Role of surface structural motifs on the stability and reflectance anisotropy spectra of Sb-rich GaSb(001) reconstructions

Conor Hogan,[1,2,*] Rita Magri,[3] and Rodolfo Del Sole[2]

[1]*Consiglio Nazionale delle Ricerche, Istituto di Struttura della Materia, via Fosso del Cavaliere 100, I-00133 Rome, Italy*
[2]*European Theoretical Spectroscopy Facility (ETSF) and Dipartimento di Fisica, Università di Roma "Tor Vergata,"*
Via della Ricerca Scientifica 1, I-00133 Roma, Italy
[3]*CNR-Istituto di Nanoscienze and Dipartimento di Fisica, Università di Modena e Reggio Emilia, Via Campi 213/A, I-41100 Modena, Italy*

The structure of the technologically important—but still mostly unknown—GaSb(001)–$c(2 \times 6)$ surface reconstruction is investigated by means of *ab initio* simulations of reflectance anisotropy spectroscopy (RAS) and total energy calculations. A large number of reconstruction models for the GaSb(001) surface in the Sb-rich coverage regime are considered. The influence of each single surface structural motif on the RAS spectra is studied in detail, as well as their role in the surface stability with regard to application of the electron counting rule (ECR). We interpret the features of the RAS data measured for this reconstruction and suggest a new model for the $c(2 \times 6)$ phase. In this model a few Sb atoms in the second layer are randomly substituted by Ga, forming surface antisite defects. When used to fulfill the ECR, this "doping" effect considerably lowers the total energy of the long chain $c(2 \times 6)$ reconstruction model, making it competitive with the more stable short-chain (4×3) reconstructions. Formation of the surface antisites occurs spontaneously in the presence of dynamical negative charge fluctuations and is favored by the excellent matching between GaSb(001) and metallic Sb and by the natural softness of the Ga–Sb bonds. Calculations of the reflectance anisotropy spectra confirm that this structure is a major component of a largely disordered surface, where motifs of the stable (4×3) reconstructions are also present.

PACS number(s): 78.68.+m, 78.20.Bh, 71.55.Eq, 73.20.−r

I. INTRODUCTION

GaSb has generated considerable interest in recent years due to its proven potential for optoelectronic applications and as a substrate for growth of low power consumption devices working in the mid-infrared.[1,2] When combined with the lattice-matched InAs, GaSb exhibits unusual broken gap alignments[3] that have been exploited to reveal fundamentally interesting phenomena such as confinement driven metal/insulator transitions[3] and excitonic insulators.[4] Moreover, the quite large spin-orbit coupling in these narrow gap systems makes them promising for high-speed spintronic devices.[5,6] These new interesting applications demand a precise control over the atomic structure of the surface and interfacial regions. Nevertheless, much controversy still exists in the literature regarding the atomic structure of the GaSb(001) surface appearing under Sb-rich growth conditions, which is the situation encountered during typical molecular beam epitaxial (MBE) growth of devices based on 6.1 Å III–V materials.

Several Sb-rich phases of GaSb(001) have been observed. Reflection high-energy electron diffraction (RHEED) measurements[7,8] on MBE-grown samples report a change in phase from (1×3) to $c(2 \times 6)$ and finally to $(n \times 5)$ as the substrate temperature is lowered from 500°C to 300°C. Based on scanning tunneling microscopy (STM) studies of the (1×3) phase of GaSb(001) and AlSb(001), Barvosa-Carter *et al.*[9] proposed that the surface is in fact locally composed of (4×3) reconstructions. While AlSb(001) clearly exhibits ordered $\alpha(4 \times 3)$, $\beta(4 \times 3)$ and $\gamma(4 \times 3)$ structures (see Fig. 1) as the Sb content increases [all of which produce (1×3) RHEED patterns], the slightly more disordered GaSb(001)–(1×3)

phase appears to be a coexistence of $\alpha(4 \times 3)$ and $\beta(4 \times 3)$ reconstructions. Total energy calculations[10] confirmed that these (4×3) structures were stable. More recently, Romanyuk and coworkers[11] explained how a (1×3) RHEED pattern can result from (4×3) unit cells by assuming a one-dimensional disorder between unit cell rows.

In the very Sb-rich regimes—where InSb(001), AlSb(001), and III–As surfaces show distinct $c(4 \times 4)$ reconstructions—the GaSb(001) surface is characterized by reconstructions featuring long Sb–Sb dimer chains arranged along the $[\bar{1}10]$ direction, as observed by STM.[7,9] In particular, a $c(2 \times 10)$ phase was reported to form when cooling the sample at a very low temperature under an Sb flux. It has been noted, however, that structural models based on long chains of Sb–Sb dimers [including the $c(2 \times 10)$ and $c(2 \times 6)$ models: see Fig. 1(c)] do not satisfy the electron counting rule (ECR).[7,10,12] Such surfaces are heavily metallic, and as a consequence have a zero temperature energy that is much higher than that of the "interrupted chain" reconstructions, such as the $c(4 \times 4)$ and (4×3) reconstructions.[13] This discrepancy was recently addressed by Houze *et al.*[13] who proposed modifications of the basic $c(2 \times 10)$ structure that lower the metallicity. Rather than breaking the dimer chains, they substituted a number of Sb atoms in the *second* atomic layer by Ga atoms. The resulting total energy is lowered, and a better agreement with the original STM images is claimed.

The $c(2 \times 6)$ phase, shown by GaSb(001) in the intermediate Sb–rich region, is in contrast poorly understood. High-resolution core-level photoemission spectroscopy combined with RHEED[8] confirmed that it is a truly distinct phase from the (1×3), and also indicated that the surface is terminated by >1 layer of Sb, with only bulklike Ga present. STM
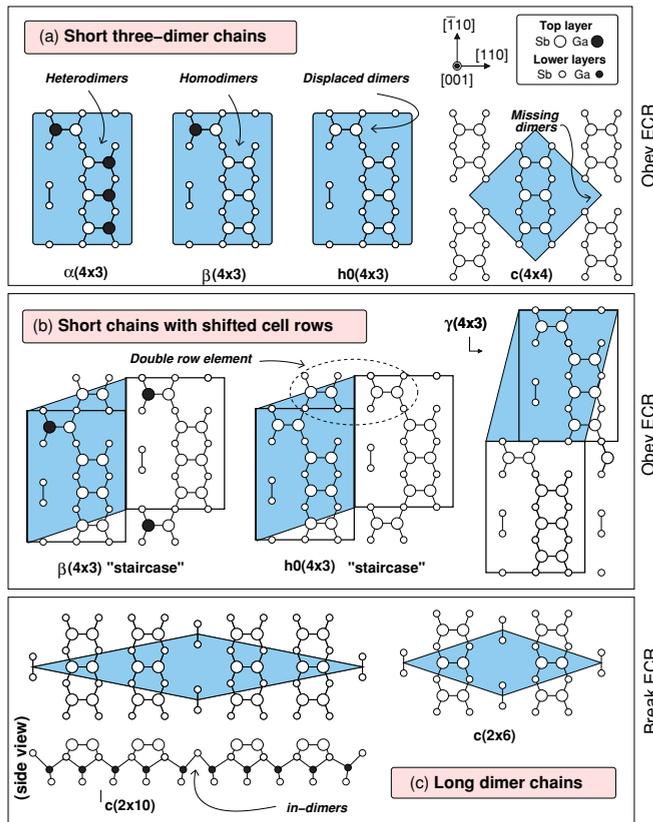
FIG. 1. (Color online) Schematic representation of GaSb(001) surface reconstruction models, grouped by dominant structural motif: (a) short dimer chains in a regular arrangement; (b) short chains in a shifted cell arrangement; (c) very long dimer chains. Other local structural motifs (e.g., in-dimers) are indicated directly. Primitive unit cells are denoted by shaded regions.

images suggest a structure based on long dimer chains[14,15] in spite of the fact that the standard $c(2 \times 6)$ model, shown in Fig. 1(c), breaks the ECR and yields relatively high total energies.[10] Although significant local disorder is observed in the STM data, low-energy electron diffraction (LEED) studies show that the long range order has a clear $c(2 \times 6)$ periodicity.[14] The recent RHEED studies[11] suggest, however, that the phase might be simply due to an aperiodic "staircase" arrangement of (4×3) unit cells, having a different pattern from the randomly placed (4×3) units constituting the (1×3) phase [part of which is represented by the models shown in Fig. 1(b)]. Moreover, various experimental reflectance anisotropy spectra have been reported in the literature[16,17] measured on surfaces presenting a clear $c(2 \times 6)$ LEED pattern. These measurements produced very similar optical signals that have never been carefully interpreted.

In this paper we present theoretical calculations of the reflectance anisotropy spectra for a number of GaSb(001) surface reconstructions in order to identify which reconstruction model is responsible for the observed optical spectra of the $c(2 \times 6)$ phase. Reflectance anisotropy spectroscopy (RAS), sometimes called reflectance difference spectroscopy (RDS), is a technique of choice for the study of surface structures since it is extremely sensitive to local structural motifs and thus can provide spectroscopic information even

in the presence of disorder. In contrast to STM it can also provide information about the bonding configurations *below* the surface layers. The RAS simulations are supported by total energy calculations performed to compare the stability of reconstructions based on different structural motifs. Based on the results of the comparison between theory and experiment we propose a structural model for the major contribution to the $c(2 \times 6)$ surface reconstruction, which is shown to yield a much better agreement with the experimental spectra. The structural model is based on the assumption that GaSb(001) stabilizes (and thus satisfies the ECR) through the random substitution of subsurface Sb atoms with Ga atoms. We have previously shown[18] that the substitution indeed occurs spontaneously on the surface in the presence of an excess of electronic charge, which can be induced thanks to the perfect match between the GaSb(001) surface lattice and the metal Sb lattice. The relative stability of this structure is confirmed by means of precise total energy calculations. We present arguments as to why the GaSb(001) surface follows such an unusual route to stability in contrast to other III–Vs. Furthermore, we investigate the influence of spin-orbit coupling on the surface optical properties within a first principles approach.

The paper is outlined as follows. In Sec. II we describe the theoretical scheme used to compute the surface structures and optical spectra, paying particular attention to the role of spin-orbit coupling and many body effects. Results of these calculations are presented and analysed in Sec. III. Section IV opens a discussion of the electron counting rule and its connection to surface structural motifs, leading to the proposal and computation in Sec. V of several new surface structures that incorporate defects. The stability of the proposed new models is discussed in Sec. VI. Finally, the driving force behind defect formation is examined in detail in Sec. VII before overall conclusions are drawn.

II. METHODOLOGY

A. Supercell approach

All calculations were carried out from first principles using density functional theory in the local density approximation (DFT-LDA),[19,20] within a plane-wave and norm-conserving pseudopotential framework. Structural relaxations and calculation of electronic occupied and unoccupied eigenstates were performed using the PWSCF code, part of the QUANTUM-ESPRESSO suite.[21] A kinetic energy cutoff of 15 Ry was chosen which yielded good agreement with structural parameters and surface optical spectra obtained at 26 Ry.

Surfaces were modelled using supercells containing thin (10–12 atomic layers) slabs separated by a vacuum region about 12Å thick. During the structural relaxations the bottom bilayer, terminated with pseudohydrogens, was kept fixed to the theoretical bulk positions (the GaSb lattice constant at 15 Ry is 6.026Å) and the remaining atoms relaxed until forces were less than 25 meV/Å. A **k**-point sampling equivalent to 144 points in the (1×1) surface Brillouin zone was used. Note that these values refer to the slabs used to compute the optical spectra: total energy calculations require a different kind of precision and are described in more detail in Sec. VI.
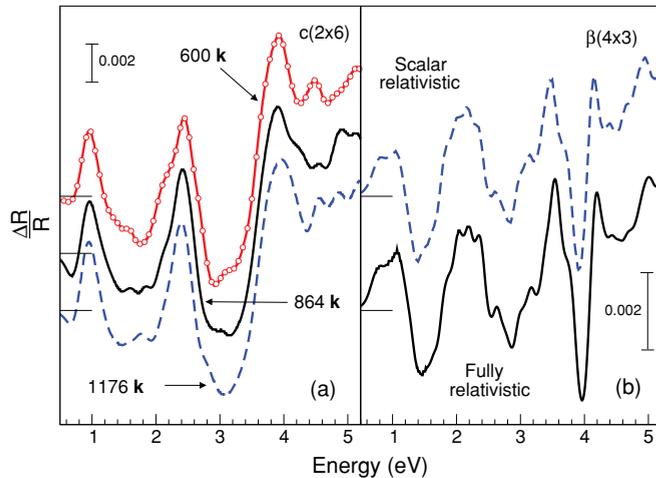
FIG. 2. (Color online) (a) Convergence of RAS spectra of the $c(2 \times 6)$ reconstruction with respect to the number of **k**-points in the (1×1) surface Brillouin zone. (b) Dependence of RAS spectra on the choice of pseudopotential, shown for the $\beta(4 \times 3)$ surface. Spin-orbit coupling is neglected (intentionally) in the fully relativistic calculation.

B. Surface optical properties

The RAS signal, $\Delta R / R$, is defined by

$$\frac{\Delta R}{R} = \frac{\Delta R_{[\bar{1}10]} - \Delta R_{[110]}}{R}, \qquad (1)$$

where $\Delta R_i / R$ ($i = x, y$) is the correction to the Fresnel reflectivity R along the i direction. Within the repeated slab/supercell scheme $\Delta R_i / R$ can be obtained from

$$\Delta R_i / R = \frac{4\omega}{c} \mathrm{Im} \left\{ \frac{4\pi \alpha_{ii}^{\mathrm{hs}}}{\epsilon_b - 1} \right\}, \qquad (2)$$

where $\alpha_{ii}^{\mathrm{hs}}$ is the diagonal tensor component of the averaged half-slab polarizability, having dimension of length. This expression is a finite-slab approximation to the formula derived by Del Sole and coworkers[22,23] for s polarized light incident on a semi-infinite crystal and is equivalent to that obtained using the three-layer model of McIntyre and Aspnes.[24] In practical calculations the half-slab polarizability is extracted from the H-terminated slab using a real-space cutoff technique.[25] Slab polarizabilities were computed at the independent-particle (so-called "RPA") level using the YAMBO code.[26] High densities of **k**-points [of the order of 864–1152 points per (1×1) cell] were used in the integration over the surface Brillouin zone. This choice yields spectra converged to better than 0.05 eV in the main peak positions, as demonstrated for the $c(2 \times 6)$ reconstruction in Fig. 2(a).

C. Pseudopotentials and spin-orbit coupling

Generally speaking, spin-orbit coupling (SOC) plays a negligible role in determining the atomic structure and phase diagram of typical semiconductor surfaces. However, if the constituent elements are heavy enough, the optical response can be substantially influenced. This is indeed the case for GaSb, whose split-off energy, $\Delta_{\mathrm{SO}} = 0.76$ eV, is in fact comparable to the direct gap of 0.812 eV.[27] SOC can modify the optical response through a number of ways, including[28]

(i) lifting of degeneracies of energy levels, (ii) modification of transition matrix elements through spin-flipping, and (iii) hybridization of states of different spin character. For instance, point (i) is manifested through the appearance of two sharp peaks at E_1 and $E_1 + \Delta_{\mathrm{SO}}$ in the absorption spectrum of bulk GaSb. Up to now, few *ab initio* calculations have been carried out that illustrate the influence of SOC on RAS spectra.[29] In the case of III-V surfaces, tight-binding calculations by Vázquez-Nava and co-workers[30] on InSb(110)-(1×1) suggested that SOC leads not only to red-shifts of up to 0.2 eV in the RAS peaks, but also to qualitative changes in the lineshape around and below the E_2 critical point energy (3–4 eV). As the value of Δ_{SO} in InSb (~ 0.8 eV) is comparable to that of GaSb, we might expect important effects due to spin orbit coupling in the present *ab initio* calculations.

As SOC is strongest in the core, the effect can be most easily incorporated within a pseudopotential framework through the use of fully relativistic (J-dependent) pseudopotentials. It is sufficient to compute the ground state density and structural relaxation without SOC, and then to include it *a posteriori* in the band structure calculations as a perturbation.[31] Fully relativistic pseudopotentials were hence developed for both Ga and Sb using the atomic code of Paolo Giannozzi.[32] Nonlinear core corrections (NLCC) were included.[33] To check their accuracy, we compared RAS spectra computed using the new pseudopotentials (with spin-orbit coupling switched off) with spectra computed using previously available scalar relativistic ones.[34] The results, shown in Fig. 2(b) for the $\beta(4 \times 3)$ reconstruction, reveal only negligible differences between the two calculations, and therefore we judge the new pseudopotentials to be sufficiently reliable to be used in all calculations of the optical spectra throughout this work.

The next step is to consider the effects of the spin-orbit coupling on the RAS. In Fig. 3 we present fully *ab initio* calculations of the RAS spectra for three kinds of GaSb(001) reconstruction with and without spin-orbit coupling, again computed using YAMBO. The overall effects are minor, but are not insignificant: additional broadening occurs over the full spectral range, and in some cases the onset of optical transitions is lowered. Nevertheless the effect is much less than that predicted by the tight binding calculations.[30] As the resulting wave functions are spinorial [point (iii) above], and the number of bands doubled [point (i)], the computation is quite demanding. Hence SOC is included only in selected calculations hereon.

D. Many-body effects

Incorporation of many-body effects, i.e., self-energy, excitonic, and local field effects, into the calculation of optical spectra is a demanding task for surfaces due to the sheer computational size (for a concise review of this issue, see Ref. 35). A small number of calculations have nevertheless been carried out including all of these effects on C(100)–(2×1)[36] and various simple silicon surfaces,[37–39] while quasiparticle corrections have been investigated on wider range of surfaces, including GaAs(110)[40] and GaP(001)–(2×4).[41] As it is not practical to thoroughly account for many-body effects in all the GaSb(001) reconstruction models studied in
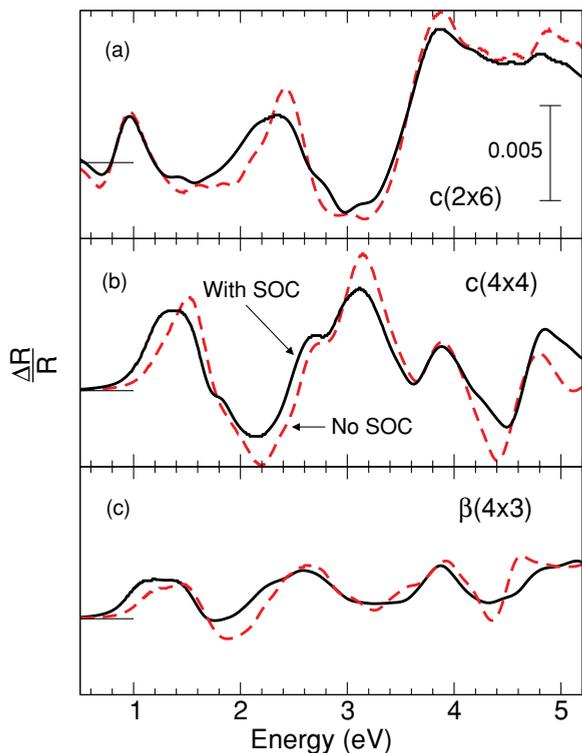
FIG. 3. (Color online) Spin orbit coupling (SOC) effects on the RAS, shown for (a) the long chain $c(2 \times 6)$ model, (b) the missing dimer $c(4 \times 4)$ model, and (c) the displaced dimer $\beta(4 \times 3)$ model. SOC effects are (not) included in the (dashed) solid curves.

FIG. 4. (Color online) Many-body effects in the surface optical spectra of GaSb(001)-$\beta(4 \times 3)$. (a) Converged RAS spectrum, with contribution of back surface of slab removed, within RPA + scissors scheme (see text). (b) Imaginary part of surface dielectric anisotropy of full slab calculated using a reduced number of **k**-points, within the RPA + scissors approach. (c) As (b), but including self-energy corrections within the GW approximation (dashed curve), and by further solution of the Bethe-Salpeter equation (solid curve).

this work, we include these effects in an approximate manner throughout this paper by applying a "scissors shift" correction of +0.4 eV to the conduction bands following the scheme proposed by Del Sole and Girlanda.[42] This "RPA + scissors" approach has been demonstrated to yield reasonable agreement with experimental spectra for other III–V (001) surfaces (see, e.g., Ref. [43]). Before continuing, however, we justify our use of this approach by means of an explicit calculation of the optical spectra including many body effects on one reconstruction [the $\beta(4 \times 3)$ model].

Self-energy corrections were computed within the so-called GW approximation, while excitonic and local field effects were accounted for by means of solving the Bethe-Salpeter equation (BSE).[44] For the GW part of the calculation, we calculated the screening matrix $\varepsilon_{\mathbf{G},\mathbf{G}'}^{-1}(\mathbf{q},\omega)$ using a plasmon pole model, using 501 **G** vectors to construct the matrix, 750 bands (corresponding to states lying 12.5 eV above the Fermi level), and 20 **q** points in the irreducible part of the Brillouin zone (IBZ). The correlation part of the self-energy Σ_c was calculated using 501 plane waves and 1000 bands, while 40000 plane waves were used in computing the exchange part Σ_x. The Bethe-Salpeter equation was solved using the Haydock recursion method.[45,46]

To keep the excitonic Hamiltonian at a manageable size, we only considered single-particle transitions with an energy $\leqslant 3.5$ eV. As this corresponds to the energetic range where transitions involving states localized at the front surface are most dominant, it also allows us to (i) use less **k**-points [12 points in the IBZ, see Fig. 4(a)] than that used in the fully

converged calculations [dashed curve in Fig. 3(c)]; (ii) avoid the difficult task of subtracting the (many-body) response of the back layers; and (iii) focus only on the anisotropy of the imaginary part of the dielectric function ($\Delta \varepsilon = \varepsilon_{[\bar{1}10]}'' - \varepsilon_{[110]}''$) rather than the more complicated RAS signal. The similarity between $\Delta \varepsilon$ and the overall RAS signal [panels (b) and (a) in Fig. 4, respectively] shows that this approach is quite sufficient for demonstrating the general influence of the many body effects on the RAS for this system in this energy range.

Our computed spectrum including self-energy corrections to the single particle states is shown as the dashed curve marked GW in Fig. 4(c). The quasiparticle opening of the DFT-LDA band gap (HOMO-LUMO energy difference) was computed to range between 0.57 (near Γ) and 0.65 eV within the IBZ. This is still less than the correction of 0.70 eV reported elsewhere for the direct gap of bulk GaSb,[47] which is to be expected as surface and bulk states experience different self-energy shifts. After subsequent solution of the Bethe-Salpeter equation we obtain the optical anisotropy shown in Fig. 4(c) (solid line, GW+BSE). Inclusion of excitonic and local field effects is seen to cause a red-shift to the GW spectra and small modifi-

cations of the peak intensities. Nevertheless, it is clear that the overall effect of the many-body corrections is to cause a rigid shift to the independent particle (RPA) spectrum, and therefore our "RPA+scissors" approach appears to be reasonable. The value assumed, +0.4 eV, is slightly higher than the calculated one as it represents an an average correction for all transitions, including those involving bulklike states lying at higher energy.

III. SURFACE OPTICAL SPECTRA

In this section we present theoretical simulations of the RAS spectra and use them to interpret features of the experimental data. Several RAS measurements on GaSb(001) have been reported in the literature and are reproduced in Fig. 5(a). These experiments were carried out on different samples prepared under different conditions and growth methods (namely MBE and MOCVD), and after preparation of the deposited layer, different techniques were then employed to avoid surface contamination.[16,17,48–50] Nevertheless, and in spite of some uncertainty in the baseline position, it is clear that the spectra are all very similar. Hence, the resulting phase is likely to be the same Sb-stabilized phase. It should be quite stable and typical of the GaSb(001) surface over a wide range of temperatures and atmospheric conditions. Spectra A' and D were explicitly reported to have a $c(2 \times 6)$ LEED pattern. The experimental spectra are characterized by a low energy positive peak (identified only in the infrared-MOCVD setup, however— spectrum A), a broad negative structure between 2 and 3 eV (showing oscillating features caused by the linear electro-optic effect[49]), and a large peak around 3.8 eV close to the E_2 critical point energy, followed immediately by a trough at about 4.4 eV.

A. Analysis of computed spectra

Computed RAS spectra are reported in Fig. 5(b) and (c). Panel (b) shows spectra for the short-chain (4×3) models

depicted in Fig. 1(a) as well as their shifted-cell counterparts shown in Fig. 1(b). At first glance, the six spectra appear to be very similar, featuring numerous peaks and troughs in a weak oscillation about the zero baseline. Panel (c) instead shows calculated spectra for the $c(4 \times 4)$ reconstruction and the long chain models of Fig. 1(c). These three spectra clearly possess a much greater peak–peak amplitude. By reporting the various data in this way, it is possible to identify and analyze trends in the computed spectra as the surface structure changes, as many of the spectra show common features. In the following, we relate the dominant features in the computed RAS spectra to the specific structural motifs which are understood to occur on GaSb(001).

First of all, every RAS spectrum exhibits a *positive* feature in the range 1.0–1.5 eV. As this energy range lies well below the computed bulk E_1 critical point energy of 2.1 eV, the feature is likely a manifestation of transitions between surface states located along the dimer *rows*. For the $\gamma(4 \times 3)$, $c(4 \times 4)$, and long chain reconstructions, a single peak is observed; in all other cases, two overlapping peaks give rise to a broader structure. In fact, the number of constituent peaks appears to be determined by the symmetry of the surface layer: a single peak is seen when both sides of the dimer chain are equivalent; conversely, a double peak appears when this local symmetry is broken by the displaced dimer. This observation is reflected in a more detailed analysis of the states involved in the optical transitions. Characteristic states are plotted in Fig. 6 for some reference structures. In the case of a single peak, a single symmetric surface state spread over the whole dimer row is active, as illustrated for the $c(2 \times 6)$ model (C3 state). Double peaks arise when two energetically distinct surface states localized on either side of the dimer row give rise to the double peaked structure [see the C1 and C2 states shown for $h0(4 \times 3)$ in Fig. 6].

Spectral features and electronic states are related in the following manner. The principal *occupied* state (V1) in the
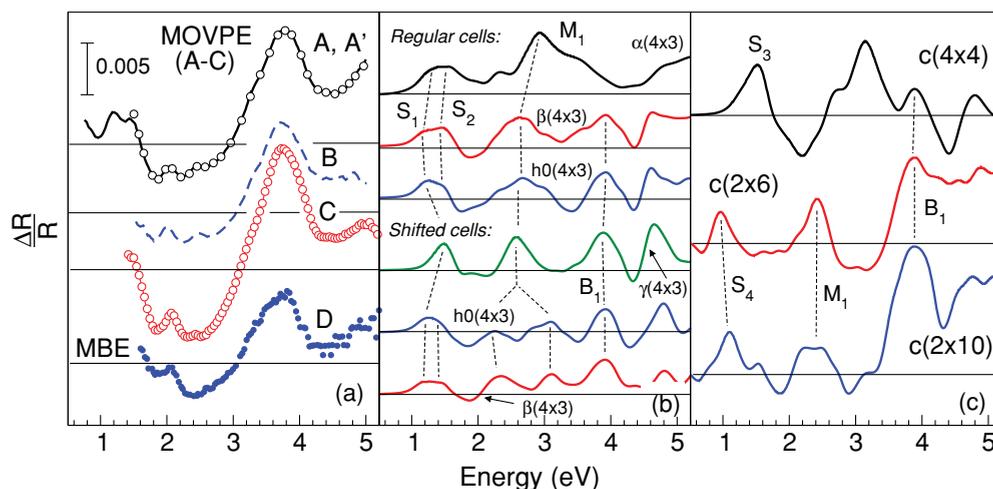
FIG. 5. (Color online) Reflectance anisotropy spectra for Sb-rich GaSb(001). (a) Experimental spectra measured on $c(2 \times 6)$ phases prepared using MOVPE (A–solid line: Moller *et al.*;[17] A'–open circles: Kollonitsch *et al.*;[48] B: Pitts *et al.*;[49] C: Weeke[50]) and MBE (D: Goletti *et al.*[16]) techniques. Data have been renormalized according to $\Delta R/R = 2\text{Re}\{\Delta \tilde{r}/\tilde{r}\}$. (b) Theoretical spectra for the short-chain surface reconstructions depicted in Fig. 1(a) and (b). (c) Theoretical spectra for the $c(4 \times 4)$ phase and the long-chain surface reconstructions depicted in Fig. 1(c). The vertical (linear) scale is equivalent in each panel. Spin-orbit effects are *not* included in the computed spectra. A Lorentzian broadening of 0.15 eV was used throughout.
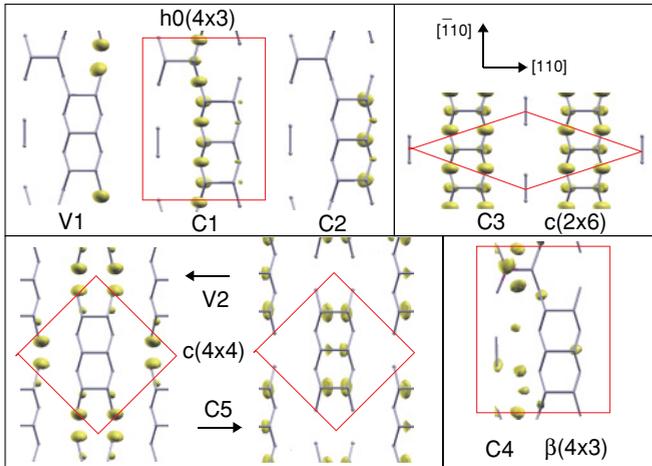
FIG. 6. (Color online) Isosurface plots of $|\psi|^2$ for typical valence (V) and conduction (C) states involved in low energy optical transitions (see text), superimposed on wire-frame models of the topmost two layers for several reconstructions of GaSb(001).

interrupted chain models consists of a lone-pair (or pairs in the case of $c(4 \times 4)$: V2) at the threefold-coordinated Sb atom occurring in the second layer at the end of each three-dimer block. This state contributes to the low energy peaks indicated S_1–S_3 in Fig. 5. No such state occurs in the long chain $c(2 \times 6)$ and $c(2 \times 10)$ models, and instead a higher energy occupied bulk or resonant state (not shown) is involved in the transition, leading to a lower energy peak (S_4). The principal *unoccupied* state is found to be a chainlike state delocalized along the ad-dimer backbonds in the $[\bar{1}10]$ direction, on one side [α, β, and $h0(4 \times 3)$ variants: C1] or on both sides (long chain models: C3) of the dimer row. This state is active in the S_1 and S_4 peaks. In the case of the $c(4 \times 4)$ reconstruction, where chainlike states are forbidden, the relevant state (C5) is a relatively-high dimer antibonding state and contributes to the formation of S_3. The remaining S_2 peaks arise from transitions to backbond chain states (C2) on the *other* side of the dimer row (i.e., with respect to the C1 states contributing to S_1); or transitions to localized Ga orbitals on displaced dimers, such as the C4 state on the $\beta(4 \times 3)$ surface.

In brief, the low-energy feature is a strong fingerprint of the top layer geometry. A double-peaked structure around 1.3 eV points to asymmetric, interrupted dimer chains; a single peak at 1.5 eV points to the missing dimer $c(4 \times 4)$ reconstruction, while a single peak at lower energy (1.0 eV) indicates the strongly metallic long chain $c(2 \times 6)$ and $c(2 \times 10)$ models. These features are generally only detectable with an infrared-RAS experimental setup, such as that used in Ref. 17.

With the exception of $\alpha(4 \times 3)$, all spectra show a positive peak at about 3.9 eV followed immediately by a negative peak or trough at 4.4 eV. This feature is reminiscent of similar ones seen near E_2 for various reconstructions of GaAs(001),[51] and its energy position coupled with its insensitivity to the surface structure suggests that it derives from bulklike states terminating at the surface that are weakly (albeit anisotropically) perturbed by the local surface potential. Since $\alpha(4 \times 3)$ is the structure most chemically disparate from the others (having 4 Ga–Sb heterodimers/surface cell), it is not surprising that the induced anisotropic strain causes a quite different pertur-

bation on the terminating bulk states than the other (4×3) models. The long chain $c(2 \times 6)$ and $c(2 \times 10)$ structures, on the other hand, exhibit the greatest long-range geometrical anisotropy. The resulting surface potential produces a strong average anisotropic perturbation on the bulk states in the surface region, and since there are many transitions due to such states falling around 3.9 eV (close to the E_0' and E_2 bulk critical points) a strong peak is formed. Hence, the strength of this peak can be interpreted as an indicator of the overall surface anisotropy.

Finally, a positive peak is seen around 2.5 eV that is common to the regular $h0$, β, and $\gamma(4 \times 3)$ structures, as well as the $c(2 \times 6)$ and $c(2 \times 10)$ reconstructions. This peak apparently splits when the cell rows are shifted in the $[\bar{1}10]$ direction ($\beta(4 \times 3)$ and $h0(4 \times 3)$ staircase structures). Nevertheless, shifting the cell rows in either direction does not lead to a substantial modification of the optical spectrum. This is consistent with the idea that RAS is a *local* probe of the atomic structure, and is not sensitive to changes in the structure on the order of the cell row, unless new motifs are created. This is particularly true for the $\gamma(4 \times 3)$ case, whereas short double-row elements are formed in the "staircase" structures, as indicated in Fig. 1(b), that give rise to the observed spectral changes. Detailed analysis reveals that the 2.5 eV ($M1$) peak is of mixed surface-bulk character, with some involvement of the in-dimer related surface states.

B. Interpretation of experiment

We now attempt to interpret the experimentally measured data on the $c(2 \times 6)$ surface based on the results and analysis presented in the previous section. Our main observation is that no single calculated spectrum is in agreement with any of the experimental data shown in Fig. 5(a). In addition, no linear combination of theoretical spectra yields a qualitative improvement over the full energy range. Hence we conclude that none of the known surface structures, *including the most stable* (4×3) *ones*, can give rise to the measured RAS spectrum of the $c(2 \times 6)$ surface of GaSb(001).

There are, however, a number of calculated spectral features that appear to correspond well with features at similar energies in the experimental data. As mentioned earlier, RAS is a probe of the local atomic structure, and therefore is sensitive to the presence of distinct surface structural motifs. In the previous section, we related several of the computed spectral features to several such motifs. Since the measured RAS signal is averaged over a spot size of 1 mm^2, the experimental data contain contributions from all such motifs that are present, even in the presence of local disorder. Two features in particular stand out. Firstly, the S_1–S_2 broad peak found in all (but one) of the (4×3) spectra is in good agreement with the low energy double-peaked structure of the MOCVD experimental data. This points strongly to the presence of interruptions in the dimer chains due to displaced dimers, as discussed in Sec. III A, or even missing dimers, as found in the $c(4 \times 4)$ reconstruction [spin-orbit coupling is important in this case: see Fig. 3(b)]. Secondly, although all spectra show a peak ($B1$) at about 3.8 eV followed by a trough, the only spectra in good agreement with the experimental lineshape and magnitude are those of the metallic $c(2 \times 6)$ and $c(2 \times 10)$ structures.

There remains, unfortunately, the clear problem of the positive feature M_1 of mixed origin, which is in strong contradiction with the negative features seen in the experiment. The $c(4 \times 4)$ spectrum is an exception, but it is anyway dominated by a strong positive peak at 3.1 eV that is also not experimentally observed. As seen in Fig. 3, the intensity of M_1 is reduced by spin-orbit coupling and appears to split in the case of a staircase alignment of (4×3) cell rows. Nevertheless, a satisfactory agreement with the experiment is still lacking. We deduce at this point that another structural motif, not present in Fig. 1, must be present that cancels out this positive feature—or better, leads to the negative one seen in the experiment. In order to deduce what this motif might be, we turn now to the surface energetics and consider what new motifs can be incorporated into our structures while maintaining, or improving, the stability.

IV. STRUCTURAL MOTIFS AND THE ECR

When considering possible structural motifs, it is desirable to examine their viability in the context of the electron counting rule (ECR) which is useful for predicting the overall structural stability.[52] First of all, the III–V surfaces tend to stabilize by lowering the total number of dangling bonds. Unpaired dangling bonds on the unreconstructed surface thus tend to pair up and form dimers, which constitute the fundamental structural motif on (001) surfaces. Secondly, most stable structures are required to fulfill the ECR that states that the lowest energy reconstructions are those having empty dangling bonds on the cations (III), and filled dangling bonds on the anions (V). The rule follows from the observation that the stablest structures are semiconducting with the valence states (the highest being localized on the anion dangling bonds) occupied and the conduction states (the lowest being localized on the cation dangling bonds) empty. This condition can be satisfied with particular arrangements of the dimers on the surface since infinite chains of dimers do not satisfy the ECR and lead to a metallic surface. Thus, the chains need to be interrupted or stabilized in some way.

For instance, under extremely V-rich conditions the $c(4 \times 4)$ reconstruction is observed [the $c(4 \times 4)$ is in fact found in all III–V(001) surfaces except GaSb(001)] in which the chains are interrupted through the formation of one missing dimer for every three dimers, with the groups of three dimers arranged as shown in Fig. 1(a). Similarly, the (4×3) structures satisfy the ECR by displacing one dimer in the chain for every three dimers and forming *in-dimers*. The $(n \times 3)$ reconstructions having in-dimers are indeed only observed on antimonide compounds (or in indium-rich $In_xGa_{1-x}As$ alloys grown on GaAs[53]) due to the particular softness of the III–V bonds in these compounds, which facilitate their formation.

As noted above, the ECR-satisfying reconstructions are predicted to be semiconducting whereas the ECR-violating ones are predicted to be metallic. To verify this, computed DFT-LDA band structures (without spin-orbit coupling) are shown in Fig. 7 for several of the GaSb(001) reconstructions considered in this work. All structural models based on short dimer chains (including the typical $\beta(4 \times 3)$ reconstruction) are indeed semiconducting. Long chain structures, on the other hand, are found to be metallic: the Fermi level of

the $c(2 \times 6)$ and $c(2 \times 10)$ models lies near the bottom of the conduction band, cutting through 2–4 low lying bands derived from surface states. Similar observations have been discussed elsewhere.[12,54]

Thus, all these motifs (missing dimers, displaced dimers, and in-dimers) help to satisfy the ECR. The connection between structural unit and ECR is quantified by the surface excess charge v, that is, the number of electrons left over when all available bonds and anion dangling bonds are filled. These excess electrons end up filling the *cation* dangling bond states that lie energetically above the valence band states, rendering infinite chain structures metallic and locating the Fermi level within the conduction band (see Fig. 7). The ECR is fulfilled and the surface is semiconducting (and hence stable) when $v = 0$, with the Fermi level lying within the gap. For V-rich III–V surface reconstructions v is given by

$$v = 2.5N_s - 6N_I - 4N_D - 8N_{MD} - 12N_{DD}. \quad (3)$$

In this expression N_s is the number of atomic sites within the last complete surface layer per unit cell (here denominated simply as subsurface), N_I is the number of in-dimers, N_D is the number of top-layer dimers, N_{MD} is the number of missing dimers, and N_{DD} is the number of displaced dimers. Increasing v to higher positive values moves the Fermi level upward into the conduction band [see Fig. 7(c)]. By changing the number of the different structural motifs in Eq. (3) we are able to change v and consequently shift the Fermi level up or down.

In this paper we also consider a new structural motif, henceforth called the surface Ga antisite defect. In contrast to bulk Ga antisites, these surface Ga antisites form two Ga–Ga bonds and two Ga–Sb bonds, and consist of replacing an Sb atom in the *second* atomic layer with a Ga atom. Such defects were proposed by Houze *et al.*[13] as a way to reduce the metallicity of the long dimer chains occurring in the $c(2 \times 10)$ surface. The most stable structure identified in that work, $c(2 \times 10)$–$s1a$, is shown in Fig. 8(a). Considering these defects, v becomes

$$v = 2.5N_s - 2N_{Ga} - 6N_I - 4N_D - 8N_{MD} - 12N_{DD}, \quad (4)$$

where N_{Ga} is the number of Ga surface antisites.

In the Sb-rich regime, where $N_{DD} = 0$, we can directly relate the number of subsurface sites to the various structural motifs by the additional condition

$$2N_I + 2N_D + 2N_{MD} = N_s, \quad (5)$$

where it is assumed that, according to the experimental observations, there are no subsurface sites having two dangling bonds. Combining Eqs. (4) and (5) yields

$$v = 0.5N_s - 2N_I - 4N_{MD} - 2N_{Ga}. \quad (6)$$

Hence we can consider two distinct and exclusive strategies that lead to fulfillment of the ECR in the Sb-rich regime: (i) formation of missing dimers, and (ii) formation of surface Ga antisites. Solving Eq. (6) for $v = 0$ in each case we find

$$(i) \quad N_{MD} = (N_s - 4N_I)/8, \qquad if \ N_{Ga} = 0; \quad (7)$$

$$(ii) \quad N_{Ga} = (N_s - 4N_I)/4, \qquad if \ N_{MD} = 0. \quad (8)$$

In the case of the $c(2 \times 6)$ surface ($N_s = 6, N_I = 1$), the conditions for ECR fulfillment are thus (i) $N_{MD} = 1/4$,
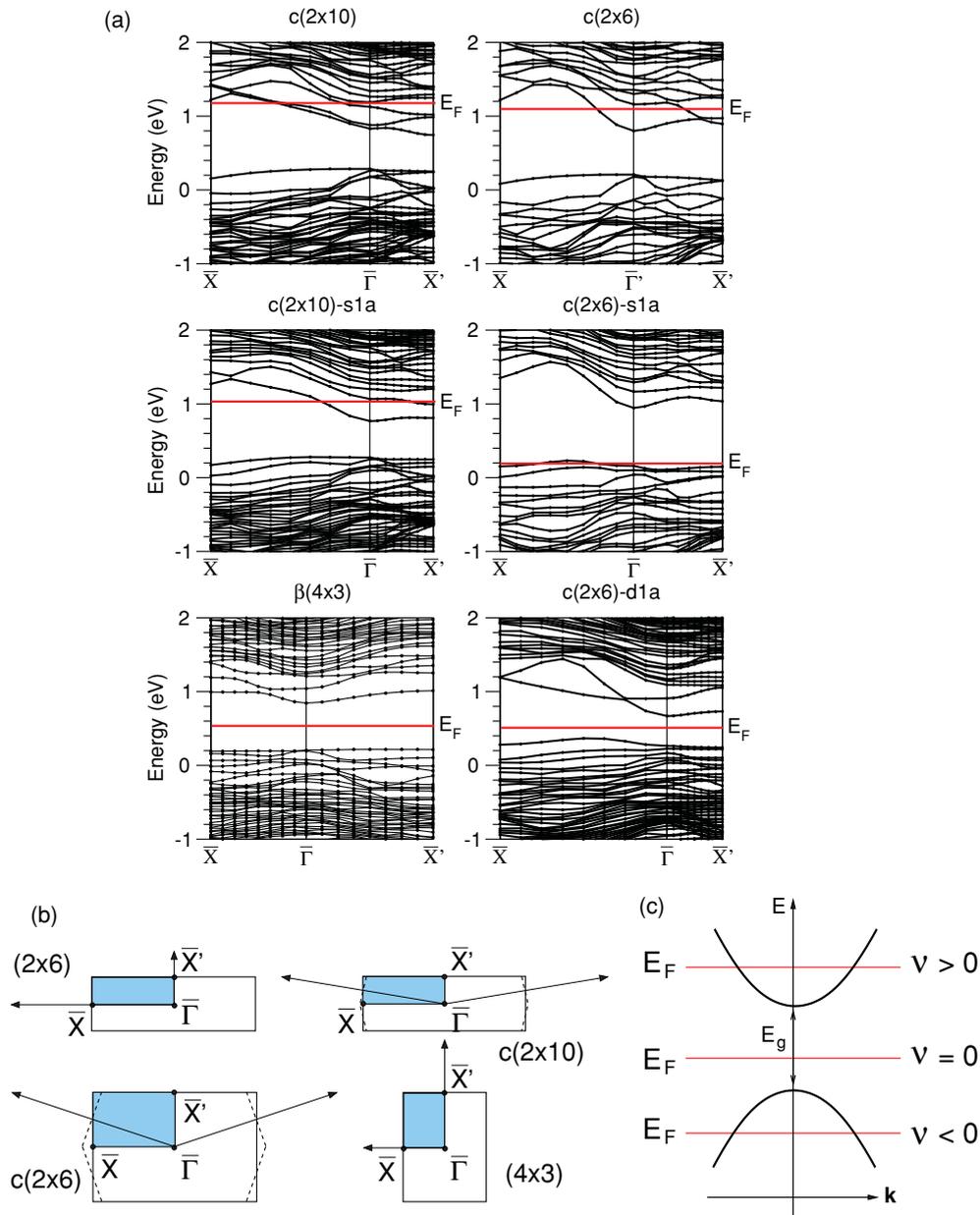
FIG. 7. (Color online) (a) DFT-LDA band structures of various GaSb(001) reconstructions, both defect-free and containing antisite defects. Fermi levels are indicated by thick horizontal lines (for semiconducting surfaces the Fermi level is conventionally located in the middle of the band gap). (b) Surface Brillouin zones (SBZ) corresponding to (a): note that the $c(2 \times 6)$–$d1a$ SBZ is simply that of the conventional (2×6) cell. (c) Relationship between the Fermi level position and the excess electron count ν with respect to the gap E_g.

that is, one dimer has to be removed from every four $c(2 \times 6)$ unit cells, or (ii) $N_{Ga} = 1/2$, i.e., one surface Ga antisite has to be introduced for every two unit cells. In a similar way, the corresponding condition for the $c(2 \times 10)$ phase ($N_s = 10, N_I = 1$) is $N_{Ga} = 3/2$, i.e., three defects for every two $c(2 \times 10)$ unit cells. This analysis explains why the structures containing one or more defects per unit cell suggested in Ref. 13 are still relatively unstable, and indeed, metallic.

In order to refer to the various different cell/defect concentration possibilities in a concise manner, we extend the naming convention of Ref. 13 such that the form "$c(2 \times 6)$–$s1a$" means an underlying $c(2 \times 6)$ lattice containing 1 substitutional defect per single (s) primitive $c(2 \times 6)$ cell area in configuration a; "$c(2 \times 10)$–$d3c$" would refer to a $c(2 \times 10)$ lattice having 3 defects per double (d) primitive $c(2 \times 10)$ cell area in configuration c, and so on. Configurations a, b, etc., indicate different inequivalent ways that subsurface Sb atoms are substituted by Ga defect atoms.

V. OPTICAL SPECTRA OF ANTISITE RECONSTRUCTIONS

We now consider the influence of these surface Ga antisite defects on the optical spectra by considering a number of defect-stabilized $c(2 \times 6)$ and $c(2 \times 10)$ surface structures.
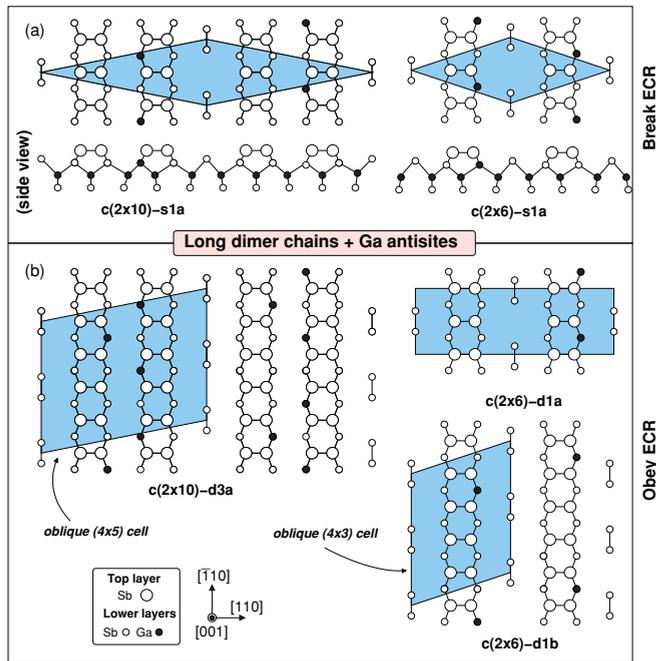
FIG. 8. (Color online) Schematic models of the $c(2 \times 6)$ and $c(2 \times 10)$ structures incorporating surface Ga antisite defects. (a) Metallic structures comprising one defect per primitive cell. (b) Possible structures containing correct density of defects needed to fulfill the ECR in each case.

Figure 8(a) shows models having one defect per primitive cell ($N_{Ga} = 1$): the aforementioned $c(2 \times 10)-s1a$ structure, and its $c(2 \times 6)$ counterpart, the $c(2 \times 6)-s1a$ structure (all other possible configurations of the latter are equivalent). As shown in Fig. 7, these models are indeed calculated to be metallic. In the $c(2 \times 6)-s1a$ case ($\nu = -1$), the Fermi level is shifted into the valence band due to a too large concentration of defects, whereas it remains in the conduction band in the $c(2 \times 10)-s1a$ case ($\nu = +1$), due to a too low defect concentration.

Fig. 8(b), on the other hand, shows structures which fulfill the ECR and are hence semiconducting. For the $c(2 \times 6)$ case with $N_{Ga} = 1/2$, two possibilities are shown having defects confined to every other dimer row [$c(2 \times 6)-d1a$] or more uniformly distributed [$c(2 \times 6)-d1b$]. Due to the large number of possibilities available in the $c(2 \times 10)$ case, we consider just one configuration for the purpose of comparison with $c(2 \times 6)$: the $c(2 \times 10)-d3a$ configuration is based on the energetic trends with respect to defect site position presented in Ref. 13, and should be one of the most stable possibilities.

Reflectance anisotropy spectra, with spin-orbit coupling included, are presented for these $c(2 \times 6)$ structures in Fig. 9(c) and compared again with experiment [panel (a)]. A corresponding calculation is shown in panel (b) for the $c(2 \times 10)-s1a$ structure in order to illustrate the influence of defect incorporation on the RAS spectra in a wider context. The spectra are found to be quite sensitive to the inclusion of defects and hence a number of conclusions can be drawn. First of all, the peak at 1.0 eV is particularly sensitive both to the density and relative position of the defects. Since this peak derives from transitions involving long chains of dimer backbond states (see Sec. III A), it is natural that the peak is strongly perturbed when the chemical composition of the backbonds changes from Sb–Sb to Sb–Ga. Secondly, the positive feature around 2.4 eV is quenched by the defect, or in some cases shifted below the x axis. This shift is most strong for the $c(2 \times 6)-s1a$ structure where the defect concentration is highest. Thirdly, the peak and trough above 4 eV remain fairly constant, and the overall peak-to-peak amplitude is maintained throughout.

Those structures which obey the ECR appear to offer the best agreement with experiment for a single model structure. The overall good agreement above 4 eV is maintained, while the peak appearing at 2.5 eV in the $c(2 \times 6)$ spectrum is quenched or made negative when defects are incorporated. Based on these results we infer that the experimental $c(2 \times 6)$ phase is composed of long-chain $c(2 \times 6)$ structures frequently interrupted by displaced or missing dimers (thus explaining the
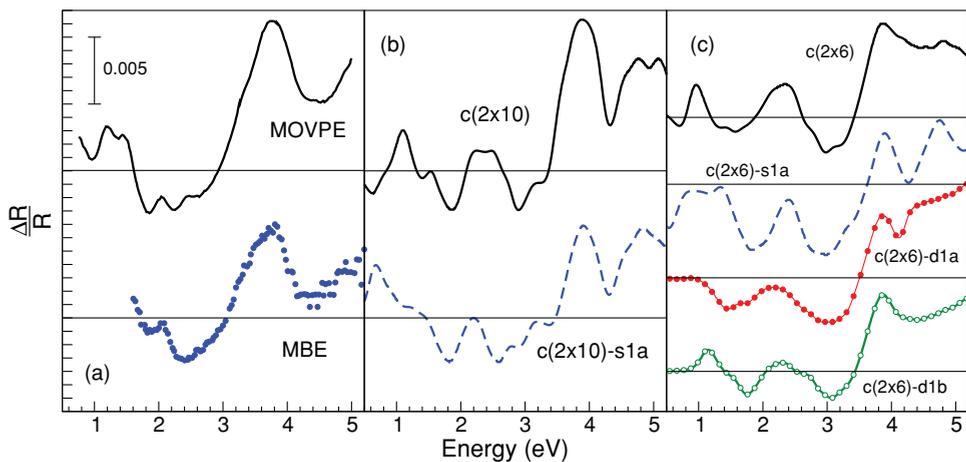
FIG. 9. (Color online) RAS spectra of GaSb(001). (a) Experimental data for typical MOVPE[17] and MBE-grown[16] surfaces showing $c(2 \times 6)$ LEED patterns. (b) Computed spectra of $c(2 \times 10)$ reconstruction, with and without one antisite defect per primitive cell (see Fig. 8). (c) Computed spectra of $c(2 \times 6)$ reconstructions, with and without different densities of antisite defects. The vertical (linear) scale is the same in each panel, although spectra have been vertically shifted for clarity (baselines are indicated). Spin orbit coupling is included in spectra of panel (c).

broad positive peak around 1.4 eV), and furthermore stabilized by the appropriate concentration of surface Ga antisite defects such that the ECR is fulfilled. Nevertheless it is as yet unknown how their energy compares with the known stable (4×3) structures. In the following we examine more closely the stability of the defect structures and discuss the mechanisms behind defect formation itself.

VI. SURFACE STABILITY

In this section we investigate the stability of the long dimer chain, defect-stabilized reconstructions by comparing their surface energies with those of the short three-dimer chain reconstructions. The surface energy per unit area γ is computed for the various structures and stoichiometries using the expression:[10,13]

$$\gamma = [E_{\text{tot}} - n_{\text{Ga}}\mu_{\text{Ga}} - n_{\text{Sb}}\mu_{\text{Sb}}]/A, \quad (9)$$

where E_{tot} is the total energy of the reconstructed *surface* of area A having n_{Ga} atoms of Ga and n_{Sb} atoms of Sb in excess of the unreconstructed Sb-terminated surface. We replace the atomic chemical potential, μ_{Sb}, with the excess chemical potential with respect to that of the elemental bulk phase: $\Delta\mu_{\text{Sb}} = \mu_{\text{Sb}} - \mu_{\text{Sb}}^{\text{bulk}}$, and make use of the standard thermodynamic relations[55]

$$\mu_{\text{Sb}} + \mu_{\text{Ga}} = \mu_{\text{GaSb}}^{\text{bulk}}, \quad (10)$$

$$\mu_{\text{GaSb}}^{\text{bulk}} - \mu_{\text{Ga}}^{\text{bulk}} - \mu_{\text{Sb}}^{\text{bulk}} = H_f \quad (11)$$

in order to rewrite γ in terms of the excess chemical potential of a single element. Here H_f is the formation enthalpy of bulk crystalline GaSb. The final expression to be computed is therefore

$$\gamma = \left[E_{\text{tot}} - n_{\text{Ga}}\mu_{\text{GaSb}}^{\text{bulk}} + (n_{\text{Ga}} - n_{\text{Sb}})\mu_{\text{Sb}}^{\text{bulk}}\right]/A$$
$$+ [(n_{\text{Ga}} - n_{\text{Sb}})\Delta\mu_{\text{Sb}}]/A. \quad (12)$$

In graphical form, this corresponds to a straight line $\gamma(\Delta\mu_{\text{Sb}})$ of slope $n_{\text{Ga}} - n_{\text{Sb}}$; $\Delta\mu_{\text{Sb}} = 0$ corresponds to the extremely Sb-rich limit.

In order to evaluate γ we computed the total energies E_{tot} within DFT-LDA for several of the structures shown in Figs. 1 and 8. These calculations differ from those of the previous sections in that we used slightly thinner slabs (seven layers), but applied more strict convergence criteria. Errors were minimized by using large supercells that fit multiple primitive cells of many different structural models, as well as equivalent **k**-point sets where possible [equivalent to 384 points in the (1×1) Brillouin zone]. A small metallic smearing (0.2 eV) was also added to consistently account for possible metallicity. Atoms (excluding the bottom bilayer) were relaxed until individual force components were less than 2.5 meV/Å. For comparison we have also calculated the surface energies of some surface reconstructions of GaAs(001) and AlSb(001). The chemical potentials of the cation and anion elements and of bulk III-V compounds were computed separately, assuming the stable α-Ga structure for elemental Ga, the fcc structure for elemental Al, and the rhombohedral A7 phases for elemental As and Sb. The III-V compounds are obviously considered in the zinc-blende structure.

The resulting GaSb(001) surface phase diagram is shown in Fig. 10, with the standard reconstructions of Fig. 1 (left panel) separated for clarity from the defect-stabilized ones of Fig. 8 (right panel). The surface formation energy is plotted in each case relative to that of the $\alpha(4 \times 3)$ phase, for which $n_{\text{Ga}} - n_{\text{Sb}} = 0$. Results for the GaSb$(4 \times 3)$ reconstructions agree quite well with those previously published[10,13] although we predict the $\beta(4 \times 3)$ to be stable across a wider range of chemical potential than previously reported. We mention that structures with shifted cell rows [Fig. 1(b)], including the $\gamma(4 \times 3)$ reconstruction and staircase variations of the $\beta(4 \times 3)$ and $\alpha(4 \times 3)$ phases, have in fact been demonstrated elsewhere to be slightly more stable than the phases plotted here,[13,56] although this does not influence the following discussion.

The influence of the surface Ga antisites on the surface formation energy is illustrated in Fig. 10(b). The $c(2 \times 6)$–$s1a$ structure lies more than 80 meV/(1×1) cell higher in energy than the $c(2 \times 6)$ phase and has been omitted from the diagram for clarity. We find that the $c(2 \times 6)$–$d1x$ structures
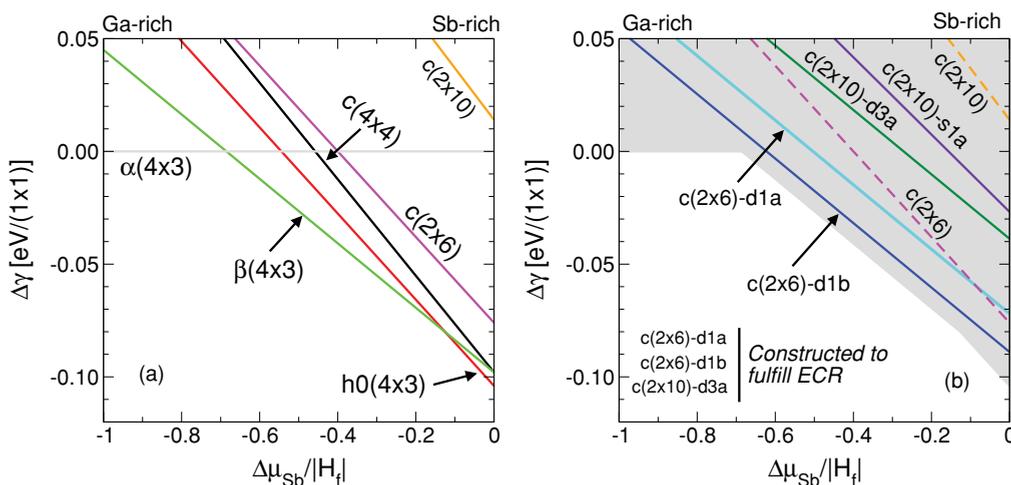
FIG. 10. (Color online) (a) Surface phase diagram for main GaSb(001) reconstructions depicted in Fig. 1. Surface energies γ are plotted relative to the $\alpha(4 \times 3)$ reconstruction. (b) Same as (a), but including reconstructions with Ga antisite defects (Fig. 8).

which obey the ECR are indeed far more stable than the metallic $c(2 \times 6)$, lowering the surface energy to within a few meV/(1×1) cell of the ground state (4×3) structures. Nevertheless, they remain less stable than the latter. Therefore, our calculations show that the phases stabilized by surface Ga defects are *metastable*, and as such are dictated by the particular environment conditions (i.e., kinetically stabilized). We note that at finite temperatures, not considered in our standard DFT calculations, entropic effects can make more disordered configurations accessible, as discussed for instance in Ref. 57.

Our calculations, as well as others (not shown) carried out in (4×6) supercells, show that (i) the antisites tend to spread out in order to occupy sites in *all* chains, so that structures like $c(2 \times 6)$–$d1a$ are significantly higher in energy than those like $c(2 \times 6)$–$d1b$; (ii) the total energies of structures having the same concentration of Ga antisites, but distributed slightly differently about the surface—subject to the fulfillment of condition (i)—are very similar. The latter point ensures that no defect configurations having a well established long-range order is favoured. Instead, a disordered (quasirandom) distribution of defects take place, such that the surface structure recovers the overall symmetry of the long-chain reconstructions [$c(2 \times 6)$ and $c(2 \times 10)$]. It is well known that a random distribution of defects do not add further spots to the diffraction pattern, since no long-range order superstructure is created, but only generate a more diffuse background. Furthermore, local fluctuations in the Ga antisite concentration may explain the weak metallicity observed with STS (a local probe) on the $c(2 \times 10)$ surface:[7] even if the ECR is globally satisfied, charge fluctuations induced by structural disorder ensure that it can be locally violated.

VII. DEFECT FORMATION

In this section we show that the Sb-rich conditions to which the surface is exposed in fact comprise the driving mechanism behind the formation of the subsurface defects. Indeed, we show that the defect formation is driven by the presence of excess charge on the surface. The charge fluctuation, in turn, is possible since the formation of Sb dimers is favored by the close lattice match between the GaSb(001) lattice and the Sb metal lattice. This is illustrated in Table I, which compares the second neighbor distance for various III–V compounds with the third neighbor distance in the rhombohedral A7 structure of As and Sb (the fact that GaSb, unique among the III–Vs, can be capped[16] with polycrystalline—i.e., not amorphous—Sb supports this idea). Other III–V compounds, in contrast, show mismatches greater than 6% (the exception being AlSb, with a small mismatch of 0.97%). This unique property of GaSb leads to electron-rich charge fluctuations on the surface occurring dynamically during growth.

A. Defect formation mechanism

As noted in Sec. III, the measured samples were exposed during and after growth to particularly Sb-rich conditions. In Ref. 18 we showed that the defect forms spontaneously, i.e., its formation energy is negative, when there is an abundance of Sb dimers on the surface. Here we review briefly the main points

TABLE I. Lattice matching between III–V compounds and the elemental bulk V (As or Sb) material: 2nn (3nn) refers to second (third) nearest neighbor. Values derived using the experimental lattice parameters.[58,59]

	2nn distance in III–V compound (Å)	3nn distance in elemental V (Å)	Lattice mismatch (%)
InAs	4.285	3.76	13.96
InSb	4.582	4.30	6.56
AlAs	4.002	3.76	6.44
GaAs	3.995	3.76	6.25
AlSb	4.342	4.30	0.97
GaSb	4.313	4.30	0.30

of that discussion for completeness. We considered extended simulation cells [see Fig. 11(a)] that are initially composed of closely packed long dimer chains [the (8×2) cells] or long dimer chains separated by in-dimers [the (8×3) cells]. The initial (8×3) structure is thus precisely equivalent to the usual $c(2 \times 6)$ model. These unbroken chain structures correspond to the largest possible value of ν for each cell type, namely $\nu = +0.166$ and $\nu = +0.5$ electrons per (1×1) unit cell for the (8×3) and (8×2) cells, respectively, and, consequently, have their Fermi levels lying inside the conduction band. We then systematically reduced ν by removing zero, one, and two dimers per cell from the dimer chain, a procedure which causes the Fermi level to move into the gap. Note that the (8×2) cell with two dimers removed from the positions shown in Fig. 11(a) simply corresponds to the (semiconducting) $c(4 \times 4)$ reconstruction. For each configuration (i.e., each value of ν) we determined the defect formation energy (DFE) of a single Ga antisite as

$$\text{DFE} = A(\gamma_{\text{phase+defect}} - \gamma_{\text{phase}})$$
$$= E_{\text{phase+defect}} - E_{\text{phase}} + \mu_{\text{Sb}} - \mu_{\text{Ga}}, \quad (13)$$

where A is the cell area. Using Eqs. (10) and (11), we arrive at the following expression:

$$\text{DFE} = E_{\text{phase+defect}} - E_{\text{phase}} + \mu_{\text{Sb}}^{\text{bulk}} - \mu_{\text{Ga}}^{\text{bulk}} - H_f + 2\Delta\mu_{\text{Sb}}. \quad (14)$$

Values for DFE$(\nu, \Delta\mu_{\text{Sb}})$ computed in this way for GaSb(001) are shown in Fig. 11(b), for both (8×2) and (8×3) simulation cells. The solid dots correspond to DFEs calculated at $\Delta\mu_{\text{Sb}}/|H_f| = 0$, while the "error" bars describe the variation of DFE over the range $-0.3 < \Delta\mu_{\text{Sb}}/|H_f| < 0$ (i.e., the high-Sb regime). It is clear from Fig. 11(b) that defect formation occurs spontaneously (DFE < 0) for GaSb(001) as soon as ν exceeds 0.14 electrons per (1×1) cell. This shows that exchange of a subsurface Sb atom with a Ga atom can be driven by a fluctuation in the surface electron number. In the figure we also present (open symbols) average computed formation energies for a single defect in the $c(2 \times 6)$–$d1a$, $c(2 \times 10)$–$s1a$, and $c(2 \times 10)$–$d3a$ models using the approximate expression

$$\langle\text{DFE}\rangle = (E_{\text{phase+defects}} - E_{\text{phase}})/n_{\text{Ga}} + \mu_{\text{Sb}} - \mu_{\text{Ga}}, \quad (15)$$

reconstructions with and without the antisite defect into various energetic contributions:

$$E_{\mathrm{phase+defect}} - E_{\mathrm{phase}} = \Delta E_{\mathrm{strain}} + \Delta E_{\mathrm{bonding}} + \Delta E_{\mu}. \tag{16}$$

To explain the different terms in this equation, we imagine that defect formation—i.e, the substitution of a subsurface Sb with a Ga atom—occurs in two separate steps. In the first step, the atoms in the lattice relax to the final positions that would be obtained if the defect was actually present. During this process, the original chemical composition is maintained. This yields a positive elastic energy contribution, $\Delta E_{\mathrm{strain}}$. In the second step, the Sb atom in its new position is substituted by the Ga defect. This gives a contribution to the energy from two terms: a term due to the creation of new chemical bonds, $\Delta E_{\mathrm{bonding}}$, and a fixed term simply given by the difference in the chemical potentials of Ga and Sb, $\Delta E_{\mu} = \mu_{\mathrm{Ga}} - \mu_{\mathrm{Sb}}$. Alternatively (and analogously), defect formation could be imagined as occurring in the opposite way: first the Sb to Ga substitution takes place without allowing the lattice to relax (yielding a positive contribution to the energy due to rebonding), and then the lattice is let to relax giving a negative contribution to the energy. Clearly, the sum of each term in Eq. (16) must give the same calculated total energy difference within both scenarios. In practice, we calculated $\Delta E_{\mathrm{strain}}$ by performing an additional calculation on the $c(2 \times 6)$ structure but with the atomic positions frozen in the sites they would occupy if the defect were present. Considering for simplicity the extreme Sb-rich condition in the limit $\Delta \mu_{\mathrm{Sb}} = 0$, Eqs. (14) and (16) when combined yield

$$\mathrm{DFE} = \Delta E_{\mathrm{strain}} + \Delta E_{\mathrm{bonding}}, \tag{17}$$

allowing $\Delta E_{\mathrm{bonding}}$ to be determined.

Figure 12 shows the strain and bonding energetic contributions to the DFE computed for GaSb, AlSb, and GaAs. First, we note that the (positive) strain energy is highest for GaSb. This is an indication that the deformation of the lattice with respect to the initial defect-free surface equilibrium configuration, which accompanies the defect formation, is *larger* for GaSb than for GaAs and AlSb. The energy gain when the substitution actually takes place ($\Delta E_{\mathrm{bonding}}$) is also much higher for GaSb than for AlSb and GaAs. These results show that the defect stability in GaSb is mainly due to the ability of GaSb to strain its lattice enough so as to form the best possible Ga–Ga and Ga–Sb bonds around the defect and thereby gain much more in bonding energy. In contrast, $\Delta E_{\mathrm{strain}}$ and $\Delta E_{\mathrm{bonding}}$ balance each other out in GaAs and AlSb, thus yielding a constant, positive DFE of about 0.5 eV. This different behavior derives from the smaller elastic constants of GaSb with respect to those of GaAs and AlSb as shown in Table II. Hence, the ultimate reason of the negative DFE lies in the smaller rigidity of the GaSb lattice.

FIG. 11. (Color online) (a) Schematic diagrams of the extended (8×3) and (8×2) surface cells used to compute the defect formation energy (DFE) of surface Ga antisites. (b) DFE computed as a function of the number of excess electrons ν. Open symbols indicate the average defect energies of the $c(2 \times 6)$–$d1b$, $c(2 \times 10)$–$s1a$, and $c(2 \times 10)$–$d3a$ structures. Numbers in parentheses indicate the amount of dimers removed from initial extended cells.

where we assume that the defects are inserted simultaneously. In all cases the defect is found to be stable. This spontaneous formation of defects on Sb-rich GaSb(001) stabilizes the long Sb-dimer chain metastable structures that are observed experimentally by pinning the dimers to the surface and thus inhibiting the formation of the missing-dimer $c(4 \times 4)$ reconstruction.

In contrast, the DFE in GaAs(001) and AlSb(001) have been elsewhere demonstrated[18] to be positive for all values of ν, implying that these surfaces stabilize by means of removing dimers—which is indeed consistent with the experimental observation of the $c(4 \times 4)$ reconstruction in the very V-rich regime.

B. The role of surface strain in the defect stabilization

In order to analyze the reasons for the defect stability in GaSb(001) [and its instability in AlSb(001) and GaAs(001)] we decompose the total energy difference between the $c(2 \times 6)$

VIII. SUMMARY AND CONCLUSIONS

In this paper we have tackled the still open problem of the identification of the $c(2 \times 6)$ reconstruction observed on Sb-rich GaSb(001) surfaces. We have performed accurate *ab initio* simulations of reflectance anisotropy spectra and

TABLE II. Elastic constants of bulk GaAs, AlSb, and GaSb. Taken from Harrison.[60] Units are 10^{11} erg/cm^3.

	C_{11}	C_{12}	C_{44}
GaAs	11.81	5.32	5.92
AlSb	8.94	4.43	4.16
GaSb	8.84	4.03	4.32

surface energy calculations for a large number of reconstruction models of the Sb-stabilized GaSb(001) surface in order to determine the structure of the $c(2 \times 6)$ phase. The calculated RAS spectra have been compared to the experimental ones previously reported in the literature for which a distinct $c(2 \times 6)$ LEED pattern was observed.[16,17]

Surface reconstruction models of Sb-rich GaSb(001) can be classified roughly into two groups: short three-dimer Sb chain reconstructions interrupted by displaced or missing dimers (with or without in-dimers), and long Sb-dimer chain reconstructions (with in-dimers). The reconstructions of the first group satisfy the ECR and are semiconducting while those of the second group do not satisfy the ECR and are metallic. Our first-principles calculation of the GaSb(001) surface phase diagram shows that the reconstructions of the first group— the (4×3) ones in particular—are much more stable than those belonging to the second group. Thus, the short chain reconstructions should be favored against the experimental observations that the very Sb-rich GaSb(001) reconstructions feature long chain reconstructions.[10,13,56]

On the other hand, the comparison between our simulated RAS and the measured data led to the conclusion that a single reconstruction model does not agree with the experiment. Instead, features belonging to both groups of reconstructions are clearly present in the experimental spectra. As a consequence, the observed $c(2 \times 6)$ reconstruction should consist of both long chain $c(2 \times 6)$ and short-chain (4×3) reconstructions, possibly with the long chains frequently interrupted by missing or displaced dimers. A similar conclusion was drawn also in

FIG. 12. (Color online) Defect formation energies (DFE) of typical III–V(001) surfaces, along with the strain and bonding contributions at $\Delta \mu_{Sb} = 0$.

Ref. [61] for the case of the Bi-stabilized InSb(001) surface, where the coexistence of metallic $c(2 \times 6)$ and semiconducting (4×3) regimes was proposed based on the results of STS/STM and photoemission spectroscopy measurements and first-principles calculations.

A mixture of $c(2 \times 6)$ and (4×3) reconstructions would nonetheless yield a strong discrepancy between theory and experiment in the energy range between 2 and 3 eV, as the computed positive peak is in stark disagreement with the negative feature observed experimentally. This positive structure persists even if spin-orbit coupling or many-body effects are considered. We further note that the large energetic instability of the long chain $c(2 \times 6)$ reconstruction model makes it difficult to explain its appearance as an actual phase of GaSb(001) at the temperatures used experimentally. These observations have led us to speculate about the existence of a new structural motif capable of improving the agreement with the experimental RAS and lowering the surface energy of the metallic long chain reconstructions.

We have determined the relation between the ECR and the number of structural motifs on the Sb-rich GaSb(001) surface, showing that the excess electron charge can be lowered not only by breaking the dimer chains, but also by inserting substitutional "electron-killing" defects, in particular by substituting Sb atoms with Ga atoms (the surface antisite defects) in the layer below the dimer chains. When inserted in the right number so as to satisfy the ECR they lead to a substantial lowering of the surface energy of the $c(2 \times 6)$ reconstruction. The new structure remains metastable, however, since its energy is still higher (albeit by a much smaller amount) than that of the stable (4×3) reconstructions. Furthermore, the occurrence of the $c(2 \times 6)$ reconstruction with defects occurs spontaneously due to the particularly Sb rich conditions to which the GaSb(001) surface is exposed during growth as we previously showed.[18]

In this paper we also show that the formation of the surface antisite defects is due to two factors which are at work for the GaSb(001) surface. The first factor is the almost perfect surface lattice match between GaSb(001) and metallic Sb which favors the occurrence of electron rich fluctuations on the surface. The second factor is the softness of the Ga–Sb bonds that allows the surface atomic arrangement to relax more efficiently around the defect and thus strengthen the bonding of the defect to the surface layer. We also find that the defects tend to should distribute randomly while trying to increase their distance from one another. Finally, the presence of the surface antisite defects noticeably improves the simulated RAS signal of the long chain reconstructions in the energy range between 2 and 3 eV making them closer to the measured spectra for the $c(2 \times 6)$ surface.

In conclusion, based on our *ab initio* study of the reflectance anisotropy signals and of the surface energies, we propose that the main reconstruction constituting the $c(2 \times 6)$ surface is a modified $c(2 \times 6)$ structure featuring Ga randomly substituted into the second row. Significant structural disorder should also be present by which the $c(2 \times 6)$ long dimer chains are interrupted by shifted dimers, thus introducing (4×3) structural motifs into the surface reconstruction.

*Corresponding author: conor.hogan@roma2.infn.it

[1] H. Kroemer, Physica E **20**, 196 (2004).

[2] R. Magri, A. Zunger, and H. Kroemer, J. Appl. Phys. **98**, 043701 (2005).

[3] R. Magri, L. W. Wang, Alex Zunger, I. Vurgaftman, and J. R. Meyer, Phys. Rev. B **61**, 10235 (2000).

[4] L. He, G. Bester, and A. Zunger, Phys. Rev. Lett. **94**, 016801 (2005).

[5] J. T. Olesberg, W. H. Lau, M. E. Flatté, C. Yu, E. Altunkaya, E. M. Shaw, T. C. Hasenberg, and T. F. Boggess, Phys. Rev. B **64**, 201301(R) (2001).

[6] Yu. D. Glinka, T. V. Shahbazyan, I. E. Perakis, N. H. Tolk, X. Liu, Y. Sasaki, and J. K. Furdyna, Appl. Phys. Lett. **81**, 220 (2002).

[7] L. J. Whitman, P. M. Thibado, S. C. Erwin, B. R. Bennett, and B. V. Shanabrook, Phys. Rev. Lett. **79**, 693 (1997).

[8] M. T. Sieger, T. Miller, and T.-C. Chiang, Phys. Rev. B **52**, 8256 (1995).

[9] W. Barvosa-Carter, A. S. Bracker, J. C. Culbertson, B. Z. Nosho, B. V. Shanabrook, L. J. Whitman, H. Kim, N. A. Modine, and E. Kaxiras, Phys. Rev. Lett. **84**, 4649 (2000).

[10] M. C. Righi, R. Magri, and C. M. Bertoni, Phys. Rev. B **71**, 075323 (2005).

[11] O. Romanyuk, V. M. Kaganer, R. Shayduk, B. P. Tinkham, and W. Braun, Phys. Rev. B **77**, 235322 (2008).

[12] G. P. Srivastava, Appl. Surf. Sci. **252**, 7600 (2006).

[13] J. Houze, S. Kim, S.-G. Kim, S. C. Erwin, and L. J. Whitman, Phys. Rev. B **76**, 205303 (2007).

[14] U. Resch-Esser, N. Esser, B. Brar, and H. Kroemer, Phys. Rev. B **55**, 15401 (1997).

[15] G. E. Franklin, D. H. Rich, A. Samsavar, E. S. Hirschorn, F. M. Leibsle, T. Miller, and T.-C. Chiang, Phys. Rev. B **41**, 12619 (1990).

[16] C. Goletti, U. Resch-Esser, J. Foeller, N. Esser, W. Richter, B. Brar, and H. Kroemer, Surf. Sci. **352–354**, 771 (1996). An opposite sign convention was used in this work.

[17] K. Möller, Z. Kollonitsch, C. Giesen, M. Heuken, F. Willig, and T. Hannappel, J. Cryst. Growth **248**, 244 (2003).

[18] C. Hogan, R. Magri, R. Del Sole, Phys. Rev. Lett. **104**, 157402 (2010).

[19] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[20] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[21] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. Fabris, G. Fratesi, S. de Gironcoli, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, J. Phys. Condens. Matter **21**, 395502 (2009).

[22] F. Manghi, R. Del Sole, A. Selloni, and E. Molinari, Phys. Rev. B **41**, 9935 (1990).

[23] R. Del Sole, Solid State Commun. **37**, 537 (1981).

[24] J. D. E. McIntyre and D. E. Aspnes, Surf. Sci. **24**, 417 (1971).

[25] C. Hogan, R. Del Sole, and G. Onida, Phys. Rev. B **68**, 035405 (2003).

[26] A. Marini, C. Hogan, M. Grüning, and D. Varsano, Comput. Phys. Commun. **180**, 1392 (2009). See [http://www.yambo-code.org].

[27] I. Vurgaftman, J. Appl. Phys. **89**, 5815 (2001).

[28] H. Ebert, Rep. Prog. Phys. **59**, 1665 (1996).

[29] R. A. Vázquez-Nava, N. Arzate, and B. S. Mendoza, Phys. Status Solidi B **247**, 1979 (2010).

[30] R. A. Vázquez-Nava, Bernardo S. Mendoza, and C. Castillo, Phys. Rev. B **70**, 165306 (2004); R. A. Vázquez-Nava, B. S. Mendoza, and N. Arzate, Phys. Status Solidi B **242**, 3022 (2005).

[31] Leonard Kleinman, Phys. Rev. B **21**, 2630 (1980).

[32] Relativistic pseudopotentials were generated using the code LD1, part of the QUANTUM-ESPRESSO distribution, with the kind support from the author, P. Giannozzi.

[33] S. G. Louie, S. Froyen, and M. L. Cohen, Phys. Rev. B **26**, 1738 (1982).

[34] For Sb, we used the Sb.pz-bhs.UPF pseudopotential available from the QUANTUM-ESPRESSO website; for Ga, we converted an NLCC pseudopotential originally generated with the FHI98PP package, M. Fuchs and M. Scheffler, Comput. Phys. Commun. **119**, 67 (1999). The lattice constant in this case is 6.054 Å at 15 Ry.

[35] W. G. Schmidt, Phys. Status Solidi B **242**, 2751 (2005).

[36] Maurizia Palummo, Olivia Pulci, Rodolfo Del Sole, Andrea Marini, M. Schwitters, S. R. Haines, K. H. Williams, D. S. Martin, P. Weightman, and J. E. Butler, Phys. Rev. Lett. **94**, 087404 (2005).

[37] Maurizia Palummo, Nadine Witkowski, Olivier Pluchery, Rodolfo Del Sole, and Yves Borensztein, Phys. Rev. B **79**, 035327 (2009).

[38] P. H. Hahn, W. G. Schmidt, and F. Bechstedt, Phys. Rev. Lett. **88**, 016402 (2001).

[39] M. Palummo, O. Pulci, A. Marini, R. Del Sole, P. H. Hahn, W. G. Schmidt, and F. Bechstedt, J. Phys. Condens. Matter **16**, S4313 (2004).

[40] O. Pulci, G. Onida, R. Del Sole, and L. Reining, Phys. Rev. Lett. **81**, 5374 (1998).

[41] W. G. Schmidt, J. L. Fattebert, J. Bernholc, and F. Bechstedt, Surf. Rev. Lett. **6**, 1159 (1999).

[42] R. Del Sole and R. Girlanda, Phys. Rev. B **48**, 11789 (1993).

[43] C. Hogan, E. Placidi, and R. Del Sole, Phys. Rev. B **71**, 041308(R) (2005).

[44] S. Albrecht, L. Reining, R. Del Sole, and G. Onida, Phys. Rev. Lett. **80**, 4510 (1998).

[45] L. X. Benedict and E. L. Shirley, Phys. Rev. B **59**, 5441 (1999).

[46] R. Haydock, Comput. Phys. Commun. **20**, 11 (1980).

[47] Xuejun Zhu and Steven G. Louie, Phys. Rev. B **43**, 14142 (1991).

[48]Z. Kollonitsch, K. Möller, H.-J. Schimper, Ch. Giesen, M. Heuken, F. Willig, and T. Hannappel, J. Cryst. Growth **261**, 289 (2004).

[49]O. J. Pitts, S. P. Watkins, and C. X. Wang, J. Cryst. Growth **248**, 249 (2003).

[50]S. Weeke, MOVPE TU-Berlin, private communication.

[51]W. G. Schmidt, F. Bechstedt, K. Fleischer, C. Cobet, N. Esser, W. Richter, J. Bernholc, and G. Onida, Phys. Status Solidi A **188**, 1401 (2001).

[52]F. Bechstedt, *Principles of Surface Physics* (Springer, Berlin, 2003).

[53]E. Penev, S. Stojković, P. Kratzer, and M. Scheffler, Phys. Rev. B **69**, 115335 (2004).

[54]K. Chuasiripattana and G. P. Srivastava, Surf. Sci. **600**, 3803 (2006).

[55]Guo-Xin Qian, Richard M. Martin, and D. J. Chadi, Phys. Rev. B **38**, 7649 (1988).

[56]O. Romanyuk, F. Grosse, and W. Braun, Phys. Rev. B **79**, 235330 (2009).

[57]O. Romanyuk, F. Grosse, and W. Braun, Phys. Status Solidi C **7**, 330 (2010).

[58]X. Gonze, J.-P. Michenaud, and J.-P. Vigneron, Phys. Rev. B **41**, 11827 (1990) and references therein.

[59]Y. Liu and R. E. Allen, Phys. Rev. B **52**, 1566 (1995).

[60]W. Harrison, *Electronic Structure and the Properties of Solids* (Dover Publications, New York, 1989).

[61]P. Laukkanen, M. P. J. Punkkinen, N. Räsänen, M. Ahola Tuomi, M. Kuzmin, J. Lång, J. Sadowski, J. Adell, R. E. Perälä, M. Ropo, K. Kokko, L. Vitos, B. Johansson, M. Pessa, and I. J. Väyrynen, Phys. Rev. B **81**, 035310 (2010).