INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O.B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONES: 224281/2/3/4/5/6
CABLE: CENTRATOM - TELEX 460392-1

SMR/111 - 9

SECOND SUMMER COLLEGE IN BIOPHYSICS

30 July - 7 September 1984

1. Separation of large DNA molecules by pulsed field gradient gel electrophoresis.
2. Fluorescence spectroscopic methods for studying conformational changes.
3. Crosslinkers as tools to study DNA packaging and repair.
4. Structure of the E. coli 16S ribosomal RNA.
5. Topological aspects of chromatin structures.

C.R. CANTOR
College of Physicians & Surgeons
Department of Human Genetics and Development
Columbia University
New York, NY 10032
U.S.A.

LECTURE NOTES: C.R. CANTOR AUGUST 1984 TRIESTE SUMMER SCHOOL

1. Separation of large DNA molecules by pulsed field gradient gel electrophoresis.

Some of the unusual properties of large DNA molecules will be described. The principles and limitations of conventional techniques for handling such molecules will be explained. Then the principles of the new pulsed field gradient technique will be described. This allows separations of DNA molecules as large as 4,000,000 base pairs. Applications of pulsed field gels to genetic mapping will be discussed in yeast, trypanosomes, and man.

References

1. D. Schwartz and C. Cantor (1984) Cell 37:67.
2. L. Van der Ploeg et al., (1984) Cell 37:77.
3. C. Cantor and P. Schimmel (1980) Biophysical Chemistry 2, pages 655-659 and 670-682.
4. M. McClelland et al., (1984) PNAS 81:983.

Most DNA molecules are too large to be separated by conventional electrophoretic means. They must be cut into fragments with restriction endonucleases but then much information is lost.

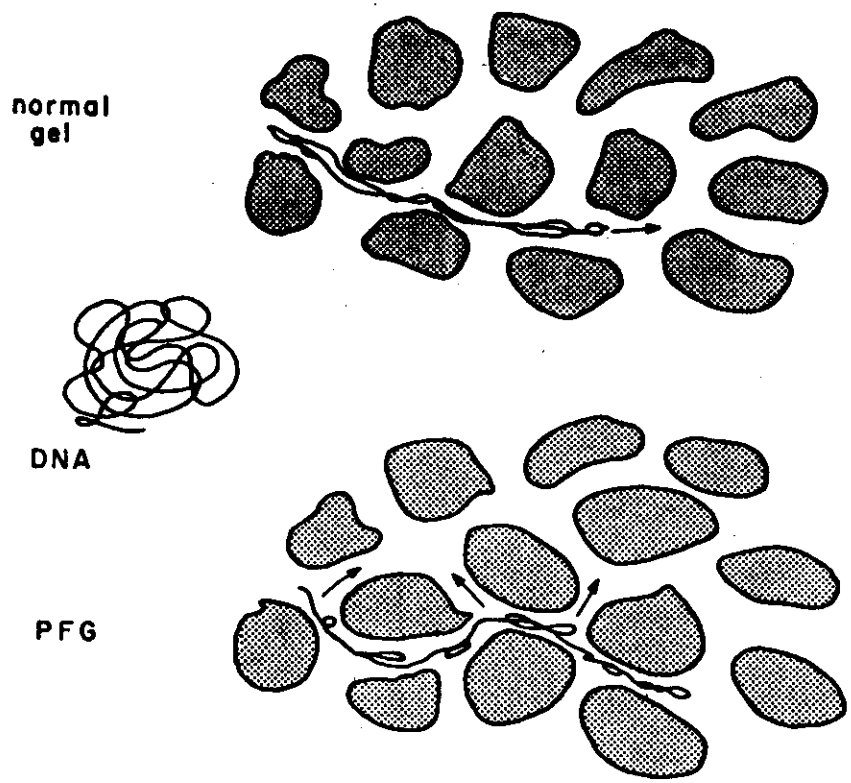
Numbers and sizes of DNA molecules

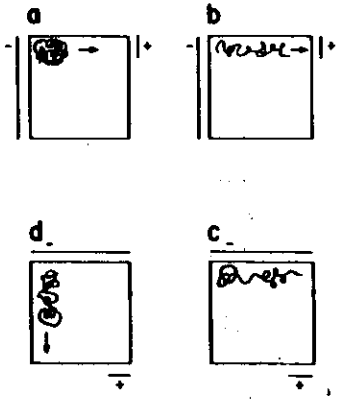
E. coli chromosome 4×10^6 bp 1-3 molecules per cell

typical bacterial plasmid $4 - 8 \times 10^3$ bp 1-1000 molecules per cell

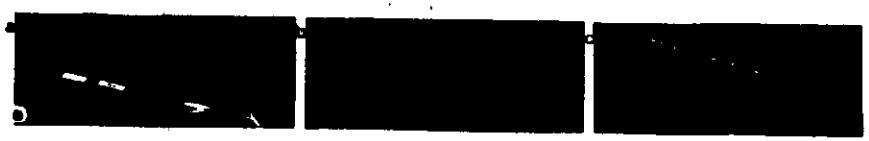
small virus	4,000 bases
lambda virus	50,000 bp
T4 virus	160,000 bp
G virus	800,000 bp
yeast chromosomes	17 per cell (haploid)
size range	300×10^3 bp to 2500×10^3 bp
human chromosomes	46 per cell (diploid)
size range	50×10^6 bp to 250×10^6 bp
mitochondria	15×10^3 bp to 75×10^3 bp 2 - 50 molecules per organelle 2 - 1000 organelles per cell (10^7 in frog egg) 1 - 99% total cellular DNA
chloroplast	$130 - 180 \times 10^3$ bp 20 - 80 molecules per organelle 2 - 40 organelles per cell 7 - 15% total cellular DNA
SV40 virus	5×10^3 bp
Adenovirus	40×10^3 bp
Herpes Simplex	150×10^3 bp

Schematic illustration of a DNA molecule in solution and in an agarose gel.

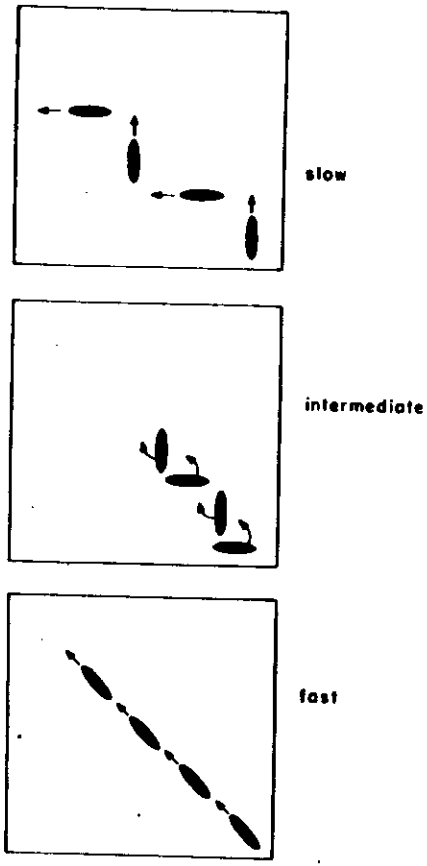




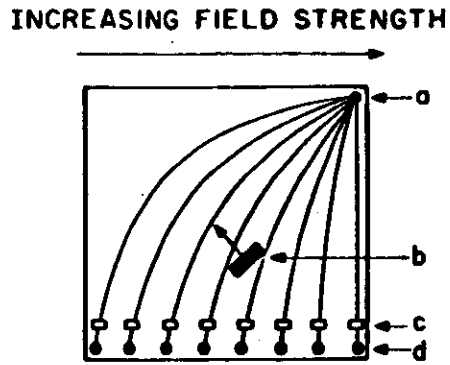
Two DNAs, one 40 kb the other 160 kb, separated with 2, 10 and 100 sec. pulses



Schematic illustration of the mechanism of PFG

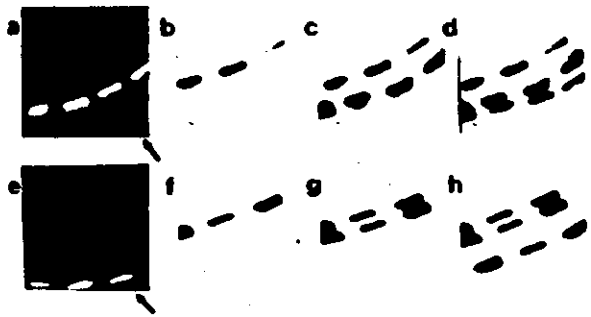


A band of DNA molecules moving in an inhomogeneous electric field



Typical PFG result for yeast chromosomal DNAs with 45 sec pulses. Markers are 48, 160 and 800 kb; alternate lanes contain two "wild type" yeast strains.

Blotting with cloned probes to assign genes to chromosomes. lanes a-d are 17 second pulses while e-h are 35 sec. Autoradiographs show chromosome 1 (b,f), 1 and 3 (c,g) 1,3, and 5 (d,h).



Some genes are controlled by shifting around DNA sequences

Trypanosoma brucei variant surface glycoproteins (VSGs)

1000 different VSG genes, scattered through the genome
expressed one at a time

imprecise but non random order of appearance

VSG sequence N----------C
variable conserved

VSG crosshybridization

No at low stringency

Expression site

promoter element coding sequence

AAA xxxxxxxxxxxxxx telomere

const. exon untranslated
variable exon translated

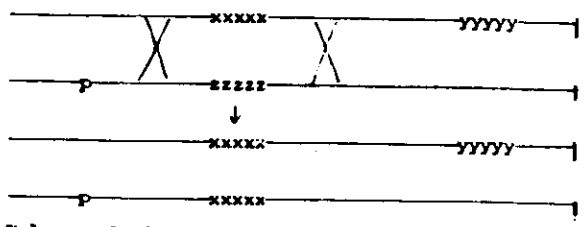
mature RNA AAAxxxxxxxx

200 copies of promoter element linked in blocks of at least ten

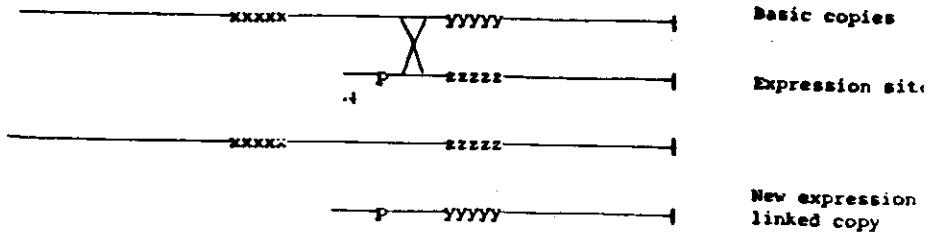
1 to 10 expression sites

VSG activation models

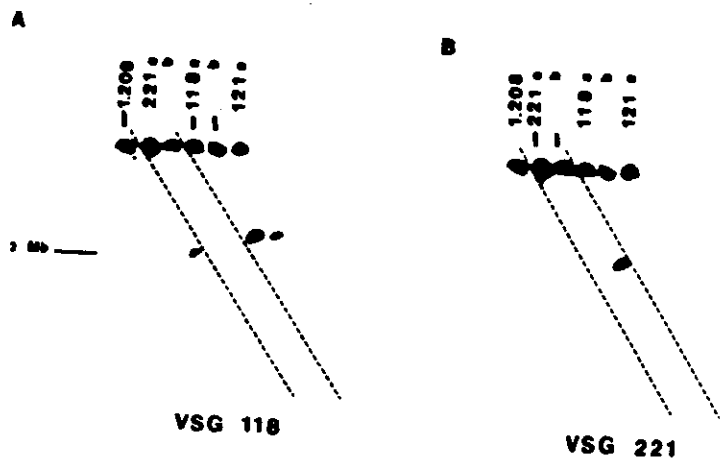
1. Gene conversion



2. Telomere Exchange



Evidence for interchromosomal gene conversion accompanying some shifts in VSG gene expression. Labels at top indicate which gene the particular sample of trypanosomes is expression; labels at bottom indicate the probe used.



For human DNA it is necessary to use specific cutting schemes to reduce intact chromosomal DNA into pieces that fall in the current size range of PFG. An enzyme with a specific recognition site n bases long will produce pieces on the average 4^n bp in size. (In practice the nonrandomness of DNA sequences makes this calculation quite approximate. For example each CpG sequence in a higher eukaryote occurs only 1/4 the expected frequency so it is as though the recognition site were actually 1 base longer.)

Conventional 6 bp specific nucleases	4,000 bp
Nuclease with 1 CpG	16,000 bp
8 bp specific nucleases like SfiI	64,000 bp
8 bp specific nuclease like Not 1 with 2 CpG's	1,000,000 bp
8 bp McClelland scheme shown below	1,000,000 bp
10 bp McClelland scheme shown below	16,000,000 bp

5' T-C-G-A-T-C-G-A 3'
3' A-G-C-T-A-G-C-T 5'
methylase ↓ *M. Taq* I
5' T-C-G-^{m}A - T-C-G-^{m}A 3'
3' ^{m}A-G-C - T-^{m}A-G-C - T 5'
restriction endonuclease ↓ *Dpn* I
5' T-C-G-^{m}A 3' 5' T-C-G-^{m}A 3'
3' ^{m}A-G-C - T 5' 3' ^{m}A-G-C - T 5'

5' A-T-C-G-A-T-C-G-A-T 3'
3' T-A-G-C-T-A-G-C-T-A 5'
methylase ↓ *M. Cla* I
5' A- T-C-G-^{m}A - T-C-G-^{m}A-T 3'
3' T-^{m}A-G-C - T-^{m}A-G-C - T-A 5'
restriction endonuclease ↓ *Dpn* I
5' A - T-C-G-^{m}A 3' 5' T-C-G-^{m}A-T 3'
3' T-^{m}A-G-C - T 5' 3' ^{m}A-G-C - T-A 5'

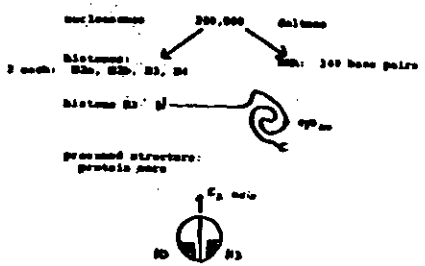
2. Fluorescence spectroscopic methods for studying conformational changes.

Nucleosomes will be used to illustrate the various ways in which fluorescence spectroscopy is employed to study the structure of macromolecular assemblies. Energy transfer provides measurements of specific distances between pairs of fluorescent labels. Anisotropy or polarization measurements provide information on size shape and flexibility. Dynamic quenching allows the accessibility of the fluorescent probe to the solvent to be determined. Excimer formation reveals when two fluorescent probes are in direct contact. Any of these measurements can also be used to monitor the kinetics of equilibria of conformational changes.

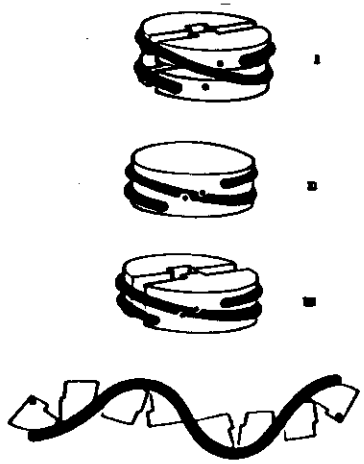
References

1. J. Daban and C. Cantor (1982) J. Mol.Biol.156:749. and J. Mol. Biol. 156:771.
2. A. Dieterich et al. (1979) J. Mol. Biol.129:587.
3. C. Prior et al. (1988) Cell 20:597.
4. C. Cantor and P. Schimmel (1988) Biophysical Chemistry 2, pages 433-465.

The unique cysteine 110 residue of histone H3 allows specific derivatives of nucleosomes to be prepared easily.



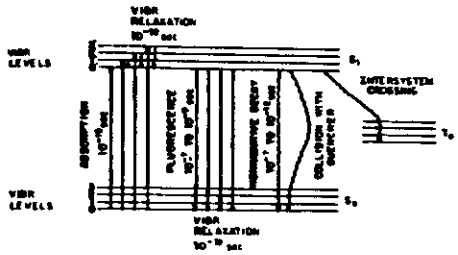
Labels on these cysteines can be used to reveal changes in nucleosome structure like those shown schematically below



They can also be used to follow nucleosome assembly both in vitro and in vivo.

The key to understanding various flourescence techniques is the time scale on which several processes compete to relax an electronically excited state.

(a)



The rate of fluorescence can be calculated from the electronic absorption intensity.

$$1/\tau_F = k_F = \text{const.}\,\bar{\nu}^2 \int \frac{\epsilon(\bar{\nu})\,d\bar{\nu}}{\bar{\nu}}$$

The quantum yield is the fraction of excited states relaxed by fluorescence. It is calculated directly from the ratio of various rates.

$$\phi = \frac{k_F}{k_F + k_{nr} + k_{st} + k_q(Q)}$$

Excited singlet states show first order decay kinetics

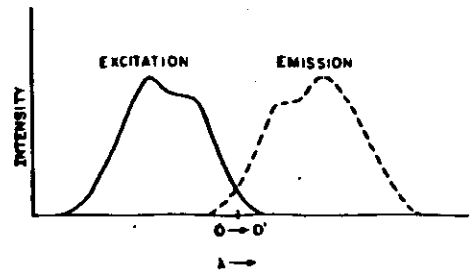
$$S_1(t) = S_1(0)\; e^{-t/\tau_D}$$

where τ_D, the observed fluorescence decay time, is just the reciprocal of the rate of all the parallel decay paths:

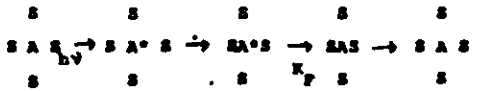
$$1/\tau_D = k_F + k_{nr} + k_{st} + k_q(Q)$$

The fluorescence spectrum in a vacuum should just be a reflection of the absorption spectrum.

(b)

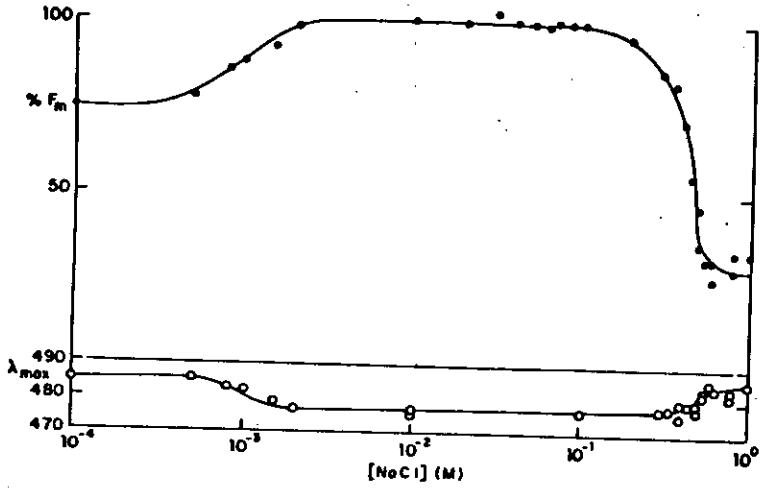


However solvent effects can change this, dramatically.



Many other chemical and physical processes can happen during the lifetime of the excited singlet such as conformational changes or proton transfer reactions. The result is that fluorescence intensities and spectra are much more environmentally sensitive than absorption spectra.

Nucleosome conformational changes are revealed by changes in intensity and spectral shape.



Collisional processes can compete to relax the excited state.

$$Q + S_1 \rightarrow Q^* \quad \text{or} \quad Q + S_o$$

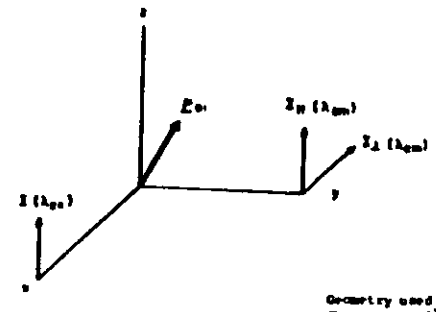
$$\phi_o/\phi = 1 + k_q(Q)/(k_F + k_{nr} + k_{st})$$

Collisional quenching rates reveal that nucleosome conformational changes lead to marked variation in the accessibility of labels on cysteine 110 to solvent.

Acrylamide Quenching of Fluorescence of IAEDANS-Nucleosomes

NaCl concentration (M)	k_q ($mole^{-1}sec^{-1}x10^7$)
1×10^{-4}	34
1×10^{-2}	8
3.5×10^{-1}	18
6.0×10^{-1}	93

When fluorescence polarization is considered, the fluorescence of a rigid system is anisotropic.



Geometry used in a fluorescence anisotropy experiment. Shown are I (λ_{ex}), the polarization direction of the exciting light, $I_\parallel$ and $I_\perp$ two polarization directions used to measure emitted light, and $\vec{\mu}_{01}$, the transition dipole moment of one chromophore of the sample.

$$I_N = 3I_\perp$$

The anisotropy, r, and polarization, P, are two related measures of this effect.

$$P = (I_N - I_\perp) / (I_N + I_\perp)$$

$$r = (I_N - I_\perp) / (I_\parallel + 2I_\perp)$$

For rigid systems P = 0.5, r = 0.4.

For non-rigid systems, the anisotropy depends on the relative rates of rotation and emission. One can measure the time dependence of the anisotropy or the average steady state value.

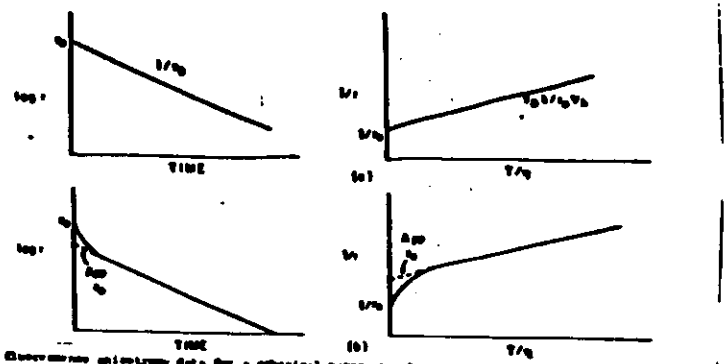
$$r(t) = r_o e^{-t/\tau_R}$$

$$\tau_R = V_h \eta / kT$$

$$\frac{1}{r} = \frac{1}{r_o}(1 + \tau_D/\tau_R)$$

$$\frac{1}{P} - \frac{1}{3} = \left[\frac{1}{P_o} - \frac{1}{3}\right](1 + \tau_D/\tau_R)$$

$$\frac{1}{r} = \frac{1}{r_o}(1 + \tau_D \frac{k}{V_h} \frac{T}{\eta})$$



Typical fluorescence anisotropy data for a spherical macromolecule. (a) A rigidly attached fluorophore: at left is the decay of fluorescence anisotropy; at right is a Perrin plot of static anisotropy versus temperature/viscosity. Other parameters are defined in equation (10). The limiting anisotropy, r_o, is also indicated. (b) A fluorophore attached by a partially flexible link. The two panels correspond to those in part (a). The apparent r_o, unaffected by flexibility is shown along with the real r_o.

Anisotropy measurements on nucleosomes show that the low and high salt forms have flexible domains while the compact intermediate salt form behaves like a rigid, somewhat assymetric object.

Fluorescence Polarization of IAEDANS-Nucleosomes

Condition	τ_D (nsec)	V_n (cm^3/mole)
0.1mM Tris, .02mM EDTA	26.4	6.57×10^6
10mM Tris, 2mM EDTA	145	3.61×10^5
0.6M NaCl	6.2	1.55×10^4
calculated	104-168	$8.2\text{-}2.9 \times 10^5$

Energy transfer can be used to measure the distance between two fluorescent molecules. It arises by a resonance mechanism.

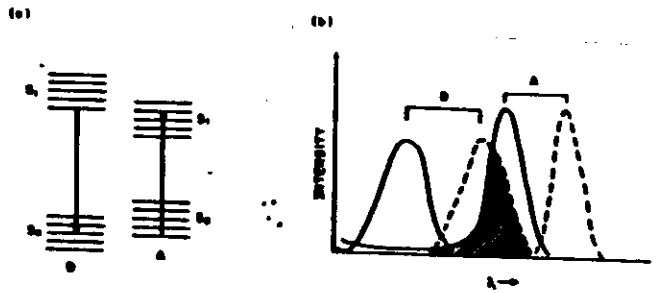


Fig. 2 - Singlet-singlet energy transfer. (a) Energy levels of a typical donor (D) and acceptor (A) pair showing one of the transitions which can be coupled by a long range resonant interaction. (b) Absorption (———) and emission (- - - -) spectra of the donor and acceptor pair showing the spectral overlap (shaded).

The rate of energy transfer is dependent on the distance, R, and the characteristic transfer distance for the dye pair, R_o.

$$k_T = \frac{1}{\tau_D^o} \frac{R_o^6}{R^6}$$

$$R_o = \text{const } (k^2 J_n^4 \phi_D^o)^{1/6}$$

The energy transfer efficiency is the fraction of excited donors relaxed by energy transfer.

$$E = k_T / (k_F + k_{nr} + k_{st} + k_T) = R_o^6 / (R_o^6 + R^6)$$

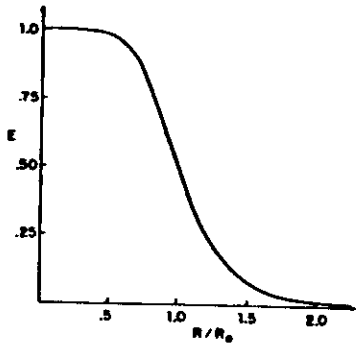


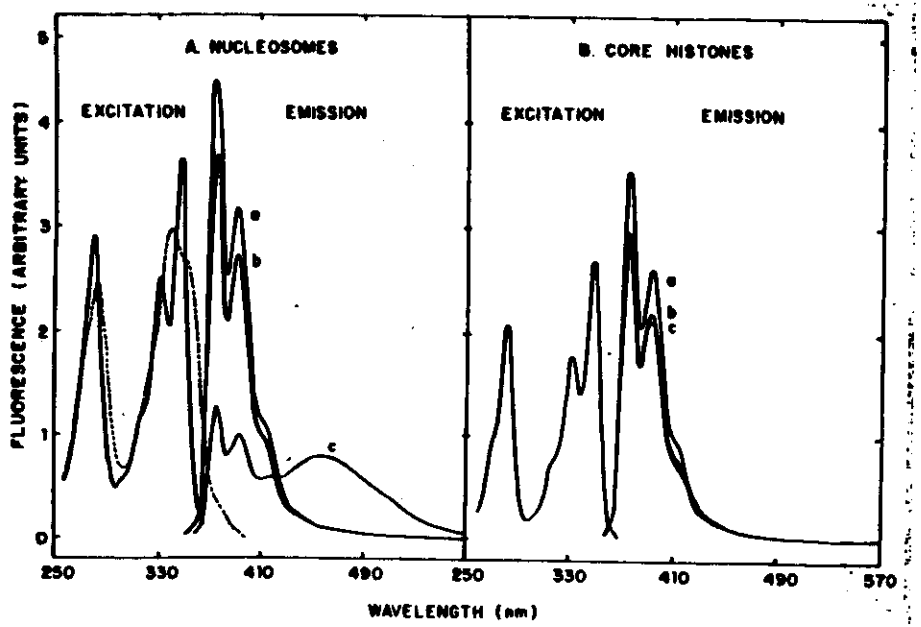
Fig. 3 - Energy transfer efficiency predicted by the Forster theory as a function of the distance between chromophores, R, and the characteristic transfer distance, R_o.

With nucleosomes, energy transfer measurements show that the distance between two labeled H3 cysteine 110 residues is markedly salt dependent.

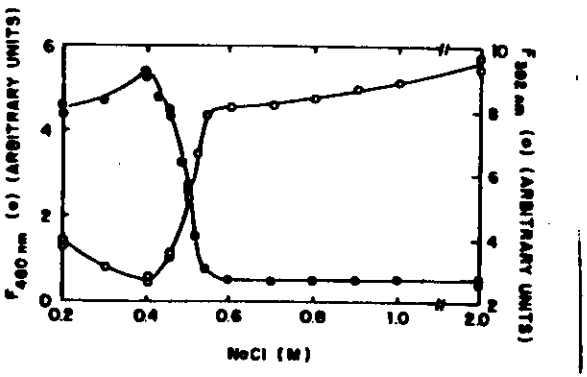
Energy Transfer Results

Condition	E_o	R/R_o	R
0.025mM EDTA	.32±.08	1.13±.07	48±3
0.025mM EDTA; 2mM NaCl	.46±.09	1.03±.07	43±3
10mM Tris; 2mM EDTA	1.00±.07	$\leq$.64	$\leq$30
0.35M NaCl	1.00±.07	$\leq$.64	$\leq$30
0.6M NaCl	0	$\geq$1.9	$\geq$70

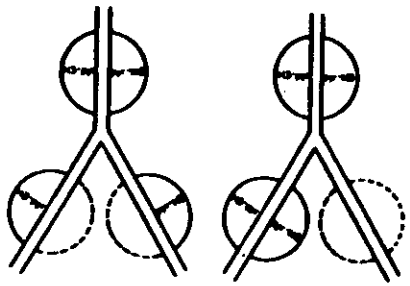
At moderate ionic strength the distance between the two cysteines is too small to measure by energy transfer. Instead the pyrene excimer can be used to show that these residues are nearly in contact.



a. 2.0M NaCl
b. 0.6M NaCl
c. 0.2M NaCl



The pyrene excimer can also be used to monitor the in vivo replication mechanism of nucleosomes.



(A) BLUE EMISSION (B) GREEN EMISSION

Figure 10.

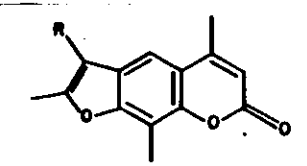
3. Crosslinkers as tools to study DNA packaging and repair.

Psoralens are useful regeants to study nucleic acid structure. Simple psoralens form monoadducts and crosslinks in double stranded helices. More complex psoralens can be used to produce DNA-protein crosslinks or links between two separate DNA regions. Techniques exist that allow the placement of site specific crosslinks. A number of different mechanisms can potentially be used to repair crosslinks in DNA in vivo. Some of these mechanisms are error prone pathways induced by severe DNA damage. The spectrum of mutations produced by these pathways is beginning to be unraveled.

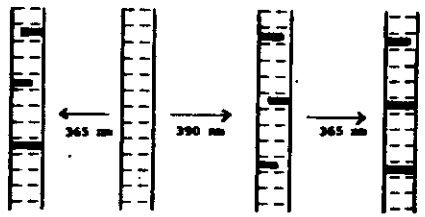
References

1. D. Schwartz, et al. (1983). Cold Spring Harbor Symp. Quant. Biol. 47: 189.
2. R. Haas et al. (1982) J. Mol. Biol. 159:71.
3. P. Chatterjee and C.R. Cantor (1978). Nucleic Acids Res. 5: 3619.
4. W. Saffran et al. (1982). PNAS 79: 4594.
5. W. Saffran and C.R. CANTOR (1984). J. Mol. Biol. in press.

Psoralens intercalate into DNA. The wavelength used for irradiation, and the local DNA sequence, will determine whether monoadducts or crosslinks are favored.



Compound, R =	Abbreviation
$HO\text{-}CH_2\text{-}$	HMT
$NH_3^+\text{-}CH_2\text{-}$	AMT



Specialized psoralens have been made that can serve as DNA-protein crosslinkers, DNA-DNA crosslinkers, or site directed DNA crosslinkers.

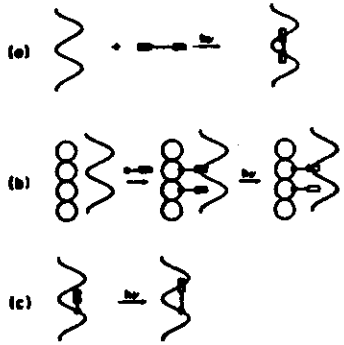


Figure 2. Schematic Illustrations of the potential uses of specialized psoralens for studies on the arrangement of packaged DNA. (a) Direct DNA-to-DNA cross-linking using BTPD; (b) cross-linking between a viral capsid (or any other protein) and adjacent DNA using HSP; (c) DNA affinity labeling using SSP.

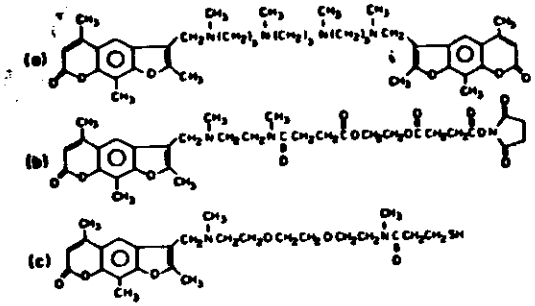


Figure 3. Three psoralens designed for specialized cross-linking studies on large DNAs. (a) BTPD, a bis-psoralen; (b) HSP, a protein–nucleic acid cross-linker; (c) SSP, a site-directed psoralen that can be attached to mercurated pyrimidines previously incorporated into DNA enzymatically.

Breakable bis psoralens seem especially promising for studies on DNA packaging since they allow the pattern of proximity of DNA regions to determined by simple, diagonal gel electrophoresis.

Site directed crosslinks can be placed into DNA double helices.

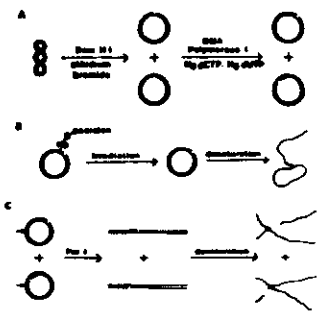


Fig. 5. Preparation of HSP crosslinks. (A) Intercalation at the BamHI site. (B) Crosslinking with HSP. (C) Cutting with Pst I.

Replication

DNA replication is not symmetric for the two strands

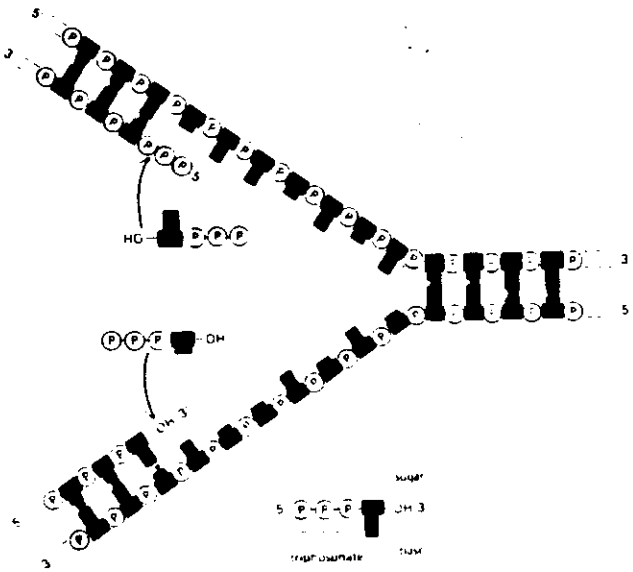


Figure 5–30 Intuitively the simplest mechanism for DNA replication would be the incorrect scheme shown here. Both daughter DNA strands would grow continuously, requiring both 5′-to-3′ and 3′-to-5′ nucleotide polymerization, as indicated.

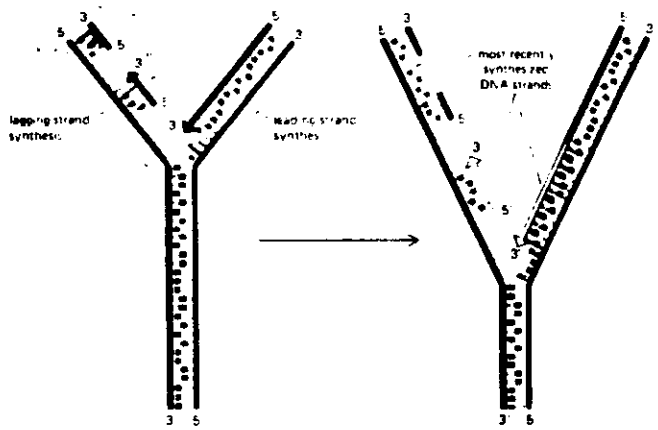


Figure 5–31 The structure of an actual replication fork, in which both daughter DNA strands are synthesized in the 5′-to-3′ direction, thereby requiring that the DNA synthesized on the lagging strand be made as a series of short pieces.

DNA polymerases can edit out errors

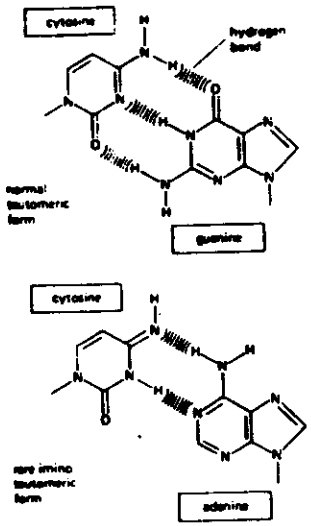


Figure 5–32 An example of an expected rare incorrect base pair, when cytosine is in its unfavored tautomeric form, it can form effective hydrogen bonds with adenine.

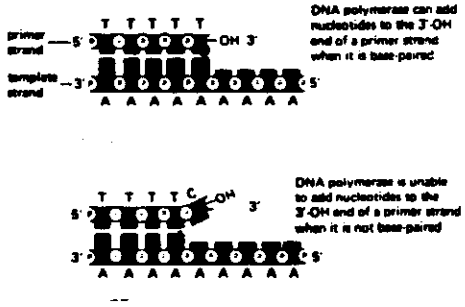


Figure 5–33 Examples of two synthetic DNA molecules that have been tested as primer templates for DNA polymerase. In such tests, the 3′-OH end of a primer strand can be extended only when it is base-paired.

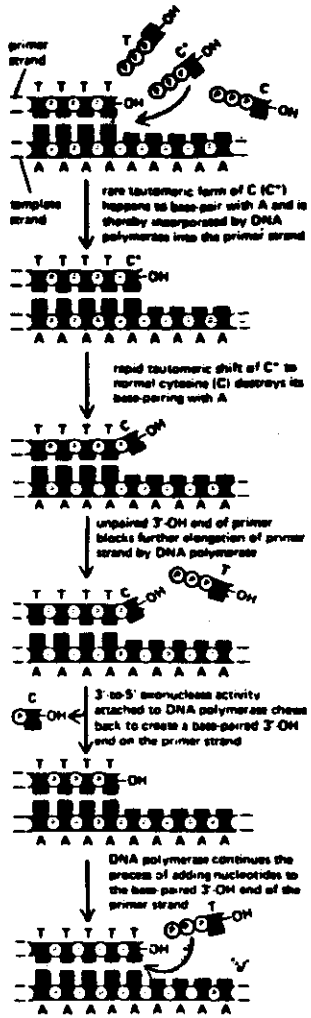


Figure 5–34 Illustration of the proofreading process known to remove errors during DNA synthesis catalyzed by procaryotic DNA polymerases. As yet, it has not been possible to demonstrate a mechanism of this type for most DNA polymerases isolated from higher eucaryotes.

Editing and initiating are incompatible. Erasable priming mechanisms are used.

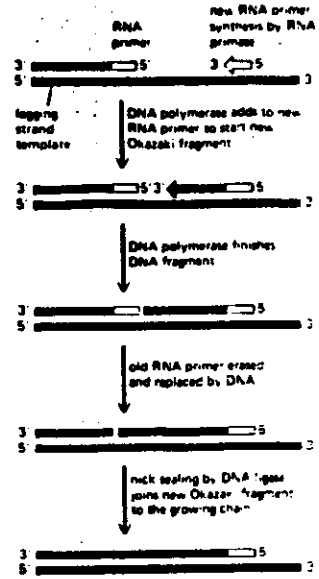
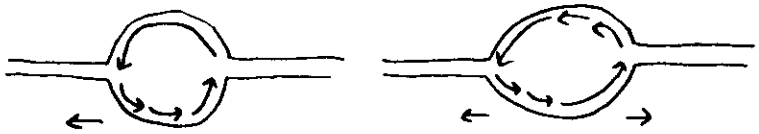
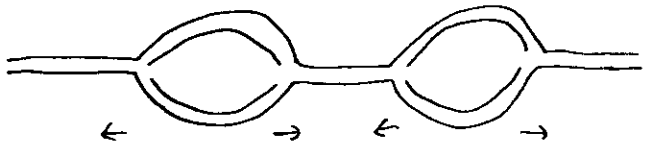


Figure 5–36 The steps involved in the synthesis of each DNA fragment on the lagging strand. In eucaryotes the RNA primers are made at intervals spaced by about 200 nucleotides on the lagging strand and each RNA primer is 10 nucleotides long. The start signals for the RNA primase have not yet been characterized, but if a specific template nucleotide sequence is involved, it must be a very short one.

Replication forks initiate at a replication origin
Forks can be unidirectional or bidirectional



Large DNA molecules can have multiple forks



DNA replication involves many different enzymes to deal with structural and topological complications

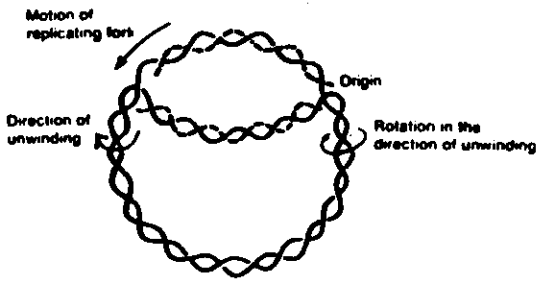


Figure 5-18
Drawing showing that the unwinding motion (curved arrows) of the daughter branches of a replicating circle lacking positions at which free rotation can occur causes overwinding of the unreplicated portion.

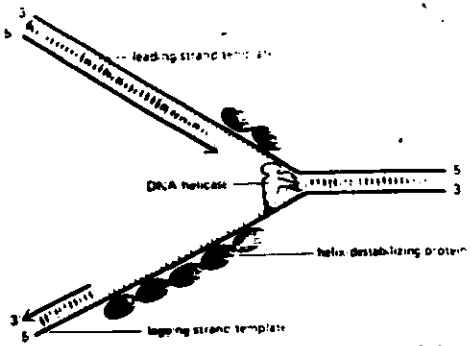


Figure 5–38 The DNA helix ahead of a replication fork is thought to be opened at a rapid rate by the combined action of a DNA helicase enzyme and helix-destabilizing proteins.

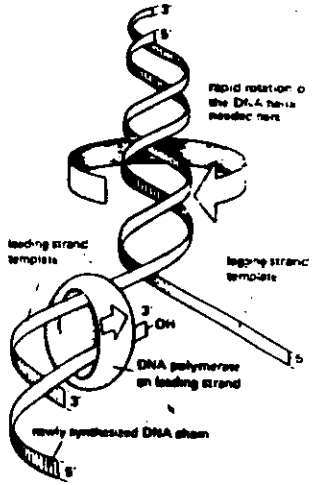
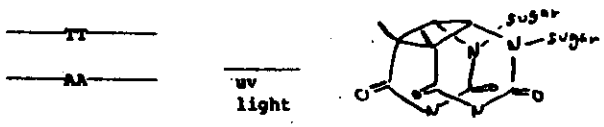


Figure 5–39 Illustration of the "winding problem" that arises during DNA replication. For a replication fork moving at 500 nucleotides per second, the parental DNA helix must rotate at 50 revolutions per second.

Repair of DNA damage

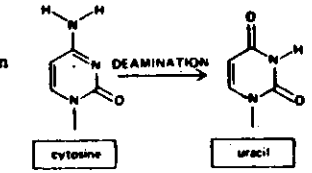
DNA is easily damaged but most common damage can be repaired

common damage



cis, syn thymine photo dimer

deamination



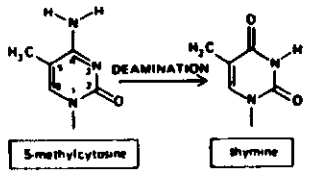
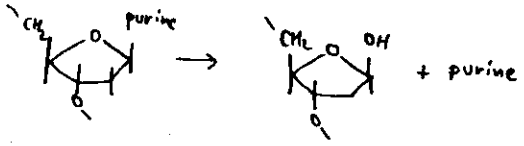


Figure 5–27 The deamination of a methylated cytosine residue in DNA produces thymine instead of uracil, which cannot be recognized and removed by uracil DNA glycosylase.

depurination

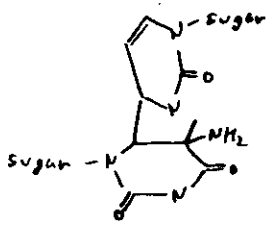


less common damage

single strand breaks

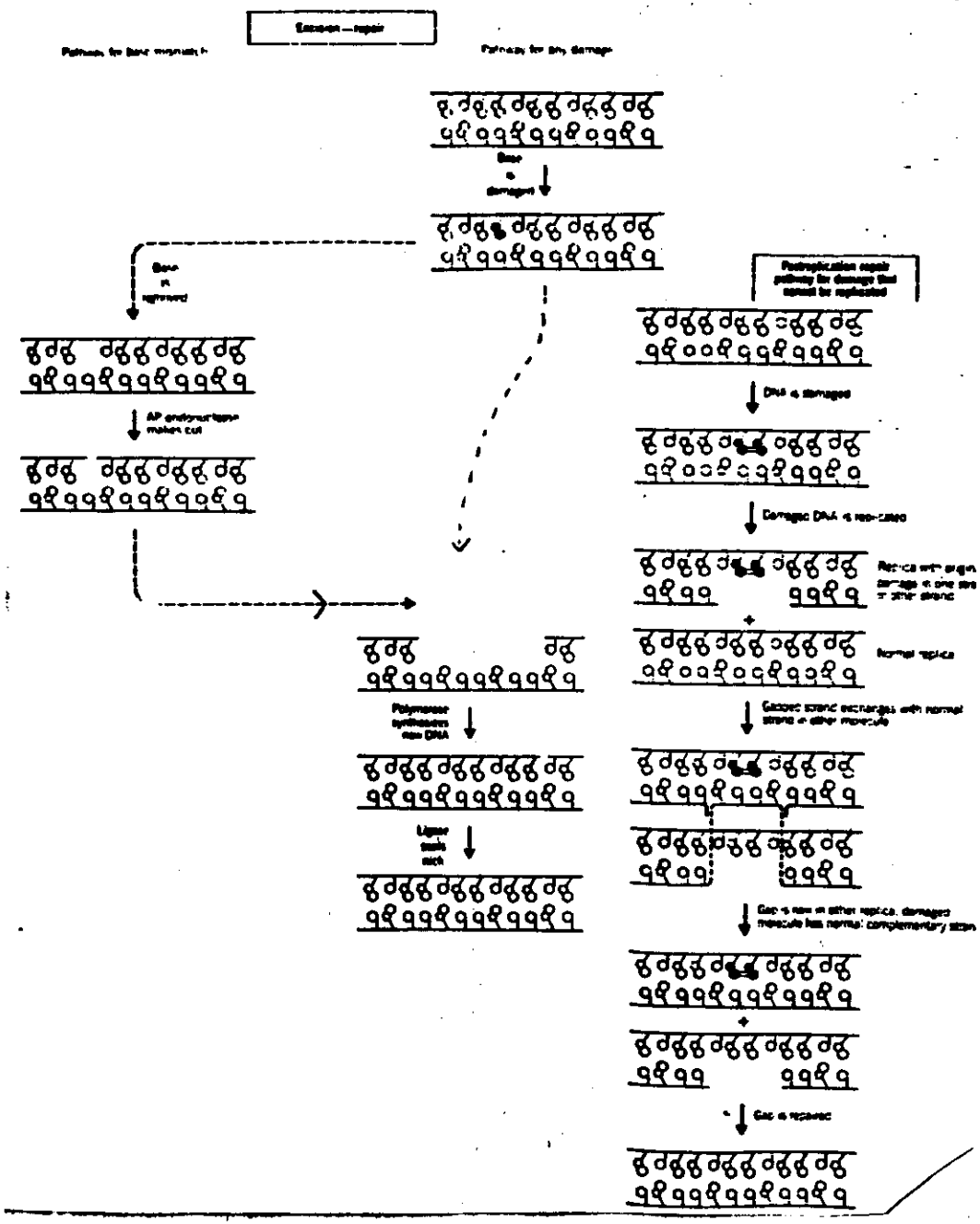
double strand breaks

TC (6-4) product, said to be



guanine O⁶ alkylation

Excision and postreplication repair allow error free correction of some damage.



Other error free repair mechanisms include photoreaction of thymine dimers

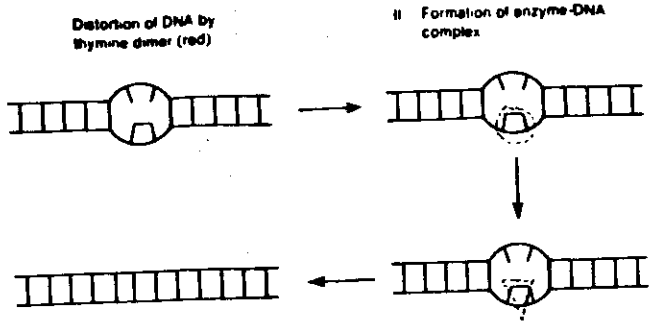
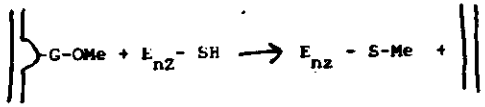


Figure 9-8
Scheme for enzymatic photoreactivation of a thymine dimer

dealkylation by O^6-methylguanine-DNA methyltransferase



enzyme works stoichiometrically rather than catalytically

Mismatch repair can sometimes, but not always be error free

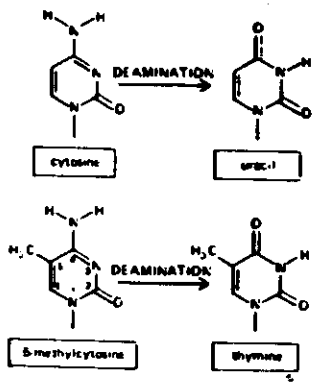
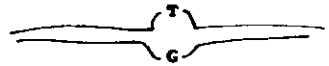
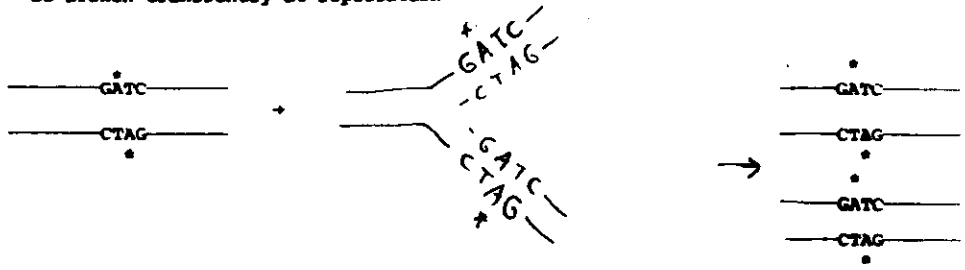


Figure 8-27 The deamination of a methylated cytosine residue in DNA produces thymine instead of uracil, which cannot be recognized and removed by uracil DNA glycosylase.

When 5-methyl C is deaminated to T producing



How can repair enzymes tell which is the "correct" base?
Occasional methylation occurs at DNA bases, particularly A^6_N in bacteria and C in eukaryotes. Pattern is symmetrical on the two strands but symmetry is broken transiently at replication



Thus repair processes can discriminate the newly synthesized strand (subject to potential unedited misincorporation).

When cells are severely damaged a set of genes is induced, SOS functions, that include error prone pathways

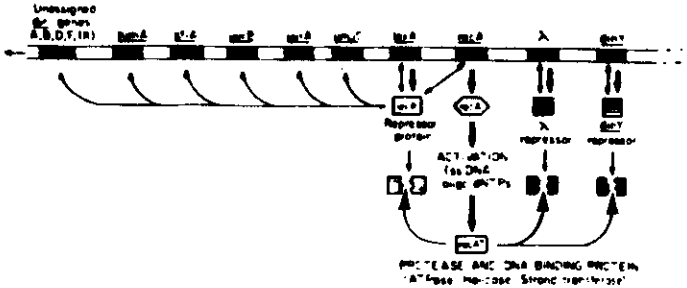


FIGURE S16.1
Regulatory scheme of recA/lexA actions in the SOS response in Salmonella. The sequence of genes does not represent their order in the chromosome. All damage-inducible genes are controlled by lexA for dsDNA and by a unique unidentified repressor for ssDNA. (Courtesy of Professor P. Hanawalt.)

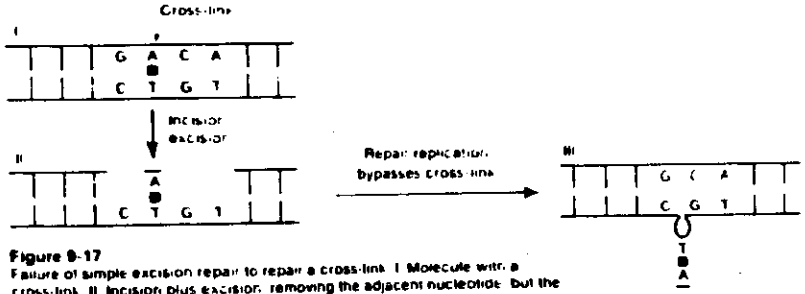
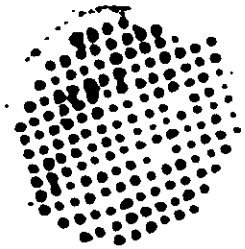
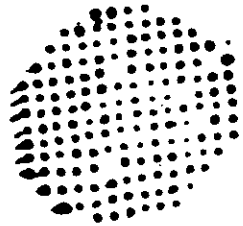


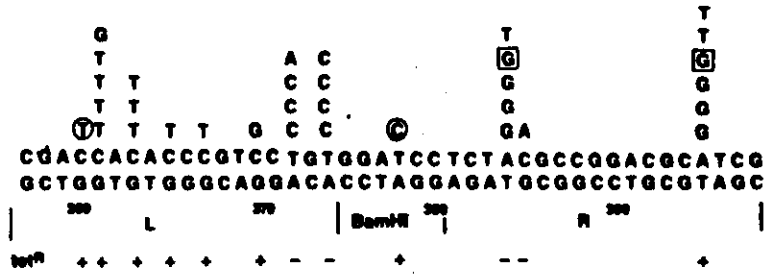
Figure 9-17
Failure of simple excision repair to repair a cross-link. I Molecule with a cross-link. II Incision plus excision, removing the adjacent nucleotide, but the nucleotide remains in the strand in which excision has occurred. III Occasionally, trans-base gap-filling occurs, yielding a single base deletion in the "repaired" strand.

For studies on the repair of psoralen crosslinks, psoralen was placed on plasmid pBR322 near the Bam HI site in the tetracycline resistance gene. E. coli was transformed with damaged plasmids. Ampicillin selection revealed those cells that successfully repaired the crosslink. These were either screened for tetracycline sensitivity or for the presence of altered sequences near the Bam HI site.





The mutations isolated to date show that both transitions and transversions occur and that expected crosslinking sites are indeed hot spots for mutagenesis.



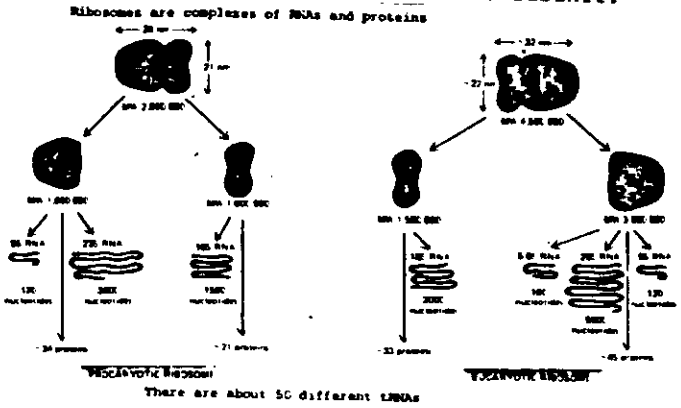
4. Structure of the E. coli 16S ribosomal RNA.

There is an enormous amount of indirect structural information on the 16S rRNA. Crosslinking is one of the few methods that can currently provide direct structural information on such a complex molecule. Both electron microscopic and gel electrophoretic techniques will be described, and how crosslinking information is used to reconstruct three dimensional structural information will be illustrated.

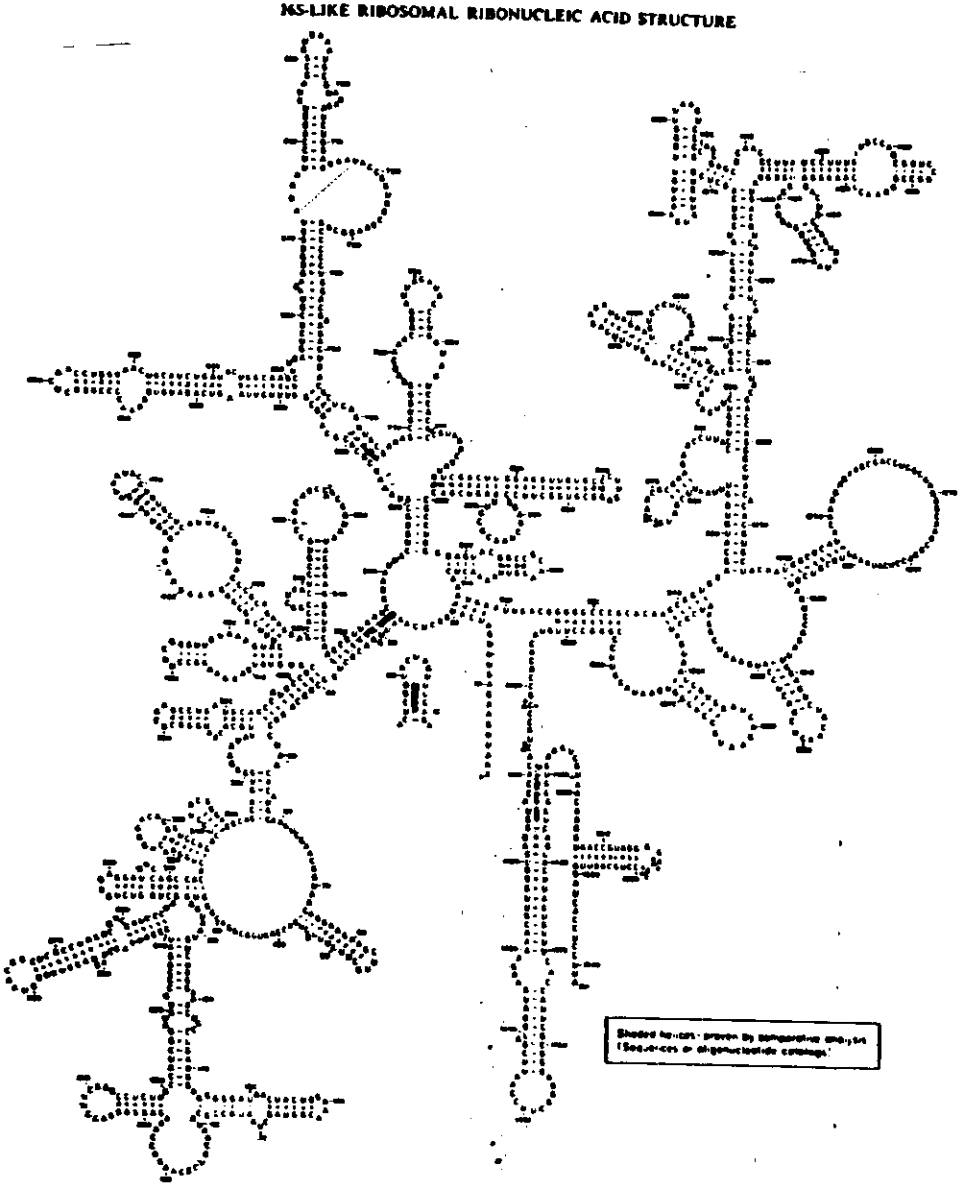
References

1. P. Wollenzien and C. Cantor (1982) J. Mol.Biol. 159:151.
2. P. Wollenzien and C. Cantor (1982) PNAS 79:3940.
3. P. Maly and R. Brimacombe (1983) Nucleic Acids Res. 11: 7263.
4. C.R. Woese et el., (1983) Microbiology Rev's. 47: 621.
5. R. Brimacombe (1984) TIBS 9:273

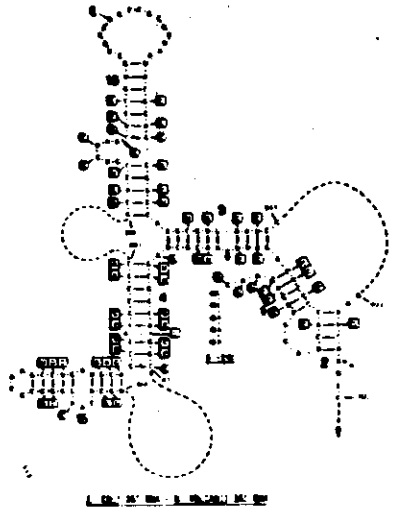
The 16S rRNA has 1542 nucleotides and forms about half the volume and 2/3 of the weight of the 30S ribosomal subunit.



Many different techniques have provided clues about the secondary structure of the 16S rRNA from E. coli and related organisms. These include phylogenetic comparisons of conserved sequences or conserved patterns of base pairing, chemical and enzymatic accessibility, and calculated folding energetics. There is a near consensus for most of the detailed secondary structure elements n the model below.



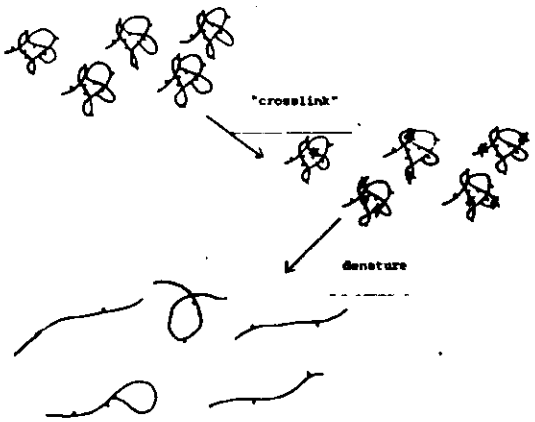
One example of how it is supported by phylogenetic data is indicated below



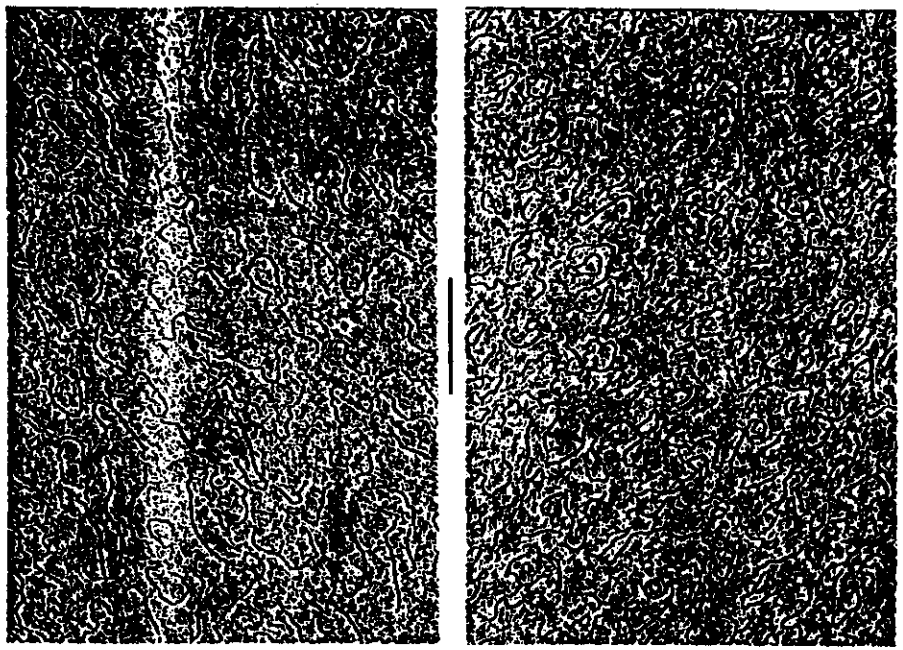
the diagrams depict the E. coli helices, with base changes in the P. polycephalum or H. volcanii sequences being denoted by the bases in boxes. Base changes in square boxes are compensating; those in round boxes are mismatching or in single-stranded regions. Solid triangles denote deletions, and bases with arrows insertions. Dotted lines or 'crossed-out' base pairs denote modified base-pairing in P. polycephalum or H. volcanii. The letters 'a' and 'b' indicate the termini of RNA fragments isolated as a base-paired complex (see text)

I will show how crosslinking can be used to test features of this model, to compare the secondary structure of the free and ribosome-bound 16S rRNA, and to determine elements of the tertiary structure.

Psoralen crosslinking directly traps secondary structure features.

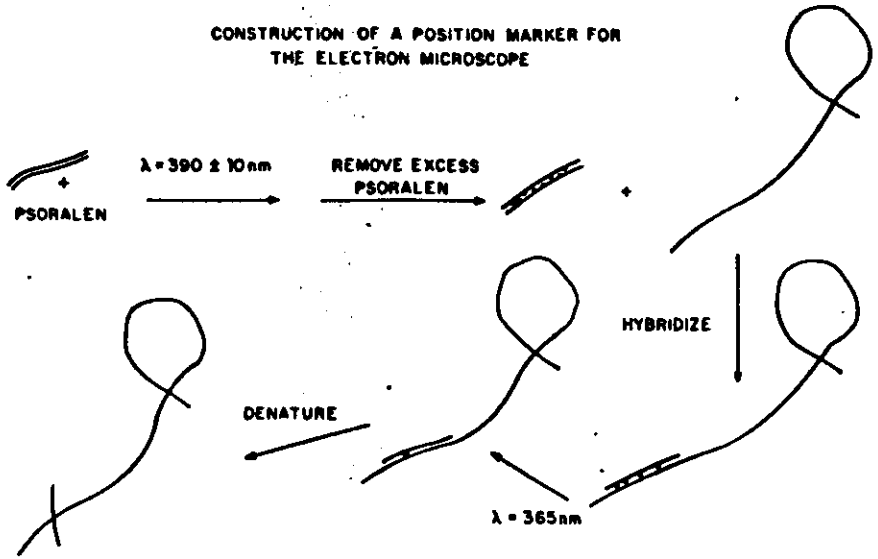


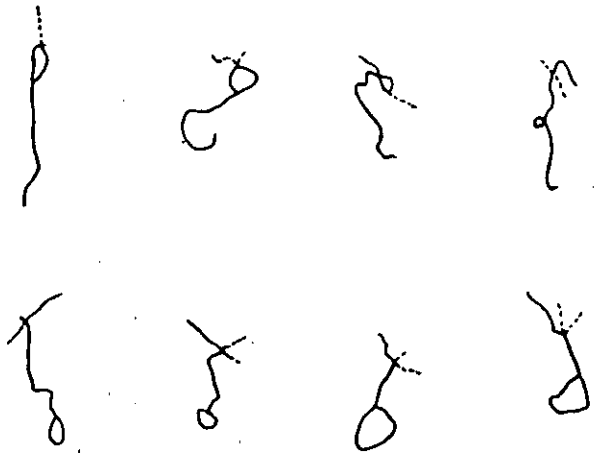
These can be visualized as loops in the electronmicroscope.



A technique has been developed to distinguish between the two ends of an RNA molecule on the e.m. grid. This takes advantage of the availability of specific cloned segments of rDNA,

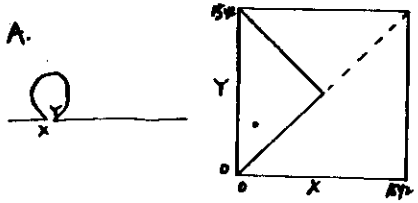
CONSTRUCTION OF A POSITION MARKER FOR THE ELECTRON MICROSCOPE



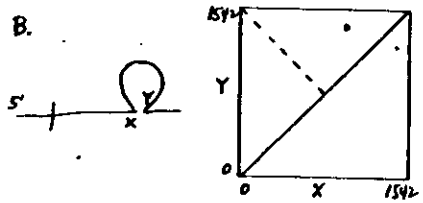


The contours of each molecule seen in the e.m. are measured. Then depending on the particular structure, and whether or not the ends are distinguished, different histogram techniques are used to collect data for statistical evaluation

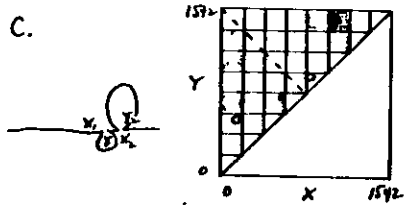
A.



B.



C.



Histograms of the psoralen crosslinks in free and ribosome bound 16S rRNA look superficially very different. In fact however most of this difference arises simply from the reduced accessibility of most of the 5' 2/3 of the 16S rRNA in the ribosome.

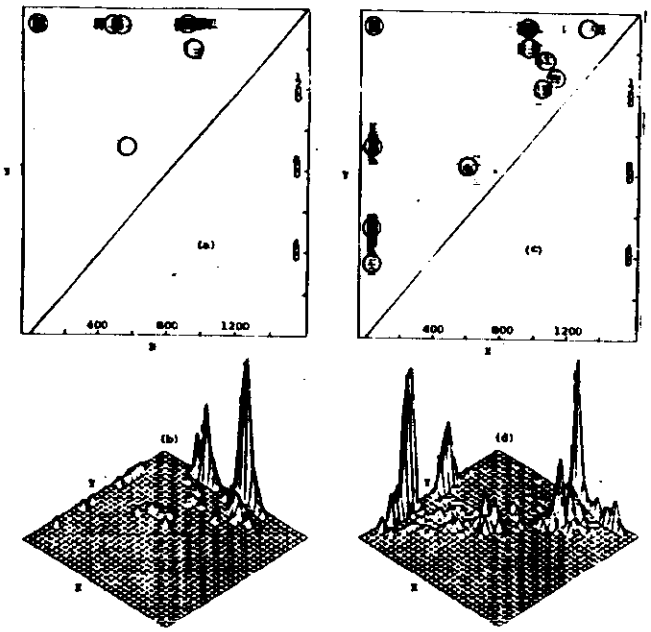


Table 1. Comparison of Secondary Structure Models with Psoralen Crosslinked 16S rRNA

Complementarity	Crosslink (in free RNA)	Crosslink (in RNA subunit)
27-37/ 547-556	EPs 5'x540 (7.9%)	not detected
17-20/ 915-918	EPs 5'x930 (6.6%)	not detected
39-47/ 394-403 52-58/ 354 359	EPs 5'x360 (4.8%)	not detected
946-955/ 1225 1235 984-990/ 1214 1221	EPs 1000x1230 (1.3%)	not detected
564 570/ 880-886 576 580/ 761-765 584-587/ 754 757	EPs 580x840 (0.9%)	EPs 550x870 (0.1%)
926 933/ 1384 1391	EPs 930x1430 (0.9%)	EPs 950x1400 (0.4%)
937-943/ 1340-1345	EPs 950x1340 (0.8%)	not detected
not predicted	EPs 920x3' (7.2%)	EPs 930x3' (21%)
not predicted	EPs 5'x3' (1.8%)[b]	EPs 5'x3' (1.5%)[b]
not predicted	EPs 1300x3' (1.6%)	not detected
not predicted	--[c]	EPs 510x3 (2.2%) EPs 450x3' (1.4%)
not predicted	EPs 1000x1200 (0.5%)	not detected

Footnotes to Table 1.

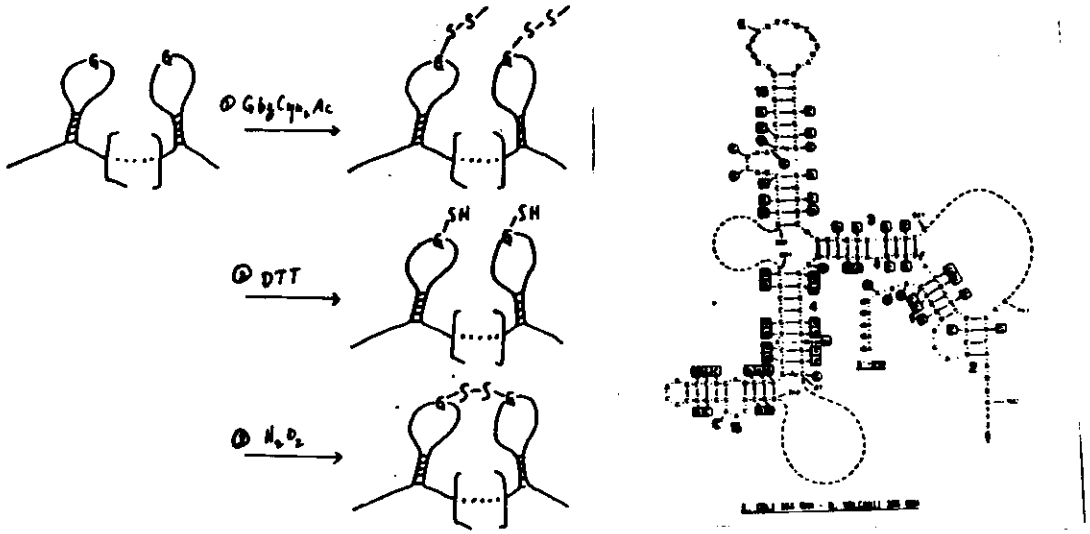
a. Results are summarized from Wollenzien et al. (1983) for psoralen crosslinked free 16S rRNA and from Wollenzien and Cantor (1982a) for 16S rRNA psoralen crosslinked in inactivated 30S subunits. The preface EPs indicates the electron microscopic identification of psoralen crosslinks. Crosslink frequencies reported are percentages of total number of molecules.

b. Molecules containing EPs 5x3' appear as circularized 16s RNA without any small tails; since they appear symmetrical they have not been included in the hybridized-oriented data sets. The position of this crosslink is indicated in Fig. 7 (a) and (c).

c. Molecules containing crosslink EPs 420x3 have been identified in histograms of unoriented molecules at a frequency of 1.6%. The tentative orientation results from the similarity to crosslinks EPs 450x3' and EPs 510x3'made in the inactivated 30S subunit.

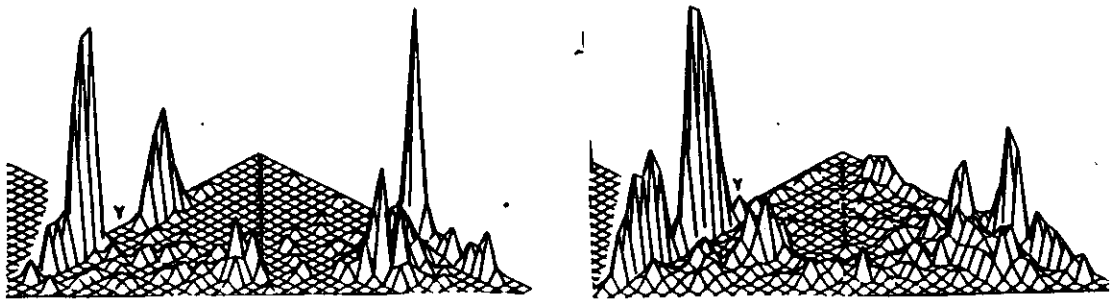
Many of the crosslinks seen match, perfectly, long distance contacts predicted by the models. Others presumably reflect tertiary structure contacts.

A second type of crosslinker is specific for near-by single stranded G residues. This reagent presumably reveals aspects of the tertiary structure.

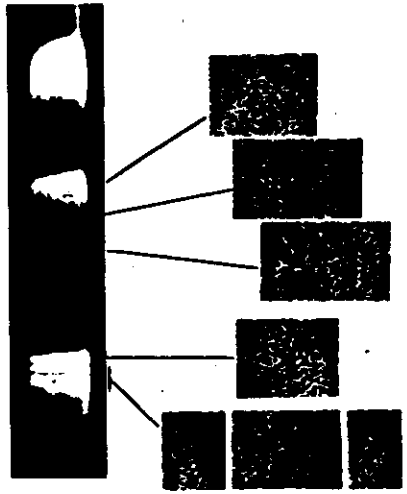


The overall patterns of crosslinking with the G-specific reagent and with psoralen are very similar. This suggests a domain structure for the 16S rRNA. The details of many of the crosslinks seen bear this out very well.

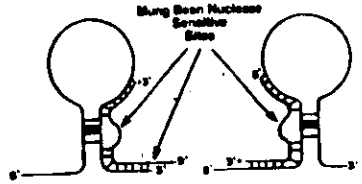
One can estimate that about 200 parameters will be needed to produce a compact coarse model of the arrangement of the 65 helices of the secondary structure model into a plausibile tertiary structure model. Thus far about 150 parameters are available but not all of these are necessarily sufficiently precise or independent.



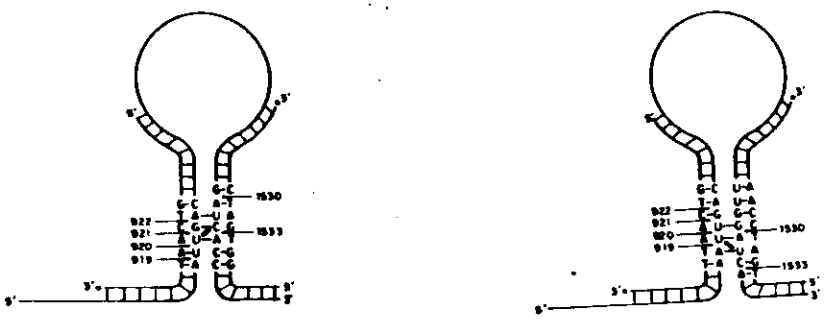
The uncertainty in e.m. localization of crosslinks is a major problem in trying to use this data. Standard methods for crosslink analysis can find locations at the sequence level butmost of these methods have no way, a priori, to pre-select interesting long distance crosslinks. However molecules containing such crosslinks can be fractionated by gel electrophoresis.



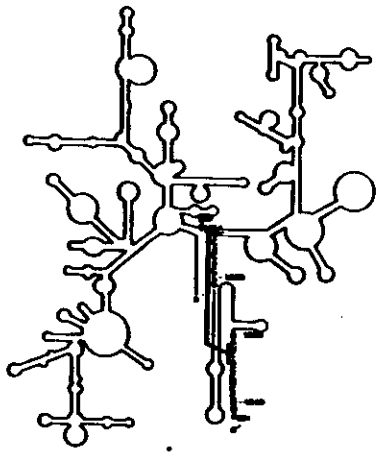
An indirect scheme has been developed for locating the crosslinks in such molecules.



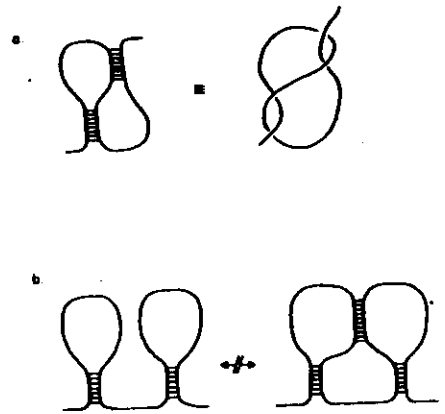
This method has been used to analyze the most frequently occurring psoralen crosslink seen in the 30S ribosome. Two possible base paired structures are consistant with the results seen but it is possible to decide between them.



The results are intriguing since they locate the mRNA recognition sequence of the 16S rRNA in close proximity to a hinge between the two major domains, near the tRNA anticodon binding site.



One special feature of crosslinking is that it traps the topology of the RNA strand interior to the crosslink. Thus one can look for the possibility of knotted RNA conformations.



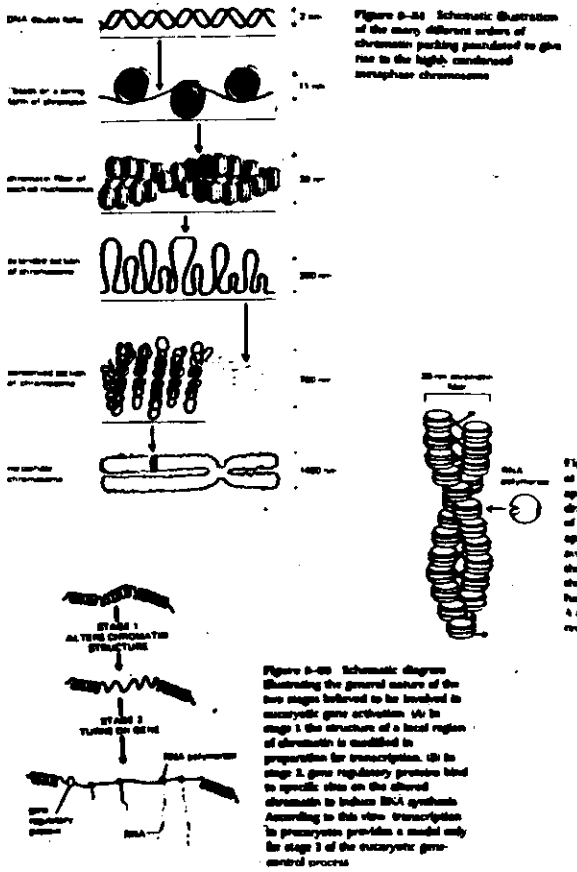
5. Topological aspects of chromatin structure.

The topological constraints on closed circular duplex DNA endow DNA molecules with some rather unique properties. They also provide ways of studying these properties. The implications of topology for the structure and properties of chromatin will be discussed. Included will be considerations of whether small regions of chromatin behave like individual domains of supercoiling and what kinds of DNA conformational changes might occur in small topological domains.

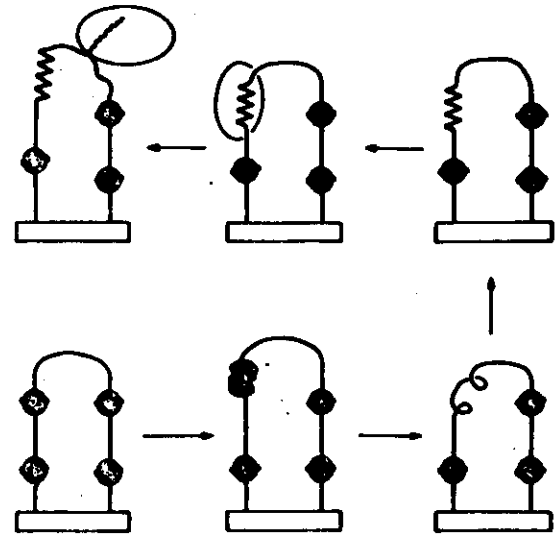
References

1. C. Cantor and P. Schimmel (1980) Biophysical Chemistry 3, pages 1265-1290.
2. W. Keller et al. (1977) Cold Spring Harbor Symp. Quant. Biol. 42: 227.
3. D. Shore and R. Baldwin (1983). J. Mol. Biol. 170: 983.
4. E. Schon et al. (1983) Cell 35:837.
5. L. Peck and J. Wang (1983) Proc. Natl. Acad. Sci. U.S. 80: 6206.
6. H. Weintraub (1983) Cell 32: 1191.
7. C. Prior, et al. (1983) Cell 34: 1033.

The hierarchical assembly of chromatin raises the possibility that a local region might be constrained to act as a topological domain.



In such a domain, gene activation might be accompanied by a number of distinct structural changes including altered nucleosome conformations, altered nucleosome position, loss of nucleosomes, local superhelical density, and altered DNA structures.



There are a number of indirect pieces of evidence that topology is important for eukaryotic gene expression. S1 hypersensitive sites in chromatin appear to match the sites seen in highly supercoiled naked closed circular DNAs. Circular plasmids injected into oocytes are vastly more active in transcription than linear plasmids.

In some actively transcribed genes, like rDNA, unusual nucleosome structures are present. These A particles are extended structures with accessible H3 cysteine 110 residues.

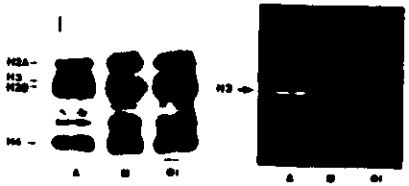


Figure 2 Specific IAF Labeling of H3 Cysteine Residues in A Particles

IAF-labeled nuclei were digested with staphylococcal nuclease (100 U/10⁷ nuclei), and chromatin subunits were purified as described in Experimental Procedures. Aliquots of each subunit (2 μg of A, 6 μg of M and CB) were electrophoresed on SDS 15% polyacrylamide gels and silver-stained as described (left). IAF-labeled H3 was visualized by UV illumination (right). Note that H3 is derivatized only in SS A particles (A) and not in 11S monomers (M) or their oligomers (CB, >tetramers).

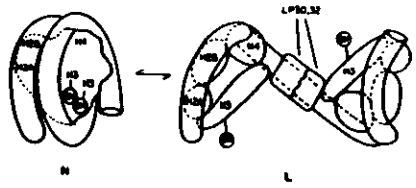
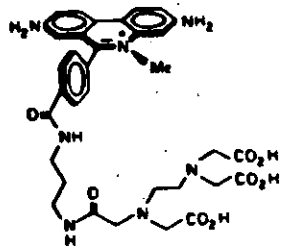


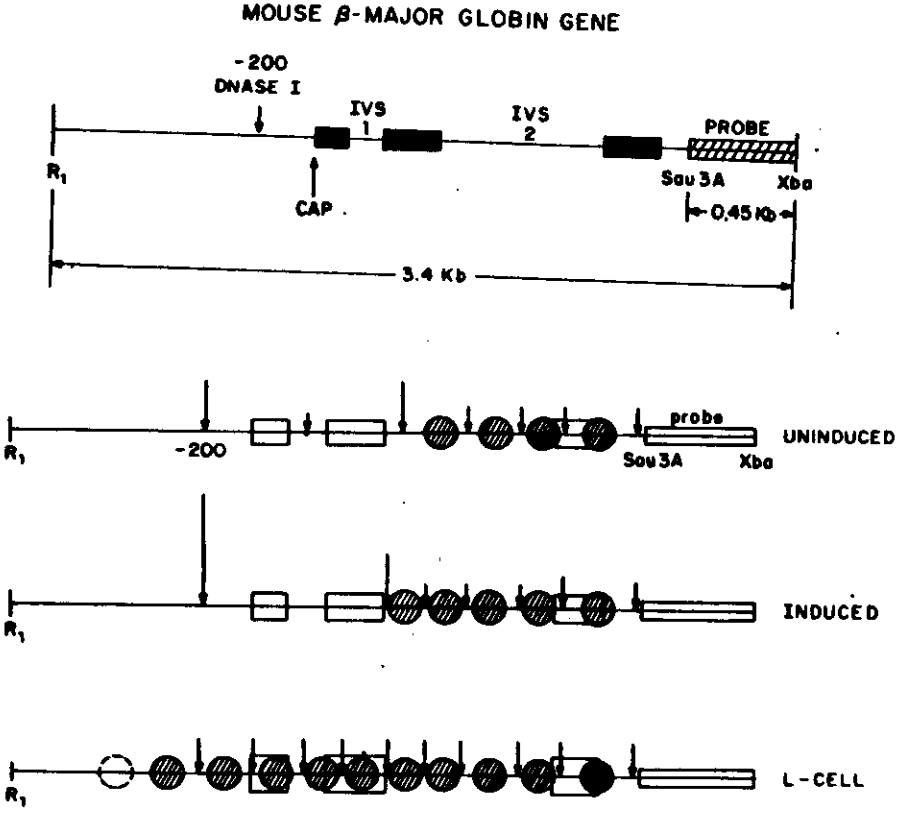
Figure 7. A Model for the Nucleosome-Lexosome Transition

The nucleosome at left (N) is a version of the three-dimensional structure of the histone core originally presented by Klug et al. (1980). Transition to the extended lexosome form (L) is hypothesized on the basis of data presented here.

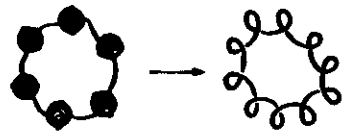
Nucleosome position is also altered in genes when they are induced or even when they are in a cell where they are capable of induction. We have compared nucleosome position on the beta globin gene in mouse L cells and MEL cells. Nucleosome position is measured by footprinting with a cleaving intercalator.



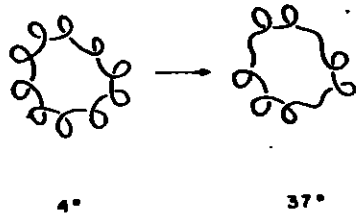
The cleavage pattern is revealed by indirect end labeling.



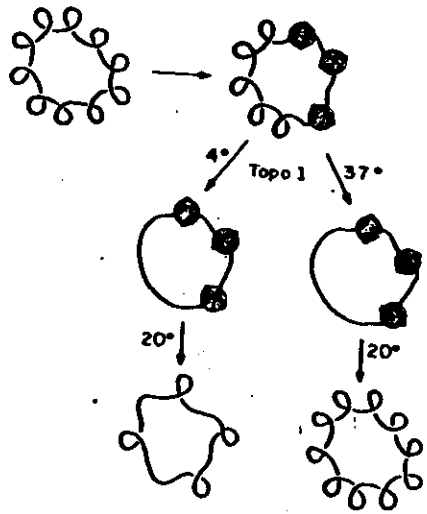
The thermal untwisting of DNA has been used as a non-perturbing probe of the torsional properties of chromatin. Natural chromatin behaves as if the only superhelicity is one negative turn of DNA sequestered in each nucleosome



Thermal untwisting in the absence of topoisomerases is revealed as a temperature dependent writhe.



To measure the thermal untwisting of chromatin we relaxed reconstituted minichromosomes at two different temperatures with topoisomerase 1.



The expected results, if core DNA is constrained and linker DNA is free, can be calculated.

pBS2 3.7 kb

Thermal unwinding 10°/kb - °c

4°c - 37°c unwinding 330°/kb

naked pBS2

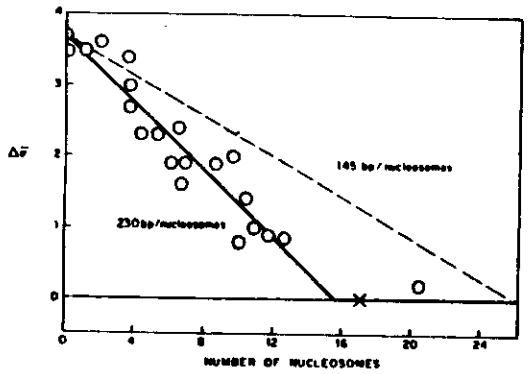
3.7 kb x 330°/360° = 3.4 supercoils lost

pBS2 with 10 nucleosomes = 2.6 kb rigid.

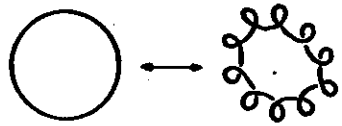
1.1 kb free linker unwinding

1.1 kb x 330°/360° = 1.0 supercoil

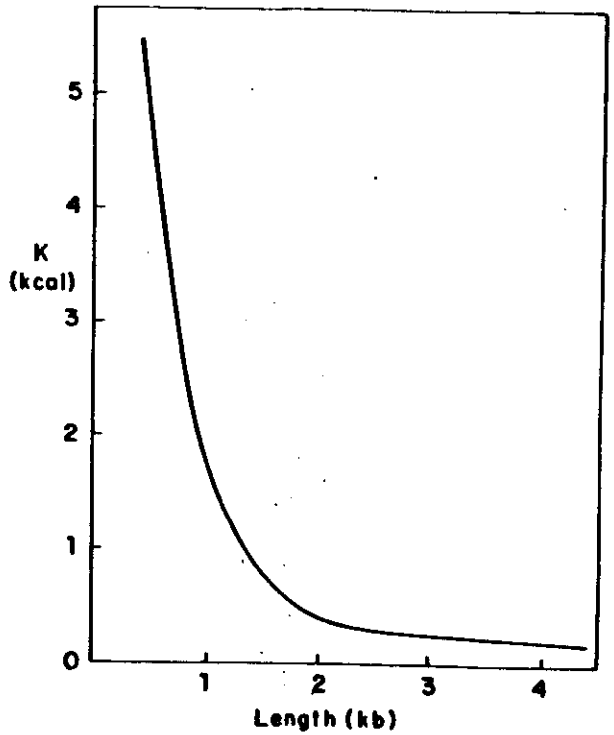
The observed results are very different and suggest that the linker is not free to twist at all.



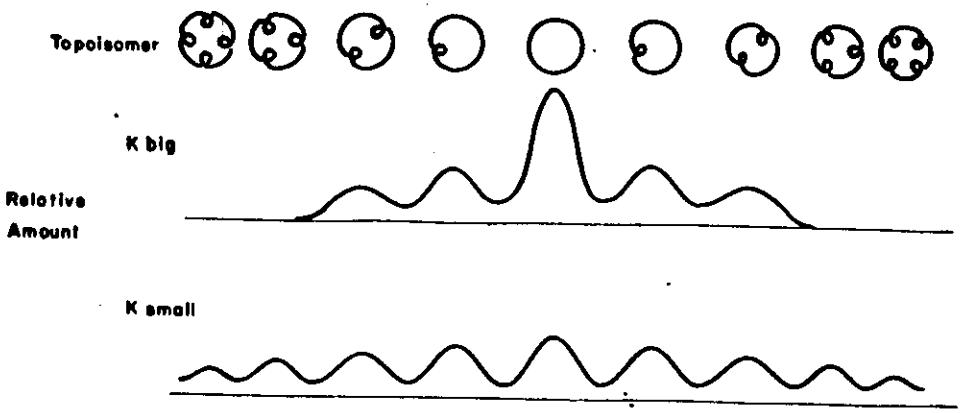
Calculations have been done to explore the kinds of structural changes likely to be induced by torsion in small DNA circles, as models for small chromatin domains. The torsional energy is depends on the square of the number of supercoils, i, and a force constant, K, that depends on the size of the circle.



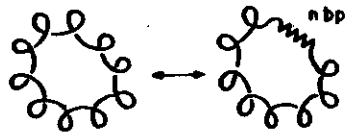
$\Delta G = Ki^2$



When K is big there are relatively few topoisomers.



A very simple thermodynamic treatment has been used to calculate the free energy change for exchanging some superhelical turns for an altered local twisted structure.



The free energy of the supercoiled form is

$$G_{init} = K(i+d)^2$$

The free energy of the form with n base pairs in an altered structure is

$$G_{fin} = \Delta G_{nuc} + n\Delta G_{gro} + K(i+d-nf)^2$$

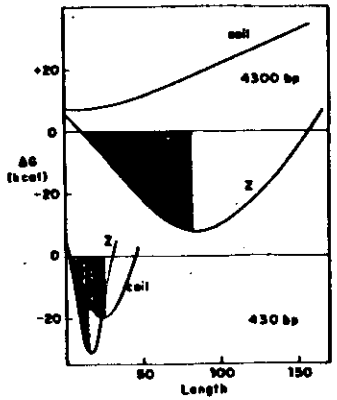
where f is the fraction of a turn unwound per base pair in the altered structure and the other parameters represent the free energy of nucleation and growth of the altered structure.

Thus the free energy change upon assuming the altered structure is

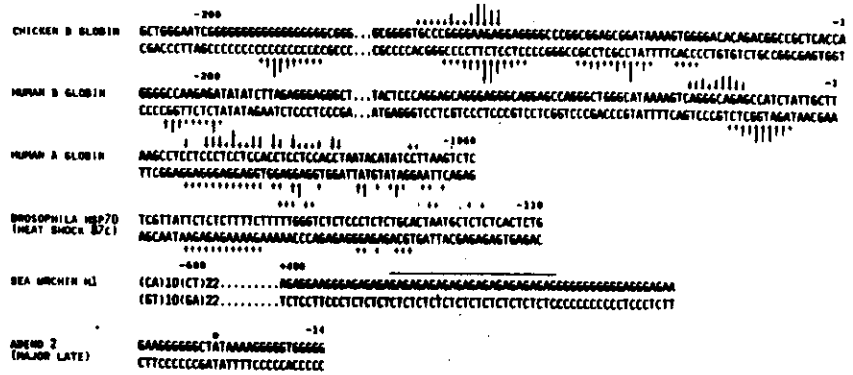
$$\Delta G_c = G_{fin} - G_{init} = \Delta G_{nuc} + n\Delta G_{gro} + K(i+d-nf)^2 - K(i+d)^2$$

$$= n^2Kf^2 + n(\Delta G_{gro} - 2K(i+d)f) + \Delta G_{nuc}$$

This is a simple quadratic in the length, n, of the altered structure. Calculated results for DNAs of 4300 bp and 430 bp with 20 and 3 supercoils are shown below.



These calculations suggest that in a small topological domain, supercoiling will always lead to DNA melting unless there is some thermodynamically more stable alternate structure. The fine structure of several S1 hypersentive sites is shown below.



It is hard to reconcile this data with DNA melting; more easy with some kind of left-handed helices.