Title
Direct observation of recombination-enhanced dislocation glide in heteroepitaxial GaAs on silicon

Permalink
https://escholarship.org/uc/item/18t072v7

Journal
PHYSICAL REVIEW MATERIALS, 2(8)

ISSN
2475-9953

Authors
Callahan, PG
Haidet, BB
Jung, D
et al.

Publication Date
2018-08-09

DOI
10.1103/PhysRevMaterials.2.081601

Peer reviewed
Direct observation of recombination-enhanced dislocation glide in heteroepitaxial GaAs on silicon

Patrick G. Callahan,1 Brian B. Haidet,1 Daehwan Jung,2 Gareth G. E. Seward,3 and Kunal Mukherjee1,*
1Materials Department, University of California Santa Barbara, Santa Barbara, California 93106, USA
2Institute for Energy Efficiency, University of California Santa Barbara, Santa Barbara, California 93106, USA
3Department of Earth Sciences, University of California Santa Barbara, Santa Barbara, California 93106, USA

We use an electron beam to induce and directly observe dislocation glide in strained GaAs thin films on silicon substrates at room temperature using electron channeling contrast imaging (ECCI). Plastic behavior in this brittle material is kinetically facilitated by a lowering of the Peierls barrier for glide due to recombination of cathodogenerated electron-hole pairs at dislocations. The high residual strain in the GaAs film provides the overall driving force for motion. The dynamics of a variety of previously inaccessible dislocation interactions are studied by tracking individual threading dislocations with high spatial and temporal resolution. The ECCI technique has an inherently high resolution and was conducted in a convenient scanning electron microscope environment. The results and methods described in this Rapid Communication provide insight into dislocation-filtering processes and failure mechanisms in technologically important III-V devices on silicon, as well as other mismatched heteroepitaxial systems.

DOI: 10.1103/PhysRevMaterials.2.081601

Dislocations are one-dimensional crystal defects composed of misordered lines of atoms. In general, they control the mechanical deformation properties of crystalline materials, but in semiconductors they also introduce additional electronic states. These electronic states trap and recombine carriers and are almost always deleterious to devices, forming a major obstacle to the heterogeneous integration of semiconductors on low-cost substrates. As an example, direct growth of GaAs on Si would enable integration of low-cost computing with datacom optoelectronics, but the 4% difference in lattice constant results in a very high density of dislocations that render devices ineffective. While the emphasis on controlling dislocations is overwhelmingly due to their negative electronic properties, dislocations are still the agents of plastic deformation, gliding, or climbing in response to external stress. Exploring this dual electronic-mechanical nature of dislocations is an active area of research [1]. In GaAs, this coupling results in a phenomenon unique to semiconductors: carrier recombination via dislocation electronic states enabling dislocation motion. Semiconductors such as GaAs are normally brittle at room temperature and dislocations do not have the energy to move. However, recombination-enhanced dislocation glide (REDG) and recombination-enhanced dislocation climb occur in many III-V devices at ambient temperatures, leading to rampant dislocation growth and device failure [2–5]. Our understanding of these processes has been inferred largely from postmortem transmission electron microscopy of failed devices [6] or low-resolution cathodoluminescence microscopy on bulk semiconductor samples [7]. REDG is phenomenologically understood as a momentary lowering of the dislocation-kink nucleation and migration barriers [5,8] which can be induced locally by carrier injection or above-band-gap illumination [5,9,10]. It is important to note that REDG is realized not simply as a heating of the sample surface [7], but as a direct coupling between the recombination of carriers and bonding at the dislocation core [8].

In this Rapid Communication, we directly observe the recombination-enhanced glide of individual dislocations in heteroepitaxial GaAs thin films on silicon using nondestructive scanning electron microscopy. As the electrons of the imaging beam undergo inelastic scattering and transfer energy to the semiconductor, electron-hole pairs are generated (along with secondary electrons, x rays, auger electrons, etc.) [11]. These electron-hole pairs recombine at dislocations and lower kinetic barriers for dislocation motion driven by residual thermal strain in the GaAs layer. The capability to simultaneously induce and observe dislocation motion provides a straightforward procedure for studying the dual electronic-mechanical nature of dislocations in these and other semiconductors towards improving the reliability of heterogeneously integrated devices.

The GaAs samples examined in this Rapid Communication were grown via molecular-beam epitaxy on on-axis (001) GaP/Si substrates (NASP III/V GmbH) and GaP substrates (ITME) in a procedure described elsewhere [12]. The films were 3.1 µm thick, unintentionally doped p-type (∼10^{16}/cm^3), and are considered fully strain relaxed at growth temperature. Upon cooling to room temperature, however, the GaAs layer on silicon is placed in a significant state of tension (0.2% strain) owing to a significant thermal mismatch; GaAs has a coefficient of thermal expansion (CTE) approximately twice that of silicon. The array of misfit dislocations at the Si/GaAs interface was nucleated during growth to relieve compressive strain, as GaAs has a larger lattice constant than Si. Therefore, these dislocations have the opposite Burgers vector necessary to relieve the residual tensile strain induced upon cooling. As a result, there exists a significant driving force to reduce the total length of interfacial misfit dislocations at room temperature,
on the order of 150–200 MPa. It is this process of residual tensile strain relaxation via recombination-enhanced reduction in misfit dislocation length that we capture in this Rapid Communication.

To resolve individual dislocations in these films, we utilized electron channeling contrast imaging (ECCI) [13,14]—a diffraction-based scanning electron microscope technique that makes it possible to observe the near-surface strain field variations arising from dislocations with minimal or no sample preparation. ECCI measurements were made using an FEI Quanta 400F scanning electron microscope. For the majority of this Rapid Communication, the (220) reflection was used. The electron-beam current varied between 3 and 5 nA (absorbed current, measured at stage), at an accelerating voltage of 30 kV. Two different backscatter detectors were used: the first was a pole piece mounted annular detector, and the second was an array of four square detectors positioned around the sample to collect low-takeoff-angle backscattered electrons. Both detectors were used interchangeably, with preference to the annular detector for high-speed measurements, and the square array for high-contrast [15] imaging. ECCI micrographs shown have been postprocessed individually to increase dislocation contrast, including contrast/brightness adjustments, bandpass filtering, and interframe noise removal.

Under excitation of the imaging electron beam, the dislocation structure in these samples changes dramatically. The primary electron beam generates electron-hole pairs (approximately $10^{17}$–$10^{18}$/cm$^3$) in the tensile-strained GaAs, which go on to recombine at dislocations. This leads to the observed dislocation glide at room temperature by REDG. An ECCI video demonstrating the dislocation motion is included in the Supplemental Material (Video 1, [16]). The plan-view ECCI micrographs in Figs. 1(a) and 1(b) show the extent of these changes in selectively exposed GaAs layers grown on silicon (large CTE mismatch) and GaP (small CTE mismatch). The CTEs at 300 K for Si, GaP, and GaAs are $2.6 \times 10^{-6}$, $4.65 \times 10^{-6}$, and $5.73 \times 10^{-6}$ $\text{C}^{-1}$, respectively [17]. The central region indicated in both cases has been exposed to a rastered electron beam for approximately 35 min at room temperature, while the outer regions have not. The most obvious change at this magnification is the appearance of light and dark lines in the [110] direction. This contrast is due to misfit dislocations, studied experimentally and verified via simulation by Carnevale [18]. In this case, these misfit dislocations originate at the heterointerface and glide via REDG towards the sample surface where they become visible to ECCI. The light and dark contrast arise from their different Burgers vectors with respect to the channeling condition used. In Fig. 1(b), there are many fewer visible misfits, owing to the much lower residual thermal strain driving dislocation motion. Interestingly, this misfit dislocation glide [most visible in Fig. 1(a)] extends 2–3 μm outside the central excited region. The region of influence is a convolution of the size of the plume of electron-hole pairs generated by the impinging electron beam (1-μm radius) and the carrier capture radius of the dislocations in question. This capture radius is governed by the carrier diffusion length, approximately 1–3 μm in GaAs [19–21]. It is worth noting that the plume of generated carriers does not extend to the heterointerface, so carrier diffusion is necessary to explain the appearance of misfit dislocations. Overall, within the region affected by the beam, REDG has allowed dislocations to glide and interact with each other, pinning, reacting, or multiplying, with the general result of increasing the complexity of the dislocation network.

The misfit dislocations visible in Fig. 1(a) only appear in one orientation because of the anisotropy in 60° dislocations in the zinc-blende structure which arises from differences in chemical bonding at their cores [22,23]. α misfits of the glide set have a [110] line direction and an As core while β misfits of the glide set have a [110] line direction and a Ga core. This disparity lends orthogonal misfit dislocations in compressively relaxed (001) films significantly different mechanical and electrical behaviors [5,24–27]. We find that residual strain in the [110] direction is relieved exclusively by the lateral glide of threading dislocations to shorten α misfits at the heterointerface [Fig. 1(c)]. Vertical glide of α misfits
was not observed to contribute to strain relief in any way. In contrast, residual strain in the [110] direction was primarily relieved by vertical glide of β misfits [Fig. 1(d)]. This process occasionally leads to misfit dislocations exiting the sample entirely, but frequently we observed this glide arrested within 100 nm of the surface (see Fig. 2), resulting in visible light and dark lines as in Fig. 1(a). Lateral glide of threading dislocations was also occasionally observed in the [110] direction, but this motion was difficult to measure, as it was typically interrupted by rising misfit dislocations.

These gliding misfit dislocations exhibited interesting behavior, and can be effectively imaged using ECCI within approximately 200–300 nm of the surface; the contrast from deeper dislocations is frequently too diffuse to measure. One such rising misfit event is shown in Fig. 2(a). As the dislocation glides towards the surface, the associated channeling contrast becomes sharper and shifts in the [110] direction. This shift is due to the dislocation gliding on the inclined (111) plane, and not straight towards the surface. This known semiconductor glide geometry allows us to couple the lateral and vertical motion of the dislocation, enabling study of dislocation behavior as a function of depth. In Fig. 2(b), a plot showing the dislocation’s [110] position versus time shows a decay in dislocation speed near the surface. This stagnation is likely the result of a local depletion of carriers due to surface recombination, inhibiting REDG. For very high beam currents (>30 nA), the density of misfit dislocations immobilized at the surface drops dramatically, indicating that increased electron-hole pair generation is sufficient to saturate the surface recombination mechanism, allowing carriers to instead contribute to REDG. This enables gliding misfit dislocations to fully exit the crystal. A video available as Supplemental Material shows two examples of rising misfit dislocations gliding on (111) planes and exiting the crystal (Video 2, [16]).

The unique structure of dislocations in heteroepitaxial thin films makes it difficult to directly apply results from bulk crystal research to this system. In bulk GaAs crystals, dislocations of the glide set are most stable as α, β, or screw type, characterized by having As, Ga, or no dangling bonds, respectively, at an ideal unreconstructed core. However, in (001) oriented thin films, the threading dislocations which terminate α and β misfits typically have an intermediate line direction [28,29], making them not screw or 60° type. The non-screw-threading dislocations terminating α and β misfits have α and β character to match their associated misfit. Previous research has suggested that in undoped and n-type bulk GaAs, α dislocations glide orders of magnitude faster than β dislocations [5,24,30]. However, the asymmetric processes we observed imply that β-misfit dislocations are more mobile here, and act as the primary strain relief mode along [110], while α-misfit dislocations are effectively stationary. There is some evidence in GaAs to show that β dislocations are more susceptible to REDG, experiencing a larger Peierls barrier lowering under excitation [5,8], and that β dislocations are much more sensitive to charged dopants [24,30]. These factors may contribute to the results we observed, but more work is necessary to determine if electronic structure, strain state, or other influences prevent dislocations in these thin films from behaving as in a bulk crystal. These differences in dislocation behavior are technologically relevant, as the favored relaxation mechanisms of REDG-prone strained layers have implications for heteroepitaxial device longevity.

Individual threading dislocation velocities were studied as a function of temperature using a variable temperature scanning electron microscopy stage between −20 and 80 °C, a range typical for semiconductor laser operation in the field. Even within this narrow range, the measured dislocation velocities varied significantly with temperature. This observation provides further evidence that the observed motion is due to a direct electronic coupling and not localized sample heating. If the sample was being heated to high enough temperatures to enable dislocation motion, a relatively small change in
the initial sample temperature would have little impact on dislocation velocity. At the low temperatures in this experiment where GaAs is normally brittle, the measured unpinned dislocation glide velocity is given by \( v = v_0 e^{-E_a/kT} \) [5,8]. Here, \( E_a \) is the activation energy for recombination-enhanced glide and \( v_0 \) is a beam-current-dependent prefactor. \( E_a \) can be expressed as \( E_a = E_d - e\tau E_d \), a function of the stress, \( \tau \), dependent Peierls barrier \( E_d(\tau) \) lowered by an energy \( \Delta E \), which is related to the detailed mechanism of REDG such as energy levels of trap states associated with recombination at the dislocation. In our GaAs/Si films, individual threading dislocations moving in the [110] and [110] directions were tracked for an average of 200 s using high-framerate time-lapse sequences; their velocities versus inverse temperature are shown in Fig. 3(a). Dislocation-dislocation interactions such as pinning are known to occur during the growth of lattice-mismatched semiconductor thin films, lowering the overall glide velocity. We directly measure the effect of pinning on REDG in our films with the capability to track individual dislocations as well as their surroundings. An activation energy, \( E_a \), of 0.2 eV for REDG of [110]-bound threading dislocations was calculated from the fastest dislocation at each temperature. Here, it was assumed that these dislocations experienced the minimum number of interactions and would yield a value closest to the unpinned limit. Despite the activation energy being low, it is sufficient to yield a threefold increase in dislocation velocity when the device temperature is raised from room temperature to 80 °C, important for modeling failure. Using cathodoluminescence (CL) measurements, Maeda et al. [5,7,8] measured the activation energies for REDG of \( \alpha \)-, \( \beta \)-, and screw-type dislocations in bulk GaAs as 0.29, 0.6, and 0.64 eV, respectively. Our results are in reasonable agreement with this report, considering the as-yet unstudied effects of the intermediate line direction of the threading dislocations.

The wide spread in observed dislocation velocities in Fig. 3(a) is the result of the dislocation scattering and pinning events occurring repeatedly in these films. ECCI enables the observation of the jerky motion of individual dislocations and provides insight into these processes. Figure 3(b) shows the position of a single threading dislocation tracked over the course of 20 min. Its motion is impeded multiple times during observation, resulting in a large spread in the instantaneous velocity [Fig. 3(b), inset]. The plateaus highlighted in the figure indicate pinning events, and correspond with interactions between the tracked gliding dislocation and other nearby threading dislocations. The mutual forces between these defects not only impede the tracked dislocation but in some cases visibly shifts the nearby “pinning site” dislocations. The dislocation tracked in Fig. 3(b) is highlighted in a video available in the Supplemental Material (Video 3, [16]). We believe the vast majority of pinning events observed in these films to be thread-thread interactions.

For a dislocation to inhibit another dislocation’s motion, it must either locally reduce the stress driving glide or remove electron-hole pairs from the system that contribute to REDG. As the limiting factor for energy dissipation at clean, impurity-free dislocations is often the recombination rate [31,32], nearby dislocations are not expected to significantly influence this. Instead, dislocations interact with the strain fields of adjacent dislocations, experiencing a reduction in the driving force for glide. Even with the surface sensitivity of ECCI, its ability to observe thick epitaxial layers with realistic strain states differentiates it from other in situ techniques. Previous work has modeled thread-misfit interactions thought to be typical during semiconductor growth [33,34], and further studies directly observed these interactions using in situ TEM [35,36]. ECCI enables the extension of this research to samples of bulk...
thickness, where thread-thread interactions might dominate dislocation pinning.

In addition to transient interactions between passing dislocations, we have also directly observed threading dislocations permanently reacting with each other and with free surfaces. Threading dislocations have been observed to fuse and annihilate, and dislocation segments have been observed to exit the sample via glide through the sample surface and edges. This subset of interactions is critical for reducing dislocation density in heteroepitaxial systems. In Fig. 4(a), two threading dislocations on parallel glide planes are drawn towards each other, react, and vanish from view. We believe this is an annihilation reaction, and the two original misfit dislocations are now connected by a short segment of dislocation at the heterointerface. Figure 4(b) shows the paths taken by these threading dislocations before reacting; the light-contrast dislocation glides in the [110] direction and the dark-contrast dislocation glides in the [−110] direction. Romanov et al. describe a theoretical critical radius of interaction, inside which dislocation annihilation is inevitable [28]. In GaAs, this radius is predicted from growth modeling to be between 50 and 100 nm. The glide planes of the two dislocations we observed are, at minimum, separated by approximately 70 nm. This separation is closed due to either climb (unlikely) or cross slip (likely) in less than one frame at the end of the sequence (≈13.5 s). These results provide direct evidence for individual dislocation annihilation processes in semiconductor thin films that have previously only been inferred indirectly.

Driving dislocations to the free surface at the edge of a sample is another mechanism often considered for threading dislocation removal. To directly observe this process, mesas were dry etched through the GaAs layer to create vertical free surfaces and REDG was used to drive dislocations out of an area of interest. Figure 4(c) shows a threading dislocation gliding in the [−110] direction. After it escapes two pinning sites, it approaches the mesa step and glides completely out of the sample. Multiple threading dislocations were observed to travel to the edge of this mesa, exiting the sample entirely and removing segments of misfit dislocation at the same time. To confirm this, an additional time-lapse sequence [Fig. 4(d)] was taken with the same mesa edge tilted at high angle towards the [110] direction while maintaining the channeling condition. In the third frame of Fig. 4(d), the dislocation appears as a diagonal white line bridging the top corner of the mesa step. This is a small dislocation segment (half loop) on the (111) glide plane, the result of an inclined threading dislocation which has already eliminated its associated misfit and is simultaneously intersecting the sample surface and side wall. By the next frame, the dislocation has exited the sample. This dislocation activity is of technological interest because it not only decreases misfit dislocation length but also
decreases the threading dislocation density near a mesa step. It is possible that a similar technique could be used to intentionally remove threading dislocations from low-dimensional regions such as nanowires or FinFET structures. Additionally, such a treatment would leave the device uniquely prepared to combat REDG-based aging, as it effectively preempts the failure mode, filtering out the dislocations most likely to glide due to nonradiative recombination during device operation. Such a filtering step could also be performed on large areas using diffuse low-energy electron flood sources or intense above-band-gap optical illumination, both capable of creating the electron-hole pairs necessary for REDG [5,9,10].

We have shown that ECCI possesses some fundamental advantages over both CL and TEM for the study of individual dislocation motion in semiconductors. ECCI enables the imaging of large sample areas without requiring sample preparation that alters the material’s strain and electronic states. Likewise, while CL is capable of resolving long-range REDG, in general, the technique is limited to light emitting semiconductors and suffers from low spatial resolution, reported as \( \approx 3 \, \mu m \) by Maeda et al. [7]. This limits its ability to resolve the finer movement and interactions between dislocations seen in this Rapid Communication, typical of higher defect density devices on silicon. With careful device and contact mask design, we propose that ECCI could even be used to track dislocation evolution in operando. This is a fascinating opportunity to explore the failure mechanisms of semiconductor devices such as GaAs/Si quantum dot lasers [37,38] and GaN [39] and SiC [40] power devices where carrier recombination or electric field-induced dislocation motion is suspected or has been observed postdegradation.

The authors thank J. E. Bowers and A. C. Gossard for their leadership of the University of California Santa Barbara quantum dot laser silicon effort. This work was supported by the NSF through the Materials Research Science and Engineering Centers (MRSEC) Program of the NSF through Grant No. DMR-1720256 (Seed Program). P.G.C. acknowledges support from the NSF DMREF Grant No. DMR 1534264. P.G.C. and B.B.H. contributed equally to this work.


