Understanding the electronic properties of defects and the ability to control them will be crucial for the performance of the future microelectronic devices [1]. Scanning Tunneling Microscopy (STM) represents a unique tool for the studies of defects as it combines atomic scale resolution with local spectroscopic capability. However, STM observation and analysis of defect states in semiconductors are complicated by surface effects such as in-gap surface states and reconstruction. These effects are avoided at the (110) surfaces of a number of III-V semiconducting systems [2], attracting extensive research [3]–[8]. A number of point defect types have been observed. However, positions of these defects with respect to the surface plane could be inferred only from indirect observations. The interpretation of such observations is complicated by the drastic effect the surface proximity may have on the defect states [1].

Modeling STM measurements of defects in semiconductors is not straightforward: Approximation of the STM images by maps of the local surface electronic density of states [11] is justified only if the charge relaxation rates of defect states significantly exceed the tunneling rate of electrons between the tip and the sample [1]. Tip-induced effects also need to be taken into account. These may include both local band bending [3], and charging of the defect states by the tunneling current, resulting in bias voltage-dependant lattice relaxation in the vicinity of the defect atoms [1]. Careful analysis is necessary to clearly separate these effects from the intrinsic defect properties, and the bulk features of the observed defect states from the surface effects.

In this paper we present cryogenic STM and scanning tunneling spectroscopy (STS) studies of the layered narrow gap semiconductor Bi$_2$Se$_3$, which can be viewed as a model system for the STM study of near-surface defect states. The bonding scheme of Bi$_2$Se$_3$ allows a direct determination of the position of a subsurface defect atom with respect to the surface: Atomic planes consisting of either Bi or Se hexagonally arranged atoms are stacked in a close-pack fcc fashion; 5 atomic planes with atomic order Se$_1$-Bi-Se$_2$-Bi-Se$_1$ (Fig. 1(a)) form a layer. The layers are weakly bound to each other by Se$_1$-Se$_1$ bonds. Both valence and conduction bands are formed almost exclusively by the 4$p$ and 6$p$ orbitals of Se and Bi respectively [2]. For each atom, the closest neighbors from the adjacent atomic planes form almost a regular octahedron, so the bonding can be roughly approximated by strongly interacting $pp\sigma$ chains of atomic $p$-orbitals (Fig. 1(b)), with a weaker $pp\pi$-type interaction between adjacent chains. A substitutional defect is therefore likely to produce a perturbation in the electronic local density of states (DOS) predominantly along the three $pp\sigma$ chains passing through the defect atom. Hence, the defect state should be observed at the surface as three spots of modified DOS around the atoms terminating these chains at the surface.

We have performed STM and STS measurements using a custom built low-temperature microscope with direct immersion in liquid He-4 [3]. Stoichiometric Bi$_2$Se$_3$ single crystal samples, as well samples doped with 2–5% excess Bi or Se were grown by a directional solidification technique. The stoichiometric as-grown Bi$_2$Se$_3$ samples are n-type with carrier concentration of about 10$^{19}$ cm$^{-3}$. Doping samples with excess Bi introduces substitutional Bi defects at the Se sites (Bi$_3$Se$_5$ antisites), which are shal-
low acceptors. However, because of the low solubility of Bi in Bi$_2$Se$_3$ [14], the Bi-doped samples are n-type due to the compensating defects. Doping samples with excess Se introduces shallow donor-type Se substitutional defects at the Bi sites (Se$_{Bi}$ antisites). The samples are sufficiently inert to obtain atomic resolution in air. However, to minimize the surface contamination, in the experiments reported here, the samples were cleaved in situ or in a glove box directly attached to the STM setup in ultrapure He gas prior to transfer to the STM with subsequent cooling to T=4.2 K.

To map out the defect states, we have performed differential current imaging tunneling spectroscopy (CITS) measurements [15]. The CITS maps were acquired by fixing the tip at each point during topographic imaging and measuring the differential conductance at various bias voltages with the feedback loop disabled. Only Bi-doped samples exhibited an appreciable density of observable defect states, as shown in Fig. 2(a) for a sample doped with 5% excess Bi (Bi$_2$Se$_{2.85}$). The map was acquired with sample bias $V_b=−0.2$ V in the feedback mode (with tunneling current set to 50 pA), and $V_b=−0.45$ V for the conductivity measurement. The defects appear as regular clover-shaped bright features, indicating areas with locally enhanced conductance at $V_b=−0.45$ V. Irregular spots in this image (mostly in the upper left corner) resulted from topographic defects. The data discussed below were obtained on a more weakly doped Bi$_2$Se$_{2.95}$ sample, where the defect density was reduced.

Fig. 2(b),(c) present topographic maps of the sample area encompassing an isolated clover-shaped defect state. Topographic image Fig. 2(b), acquired at a sample bias voltage $V_b=−0.3$ V, shows a periodic atomic structure, indicating no significant structural variation associated with the defect. The height of the atomic corrugations in Fig. 2(b) is about 30 pm. Fig. 2(c) shows a topographic image of the same area acquired at $V_b=−0.6$ V, where the largest corrugations, locally enhanced by tunneling through the defect state, are about 100 pm high. The highest amplitude of the defect state correlates with positions of three surface Se atoms (marked with larger black circles in schematic Fig. 2(d)), forming a regular triangle. These atoms terminate three p$p\sigma$-bonded chains passing through the Se1 site five atomic layers below the surface, for which we also use notation Se3. Since the observed defects appear only in Bi-doped samples and they originate from Se sites, we attribute them to the Bi$_{Se3}$ antisites.

Fig. 2(a) presents a series of differential conductance spectra acquired in the vicinity of an isolated defect. The spectra were obtained by numerical differentiation of 60 I-V curves with setpoint parameters $V_b=−0.3$ V, $I=0.8$ nA. At the measurement temperature of 4.2 K, thermal broadening is negligible on the displayed bias voltage scale. The defect state appears as a broad resonance (indicated by an arrow) in the energy range where the differential conductance is suppressed away from the defect. The established semiconducting gap value is
about 0.3 eV \[12, 13\], therefore the defect levels appear inside the valence band. Theoretical modeling is necessary to understand this spectroscopic feature, as well as why only Bi\textsubscript{Se} antisites five atomic planes below the surface are observed.

First, we performed \textit{ab initio} calculations in the full potential relativistic LAPW formalism \[16\] within LDA approximation. To model the surface, a supercell geometry was used, with distance between slabs (consisting of 15 atomic planes, or 3 layers, each) increased by 0.5 to 1.5 nm as compared to the bulk crystal structure. The calculated band structure did not exhibit significant variation for the slab separation larger than 0.4—0.5 nm, therefore we found the slab separation of 0.7 nm, used for the calculations presented below, sufficient for modeling the surface. The differential conductance spectra were approximated by the local DOS in the center of the gap between the slabs \[10\]. The calculation presented in Fig. 3(b) was performed for a position above a Se\textsubscript{1} atom, although we found the variation of the calculated spectra with position respective to the surface atoms to be insignificant. The calculation reproduces both the finite conductance in the bulk semiconducting gap, and the suppressed conductance just below the gap. In Fig. 3(c), the total DOS calculated in the slab geometry is compared to the calculation of the bulk DOS, which reproduces the accepted semiconducting gap value of 0.3 eV \[13\]. Band structure analysis indicates that the highest valence band (HVB) states are predominantly Se\textsubscript{1}-Se\textsubscript{1} antibonding type. As the Se\textsubscript{1}-Se\textsubscript{1} bonds are broken at the surface, the splitting of these states is reduced, resulting in the observed suppression of the differential conductance in the HVB energy range (Fig. 4(b)). The states that appear in the gap have high dispersion along the surface. They originate from the rehybridization of the surface Se\textsubscript{1} valence with Bi conduction states, bringing the latter down below the bulk conduction band minimum.

\textit{ab initio} calculations of a single defect state are complicated by the large cluster or supercell size necessary to model the impurity states without introducing artificial interaction between defects. Instead we use a LCAO approximation as a simple model of the system \[12\]. This model presents just a qualitative argument and is not capable of reproducing the detailed electronic structure or the semiconducting gap value. However, it gives a surprisingly good qualitative agreement with the experimental observations and first principles calculations. In a tight binding formalism

\[ H\psi = H\sum_i u_i\phi_i = \sum_i E_i u_i\phi_i + \sum_i \sum_{i\neq j} V_{ij} u_i\phi_j, \]

where \( \phi_i \) are atomic wave functions and \( V_{ij} \) are off-diagonal matrix elements of \( H \). We approximate \( E_i \) by the atomic term values \[15\] of Bi and Se, and take into account only \( p\pi \sigma \) interaction between the closest neighbors, as shown in Fig. 4(b). Thus the problem is reduced to a system of noninteracting one-dimensional chains, with three matrix elements \( V_1, V_2, \) and \( V_3 \), corresponding to Se\textsubscript{1}-Bi, Se\textsubscript{2}-Bi, and Se\textsubscript{1}-Se\textsubscript{3} \( p\pi \sigma \) bonds. Consider first a 5-atom chain Se\textsubscript{1}-Bi-Se\textsubscript{2}-Bi-Se\textsubscript{3} (we also call it a unit), representing a single layer in our approximation. Here Se\textsubscript{3} position is equivalent to Se\textsubscript{1}. The highest energy filled state is nonbonding

\[ \psi_0 = \frac{1}{\sqrt{2 + (V_1/V_2)^2}}(\phi_{S1} + \phi_{S3} - V_1/V_2\phi_{S2}). \]

We model the bulk by a long chain of Se\textsubscript{1}-Bi-Se\textsubscript{2}-Bi-Se\textsubscript{3} units, with the interaction between layers expressed by \( V_3 \). Due to the Se\textsubscript{1}-Se\textsubscript{1} interaction, the nonbonding level is split, and the HVB states become antibonding in the sense of Se\textsubscript{1}-Se\textsubscript{1} bond character. The splitting is large, because the state Eq. 3 has a significant weight on the interlayer Se\textsubscript{1} atoms. As the interlayer bonds are broken at the surface, the HVB states of a chain decay at the surface (the end of the chain), as shown in Fig. 4(a), where the antibonding character of the valence band maximum (VBM) state can also be seen. In semiclassical terms, as illustrated in Fig. 4(b), the surface gap is larger than the bulk value. This effect is in agreement with the more accurate first principles calculations and spectroscopic measurements, Fig. 3(b), (c).

To highlight the importance of these surface effects for the observation of the defect states, in Fig. 4(c) we plot the calculated dependence of the near-gap energy levels on the position of the Bi\textsubscript{Se} antisite. The second layer (positions 4–6) is only weakly affected by the proximity of the surface, so the defect level is split from the VBM and its energy is only weakly dependent on the position. As the antisite position approaches the surface (positions 1–3), the defect level energy is reduced as the surface gap opens up, so that the defect state merges with the bulk valence states, forming a resonance inside the bulk valence band. This behavior is supported by the semiclassical picture shown in Fig. 4(b). Only the Bi\textsubscript{Se\textsubscript{3}} state is observed in the experiment, because, as our model suggests, Bi\textsubscript{Se\textsubscript{1}} and Bi\textsubscript{Se\textsubscript{2}} states are so much lowered in energy by the proximity of the surface that they form small amplitude broad resonances in the valence band. STM images do not exhibit defect features associated with Bi\textsubscript{Se}, antisite in the second layer. Our model suggests, that they form bound states in the bulk gap, which cannot sustain STM current. Surface effects thus provide a mechanism for the charge relaxation of near-surface defect states through the bulk valence band.

Similar surface effects should be observable in other semiconductors, e.g. at (110) surface of GaAs, where the valence band is, like in Bi\textsubscript{Se\textsubscript{3}}, suppressed at the surface \[4\]. As a result, in-gap impurity states may become resonances in the valence band, if the originating impuri-
We thank J. Nogami, Norman O. Birge, M. I. Dykman, T. Hogan, and S. Lal for helpful discussions. This work was supported in part by NSF (DMR-0075230) and ONR/DARPA (N00014-01-1-0728). SHT acknowledges support of the Alfred P. Sloan Foundation.

[1] See, e.g. B. E. Kane, Nature 393, 133 (1998).
[2] Xuejun Zhu et al., Phys. Rev. Lett. 63, 2112 (1989).
[3] J. F. Zheng et al., Phys. Rev. Lett. 72, 1490 (1994).
[4] R. M. Feenstra, J. M. Woodall, and G. D. Pettit, Phys. Rev. Lett. 71, 1176 (1993).
[5] B. Grandidier et al, Appl. Phys. Lett. 72, 2454 (1998).
[6] Chr. Wittneven it et al., Phys. Rev. Lett. 81, 5616 (1998).
[7] B. Grandidier et al., Appl. Phys. Lett. 74, 1439 (1999).
[8] Ph. Ebert, K. Urban, and M. G. Lagally, Phys. Rev. Lett. 72, 840 (1994).
[9] R. B. Capaz, K. Cho, and J. D. Joannopoulos, Phys. Rev. Lett 75, 1811 (1995).
[10] J. Tersoff and D. R. Hamann, Phys. Rev. Lett. 50, 1998 (1983).
[11] X. de la Broise et al., Phys. Rev. B 61, 2138 (2000).
[12] S. K. Mishra, S. Satpathy, and O. Jepsen, J. Phys.: Cond. Matt. 9, 461 (1997).
[13] S. H. Tessmer, D. J. Van Harlingen, and J. W. Lyding, Rev. Sci. Instrum. 65, 2855 (1994).
[14] D. R. Lovett, Semimetals and Narrow-Bandgap Semiconductors (Pion Limited, London, UK, 1977).
[15] J. A. Stroscio and R. M. Feenstra, in Scanning Tunneling Microscopy, Methods in Experimental Physics Vol. 27 (Academic Press, New York, 1993).
[16] David J. Singh, Plane waves, Pseudopotentials and the LAPW method, (Kluwer Academic Publishers, Boston, 1994).
[17] Although LDA approximation usually underestimates the semiconducting gap value, it has been rather successful in application to chalcogenide semiconductors. See, e.g. P. Larson, S. D. Mahanti, and M. G. Kanatzidis, Phys. Rev. B 61, 8162 (2000).
[18] W. A. Harrison, Electronic Structure and the Properties of Solids (W. H. Freeman and Co., San Francisco, 1980).