

Nature of bonding forces between two hydrogen-passivated silicon wafers

K. Stokbro and E. Nielsen

Mikroelektronik Centret (MIC), Technical University of Denmark, Building 345E, DK-2800 Lyngby, Denmark

E. Hult, Y. Andersson, and B. I. Lundqvist

Department of Applied Physics, Chalmers University of Technology and Göteborg University, S-41296 Göteborg, Sweden

(Received 8 September 1998)

The nature and strength of the bonding forces between two H-passivated Si surfaces are studied with the density-functional theory, using an approach based on recent theoretical advances in understanding of van der Waals forces between two surfaces. Contrary to previous suggestions of van der Waals attraction between H overlayers, we find that the attraction is mainly due to long-range van der Waals interactions between the Si substrates, while the equilibrium separation is determined by short-range repulsion between occupied Si-H orbitals. Estimated bonding energies and Si-H frequency shifts are in qualitative agreement with experiment. [S0163-1829(98)06448-0]

Direct bonding of oxide-free Si wafers is a promising technique for the fabrication of materials with abrupt dopant profiles,^{1,2} as well as for the integration of micromechanical and microelectronic components.³ A standard way of preparing Si wafers prior to bonding is to rinse them in a H-fluoride solution, thus forming H-passivated (hydrophobic) surfaces. When the two surfaces are brought into contact they bond spontaneously, but weakly, with bonding energies in the range 10–30 mJ/m².^{4–7} To strengthen this bond, the wafers are subsequently annealed at high temperature, whereby hydrogen desorbs from the interlayer.^{1,5}

There have been several proposals for the origin of the initial room-temperature bond between the two H-passivated surfaces, including intrinsic bonding due to van der Waals forces between H atoms^{4,5} and bonding caused by adsorbed OH groups⁸ or F atoms.⁹ In this paper, we investigate the nature of the bonding forces using first-principles techniques. Our theoretical method is based on density-functional theory (DFT), and we compare calculations where the exchange-correlation energy is calculated in the local-density approximation (LDA),¹⁰ the generalized gradient approximation (GGA),¹¹ and an approach based on recent advances in the description of van der Waals forces in DFT.^{12–15} To our knowledge, this is the first time such an approach has been used to calculate binding energies.

Calculations are performed for two Si(100)-H(2×1) and two Si(111)-H(1×1) surfaces. In both cases we find an attractive interaction. We show, contrary to previous proposals, that the long-range attraction is mainly due to van der Waals forces between the underlying Si substrates, with a weak dependence on the Si-H bond length. The latter vanishes at large separations, but at the equilibrium separation it gives rise to a small elongation of the Si-H bond. The equilibrium separation is determined by the competition between the attractive van der Waals force and kinetic-energy repulsion between occupied Si-H orbitals.

The DFT calculations are based on a plane-wave electronic-structure program, where H and Si atoms are described by ultrasoft pseudopotentials.¹⁶ We determine the Si lattice constant to be $a_0^{\text{GGA}} = 5.46 \text{ \AA}$ in the GGA and $a_0^{\text{LDA}} = 5.40 \text{ \AA}$ in the LDA, both in good agreement with the experimental lattice constant $a_0^{\text{Exp}} = 5.43 \text{ \AA}$.¹⁷ To study surface-surface interactions, we use a supercell with two slabs, where one is a mirror image of the other (see Fig. 1). We vary the surface-surface separation d defined as the distance between the top layers of Si atoms, as well as the relative parallel displacement \mathbf{R} measured in units of the surface-lattice constant. We calculate the interaction energy, $E_{\text{int}} = E_2 - 2E_1$, where E_2 is the total energy of the two interacting slabs in the supercell, and E_1 is the total energy of an isolated slab in the supercell. By using the same supercell length for both calculations, we obtain faster convergence of the interaction energy as a function of the plane-wave basis set. Detailed convergence tests¹⁸ show that well-converged interaction energies can be obtained by using two four-layer Si slabs with H-passivated back sides, a total of 17 Å empty space in the supercell, special-point k -grid spacings $< 0.4 \text{ \AA}^{-1}$, and plane-wave basis sets with kinetic-energy cutoff 23 Ry. The geometry of the isolated surfaces is ob-

= 5.40 Å in the LDA, both in good agreement with the experimental lattice constant $a_0^{\text{Exp}} = 5.43 \text{ \AA}$.¹⁷ To study surface-surface interactions, we use a supercell with two slabs, where one is a mirror image of the other (see Fig. 1). We vary the surface-surface separation d defined as the distance between the top layers of Si atoms, as well as the relative parallel displacement \mathbf{R} measured in units of the surface-lattice constant. We calculate the interaction energy, $E_{\text{int}} = E_2 - 2E_1$, where E_2 is the total energy of the two interacting slabs in the supercell, and E_1 is the total energy of an isolated slab in the supercell. By using the same supercell length for both calculations, we obtain faster convergence of the interaction energy as a function of the plane-wave basis set. Detailed convergence tests¹⁸ show that well-converged interaction energies can be obtained by using two four-layer Si slabs with H-passivated back sides, a total of 17 Å empty space in the supercell, special-point k -grid spacings $< 0.4 \text{ \AA}^{-1}$, and plane-wave basis sets with kinetic-energy cutoff 23 Ry. The geometry of the isolated surfaces is ob-

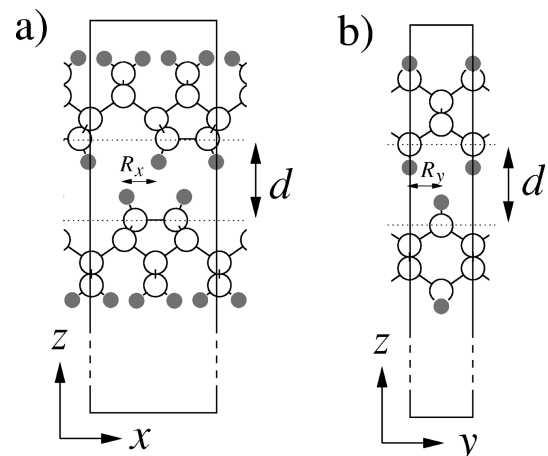


FIG. 1. Geometry of supercell with two Si(100)-H(2×1) slabs. (a) Side view of the x direction. (b) Side view of the y direction. The components R_x and R_y of the relative parallel displacement between the two slabs are indicated.

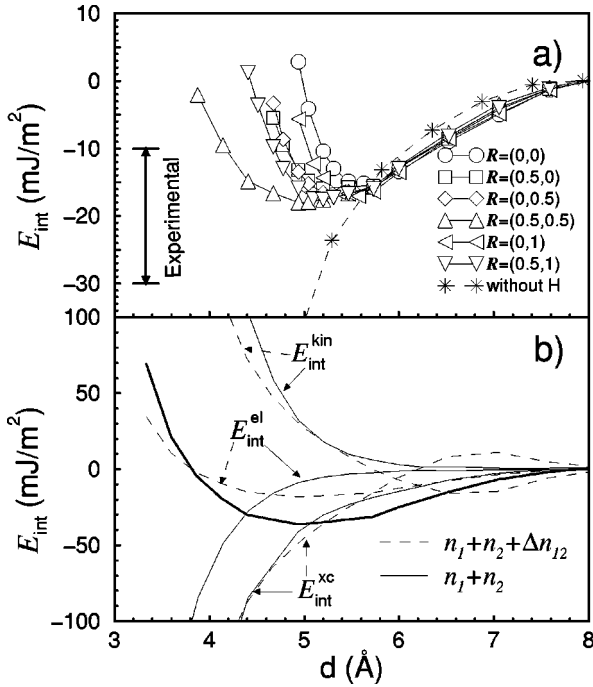


FIG. 2. (a) Interaction energies calculated within GGA of two Si(100)-H(2×1) slabs, as a function of their separation d . Different curves are for varying parallel displacement \mathbf{R} . The dashed line shows the interaction energy when H is removed from the two surfaces. (b) The interaction energy for displacement $\mathbf{R}=(0.5,0.5)$ separated into kinetic energy $E_{\text{int}}^{\text{kin}}$, electrostatic energy $E_{\text{int}}^{\text{el}}$, and exchange-correlation energy $E_{\text{int}}^{\text{xc}}$. Solid lines show calculations using a non-self-consistent density, and dashed lines show self-consistent calculations.

tained by relaxing the H atoms and the first two Si layers, and the resulting bond lengths and bond angles are similar to other first-principles calculations.^{19–23} Test calculations show that surface relaxations due to interaction between the wafers change the interaction energies by less than 5%, and we only include such effects for the calculation of Si-H frequency shifts.

Figure 2(a) shows the interaction energy calculated within GGA for the two Si(100)-H(2×1) surfaces as a function of d and \mathbf{R} . At large d , the interaction energy is nearly independent of \mathbf{R} and coincides with the interaction energy of two Si(100) surfaces without H passivation (dashed line). At smaller d , the interaction between the passivated surfaces becomes repulsive, whereas it continues to become more attractive for the nonpassivated surfaces. The value of d for which the interaction energy is a minimum depends on the relative displacement \mathbf{R} . The bonding energy is smallest for the relative geometry where the H atoms on the two surfaces are in registry $\mathbf{R}=(0,0)$ and largest, where they are laterally furthest apart $\mathbf{R}=(0.5,0.5)$.

We next investigate the microscopic nature of the short-range repulsive and long-range attractive forces. For this purpose, we divide the interaction energy into contributions from kinetic energy $E_{\text{int}}^{\text{kin}}$, electrostatic energy $E_{\text{int}}^{\text{el}}$, and exchange-correlation energy $E_{\text{int}}^{\text{xc}}$. Van der Waals interactions arise from density-density correlations and are part of the exchange-correlation contribution. However, long-range density-density correlations are not properly accounted for in

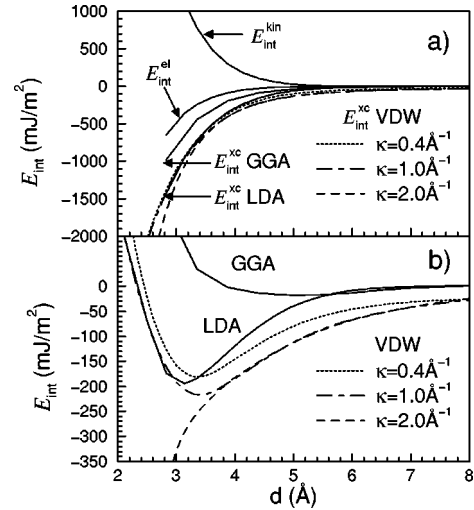


FIG. 3. The interaction energy of two Si(100)-H(2×1) slabs with displacement $\mathbf{R}=(0.5,0.5)$. (a) Solid lines show the interaction energy separated into kinetic energy $E_{\text{int}}^{\text{kin}}$, electrostatic energy $E_{\text{int}}^{\text{el}}$, and exchange-correlation energy $E_{\text{int}}^{\text{xc}}$ calculated within LDA and GGA. The dotted, dot-dashed, and dashed lines show the VDW approximation for $E_{\text{int}}^{\text{xc}}$ calculated with $\kappa=0.4$, 1, and 2\AA^{-1} , respectively. (b) Total interaction energies with the exchange-correlation part calculated using the GGA, LDA, and the three different VDW approximations.

local or semilocal approximations, like LDA or GGA. When a long-range exchange-correlation attraction is mimicked in these approximations,²⁴ it is often due to an unphysical polarization of the electron density. We, therefore, make a separation of the self-consistent density of the interacting system $n_{12}^{\text{sc}}=n_1+n_2+\Delta n_{12}$ where n_1 and n_2 are the densities of the single slabs, and Δn_{12} is the polarization. We define $n_{12}=n_1+n_2$ for the nonpolarized density of the interacting system.

In Fig. 2(b) we show the separate terms of the interacting energy of the two Si(100)-H(2×1) surfaces displaced by $\mathbf{R}=(0.5,0.5)$ calculated using the Harris functional²⁵ with n_{12}^{sc} (dashed lines) and n_{12} (solid lines) as input densities. The total energy is variational and the two density approximations give roughly the same total interaction energy. In the range $3\text{\AA} < d < 5.5\text{\AA}$, the values of $E_{\text{int}}^{\text{xc}}$ and $E_{\text{int}}^{\text{el}} + E_{\text{int}}^{\text{kin}}$ ($= E_{\text{int}} - E_{\text{int}}^{\text{xc}}$) are similar in the two approximations. The repulsion between the two surfaces is from the kinetic energy, and it arises mainly from the overlap between the occupied Si-H orbitals. The long-range attraction is from the exchange-correlation energy and thus due to van der Waals forces. For $d > 5.5\text{\AA}$, the exchange-correlation energy calculated using n_{12}^{sc} becomes positive and the kinetic energy becomes negative. This illustrates the fact that for $d > 5.5\text{\AA}$ the GGA approximation is no longer valid and the calculated polarization $\Delta n_{12}^{\text{sc}}$ is unphysical. At the wafer separations relevant for wafer bonding we expect that the polarization will be small, and in the following we approximate the density of the interacting system by n_{12} .

In Fig. 3(a) we show the separate terms calculated in the LDA and the GGA. Since LDA and GGA densities are almost identical, the two approximations give similar values for the kinetic and electrostatic energies, while values for $E_{\text{int}}^{\text{xc}}$ are very different. For example, defining the bonding energy,

TABLE I. The maximum/minimum bonding energy E_B of all different relative displacement of two Si(100)-H(2×1) and two Si(111)-H(1×1) surfaces, calculated within the LDA, GGA, and VDW ($\kappa=1 \text{ \AA}^{-1}$) approximations, and the separations of the wafers d_{eq} in the bonding geometry, as measured between the topmost Si layers. The relative displacement with minimum bonding energy is in both cases $\mathbf{R}=(0,0)$, while $\mathbf{R}=(0.5,0.5)$ is the displacement with maximum bonding energy. The two Si(111)-H(1×1) surfaces are rotated by 180° in the minimum bonding energy configuration. For the two Si(111)-H(1×1) surfaces we also list the change in the Si-H bond length Δz_H and the shift in the frequency of the asymmetric Si-H stretch mode $\Delta\omega_{\text{as}}$.

Si(100)-H(2×1)	LDA	GGA	VDW	Expt.
E_B (mJ/m ²)	198/57	18/16	217/94	10–30 ^{a-d}
d_{eq} (Å)	3.1/4.9	4.9/5.7	3.4/5.2	
Si(111)-H(1×1)	LDA	GGA	VDW	Expt.
E_B (mJ/m ²)	151/68	17/17	121/88	
d_{eq} (Å)	3.9/5.0	5.8/5.8	4.5/5.3	
Δz_H (0.0001 Å)	2/50	-1/4	-5/12	
$\Delta\omega_{\text{as}}$ (cm ⁻¹)	-12/-21	-2/-2	-5/-8	-17 ^e

^aReference 4.

^bReference 5.

^cReference 6.

^dReference 7.

^eReference 28.

E_B , as the minimum interaction energy for a given \mathbf{R} , we find for $\mathbf{R}=(0.5,0.5)$ that the LDA gives $E_B=200 \text{ mJ/m}^2$, and the GGA gives $E_B=20 \text{ mJ/m}^2$. Furthermore, the van der Waals attraction between two parallel surfaces is known to have a power law decay at large distances,^{26,13,14} while the interaction energy calculated within LDA or GGA has an exponential decay. It is therefore clear that neither approximation is usable for investigating the large d limit.

Recently, Refs. 12–14 have proposed van der Waals density functionals, with explicit expressions in the asymptotic limits. To separate out the long-range contribution, we follow Kohn *et al.*¹⁵ and divide the exchange-correlation energy into a short-range chemical part and a long-range van der Waals part. We calculate the short-range part using the exchange-correlation energy of a homogeneous electron gas with short-range electron-electron interaction $U_{\text{sr}}(r) = e^{-\kappa r}/r$ and call this approximation srLDA.²⁷ Inspired by Refs. 12–14, we propose a parametrization of the long-range contribution,

$$E_{\text{int}}^{\text{lrxc}}(d) = -C_2^{12} \left(\frac{1 - e^{-\kappa(d-2z_0)}}{d-2z_0} \right)^2, \quad (1)$$

where C_2^{12} is called the van der Waals coefficient and z_0 the van der Waals plane measured outwards relative to the first Si plane. This equation has the correct asymptotic form for the van der Waals attraction between two parallel surfaces,¹³ and we have added the exponential term in the numerator in order to account for the modified long-range Coulomb potential, $U_{\text{lr}}(r) = (1 - e^{-\kappa r})/r$. Andersson *et al.*¹³ have proposed a model for calculating asymptotic values for C_2^{12} and z_0 , using the average density perpendicular to the slab and the centroid of the surface charge induced by a static uniform external electric field oriented perpendicular to the surface. We have calculated the average density and the centroid of the Si(100)-H(2×1) surface within the LDA, and using

those values we obtain $C_2^{12}=0.064 \text{ eV}$ and $z_0=1.1 \text{ \AA}$. The corresponding values for the Si(111)-H(1×1) surface are $C_2^{12}=0.064 \text{ eV}$ and $z_0=1.0 \text{ \AA}$. The main assumptions in these calculations are the neglect of surface corrugation, the use of local jellium approximations for the dielectric tensor, and a cutoff.¹³

We now define an approximation, named VDW, for the exchange-correlation energy as the long-range part calculated using Eq. (1) and the short-range part calculated using the srLDA. The separation of exchange-correlation energy into these two fractions is determined by the value of κ . The value of κ must be such that the short-range part dominates at distances where chemical interactions are important, and the long-range part dominates at distances where van der Waals interactions are most important. To fulfill this we must have $\kappa \sim 1 \text{ \AA}^{-1}$. Furthermore, the asymptotic value of the VDW exchange-correlation energy at small separations is determined by the value of κ ($\lim_{d \rightarrow 2z_0} E_{\text{int}}^{\text{lrxc}} = \kappa^2 C_2^{12}$), and must approach the LDA energy which we assume to be accurate at such separations. In Fig. 3(a) we show the VDW exchange-correlation energy calculated with $\kappa=0.4, 1$, and 2 \AA^{-1} , and we see that the $\kappa=0.4$ and $\kappa=1 \text{ \AA}^{-1}$ approximations have the correct asymptotic behavior, while the $\kappa=2 \text{ \AA}^{-1}$ approximation is too large at small separations. For the $\kappa=1 \text{ \AA}^{-1}$ approximation the crossover from dominant short-range to long-range exchange-correlation takes place at $d \sim 3.5 \text{ \AA}$. The fact that the LDA and VDW results coincide in this regime shows that LDA mimics van der Waals forces rather well at medium distances. In Fig. 3(b) we show the total energies obtained by adding the kinetic and electrostatic energy to the exchange-correlation energies. The $\kappa=0.4$ and $\kappa=1 \text{ \AA}^{-1}$ approximations predict bonding energies similar to the LDA value, while the $\kappa=2 \text{ \AA}^{-1}$ approximation has no minimum.

In Table I, we summarize the results of calculations for

the bonding energy of two similar H-passivated Si surfaces using the LDA, GGA, and the VDW approximation with $\kappa = 1 \text{ \AA}^{-1}$. The table shows the maximum and minimum bonding energies obtained when varying the relative displacement of the two slabs. The VDW approximation gives bonding energies similar to the LDA values, while the GGA values are one order of magnitude smaller. In all three approximations the minimum bonding energy is obtained for the geometry where H atoms on the two surfaces are in registry, and the maximum bonding energies obtained in the geometry where H atoms are furthest apart. An arbitrary rotation of the two wafers relative to each other will in general result in a geometry somewhere between these two extreme situations, and we therefore expect the corresponding bonding energy to be intermediate to the values listed in Table I.

For the Si(111)-H(1×1) surfaces, we also list the change in the Si-H bond length Δz_H and the asymmetric Si-H stretch mode ω_{as} due to the presence of the other wafer at the equilibrium wafer-wafer separation. Similar to previous LDA calculations²⁸ we find a small outward relaxation of the H atoms and a softening of ω_{as} during bonding. To obtain the VDW result we calculated each term of the total energy as function of both d and z_H . At the equilibrium separation the outward relaxation of the H atoms is vanishing, however, the wafer interactions change the curvature of the H potential and thereby the frequency ω_{as} .

It is instructive to compare the theoretical results with experiment, bearing in mind that the experimental situation may be considerably more complex due to steps, multiple phases, adsorbed species, and elastic deformation of the crystals. The VDW approximation with $\kappa = 1 \text{ \AA}$ has the most correct description of the van der Waals interactions and therefore the best theoretical model. For the two Si(111)-H(1×1) surfaces, the VDW shift of $\Delta\omega_{as}$ is in qualitative agreement with the measured shift $\Delta\omega_{as} \approx 17 \text{ cm}^{-1}$.²⁸ For H-passivated Si(100) the experimental estimates of the bonding energies obtained by the crack opening method give $E_B = 10\text{--}30 \text{ mJ/m}^2$.⁴⁻⁷ We should compare this value with the average value for the different registries of two Si(100)-H(2×1) surfaces. The VDW approximation predicts an average bonding energy of $\approx 150 \text{ mJ/m}^2$. It is certainly reasonable that theory will tend to overestimate E_B , as corrugation effects such as steps and other defects will tend to lower the effective contact area, and hence the measured bonding energy. However, even taking such effects into account the VDW bonding energy seems too large compared with experiment. The GGA bonding energy is much smaller, suggesting that an improved description of the system can be obtained by making a VDW approximation based on a short-range GGA functional.

We now discuss the nature of the bonding forces. The attraction is from van der Waals surface-surface forces,

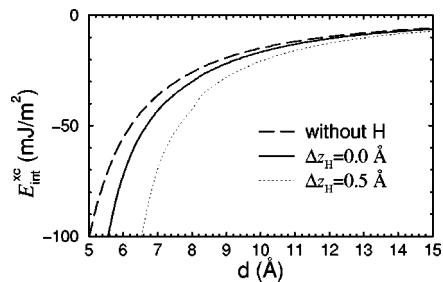


FIG. 4. The exchange correlation part of the interaction energy E_{int}^{xc} for two Si(111) surfaces without H passivation (dashed), with H passivation (solid), with H passivation where the Si-H bond length is stretched 0.5 \AA (dotted).

which are parametrized using z_0 and C_2^{12} . In Fig. 4, we show the attractive part of the interaction between two Si(111) surfaces for different H passivations. At large separations the attraction is determined mainly by C_2^{12} . This parameter is independent of the position of the H atoms and similar to C_2^{12} of a nonpassivated silicon surface. The asymptotic long-range attraction is therefore mainly from the Si substrate, as suggested by Fig. 4 and the GGA results of Fig. 2(a). However, the van der Waals plane z_0 does depend on the position of the H atoms. This gives rise to the vertical shift of the curves in Fig. 4 and is the origin of the outward relaxations of the H atoms. We emphasize that this effect is not due to direct H-H van der Waals interactions, since at the equilibrium wafer separations the van der Waals interaction (parametrized using values of Ref. 29) between two isolated H overlayers is much smaller than the wafer-wafer van der Waals interaction and cannot account for the bonding energy. Although we cannot exclude that other effects may contribute to the bonding energy, such as OH groups or F atoms adsorbed on the surfaces,^{8,9} our theoretical results show that such a hypothesis is not necessary to explain the measured bonding energy.

In conclusion, we have calculated the bonding energy between two H-passivated Si surfaces using the LDA, GGA, and an approach based on recent advances in the description of van der Waals forces in DFT. The calculations show that the force that leads to spontaneous bonding is a long-range van der Waals force between the Si samples, while the equilibrium separation of the wafers is determined by short-range kinetic energy repulsion between occupied Si-H orbitals.

We acknowledge F. Grey, A. P. Jauho, and K. Hermanson for many valuable discussions and careful reading of the manuscript. The use of national computer resources was supported by the Danish Research Councils.

¹S. Bengtsson and O. Engström, J. Appl. Phys. **66**, 1231 (1989).

²K. Ljungberg *et al.*, J. Electrochem. Soc. **141**, 2829 (1994).

³A. Nakagawa, K. Imamura, and K. Furukawa, Toshiba Review (Int. Ed.) **161**, 34 (1987).

⁴G. A. C. M. Spierings and J. Haisma, in *The Electrochemical*

Society Proceedings Series, edited by U. Gösele, J. Haisma, T. Abe, and M. A. Schmidt (Springer-Verlag, Berlin, 1992), Vol. PV 92-7, p. 18.

⁵Y. Bäcklund, K. Ljungberg, and A. Söderbärg, J. Micromech. Microeng. **2**, 158 (1992).

- ⁶K. Ljungberg, A. Söderbärg, and Y. Bäcklund, *Appl. Phys. Lett.* **62**, 1362 (1993).
- ⁷Q.-Y. Tong and U. Gösele, *J. Microelectromech. Syst.* **3**, 29 (1994).
- ⁸H. Himi, M. Matsui, S. Fujino, and T. Hattori, *Jpn. J. Appl. Phys., Part 1* **33**, 6 (1994).
- ⁹Q.-Y. Tong, E. Schmidt, U. Gösele, and M. Reiche, *Appl. Phys. Lett.* **64**, 625 (1994).
- ¹⁰J. P. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).
- ¹¹J. P. Perdew *et al.*, *Phys. Rev. B* **46**, 6671 (1992).
- ¹²E. Hult, Y. Andersson, B. I. Lundqvist, and D. C. Langreth, *Phys. Rev. Lett.* **77**, 2029 (1996).
- ¹³Y. Andersson, E. Hult, P. Apell, D. C. Langreth, and B. I. Lundqvist, *Solid State Commun.* **106**, 235 (1998).
- ¹⁴Y. Andersson, Ph.D. thesis, Chalmers University of Technology, S-412 96 Göteborg, Sweden, 1996.
- ¹⁵W. Kohn, Y. Meir, and D. E. Makarov, *Phys. Rev. Lett.* **80**, 4153 (1998).
- ¹⁶D. Vanderbilt, *Phys. Rev. B* **41**, 7892 (1990).
- ¹⁷*CRC Handbook of Chemistry and Physics*, 75th ed., edited by D. R. Lide (CRC Press, New York, 1994).
- ¹⁸E. Nielsen, Master thesis, Danish Technical University, Lyngby, 1997.
- ¹⁹P. Kratzer, B. Hammer, and J. K. Nørskov, *Phys. Rev. B* **51**, 13 432 (1995).
- ²⁰M. R. Radeke and E. A. Carter, *Phys. Rev. B* **54**, 11 803 (1996).
- ²¹J. E. Northrup, *Phys. Rev. B* **44**, 1419 (1991).
- ²²E. Kaxiras and J. D. Joannopoulos, *Phys. Rev. B* **37**, 8842 (1988).
- ²³X. P. Li and D. Vanderbilt, *Phys. Rev. Lett.* **69**, 2543 (1992).
- ²⁴N. D. Lang, *Phys. Rev. Lett.* **46**, 842 (1981).
- ²⁵J. Harris, *Phys. Rev. B* **31**, 1770 (1985).
- ²⁶N. G. van Kampen, B. R. Nijboer, and K. Schram, *Phys. Lett. A* **26**, 307 (1968).
- ²⁷A. Savin and H. Flad, *Int. J. Quantum Chem.* **56**, 327 (1995).
- ²⁸M. K. Weldon *et al.*, *J. Vac. Sci. Technol. B* **14**, 3095 (1996).
- ²⁹Y. Andersson, D. C. Langreth, and B. I. Lundqvist, *Phys. Rev. Lett.* **76**, 102 (1996).