antibody to chicken Cμ (M1) (Southern Biotechnology Associates, Birmingham, AL) and then with polyclonal fluorescein isothiocyanate–conjugated goat antibodies to mouse IgG (Fab)₂ (Sigma). Predominantly sIgM(+) subclones were excluded from the analysis, because they most likely originated from cells that were already sIgM(+) at the time of subcloning.

23. For Ig light chain sequencing, PCR amplification and sequencing of the rearranged light chain V segments were performed as previously described

(19), except that high-fidelity PfuTurbo polymerase (Stratagene) was used with primer pair Vλ1/Vλ2 for PCR, and primer Vλ3 was used for sequencing (17). Only one nucleotide change, which most likely reflects a PCR-introduced artifact, was noticed in the V-J-3′ intron region in a total of 80 0.5-kb-long sequences from AID^{−/−E} cells.

24. We thank M. Reth and T. Brummer for kindly providing the MerCreMer plasmid vector; P. Carnici and Y. Hayashizaki for construction of the riken1 bursal

Capturing Chromosome Conformation

Job Dekker,[1]* Karsten Rippe,[2] Martijn Dekker,[3] Nancy Kleckner[1]

We describe an approach to detect the frequency of interaction between any two genomic loci. Generation of a matrix of interaction frequencies between sites on the same or different chromosomes reveals their relative spatial disposition and provides information about the physical properties of the chromatin fiber. This methodology can be applied to the spatial organization of entire genomes in organisms from bacteria to human. Using the yeast *Saccharomyces cerevisiae*, we could confirm known qualitative features of chromosome organization within the nucleus and dynamic changes in that organization during meiosis. We also analyzed yeast chromosome III at the G₁ stage of the cell cycle. We found that chromatin is highly flexible throughout. Furthermore, functionally distinct AT- and GC-rich domains were found to exhibit different conformations, and a population-average 3D model of chromosome III could be determined. Chromosome III emerges as a contorted ring.

Important chromosomal activities have been linked with both structural properties and spatial conformations of chromosomes. Local properties of the chromatin fiber influence gene expression, origin firing, and DNA repair [e.g., (1, 2)]. Higher order structural features—such as formation of the 30-nm fiber, chromatin loops and axes, and interchromosomal connections—are important for chromosome morphogenesis and also have roles in gene expression and recombination. Activities such as transcription and timing of replication have been related to overall spatial nuclear disposition of different regions and their relationships to the nuclear envelope [e.g., (3–6)]. At each of these levels, chromosome organization is highly dynamic, varying both during the cell cycle and among different cell types.

Analysis of chromosome conformation is complicated by technical limitations. Electron microscopy, while affording high resolution, is laborious and not easily applicable to studies of specific loci. Light microscopy

affords a resolution of 100 to 200 nm at best, which is insufficient to define chromosome conformation. DNA binding proteins fused to green fluorescent protein permit visualization of individual loci, but only a few positions can be examined simultaneously. Multiple loci can be visualized with fluorescence in situ hybridization (FISH), but this requires severe treatment that may affect chromosome organization.

We developed a high-throughput methodology, Chromosome Conformation Capture (3C), which can be used to analyze the overall spatial organization of chromosomes and to investigate their physical properties at high resolution. The principle of our approach is outlined in Fig. 1A (7). Intact nuclei are isolated (8) and subjected to formaldehyde fixation, which cross-links proteins to other proteins and to DNA. The overall result is cross-linking of physically touching segments throughout the genome via contacts between their DNA-bound proteins. The relative frequencies with which different sites have become cross-linked are then determined. Analysis of genome-wide interaction frequencies provides information about general nuclear organization as well as physical properties and conformations of chromosomes. We have used intact yeast nuclei for all experiments. Although the method can be performed using intact cells, the signals are considerably lower, making quantification difficult (9). The general nuclear organization

of purified nuclei is largely intact, as shown below.

For quantification of cross-linking frequencies, cross-linked DNA is digested with a restriction enzyme and then subjected to ligation at very low DNA concentration. Under such conditions, ligation of cross-linked fragments, which is intramolecular, is strongly favored over ligation of random fragments, which is intermolecular. Cross-linking is then reversed and individual ligation products are detected and quantified by the polymerase chain reaction (PCR) using locus-specific primers. Control template is generated in which all possible ligation products are present in equal abundance (7). The cross-linking frequency (X) of two specific loci is determined by quantitative PCR reactions using control and cross-linked templates, and X is expressed as the ratio of the amount of product obtained using the cross-linked template to the amount of product obtained with the control template (Fig. 1B). X should be directly proportional to the frequency with which the two corresponding genomic sites interact (10).

Control experiments show that formation of ligation products is strictly dependent on both ligation and cross-linking (Fig. 1C). In general, X decreases with increasing separation distance in kb along chromosome III ("genomic site separation"). Cross-linking frequencies for both the left telomere and the centromere of chromosome III with each of 12 other positions along that same chromosome (Fig. 1, C and D) were determined using nuclei isolated from exponentially growing haploid cells. Interestingly, the two telomeres of chromosome III interact more frequently than predicted from their genomic site separation, which suggests that the chromosome ends are in close spatial proximity. This is expected because yeast telomeres are known to occur in clusters (11, 12).

We next applied our method to an analysis of centromeres and of homologous chromosomes ("homologs") during meiosis in yeast (7). In mitotic and premeiotic cells, centromeres are clustered near the spindle pole body (13, 14) and homologous chromosomes are loosely associated (15–17). These features change markedly when cells enter meiosis (13). The centromere cluster is rapidly lost and is not restored until just before the first meiotic division. Loose interactions be-

[1]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA. [2]Molekulare Genetik (H0700), Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, and Kirchhoff-Institut für Physik, Physik Molekularbiologischer Prozesse, Universität Heidelberg, Schröderstrasse 90, D-69120 Heidelberg, Germany. [3]2e Oosterparklaan 272, 3544 AX Utrecht, Netherlands.

*To whom correspondence should be addressed. E-mail: jdekker@fas.harvard.edu

tween homologs are transiently lost during S phase, but are immediately restored and then become increasingly robust during prophase when synaptonemal complex is formed [reviewed in (18, 19)].

Centromere relationships were probed by analyzing the frequencies with which the centromere of chromosome IV (CEN4) became cross-linked to each of 10 sites along the length of chromosome III. In premeiotic cells, CEN4 interacted strongly only with the chromosome III centromere (CEN3; primer pair 6+14, Fig. 2A). Identical results were obtained in exponentially growing diploid cells and in both MATa and MATα haploid cells (9). In contrast, at 4 and 5 hours after the onset of meiosis, the interaction frequency between the two centromeres was reduced by a factor of 4 to 5 (Fig. 2A), in good agreement with the timing and extent of reduction observed in cytological studies (12, 14). Interactions between CEN4 and sites on chromosome III distant from CEN3 were little or not at all reduced. A low frequency of centromere interactions was still observed at the later time points. This signal likely represents the 10 to 15% of cells that typically fail to enter meiosis in such experiments (20).

Relationships between homologs were analyzed using maternal and paternal versions of chromosome III marked differentially with Xho I restriction site polymorphisms that flank a meiotic recombination hotspot (HIS4LEU2) (21). This hotspot is located in the middle of the 106-kb left arm (Fig. 2B). Cross-linking and ligation of the Xho I hotspot fragments from the two homologs yields unique ligation products that are not formed after ligation of fragments from sister chromatids. As a control, we analyzed nonhomologous interactions between each of the chromosome III hotspot fragments and analogously positioned sites on chromosome VI. The latter sites were located in the middle of the chromosome VI right arm, which is also ~100 kb. As a result, in the control experiment, homologous and nonhomologous interactions should be closely similar with respect to juxtaposition mediated by clustering of telomeres or centromeres.

The level of homologous interactions was low in premeiotic cells (Fig. 2B, lanes 1 and 5) but increased by a factor of ~10 after 4 and 5 hours, times at which homologs are known to be maximally juxtaposed (20, 21). In contrast, nonhomologous interactions were infrequent in premeiotic cells and did not increase during meiosis (Fig. 2B, lanes 2 to 4 and 6 to 8). Interestingly, even in premeiotic cells homologous interactions were slightly more frequent than nonhomologous interactions, consistent with the loose pairing of homologs detected at this stage by FISH (15–17).

These observations demonstrate that the 3C assay reliably detects important qualitative features of chromosome organization. In addition, these results suggest that nuclear organization is not markedly affected during nuclei purification.

The 3C assay also permits detailed quantitative analysis of chromosome structure. The spatial disposition of the chromatin fiber is determined by its flexibility and by additional constraints on its path. These parameters together determine the interaction frequencies of different sites. When a large number of cross-linking frequencies is determined, the relationship between cross-linking frequency and genomic site separation can be interpreted using polymer models that describe this relationship in terms of flexibility and other structural parameters that relate to chromosome conformation (22–26).

The cross-linking frequency between two loci of the chromosome with a site separation distance s (in kb), $X(s)$, is directly proportional to the local concentration $j_M(s)$, which is the concentration of one site of the polymer in proximity to the other site (26):

$$X(s) = k \times j_M(s) \qquad (1)$$

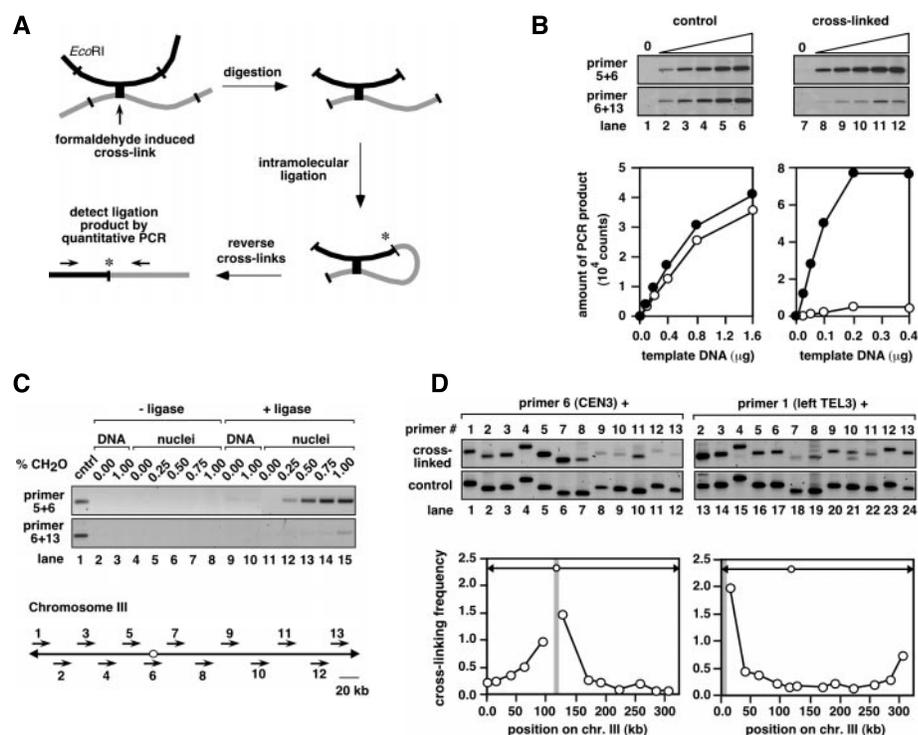The proportionality constant k reflects the efficiency of the cross-linking reaction and

Fig. 1. The Chromosome Conformation Capture (3C) methodology. (**A**) Schematic representation of the assay: formaldehyde cross-linking, Eco RI digestion, intramolecular ligation, and PCR-mediated detection of ligation products after reversal of the cross-links. The asterisk indicates the newly formed restriction site. (**B**) Determination of the cross-linking frequency of two loci. The linear range for the quantitative PCR reactions was determined by titrating the cross-linked and control templates, and the products were run on agarose gels. The graphs show the quantitation of PCR product that was obtained with primer pair 5+6 (solid circles) and primer pair 6+13 (open circles) [see (C)]. PCR product formation was linear up to 0.15 μg of cross-linked template and up to 0.4 μg of the control template. In all subsequent experiments, 35 ng of cross-linked template and 150 ng of control template were used. (**C**) Lower panel: Positions of 13 primers along chromosome III are indicated. The open circle indicates the centromere; arrowheads indicate the telomeres. Upper panel: Control experiments. Primer pairs 5+6 and 6+13 were used to detect ligation products on various templates. Similar amounts of ligation products were detected when the control template was used (lane 1). No PCR products were obtained on any of the templates when the ligation step was omitted (lanes 2 to 8). Treating purified genomic DNA with formaldehyde before digestion followed by dilution and ligation did not result in increased ligation product formation (lanes 9 and 10). Using nuclei, no random intermolecular ligation products were detected (lane 11), because no products were detected when the formaldehyde treatment was left out (44). Ligation product formation increased linearly with formaldehyde concentration (lanes 12 to 15). In all other experiments, 1% formaldehyde was used. (**D**) The method was applied to nuclei isolated from an exponentially growing haploid culture. The cross-linking frequency of the centromere (primer 6, lanes 1 to 12) and the left telomere (primer 1, lanes 13 to 24) with the 12 other sites along chromosome III was determined and plotted. A schematic representation of chromosome III is indicated in which the gray vertical bars represent the position of either the centromere or the left telomere.

can slightly vary from experiment to experiment. Within a given experiment, variation in *k* can reflect real local differences in susceptibility to cross-linking.

The value of $j_M(s)$ can be expressed in terms of two key parameters, *l* and *c*, by a numerical expression that combines the freely jointed chain (FJC) and the Kratky-Porod (elastic rod) polymer description (*25, 26*). The parameter *l* reflects the flexibility of the fiber; *l* is the statistical segment length that equals two times the persistence length in nm. The parameter *c* describes the presence of constraints on the random walk of the polymer. In the absence of such constraints, a linear fiber displays a maximum value of *X* at a site separation distance of $l \sim 1.7$ segments, followed by a gradual decrease in interaction frequency with increasing site separation. The presence of constraints is detected as a deviation from this behavior [see also (*27*)]. Formally, such deviations may result in apparent circular behavior of an intrinsically linear polymer. The parameter *c* gives the apparent circle size in kb. For a linear polymer that is unconstrained, *c* is infinitely large. The relation between the cross-linking frequency $X(s)$ and the site separation *s* can be expressed as

$$X(s) = k \times 0.53 \times \beta^{-3/2} \times$$

$$\exp\left(-\frac{2}{\beta^2}\right) \times l^{-3} \frac{nm^3\ mol}{liter} \qquad (2)$$

where $\beta = 11.1\ nm/kb \times (s/l) \times [1 - (s/c)]$. For a linear polymer, *s/c* approaches zero and Eq. 2 becomes the simpler equation that describes a linear polymer (*26*).

Equation 2 includes an estimate of the packing ratio of the chromatin fiber. In yeast, a packing ratio of 11.1 nm/kb was estimated from a mass density of 6 nucleosomes per 11 nm of fiber (*28*), corresponding to a 30-nm fiber and a nucleosome repeat length of 165 base pairs (*29*). For the following analysis we made the assumption that β is constant and that there is no higher level of organization of the chromatin fiber. Also, we assumed that the fiber is not confined to a limited volume. Equation 2 can be applied to chromosomes of any organism by replacing the value of 11.1 in the expression for β with a suitable fiber packing ratio.

We applied Eq. 2 to the modeling of the ~320-kb yeast chromosome III as it occurs in the G_1 phase of the cell cycle. Nuclei were purified from haploid cells arrested in G_1 with α-factor (*7*), and cross-linking frequencies were determined for 78 different site pairs. The pairs examined were selected to represent the full range of genomic site separation distances across the entire length of the chromosome (all pairwise combinations of positions 1 to 13 shown in Fig. 1C). Interaction frequencies were plotted against the genomic site separation *s* (Fig. 3A) and optimal values for *l*, *c*, and *k* were found by fitting the data to Eq. 2, yielding $l = 56$ nm, $c = 363$ kb, and $k = 4.0 \times 10^6$ M^{-1} ($R^2 = 0.86$). Although any particular model can only be an approximation for real chromatin, the current results suggest that polymer models are useful for this purpose and that the
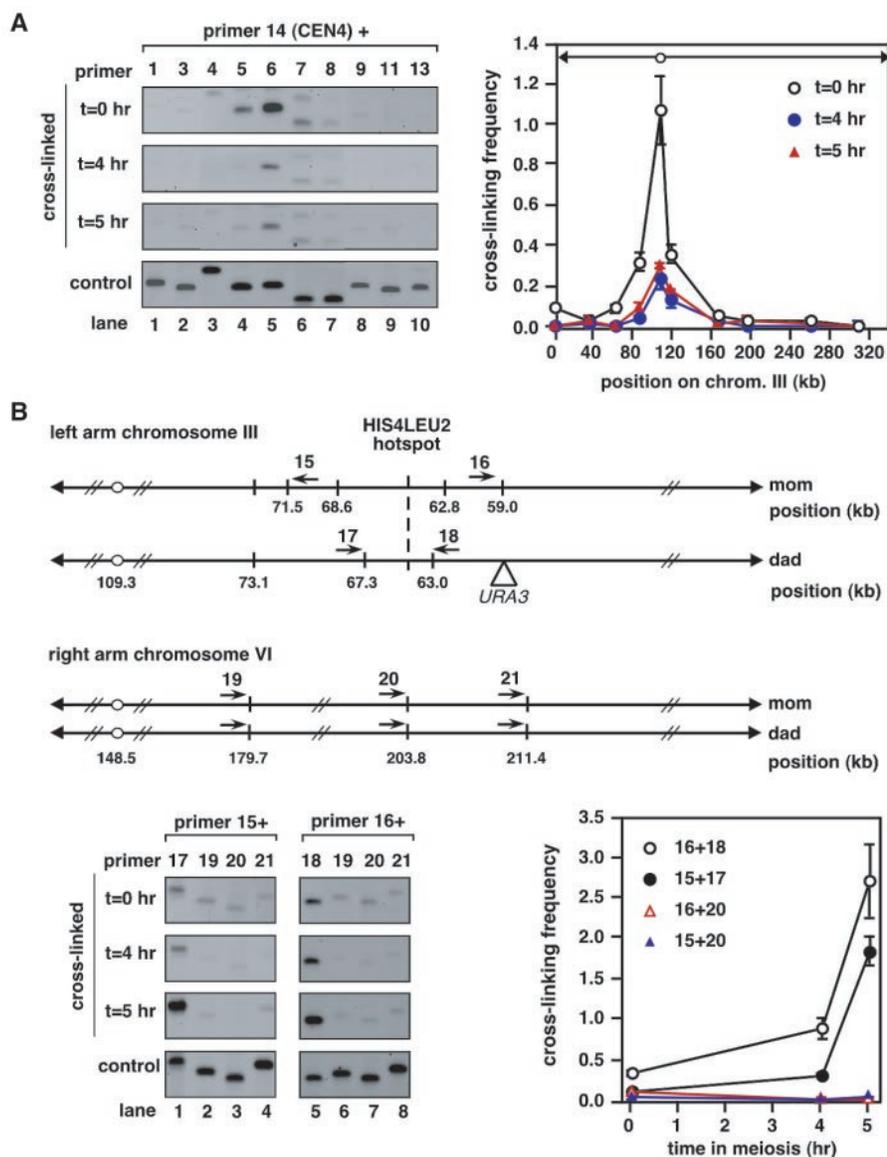

Fig. 2. Analysis of nuclear dynamics during meiosis. (**A**) Loss of centromere clustering during meiosis. Nuclei were isolated at 0, 4, and 5 hours after induction of meiosis (*7*). Eco RI was used to generate cross-linked and control templates. Primer 14 was specific for the Eco RI fragment containing CEN4. The cross-linking frequency of CEN4 with 10 positions along chromosome III was determined (lanes 1 to 10). Cross-linking frequencies were determined in triplicate; the average and standard error of the mean are plotted in the graph. Chromosome III is schematically depicted at the top to indicate the position of CEN3. (**B**) Upper panel: Schematic representation of the left arm of chromosome III and the right arm of chromosome VI. The open circles indicate the centromeres; arrowheads indicate the telomeres. Small vertical bars indicate Xho I sites. Arrows indicate the positions of the primers (15 to 21) used to detect chromosomal interactions. The Xho I restriction fragment length polymorphism on the left arm of chromosome III contains the *HIS4LEU2* hotspot [indicated by the dashed line (*21*)]. Lower panels: Interactions between homologous and nonhomologous chromosomes were analyzed with the same nuclei preparations as used for the experiment shown in (A), but here Xho I was used to generate cross-linked and control templates. Interactions between allelic sites were analyzed using primers 15 and 17 and primers 16 and 18 (lower left, lanes 1 and 5). Interactions between chromosome III and VI were studied using primers 15 and 16 in combination with primers 19, 20, and 21 (lanes 2 to 4 and 6 to 8). Cross-linking frequencies were determined in triplicate, and the average and standard error of the mean are plotted over time (lower right). Only two of the six nonhomologous interactions are plotted; the other four were similar to the ones graphed and are not plotted.

particular model we used is a good first approximation for chromosome III under the cellular circumstances of these experiments. We did not need to include terms for volume exclusion in this analysis. Deviations from the fitted curve may be due to noise but may also reflect real differences (see below).

A value of c = 363 kb indicates that this linear chromosome is constrained in a circular conformation whose apparent circle size is only slightly larger than the actual length of the chromosome (~320 kb). This feature is most prominently apparent in the tendency for cross-linking frequencies to increase for larger site separations (s > 200 kb; Fig. 3A).

A value of l = 56 nm corresponds to a persistence length of 28 nm, which in turn corresponds to 2.5 kb. These results suggest that yeast chromatin is quite flexible [see also (30)]. The persistence length, expressed in kb, of chromatin is considerably larger than that of naked DNA (2.5 kb versus 150 base pairs). Therefore, this property affects interactions between sites separated by larger genetic distances in the case of chromatin than in the case of naked DNA. Hence, chromosome III is quite compact, with sites separated by large genomic distances being relatively close to one another in three-dimensional (3D) space. The highest relative local concentration of two sites occurs at a genomic site separation distance of $s \sim 9$ kb [$l \sim 60$ nm (31)]. At this site separation, the relative local concentration of the two sites is 6×10^{-7} M. The local concentration remains relatively high, at least 2×10^{-8} M, for all pairs of sites along chromosome III. These values imply that interactions between proteins present at nanomolar concentrations will be strongly facilitated when they are bound to this chromosome, even if their binding sites are separated by a relatively large genomic distance.

The value of l also has consequences for the higher order organization of yeast chromosomes. Although some physical properties of mitotic and meiotic chromosomes are likely to be different from the properties we determined here, the flexibility of the chromatin fiber in G_1 is at least consistent with the observed loop sizes of ~20 kb of meiotic chromosomes (32) and estimates of loop sizes of 9 to 15 kb in mitotic chromosomes (33, 34). These relationships suggest that chromatin flexibility may be an important parameter of loop formation and function [see also (19)]. The value of l will also be important for chromosomal processes that involve looping and other long-range interactions along the chromatin fiber.

Yeast chromosomes comprise GC-rich and AT-rich domains, or isochores, of about 50 to 100 kb (35, 36). Chromosome III has one AT-rich domain (located on the right arm from 100 to 190 kb) and two GC-rich domains (located on the left arm from 25 to 100 kb and on the

right arm from 190 to 280 kb). These domains exhibit functional differentiation, as is most clearly illustrated by the fact that the GC-rich domains display high levels of meiotic recom-

bination relative to the AT-rich domain (37, 38). These domains are likely analogous to the R- and G-bands found in larger eukaryotes (19, 39). We were interested in the possibility that
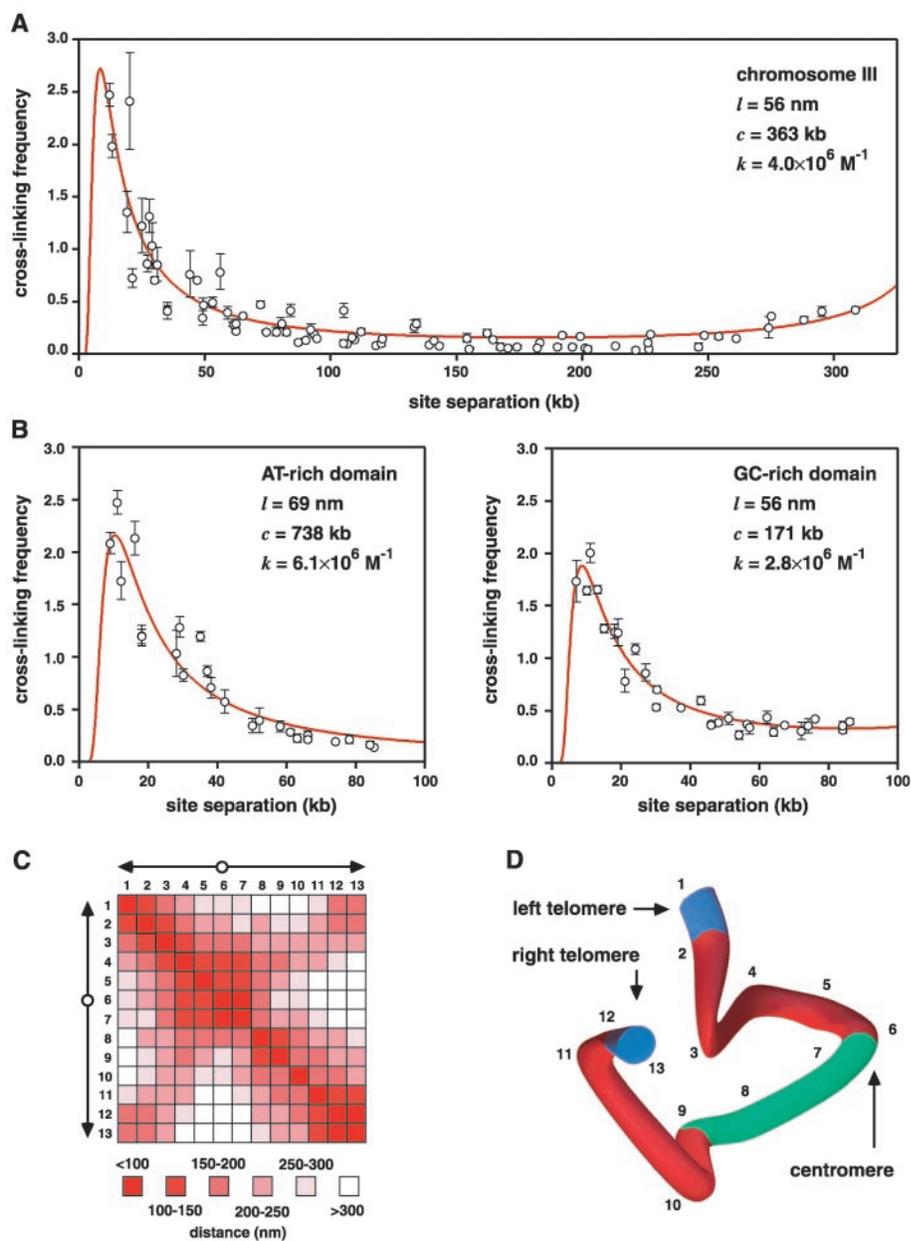

Fig. 3. Analysis of the structure of chromosome III during interphase. The 3C technology was applied to nuclei purified from haploid NKY2997 cells arrested in G_1 with α-factor. (**A**) All pairwise cross-linking frequencies between positions 1 and 13 shown in Fig. 1C were determined in triplicate, and the average and standard error of the mean are plotted against site separation. The fit to Eq. 2 is indicated by the continuous line, and the values for l, c, and k are indicated. (**B**) Cross-linking frequencies between a large number of sites located in the AT-rich domain (left panel) and the GC-rich domain (right panel) of the right arm of chromosome III were determined in triplicate, and the average and standard error of the mean are plotted against site separation. Fits to Eq. 2 are indicated by continuous lines, and the values for l, c, and k are indicated. (**C**) Schematic representation of the complete distance table of chromosome III. Distances were calculated for all 78 pairwise combinations of sites 1 to 13 (Fig. 1C) using cross-linking frequencies shown in (A). (**D**) Population-average 3D model of chromosome III, drawn with Truespace software. The model was calculated using the set of 78 distances shown in (C). The numbers correspond to the positions shown in Fig. 1C. The AT-rich region in the right arm (positions 6 to 9) is indicated in green, the GC-rich domains (position 2 to 6 and 9 to 12) are indicated in red, and the subtelomeric regions are indicated in blue.

these isochore domains might also differ in their basic structural properties. We analyzed a large number of interactions, for sites separated by 7 to 85 kb, within each given region. Results for the AT-rich and GC-rich domains of the right arm are shown in Fig. 3B. Fitting the data to Eq. 2 yielded very good fits with $l = 69$ nm, $c = 738$ kb, and $k = 6.1 \times 10^6$ M^{-1} for the AT-rich domain ($R^2 = 0.90$) and $l = 56$ nm, $c = 171$ kb, and $k = 2.8 \times 10^6$ M^{-1} for the GC-rich domain ($R^2 = 0.96$). The 75-kb GC-rich domain within the left arm of chromosome III was analyzed in a similar way, and values of $l = 62$ nm, $c = 117$ kb, and $k = 3.0 \times 10^6$ M^{-1} were obtained (9). Essentially the same results were obtained in exponentially growing *MAT***a** and *MAT*α cells (9).

The AT-rich domain exhibits little curvature ($c = 738$ kb) and thus behaves similarly to an unconstrained linear polymer. In contrast, the GC-rich domains exhibit apparent circular conformations ($c = 171$ kb and $c = 117$ kb). This difference is also reflected in a qualitative difference in the data of the two domains. In the case of the GC-rich domain, cross-linking frequencies of sites separated by more than 55 kb were higher (by a factor of 2 to 3) than expected for an unconstrained fiber, and they do not decrease with increasing site separation. In contrast, the interaction frequencies measured in the AT-rich domain continue to decrease with increasing site separation. These results show that the degree of apparent curvature varies by domain along the chromosome, in correlation with variation by domain in average base composition. The constraints responsible for the apparent curvature of GC-rich regions could be internal, via intrinsic bias in fiber path, or external (e.g., by tethering of parts of the chromosome to the nuclear envelope).

A twofold difference in the value of k was found when the AT- and GC-rich domains were compared, whereas an intermediate value of k was determined for the entire chromosome ($k = 4.0 \times 10^6$ M^{-1}). These data sets were obtained using the same DNA sample as PCR template. Thus, the different values of k may reflect real differences in the internal structure of the fiber (e.g., nucleosome density). We assume a homogeneous packing ratio, but small differences in packing ratio between these domains may exist, and in the current analysis this will also result in differences in the value of k. The AT-rich domain may be slightly stiffer than the rest of the chromosome ($l = 69$ nm versus 56 nm), but given the intrinsic uncertainties of a three-parameter fit with an error in l of about 10%, we do not know whether the difference in l between the AT- and GC-rich domains is significant.

Given a complete set of spatial distances between a number of sites along a chromosome, a 3D model of the chromosome can be developed. Such distances can be obtained from cross-linking frequency data in two steps. First, each intersite cross-linking frequency can be related to a relative local concentration by Eq. 1. Second, because the chromatin fiber is accurately described by the polymer model outlined above, each relative local concentration can be related to the squared average spatial distance $<r^2>$ (in nm^2) between two sites. For the FJC model this relation is given by

$$j = \left(\frac{3}{2\pi\langle r^2 \rangle} \right)^{3/2} \text{ or}$$

$$\sqrt{\langle r^2 \rangle} = 0.82 \times j_M^{-1/3} \left(\frac{\text{mol}}{\text{liter}} \right)^{1/3} \text{ nm} \quad (3)$$

which holds for both linear and circular polymers. Given a complete distance table for a number of sites along the length of the chromosome, a population-average 3D model of the chromosome can be calculated using Choleski decomposition of symmetric matrices (40).

The 78 measured cross-linking frequencies shown in Fig. 3A were converted to 78 average intrachromosomal distances [using $k = 4.0 \times 10^6$ M^{-1}; see (41)] and are schematically indicated in the distance matrix shown in Fig. 3C. These 78 distances constitute a complete distance table for the 13 positions shown in Fig. 1C and were used to calculate a population-average 3D model of chromosome III (Fig. 3D) [see (40)]. The AT-rich domain of the chromosome is indicated in green (positions 6 to 9), the GC-rich domains (positions 2 to 6 and 9 to 12) are indicated in red, and the subtelomeric regions are indicated in blue. The difference between each measured distance and the corresponding modeled distance was on average 6%, indicating that the 3D model fits the data very well.

The model clearly shows the circular behavior of the chromosome, with the two telomeres closely juxtaposed. The two telomeres are on average 170 nm apart. The GC-rich domains appear more strongly curved than the AT-rich domain, although the resolution of the model is not sufficient to make these domain-related constraints clearly visible. The model also indicates the presence of a sharp bend around the centromeric region (position 6), partly because interaction frequencies between positions 7 and 8 and positions 3 and 4 were higher than the corresponding fitted values in Fig. 3A. As a result of these deviations, the chromosome appears as a contorted ring and not as a homogeneous circular polymer.

The size of the model chromosome is in accord with several expectations. First, the largest distance is that between the centromere and the right telomere. Those sites are on average 330 nm apart, which is considerably less than the diameter of a haploid yeast nucleus (~1.3 to 1.8 μm). Second, the volume of chromosome III, estimated by assuming the chromosome to be a sphere with a diameter of 330 nm, is at most 2% of that of the nucleus, in agreement with the fact that chromosome III contains ~2.6% of the genome. Third, with a value of $l = 56$ nm (Fig. 3A), the longest chromosome arm of yeast, which is around 1 Mb, would have a maximum average extension of ~800 nm and thus would fit in the nucleus without the need to loop back. These observations indicate that the estimate of a packing ratio for yeast chromatin of 11.1 nm/kb is probably accurate. We also determined 3D models for chromosome III from exponentially growing *MAT***a** and *MAT*α cells, and similar models were obtained (9).

It is important to emphasize that this model is the population-average path of a very flexible chromosome. The actual trajectory of an individual chromosome at a given time will fluctuate around conformations similar to the one depicted in Fig. 3D.

The approaches described provide detailed information about the 3D organization of chromosomes and the nucleus in general. Potentially, microarray technology can be used to detect ligation products. Moreover, 3C technology can be applied to any organism for which genomic sequence information is available. When combined with other approaches, information about spatial organization can be related to specific chromosome functions.

References and Notes

1. C. L. Woodcock, S. Dimitrov, *Curr. Opin. Genet. Dev.* **11**, 130 (2001).
2. A. Wolffe, *Chromatin, Structure and Function* (Academic Press, San Diego, CA, 1998).
3. E. D. Andrulis, A. M. Neiman, D. C. Zappulla, R. Sternglanz, *Nature* **394**, 592 (1998).
4. T. Cremer, C. Cremer, *Nature Rev. Genet.* **2**, 292 (2001).
5. S. M. Gasser, *Cell* **104**, 639 (2001).
6. D. M. Cimbora, M. Groudine, *Cell* **104**, 643 (2001).
7. Supplementary details of experimental procedures are available on *Science* Online at www.sciencemag.org/cgi/content/full/295/5558/1306/DC1.
8. S. Keeney, N. Kleckner, *Genes Cells* **1**, 475 (1996).
9. J. Dekker, K. Rippe, M. Dekker, N. Kleckner, data not shown.
10. The cross-linking is mediated by DNA bound proteins, and thus the local protein density will affect the cross-linking efficiency. As is clear from Fig. 3, the parameter k (in Eq. 1), which reflects this efficiency, appears to vary slightly along the length of chromosome III. In addition, a high local digestion efficiency or a favorable geometry of the cross-linked intermediate will also influence the amount of ligation product formation and thereby the apparent cross-linking frequency. However, as is apparent in Figs. 2 and 3, these effects are minor.
11. M. Gotta *et al.*, *J. Cell Biol.* **134**, 1349 (1996).
12. E. Trelles-Sticken, M. E. Dresser, H. Scherthan, *J. Cell Biol.* **151**, 95 (2000).
13. Q. W. Jin, E. Trelles-Sticken, H. Scherthan, J. Loidl, *J. Cell Biol.* **141**, 21 (1998).
14. Q. W. Jin, J. Fuchs, J. Loidl, *J. Cell Sci.* **113**, 1903 (2000).
15. B. M. Weiner, N. Kleckner, *Cell* **77**, 977 (1994).
16. S. M. Burgess, N. Kleckner, B. M. Weiner, *Genes Dev.* **13**, 1627 (1999).

17. S. M. Burgess, N. Kleckner, *Genes Dev.* **13**, 1871 (1999).
18. D. Zickler, N. Kleckner, *Annu. Rev. Genet.* **32**, 619 (1998).
19. _____, *Annu. Rev. Genet.* **33**, 603 (1999).
20. R. Padmore, L. Cao, N. Kleckner, *Cell* **66**, 1239 (1991).
21. N. Hunter, N. Kleckner, *Cell* **106**, 59 (2001).
22. O. Kratky, G. Porod, *Rec. Trav. Chim.* **68**, 1106 (1949).
23. P. J. Flory, *Statistical Mechanism of Chain Molecules* (Wiley, New York, 1969).
24. K. Rippe, P. H. von Hippel, J. Langowski, *Trends Biochem. Sci.* **20**, 500 (1995).
25. L. Ringrose, S. Chabanis, P. O. Angrand, C. Woodroofe, A. F. Stewart, *EMBO J.* **18**, 6630 (1999).
26. K. Rippe, *Trends Biochem. Sci.* **26**, 733 (2001).
27. P. Hahnfeldt, J. E. Hearst, D. J. Brenner, R. K. Sachs, L. R. Hlatky, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 7854 (1993).
28. S. E. Gerchman, V. Ramakrishnan, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7802 (1987).
29. K. E. van Holde, *Chromatin* (Springer, Heidelberg, 1989).
30. The flexibility of chromatin in higher eukaryotes is not firmly established. FISH analysis of mammalian chromosomes yielded values of l [l = 137 to 440 nm (*27*) and l = 196 to 272 nm (*42*)], whereas detailed in vitro experiments with single chromatin fibers from chicken erythrocytes yielded lower values of l [l = 60 nm (*43*)].
31. The local concentration is maximal for sites separated by ~1.7 times the statistical segment length l. We estimate that the contour length of the yeast chromatin fiber is 11.1 nm/kb. A segment length of 56 to 69 nm (as determined in Fig. 3, A and B) will then correspond to 5.0 to 6.2 kb. The local concentration is therefore maximal for sites separated by 8.5 to 10.5 kb.
32. P. B. Moens, R. E. Pearlman, *Bioessays* **9**, 151 (1988).
33. Y. Blat, N. Kleckner, *Cell* **98**, 249 (1999).
34. S. Laloraya, V. Guacci, D. Koshland, *J. Cell Biol.* **151**, 1047 (2000).
35. P. M. Sharp, A. T. Lloyd, *Nucleic Acids Res.* **21**, 179 (1993).
36. B. Dujon, *Trends Genet.* **12**, 263 (1996).
37. F. Baudat, A. Nicolas, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5213 (1997).
38. J. L. Gerton *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11383 (2000).
39. G. P. Holmquist, *Am. J. Hum. Genet.* **51**, 17 (1992).
40. Given a table of all distances between n + 1 points, one can find a constellation of n + 1 points in three dimensions, satisfying this distance table, with the use of the following algorithm, which is based on Cholesky decomposition of symmetric matrices. The distance table can be converted to a symmetric n × n matrix **M** that contains all the inner products of the n + 1 points in n-space. If the distance table describes a 3D structure, **M** should have rank 3. In general, however, **M** does not have rank 3 because of noise in the data. By slightly moving **M** in n^2-space, we obtain a matrix **M′**, close to **M**, which does have rank 3. This is done by setting all but the three largest eigenvalues to zero. Indeed, three eigenvalues we determined were considerably larger than zero, and the other had very low values, which indicates that the **M** is very close to a matrix with rank 3. This is strong support for the accuracy of the distance calculation. Cholesky decomposition of **M′** will yield the positions of the n + 1 points in three dimensions, and the distances between those points will nearly satisfy the experimental distance table. By using **M′** instead of **M** we introduce an error, which is related to the error in the data of the experimental distance table. The distance between **M** and **M′** is a measure for this error and was found to be only about 6%.
41. The FJC model will give a reasonable approximation of j_M for values of s larger than 3 to 4 times the statistical segment length, which, for a value of l = 56 to 69 nm and a contour length of 11.1 nm/kb, corresponds to 15 to 20 kb. For sites separated by less than 3 but more than 1.7 statistical segments (corresponding to site separation distances of 8.5 to 15 kb), Eq. 3 will yield a value of r that is slightly too high, but the highest deviation is about 20% (*9*). Because only 2 of the 78 site separations were smaller than 15 kb (12 and 13 kb), the set of 78 cross-linking frequencies presented in Fig. 3A could be used to estimate the average spatial distances between the respective sites according to Eqs. 1 and 3 with good accuracy. For k the chromosome-wide average value of 4.0 × 10⁶ M^{-1} was used, which was determined by fitting the chromosome-wide data set to Eq. 2. Because the analysis of AT- and GC-rich domains (Fig. 3B) revealed small local variations in the value of k, this might introduce small errors (at most 10%) in certain regions of the chromosome.
42. J. Y. Ostashevsky, C. S. Lange, *J. Biomol. Struct. Dyn.* **11**, 813 (1994).
43. Y. Cui, C. Bustamante, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 127 (2000).
44. In some cases a very small amount of ligation product was detected in the absence of cross-linking, but this was always 5% or less than the amount detected after cross-linking with 1% formaldehyde.
45. We thank Kleckner lab members for discussions. We are especially grateful to J. M. Perry and A. J. M. Walhout for valuable comments on the experiments and the manuscript. Part of the work by K.R. was done at the division "Biophysik der Makromoleküle" of the Deutsches Krebsforschungszentrum in Heidelberg. Supported by EMBO long-term fellowship ALTF-223-1998 and a fellowship from the Medical Foundation/Charles A. King Trust (J.D.), the Volkswagen Foundation program "Junior Research Groups at German Universities" (K.R.), and NIH grants GM44794 and GM25326 (N.K.).

A Class of Potent Antimalarials and Their Specific Accumulation in Infected Erythrocytes

Kai Wengelnik,[1] Valérie Vidal,[2] Marie L. Ancelin,[1] Anne-Marie Cathiard,[1] Jean Louis Morgat,[3] Clemens H. Kocken,[4] Michèle Calas,[2] Socrates Herrera,[5] Alan W. Thomas,[4] Henri J. Vial[1]*

During asexual development within erythrocytes, malaria parasites synthesize considerable amounts of membrane. This activity provides an attractive target for chemotherapy because it is absent from mature erythrocytes. We found that compounds that inhibit phosphatidylcholine biosynthesis de novo from choline were potent antimalarial drugs. The lead compound, G25, potently inhibited in vitro growth of the human malaria parasites *Plasmodium falciparum* and *P. vivax* and was 1000-fold less toxic to mammalian cell lines. A radioactive derivative specifically accumulated in infected erythrocytes to levels several hundredfold higher than in the surrounding medium, and very low dose G25 therapy completely cured monkeys infected with *P. falciparum* and *P. cynomolgi*.

The widespread resistance of malaria parasites to most common antimalarials, and cross-resistance to structurally unrelated drugs, emphasize the need for new chemotherapeutic targets (*1*). When parasitizing red blood cells, the parasite invests heavily in membrane biogenesis, making phospholipid (PL) biosynthesis essential for intraerythrocytic development (*2*), whereas these activities are absent from mature uninfected erythrocytes (*3*). Phosphatidylcholine (PC) is the major PL present in infected erythrocytes. It is mainly synthesized from plasma-derived choline by the parasite enzymatic machinery and provides an attractive target for antimalarial chemotherapy (*2*).

To inhibit PC synthesis in malaria parasites, we designed and synthesized a series of compounds based on their capacity to mimic choline structure (*4*). They were optimized for in vitro antimalarial activity, and there was a very close correlation between the inhibition of parasite growth in vitro and specific inhibition of parasite PC biosynthesis de novo from choline (*5*). The compounds do not interfere with other types of PL synthesis, such as that of phosphatidylethanolamine or macromolecules like DNA (*5*). Toxic activity is highest against mature parasites, when PL synthesis is maximal (*5*). We hypothesize that the most likely drug target is the choline transporter, either on the surface of the infected erythrocyte or in the plasma membrane of the intracellular parasite (*6*). Structure-activity relation (SAR) studies revealed the importance of the cationic group volume. Duplication of this group (bisquaternary ammonium salts) separated by a sufficiently long alkyl chain markedly increases antimalarial activity to within the low-nanomolar range (*7*). This defined the lead compound, G25 [1,16-hexadecamethylenebis(*N*-methylpyrro-

[1]CNRS UMR 5539, CP 107, [2]CNRS UMR 5810, CP 22, Université Montpellier II, Place E. Bataillon, 34095 Montpellier Cedex 5, France. [3]CRBA CNRS UMR 5473, Université Montpellier I, Faculté de Pharmacie, 15 Av. Charles Flahault, 34060 Montpellier, France. [4]Department of Parasitology, Biomedical Primate Research Centre, Postbox 3306, 2280 GH Rijswijk, Netherlands. [5]Fundacion Centro de Primates, Universidad del Valle, Cali, Colombia.

*To whom correspondence should be addressed. E-mail: vial@univ-montp2.fr

Capturing Chromosome Conformation
Job Dekker, Karsten Rippe, Martijn Dekker and Nancy Kleckner
(February 15, 2002)
Science **295** (5558), 1306-1311. [doi: 10.1126/science.1067799]

Editor's Summary

Article Tools	Visit the online version of this article to access the personalization and article tools: http://science.sciencemag.org/content/295/5558/1306
Permissions	Obtain information about reproducing this article: http://www.sciencemag.org/about/permissions.dtl