Experimental Realization of a Topological p–n Junction by Intrinsic Defect Grading

Thomas Bathon, Simona Achilli, Paolo Sessi,* Vladimir Andreevich Golyashov, Konstantin Aleksandrovich Kokh, Oleg Evgenievich Tereshchenko, and Matthias Bode

The recent discovery of topologically protected surface states in the binary chalchogenides Bi$_2$Te$_3$[1] Bi$_2$Se$_3$[2] and Sb$_2$Te$_3$[3] has tremendously revitalized the interest in these narrow gap semiconductors that have been studied for decades because of their excellent thermoelectric properties.[4] Topological insulators (TIs) are materials insulating in the bulk but metallic on their surface due to the existence of linearly dispersing gapless states which, contrary to the trivial surface states usually found in metals and semiconductors, are protected by time reversal symmetry.[5] The strong spin–orbit coupling characterizing these materials perpendicularly locks the spin to the momentum, resulting in a chiral spin texture which forbids backscattering[6,7] and makes spin currents intrinsically related to charge currents.[8]

These unconventional properties make TIs an ideal platform to realize exotic states of matter, such as Majorana fermions[9] and magnetic monopoles.[10] Furthermore, they appear to be ideal candidates to realize magneto-electric and spintronics devices with low power consumption.[11] Within this framework, the creation of p–n junctions (PNJ), which are the building blocks of several semiconducting devices, such as diodes, sensors, solar cells, or transistors, in topological materials would represent the first step toward the direct application of this fascinating class of materials.[12] In TIs, the massless character of surface carriers and the helical spin texture inversion at the p–n interface are predicted to result in several unconventional phenomena, e.g., gapless chiral edge mode,[12] Klein tunneling,[13] Veselago lenses,[14] exciton condensation, and charge fractionalization.[15]

Experimentally, all these fascinating properties remained unexplored so far, mainly because prototypical topological PNJs were not readily available. One issue toward the successful fabrication of topological PNJs is the presence of defects which are intrinsically incorporated in TI materials during the growth process. Thereby, the resulting TI crystals are heavily doped at a relatively high bulk carrier concentration level.[16,17] Since bulk carriers give rise to leakage currents that make capacitive charging impossible their existence represent a severe obstacle toward the creation of topological PNJs by applying a gate bias between opposite surfaces of a TI thin film.[15]

Recent studies showed, however, that—depending on the growth conditions—large deviations from the nominal composition can be obtained.[18] These defects, which can be antisites or vacancies, are characterized by different formation energies. Here, we demonstrate that their inevitable presence can be turned into a positive effect by thermodynamically controlling their distribution across the crystal. In particular, we show how appropriate growth conditions result in the creation of bulk crystals that intrinsically contain topological PNJs at their surface, thereby avoiding the complicated fabrication of heterostructures and the problems related to the creation of interfaces between different materials.[19–21] Atomic scale scanning tunneling microscopy (STM) and spectroscopy (STS) measurements combined with ab initio calculations evidence that Te segregation[17,22,23] results in well-separated Te- and Bi-rich regions that display p- and n-transport character, respectively. Spatially resolved Hall and Seebeck measurements confirm the transition from p- to n-like transport. At the p–n interface the oppositely charged defects, i.e., donors or acceptors, compensate, resulting in a substantial drop of conductivity by almost two orders of magnitude. Scanning tunneling microscopy reveals a built-in potential of about 110 meV, thereby considerably shifting the Dirac point in between the p- and n-region.
Spectroscopic data obtained within the p–n compensation region indicate that the junction width amounts to several 10 nm only.

Figure 1 schematically presents some essential background information related to our approach. Within the hatched box of Figure 1 the crystal structure of a quintuple layer of the binary chalcogenide Bi$_2$Te$_3$ is shown. It consists of alternating Te and Bi layers. Adjacent quintuple layers are weakly bound by van der Waals forces thereby offering a natural cleavage plane, a situation highly advantageous for STM experiments. The bottom panel of Figure 1 shows a photographic image of a Bi$_2$Te$_3$ crystal grown by the modified Bridgman technique (see the Experimental Section for details). The crystal has a nominal composition with a Te content of 61 mol%. As recently shown, the conduction of Bi$_2$Te$_3$ crystals changes from p- to n-type at approximately this Te concentration. Due to Te segregation, however, the stoichiometry will not be constant but continuously change during the growth process from Bi-rich to Te-rich, as indicated by different color shadings in the bottom panel of Figure 1. Correspondingly, we expect Dirac points which are energetically located below and above the Fermi level, respectively, with opposite rotational sense of the spin polarization. As we move from p- to n-doped surface areas we expect to pass through a transition region with equal concentrations of Bi- and Te-induced charge carriers where the carrier concentration becomes minimal, thereby realizing an intrinsic TI.

In fact, the differently doped areas can clearly be distinguished by STM and STS. Figure 2a displays an overview topographic STM image obtained at the (0001) surface of the Bi-rich crystal region. Four different defects labeled (i–iv) are visible on the surface. Since the defects symmetry and extension reflects the perturbation introduced by the bonding structure within the crystal\cite{25} (see bond geometry sketched in the inset of Figure 1), a detailed analysis of the atomically resolved images reported in the top row of Figure 2b allows to identify their location within the quintuple layer structure of Bi$_2$Te$_3$, i.e., Te(I)-Bi(II)-Te(III)-Bi(IV)-Te(V), where Te(I,V) and Te(III) represent two non-equivalent Te planes. Following this procedure the defects (i–iv) are all located within Te layers, and they can be Te vacancies ($V_{\text{Te}}$) or antisites (Bi$_{\text{Te}}$). Due to the lack of chemical sensitivity of STM a definite assignment can only be achieved by a comparison of the experimental data with simulated STM images obtained from ab initio calculations (bottom row of Figure 2b). Based on the good general agreement achieved between the experimental and theoretical images we can make the following defect assignment: defect (i) is a Bi$_{\text{Te}(i)}$ antisite; defect (ii) is a Te...
surface vacancy (V_{TeIII}); defect (iii) is a Te vacancy in the third layer (V_{TeIII}); and defect (iv) is a Bi antisite in the fifth layer (Bi_{TeIV}). Since all defects are located in Te layers this region is expected to be Te-poor. This is confirmed also by theoretical calculations which show that these defects are characterized by the lowest formation energy in Bi-rich regions. The electronic properties of this sample region can be characterized by an investigation of the local density of states (LDOS) as inferred by the STS spectra reported in Figure 2c. Following the energy level positioning scheme adopted in Ref. [16], we obtain a valence band (VB) maximum close to the Fermi level, thereby indicating the p-doped character of this sample region.

Investigation of the opposite side of the crystal reveals four other defects (v–viii) (Figure 2d), the STM appearance of which is quite different from the data presented in Figure 2a before. Again the symmetry and lateral extension of the various defects in atomic resolution data were used to estimate the position of the defects. Based on these estimations ab initio density functional theory (DFT) calculations were performed. The comparison presented in Figure 2e reveals a very reasonable agreement. Our results show that all four defects are located in the two Bi layers of the Bi_{2}Te_{3} quintuple layer, in agreement with the expected Te-rich stoichiometry. Namely, the following assignments were concluded: defect (v) is a Te_{Bi} antisite; defect (vi) is a Bi vacancy (V_{BiII}); defect (vii) is a Te_{Bi} (V_{BiIV}) antisite; and defect (viii) is a Bi vacancy, V_{BiIII}. As for the p-doped region, our defects identification is consistent with the formation energies calculated by Scanlon et al., [16] which show that in Bi_{2}Te_{3} grown under Te-rich conditions the lowest energy defects are Te_{Bi} antisites and Bi vacancies, with the last ones playing a much less significant role. Comparison of the STS spectra shown in Figure 2, panels c,f reveals that the different types of defects present in Bi- and Te-rich areas strongly influence the electronic properties, i.e., Te-rich regions show a rigid spectral shift toward negative energies with respect to Bi-rich areas. This negative shift amounts to approximately 140 meV between the p- and n-doped surface areas of the Te-rich and Bi-rich parts of the crystal, respectively. This observation implies that a topological PNJ is naturally present within our crystals.

Figure 3a shows the room-temperature Seebeck coefficient $S$ of a Bi_{2}Te_{3} single crystal measured at various positions (0–60 mm) along the crystal growth direction x. It continuously and slowly changes along the crystal rod with a sign change occurring at a position of 35 mm. A positive Seebeck coefficient in the range of 250–100 $\mu$V K$^{-1}$ is obtained for the left part of the crystal boule (starting from x = 0). Further right to this region, i.e., between 30 and about 35 mm, the Seebeck coefficient rapidly drops and eventually becomes negative, indicating that the crystal conductivity changes from p- to n-type at the point where $S = 0$ $\mu$V K$^{-1}$. In the region of the p–n transition ($S = 0$ $\mu$V K$^{-1}$) a potential Seebeck microprobe demonstrates the change of sign of the Seebeck coefficient on the 200 mm distance, limited by the resolution of the method. At x = 47 mm a high value $S = -390$ $\mu$V K$^{-1}$ is obtained.

This interpretation is also corroborated by Hall measurements presented in Figure 3b which were taken at opposite sides of the boule. They show opposite signs of Hall coefficients $R_{H}$, thereby confirming the presence of both p- and n-type regions. The left part of the boule (0–30 mm) shows metallic behavior with small positive Hall coefficients leading to a weakly temperature-dependent p-type carrier concentration, $p = (5–10) \times 10^{18}$ cm$^{-3}$. At the right part a similar metallic-like behavior was observed for n-type carriers ($n = 10$ $\times 10^{19}$ cm$^{-3}$). The large carrier concentration and its weak temperature dependence indicate that opposite sides of the boule are heavily doped by acceptors and donors, respectively. Due to a strong reduction of the carrier concentration by two orders of magnitude ($10^{17}$ cm$^{-3}$) in the vicinity of the p–n junction, the Hall coefficient changes sign and drops down to 45 cm$^{-1}$ C$^{-1}$ at low temperatures. These data indicate that an intrinsic binary chalcogenide TI can be obtained by a minimization of the carrier concentration through a careful control of the alloy composition between Bi and Te.

Electrical resistivity temperature dependencies were measured on the samples taken from positions A, B, and C (Figure 3a), i.e., within a distance of 2 mm of the sign change of the Hall
coefficient. In all the points resistivity decreases for temperatures from RT to $T = 30$ K, and then remains almost constant down to the liquid helium temperatures for p- and n-type samples, while for point B resistivity increases, demonstrating semiconducting behavior. In both p- and n-parts of the crystal, the mobilities show a high average value of $10^4$ cm$^2$ V$^{-1}$ s$^{-1}$ at low temperatures (Figure 3c). At $x = 45$ nm in the n-type region, i.e. close to the p–n transition, $\mu$ reaches the maximum value of about $10^5$ cm$^2$ V$^{-1}$ s$^{-1}$. At the same position a maximum Seebeck coefficient $S = -400 \mu$V K$^{-1}$ is observed, meaning that near the p–n transition there is the suppression of the defects formation.

In order to investigate the minimal width of intrinsic TI p–n transitions that result from the slightly graded stoichiometry during the crystallization process, we have performed STM/STS in the zone where the transition was detected in the transport measurements described above. An overview scan is displayed in Figure 4a. While the surface region in the left and the bottom right part of the image is relatively rough (marked R) because of numerous small islands and clusters, we also find regions with extended atomically flat terraces separated by unit-cell step edges (F). As we zoom into such an atomically flat region (Figure 4b) we can observe the coexistence of defects which we identify as being characteristic for n- and p-type Bi$_2$Te$_3$, i.e., defects (i), (ii), and (iii) as well as (v) and (vi), respectively (see above). This microscopic picture explains the drop of the Hall coefficient observed in transport, see Figure 3 and Ref. [18,23]. It is also in agreement with the mobility drop detected in the p–n junction region, which can be thus directly linked to the increased number of total defects. Overall, these findings have deep implications going beyond the scope of the present work. Minimizing bulk conductivity has always been considered as a powerful strategy to let topological surface states dominate the scene. However, it was not clear up to now if this effect should be ascribed to defects suppression or compensation. The combination of transport measurements with atomic scale characterization techniques provides thus a detailed microscopic picture solving this problem, identifying defect compensation as the responsible mechanism. Defect suppression can be safely ruled out since it would require a significant reduction of the defects concentration, a condition not consistent with our results.

Local spectroscopic data show that these regions contain several PNJs which are very sharp. This is corroborated by the color-coded STS data shown in Figure 4c measured with a 3 nm increment along a 60 nm long line in a flat region. The minimum position is marked by a red dot in each spectrum. Indeed we observe a shift of the minimum from close to the Fermi level $E_F$ at $x \leq 0$ nm to $E - E_F = 110$ meV at $x \geq 50$ nm, i.e., the transition from p- to n-doped Bi$_2$Te$_3$ takes place over the remarkably short distance of about 40 nm, although the p- or n-type character observed at the two ends of the boule (Figure 2c,f) is not fully reached.

Experimental data obtained on a larger length scale indicate that the doping gradient is not monotonic but exhibits some fluctuations which lead to the coexistence of several p–n transition regions. Their distance typically amounts to $\approx 1$ nm, a value also matching the length scale of flat (F) and rough (R) surface patches in Figure 4a. Both behaviors potentially reflect the complicated convection processes that take place during the slow cooling and crystallization process of the boule. Their averaged random distribution results in an Ohmic characteristic preventing spatially averaging techniques from verifying the existence of a rectifying behavior, an aspect which may be directly tackled by using nanoscale four probes techniques. Although their creation is not yet fully controlled, our data indicate that p–n transition regions naturally resulting from the slightly graded stoichiometry of crystals may be very narrow. Further research may lead to optimized procedures which better control their position and width to make them useful for application.

Figure 4. a) Overview topographic STM image of the Bi$_2$Te$_3$ crystal taken in the p–n transition zone (scan range: $1.5 \mu$m $\times$ $1.5 \mu$m). The surface shows two characteristic regions each of which is several hundred nm wide; while the rough surface region (R) contains numerous islands and clusters the other, the flat region, exhibits atomically flat terraces (F) that are separated by unit cell step edges. The zoomed in image (b) was taken within a flat region. Defects characteristic for both, n- (i–iv) and p-type (v–viii) Bi$_2$Te$_3$ can be found. c) Sequence of color-coded tunneling spectra taken along a 60 nm long line (3 nm increment) within a flat region. The energetic position of the minimum, which is marked by a red point in each spectrum, shifts from just above the Fermi level (bottom) to about 110 meV, indicative for a p–n transition with a width of about 40 nm. The hatched lines are guides to the eye.
In summary, we demonstrated that, contrary to other materials, the ambipolar behavior of Bi$_2$Te$_3$ makes possible, by taking advantage of the intrinsic defects which are unavoidably introduced during the growth process, to create topological p–n junctions. Scanning tunneling spectroscopy reveals that their width amounts to 40 nm only. The related potential drop of approximately 110 meV shall allow for direct applications of PNJs in room temperature devices. Additionally, the defect compensation naturally achieved at the p–n interface results in a strong reduction of the bulk carriers concentration, thus paving the way to explore the PNJ topological properties.

**Experimental Section**

Bi$_2$Te$_3$ crystals have been grown by the modified Bridgman technique with a temperature gradient of about 10 K cm$^{-1}$ at the front of crystallization.$^{[24]}$ As-grown ingots had a single crystalline structure and were split into two parts along the cleavage plane (0001) oriented along the growth direction (Figure 1). One part of each crystal was cut perpendicular to the growth axis into 0.5–1 mm samples. Indium solder Ohmic contacts were used for transport measurements. The Hall resistance $R_H$ and the resistance $R_{xx}$ were measured in the Hall bar geometry using a standard six-probe method on rectangular samples. A potential Seebeck microprobe was used to investigate the room-temperature Seebeck coefficient in the region of p–n transition with a spatial resolution of 200 µm. The Hall mobility was calculated from the measured conductivity and calculated carrier concentration, which was extracted from the measured Hall coefficient

$$\eta = \frac{1}{R_H \times \epsilon}$$

The carrier concentration dependence along the crystal is shown in Ref. [24].

STM measurements were performed at a tip and sample temperature $T = 4.8$ K with electro-chemically etched tungsten tips. After transfer into the ultra-high vacuum system, samples were cleaved at room temperature at a base pressure $p < 2 \times 10^{-11}$ mbar and immediately inserted into the cryogenic STM. The PNJ was located by bringing the tip into tunneling distance from the surface and recording a local tunneling spectrum indicative for p- or n-type doping (Figure 2c,f). After retracting the tip the sample was moved with an x–y-stage (initially by several tens to hundreds of µm). This procedure was repeated until a spectral shift signaled that the boundary to the region governed by the other dopant had been crossed. Then a refined procedure with smaller x–y-movements was performed.

**Theoretical Calculation:** Ab initio theoretical calculations were performed by density functional theory, using a pseudo-potential representation of the electron–ion interaction and local orbital basis sets, as implemented in the SIESTA code.$^{[27]}$ A generalized gradient approximation (GGA) was used for the exchange-correlation functional$^{[28]}$ and a plane wave cutoff equal to 250 Ry. The experimental lattice constant for the hexagonal cell of Bi$_2$Te$_3$ was used$^{[29]}$ and a constant-distance STM images have been simulated in the Tersoff–Hamann model$^{[30]}$ by calculating the Kohn–Sham local density of states in the energy interval between the Fermi level $E_F$ and $E_F + eV_{x}^{\text{TH}}$ where the applied bias $eV_{x}$ was fixed in agreement with the experimental setup. In particular the moderately large value of $eV_{x}$, of the order of 0.4 eV, allows us to obtain numerically stable results and to reasonably neglect the small inaccuracy of the LDOS around the Fermi level due to the absence of spin–orbit interaction in the calculation. A tip–surface distance equal to 3 Å was considered, although it was verified that the result of the STM simulations were robust against small variations of this value. The extension of the tip was taken into account in the STM simulation by considering the mean value of LDOS within a radius of 4 Å around the tip position.

**Acknowledgements**

The authors gratefully acknowledge stimulating discussions with Mario Italo Trioni. This work was supported by the Deutsche Forschungsgemeinschaft within SPP 1666 (Grant No. BO1468/21-1). K.A.K. and O.E.T. acknowledge the financial support by the RFBR (Grant Nos. 13-02-92105 and 14-08-31110).

Received: September 28, 2015  
Revised: December 3, 2015  
Published online: January 18, 2016

[1] Y. L. Chen, J. G. Analytis, J.-H. Chu, Z. K. Liu, S.-K. Mo, X. L. Qi, H. J. Zhang, D. H. Lu, X. Dai, Z. Fang, S. C. Zhang, I. R Fisher, Z. Hussain, Z.-X. Shen, Science 2009, 325, 178.
[2] Y. Xia, D. Qian, D. Hsieh, L. Wray, A. Pal, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, M. Z. Hasan, Nat. Phys. 2009, 5, 398.
[3] Y. Jiang, Y. Wang, M. Chen, Z. Li, C. Song, K. He, L. Wang, X. Chen, X. Ma, Q.-K. Xue, Phys. Rev. Lett. 2012, 108, 016401.
[4] R. Venkatasubramanian, E. Sivola, T. Colpitts, B. O’Quinn, Nature 2001, 413, 597.
[5] M. Z. Hasan, C. L. Kane, Rev. Mod. Phys. 2010, 82, 3045.
[6] P. Roushan, J.seo, C. V. Parker, Y. S. Hor, D. Hsieh, D. Qian, A. Richardella, M. Z. Hasan, R. J. Cava, A. Yazdani, Nature 2009, 460, 1106.
[7] T. Zhang, P. Cheng, X. Chen, J.-F. Jia, X. Ma, K. He, L. Wang, H. Zhang, X. Dai, Z. Fang, X. Xie, Q.-K. Xue, Phys. Rev. Lett. 2009, 103, 266803.
[8] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, S.-C. Zhang, Science 2007, 318, 766.
[9] L. Fu, C. L. Kane, Phys. Rev. Lett. 2008, 100, 096407.
[10] X.-L. Qi, R. Li, J. Zhang, S.-C. Zhang, Science 2009, 323, 1184.
[11] I. Garate, M. Franz, Phys. Rev. Lett. 2010, 104, 146802.
[12] J. Wang, X. Chen, B. F. Zhu, S.-C. Zhang, Phys. Rev. B 2012, 85, 235131.
[13] M. I. Katsnelson, K. S. Novoselov, A. K. Geim, Nat. Phys. 2006, 2, 620.
[14] V. V. Cheianov, V. Fal’ko, B. L. Altshuler, Science 2007, 315, 1252.
[15] B. Seradjeh, J. E. Moore, M. Franz, Phys. Rev. Lett. 2009, 103, 066402.
[16] D. O. Scanlon, P. D. C. King, R. P. Singh, A. de la Torre, S. McKeown Walker, G. Balakrishnan, F. Baumberger, C. R. A. Catlow, Adv. Mater. 2012, 24, 2154.
[17] L.-L. Wang, M. Huang, S. Thimmaiah, A. Alam, S. L. Bud’ko, A. Kaminski, T. A. Lograsso, P. Canfield, D. D. Johnson, Phys. Rev. B 2013, 87, 125303.
[18] J.-L. Mi, M. Brenholm, M. Bianchi, K. Borup, S. Johnsen, M. Søndergaard, D. Guan, R. C. Hatch, P. Hofmann, B. B. Iversen, Adv. Mater. 2013, 25, 889.
[19] Z. Zeng, T. A. Morgan, D. Fan, C. Li, Y. Hirono, X. Hu, Y. Zhao, J. S. Lee, J. Wang, Z. M. Wang, S. Yu, M. E. Hawridge, M. Benamara, C. J. Salamo, AIP Adv. 2013, 3, 072112.
[20] Y. Zhao, C.-Z. Chang, Y. Ji, A. DaSilva, Y. Sun, H. Wang, X. Xing, Y. Wang, K. He, X. Ma, Q.-K. Xue, J. Wang, Sci. Rep. 2013, 3, 3060.
[21] Y. Zhao, H. Liu, X. Guo, Y. Jiang, Y. Sun, H. Wang, Y. Wang, H.-D. Li, M.-H. Xie, X.-C. Xie, J. Wang, Nano Lett. 2014, 14, 5244.
[22] G. Wang, X.-G. Zhu, Y.-Y. Sun, Y.-Y. Li, T. Zhang, J. Wen, X. Chen, K. He, L.-L. Wang, X.-C. Ma, J.-F. jia, S. B. Zhang, Adv. Mater. 2011, 23, 2929.
[23] S. Jia, H. Beidenkopf, I. Drozdov, M. K. Fuccillo, J. Seo, J. Xiong, N. P. Ong, A. Yazdani, R. J. Cava, Phys. Rev. B 2012, 86, 165119.
[24] K. A. Kokh, S. V. Makarenko, V. A. Golyashov, O. A. Shegai, O. E. Tereshchenko, CrystEngComm. 2014, 16, 581.
[25] Y. Jiang, Y. Sun, M. Chen, Y. Wang, Z. Li, C. Song, K. He, L. Wang, X. Chen, Q.-K. Xue, X. Ma, S. B. Zhang, Phys. Rev. Lett. 2012, 108, 066809.
[26] P. Sessi, M. M. Otrokov, T. Bathon, M. G. Vergniory, S. S. Tsirkin, K. A. Kokh, O. E. Tereshchenko, E. V. Chulkov, M. Bode, Phys. Rev. B 2013, 88, 161407.
[27] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, Pablo Ordejón, J. Phys.: Cond. Matter 2002, 14, 2745.
[28] J. P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. 1996, 77, 3865.
[29] G. Wang, T. Cagin, Phys. Rev. B 2007, 76, 075201.
[30] J. Tersoff, D. R. Hamann, Phys. Rev. B 1985, 31, 805.
[31] F. Donati, P. Sessi, S. Achilli, A. Li Bassi, M. Passoni, C. S. Casari, C. E. Bottani, A. Brambilla, A. Picone, M. Finazzi, L. Duò, M. I. Trione, F. CicacciPhys. Rev. B 2009, 79, 195430.