11. Pfeffer, P. Effect of inversion asymmetry on the conduction subbands in GaAs-$Ga_{1-x}Al_xAs$ heterostructures. *Phys. Rev. B* **59,** 15902–15909 (1999).
12. Lommer, G., Malcher, F. & Rossler, U. Spin splitting in semiconductor heterostructures for B → 0. *Phys. Rev. Lett.* **60,** 728–731 (1988).
13. Das, B., Datta, S. & Reifenberger, R. Zero-field spin splitting in a two-dimensional electron gas. *Phys. Rev. B* **41,** 8278–8287 (1990).
14. Luo, J., Munekata, H., Fang, F. F. & Stiles, P. J. Effects of inversion asymmetry on electron energy band structures in GaSb/InAs/GaSb quantum wells. *Phys. Rev. B* **41,** 7685–7693 (1990).
15. Dresselhaus, P. D., Papavassiliou, C. M. A., Wheeler, R. G. & Sacks, R. N. Observation of spin precession in GaAs inversion layers using antilocalization. *Phys. Rev. Lett.* **68,** 106–109 (1992).
16. Knap, W. *et al.* Weak antilocalization and spin precession in quantum wells. *Phys. Rev. B* **53,** 3912–3924 (1996).
17. Jusserand, B., Richards, D., Peric, H. & Etienne, B. Zero-magnetic-field spin splitting in the GaAs conduction band from Raman scattering on modulation-doped quantum wells. *Phys. Rev. Lett.* **69,** 848–851 (1992).
18. Nitta, J., Akazaki, T., Takayanagi, H. & Enoki, T. Gate control of spin-orbit interaction in an inverted $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ heterostructure. *Phys. Rev. Lett.* **78,** 1335–1338 (1997).
19. Kalevich, V. K. & Korenov, V. L. Effect of electric field on the optical orientation of 2D electrons. *JETP Lett.* **52,** 230–235 (1990).
20. Gossard, A. C. Growth of microstructures by molecular beam epitaxy. *IEEE J. Quant. Electron.* **QE-22,** 1649–1655 (1986).
21. Crooker, S. A., Awschalom, D. D., Baumberg, J. J., Flack, F. & Samarth, N. Optical spin resonance and transverse spin relaxation in magnetic semiconductor quantum wells. *Phys. Rev. B* **56,** 7574–7588 (1997).
22. Kikkawa, J. M. & Awschalom, D. D. Lateral drag of spin coherence in gallium arsenide. *Nature* **397,** 139–141 (1999).
23. Kikkawa, J. M. & Awschalom, D. D. Resonant spin amplification in *n*-type GaAs. *Phys. Rev. Lett.* **80,** 4313–4316 (1998).
24. Flatté, M. E. & Byers, J. M. Spin diffusion in semiconductors. *Phys. Rev. Lett.* **84,** 4220–4223 (2000).
25. La Rocca, G. C., Kim, N. & Rodriguez, S. Effect of uniaxial stress on the electron spin resonance in zinc-blende semiconductors. *Phys. Rev. B* **38,** 7595–7601 (1988).
26. Rashba, E. I. & Sheka, V. I. Combinational resonance of zonal electrons in crystals having a zinc blende lattice. *Sov. Phys. Solid State* **3,** 1257–1267 (1961).
27. Rashba, E. I. & Efros, A. L. Orbital mechanisms of electron-spin manipulation by an electric field. *Phys. Rev. Lett.* **91,** 126405 (2003).
28. Kato, Y. *et al.* Gigahertz electron spin manipulation using voltage-controlled g-tensor modulation. *Science* **299,** 1201–1204 (2003).
29. Salis, G. *et al.* Electrical control of spin coherence in semiconductor nanostructures. *Nature* **414,** 619–622 (2001).
30. Gupta, J. A., Knobel, R., Samarth, N. & Awschalom, D. D. Ultrafast manipulation of electron spin coherence. *Science* **292,** 2458–2461 (2001).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements We thank A. M. Andrews, E. L. Hu, P. M. Petroff and J. S. Speck for discussions. This work was supported by the DARPA SPINS and QuIST programmes.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to D.D.A. (awsch@physics.ucsb.edu).

The interface between silicon and a high-*k* oxide

Clemens J. Först[1,2], Christopher R. Ashman[1], Karlheinz Schwarz[2] & Peter E. Blöchl[1]

[1]*Clausthal University of Technology, Institute for Theoretical Physics, Leibnitzstrasse 10, D-38678 Clausthal-Zellerfeld, Germany*
[2]*Vienna University of Technology, Institute for Materials Chemistry, Getreidemarkt 9/165-TC, A-1060 Vienna, Austria*

The ability of the semiconductor industry to continue scaling microelectronic devices to ever smaller dimensions (a trend known as Moore's Law[1]) is limited by quantum mechanical effects: as the thickness of conventional silicon dioxide (SiO_2) gate insulators is reduced to just a few atomic layers, electrons can tunnel directly through the films. Continued device scaling will therefore probably require the replacement of the insulator with high-dielectric-constant (high-*k*) oxides[2], to increase its thickness, thus preventing tunnelling currents while retaining the electronic properties of an ultrathin SiO_2 film. Ultimately, such insulators will require an atomically defined interface with silicon without an interfacial SiO_2 layer for optimal performance. Following the first reports of epitaxial growth of AO and ABO_3 compounds on silicon[3–7], the formation of an atomically abrupt crystalline interface between strontium titanate and silicon was demonstrated[8–10]. However, the atomic structure proposed for this interface is questionable because it requires silicon atoms that have coordinations rarely found elsewhere in nature. Here we describe first-principles calculations of the formation of the interface between silicon and strontium titanate and its atomic structure. Our study shows that atomic control of the interfacial structure by altering the chemical environment can dramatically improve the electronic properties of the interface to meet technological requirements. The interface structure and its chemistry may provide guidance for the selection process of other high-*k* gate oxides and for controlling their growth.

Interfacing a new oxide with silicon is a major challenge. A gate oxide has to fulfil a number of requirements in addition to intrinsic properties such as the high dielectric constant and low defect concentrations. The gate oxide must also be chemically stable in contact with silicon, it must have a sufficiently large injection barrier, and it must not have interface states in the bandgap of silicon. Oxides related to silicon dioxide, used in today's transistors, uniquely fulfil these requirements owing to their strong bonds to silicon and their flexible bonding network. They fail only because of their low dielectric constant. Retaining the same beneficial properties for high-*k* oxides has turned out to be very difficult. The first high-*k* oxides introduced technologically are likely to be amorphous oxides or silicates of Hf and Zr with an interfacial SiO_2 layer. By around 2010, however, the projected miniaturization of transistors will mean that an interfacial SiO_2 layer will no longer be tolerable and oxides with an atomically well-defined interface with silicon will be required.

Before discussing the formation of the interface, we need to review the clean (001) surface of silicon and describe its changes due to Sr adsorption. On the unreconstructed silicon surface the atoms form a square array. Owing to a lack of upper bonding partners, each atom has two singly occupied dangling bonds pointing out of the surface. Pairs of silicon atoms dimerize, using up one dangling bond per atom to form the dimer bond. This is called the 'dimer row' reconstruction. A second rearrangement leads to the 'buckled dimer' reconstruction, in which one atom of each dimer lifts up and the other shifts down, resulting in a buckled dimer. This buckling causes both electrons to localize in the upper silicon atom of a dimer, whereas the other silicon atom with the empty dangling bond prefers a more planar arrangement.

By depositing various amounts of Sr atoms onto the Si(001) surface we explored the structural complexity of Sr adlayers[11]. This enables us to attribute atomic structures to the periodicities observed experimentally[10,12]. Here we summarize only the findings relevant for the present topic: initially each Sr atom donates two electrons into the empty dangling bonds of the surface. As Sr is added, the dimer buckling vanishes because both dangling bonds of a Si dimer become filled with electrons. Similar to Ba (refs 13, 14), Sr first occupies the trough between the dimer rows, in the centre of four dimers. At a coverage of half a monolayer all positions in the trough are occupied and each dimer dangling bond is filled with two electrons. This (2×1) structure is the only Sr-covered surface without surface states in the bandgap of silicon. Therefore, it is a suitable building block for an interface without states in the gap, as required for device applications. The finding of a sizeable bandgap also explains why this surface is fairly resistant to oxidation[15].

Thus we found that the substrate surface can be chemically

saturated by half a monolayer of Sr. Such a surface is 'isoelectronic' to an H-terminated silicon surface. Hydrogen is known to be very effective at passivating silicon. In the following, we therefore refer to the Si surface covered with half a monolayer of Sr as the Sr-passivated substrate.

After having gained insight into the metal overlayers, we investigated the formation of an oxide layer. We start from the Sr-passivated substrate and simulate the deposition of one layer of SrO. During a heating cycle to 600 K this single oxide layer reconstructs significantly. However, after placing two or more layers of SrO or $SrTiO_3$ on top of the reconstructed SrO layer, the oxide layers crystallize into the perfect bulk structure. Thus we obtain an atomically abrupt interface between the silicon substrate and the high-*k* oxide. This interface structure, denoted A and shown in Fig. 1, corresponds to the Sr-passivated silicon surface matched to the nonpolar SrO layer of the oxide.

After the chemical saturation of the substrate surface with half a monolayer of Sr, the second-most-important property is the matching of the charge patterns of the oxide and the Sr-passivated substrate surfaces joined at the interface. Whereas the SiO_2/Si interface relies on strong covalent bonds across the interface and a flexible bond network of the oxide, the interface described here is based on the chemical saturation of the silicon surface with an alkaline earth metal, so that a template for the deposition of a matching oxide is obtained.

In a device the interface is exposed to a number of chemical influences that affect the stability of the stack. The most critical question is the stability of the interface with respect to oxidation. Oxygen ions can diffuse out of the gate oxide to the interface.

To explore how the interface changes upon oxidation, we have added oxygen atoms to a wide range of different sites. Oxygen first attacks the surface silicon atoms at their vacant coordination sites. After introducing one monolayer of oxygen into the interface all these sites are consumed. The resulting structure, denoted B, is shown in Fig. 1. Additional oxygen atoms, up to a total oxygen content of 1.5 monolayers, insert into the dimer bonds. As explained below, structure B and the dimer-oxidized variant of structure B are the optimum choices for device applications.

Figure 2 illustrates the phase boundaries of the interfaces as a function of the oxygen chemical potential. The chemical potential is defined as the energy required to add a single oxygen atom to the system. It is the driving force for oxidation, which can be controlled externally, for example, by choosing the appropriate temperature

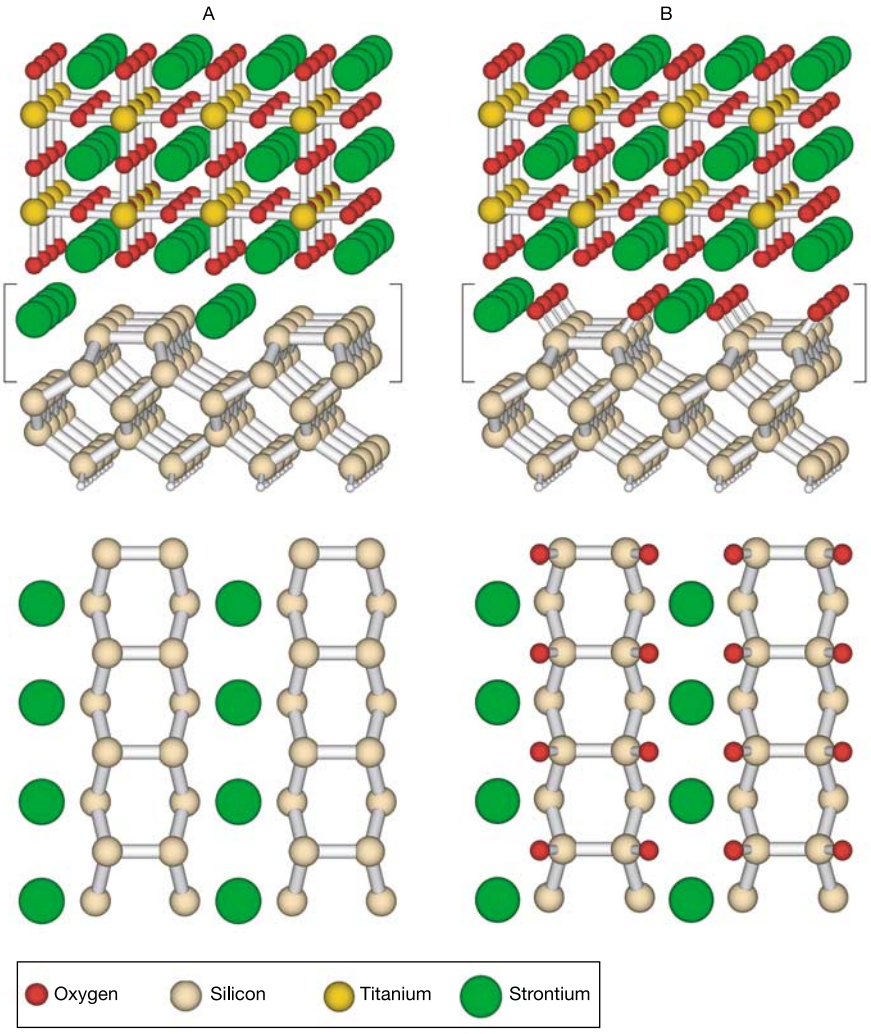

Figure 1 Atomic structures of the $SrTiO_3$/Si(001) interfaces: structure A (on left), unoxidized interface; structure B (on right), oxidized interface. Top, view slightly off the [110] direction of silicon, which is parallel to the [100] direction of the oxide. The topmost layer corresponds to the oxide surface of our slab calculation. Bottom, view along the interface normal of the interface layer.

and oxygen partial pressure during growth. In thermal equilibrium the chemical potential is related to the partial pressure and temperature of the growth chamber. It should be noted, however, that the formation of the stack mostly involves non-equilibrium processes, so that the chemical potential at the interface lags behind the value reached in the growth chamber. The oxygen chemical potential needs to be sufficiently low to avoid formation of an interfacial SiO_2 layer but high enough to avoid a large oxygen vacancy concentration in the gate oxide, which may cause trap-assisted leakage currents.

Our results indicate that structures A and B can be formed selectively in the absence of an interfacial SiO_2 layer. The processing window—the stability region of interface B—extends from −0.24 to 0.05 eV and corresponds to a range of partial pressures of nearly three orders of magnitude at 1,000 K. This region lies almost entirely below the coexistence line of Si and SiO_2.

In addition, we find indications of sizeable thin-film effects, which delay the formation of an interfacial SiO_2 layer. This can be inferred from the fact that the oxidation of all dimer bonds requires a chemical potential of 0.19 eV, and subsurface oxidation starting from interfaces A or B requires energies in excess of 0.91 eV above the coexistence line of Si and SiO_2.

The interface structure proposed in this work is quite different and much simpler than previously assumed: it was believed that an interfacial silicide[8,9] or silicate[16,17] layer must be formed. Such a layer is not present in the interface that emerged from our simulations. Nevertheless, our simulated $SrTiO_3$/Si(001) interfaces reproduce the undisputed features of the *Z*-contrast images of ref. 8, such as the pattern of interfacial Sr atoms and the oxide-substrate registry. Thus the lateral alignment of the columns of Sr and Ti atoms relative to the Si substrate and the (2×1) periodicity can clearly be identified.

We also performed calculations on the interfaces proposed by McKee and co-workers[8,9,10] and by Wang and co-workers[16,17]. As also shown previously[18], the former interface[8,9] is metallic, which is detrimental for device applications. The same applies for the latter structural proposal[16,17], which reconstructs significantly upon relaxation.

On the basis of electron-count arguments, Robertson and Peacock recently proposed a structure[18] that is related to our dimer-oxidized variant of structure B. It differs in that the oxide starts with the TiO_2 layer instead of a SrO layer and it is derived from a $c(2 \times 2)$ dimer reconstruction of the silicon surface. If we modify our interface by terminating the oxide with a TiO_2 layer instead of a SrO layer, it is more stable than Robertson's proposal by 0.19 eV per (1×1) surface unit cell. Having an oxide terminated by SrO is, however, favourable compared to TiO_2 terminated oxides, because the TiO_2 layer and the substrate in direct contact are expected to react, as pointed out[18]. Conceptually similar interfaces with a TiO_2 interface layer have been investigated[19], ruling out their use in devices on the basis of their electronic properties.

A critical parameter for gate stacks is the injection barrier, which is the offset between the conduction band edges of the silicon substrate and the oxide. It prevents electrons from entering the oxide conduction band, where they can cross the gate oxide. For device applications the injection barrier should be larger than 1 eV (ref. 3). There are indications[20,21] that the injection barriers for most high-*k* oxides are too low.

Before we discuss our results on the band offsets we need to briefly touch upon the bandgap problem of density functional theory (DFT)[22,23]: typically the one-particle energies obtained in these calculations underestimate the bandgap. Therefore, there is an uncertainty in our calculated injection barriers. Assuming that the error of the valence band edge is negligible, as required in exact DFT, and using the experimental bandgaps of silicon and $SrTiO_3$ (ref. 24), we anticipate that our calculations underestimate the injection barrier by 0.7–0.8 eV.

For interface A we obtain an injection barrier that is negative by 0.6 eV. Including our correction, we estimate the injection barrier to lie at 0.1–0.2 eV. The injection barrier lies below the technologically required minimum.

For interface B, however, we obtain a positive injection barrier of 0.5 eV. Adding the correction, our final estimate yields a positive injection barrier of 1.2–1.3 eV, which fulfils the criterion. The margin is sufficiently large that a reasonable error in the band-offset correction does not lead to an unacceptably low injection barrier. Most important is the ability to influence the injection barrier by carefully choosing the processing conditions.

Note that band offsets are frequently derived from the properties of the two bulk materials alone[21], disregarding the interface structure and composition in detail. The interface between silicon and $SrTiO_3$ is an example where the band offset can be engineered by controlling the chemical environment. The change of the band offset due to oxidation is about 1.1 eV, and is thus sizeable. It results from a dipole created when the electrons are transferred from the filled dangling bonds of the surface silicon atoms to the oxygen atoms that attach to the vacant coordination sites.

The injection barrier of interface B is fairly insensitive to additional oxidation of the dimer bonds. An amorphous interfacial SiO_2 layer is likely to destroy these restrictions and thus lead to a lower injection barrier. □

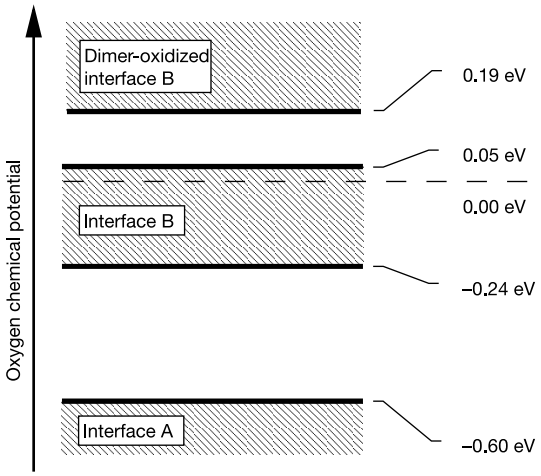

Figure 2 Phase diagram for interface oxidation. Shaded areas indicate the stability regions of the defect-free interfaces A and B and the dimer-oxidized variant of interface B. The blank regions separating them correspond to disordered structures with an oxygen content that increases with increasing chemical potential. The external parameter is the oxygen chemical potential. The zero of the chemical potential corresponds to the coexistence of bulk Si and SiO_2 (α-quartz) in thermal equilibrium.

Methods

We performed state-of-the-art electronic structure calculations and *ab initio* molecular dynamics simulations[25] based on density functional theory[22,23,26] and the projector augmented wave method[27].

The calculations have been done on 5-layer slabs of silicon. The slab calculations included a vacuum region of at least 6 Å between repeated slabs. The relevant calculations of the interfaces are done in a (2×2) supercell. All structures are relaxed without symmetry constraints. The hydrogen-terminated silicon back plane has been kept frozen.

Our calculations used a plane wave cutoff of 30 Ry for the plane wave part of the wave function. We used the frozen-core approximation. Semi-core states of Sr and Ti have, however, been treated as valence electrons. We used the following sets of projector functions per angular momentum: 2*s*2*p*1*d* for oxygen, 2*s*2*p*1*d* for silicon, 3*s*2*p*2*d* for strontium and 2*s*2*p*2*d* for titanium.

For all calculations of Sr adsorption we used a grid with about 64 lateral *k*-points per

(1 × 1) surface unit cell. For the interfaces we used a grid corresponding to 16 *k*-points per (1 × 1) unit cell. For metallic structures we used the Mermin functional with a temperature of 1,000 K and the extrapolation to zero Kelvin proposed in ref. 28.

The band offsets have been derived by relating the plane wave part of the potential, averaged laterally, to band edges. The relative displacement between potential and band edges has been obtained from the epitaxially strained bulk materials.

The phase diagram in Fig. 2 was obtained from the total energies as a function of the oxygen chemical potential. A number of different, stoichiometric and non-stoichiometric, interface structures with varying oxygen content have been considered. The regions where one of the stoichiometric interfaces (A, B or the dimer-oxidized variant of B) is most stable are shaded. Tests for the upper boundary of interface A with larger, that is (4 × 4), unit cells confirmed that the phase boundaries obtained in (2 × 2) unit cell are reliable to about 0.03 eV.

Received 28 February; accepted 10 November 2003; doi:10.1038/nature02204.

1. Moore, G. E. Cramming more components onto integrated circuits. *Electronics* **38,** 114–117 (1965).
2. Semiconductor Industry Association. *International Technology Roadmap for Semiconductors* at ⟨http://public.itrs.net/⟩ (International SEMATECH, Austin, TX, 2001).
3. Ishiwara, H. & Azuma, K. Oriented growth of $SrTiO_3$ films on Si(100) substrates using in-situ cleaning by excited hydrogen. *Mater. Res. Soc.* **116,** 369–374 (1988).
4. Kado, Y. & Arita, Y. Heteroepitaxial growth of SrO films on Si substrates. *J. Appl. Phys.* **61,** 2398–2400 (1987).
5. Mori, H. & Ishiwara, H. Epitaxial growth of $SrTiO_3$ films on Si(001) substrates using a focused electron beam evaporation method. *J. Appl. Phys.* **30,** 1415–1417 (1991).
6. Kim, T. W. *et al.* Structural properties and interfacial layer formation mechanisms of $PbTiO_3$ thin films grown on p-Si substrates. *Appl. Phys. Lett.* **64,** 2676–2678 (1994).
7. Alexe, M. Measurement of interface trap states in metal ferroelectric silicon heterostructures. *Appl. Phys. Lett.* **72,** 2283–2285 (1998).
8. McKee, R. A., Walker, F. J. & Chisholm, M. F. Crystalline oxides on silicon: the first five monolayers. *Phys. Rev. Lett.* **81,** 3014–3017 (1998).
9. McKee, R.A. & Walker, F.J. Process for depositing an oxide epitaxially onto a silicon substrate and structures prepared with the process. US patent 5,225,031 (United States Patent Office, 1993).
10. McKee, R. A., Walker, F. J. & Chisholm, M. F. Physical structure and inversion charge at a semiconductor interface with a crystalline oxide. *Science* **293,** 468–471 (2001).
11. Ashman, C., Först, C. J., Schwarz, K. & Blöchl, P. E. First-principles calculations of strontium adsorption on Si(001). *Phys. Rev. B* **69** (in the press).
12. Fan, W. C., Wu, N. J. & Ignatiev, A. Observation of ordered structures of Sr on the Si(001) surface. *Phys. Rev. B* **42,** 1254–1257 (1990).
13. Yao, X. *et al.* Initial stages of Ba adsorption on the Si(100)-(2 × 1) surface at room temperature. *Phys. Rev. B* **59,** 5115–5119 (1999).
14. Wang, J. *et al.* Bonding and diffusion of Ba on a Si(001) reconstructed surface. *Phys. Rev. B* **60,** 4968–4971 (1999).
15. Liang, Y., Gan, S. & Engelhard, M. First step towards the growth of single-crystal oxides on Si: formation of a two-dimensional crystalline silicate on Si(001). *Appl. Phys. Lett.* **79,** 3591–3593 (2001).
16. Wang, J., Ooms, W. J. & Hallmark, J. A. Semiconductor structure having a crystalline alkaline earth metal oxide interface with silicon. US patent 6,238,459 (United States Patent Office, 2001).
17. Droopad, R. *et al.* Epitaxial oxides on silicon grown by molecular beam epitaxy. *J. Cryst. Growth* **227–228,** 936–943 (2001).
18. Robertson, J. & Peacock, P. W. Atomic structure, band offsets and hydrogen in high k oxide:silicon interfaces. *Mater. Res. Soc. Symp. Proc.* **747,** 99–111 (2002).
19. Zhang, X., Demkov, A. A., Li, H., Hu, X. & Wei, Y. Atomic and electronic structure of the $Si/SrTiO_3$ interface. *Phys. Rev. B* **68,** 125323 (2003).
20. Chambers, S. A. *et al.* Band offset and structure of $SrTiO_3$/Si(001) heterojunctions. *J. Vac. Sci. Technol. A* **19,** 934–939 (2001).
21. Robertson, J. Band offsets of wide-band-gap oxides and implications for future electronic devices. *J. Vac. Sci. Technol. B* **18,** 1785–1791 (2000).
22. Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Phys. Rev. B* **136,** 864–871 (1964).
23. Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev. A* **140,** 1133–1138 (1965).
24. Noland, T. A. Optical absorption of single-crystal strontium titanate. *Phys. Rev.* **94,** 724 (1954).
25. Car, R. & Parrinello, M. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* **55,** 2471–2474 (1985).
26. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77,** 3865–3868 (1996).
27. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50,** 17953–17979 (1994).
28. Gillan, M. G. Calculation of the vacancy formation energy in aluminum. *J. Phys. Condens. Matter* **1,** 689–711 (1989).

Acknowledgements We thank S. Chambers, M. Chisholm, W. Daum, A. Dimoulas, J. Fompeyrine, J.-P. Loquet, R.A. McKee, G. Norga and S. Stemmer for discussions. This work has been funded by the European Commission in the project 'Integration of Very High-k Dielectrics with CMOS Technology' (INVEST) and by the AURORA project of the Austrian Science Fund. Parts of the calculations have been performed on the computers of the 'Norddeutscher Verbund für Hoch- und Höchstleistungsrechnen' (HLRN).

Competing interests statement The authors declare competing financial interests: details accompany the paper on **www.nature.com/nature**.

Correspondence and requests for materials should be addressed to P.E.B. (peter.bloechl@tu-clausthal.de).

High-latitude controls of thermocline nutrients and low latitude biological productivity

J. L. Sarmiento[1], N. Gruber[2], M. A. Brzezinski[3] & J. P. Dunne[4]

[1]*Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, New Jersey 08544, USA*
[2]*IGPP and Department of Atmospheric Sciences, University of California at Los Angeles, Los Angeles, California 90095, USA*
[3]*Department of Ecology, Evolution and Marine Biology and the Marine Science Institute, University of California, Santa Barbara, California 93106, USA*
[4]*NOAA/Geophysical Fluid Dynamics Laboratory, PO Box 308, Forrestal Campus B Site, Princeton, New Jersey 08542, USA*

The ocean's biological pump strips nutrients out of the surface waters and exports them into the thermocline and deep waters. If there were no return path of nutrients from deep waters, the biological pump would eventually deplete the surface waters and thermocline of nutrients; surface biological productivity would plummet. Here we make use of the combined distributions of silicic acid and nitrate to trace the main nutrient return path from deep waters by upwelling in the Southern Ocean[1] and subsequent entrainment into subantarctic mode water. We show that the subantarctic mode water, which spreads throughout the entire Southern Hemisphere[2,3] and North Atlantic Ocean[3], is the main source of nutrients for the thermocline. We also find that an additional return path exists in the northwest corner of the Pacific Ocean, where enhanced vertical mixing, perhaps driven by tides[4], brings abyssal nutrients to the surface and supplies them to the thermocline of the North Pacific. Our analysis has important implications for our understanding of large-scale controls on the nature and magnitude of low-latitude biological productivity and its sensitivity to climate change.

The classical explanation for the observed nutrient distribution of the ocean in the low latitudes is that the downward flux of biogenic material from the surface of the ocean is balanced by upwelling of dissolved inorganic nutrients driven by vertical mixing in the main thermocline. This essentially one-dimensional view grew out of early theories of thermocline and thermohaline circulation that are no longer tenable. Estimates of the magnitude of vertical mixing in the main thermocline are about an order of magnitude smaller than required to explain the vertical profiles of tracers within this feature[5]. In addition, ocean model simulations of radiocarbon distribution show that balancing the formation of deep waters by upwelling through the main thermocline gives results that are inconsistent with observations[1]. Instead, these studies suggest that a more likely return path for the deep water to the surface is in the Southern Ocean[1].

Subantarctic Mode Water (SAMW) has been identified as the main conduit of nutrients from the Southern Ocean to the upwelling regions of the equatorial Pacific and off South America[6]. The SAMW is a pycnostad (a layer of relatively uniform density) that originates in the thick wintertime mixed layers that ring the Southern Ocean[7] (Fig. 1d). This belt, which is particularly strong eastward of the central Indian Ocean to the western South Atlantic, coincides with the Subantarctic Zone (SAZ) between the Subtropical Front at about 40–45° S and the Subantarctic Front at about 45–55° S, and appears to also include the Polar Front Zone (PFZ) between the Subantarctic Front and Polar Front just to the south (Fig. 1d). The SAMW pycnostad increases in density from $\sigma_\theta = 26.5$ (equivalent to 1,026.5 kg m^{-3}) in the western Atlantic to $\sigma_\theta = 27.1$ in the southeast Pacific[8] as it flows in an eastward circuit around the Southern Ocean.