

selectively retained in the gonads of male mice, thereby producing a considerable radiation dose, as measured locally by using both miniature dosimeters and determination of the yttrium-90 concentrations in the testicular tissue⁶. The β -radiation from yttrium-90 causes an initial dose rate in the testes of about 0.5 mGy h⁻¹ after injection of 37.5 kBq of yttrium-90 or strontium-90.

Radionuclides with chemical and physiological properties similar to those of yttrium are likely to behave in an analogous way. Such radionuclides include fission products in the form of rare-earth elements which, like yttrium-90, are all β -emitters, and the actinides, which belong to the same group as the rare earths. The actinides include uranium and plutonium, which are α -emitters. If such α -emitting radionuclides were to be concentrated in the testes in the same way as yttrium-90 this would result in radiation effects far greater than those of β -emitters because the α -particle energy is normally about one order of magnitude greater than that of the β -particles and produces about ten times greater internal doses for the same activity. Furthermore, the relative biological effectiveness of α -emitters per Gy absorbed dose is 10–20 times greater than that of β -emitters, resulting in an increased mutagenic effect of this order.

WALTER A. MÜLLER

GSF-Institut für Pathologie,
D-8042 Neuherberg, FRG

1. Gardner, M.J. *et al.* *Br. Med. J.* **300**, 423–429 (1990).
2. Dunster, J.H. *Nature* **344**, 98 (1990).
3. Lüning, K.G. *et al.* *Nature* **197**, 304–305 (1963).
4. Riley, P.A. & Willson, R.L. *Br. Med. J.* **300**, 676 (1990).
5. Evans, J.H. *Nature* **345**, 16–17 (1990).
6. Müller, W.A. *Nature* **214**, 931–933 (1967).

Impropriety in proteins

SIR—The production of recombinant proteins from heterologous cells is limited in many cases by the failure of the protein to fold correctly into its active conformation; a common problem is that the recombinant protein forms insoluble inclusion bodies. Darby and Creighton recently suggested in *Scientific Correspondence*¹ that one reason for this behaviour is the interaction of the foreign polypeptide with some component of the cell to form a tight complex which impedes correct folding.

This suggestion is consistent with our more general proposal that several fundamental cellular processes suffer from the problem of incorrect interactions². These are defined as interactions between parts of polypeptide chains and other chemicals (either parts of the same or other polypeptides, or other chemicals such as nucleic acids or small metabolites) which lead to structures that are non-functional in the biological sense. This problem arises

because cellular processes such as protein synthesis, protein transport and the assembly and function of oligomers involve the transient exposure of interactive surfaces to the cellular environment.

The hypothesis of self-assembly³ supposes that all the interactions between surfaces exposed during such processes are both necessary and sufficient to produce the correct structure. It seems more likely that in any given interaction there will be a certain probability that incorrect structures may form. Where this probability is high the cell must reduce the formation of too large a proportion of such structures. One way that cells do this is to produce molecular chaperones, proteins that recognize and bind to transient exposed interactive surfaces in such a way that incorrect interactions are inhibited (see ref. 4).

The take-home message for biotechnologists experiencing difficulty in produc-

ing recombinant proteins in the required active form is that they should re-examine the biogenesis of the protein of interest in the cells in which it occurs naturally to determine whether molecular chaperones are involved. Where this is the case, it is worth exploring the possibility that the yield of active protein could be increased either by the co-expression of the required chaperone in the heterologous cell or by the *in vitro* addition of the chaperone to the protein as it refolds on removal of solubilizing denaturants⁵.

R. JOHN ELLIS

Department of Biological Sciences,
University of Warwick,
Coventry CV4 7AL, UK

1. Darby, N.J. & Creighton, T.E. *Nature* **344**, 715–716 (1990).
2. Ellis, R.J. & Hemmingsen, S.M. *Trends biochem. Sci.* **14**, 339–342 (1989).
3. Anfinsen, C.B. *Science* **181**, 223–230 (1973).
4. Ellis, R.J. *Sem. Cell Biol.* **1**, 1–72 (1990).
5. Goloubinoff, P., Christeller, J.T., Gatenby, A.A. & Lorimer, G.H. *Nature* **342**, 884–889 (1989).

Relationships from gene sequences

SIR—Analysis of gene-sequence data has led to new insights into the evolutionary relationships between organisms. These data are interpreted by inferring hypothetical point mutations and short insertions and deletions between the present-day organisms and their most recent common ancestors. The resulting phylogenetic trees represent the true evolutionary history of these organisms to the extent that one or two gene families are fully representative of the evolution of the entire genome. Clearly, a more complete description of this history requires the comparison of full genomes or at least an analysis of possible genomic evolutionary events such as transposition, duplications and other recombination-based processes.

We have recently proposed a statistical method¹, using published genetic maps and gene designations², for the comparison of gene order among bacteria. In the course of this work, we frequently encountered incompatibilities among the data sets from different bacteria. A gene from one bacteria has the same genetic designation as a completely different one from another, or genes apparently encoding homologous proteins from different bacteria have different names. Some effort has been expended to reconcile the differences between *Escherichia coli* and *Salmonella typhimurium*^{3–5}, but inconsistencies remain. Others become apparent when comparing more distant species of bacteria.

We have addressed this problem by creating a normalized gene designation database (NGDD) for the five bacteria we studied (*E. coli*, *S. typhimurium*, *B. subtilis*, *Caulobacter crescentus* and *Pseudomonas aeruginosa*).

Our work was greatly assisted by the use of the *Salmonella* gene database graciously provided by K. Sanderson⁴. In our database, containing more than 2,500 fully documented entries, gene designations are normalized to a four-letter code (three letters for the locus and one for the gene), and inconsistencies in nomenclature are cross-referenced. NGDD is available from the GenInfo Data Repository, National Center for Biotechnology Information, Building 38A, National Institutes of Health, 8600 Rockville Pike, Bethesda, Maryland 20894, USA*.

We believe that this database will be an important tool for naming newly sequenced genes from different organisms, and its availability should stimulate the reconciliation of genetic nomenclature, a problem that will only become more acute as laboratories dedicated to producing sequences on a large scale begin to produce more data.

ROBERT CEDERGREN

Department of Biochemistry,

DAVID SANKOFF

YVON ABEL

Center for Mathematical Research,
University of Montreal,
Montreal,
Quebec H3C 3J7,
Canada

1. Sankoff, D., Cedergren, R. & Abel, Y. *Meth. Enzym.* **183**, 428–438 (1990).
2. O'Brien, S. *Genetic Maps* Vol. 4 (Cold Spring Harbor Laboratory, 1987).
3. Iino, T. *et al.* *Microbiol. Rev.* **52**, 533–535 (1988).
4. Sanderson, K.E. & Roth, J.R. *Microbiol. Rev.* **52**, 485–532 (1988).
5. Bachmann, B. in *Escherichia coli and Salmonella typhimurium* Vol. 2, 807–876 (American Society of Microbiology, 1987).

* Users of Internet can access the database directly at the address: NCBI.NLM.NIH.Gov. User name: Anonymous. Password: your name.