er, late last year Bernstein, Phillips, and their colleagues, a second group including Gordon Keller of the Basel (Switzerland) Institute for Immunology and Erwin Wagner of the European Molecular Biology Laboratory in Heidelberg, and a third group consisting of Anderson and his colleagues, all reported success at introducing a functional *neo* gene into the bone marrow of mice. Not only was the gene product made in relatively high concentrations, but the results established that the transferred gene entered the bone marrow stem cells.

Why these three groups have obtained better expression of transferred genes in living animals than other investigators is currently not clear, although the design of the viral vectors may be important. Both the Anderson and the Keller-Wagner groups used the Gilboa vector. The Toronto workers used one that they had made but its design is similar to that of Gilboa's vector.

More recently, Anderson, with Arthur Nienhuis of NHLBI and Richard O'Reilly of Memorial Sloan-Kettering Cancer Institute, has gone on to use the Gilboa vector to introduce the the human ADA gene into the bone marrow of living monkeys. The gene product is made in bone marrow–derived cells of five of the seven monkeys treated, although the levels of expression are no more than 10% of those thought to be needed for successful gene therapy of the ADA deficiency. "We are basically making the observation that we have a long way to go [to human trials]," Anderson says.

Williams, Orkin, and Mulligan attempted a similar transfer of the human ADA gene into mice with their vector but were unable to detect any expression of the gene even though the animals' cells carry the new DNA in intact and unrearranged form. "The gene works magnificently in cultured cells," Orkin says, "but when you infect the animal, it [the gene] is nearly dead."

Persistent expression of a transferred gene in the living animal requires that the gene be functional in the bone marrow stem cells, which can then continually replenish the body's red and white cell populations. Differences in the properties of the stem cells and of ordinary cells grown in culture may account for the trouble that investigators have encountered in obtaining good expression of transferred genes in living animals. "The genes are not rearranged. It seems to be a biological problem," Orkin notes.

Work by the Bernstein-Phillips group indicates that the transferred genes are expressed shortly after their entry into stem cells but are later inactivated. The reason why the transferred genes are turned off in stem cells is currently unclear, but there are indications that the cells can specifically inactivate genes regulated by the viral control sequences, as has been the case with the first-generation vectors.

The hereditary anemias, which include sickle cell disease and the thalassemias, are caused by defective genes coding for the hemoglobin proteins. The anemias would be good candidates for gene therapy, except that transferred globin genes have so far not worked particularly well even in cultured cells. Their expression is often low and not subject to the normal regulatory influences in the recipient cells.

According to Nienhuis and Stephan Karlsson of NHLBI, that situation may also be changing. When they introduced a hybrid globin gene with its own regulatory sequences into mouse erythroleukemia cells, the amount of messenger RNA (mRNA) made by the transferred gene was about 10 to 40% of the amount of mRNA made by the natural globin genes. In addition, control of the transferred gene paralleled that of the endogenous genes. "We are encouraged by these results," Nienhuis says. "Previously investigators hadn't seen much mRNA made." Nevertheless, and for reasons that are still poorly understood, the efficient production of mRNA by the transferred genes translates into very little protein synthesis.

Investigators clearly have to do more tinkering with their vectors to improve the expression of transferred genes. At the same time, they have also begun addressing a potential safety problem. In animals at least, insertion of the viral regulatory sequences in the genome may cause a cell to become cancerous if they go in near an oncogene and activate it.

No one knows whether this will also happen in bone marrow cells that have been infected with a viral vector, but Gilboa, Bernstein, and Mulligan, among others, are designing their newer vectors to minimize the possibility. The idea is to modify the vector in such a way that the gene-activating sequences will be lost during integration. Genes carried by such vectors would then have to be controlled by regulatory sequences other than those of the viruses, but this may well be a good idea if some of the expression problems are caused by inactivation of genes under viral control. ■

JEAN L. MARX

ADDITIONAL READING

W. French Anderson, "Prospects for human gene therapy," *Science* **226**, 401 (1984).

R. A. Hock and A. D. Miller, "Retrovirus-mediated transfer and expression of drug resistance genes in human haematopoietic progenitor cells," *Nature (London)* **320**, 275 (1986).

D. A. Williams, S. H. Orkin, R. C. Mulligan, "Retrovirus-mediated transfer of human adenosine deaminase sequences into cells in culture and into murine hematopoietic cells *in vivo*, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 2566 (1986).

Briefing:

Unexpected Size Pattern in Bacterial Proteins

While he was examining data on *Escherichia coli* proteins, Michael Savageau, a biologist at the University of Michigan, came across an unexpected pattern. Rather than being spread over some kind of continuous size distribution, the proteins instead were bunched into groups, which were multiples of a molecular mass of 14,000 daltons (14 kD).

Perhaps this 14-kD unit represents some kind of previously unrecognized fundamental structure, imposed by some aspect of gene architecture or protein chemistry? In which case, suggests Savageau, it might have functional and evolutionary implications.

The observation undoubtedly falls into the "If true, then interesting" category. But how important it might be must then depend on how extensive the pattern is and what exactly underlies it, according to Gregory Petsko of the Massachusetts Institute of Technology.

The idea that proteins might be assembled from basic structural units with a molecular mass of about 18,000 goes back to the 1930's, when Svedberg and others proposed the existence of an underlying unit architecture. Known as the cylcol hypothesis, the suggestion was overwhelmed by strong criticism and eventually was abandonned. If Savageau is correct in his current observations, it may be that he is seeing more clearly what others had perceived earlier but had been unable to substantiate.

Savageau's data come from the extensive catalog of protein information that now exists for *E. coli*. Of the 3000 or so proteins that this bacterium probably manufactures, details of about 1000 are now in the catalog and therefore offer a reasonable sample for statistical inference. A frequency distribution of these proteins shows the expected skew to smaller sizes. But it also reveals an unexpected "fine structure," which analysis shows to be a 14-kD periodicity. "This period of 14 kD is evident for four cycles and then disappears as the data at higher molecular masses become sparse," notes Savageau.

In order to check whether the periodicity might be an artifact of the gel methods used to measure the protein size, Savageau looked at data for *E. coli* genes. GenBank contains data on 146 *E. coli* genes suitable for analysis, and, according to Savageau, these cluster at "lengths of 400, 800, 1200, and 1600 base pairs, corresponding approximately to proteins of 14.6, 29.3, 44.0, and

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.