III–V compound semiconductor (001) surfaces

W.G. Schmidt*

Computational Materials Science Group, Institut für Festkörpertheorie und Theoretische Optik, Friedrich-Schiller-Universität, Max-Wien-Platz 1, 07743 Jena, Germany

Received: 18 May 2001/Revised version: 23 July 2001/Published online: 3 April 2002 – © Springer-Verlag 2002

Abstract. There has been renewed interest in the structure of III–V compound semiconductor (001) surfaces caused by recent experimental and theoretical findings, which indicate that geometries different from the seemingly well-established dimer models describe the surface ground state for specific preparation conditions. I review briefly the structure information available on the (001) surfaces of GaP, InP, GaAs and InAs. These data are complemented with first-principles total-energy calculations. The calculated surface phase diagrams are used to explain the experimental data and reveal that the stability of specific surface structures depends largely on the relative size of the surface constituents. Several structural models for the Ga-rich GaAs (001)(4×6) surface are discussed, but dismissed on energetic grounds. I discuss in some detail the electronic properties of the recently proposed cation-rich GaAs (001)(4×2) geometry.

PACS: 68.35.Bs; 68.35.Md; 73.20.At

The (001)-oriented substrates of III–V semiconductors are commonly used in growth technologies like molecular beam epitaxy (MBE) or metal-organic vapour phase epitaxy (MOVPE). Apart from this application aspect, much interest in the structure and stability of the (001) surfaces has also been caused by the rich variety of surface reconstructions which can be observed depending on the surface preparation. For instance the GaAs (001) surface shows a c(4×4) symmetry for As-rich conditions, but changes its periodicity to (2×4)/c(2×8) and finally (4×2)/c(8×2) as the surface gets more cation-rich. In addition, a large variety of transient structures or apparent symmetries such as (2×3), (2×1), (3×1), (3×6), (2×6), (6×6), (4×6), etc., are reported [1–4].

The large number of surface structures observed on III–V(001) surfaces have early on prompted attempts to classify and understand them from a more general point of view. For example a so-called electron-counting rule (ECR) was recognised to govern many reconstructions [5]. A surface structure satisfies this model if it is possible to have all the dangling bonds on the electronegative element occupied and all the dangling bonds on the electropositive element empty, given the number of available electrons. This condition results necessarily in there being no net surface charge. The ECR together with the dimerisation as the major mechanism to reduce the number of dangling bonds on III–V (001) surfaces were able to explain successfully a large number of surface geometries. This success prompted Mönch to write [2]: “The electronic energy of such surfaces will be lowest when i. surface atoms in the top layer are forming dimers and ii. dangling bonds are filled on surface anions and are empty at the surface cations”.

While the ECR gives an indication of which structures might be stable and which should not be, it does not allow us to discriminate energetically between two structures complying with the electron-counting principle. A major progress in that respect was the work by Northrup and Froyen on the role of electrostatic interactions between the surface structural units [6]. They showed that Coulomb repulsion between negatively charged anion dimers favours the two-dimer β(2×4) structure over the three-dimer β(2×4) structure for GaAs. Based on similar electrostatic arguments, Schmidt et al. [7] predicted recently that a one-dimer α(2×4) structure should be lower in energy than the hitherto accepted α(2×4) model.

The estimation of relative surface energies from empirical concepts was further refined by Zhang and Zunger [8], who observed that the large collection of equilibrium surface structures is built from a limited number of recurring local structural motifs, such as threefold-coordinated pyramidal anions or surface dimers. The energies of these structural motifs can be obtained from a fit to the results of ab initio calculations on flat surfaces and bulk defects. Based on motif energies obtained in such a way, Zhang and Zunger estimated the surface energies for reconstructions larger than what could be dealt with from first principles at that time. The accuracy of their method is limited, however, as I will show below, using the GaAs (001)(2×6) surface as an example.

A related approach has been followed by Mirbt et al. [9]. Starting from atomic energy levels and bonding energies they derived a very simple expression called a surface reconstruc-

* (Fax: +49-3641/947152; E-mail: w.g.schmidt@ifto.physik.uni-jena.de)
tation parameter in order to assess the stability of reconstruction models. According to this parameter, for the stability of III–V surfaces it is required that under cation-rich conditions the sum of anion dangling bonds and anion–anion bonds should be minimised. For anion-rich conditions the expression \( N_a - N_{a-a} + 2N_c-c \) needs to be minimised, where \( N_a \) is the number of anion dangling bonds and \( N_{a-a} \) and \( N_{c-c} \) denote the numbers of anion–anion and cation–cation bonds. Later, the same group pointed out how local stress may modify these rules [10].

The GaAs dimer models were extrapolated and used to explain the (001) surface geometries of materials like GaP, InP and InAs. In many cases these dimer geometries allowed us indeed to explain the experimental findings and a consistent picture emerged [3]. In the last three to four years, however, it became more and more clear that our understanding of III–V (001) surface reconstructions is still far from being complete and that the complexity of the forces driving the formation of certain reconstructions is not captured alone by the simple rules given above. Symmetric dimers, for example, fail to explain all (001) surface structures, as became obvious from the scanning tunnelling microscopy (STM) images of the In-rich InP (001) \((2 \times 4)\) surface [11]. Mixed hetero-dimers consisting of anions and cations were proposed instead for cation-rich surfaces of InP [12] and GaP [13]. Moreover, \((2 \times 1)\) and \((2 \times 2)\) reconstructions were found for P-rich InP surfaces [14], while a \((2 \times 1)\) surface was reported for GaP (001) [15]. It is known, however, that the \((2 \times 4)\) reconstruction is the smallest unit cell which complies with the ECR [5]. Violations of the electron-counting rule were also reported for AlSb and GaSb (001) surfaces [16]. The GaAs (001) surface is arguably the most intensively investigated compound semiconductor surface. Its stoichiometry-dependent surface structure seemed to be well-established from both experiment and theory. Nevertheless, even for GaAs, recent theoretical work questions the validity of the ‘traditional’ dimer model to explain the cation-rich \((4 \times 2)/c(8 \times 2)\) reconstruction [17].

In the present paper I will briefly review the experimental and theoretical findings on the structure of the (001) surfaces of GaP, InP, GaAs and InAs. These findings are explained and put in perspective using surface phase diagrams derived from first-principles-total-energy calculations. The calculations presented here are well-converged and on the same footing for all materials considered. They can thus be used to identify chemical trends.

1 Method

Density-functional theory [18, 19] in the local-density approximation [20, 21] (DFT–LDA) together with nonlocal norm-conserving pseudopotentials [22] is used to determine the structurally relaxed ground states of the surface structures. The Ga 3d and In 4d electrons, respectively, are partially taken into account by means of a nonlinear core correction to the exchange and correlation energies. A massively parallel, real-space finite-difference method [23] is used to deal efficiently with the large unit cells needed to describe the surface. A multigrid technique is employed for convergence acceleration. The spacing of the finest grid used to represent the electronic wave functions and the charge density was determined through a series of bulk calculations. We find that structural and electronic properties are converged for a spacing corresponding to 4% of the bulk lattice constant. The calculations were performed using the theoretical equilibrium lattice constants of 5.39, 5.84, 5.57 and 6.00 Å, respectively, for GaP, InP, GaAs and InAs. They are slightly smaller than the corresponding experimental values measured at room temperature (5.45, 5.87, 5.65 and 6.06 Å [24]) due to the overestimation of the bond strengths typical for LDA calculations and the neglect of temperature effects.

The surfaces are modelled by using periodic supercells. They contain material slabs about 12 Å thick, separated by 12 Å of vacuum. The surface dangling bonds at the bottom layer are saturated with fractionally charged pseudohydrogens. The investigated geometries were relaxed until all calculated forces were below 20 meV Å\(^{-1}\). The atoms in the lowest bilayer were kept frozen in the ideal bulk configuration. Integrations in the surface Brillouin zone are performed over four special \(k\) points in its irreducible part.

In order to compare energetically surface structures representing different stochiometries, one has to take into account the chemical potentials \(\mu\) of the surface constituents. Since the surface is in equilibrium with the bulk material, they are related to each other: their sum equals the chemical potential of the bulk semiconductor. Consequently, the surface formation energy may be written as a function of a single variable, which I take to be the relative chemical potential of the cation with respect to its bulk phase, \(\Delta\mu\) (Ga, In). The computational accuracy in determining the chemical potentials is of the order of 0.1 eV [25]. The uncertainty of the calculated surface energies is less than 0.01 eV per surface atom.

2 Results and discussion

2.1 GaP

Experimentally, \((4 \times 2)\) [15, 26–32], \((2 \times 4)\) [13, 26, 33–38], \((2 \times 1)/(2 \times 2)\) [13, 33, 37, 38] and \((4 \times 4)\) reconstructions [35] were observed for GaP (001) surfaces with increasing P content. First-principles total-energy calculations confirmed the stability of \((2 \times 4)\) [10, 13] and \((4 \times 4)\) reconstructions [39]. They gave no indication, however, for a \((4 \times 2)\) symmetry. Recent experimental studies [13, 37] also suggest that the reported \((4 \times 2)\) periodicities are in fact \((2 \times 4)\) reconstructions, provided the crystal axes are defined according to the usual notation. The wrong symmetry assignment was probably caused by the fact that cation-rich GaAs (001) surfaces, often serving as a model case, reconstruct \((4 \times 2)\).

Pulci et al. [40] calculated the surface energy for a large variety of structures put forward to explain the GaP \((2 \times 1)/(2 \times 2)\) surfaces observed for P-rich preparation conditions. The surface phase diagram for GaP shown in Fig. 1 contains the energetically most favourable \((2 \times 1)/(2 \times 2)\) reconstructions as found by Pulci et al. [40] in conjunction with earlier total-energy results by Schmidt and co-workers [13, 39]. The corresponding equilibrium structures are shown in Fig. 2.

The calculated phase diagram predicts two \((2 \times 4)\) reconstructed surface structures which should occur for Ga-