58.6 kD." In other words, the data on genes appear to corroborate the pattern derived from proteins.

"The most straightforward interpretation of these results is that a basic structural unit of about 14 kD exists," concludes Savageau. One explanation, he says, is that the modern population of proteins derived from an ancestral 14-kD polypeptide by a combination of divergence and duplication. In which case, one might expect a degree of homology to be detectable among the modern sequences. There are no convincing data to support this suggestion, says Savageau.

A second possible explanation is that "the clustering of sizes could be the result of selective pressures operating at the DNA or protein levels." The obvious question then is, What determines this particular size? And the obvious implication is that "a strongly conserved structural unit for proteins would have important implications for evolution." For instance, Savageau suggests that modification of amino acids within the structural unit would constitute microevolutionary change. Moreover, modifications that pushed the protein beyond the bounds of unit stability might trigger a macroevolutionary leap: the original polypeptide becomes a dimer of two 14-kD units, and so on.

Petsko and his students have scrutinized Savageau's data and, says Petsko, conclude that "he is clearly seeing something." However, they say that the effect may be marginal, although it is too early to be certain about this. A more rigorous statistical analysis, including Fourier transform, on a wider set of protein data is necessary to determine how robust and general the periodicity is.

The most likely effect at work here, says Petsko, is a preferred surface area to volume ratio. There is some indication that globular proteins do form spherical units composed of about 130 amino acids, which represents units of 14-kD molecular mass. "This might be telling you something about the way that protein chains prefer to fold, other things being equal," says Petsko. But proteins can also form extensive beta-sheets and alpha-helices, which would have the same effect. This would explain why the 14-kD periodicity does not stand out as sharply as it might otherwise have done.

Although Petsko considers that the significance of Savageau's observation "might not be as great as he implies," he firmly agrees that it is extremely interesting and worthy of further scrutiny. ■ ROGER LEWIN

ADDITIONAL READING

M. S. Savageau, "Proteins of *Escherichia coli* come in sizes that are multiples of 14 kDa: Domain concepts and evolutionary implications," *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1198 (1986).

A Silicon Solution for Gallium Arsenide IC's

Epitaxial growth of crystalline gallium arsenide layers on silicon wafers could combine the best properties of both semiconductors in future high-speed microelectronic chips

GALLIUM arsenide transistors switch faster than those made of silicon, a distinct advantage as the speed-conscious, high-tech world of microelectronics looks to future generations of integrated circuits. Moreover, gallium arsenide and closely related compounds efficiently emit near-infrared and visible light, another plus as interest grows in combining electrical and optical functions on the same chip. In almost every other respect, however, gallium arsenide is inferior to silicon. In particular, materials scientists and electrical engineers have yet to master the more complicated gallium arsenide technology sufficiently to make high-quality, low-cost material that is competitive with silicon.

If the excitement demonstrated at the Spring Meeting of the Materials Research Society last month is any indication, epitaxial growth of crystalline gallium arsenide layers on silicon substrates may not only boost the fortunes of gallium arsenide but also combine the best of both worlds.* An overflow crowd forced meeting organizers to switch the symposium titled "Heteroepitaxy on Silicon Technology" to a larger room that had been scheduled for less well-attended sessions on silicon integrated circuits. "We made our plans 6 months ago and didn't realize how fast interest was growing," explained cochairman of the heteroepitaxy symposium, John C. C. Fan, a former M.I.T. Lincoln Laboratory researcher who has founded the Kopin Corporation, a new firm in Taunton, Massachusetts, that specializes in thin-film, next-generation integrated circuit concepts, including gallium arsenide on silicon.

At the meeting, researchers from the United States and Japan reported that they can now make heteroepitaxial gallium arsenide layers on silicon whose quality is nearly on a par with that of conventional gallium arsenide. They also described individual transistors of several types (metal-semiconductor field effect, heterojunction bipolar, and modulation doped) with electrical properties comparable to those of existing devices. But light sources have more demanding requirements for materials quality and fabrication skill and, while demonstrating potential for the future, their performance is still well below the state of the art. Despite the progress, however, fundamental understanding of what takes place during the gallium arsenide growth process lags empirically developed methods for making high-quality material.

If gallium arsenide is a problem by itself, how can the added complication of mating it with silicon help? Why muck up two materials was the complaint not so long ago, when the field was new. Part of the motivation for taking on the extra complexity lies in the better mechanical and thermal properties of silicon as compared to those of gallium arsenide. In short, silicon is lighter and less brittle, which makes it easier to handle without breakage during the multistage integrated circuit fabrication process, a crucial factor in determining economic viability.

Moreover, partly because the higher thermal conductivity of silicon results in a more uniform removal of heat from solidifying crystals, engineers can grow silicon crystals 6 inches in diameter and are heading toward 8 and 10 inches, while the maximum for gallium arsenide is currently 3 inches, another important cost-lowering factor. The higher thermal conductivity also means faster dissipation of the heat generated by transistors, which permits more devices to be crammed onto a chip.

In an evening panel discussion on the future of gallium arsenide on silicon, Richard Koyama of TriQuint Semiconductor, a Beaverton, Oregon, manufacturer of gallium arsenide integrated circuits, summed up these benefits. Koyama argued that before long the highest performance commercial gallium arsenide integrated circuits would be of the type requiring fabrication in epitaxial layers grown on the wafer that is sliced from a crystal rather than directly on the

*1986 Spring Meeting of the Materials Research Society, Palo Alto, California, 15–19 April 1986. Proceedings of Symposium A, "Heteroepitaxy on Si Technology," are to be published as volume 67 of the Materials Research Society Symposia Proceedings, Pittsburgh, Pennsylvania.

wafer as at present. Noting the dominating role of economics in the adoption of new chip technologies, he concluded that manufacturers would prefer silicon to gallium arsenide wafers when the switchover takes place.

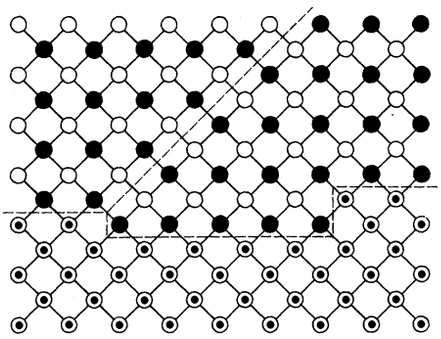
This does not imply that the road is open for gallium arsenide on silicon simply to replace gallium arsenide in existing applications and take over from there. Another panel member, Masahiro Akiyama from the Oki Electric Company Limited in Tokyo, argued that, despite the advantages of silicon wafers, gallium arsenide on silicon chips would have to offer something more than gallium arsenide alone.

An example of a new capability offered by gallium arsenide on silicon that was repeatedly mentioned in talks by Lincoln Laboratory researchers is the integration on a single chip of transistors and light sources. The motivation is the increasingly congested communication path between chips as the transistor count per chip rises into the millions. Optical communication, which is inherently a high-data-rate, interference-free process, is a proposed solution. In one scenario, a gallium arsenide light source, such as a laser diode or a light-emitting diode, that is driven by electrical signals from circuits on an otherwise all-silicon chip sends an optical signal through an optical fiber to a gallium arsenide or silicon photo-detector on a neighboring chip, where it is converted back to electrical form.

It is also possible to mix silicon and gallium arsenide transistors on the same chip, a small section of which might be dedicated to gallium arsenide for high-speed signal processing, while silicon devices attend to more mundane affairs. These are just some of the options that the three-dimensional materials engineering technique of heteroepitaxy makes possible.

In an invited talk leading off the symposium Herbert Kroemer of the University of California at Santa Barbara reviewed some of the problems that have plagued gallium arsenide on silicon in the past and some that remain to be solved. The problems arise because (i) gallium arsenide is a partially ionic (polar) semiconductor while silicon is covalent (nonpolar), (ii) the lattice constant of gallium arsenide is about 4% larger than that of silicon, and (iii) the thermal expansion coefficients of the two materials are not the same. Panelists at the evening session mostly agreed that it will be at least 2 years before even simple commercial products appear, provided that the remaining problems are overcome.

To appreciate the difficulty of growing defect-free polar epitaxial layers on a nonpolar substrate, first consider the crystal structures of silicon and gallium arsenide. Silicon has a diamond cubic structure in which each silicon atom is surrounded by four equally spaced nearest neighbors, giving the characteristic covalent tetrahedral bonding geometry. The lower part of the drawing shows a projection of this structure along one of the edges of the cube. Gallium arsenide has the same geometry except that each gallium is surrounded by four arsenics and each arsenic by four galliums, as in the upper half of the drawing.



GaAs/Si interface

This cross-section view of the (100) interface between silicon (bottom) and gallium arsenide (top) shows that a single surface step generates an antiphase domain (delineated by the diagonal dashed line), while a double step does not. The white, black, and dotted circles represent gallium, arsenic, and silicon atoms.

So far, so good because it would seem that the gallium arsenide could grow by simply taking over where the silicon leaves off. However, gallium and arsenic atoms more or less arrive randomly at the surface, but the first layer on the silicon must be either all gallium or all arsenic. If it is not, so-called antiphase domains form. In the domain, each gallium and arsenic maintains the correct bonding to its neighbors, but across the antiphase domain boundary the bonds are between gallium atoms and between arsenic atoms. It is not known exactly why antiphase domains are harmful, but researchers find that the best-quality material is most free of these defects.

Researchers grow the epitaxial layers by one of two techniques, molecular beam epitaxy (a research-oriented, ultrahigh-vacuum technique) and metal-organic chemical vapor deposition. In either case, it is possible to put down a monolayer of pure arsenic or gallium before commencing gallium arsenide growth and thereby to avoid the formation of antiphase domains, although another effect suggests the matter ought not to be so simple.

The idea is that crystal surfaces are seldom atomically flat; instead, they comprise flat terraces joined by sharp steps one or more atomic layers in height. The drawing shows two such steps, a single and a double step on a silicon (100) surface; that is, a surface corresponding to a face of the cubic unit cell. In the presence of a single step, a monolayer of arsenic or gallium on the surface does not prevent antiphase domain formation, whereas the double step prevents their appearance.

The unavoidable presence of steps suggests that growing defect-free gallium arsenide on (100) silicon surfaces (the orientation silicon device makers prefer and hence, most readily available from commercial suppliers) should be difficult. Although Kroemer and Hadis Morkoç of the University of Illinois, among others, have proposed models for the experimentally verified absence of antiphase domains even when gallium or arsenic pre-layers are not used, there is little evidence for their ideas at the moment. This is one example of the gap between fundamental understanding and empirically derived methods.

Kroemer, who advocates the use of wafers having a (211) orientation because antiphase domains cannot occur on this surface, proclaimed the problem of antiphase domains to be largely solved for either orientation. An issue that is yet to be fully worked out is that of the formation of defects called dislocations at the silicon–gallium arsenide interface. Dislocations arise because of the mismatch between the lattice parameters of the two materials. With a 4% mismatch, for every 25 rows of silicon atoms intersecting the interface, there are only 24 rows in the gallium arsenide. For epitaxial growth, the rows must line up across the interface (see photograph). So, the mismatch is concentrated in a dislocation roughly once every 25 rows or about 10^{12} dislocations per square centimeter of surface.

In the simplest case, a pure edge dislocation in the jargon, the distortion this entails is confined to the interface region. Typical thicknesses of gallium arsenide layers on silicon wafers are in the range of a few micrometers so these dislocations do not adversely affect the performance of transistors, laser diodes, and so on that reside in the upper portion of the layer. However, other types of dislocations bend upward and propagate through the gallium arsenide to the top surface where they wreak havoc if present in large numbers.

Until recently, the dislocation density at the top surface has been in the range of 10^8 to 10^9 per square centimeter, which is far too high. At the meeting, Morkoç, Akiyama, and Shiro Sakai of the Nagoya Institute of Technology in Japan described gallium arsenide on silicon made in their laboratories with only a few thousand dislo-

cations per square centimeter, which is "getting there," according to John Poate of AT&T Bell Laboratories, the symposium's other cochairman. These reports were somewhat controversial, however, partly because different methods of measuring the dislocation density give different answers even at high densities for which they are considered to be reliable, and partly because protocols for low densities are yet to be agreed upon.

To get an idea for how scientifically guided intuition and empirical methods combine in the successful growth of gallium arsenide on silicon, consider the technique the Illinois group uses to minimize dislocations, an extension of the well-known trick of using silicon wafers that are slightly misoriented; that is, they are not quite (100) surfaces. Interestingly, use of off-oriented (100) wafers was originally thought of as a way to minimize antiphase domains.

The explanation proposed for the role of the misorientation involves generating predominantly pure edge dislocations at the silicon–gallium arsenide interface. An atomic-resolution transmission electron microscopy study reported by Nobuo Otsuko of Purdue University, who collaborated with researchers from Purdue, Illinois, and the Tokyo Institute of Technology, showed that the type of dislocations formed strongly depended on orientation. In particular, edge dislocations prefer to run parallel to steps on the surface.

Accordingly, the orientation used by the Illinois group requires, first, that surface steps occur roughly as often as dislocations do (that is, at least every 25 lattice spacings) to encourage the formation of edge dislocations at every site where dislocations are required by the lattice mismatch. And, second, the steps must run in two perpendicular directions on the surface. If steps ran in only one direction, they would line up with the dislocations required to accommodate the lattice mismatch in only one direction, whereas mismatch and the associated dislocation generation occurs in both directions.

Whether this explanation is correct remains to be verified. Other researchers, including Jhang Lee of Texas Instruments, reported attempts to find the optimum combination of wafer orientation and growth temperature that suggested that good material is obtainable in a variety of ways. Although most of the investigators, including Lee, advocate off-oriented (100) substrates, it appears possible to succeed with wafers that are not deliberately misoriented at all.

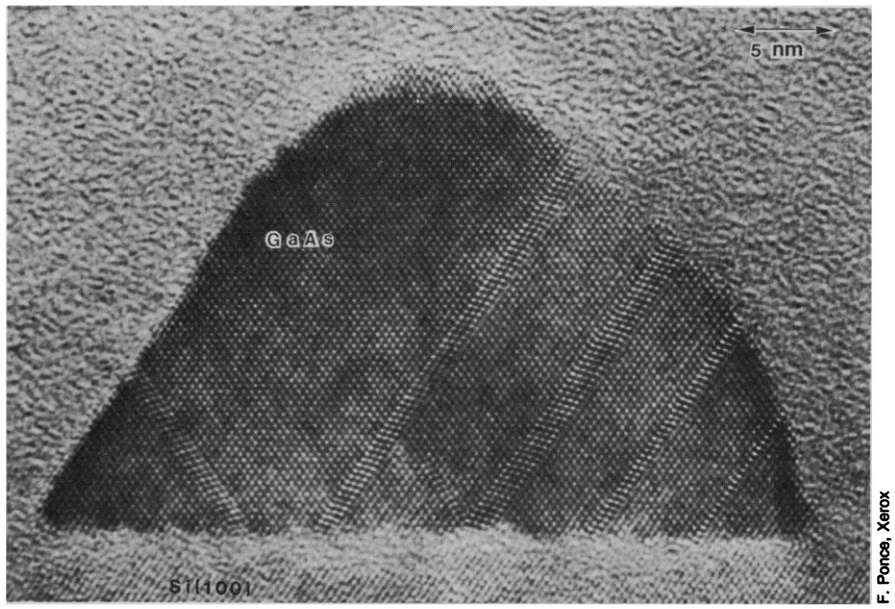
Finally, the different thermal expansion coefficients of silicon and gallium arsenide mean that, during cooling from the growth temperature of several hundred degrees Celsius, enormous stresses are generated in the materials. In particular, since the thermal expansion coefficient of gallium arsenide is the greater of the two, the epitaxial layer wants to contract the more during cooling. As discussed by Sakai, for a thin layer on a thick wafer, the wafer wins, as it were, and remains flat, but the dislocations penetrate throughout the gallium arsenide. For a thick layer, the wafer loses and bows, but the dislocations in the gallium arsenide are confined to the lower half of the layer, where they are less harmful. Unresolved issues include whether bowing always occurs and, if so, how harmful it is. Current commercial specifications list wafer flatness as an important item.

Sakai's report also deals with the issue of the usefulness of a buffer layer between the silicon and the gallium arsenide. The lower portion of a thick gallium arsenide layer in effect acts as a buffer because most of the defects generated by the lattice mismatch are concentrated there. Too thick a layer is economically unattractive because it slows the production process. Another alternative is to deliberately incorporate a third material as a buffer, although this complicates the growth process.

The Nagoya researchers use what is called a strained-layer superlattice, in their case a sequence of ten alternating layers of gallium phosphide and gallium arsenide phosphide, then a sequence of eight gallium arsenide phosphide and gallium arsenide layers, each 200 angstroms thick. Between the superlattice and the silicon is a gallium phosphide layer 500 angstroms thick because the lattice spacing of gallium phosphide nearly matches that of silicon. The superlattice accommodates the mismatch between gallium phosphide and gallium arsenide without the generation of dislocations that propagate upward into the final gallium arsenide layer that resides on top of the superlattice. However, as already mentioned, even the superlattice cannot prevent the generation of all dislocations.

In the end, the question that confronts all new technologies is whether laboratory accomplishments can be translated into the production environment. A report from Ford Microelectronics, Inc., in Colorado Springs given by Chris Ito gave cause for hope. Researchers at Ford customized for gallium arsenide growth a high-volume chemical vapor deposition reactor that could hold 39 wafers each 3 inches in diameter or smaller numbers of larger wafers up to 8 inches in diameter. Material grown on 3-inch silicon wafers had dislocation densities of about 10^7 per square centimeters and in other respects was also of lower quality than the best reported material grown in laboratory-scale, low-volume machines. Material on 5-inch wafers was a little poorer in quality than that on 3-inch wafers. Nonetheless, progress has been remarkably rapid. ■

ARTHUR L. ROBINSON



F. Ponce, Xerox

Gallium arsenide island growing on silicon (100)

This high-resolution transmission electron micrograph by Fernando Ponce of the Xerox Palo Alto Research Center shows some of the structural characteristics of the heteroepitaxial growth process including the coherent nature of the interface; that is, the atomic rows in the gallium arsenide match those in the silicon below. The striped regions in the gallium arsenide are defects called stacking faults that do not occur in high-quality material. Evidently, growth proceeds by the formation of islands, such as this one, which then coalesce into a continuous film.