$\mathbf{r}$ between the markers,

$$P(\mathbf{r}) = \left(\frac{3}{2\pi N' a^2}\right)^{3/2} e^{-3(\mathbf{r}-\mathbf{R})^2/2N'a^2}, \tag{8.35}$$

where N' is now the number of Kuhn segments between the second tether and the second marker. This formula follows simply from Equation 8.34 when applied to the distribution of distances $\mathbf{r} - \mathbf{R}$ between the second tether and the second marker. It is interesting to note that mathematical properties of Gaussian distributions, like the one that says that a convolution of two Gaussian distributions is a Gaussian distribution, dictate that *any* tethering configuration will result in a displaced Gaussian distribution of distances.

The implicit assumption we have made in writing Equations 8.34 and 8.35 is that chromosome configurations can be described by random walks. In light of the dense packing of chromosomes in cells, this might seem like an overly zealous use of a simple physical model. However, as we demonstrate using several examples later in this section, this model captures key features of experimental data on chromosomes and, more importantly, it makes falsifiable predictions suggesting new directions for experimentation. As a result, this model is a good starting point for quantitative investigations of chromosome geography. This idea is further bolstered by the Flory theorem, which states that for dense polymer systems, such as chromosomes confined to cells, distributions of distances between monomers are described by random walk statistics.

The contour length of the chromosome between the two tagged loci, Na, can be expressed in terms of the genomic distance between the two fluorescent markers as $Na = N_{bp}/\nu$, where ν is the linear packing density of DNA in chromatin. For example, two genomic loci $N_{bp} = 100\,\text{kb}$ apart would be separated by a 30 nm fiber, which is $100\,\text{kb}/(100\,\text{bp/nm}) = 1\,\mu\text{m}$ in contour length. Assuming that the chromatin structure is that of a 10 nm fiber the contour distance along the fiber between the loci would be 10 times as large given the 10 times smaller packing density.

The end-to-end distribution function for a random walk polymer is determined by a single parameter Na^2, the mean end-to-end distance squared. Since the contour length $Na = N_{bp}/\nu$, the mean end-to-end distance squared can also be written as $\langle R^2 \rangle = N_{bp} a/\nu$. Therefore the material parameter that characterizes the random walk model of chromosomes is the ratio of the Kuhn length and the packing density. This parameter can be determined from measurements of the average distance squared between two labeled regions of the chromosome as a function of their genomic distance. The results of such a measurement on human chromosome 4 are shown in Figure 8.11. The fit to the data yields an estimate of $a/\nu = 2\,\text{nm}^2/\text{bp}$, which is nothing but the initial slope of the linear portion of the data. The fact that the data level off at large genomic distance can be attributed to the effect of chromosome confinement within the cell nucleus. Below we analyze this confining effect using a random walk model for chromosome configurations in the bacterium *Vibrio cholerae*.

With a measurement of the chromatin material parameter a/ν in hand, we can compute the expected probability distribution of distances between fluorescently tagged loci on the chromosome. Typically, due to random orientations of cells in the microscope, experiments with tagged chromosomes only yield information about the magnitude r of the distance vector $\mathbf{r}$ between the two marked spots on the chromosome. Probability distributions for this quantity follow

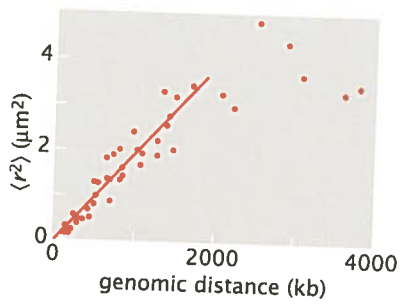

Figure 8.11: Physical distance between two fluorescently labeled loci on human chromosome 4 as a function of the genomic distance. The physical distance is measured in terms of the average squared distance between the two labels (dots). The curve corresponds to a linear fit as discussed in the text. (Adapted from G. van den Engh et al., *Science* 257:1410, 1992.)

from Equations 8.34 and 8.35 by integrating out the angular variables θ and ϕ associated with the vector $\mathbf{r}$. This procedure yields

$$P(r) = \left(\frac{3}{2\pi N a^2}\right)^{3/2} 4\pi r^2 e^{-3r^2/2Na^2}, \tag{8.36}$$

for the free-polymer case, and

$$P(r) = \left(\frac{3}{2\pi N' a^2}\right)^{1/2} \frac{r}{R}\left(e^{-3(r-R)^2/2N'a^2} - e^{-3(r+R)^2/2N'a^2}\right) \tag{8.37}$$

when the polymer is tethered. Note that tethering gives a different functional form for the distribution of distances. This provides us with a mathematical tool with which to detect tethering of chromosomes in cells.

Measurement of the distribution of distances between tagged regions on yeast chromosome III suggests that this difference in distributions can be observed *in vivo*. In Figure 8.12, we show the distance distribution measured between two fluorescent tags, one placed near the so-called HML region of chromosome III of budding yeast and the other on the spindle pole body, which is at a fixed location on the nuclear periphery and essentially marks the location of the centromere. The measured distribution is poorly fitted by the free-polymer formula, Equation 8.36, while the tethered-polymer formula, Equation 8.37 does the job well.

The fit to the tethered-polymer distribution yields two quantities that characterize the model, namely, the mean-squared distance between the tether and the fluorescent marker at HML, $N'a^2 = 0.5\,\mu\text{m}^2$, and $R \approx 0.9\,\mu\text{m}$, the distance from the spindle pole body to the tethering point. Note that in order to compute the genomic location of the putative tethering point, we need the quantity a/ν that characterizes chromatin structure. For that, measurements like those leading to Figure 8.11 for human chromosome 4 are needed.

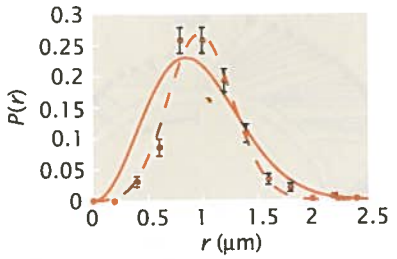

Figure 8.12: Statistics of yeast chromosome III. Distribution of distances between two fluorescent tags placed in proximity of the centromere and the HML region on yeast chromosome III. These two regions are separated by approximately 100 kb in genomic distance. The full line is a fit to the free-polymer distance distribution, Equation 8.36, while the dashed line is a fit to the tethered-polymer formula, Equation 8.37. (Courtesy of S. Gordon-Messer, J. Haber, and D. Bressan.)

Chromosome Territories Have Been Observed in Bacterial Cells

Bacterial chromosomes used to be thought of as unstructured and random. This view has been seriously challenged by experiments that utilize fluorescent markers placed at different genomic locations, as shown in Figure 8.13. In this experiment, 112 different mutants of *Caulobacter crescentus* were created with fluorescent tags placed at 112 different locations covering the length of its circular chromosome. Measurements of the average position of the marker along the length of the cell revealed a linear relationship between the genomic distance from the origin of replication and the physical distance away from the pole of the bacterium. This is not to be expected assuming a simple model of the 4 Mb circular chromosome as a polymer loop confined to the cell.

Estimate: Chromosome Organization in *C. crescentus*

ESTIMATE

Another measure of the organization of the chromosome in *C. crescentus* is provided by the width of the distribution of positions of the marked regions. As shown in Figure 8.13, the standard deviation of the position is independent of genomic distance from the origin of replication, and is approximately $0.2\,\mu\text{m}$ (cell length $L \approx 2\,\mu\text{m}$). We can rationalize this measurement within a simple model where the chromosome is partitioned into loops. This can be effected by proteins that make contact between different locations on the chromosome (H-NS