We have been able to identify at least four distinct types of lipidic particle from electron micrographs of freeze-fracture replicas of aqueous dispersions of mixed galactolipids. Figure 2 provides a composite view of some of these. The particles shown in Fig. 2*a* and *b* are regular quasi-crystalline arrangements induced by the addition of ethylene glycol and by heat, respectively. Figure 2*c* shows isolated pits formed from particles of the type seen in Fig. 2*a*. A two-dimensional array of particles located within a single fracture-face is shown in Fig. 2*d* together with a number of rather larger particles that resemble structures attributed by some workers to inter-bilayer attachment sites[1–4].

The quasi-crystalline structures, as we pointed out in our original communication[5], seem to correspond to aggregates of inverted lipid micelles. Hui and Boni have suggested that such structures correspond to isotropic cubic phases of a type previously reported in X-ray diffraction studies of soaps and detergents[6] and simple monoglycerides[7,8]. In our experience, a variety of quasi-crystalline structures characterized by different particle sizes and symmetries can be induced to form in mixed galactolipid dispersions. We are, consequently, less willing than Hui and Boni to commit ourselves to any definite symmetry assignments. We do not exclude the possibility that such structures are related to cubic phases similar to those referred to by these authors but, as we have emphasized elsewhere[9], more evidence based on other techniques is required before any realistic attempt can be made to analyse the detailed structure of the aggregates we have reported.

Hui and Boni further suggest that the occurrence of cusps formed by groups of pits associated with inverted lipid micelles present in these aggregates are directly related to the structures attributed to inter-bilayer attachment sites[1–4]. This resemblance is in our opinion purely fortuitous and merely reflects the fact that such aggregates contain a continuous three-dimensional hydrocarbon matrix. We believe that the various types of lipidic particles reported in the literature can be accounted for in terms of a series of intermediary states made up of inverted micelles at one extreme and bilayer arrangements in which the lipids forming these micelles have been re-incorporated into the bilayer phase at the other.

This reorganization, as detailed elsewhere[10], involves an appreciable expansion of the surface area of bilayer. In the initial stages, this leads to blebs of the type seen in the fracture-face shown in Fig. 2*c*, and then to a buckling of the fracture-face as shown in Fig. 2*d*. This process can account both for the formation of tubular–micelle structures and particle–strings of the type shown in Fig. 1 of our original report[5] and the characteristic arrangements of ridges and particle–strings reported in other studies.

W. P. WILLIAMS*
A. SEN†
A. P. R. BRAIN‡
P. J. QUINN†
*Departments of *Biophysics and*
†Biochemistry and the
‡Electron Microscopy Unit,
Chelsea College,
University of London,
London SW3 6LX, UK

M. J. DICKENS
MRC Cell Biophysics Unit,
King's College, University of London,
26 Drury Lane,
London WC2R 2LS, UK

1. Miller, R. G. *Nature* **287,** 166–167 (1980).
2. Hui, S. W. & Stewart, T. P. *Nature* **290,** 427–428 (1981).
3. Rand, R. P., Reese, T. S. & Miller, R. G. *Nature* **293,** 237–238 (1981).
4. Hui, S. W., Stewart, T. P., Boni, L. T. & Yeagle, P. L. *Science* **212,** 921–923 (1981).
5. Sen, A., Williams, W. P., Brain, A. P. R., Dickens, M. J. & Quinn, P J. *Nature* **293,** 488–490 (1981).
6. Luzzati, V. & Reiss-Husson, F. *Nature* **210,** 1351–1352 (1966).
7. Lindblom, G., Larsson, K., Johansson, L., Fontell, K. & Forsen, S. *J. Am. chem. Soc.* **101,** 5465–5470 (1979).
8. Larsson, K., Fontell, K. & Krog, N. *Chem. Phys. Lipids* **27,** 321–328 (1980).
9. Sen, A., Williams, W. P., Brain, A. P. R. & Quinn, P. J. *Biochim. biophys. Acta* (in the press).
10. Sen, A., Williams, W. P., Brain, A. P. R. & Quinn, P. J. *Biochim. biophys. Acta* (in the press).

Gene duplication and the birthday problem

EDLUND and Normark[1] have presented sequence and distribution data of tandem duplications of the *ampC* gene of *Escherichia coli* K-12, which suggest that unequal recombination between homologous sequences of the order of 12 base pairs (bp) in length may be a general mechanism for the generation of tandem duplications. In support of this conclusion they present calculations showing the probability of occurrence of homologies 10–14 bp long in a DNA sequence of given length. For example they state[1] that "on average two homologous sequences of any composition will be found within a random DNA sequence of 4 kilobases (kb)".

It is more realistic to calculate the length of DNA required such that the *expected* or average number of homologies is one. The number of homologies will be expected to be approximately Poisson distributed and the expected number of homologies, λ, will be given by[2]

$$\lambda = n\frac{e^{-r/n}}{k!}\left(\frac{r}{n}\right)^k \qquad (1)$$

where r is the length of the DNA segment, n is the number of possible different sequences of length m, equal to 4^{12} for a sequence 12 bp long, and k is the number of copies of a homology. Thus, on the average, one pair of homologous sequences 12 bp long will occur in a DNA sequence of length 5.8 kb rather than 4 kb as stated by Edlund and Normark. Allowing for the 1 kb sequence of the *ampC* gene, this value becomes 6.8 kb. The corrected values for Table 1 in ref. 1 may be obtained easily from equation (1).

Table 1 (ref. 1) does not make clear that the probabilities of pairs of homologous sequences are being calculated, rather than the total number of sequences which are homologous in a DNA segment of given length. Although the size of DNA required to obtain an average of four homologous pairs of length 12 bp is 11.6 kb (not allowing for the *ampC* gene), the length of DNA required to obtain an average of one sequence occurring four times is 580 kb! This point is significant as the opportunity for unequal recombination is much greater if multiple copies of a sequence are present than if several pairs of different sequences are present.

A further complication is that the calculation of the probability of occurrence of a sequence homology assumes that the DNA segment of length r contains $r-m+1$ (approximately r) independent sequences of length m. This is clearly incorrect as neighbouring sequences overlap and their probabilities of occurrence will depend on each other. If this dependency is ignored, then the frequency of sequence homologies will be overestimated.

These calculations do not alter Edlund and Normark's conclusion that unequal recombination between random homologies could be a mechanism for generating tandem duplications of the size they observe. This size, 9.8 kb, is still larger than the estimate of the minimum length of DNA required for a homology of length 12 bp to occur an average of once, even if this estimate is biased. However, the importance of such random homologies for generating tandem duplications significantly shorter in length may be questionable.

JULIAN ADAMS
Division of Biological Sciences,
The University of Michigan,
Ann Arbor,
Michigan 48109, USA

1. Edlund, T. & Normark, S. *Nature* **292,** 269–271 (1981).
2. Feller, W. *An Introduction to Probability Theory and its Applications* 3rd edn (Wiley, New York, 1968).

NORMARK REPLIES—Edlund and I recently suggested[1] that unequal recombination between perfect homologies of the order of 10–14 bp may be a general mechanism for generating duplications in the size range of 10 kb. This suggestion was based on DNA sequencing of one duplication in the *ampC* region of the *E. coli* chromosome, on end point determinations of a number of duplications and on probability calculations of occurrence of homologies.

Our calculations that are questioned by Adams must be regarded as approximations. They were made in order to see whether or not our distribution of duplications could be explained by recombination betweeen short homologies of any base composition. Our calculations

0028–0836/82/100176—02$01.00