Proceedings of ICRC 2001: 2070 © Copernicus Gesellschaft 2001

ICRC 2001

Galactic cosmic rays from supernova remnants: myth or reality?

E. Parizot[1], **J. Paul**[2], **and A. Bykov**[3]

[1]Institut de Physique Nucléaire d'Orsay, 91406 Orsay, France
[2]Service d'Astrophysique, CEA-Saclay, 91191 Gif-sur-Yvette, France
[3]A.F. Ioffe Institute for Physics & Technology, 194021, St. Petersburg, Russia

Abstract. Arguments in favour and against a supernova remnant origin of the bulk of Galactic cosmic rays are reviewed and discussed. We analyse the current observational and theoretical status of the model and argue that some long-standing problems do not appear to be solvable in the current state of knowledge.

1 Introduction

Most of the reviews written in the last twenty years on cosmic ray astrophysics begin with a sentence like *supernova remnants (SNRs) are the favourite candidates for the source of Galactic cosmic rays (GCRs)*, or *there is now broad consensus that GCRs are accelerated at SNR shocks*, or *it is generally believed that...* Most of these papers are fair enough to also acknowledge that the question is not settled yet, in spite of decades of efforts on both observational and theoretical sides. Many developments and refinements of the original theory have been considered over the years, such as the self generation of turbulence upstream of the shock or the retroaction of the accelerated particles on the shock structure and dynamics. Many instruments and satellites have also been operated to track the cosmic rays down to their alleged sources. Still in vain, or at least not convincingly.

The theory of diffusive shock acceleration has been studied in great detail, both analytically and numerically, through Monte Carlo simulations as well as using kinetic theory formulations, and the physical mechanisms at work are quite well understood: numerical results are in very good agreement with the analytical calculations and theoretical expectations, and good agreement is also found with direct observational data whenever they are available, at interplanetary shocks in the solar system or at the Earth bow shock. There is no doubt that shocks do accelerate particles, because we see the number of energetic particles increase as a spacecraft crosses a shock front (e.g. Ellison et al., 1990; Baring et al., 1997). As for SNR shocks, we clearly see the sharp radio rings indicating the presence of electrons accelerated in situ. It is therefore quite natural to believe that SNR shocks also accelerate protons and nuclei from the ambient medium, and this is indeed predicted by all the model as a direct physical consequence of the interaction of charged particles with the magnetic fields on either side of the shock front, moving relative to one another.

However, in spite of a considerable observational effort, a clear proof of the presence of energetic nuclei in SNRs at a significant level is still missing, as the γ-rays above ~ 100 MeV due to the decay of π^0 mesons have not been identified. Moreover, the fact that SNR shocks accelerate *some* particles, which we shall call 'SN energetic particles' (or SN-EPs) does not imply that these particles *are* the Galactic cosmic rays. The identification of the GCRs with the SN-EPs is the thesis which we further investigate in this paper.

Historically, it has been based on three main arguments: 1) the energy input rate required to power the GCRs observed at Earth is comparable to a reasonable fraction (10–30%) of the kinetic energy input rate in the interstellar medium (ISM) from SN explosions, 2) radio observations of synchrotron emission from relativistic electrons at SNRs, and 3) simple 'test particle' calculations of diffusive shock acceleration predict a power law spectrum with a universal slope, so that the contributions of many different sources in the ISM can accumulate to produce the observed power law spectrum. The apparent discrepancy between the predicted logarithmic slope, namely 2.0, and that observed at Earth, namely 2.7, could be attributed to the energy dependent confinement time of the energetic particles in the Galaxy. In a simple leaky-box model, a confinement time in $E^{-0.7}$ would indeed turn a *source spectrum* in p^{-2} into a *propagated spectrum* in $p^{-2.7}$ above ~ 1 GeV/n.

In spite of this promising debut, the question of the origin of GCRs is not yet clarified. Contrary to what one would have hoped, the increase amount of work on a SNR origin of GCRs has not resulted in an increased confidence in the model. Two decades of efforts have certainly not been vain, as diffusive acceleration at SNR shocks *is* the best understood model for particle acceleration in the Galaxy. But its very preciseness and trustworthiness makes its predictions more and more inescapable, so that any disagreement with the GCR data pleads all the more strongly against an identification of the SN-EPs with the observed GCRs. At the time being, all the efforts seem to have failed to clearly confirm

Correspondence to: E. Parizot (parizot@ipno.in2p3.fr)

the ’widely accepted’ model for CR origin, and several arguments cast doubt on whether SNRs have something to do with GCRs at all. In this paper, we sum up the observational and theoretical arguments relevant to this question, and discuss the possibility of considering alternative scenarios.

2 Arguments from cosmic-ray propagation

The increase in the amount of cosmic-ray data allows one to severely constrain the propagation models at least at energies lower than a few TeV, more or less independently of the CR source model, i.e. with no *a priori* assumption about the CR source spectrum, composition and spatial distribution in the Galaxy. It is quite remarkable that a detailed, but conceptually simple propagation model can fit the available data with standard physical and astrophysical inputs and sensible values of the parameters. Strong and Moskalenko (1998, 2001) have shown that all the measured secondary-to-primary abundance ratios among the EPs, the radioactive daughter-to-parent nuclei ratios, and the positron-to-electron ratio can be reproduced by assuming a power-law CR source spectrum in $p^{-2.35}$, a confinement time in $p^{0.36}$, reasonably close to the Kolmogorov value, and a simple, minimal reacceleration model. It should be stressed that the number of free parameters here is much less than the number of observables, as each of the above-mentioned ratios does not provide only one constraint, but must be fit *as a function of energy*. Moreover, the observed CR anisotropy at all energies is also consistent with the model, whereas it would not be the case if one assumed a diffusion coefficient in $E^{-0.6}$ (Ptuskin, 1997).

This success means that, as it stands, the transport of the GCRs in the Galaxy is better understood than their acceleration, and the implications of the propagation model should be taken seriously. An important one is that the CR source spectrum is steeper than expected from test particle diffusive shock acceleration, which would indicate that GCRs cannot be SN-EPs. Another important outcome of the propagation models is that the SNR distribution as a function of galactocentric distance appears inconsistent with the data on both CR anisotropy and γ-ray emission from π^0 decay. Indeed, Strong and Moskalenko (1998) have shown that if the GCR sources were distributed similarly to the SN remnants in the Galaxy, the resulting CR gradient would be much steeper than that deduced from the EGRET data above 100 MeV, even in the extreme case of a very extended CR halo. In other words, if the abovementioned gamma-ray emission is dominated by the π^0-decay component generated by GeV nucleons, then the source of these CRs *cannot* be SNRs. Likewise, a CR source distribution concentrated towards the Galactic molecular ring, like the SNR distribution, would produce a much stronger anisotropy than observed in the CR fluxes around 10^{14} eV (by at least one order of magnitude; Gaisser et al., 1995; Ptuskin et al., 1997), and therefore appears inconsistent with the anisotropy measurements as well. Note that both problems would be consistently solved with a smoother CR source distribution.

Of course, one can always invoke a much more complicated propagation model to reconcile the data with the calculations, but one should realise that the Ockham’s razor criterium here is strongly in favour of keeping a simple propagation model which not only fits the data but does so in conformity with theoretical expectations (e.g. a Kolmogorov spectrum for the invoked magnetic turbulence), rather than trying to build a new propagation model with no obvious theoretical justification (see however Ptuskin et al., 1997) so as to keep as is an acceleration model which in other respects suffers from various problems anyway, as we now discuss.

3 Arguments from particle acceleration

As recalled above, an important argument originally supporting SNR shock acceleration as the source of GCRs has been the prediction of a universal power-law spectrum, enabling different independent contributions to be added up smoothly. However, the predicted power-law index from test-particle calculations, $\alpha = 2.0$ (or even 2.1 or 2.2), appears too small to be consistent with the data. But in any case, test particle calculations are clearly not relevant here, since if the SNR shocks are to be responsible for GCR acceleration, they must impart a substantial fraction of their energy to energetic particles, which implies that the CR *pressure* at the shock front must play a significant dynamical role. Non-linear models for diffusive shock acceleration have been developed to include the back-reaction of the accelerated particles on the shock structure, density profile and overall compression ratio. Although a time-dependent, fully consistent model of non-linear acceleration in SNRs is not yet available, several attempts have been proposed, both analytical and numerical, which consistently find that the resulting energy spectrum should not be a power-law at all, but rather exhibit a characteristic concavity, the spectrum being flatter and flatter as the energy increases, up to a cut-off energy where the logarithmic slope is as low as 1.5 or even less (Ellison et al. 1996; Berezhko et al., 1999; Ellison et al. 2000). This represents an additional problem for the standard CR origin model.

As far as the maximum energy is concerned, the situation is even more problematic. It had first been hoped that SNRs could be the source of CRs up to the ‘knee’ observed in the CR energy distribution, around 3×10^{15} eV, since the absence of any feature in the spectrum below this energy (e.g. Asakimori et al., 1998) strongly suggests that it consists of one and only component of energetic particles. Unfortunately, as was early emphasized by Lagage and Cesarsky (1983), SNRs do not appear to be large enough nor to live long enough to accelerate particles up to more than a few 10^{14} eV. Despite nearly two decades of efforts, this remains one of the main problems of the GCR-SNEPs identification. But at any rate, we emphasize that a plausible model for GCR origin should actually be able to account for the CR distribution not only up to the knee, but right up to the ‘ankle’ ($\gtrsim 10^{18}$ eV)! Indeed, it has been noted long ago that the observed smooth distribution of CRs across the knee could not be obtained without fine tuning if two different CR components were involved. We stress here that such a solution would actually require a *double* fine tuning. Indeed, the smooth joining of two power-law distributions of increasing index (2.7 and 3.0 respectively below and above the knee) requires 1) that the first one finishes exactly where the second one starts, and 2) that at the energy where one stops and the other begins, they have exactly the same level (see also Kazanas and Nicolaidis, 2001). This double fine tuning seems highly improbable and speaks very strongly against a SNR origin of CRs, as it seems quite clear that an energy as high as a few 10^{18} eV can certainly not be achieved at SNR shocks.

4 Arguments from observation and phenomenology

From the observational point of view, it should be reminded that no clear identification of energetic CR nuclei in SNRs has been reported yet. Although there may be some good reasons to that after all (notably the predominance of electron-induced radiation), the fact remains that the long hoped direct confirmation of proton acceleration in SNRs is still being awaited. However, the question addressed here is not to know whether nuclei are accelerated in SNRs or not: we do trust that SNEPs exist, and that the theoretical models we have are fairly accurate. But what we want to assess is whether these SNEPs can be the CRs observed at Earth. Judging from the maximum energy attainable in SNRs according to the models, this does not seem to be case. And in this respect, observations seem to agree fairly well with the theoretical expectations. Indeed, not only do we not see the high energy CRs (between 10^{13} and 10^{15} eV, say) in SNRs, but we actually see that they are not there. In a very convincing analysis of the X-ray emission from radio-bright remnant shells, Reynolds and Keohane (1999) put upper limits on the maximum energy achieved by electrons in SNRs, at a few 10^{13}eV at most, with only one exceptions at $2\,10^{14}$ eV. These are conservative, model-independent upper limits. As noted by the authors, these limits also apply for protons and nuclei, since the physical conditions in the remnants considered make it very improbable that the electron spectrum is cut at high energy by synchrotron or inverse-Compton energy losses. As a consequence, the electron and proton rigidity spectra should be identical (except for a different normalization due to different 'injection' processes).

Another difficulty with the GCR-SNEP connection arises from astronomy itself. It is well known that (core-collapse) supernova explosions are induced by massive stars, which are mostly found in associations. Up to 90% of the SN progenitors are members of so-called OB associations which do not disperse before the explosions occur, because of the short lifetime of the most massive stars (e.g. Higdon et al., 1998). As a consequence, many SNe explode one after the other at about the same place in the ISM, leading to the formation of a large structure observed as a superbubble, instead of a collection of individual SNRs. This means that, from the observational point of view, individual SNRs are relatively marginal objects, which should not be expected to play more than a marginal role in the Galactic energy balance and the generation of energetic particles. To state this loosely, if only 10% of the SNe explode in isolation and are responsible for the GCR acceleration, then each of them must impart 10 times more energy to the energetic particles than previously estimated. In other words, the acceleration efficiency must be close to 100%, rather than the usually assumed 10% – a rather extreme value contradicting SNR observations.

Interestingly enough, the collective effect of many SN explosions inside superbubbles has been recently considered as a possible solution of the light element nucleosynthesis problem. Observations of unexpectedly high beryllium and boron abundances in very metal-poor stars in our Galaxy have revealed that CRs accelerated out of the average ISM cannot be responsible for the production of light elements by spallation, as had been thought for more than two decades (see e.g. Vangioni-Flam et al., 2000). However, it was shown that a complete revision of the nucleosynthesis models was unnecessary, and that the data could be easily explained if one assumed that particle acceleration occured inside superbubbles, where most of the energy and the freshly synthesized heavy nuclei were released by SN explosions, in agreement with the theoretical expectations (Bykov, 1999; Parizot and Drury, 1999; Parizot, 2000).

Another important conclusion of these studies is that a power-law energy spectrum extending down to thermal energies, as predicted by single shock acceleration models, is not efficient enough in spallatively producing light elements to account for the amount of Be and B observed in the Galaxy (Parizot, 2000). This therefore puts into question a scenario in which GCRs are the dominant energetic component in the Galaxy and are accelerated at the shock of *isolated* SNRs. This difficulty is reinforced by the fact that the demodulated CR spectrum in the local ISM (Webber, 1998) also shows evidence for a flatter low-energy part. Indeed, the very spectrum which proves to solve the Li-Be-B origin problem is also found to provide a good fit of the inferred CR propagated spectrum (Parizot and Reeves, 2001), which is not the case for a single power-law spectrum. One may conclude from this section that, in many respects, the GCR spectrum does not fit with that expected from SNR shock acceleration (and the same is true for the CR chemical composition, at least in the early Galaxy).

5 Discussion

We can summarize the above arguments in the following way. The remarkable successes of the available CR propagation models encourage one to take their implications seriously. Among them are our two first arguments against a SNR origin of GCRs: 1) the spatial distribution of SNRs is incompatible with the low CR anisotropy measured at energies of 10^{13}–10^{14} eV, and 2) it is also incompatible with the longitudinal profile of γ-ray emission above 100 MeV observed by EGRET (if attributed to π^0 decay). The third problematic implication is that 3) the CR source spectrum is significantly steeper than that predicted from SN shock acceleration models. In addition, 4) if SNRs are the source of GCRs, the shocks must be CR-modified and actually generate a non power-law spectrum, significantly harder than even the 2.0 slope spectrum at energies close to the highest achieved energies. Now in fact, 5) the maximal energy achieved in SNR shock acceleration falls far too short from the 'knee' in the observed GCR energy distribution. Besides, even if the knee could be attained, it would be *extremely* difficult to explain the continuity of the spectrum accross this feature, so that 6) the SNRs should actually accelerate CRs up to the 'ankle', at a few 10^{18} eV, which is virtually impossible!

In addition to these problems at high energy, 7) the low energy spectrum of GCRs appears incompatible with a simple extrapolation of the high-energy power law, as would be expected from SNR shock acceleration models. Besides, 8) the same conclusion is obtained from the study of the independent problem of Li, Be and B nucleosynthesis, from the point of view of both energy spectrum and chemical composition. Moreover, the study of Li-Be-B production reminded us that 9) isolated SNRs are marginal objects anyway, as most of the SN explosion energy (80–90%) is released inside larger structures referred to as superbubbles. Finally, 10) evidence for energetic nuclei inside SNRs could not yet be obtained, although γ-ray fluxes above current detection thresholds had been expected. And even though this non detection can probably be reconciled with the models, 11) X-ray emission from SNRs indicates that the electron spectrum cannot extend up to energies higher than $\sim 10^{14}$ eV. Since electron-specific energy losses are probably not occuring in these SNRs, the maximum energy of nucleons should be comparable to that of electrons, which argues against the SNEP-GCR identification.

As it stands, the case for a SNR origin of GCRs thus seems rather weak. We are certainly not claiming that none of the abovementioned problems can be solved by specific assumptions or re-examinations of the involved theories. Several amendments have been proposed in the literature, and it may well turn that some of the problems can be overcome with quite reasonable assumptions. However, in view of the number of difficulties, we suggest that it may be unlikely for one to solve *all* the problems by a series of modifications and save a standard model for GCR acceleration which would then not look quite standard anymore...

As the original spectrum argument (see Sect. 1) supporting GCR-SNEP identification has collapsed, one should have a look at the last argument: CR energetics. It is still quite impressive that the total power of GCRs in the Galaxy is comparable to the SN power. But it should be noted that this only argues in favour of a model in which the CR *energy* originates from SN explosions. This does not necessarily mean that the GCRs are accelerated *at SNR shocks*. For example, in the abovementioned 'superbubble model' the energy eventually imparted to the energetic particles also originates primarily from SN explosions, but it would consist in a quite distinct CR-origin scenario. As far as energetics is concerned, it should also be stressed that the situation has significantly changed in the last few decades: we are now aware of several sources of mechanical energy which were basically unknown when the SNR model was first put forward, such as jets, plasmoids, gamma-ray bursts, black-hole or neutron-star accretion power.

In conclusion, in the light of the series of problems faced by the standard CR-origin scenario, it may be wise to keep one's mind open and invest some theoretical and observational efforts in alternative scenarios. Nevertheless, the theory of diffusive shock acceleration remains one of the most important fields in high energy astrophysics. It is certainly the key of detailed SNR modelling, and it should also be kept in mind that shocks exist at many different scales in the universe, from interplanetary to galactic and even cluster-size shocks. Such 'non stellar' shocks may prove to play a major role in the acceleration of GCRs after all, and maybe also of the ultra-high energy cosmic rays.

From the observational point of view, several experiments in the near future should provide precious information about both SNR modelling and CR origin, notably INTEGRAL and GLAST. The observation of gamma-ray lines from superbubbles, for instance, may shed a new light on the CR-origin problem. Neutrino astronomy should also provide important constraints. The detection of neutrinos from π^0 decay, consecutively to CR interactions in dense molecular clouds, could be detected with the forthcoming km^3 neutrino detectors and provide the first, clear and still awaited confirmation of the presence of energetic hadrons in the Galactic disk.

References

Asakimori, K., et al., ApJ, 502, 278, 1998.

Baring, M. G., Ogilvie, K. W., Ellison, D. C. and Forsyth, R. J., ApJ 476, 889, 1997.

Berezhko, E. G. and Ellison, D. C., ApJ, 526, 385–399, 1999.

Bykov, A. M., in 'LiBeB, Cosmic Rays, and Related X- and Gamma-Rays', eds. Ramaty, R., et al., ASP Conf. Ser., Vol. 171, 146, 1999.

Ellison, D. C., Möbius, E. and Paschmann, G., ApJ, 352, 376, 1990.

Ellison, D. C., Baring, M. G. and Jones, F. C., ApJ, 473, 1029, 1996.

Ellison, D. C., Berezhko, E. G. and Baring, M. G., ApJ, 540, 292–307, 2000.

Higdon, J. C., Lingenfelter, R. E. and Ramaty, R., ApJ Letters, 509, L33–L36, 1998.

Kazanas, D. and Nicolaidis, A., submitted to PRL, 2001 `astro-ph/0103147`.

Lagage, P. O. and Cesarsky, C. J., A&A, 125, 249–257, 1983.

Parizot, E., A&A, 362, 786–798, 2000.

Parizot, E. and Drury, L., A&A, 349, 673–684, 1999.

Ptuskin, V. S., Adv. Space Res., 19, 697–705, 1997.

Ptuskin, V. S., Voelk, H. J., Zirakashvili, V. N. and Breitschwerdt, D., A&A, 321, 434–443, 1997.

Strong, A. W. and Moskalenko, I. V., ApJ, 509, 212–228, 1998.

Strong, A. W. and Moskalenko, I. V., to appear in Adv. Space. Res., 2001 (`astro-ph/0101068`).

Vangioni-Flam, E., Cassé, M. and Audouze, J., Phys. Reports, 333, 365–387, 2000.

Webber, W. R., ApJ, 506, 329–334, 1998.