to the production of an intense, tightly collimated beam of photons at the frequency of the transition, which is only limited by the rate at which the inversion can be maintained by the processes responsible for the supply of molecules in the upper state.

Population inversion of molecules is occasionally achieved under natural conditions in interstellar space, creating bright, compact spectral line sources that are easily observed at radio wavelengths. The first interstellar molecule detected at radio frequencies (18), OH, was also the first to be detected as a maser (19). Other important species that have masing transitions are H_2O , SiO, and CH_3OH (20). All

occur in the interstellar medium, and all but CH_3OH occur in the envelopes of evolved stars. The brightness and compactness of maser spots can be exploited to yield kinematic information about interstellar gas via the Doppler effect and observations of proper motion. Perhaps the most spectacular example has been the determination of supermassive black hole masses using the motions of H_2O megamasers around active galactic nuclei (21). In addition, masers have been used to track outflows and the rotation of protostellar discs in star-forming regions, and to determine stellar kinematics and wind velocities from OH and SiO masers arising in dense stellar winds from evolved stars (20). The intensity of maser emission allows the measurement of weak circular polarization across the maser line induced by the Zeeman effect, which permits a determination of the line-of-sight magnetic field. If both linear and circular polarization can be detected, then the strength and orientation of the field can, in principle, be determined. The field strengths derived from maser measurements have been controversial because of the possibility that the masers are saturated, in which case the inferred field strengths may be overestimated (22, 23), but this does not appear to be a serious problem for the 1720-MHz OH masers associated with SNRs (24).

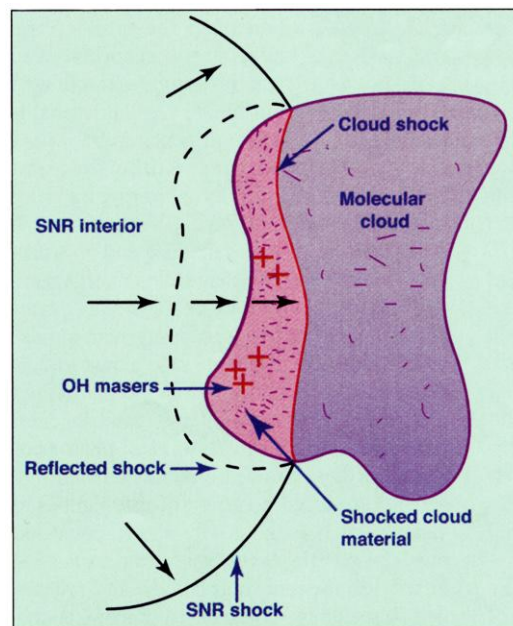


Fig. 1. Schematic of an expanding supernova remnant (SNR) interacting with an adjacent molecular cloud. Black arrows indicate velocity.

ticular radiative transition of a molecule (or, more generally, an ion or atom) when absorption of photons from the radiation field by molecules in the lower level is dominated by stimulated emission of photons by molecules in the upper state. This requires that the upper state be overpopulated relative to the lower

¹Research Centre for Theoretical Astrophysics, School of Physics A28, University of Sydney, NSW 2006, Australia. ²Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208, USA.

*To whom correspondence should be addressed. E-mail: wardle@physics.usyd.edu.au

OH is one of the most important maser species because it is abundant and has transitions near 1.6 GHz that are easily observed using the techniques of radio astronomy. The rotation ladder is split several times because of a single unpaired electron sitting in a $2p$ orbital in the O atom. The spin-orbit coupling of this electron first splits the rotational ladder into $^2\Pi_{1/2}$ and $^2\Pi_{3/2}$ ladders. Each rotational level within these ladders is further split by the small difference in the moment of inertia when the electron orbital is parallel or perpendicular to the molecule's rotation axis (lambda doubling), and again by the hyperfine interaction between the electron spin and the spin of the proton in the H atom. The energy splittings are such that the rotational transitions occur in the far- and mid-infrared, whereas the transitions within a rotational level are at microwave frequencies. These transitions are potentially the easiest to invert and may occur naturally. Indeed, astronomical masers are observed in all four of the transitions within the ground rotational state (at 1612, 1665, 1667, and 1720 MHz) and in many of the transitions within higher rotational levels as well.

Most observations have focused on the ground-state masers. Although any of these may occur alone or closely associated with any of the other three, masers at 1665 and 1667 MHz are generally associated with each other and with star-forming regions, whereas 1612-MHz masers are associated with evolved stars (16, 20). This is understood in terms of the different physical conditions that lead to inversion of each of the transitions. In particular, the 1612-, 1665-, and 1667-MHz masers are usually pumped by a far-infrared (FIR) radiation field.

Until recently, a class of relatively rare, weak masers in the 1720-MHz OH line associated with SNRs was neglected as a curiosity. By the late 1960s, it was known that some SNRs—notably W28, W44, and IC443—had these masers (25). In the following decade, it was pointed out that 1720-MHz masers could occur in warm molecular gas without the need for pumping by a FIR radiation field (26), suggestive of shocked gas, and that they might be associated with SNR-driven shock waves (27).

However, when interferometric observations of W28 with the Very Large Array

(VLA) showed many discrete maser spots around the remnant where it was believed to be interacting with a molecular cloud, it was realized that the masers were associated with shock waves driven into the cloud by the interaction (11). Follow-up studies showed that the masers were indeed associated with warm, shocked molecular gas, suggesting that the masers can be used as “signposts” of SNR–molecular cloud interactions (28, 29). Surveys of galactic SNRs for 1720-MHz OH masers (12–14, 30) have shown that 10% have maser emission. Apart from identifying these new cases of SNR–cloud interactions, 1720-MHz OH masers have proved useful in untangling the crowded inner 100 pc of the Galaxy and clarifying the relationship among several interacting components (31, 32).

The maser kinematics allow the distance to a SNR to be estimated from models of the differential rotation of the galactic disc (11). The key notion is that the radial velocities of the masers are characteristic of the galactic

neighborhood of the SNR. The difference in line-of-sight velocities between maser spots in a particular remnant is generally small even when they are distributed around a substantial fraction of the shell, which implies that the maser velocities are close to the line-of-sight velocity component of an unshocked cloud complex that envelops a large portion of the remnant. The shock waves at those locations are presumed to be traveling perpendicular to the line of sight, as an edge-on shock geometry favors the generation of masers by increasing the column density and reducing the line-of-sight velocity gradient.

Zeeman observations of 1720-MHz maser spots have measured line-of-sight components of the magnetic field (in the milligauss range) in the shocked gas adjacent to several remnants (33, 34). These are the highest signal-to-noise measurements to date of magnetic fields at gas densities typical of molecular clouds, 10^4 to 10^5 cm $^{-3}$, which otherwise have been obtained from Zeeman OH absorption-line studies (35). This information is

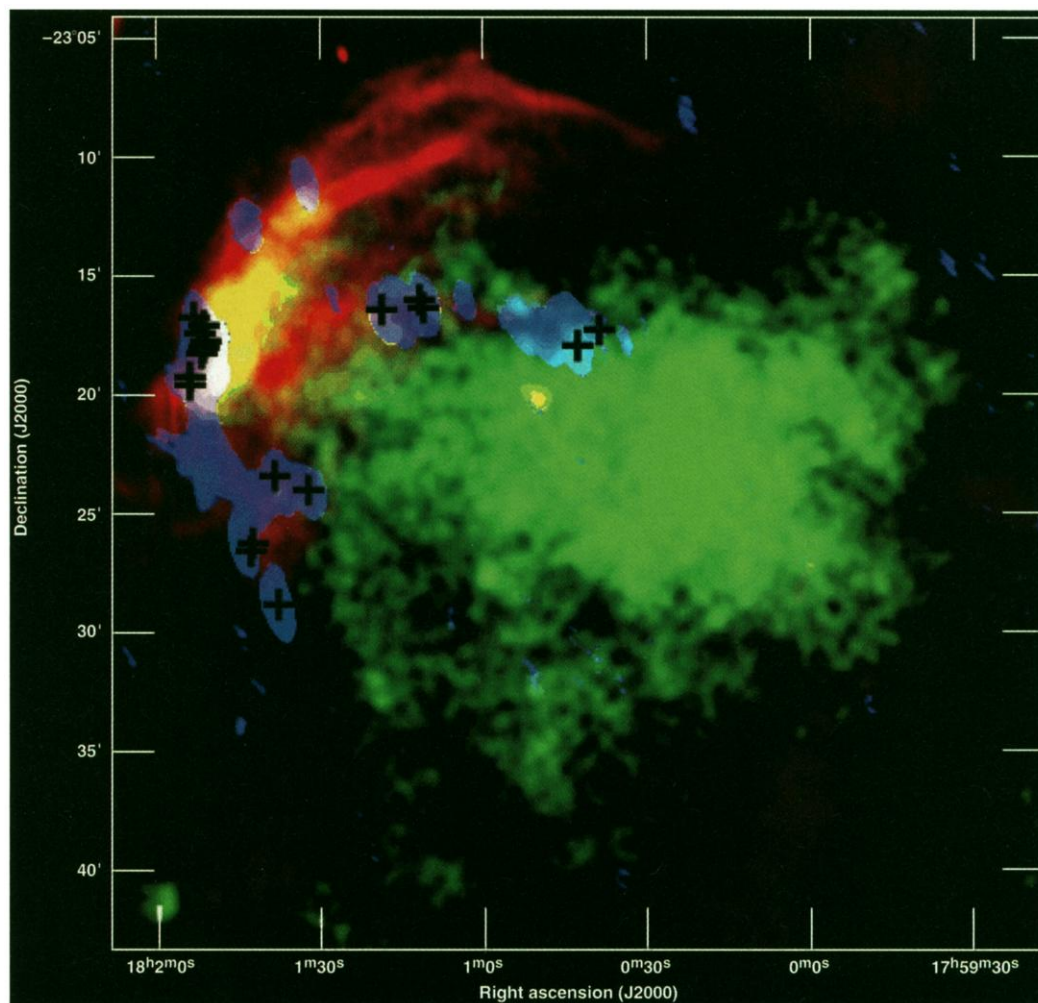


Fig. 2. Extended maser emission (blue) superposed on a 20-cm continuum image of SNR W28 (red) with a resolution of $28'' \times 16''$ (62). The center-filled morphology of the remnant in Röntgen Satellite (ROSAT) x-ray observations (green) contrasts with the shell morphology at radio wavelengths. Crosses correspond to the positions of detected compact maser sources (33).

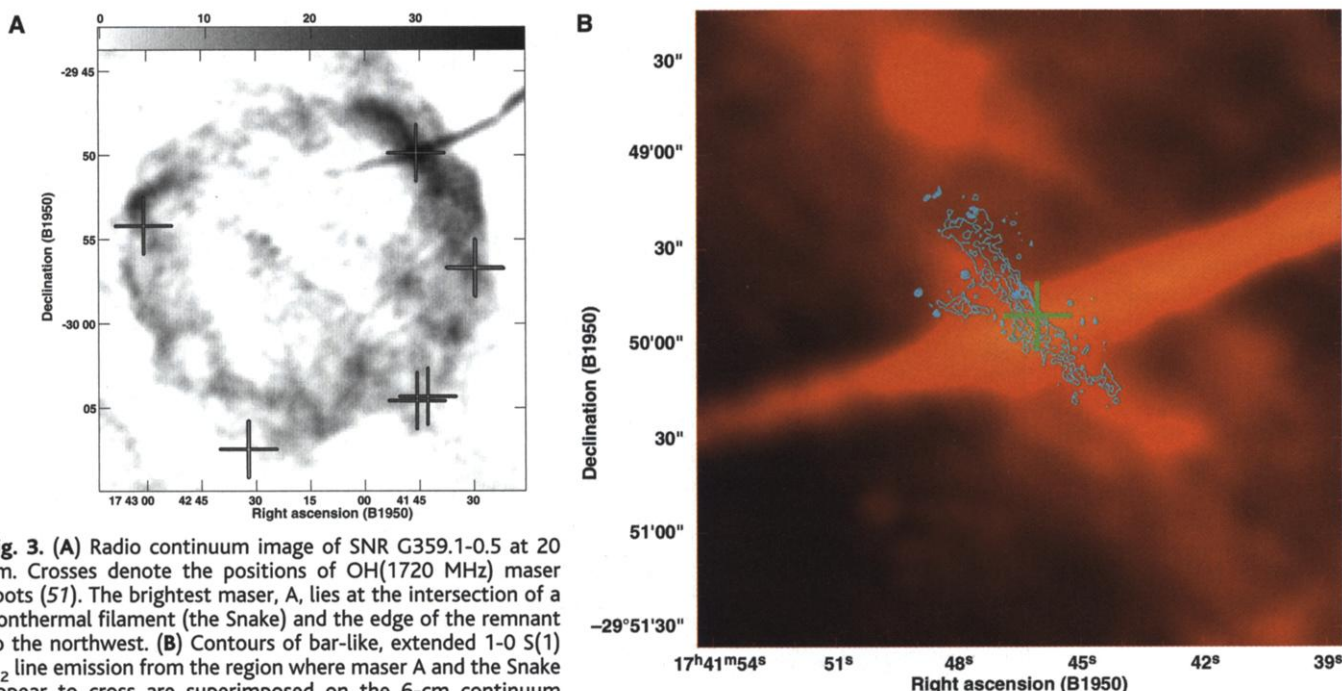


Fig. 3. (A) Radio continuum image of SNR G359.1-0.5 at 20 cm. Crosses denote the positions of OH(1720 MHz) maser spots (51). The brightest maser, A, lies at the intersection of a nonthermal filament (the Snake) and the edge of the remnant to the northwest. (B) Contours of bar-like, extended 1-0 S(1) H₂ line emission from the region where maser A and the Snake appear to cross are superimposed on the 6-cm continuum image.

fundamental to understanding the internal dynamics and evolution of molecular clouds because the magnetic pressure dominates thermal pressure by two orders of magnitude. Further, the magnetic field pressure in the shocked gas is approximately equal to the pressure in the SNR's interior, a key quantity in models of SNR evolution.

Thus, 1720-MHz OH masers signal interaction sites between SNRs and molecular clouds, probe kinematics and magnetic field strengths, and determine two key parameters—the remnant's distance and internal pressure. Much more can be inferred from the masers if the conditions required to invert the level population and produce strong maser action are understood.

Theory

The population of internal states is determined by absorption, stimulated emission, spontaneous emission of a photon, and excitation and deexcitation between all of the states available to the molecule by collisions with other species. In particular, the inversion is sensitive to the gas temperature and density (which determine collisional excitation and deexcitation) and the radiation field (absorption to higher lying levels that may eventually cascade down to the upper level of the transition under consideration). Maser amplification also requires a sufficient column of molecules with a small velocity gradient; otherwise, the Doppler shifting of the transition frequency prevents amplification from occurring along the entire column.

Population inversion and amplification in the 1720-MHz line alone requires densities of

$\sim 10^5 \text{ cm}^{-3}$, temperatures in the range 30 to 125 K, an OH column density in the range 10^{16} to 10^{17} cm^{-2} , and an absence of a strong FIR continuum (26, 36). The latter constraint explains why 1720-MHz masers are rare in star-forming regions, where the ultraviolet (UV) radiation from hot, massive young stars is absorbed by dust grains and reradiated in the FIR. Molecular gas cools very efficiently, so that the required column of warm OH implies that the thickness of the shock wave is consistent with a nondissociative C-type shock wave rather than a thin, dissociative J-type transition (36). The former type of shock is expected to be prevalent in SNR–molecular cloud interactions.

Although this finding confirms that shocked molecular gas supplies the physical conditions necessary to create an inversion of the 1720-MHz transition, it does not easily explain how the OH itself is produced. Quiescent molecular gas consists of H₂, He, CO, and either atomic or diatomic oxygen (at a level of a few parts in 10^4), with a very low OH abundance. OH is produced by the endothermic reaction $\text{O} + \text{H}_2 \rightarrow \text{OH} + \text{H}$ at temperatures above 400 K within a C-type shock wave; however, this OH is consumed by the faster reaction $\text{OH} + \text{H}_2 \rightarrow \text{H}_2\text{O} + \text{H}$, which is less endothermic. Shock waves convert all of the preshock O and O₂ into water rather than OH (37, 38).

One way of creating OH is by dissociation of this water by UV radiation from young, massive stars. This process is responsible for the OH in star-forming regions (39, 40), but in the SNR context it is difficult to supply the required flux to the postshock gas without heat-

ing the intervening dust grains and producing a strong FIR continuum, which suppresses inversion of the 1720-MHz line. In addition, there are generally no obvious strong UV sources associated with the SNR masers.

An alternative dissociation mechanism is provided by the strong x-ray flux due to bremsstrahlung from the hot gas filling the interior of the adjacent SNR (41, 42). The x-rays photoeject electrons from molecules and dust grains throughout the molecular cloud. The runaway electrons thermalize as a result of collisions in the gas, and during this process molecular hydrogen can be impact-excited to a higher electronic state in the Lyman and Werner bands. The subsequent radiative decay produces a UV photon capable of dissociating water. The ionizations associated with cosmic-ray interactions with the molecular gas produce a similar effect, which has previously been incorporated in models of cloud chemistry (43) and shock waves (44). However, the ionization rate per H nucleus in these contexts is $\sim 10^{-17} \text{ s}^{-1}$, far less than the $\sim 10^{-15} \text{ s}^{-1}$ expected for the molecular gas adjacent to SNRs. When this is incorporated into C-type shock models, the far-UV flux that is produced locally by impact excitation is increased by a factor of 30 to 100, so that dissociation of H₂O and OH is more rapid and a sufficient column of OH is produced in the cooling gas behind the shock front (42).

The resulting picture, then, is as follows (Fig. 1). A SNR drives a shock wave at $\sim 25 \text{ km s}^{-1}$ into the molecular cloud. The shock converts atomic and diatomic oxygen into water, which is subsequently dissociated to OH (and then to O) by the secondary far-UV flux induced throughout the cloud by the x-ray flux

from the SNR interior. Behind the shock wave, conditions are right to invert the 1720-MHz OH line, and there is a sufficient OH column density to give rise to the OH(1720 MHz) masers.

Confrontation with Observations

This model explains the empirical association between the masers and SNR–molecular cloud interactions. But is it correct? There are a couple of good tests. First, OH(1720 MHz) masers should be associated with C-type shock waves, which are strong sources of emission in the 1-0 S(1) line of H_2 at 2.12 μm (37, 45, 46). Second, the OH column produced by the dissociation process should be observable in absorption at radio wavelengths against the bright radio continuum of the SNR if it lies on the near side of the remnant (42).

These “predictions” are generally consistent with what was already known for the existing well-studied cases. For example, the SNR IC443 shows both H_2 emission (47, 48) and OH absorption (27), and OH absorption was known in SNR W28 (49) (Fig. 2). Subsequent interferometric observations of W28 (50) have shown that the OH absorption nicely overlies the shocked component of CO (29) lying adjacent to the unshocked component.

More compellingly, the model is supported by observations toward less-studied remnants with OH(1720 MHz) masers. G359.1-0.5 is a classic example of an OH(1720 MHz) SNR with several masers lying around the edge of its radio shell (Fig. 3) (51). This remnant is distinguished by an enigmatic nonthermal filament, the “Snake,” that crosses the brightest section of the shell in the northwest (52). A search for emission from shock-heated H_2 was conducted toward the OH(1720 MHz) maser positions to test the C-type shock model of the emission (53). The bright compact maser emission is located near the center of an elongated bar of H_2 emission that is aligned with the edge of the remnant. The intensity of the 2-1 S(1) and 1-0 S(1) H_2 line emission is consistent with shock-heated gas rather than fluorescence, and molecular-line observations have identified the preshock and shocked gas on either side of the bar (54). G349.7+0.2, a distant (20 kpc) SNR lying on the far side of the Galactic Center, confirms the theory with detections of CO, shocked H_2 , and OH absorption overlying one another and the group of previously detected masers sitting in the remnant (53) (Fig. 4). The elongated radio continuum source G357.7-0.1 (the Tornado Nebula) has a spectral index consistent with a shell-type SNR (55, 56). Frail *et al.* (12) detected an OH(1720 MHz) maser at

-12.4 km s^{-1} on the western edge of the nebula, and H_2 observations have detected emission from shocked molecular hydrogen with a striking filamentary structure similar to the nonthermal radio morphology (57).

Implications and Future Directions

The model for the production of OH masers associated with SNRs has several implications. First, the shock wave must be “C-type”—the dissipation within the front is inefficient, enabling cooling by molecules to keep the temperature below 3000 K throughout the shock front. These broad, cool shock waves were predicted to exist three decades ago (58) and are thought to produce intense H_2 emission from shock-heated gas in interstellar clouds (45, 46). Problems in detailed modeling of observed line ratios have led to doubts about this interpretation (48, 59), so the success of the maser model is an important validation of the theory of shock waves in molecular clouds. Second, the model highlights the unappreciated role that thermal x-rays from the remnant’s interior play in the ionization and chemistry of the adjacent molecular cloud. The ionization level, which is increased by a factor of 5 to 10, determines the coupling between the strong magnetic field and the gas and controls the structure of shock waves (60), the damping of magnetohydrodynamic waves, and the diffusion of the magnetic field during gravitational

collapse to form stars (61). Third, the soft thermal (rather than hard power-law) nature of the x-ray spectrum is a crucial factor because the interaction cross section for x-rays with matter decreases at high incident photon energies. SNRs with soft thermal emission are generally shell-like at radio wavelengths and show a filled, centrally peaked x-ray morphology with a thermal spectrum and a temperature below about 1 keV—the “mixed morphology” remnants (6). The emission is thought to arise via the incorporation of dense material overrun by the expanding remnant into its interior (9). Thus, one expects that all of the maser remnants should display mixed morphology. Fourth, OH absorption can be used to study the distribution and kinematics of shocked gas (subject to the limitations imposed by the requirement that the OH-rich gas overlie a sufficiently strong radio continuum). Finally, weak inversion of the 1720-MHz line occurs over a broad range of physical parameters and produces extended bright emission in the 1720-MHz line surrounding the compact maser spots in several remnants: G359.1-0.5, G357.7+0.3 (the Square Nebula), and G357.7-0.1 (the Tornado Nebula) (30, 51). The emission surrounding the maser spots is important because it traces the molecular gas where the abundance of the OH molecule has been enhanced by a shock wave (27, 42) and is bright enough to permit

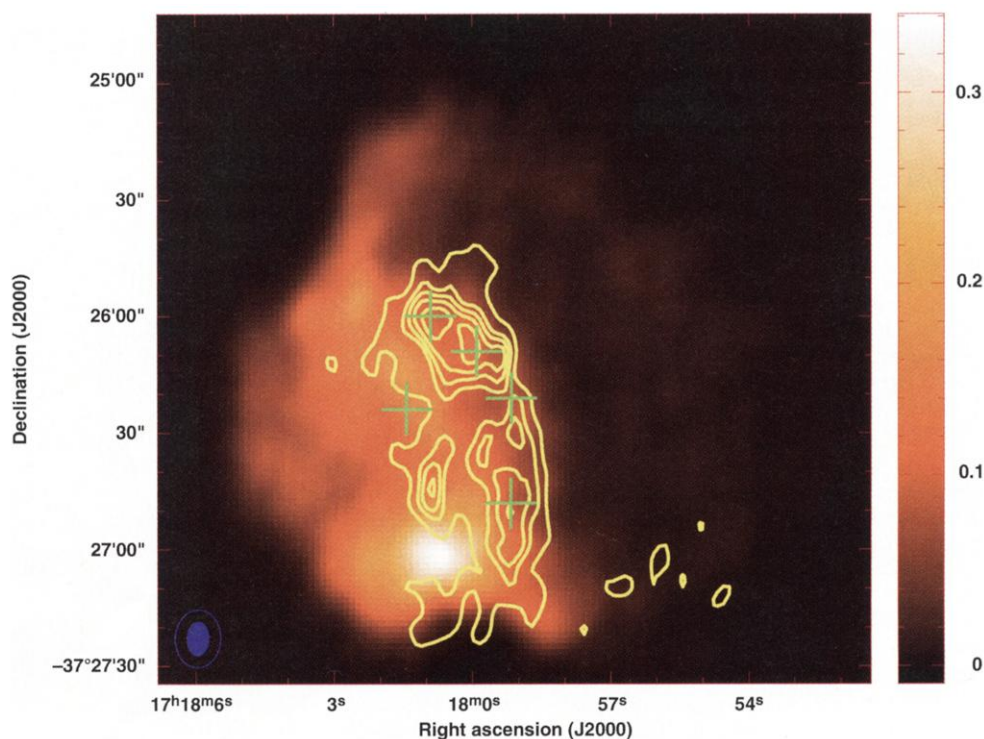


Fig. 4. OH absorption (contours) against the 18-cm radio continuum emission (color) from SNR G349.7+0.2. Crosses indicate OH(1720 MHz) maser positions (12). The contour levels are -0.11 , -0.14 , -0.16 , -0.19 , -0.22 , -0.23 , -0.24 , and $-0.27 \text{ Jy per } 14.9'' \times 11.8'' \text{ beam}$.

Zeeman mapping of the magnetic field over extended regions (30).

References and Notes

1. A. G. W. Cameron, *Publ. Astron. Soc. Pac.* **69**, 201 (1957).
2. C. F. McKee, J. P. Ostriker, *Astrophys. J.* **218**, 148 (1977).
3. C. F. McKee, L. L. Cowie, J. P. Ostriker, *Astrophys. J.* **219**, L23 (1978).
4. T. W. Hartquist, *Astrophys. Space Sci.* **216**, 185 (1994).
5. K. W. Weiler, R. A. Sramek, *Annu. Rev. Astron. Astrophys.* **26**, 295 (1988).
6. J. Rho, R. Petre, *Astrophys. J.* **503**, L167 (1998).
7. R. A. Chevalier, *Astrophys. J.* **511**, 798 (1999).
8. D. P. Cox et al., *Astrophys. J.* **524**, 179 (1999).
9. R. L. Shelton et al., *Astrophys. J.* **524**, 192 (1999).
10. K. J. Borkowski, J. Rho, S. P. Reynolds, K. K. Dyer, *Astrophys. J.* **550**, 334 (2001).
11. D. A. Frail, W. M. Goss, V. I. Slysh, *Astrophys. J.* **424**, L111 (1994).
12. D. A. Frail et al., *Astron. J.* **111**, 1651 (1996).
13. A. J. Green, D. A. Frail, W. M. Goss, R. Otrupcek, *Astron. J.* **114**, 2058 (1997).
14. B. Koralesky, D. A. Frail, W. M. Goss, M. J. Claussen, A. J. Green, *Astron. J.* **116**, 1323 (1998).
15. F. Yusef-Zadeh, D. A. Roberts, W. M. Goss, D. A. Frail, A. J. Green, *Astrophys. J.* **512**, 230 (1999).
16. M. Elitzur, *Annu. Rev. Astron. Astrophys.* **30**, 75 (1992).
17. M. Gray, *Philos. Trans. R. Soc. London Ser. A* **357**, 3277 (1999).
18. S. Weinreb, A. H. Barrett, M. L. Meeks, J. C. Henry, *Nature* **208**, 29 (1963).
19. H. Weaver, D. Williams, N. Dieter, W. Lum, *Nature* **208**, 29 (1965).
20. M. J. Reid, J. M. Moran, *Annu. Rev. Astron. Astrophys.* **19**, 231 (1981).
21. M. Miyoshi et al., *Nature* **373**, 127 (1995).
22. M. Elitzur, *Astrophys. J.* **457**, 415 (1996).
23. W. D. Watson, H. W. Wyld, *Astrophys. J.* **558**, L55 (2001).
24. M. Elitzur, *Astrophys. J.* **504**, 390 (1998).
25. W. M. Goss, B. J. Robinson, *Astrophys. Lett.* **2**, 81 (1968).
26. M. Elitzur, *Astrophys. J.* **203**, 124 (1976).
27. L. K. DeNoyer, *Astrophys. J.* **228**, L41 (1979).
28. D. A. Frail, G. F. Mitchell, *Astrophys. J.* **508**, 690 (1998).
29. Y. Arikawa, K. Tatematsu, Y. Sekimoto, T. Takahashi, *Publ. Astron. Soc. Jpn.* **51**, L7 (1999).
30. F. Yusef-Zadeh, W. M. Goss, D. A. Roberts, B. Robinson, D. A. Frail, *Astrophys. J.* **527**, 172 (1999).
31. F. Yusef-Zadeh, F. Melia, M. Wardle, *Science* **287**, 85 (2000).
32. F. Yusef-Zadeh, S. R. Stolovy, M. Burton, M. Wardle, M. C. B. Ashley, *Astrophys. J.* **560**, 749 (2001).
33. M. J. Claussen, D. A. Frail, W. M. Goss, R. A. Gaume, *Astrophys. J.* **489**, 143 (1997).
34. C. L. Brogan, D. A. Frail, W. M. Goss, T. H. Troland, *Astrophys. J.* **537**, 875 (2000).
35. T. L. Bourke, P. C. Myers, G. Robinson, A. R. Hyland, *Astrophys. J.* **554**, 916 (2001).
36. P. Lockett, E. Gauthier, M. Elitzur, *Astrophys. J.* **511**, 235 (1999).
37. B. T. Draine, W. G. Roberge, A. Dalgarno, *Astrophys. J.* **264**, 485 (1983).
38. M. J. Kaufman, D. A. Neufeld, *Astrophys. J.* **456**, 611 (1996).
39. T. W. Hartquist, A. Sternberg, *Mon. Not. R. Astron. Soc.* **248**, 48 (1991).
40. T. W. Hartquist, K. M. Menten, S. Lepp, A. Dalgarno, *Mon. Not. R. Astron. Soc.* **272**, 184 (1995).
41. M. Wardle, F. Yusef-Zadeh, T. R. Geballe, in *The Central Paradoxes of the Galaxy*, H. Falcke, A. Cotera, W. Duschl, F. Melia, M. Rieke, Eds. (Astronomical Society of the Pacific, San Francisco, 1999), p. 432.
42. M. Wardle, *Astrophys. J.* **525**, L101 (1999).
43. S. S. Prasad, S. P. Tarafdar, *Astrophys. J.* **267**, 603 (1983).
44. G. Pineau des Forêts, E. Roueff, D. R. Flower, *Mon. Not. R. Astron. Soc.* **223**, 743 (1986).
45. B. T. Draine, W. G. Roberge, *Astrophys. J.* **259**, L91 (1982).
46. D. F. Chernoff, C. F. McKee, D. J. Hollenbach, *Astrophys. J.* **259**, L97 (1982).
47. M. G. Burton, D. J. Hollenbach, M. R. Haas, E. F. Erickson, *Astrophys. J.* **355**, 158 (1990).
48. M. J. Richter, J. R. Graham, G. S. Wright, D. M. Kelly, J. H. Lacy, *Astrophys. J.* **449**, L83 (1995).
49. M. I. Pastchenko, V. I. Slysh, *Astron. Astrophys.* **35**, 153 (1974).
50. A. J. Green, M. Wardle, J. S. Lazendic, paper presented at the 24th meeting of the IAU, Joint Discussion 1, Atomic and Molecular Data for Astrophysics: New Developments, Case Studies and Future Needs, Manchester, England, August 2000.
51. F. Yusef-Zadeh, K. I. Uchida, D. Roberts, *Science* **270**, 1801 (1995).
52. A. D. Gray, J. Nicholls, R. D. Ekers, L. E. Cram, *Astrophys. J.* **448**, 164 (1995).
53. J. S. Lazendic et al., *Mon. Not. R. Astron. Soc.* **331**, 537 (2002).
54. J. S. Lazendic et al., in preparation.
55. P. A. Shaver et al., *Astron. Astrophys.* **147**, L23 (1985).
56. R. T. Stewart, R. F. Haynes, A. D. Gray, W. Reich, *Astrophys. J.* **432**, L39 (1994).
57. J. S. Lazendic et al., in preparation.
58. D. J. Mullan, *Mon. Not. R. Astron. Soc.* **153**, 145 (1971).
59. P. W. J. L. Brand et al., *Astrophys. J.* **334**, L103 (1988).
60. B. T. Draine, C. F. McKee, *Annu. Rev. Astron. Astrophys.* **31**, 373 (1993).
61. E. G. Zweibel, *Phys. Plasmas* **6**, 1725 (1999).
62. G. M. Dubner, P. F. Velásquez, W. M. Goss, M. A. Holdaway, *Astron. J.* **120**, 1933 (2000).
63. We thank G. Dubner, J. Hewitt, and J. Rho for assistance with Fig. 2 and J. Lazendic for assistance with Figs. 3 and 4. F. Y.-Z. acknowledges support by NASA grant NAG-5-9188.

REVIEW: AIDS

Diversity Considerations in HIV-1 Vaccine Selection

Brian Gaschen,¹ Jesse Taylor,¹ Karina Yusim,¹ Brian Foley,¹ Feng Gao,² Dorothy Lang,¹ Vladimir Novitsky,³ Barton Haynes,² Beatrice H. Hahn,⁴ Tanmoy Bhattacharya,¹ Bette Korber^{1,5*}

Globally, human immunodeficiency virus–type 1 (HIV-1) is extraordinarily variable, and this diversity poses a major obstacle to AIDS vaccine development. Currently, candidate vaccines are derived from isolates, with the hope that they will be sufficiently cross-reactive to protect against circulating viruses. This may be overly optimistic, however, given that HIV-1 envelope proteins can differ in more than 30% of their amino acids. To contend with the diversity, country-specific vaccines are being considered, but evolutionary relationships may be more useful than regional considerations. Consensus or ancestor sequences could be used in vaccine design to minimize the genetic differences between vaccine strains and contemporary isolates, effectively reducing the extent of diversity by half.

multiple epidemiologically unlinked individuals. Currently, strains belonging to the same subtype can differ by up to 20% in their envelope proteins, and between-subtype distances can soar to 35%. Moreover this diversity is continually growing. The need for frequent changes in the annual influenza vaccine puts into perspective the implications of such diversity—less than 2% amino acid change can cause a failure in the cross-reactivity of the polyclonal response to the influenza vaccine and necessitates changing the vaccine strain (4).

Since HIV-1 M group began its expansion in humans roughly 70 years ago (1, 2) it has diversified rapidly (3), now comprising a number of different subtypes and circulating recombinant forms (CRFs). The HIV-1 M group is the set of diverse viruses that dominates the global AIDS epidemic. Subtypes are genetically

ly defined lineages that can be resolved through phylogenetic analysis of the HIV-1 M group as well-defined clades, or branches, in a tree. Recombination occurs frequently, and a CRF carries sections of two or more subtypes in a mosaic genome; a recombinant lineage is designated a CRF when related forms are found in

¹Los Alamos National Laboratory, Los Alamos, NM 87545, USA. ²Duke University AIDS Center, Durham, NC 27710, USA. ³Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, MA 02115, USA. ⁴University of Alabama at Birmingham, Birmingham, AL 35294, USA. ⁵Santa Fe Institute, Santa Fe, NM 87501, USA.

*To whom correspondence should be addressed. E-mail: btk@t10.lanl.gov