Isolation and characterization of origin-enriched sequences from early-replicating human cells

by

Michael A. McAlear

A thesis submitted to the Faculty of Graduate Studies and Research of McGill University in partial fulfilment of the requirement for the degree of Doctor of Philosophy

Department of Medicine
Division of Experimental Medicine
McGill University
Montreal, Canada

ABSTRACT

The objective of this thesis is to research at the molecular level the mechanisms involved in the initiation of mammalian DNA replication. Human WI-38 cells were synchronized to the G1/S border by serum starvation and aphidicolin block. Cells were relased from arrest and followed as they progressed into S phase by microfluorometric analysis. Early replicating DNA was extruded from replication bubbles, purified from the high molecular weight parental DNA and cloned into the NruI site of pBR322. The recombinant plasmids were surveyed for properties previously associated with origins of replication. In a random sample of 10 human origin enriched sequences (hors) that were analyzed, 5 were capable of autonomous replication in a transient BrdU substitution assay. Two clones contained DNA fragments that migrate anomalously on acrylamide gels suggestive of bent DNA. One clone contained a weak DNA unwinding element as judged by sensitivity to the single strand specific enzyme mung bean nuclease. Primary sequence analyses of five of the hors clones (hors 98, 106, 112, 129 and 133) revealed that they were enriched for the AP2-A, NF-1 related and iron response consensus sequences. The replicating clones also contained potential cruciform structures within 50 bp of an A/T rich region. A DNA binding activity was identified in HeLa nuclear extracts that binds to a sub-fragment of one of the replicating clones (hors106) by bandshift assays and it was partially purified by DEAE and PC-11 column chromatography.

RESUME

La synchronization de cellules WI-38 à la phase G1/S du cycle cellulaire a été effectuée par carence de sérum suivie d'un traitement avec aphidicoline. Le cycle cellulaire a été remis en marche par remplacement du milieu de culture, et la progression des cellules en phase S a été suivie par microfluorométrie. Les fragments d'ADN répliqués au début de la phase S ont été extraits des bulles de réplication, séparés de l'ADN de haut poids moléculaire provenant des cellules-mères, et clonés dans le site Nru I de pBR322. Nous avons vérifié les propriétés réplicatives des plasmides recombinants. Sur 10 clones sélectionnés au hasard, 5 se sont révélés capables de réplication autonome lors d'expérience de substitution par BrdU (un analogue lourd de la base thymidine) au cours de 24 heures. Deux de ces clones contiennent des fragments d'ADN migrant de façon anormale sur gel de polyacrylamide, ce qui suggère une structure courbée ("bent DNA"). L'un des clones contient un fragment d'ADN sensible à la nucléase de la fève mung qui est spécifique à l'ADN simple brin, suggèrant que ce fragment peut exister dans une configuration ouverte de manière transitoire ("unwinding element"). L'analyse de la séquence de ces cinq clones hors (human origin enriched sequence, hors 98, 106, 112, 129, et 133) n'a révélé aucune homologie significative parmi ces clones. Toutefois, ces clones contiennent des séquences AP2-A, la séquence consensus pour la réponse au fer ("iron response consensus sequence"), ainsi que des séquences apparentées à NF-1. Toutes ces séquences contiennent également, à l'intérieur d'une région de 50pb riche en A/T, des régions pouvant potentiellement former des

structures cruciformes. Nous avons démontré par déplacement de bande sur gel de polyacrylamide, la présence d'un site de liaison dans un fragment de hors 106 (un des clones capables de réplication autonome) pour une ou des protéines présente(s) dans un extrait nucléaire provenant de cellules HeLa. La purification partielle de ce(s) factor(s) a été réalisée par chromatographie sur DEAE et PC-11.

ACKNOWLEDGEMENTS

There are so many people to whom I owe a dept of gratitude, that it is not possible to name them all. I would like to thank first of all Dr. Maria Zannis-Hadjopoulos for supervising and directing my research over the course of these studies. I have appreciated and benefitted from her suggestions, encouragement and her continual openness and accessibility for discussion. I would also like to thank Dr. Jerry Price and the many members of the McGill Cancer Centre with whom I have had the pleasure of working over the last five years. Cunle Wu, Suzanne Landry, Lori Frappier, Barry Sleno, Christopher Pearson, David Mah, and Marcia Ruiz are all valued friends who have made my studies both richer and more enjoyable.

I am appreciative towards The Cancer Research Society and The Alexander McPhee fellowship for supporting me over the course this research.

I am also deeply grateful to my family and friends for their support over the course of my studies. I especially appreciate the importance of education emphasized by my parents. Finally, I would like to thank Kathleen. She has shared my life through the entirety of this research and has given me comfort, understanding, encouragement and love, without which I don't know what I would have done.

PREFACE

The research presented in this thesis was carried out under the supervision of Dr. Maria Zannis-Hadjopoulos. The cell synchrony experiments and microfluorometric analysis were performed with the assistance of Dr G. B. Price and Richard McKenzie. The somatic cell hybrid blot was provided by Dr. J.L. Hamerton and was probed with the assistance of Dr. Awatef Shihab-El-Deen. The protein purification and in-vitro replication experiments were carried out with the assistance of Christopher Pearson and Marcia Ruiz.

LIST OF ABBREVIATIONS

A^r	Ampicillin resistant
ARS	Autonomously replicating sequence
ATP	Adenosine ribonucleoside triphosphate
bp	Base pairs (of DNA)
BrdU	Bromodeoxyuridine
D	Purine or thymine
dATP	Deoxyadenosine ribonucleoside triphosphate
dCTP	Deoxycytosine ribonucleoside triphosphate
dGTP	Deoxyguanosine ribonucleoside triphosphate
DHFR	Dihydrofolate reductase
dTTP	Deoxythimidine ribonucleoside triphosphate
DMEM	Dulbecco's modified Eagle medium
FACS	Fluorescence activated cell sorter
FCS	Fetal calf serum
GTP	Guanosine ribonucleoside triphosphate
H	Adenine or pyrimidine
Hg-dUTP	5-mercuri-deoxyuridine ribonucleoside triphosphate
K	Guanine or thymine
kb	Thousand base pairs
N	Any base

nt.	Nucleotide(s)
ORS	Origin enriched sequence
PBS	Phosphate buffered saline
SV40	Simian virus 40
TCA	Trichloroacetic acid
T^s	Tetracyclin sensitive

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER FOUR PROPERTIES OF ORIGIN ENRICHED SEQUENCES (HORS)

CHAPTER ONE

INTRODUCTION

I. Basics of DNA Replication

(i) Perspectives

"We wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest... It has not escaped our attention that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material." (Watson and Crick, 1953).

The complexity of the eukaryotic genome and the fidelity with which it is duplicated have long been objects of intense research. Indeed, the history of molecular and cellular biology has followed, to a large degree, the progress in our comprehension concerning the details of DNA replication, structure and function So fundamental is the process of the DNA synthesis, that the proliferating cell cycle, one of the conceptual foundations of cell biology, is defined in relation to the landmarks of DNA replication (S phase) and chromosome segregation (mitosis) (Howard and Pelc, 1951). The act of DNA replication is a complex and universal feature of the biology of all organisms. Research revealing the variety of strategies used by different organisms to replicate their genomes has not only broadened our appreciation of the diversity of life, it has led to the development of a number of tools enabling us to better study the details of molecular biology. Improved vectors, efficient DNA sequencing, and the polymerase chain reaction are but three of the

technological advances that have emerged from basic research on DNA replication.

The study of DNA replication touches on a multiplicity of topics in biology ranging from the origins and evolution of life to the issue of ageing and cell senescence. One such topic of considerable medical relevance is the study of the molecular mechanisms that regulate cell division. This topic is one of the broadest areas of research in molecular and cellular biology today and is closely linked to our understanding of the causes and possible cures of the diseases of cell proliferation. An increasing number of cellular oncogenes thought to function in the cell division cycle have been implicated in the development of a range of human cancers (Der, 1989). The relationship between the process of DNA replication and cancer has been exploited by the many chemotherapeutic agents currently in use that have their effect at the level of disruption of DNA synthesis (Borsa and Whitmore, 1969; Glazer and Knode, 1984). An increased understanding of the molecular mechanisms that regulate and execute DNA replication may permit us to formulate new more specific agents capable of controlling errant cell division.

Whereas the majority of research on DNA replication carried out to date has been on viral and prokaryotic systems, this thesis deals with DNA replication in human cells. Although the basic process of initiation and elongation is now well defined in a few of these model systems, the corresponding processes in higher eukaryotes remain to be

solved. In addition to the complex enzymatic mechanisms that serve to synthesize DNA, eukaryotic cells have a sophisticated regulatory mechanism that oversees the progression of S phase. Eukaryotic DNA replication is likely to emerge as a process that is very much more complicated than that revealed by the simple prokaryotic and viral models. The focus of this project to research DNA replication in human cells at the level of initiation. To this purpose, I have investigated both the cis and trans acting elements involved in the initiation of DNA replication in human cells.

(ii) The Replicon

One of the ways in which the complex process of DNA replication has been organized is by dividing DNA molecules into small functional segments known as replication units or replicons. A replicon is a section of DNA that is replicated from a single origin of DNA replication (Jacob et al., 1963). Although this term was first used to describe the replication of prokaryotic genomes, it has become an important organizing concept regarding our understanding of eukaryotic DNA replication. One of the first indications that DNA molecules were organized as multiple discreet replication units came from experiments that involved the labelling of replicating cells with ^{3}H-thymidine (Cairns, 1966). Such DNA fibre autoradiography studies indicate that eukaryotic DNA replication initiates at origins of replication and continues bidirectionally outwards along the DNA molecule (Huberman and Riggs, 1968). These early experiments allowed researchers to measure both the frequency of initiation and the range in sizes of eukaryotic replicons. It was estimated that a typical mammalian cell possesses in the order of 10^{5} origins of replication serving replicons ranging in size from 50-250 kb (Hand and Tamm, 1974; Hand, 1978). It was also noted that clusters of replicon units were replicated at the same time in S phase (Huberman and Riggs, 1968). In fact, without this coordinate replication of adjacent replicons it would not be possible to measure the distance between functional origins of replication. If the activation of replicons was a random event through S phase, then the probability of two adjacent replicons being

labelled during the same short pulse would be low. Recently, there have been more sophisticated labelling experiments aimed at characterizing the nature of this organized activation of replicon clusters. The visualization of intact nuclei, rather than DNA fibres, has enabled researchers to investigate the process of replication at a higher structural level. The intranuclear structures corresponding to replicon clusters can be visualized by immunofluorescent staining of nuclei labelled with 5-bromodeoxyuridine (Nakamura et al., 1986) Roughly 100 ring-like structures are observed per S phase nuclei and the intensity and kinetics of staining suggest that each consists of at least ten replicons. These structural components within nuclei that are thought to play a role in replication form part of the nuclear matrix.

One of the central issues in the field of DNA replication involves the nature of the sequences at which DNA replication is initiated (Callan, 1973). Are there specific DNA sequences that interact with the replication machinery and serve as origins of DNA replication, or is the initiation of replication a random event? If specific origin sequences exist, what are the features of these sequences, how do they interact with the enzymes of replication, and do they consistently regulate replication from cell cycle to cycle and in different cell types? The large size and variety of cellular replicons have made the identification of discreet cellular DNA replication origins difficult. Although studies on prokaryotes (Meijer, et al., 1979) and eukaryotic viruses (Deb et al., 1986)

suggest that origins of replication are in fact discreet sequences, until recently there were few candidates for eukaryotic chromosomal origins.

In an attempt to address this question, several researchers have focused on the regulation of DNA replication during the process of development. It has long been known that the cells of embryos have a shortened cell cycle time (Painter and Schaefer, 1969). Whereas a normal differentiated cell might have an S phase lasting 8 hours, in embryos the process of DNA replication may take as little time as 4 minutes (Callan et al., 1972). DNA fibre autoradiography experiments revealed that the shortened length of S phase in embryonal cells corresponded to a smaller size of the average replicon (Blumenthal et al., 1974). This apparent increase in the number of origin sequences is thought to be necessary to replicate the genome within the short S phase of these rapidly dividing cells. If origin sequences are specific, then there are many such potential sequences and although they may all be functional in embryonal cells, the majority of them are not activated in adult tissues. This data is supported by the observation that any piece of DNA, regardless of species, that is microinjected into Xenopus eggs will be replicated (Harland and Laskey, 1980). It is likely that since the cells of embryos have a particular requirement for rapid DNA synthesis, they may not regulate DNA replication in the same way as differentiated cells.

Another method used to assay the control of DNA replication has

been to monitor the replication timing of specific sequences among different cell types. Several investigators have found a link between the transcriptional activity of regions of DNA and the time within S phase that these sequences are replicated. In chicken embryonic erythrocytes, the transcriptionally active histone H5 genes are replicated from an origin located 5' to the gene, whereas inactive H5 genes are replicated in the opposite direction from a downstream origin (Trempe et al., 1988). The 140 kb human B-globin gene domain also shows a correlation between transcriptional activity and replication status. In K562 erythroleukemia cells, in which the embryonic and fetal globin genes are expressed, these sequences replicate during early S phase. In HeLa cells, where these genes are transcriptionally silent, these sequences replicate during late S phase (Dhar et al., 1988). These studies indicate that the activation of origins of replication is affected by cell type. The mechanisms by which this control is exerted is unclear although it may relate to such factors as the physical state of chromatin (Dhar et al., 1988).

(iii) Supercoiled Loops and the Nuclear Matrix

The large size of mammalian genomes imposes an immense topological problem to the cell. Some 50 cm of DNA must be packaged into a nucleus of 10 um diameter in such a way as to accommodate the requirements for complex cellular processes such as transcription, replication and chromosome segregation. Electron micrographic studies coupled with the identification of some of the proteins involved in this packaging (i.e. histones) has led some investigators to suggest a hierarchy of chromosomal organization leading from the naked double helix to the condensed metaphase chromosome (reviewed in Nelson, et al., 1986). One of the key orders of organization in these models is the supercoiled loop. Supercoiled loops are topologically constrained segments of DNA estimated to range in size from 10 to 200 kb (Earnshaw and Laemmli, 1985; Hancock and Boulikas, 1982). Experiments demonstrating the existence of supercoiled domains of DNA have been performed on a variety of eukaryotic cells using the intercalating dye ethidium bromide (Paulson and Laemmli, 1977; Vogelstein et al., 1980). These domains of DNA correspond roughly to the average size of replicons and are thought to be important in the regulation of transcription as well as replication.

The independent and supercoiled nature of these DNA loops is thought to depend on their attachment to the nuclear matrix. The nuclear matrix is an insoluble, skeletal framework of the nucleus that is revealed following digestion with DNAse and extraction with non-

ionic detergents and hypertonic salt buffers (Berezney and Coffey, 1977). The residual structure that remains resembles the intact nucleus in size and shape although it contains only 10% of the total nuclear protein mass (Berezney and Coffey, 1976). Considerable effort has been placed into identifying both the protein constituents that form the nuclear matrix as well as the DNA sequences that serve to anchor the loops. Some of the proteins contained in the nuclear matrix include the scaffold proteins (Lebkowski and Laemmli, 1982), the lamin proteins (Fischer et al., 1982: Gerace et al., 1978; Gerace and Blobel, 1980), and c-myc proteins (Eisenman et al., 1985). There is evidence that the enzymes involved in DNA replication are also localized on the nuclear matrix (Berezney and Buchholtz, 1981; Jackson and Cook, 1986; Cook, 1991). Both DNA polymerase alpha and DNA primase were localized to the internal region of the nuclear matrix (Wood and Collins, 1986; Smith et al., 1984). DNA topoisomerase II is associated with the nuclear matrix and may play a dual catalytic/structural role (McConaughy et al., 1981; Berrios et al., 1985; Earnshaw et al., 1985). A model of replication has been proposed in which DNA loops reel through these large replisome complexes as they are replicated (McCready et al., 1980, Dingman, 1974).

Autoradiography experiments in BHK cells suggest that DNA replication origins are attached to the nuclear matrix throughout S phase (Dijkwel et al., 1986). Brief pulses of (^{3}H)-thymidine administered at the onset of S phase in synchronized cells labelled

DNA that remained associated with the nuclear matrix. In contrast, if the pulse was used to label DNA in non-synchronized cells, then one could see the label move out from the matrix into the surrounding halo of DNA. Similar results were obtained with (^{3}H)deoxyadenosine labelling studies in Xenopus laevis cells synchronized with 5-fluoro-2'-deoxyuridine (Carri et al., 1986). These data suggest that non-origin DNA sequences associate with the nuclear matrix only when they are part of a replication fork, whereas origins of DNA replication remain permanently associated with the nuclear matrix.

(iv) The Enzymology of DNA Replication

There are many complex and interacting biochemical pathways and cascades involved in the progression of cells into S phase. The initiation of DNA replication constitutes a major transition in the eukaryotic cell cycle, and as such it is subject to a variety of regulatory controls (Murray, 1987). Once DNA replication is initiated, the cell is committed to complete a round of cell division (Hartwell et al., 1974). The proliferative response to extracellular stimuli is emerging as a complex pathway linking receptors to messengers and effectors. A number of growth factors have been identified that can trigger an orderly sequence of events that culminate in DNA synthesis and ultimately cell division (Holley, 1975). These mitogens stimulate the synthesis of some of the proteins involved in the proliferative pathway including cyclin (PCNA), dividin, and progressin (Bravo et al., 1981; Celis and Madsen, 1986: and Celis et al., 1987). In addition to these complex pathways of signal transduction, there are the basic metabolic and biosynthetic pathways that are required for the production of the molecular precursors (nucleotides, histones etc.) necessary for DNA synthesis. The duplication of the eukaryotic genome constitutes a major biosynthetic process and it is not surprising that cell division is also controlled in response to nutrient limitations. Much of what we know about the molecular mechanisms that regulate and execute DNA replication has come out of the study of prokaryotic (reviewed by Zyskind and Smith, 1986), viral (reviewed by Depamphilis and Bradley,

1986), and lower eukaryote systems (reviewed by Campbell, 1986). These simple models have been useful in characterizing both the cis and trans acting elements required for DNA replication in a variety of organisms.

One productive approach towards deciphering the enzymology of DNA replication has been to characterize the proteins required to catalyze SV40 virus DNA replication in-vitro (Li and Kelly, 1984). SV40 is a member of the papovavirus family, has a double stranded DNA genome of 5.2 kb and contains a single well characterized origin of replication (Li et al., 1986). This system has been dissected by both genetic and biochemical approaches, revealing considerable detail of the mechanism of initiation and elongation of DNA replication (reviewed in Stillman, 1989). The initiator protein T antigen is the only virus encoded protein required for DNA replication, and as such this has proven to be a powerful system in which to identify cellular replication proteins.

Human cell extracts have been fractionated and assayed for replicative activity with plasmids containing the SV40 origin of replication in the presence of exogenously added T antigen. Two essential cellular factors that have been identified are RF-A, a single stranded DNA binding protein (Wobbe et al., 1987; Wold and Kelly, 1988; Fairman and Stillman; 1988), and RF-C, a multi-subunit protein that seems to be required for efficient elongation of DNA replication (Tsurimoto and Stillman, 1939). In addition, there is a

requirement for the DNA polymerase alpha/primase complex, topoisomerase II and PCNA (proliferating cell nuclear antigen), also known as cyclin (Bravo et al., 1987; Prelich et al., 1987).

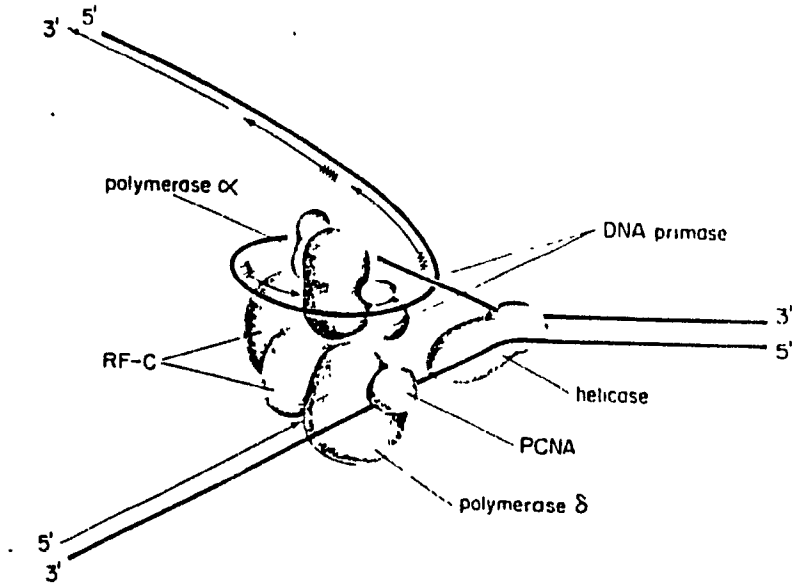


Figure 1. Diagram of a replication fork from Stillman, 1989.

PCNA is the auxiliary protein of DNA polymerase delta and is required for increased enzyme processivity (Tan et al., 1987). DNA polymerase delta is one of the five DNA polymerases that have been identified in human cells. The different polymerases can be differentiated by their sizes, sub-unit structures and susceptibility to certain drugs (ie. aphidicolin). Each of these enzymes catalyses the polymerization of nucleotides in the 5'-3' direction and each has a requirement for a nucleotide primer of either DNA or RNA. There are however a number of other activities associated with the various polymerases, and these reveal clues as to the roles of the distinct DNA polymerases in replication, repair and recombination. A primase activity co-purifies with DNA polymerase alpha (Conaway and Lehman,

1982), and it is thought to be the polymerase associated with lagging strand synthesis (reviewed in Burgers, 1989). DNA polymerase delta, when associated with PCNA has a high processivity and is thought to be the leading strand polymerase (Tan et al., 1986).

Adenovirus is another favourite system for the study of DNA replication. Although the replication of the 36,000 bp adenovirus genome does not reflect the process of chromosomal replication, it has provided us with insight into a novel replication strategy. The origins of replication are located at the termini of the genome and initiation involves a terminal protein (pTP) that is covalently linked to one strand of DNA (Robinson et al., 1973; Rekosh et al., 1977). In addition to the viral replication proteins, there is a requirement for several cellular proteins, nuclear factors I, II, and III (NF-I,II,III). NF-1 is a DNA binding protein that stimulates the formation of the pTP-dCMP complex (Nagata et al., 1982). NF-II is required for the replication of full length molecules only and may remove positive superhelical tension ahead of the replication fork (Guggenheimer et al., 1983). NF-III is a DNA binding protein that interacts with cellular enhancers and promoters as well as the terminal repeat sequences of the adenovirus genome (Pruijin, et al., 1988; O'Neil and Kelly, 1988). While these studies have been invaluable in characterizing the replicative machinery, they are of limited relevance to the initiation of eukaryotic DNA replication in that they are based on the interaction of viral origin sequences with viral initiator proteins.

II. Origins of DNA Replication from Saccharomyces cerevisiae.

(i) Autonomously replicating sequences

The yeast Saccharomyces cerevisiae has proven to be a useful organism in which to study the molecular details of DNA replication. The relative simplicity of the yeast genome and the ease with which yeast cells may be synchronized to various points of the cell cycle has lead to the development of techniques which are not easily duplicated in higher eukaryotes (Campbell, 1986). Many of the advances in this field have exploited the ease in which both natural and recombinant plasmids may be manipulated in yeast (Struhl et al., 1979). Early on, DNA sequences that permit autonomous replication of selectable plasmids were isolated as candidates for chromosomal origins of replication (Stinchcomb et al., 1979). These sequences had many of the properties one would expect of origin sequences and they were found in the expected frequency for chromosomal origins (Chan and Tye, 1980). Detailed sequence comparisons of ARS elements has identified an 11 bp core consensus sequence that is present in all ARS elements (Broach et al., 1982). This core consensus sequence is the only domain of ARS elements that can be inactivated by point mutations (Kearsey, 1984). A single ARS consensus sequence is not sufficient to promote autonomous replication on its own. ARS function requires a copy of this consensus as well as an additional near match to this sequence 3' to the T-rich strand (Palzkill and Newlon, 1988). This 'replication enhancer' is effective only when it is located 3' to the

consensus core and may involve DNA bending (Snyder et al., 1986). Proteins that bind to ARS1 have been identified in yeast extracts (Diffley and Stillman, 1988), and although they do induce DNA bending, their role in replication has yet to be determined.

The powerful genetic and biochemical techniques available in yeast have been applied to the study of DNA replication on a number of levels. Analysis of temperature sensitive mutants has permitted the identification of a number of the replicative enzymes. Yeast artificial chromosomes (YACS) containing cloned centromeres, telomeres, and ARS elements have been constructed and studied as models for whole chromosome replication. It appears that only one arm of such a 60 kb minichromosome requires an ARS element implying that replication can proceed through a functional centromere (Wellinger and Zakian, 1989). Not only are these artificial chromosomes valuable models for replication, YACS have also proven to be very useful tools in molecular biology for the manipulation of large sequences.

(ii) Origin mapping methods

Recently, methods have been developed to determine whether or not ARS elements correspond to cellular replication origins (Brewer and Fangman, 1987). A two-dimensional gel assay exploiting the altered mobilities of branched and bubble containing DNA molecules was used to map the initiation region of DNA replication on yeast plasmids. DNA replication intermediates containing either bubbles or Y structures are resolved by electrophoresis in the first dimension relative to their mass. These molecules are electrophoresed in the second dimension under conditions that favour migration rates relative to their shape and then detected by southern blotting. The hybridization pattern observed from different digests of the replicating DNA molecules can be used to map the site of initiation. Some of these ARS elements have been mapped on the yeast genome and interestingly, not all sequences that promote plasmid replication actually serve as origins of DNA replication in their chromosomal context. A detailed replication profile of 200 kb from chromosome III suggested that replication was initiated at three out of a possible 10 ARS elements within the region (Reynolds et al., 1989). These data support the hypothesis that only a subset of the total number of potential origins are active in a given cell at a given time.

A similar two dimensional gel assay was used to map the direction of replication fork movement through cloned sequences of yeast DNA (Huberman et al., 1987). In this method, one isolates DNA

from dividing cells and selects for replicating DNA by benzoylated napthoylated DEAE-cellulose chromatography (Hay and DePamphilis, 1982). These fragments are digested with restriction enzymes and electrophoresed through a neutral agarose gel. Under these electrophoresis conditions, fragments are separated on the basis of size and structure. A lane from this gel is embedded in a second gel of higher agarose percentage and run under alkaline conditions. This gel separates individual DNA strands and is transferred by southern blotting to a nitrocellulose filter. Hybridization of these filters with probes spanning the sequences of interest yields hybridizations patterns that can identify the position and direction of fork movement through the sequence of interest. This method was used to map the origin of DNA replication in the 2um plasmid of yeast to within a 200 bp sequence that corresponded to the previously identified ARS sequence as determined by genetic analysis (Broach et al., 1982).

In yeast, the rDNA consists of some 120 tandemly repeated units arranged head to tail on chromosome XII (Phillipsen et al., 1978), and each transcription unit is 9.1 kb long (Bayev et al., 1980, and Bell et al., 1977). Within each repeat there are two nontranscribed spacer regions (NTS) that separate the 37S precursor RNA and 5S RNA regions. A weak ARS element has been mapped within NTS2 (Skryabin et al., 1984). This large array of repeats has been subjected to investigation as to how this unusual set of sequences is replicated by the two dimensional gel mapping technique (Linskens and Huberman, 1988). It was found that DNA replication does in fact originate

within NTS2, but that not all of these ARS elements act as origins in-vivo. The average replicon size is at least several repeats in size and the progress of DNA replication is primarily unidirectional. The direction of replication follows the direction of transcription of the 37S precursor, and forks that move in the direction opposing this transcript, stop near the termination site of transcription. Although these origin mapping methods have been successful in yeast, the added complexity of higher eukaryotic DNA replication has hindered their use in mammalian cells.

III. Mammalian Origins of DNA Replication

(i) The DHFR Replicon

The complexity of the mammalian genome has long been an obstacle to the isolation of a single copy origin of DNA replication. It is difficult to follow the replication of a single copy sequence throughout the long and complicated process of DNA replication. This problem has led several investigators to focus on repetitive genetic elements in an effort to identify and characterize replicons that may be present in more than one copy per haploid genome. The characterization of naturally occurring repetitive elements, ie. Alu sequences and rRNA repeats, or selection for gene amplifications has enabled researchers to work on replication origins that are present in many copies per cell.

Chinese hamster ovary cells were cultured in increasing concentrations of methotrexate to select for the amplification of the dihydrofolate reductase gene. One such methotrexate resistant line (CHOC 400) was isolated that contains 1000 copies of the DHFR amplicon (Milbrandt et al., 1981). In CHOC 400 cells, DHFR constitutes 30% of the cytoplasmic protein and the DHFR mRNA represents 25% of the in-vitro translatable message. Discreet bands from the amplified sequences can be visualized on agarose gels when genomic CHOC 400 DNA is digested with restriction enzymes, and stained with ethidium bromide. This unusual molecular organization has allowed researchers

to study the replication of a normally unique sequence when it is present as a large tandem array of replicons.

This high copy number of the DHFR amplicon has made it possible to identify those subsequences within the amplified region that replicate first (Heintz and Hamlin, 1982). The initiation of DNA replication was investigated by radiolabelling the DNA synthesized in early S phase of synchronized CHOC 400 cells. It was found that DNA replication initiates in a small subset of EcoR1 fragments derived from the amplified sequences. These fragments are replicated consistently in early S phase whether cells are synchronized by either aphidicolin or hydroxyurea block. The other fragments from the amplicon were found to replicate later in S phase. These data suggest that replication of the amplified units is initiated in a discreet and conserved locus that may be similar to the chromosomal origin that functions in the normal unamplified cells. Two such early-labelled fragments (ELF-C, 11.6 kbp; ELF-F, 6.1kbp) were used as probes to isolate clones from a cosmid library made from CHOC 400 DNA (Heintz et al., 1983). A cosmid that contained 28 kbp of genomic DNA was mapped and shown to contain both of these early replicating fragments. The entire amplicon has been cloned in overlapping cosmids (Milbrandt et al., 1983, Looney and Hamlin, 1988), permitting researchers to probe the sequences associated with the DHFR replicon.

High resolution mapping of replication fork movement through the DHFR replicon was achieved by use of an in-gel renaturation technique

(Leu and Hamlin, 1989). This method eliminates background labelling from single copy sequences and allows for a more sensitive quantisation of fragment labelling from the amplified DHFR domains. Initiation of DNA replication was found to occur at two distinct sites within the 243 kb amplified domain. These two initiation zones are separated by some 22 kb of DNA and are located downstream of the DHFR gene. These results were supported by experiments using nascent DNA as hybridization probes on cosmid digests that spanned the amplified domain (Anachkova and Hamlin, 1989), and confirmed that replication initiated within the previously identified fragments ELF-C and ELF-F.

In an effort to better define the DNA sequence elements required for initiation of DNA replication, a 6.2 kb fragment from ELF-F was cloned, sequenced and characterized (Caddle et al., 1990). It contains a region of DNA that has been shown to act as an ARS in yeast (Hamlin et al., 1988), but has not been shown to promote replication of plasmids in transfected CHO cells. This fragment contains a number of non-B form DNA structures and sequence motifs, some of which have been implicated in origin activation. Several DNA unwinding elements were identified in this sequence as determined by mung bean nuclease hypersensitivity assays. A region of bent DNA was found as well as an A-G cluster that can form triplex DNA when under torsional stress at low pH. This sequence contains both ALU sequences and a novel repetitive element as well as several simple repeating sequences ie. $(A\text{-}C)_{18}$, $(A\text{-}G)_{27}$, and $(CAGA)_4$. A 60-kDa polypeptide (RIP60) that binds to the stably bent DNA fragment was purified from HeLa cell nuclear

extracts (Dailey et al., 1990). This protein also binds to the B domain of the yeast autonomously replicating sequence ARS1. Binding of RIP60 to the DHFR fragment alters the DNA structure and increases DNA bending (Caddle et al., 1990). A 100-kDa polypeptide (RIP100) co-purifies with RIP60 and has an ATP dependant helicase activity. The exact role these proteins may play in the initiation of DNA replication has yet to be determined.

During the initiation of DNA replication, there is a transition from discontinuous to continuous DNA synthesis at the origin of bidirectional DNA replication. This transition can be identified by measuring the production of Okazaki fragments produced on each strand. Labelled Okazaki fragments were hybridized to single stranded DNA templates representing single copy sequences of the DHFR locus (Burhans et al., 1990). These experiments were carried out in CHO cells which carry the original unamplified DHFR gene. It was found that at least 80% of the replication forks detected originated within a 0.45 kb sequence approximately 17 kb downstream of the DHFR gene. This sequence containing the origin of bidirectional replication did not contain either the triplex DNA, the bent DNA, or the ARS homology region previously identified in this zone, and it is unclear if these structures actually function in the initiation of replication at this origin.

(ii) The human c-myc replicon

The c-myc proto-oncogene is the cellular homolog of the transforming gene from the avian acute leukemia virus MC-29 (Deusberg et al., 1977). Expression of c-myc is elevated in proliferating cells (Kelly et al., 1983) and amplification of the c-myc gene has been associated with a number of tumour cell lines (Collins and Groudine, 1982). Although the c-myc proto-oncogene is thought to play a role in the proliferating cell cycle, little is known of how this protein functions at the molecular level. One controversial and unconfirmed suggestion is that c-myc acts as an initiator protein by binding to cellular origins of replication. It has been reported that c-myc can substitute for T antigen in an in-vitro SV40 replication system (Iguchi-Ariga et al., 1987a). DNA sequences that contain binding sites for the c-myc protein were purified from human cells and it is reported that nearly 90% of them could replicate autonomously in mouse and human cells (Iguchi-Ariga et al., 1987b). The same authors claim that the c-myc protein binds to an origin of replication that is located 2 kb upstream of the c-myc gene itself (Iguchi-Ariga et al., 1988a), and that the tumour antigen p53 similarly binds to and activates another chromosomal origin (Iguchi-Ariga et al., 1988b). Although much of this data remains unconfirmed and is considered as suspect there is independent evidence for an origin of DNA replication upstream of the human c-myc gene. A run-off replication assay involving in-vitro extension of initiated replication forks has mapped an origin of bidirectional replication to a 3.5 kb domain 5' to the

c-myc gene (McWhinney and Leffak, 1990). A 2.4 kb sub-fragment from this region was cloned into a selectable vector and transfected into HeLa cells. The recombinant plasmid was retained under selection for more than 300 cell generations, and some 1000 times the original plasmid mass could be recovered from the culture. Short term bromodeoxyuridine substitution assays demonstrated that the plasmid replication was semiconservative and followed similar controls that governed chromosomal DNA replication.

(iii) Monkey CV-1 Origin enriched sequences (ors)

Stable replication intermediates can be isolated from proliferating cells if one permits only a limited amount of DNA replication to proceed after initiation at an origin occurs. If the replication is bidirectional, these intermediates consist of a replication bubble centred at the origin of DNA replication. Although these replicative intermediates are stable when the parental DNA is superhelical, the nascent strands are readily extruded from these bubbles when the parental DNA is nicked or made linear (Zannis-Hadjopoulos et al., 1981). The nascent strands can be recovered as double stranded DNA presumably via a mechanism involving DNA breathing, annealing and branch migration. Nascent strands isolated in this way can be recovered from replicating molecules of the SV40 genome were used to map the SV40 origin of replication (Zannis-Hadjopoulos et al., 1984).

Origin enriched nascent DNA was isolated from mouse, rat and monkey cells that were growing synchronously or asynchronously (Zannis-Hadjopoulos et al., 1984). These fragments had an average size of 2 kb and were enriched for sequences that could rapidly reassociate when conditions were changed from denaturing to renaturing. The presence of these snap-back sequences was assayed by hydroxyapatite chromatography. Hydroxyapatite will retain both double stranded DNA and single stranded DNA that has at least 50 bases in the double stranded form (Wilson and Thomas, 1973). Greater than 50 % of

these fragments bound to a hydroxyapatite column suggesting that some origins of replication contain or lie near inverted repeat sequences.

Using the strand extrusion method, early replicating DNA was isolated and cloned from CV-1 cells synchronized to the onset of S phase (Kaufmann et al., 1985). This nascent DNA was cloned into the NruI site of pBR322, producing a library of origin enriched sequences (ors). These clones contained DNA inserts ranging from several hundred to over a thousand base pairs in length and should contain origins of DNA replication at or near their centres. Four out of the five clones tested hybridized preferentially with early replicating fractions of DNA derived from CV-1 cells that were synchronized and labelled with bromodeoxyuridine. A random sample of twelve of these ors clones were chosen and characterized in detail. Two of these clones consisted of unique sequences (ors1 and ors 8); the rest containing either middle or highly repetitive sequences (Zannis-Hadjopoulos et al., 1985). These twelve ors clones were assayed for the ability to promote transient episomal replication when transfected CV-1, COS-7, and HeLa cells (Frappier and Zannis-Hadjopoulos, 1987). Plasmid DNA was recovered from transfected cells and screened by the DPN1 assay (Wirak et al., 1985; Peden et al., 1980) or by measuring the incorporation of bromodeoxyuridine. Four of the twelve ors clones (ors3, 8, 9, and 12) could support plasmid replication in mammalian cells. Ors replication was semiconservative, regulated, and most efficient in HeLa cells. Mapping of replication bubbles by electron microscopy of replicating molecules from ors8 and ors12 plasmids

confirmed that the initiation of DNA replication commenced within the ors insert.

Primary sequence analysis was determined for eight of the ors clones (Rao et al., 1990). Statistical analysis of these sequences did not detect any extensive shared regions or significant consensus sequences with exception of long assymetrically distributed A-T rich stretches. Comparison of these sequences with previously reported consensus sequences from other reports revealed that the ors clones were enriched for the CACCC control consensus described by Dierks et al. (1983), the scaffold attachment T consensus (Gasser and Laemmli, 1986), and the minimal ARS consensus (Palzkill and Newlon, 1988). Recently, a further 23 clones from the ors library were screened for the ability to replicate autonomously in transfected HeLa cells (Landry and Zannis-Hadjopoulos, 1991). Five of these clones were chosen based on homology in dot blot hybridizations to the previously characterized ors3 -8, -9, and 12 sequences. The remaining 18 clones were chosen at random. In all thirteen of the clones (four out of five in the non-random and nine out of the eighteen random) were judged to promote autonomous replication by the DpnI and bromodeoxyuridine assay. Sequence analysis of eleven of these newly identified functional ors sequences revealed the presence of inverted repeats and A-T rich regions. As before there were not any extensive sequence homologies amongst these sequences although several of them did contain either highly or middle repetitive elements.

(iv) Other putative human origins

In prokaryotes and lower eukaryotes one can select for potential origins of DNA replication by cloning random genomic fragments into a plasmid containing a selectable marker (reviewed in Campbell, 1986). Transformation of cells with a library of such plasmids and subsequent selection of resistant clones identifies candidate origin sequences. Although this method has proven to be a reliable technique for isolating autonomously replicating sequences in yeast (Stinchcomb et. al. 1979), similar experiments in higher eukaryotes have not been as successful. One reason why this strategy has not been as effective in higher eukaryotes may be that the sequences that promote plasmid replication are not sufficient to ensure plasmid maintenance over the course of many cell divisions. A variation of this technique designed to circumvent this problem of plasmid stability has been to clone random human genomic fragments into a defective Epstein-Barr virus vector (Krysan et al. 1989). Plasmids containing the origin of DNA replication from Epstein-Barr virus (oriP) replicate autonomously and stably in human cells. Both plasmid replication and maintenance is dependent on the presence of the Epstein-Barr nuclear antigen (EBNA-1) (Yates et al. 1985). The defective vector chosen for these experiments lacks the dyad region of oriP, retains the viral sequences required for nuclear retention, and contains the hygromycin resistance gene for selection. A random human genomic library (average size of 6 kb) was ligated into this vector, transfected into human 293S cells, and after selection for 2 months, plasmids were rescued from several

resistant clones. The average size of the replication promoting human fragments was 12 kb, considerably larger than the size determined for any previous replication origin. A more detailed analysis of these fragments revealed that replication efficiency (as judged by sensitivity to digestion with MboI) was correlated with the size of the fragment (Heinzel et al. 1991). Large fragments were found to replicate more efficiently than smaller fragments in 293S, CV-1 and NC-37 cells. A two-dimensional gel mapping method revealed that initiation of DNA replication occurred throughout the sequence rather than at a specific site (Krysan and Calos, 1991). These data have led the authors to suggest that the initiation of DNA replication in human cells may not be as precise and sequence dependent as it is in lower eukaryotes, prokaryotes and viral systems.

IV. DNA Structures Associated with Replication Origins

(i) DNA unwinding elements

One of the first and most important predictions arising from the double helical model of DNA structure was that the replication of a DNA molecule could be achieved by unwinding the double helix with subsequent synthesis of the complementary strand by using each single strand as a template. Unwinding of the helix not only divides the molecule into two complementary halves (each of which carries an equivalent genetic content), it gives the replication machinery access to the DNA templates. This unwinding of DNA can occur in naturally negatively supercoiled DNA in order to relieve torsional stress (Vinograd et al., 1968; Gellert, 1981). Another prediction derived from the energetics of the base pairing that holds the DNA strands together is that unwinding of the DNA is more likely to occur at A-T rich sequences. The enzyme mung bean nuclease has been used as a probe to identify unwound regions of DNA (Lilley, 1980). Mung bean nuclease cleaves single stranded DNA and was used to identify an unwound A-T rich sequence in the plasmid PM2 that co-localized with the zone of initiation of DNA replication (Shelfin and Kowalski, 1984).

Detailed deletion analysis of the ARS associated with the histone H4 gene has revealed that there are sequences flanking the core consensus that are required for H4 ARS function (Bouton and

Smith, 1986). Although no required primary specific sequences could be identified with 10 bp linker-scanning substitutions, H4 ARS function varied with position and orientation to the flanking sequences. These flanking sequences were assayed for DNA unwinding properties by a mung bean nuclease hypersensitivity assay (Umek and Kowalski, 1988). A readily unwound sequence was found adjacent to the core consensus. Unwinding activity and ARS function was conserved in linker-scanning and linker-deletion derivatives, while large deletions eliminated both unwinding and origin activity. Interestingly, ARS activity in the deletion mutants could be rescued by inserting an unrelated DNA sequence that had similar DNA unwinding properties. A 147 bp fragment of the Amp^r gene terminator from pBR322 that had previously been shown to have unwinding properties (Sheflin and Kowalski, 1985) was able to restore ARS activity. Although easily unwound sequences have a high A-T content, this property alone can not be used to identify mung bean nuclease hypersensitive sites (Umek and Kowalski, 1987). The detection of unwinding elements requires a biochemical assay either in-vitro or in-vivo.

(ii) Bent DNA

The history of DNA curvature dates back to the first published DNA restriction digest (a HindII + HindIII digest of the SV40 genome), although the bent molecules within this digest may not have been appreciated at the time (Danna and Nathans, 1971). Some years later, several reports identified DNA sequences that had reduced mobilities on acrylamide gels compared to what one would expect from their actual size (Mertz and Berg, 1974; Maniatis et al., 1975; Marini, 1977). It was noted that the degree of this electrophoretic anomaly (i.e. reduced mobility) decreased as the temperature of the electrophoresis was increased. Restriction digests of mitochondrial minicircle DNA from trypanosomatids also revealed bent DNA fragments (Simpson, 1979). The sequence of one such fragment was reported, and it was suggested that the reduced electrophoretic mobility was due to a curvature of the helix axis of the molecule (Marini et al., 1982). This observation is in agreement with the model proposed by Trifonov and Sussman (1980) that correlated the spacing of AA dinucleotides and the helix repeat of 10.5 bp/turn. They proposed that AA dinucleotides have noncoplanar base pair planes, and that these distortions might shift the helix axis, resulting in DNA curvature. The sum of several such "wedges" spaced along the repeat length of the helix could contribute to significant DNA bending. These early studies laid the framework for the subsequent experiments linking the curvature of DNA fragments to specific nucleotide sequences.

The origin of DNA replication in bacteriophage lambda contains a region of DNA that exhibits pronounced curvature in solution (Zahn and Blattner, 1985). The initiator O protein binds to this sequence and enhances bending of the origin DNA around the O proteins forming a condensed O-some (Dodson et al., 1985). Molecules of only 84 residues are capable of forming small circles of the appropriate radius of curvature (Zahn and Blattner, 1987). An A-T rich region within the SV40 core origin contains a sequence that directs DNA bending (Deb et al., 1986). This 20 base pair region contains a tract of eight adenines that are conserved among polyomaviruses. Single base mutations within this region decrease the extent of fragment bending and also reduce origin efficiency. It is interesting to note that the sequence requirements for bending and replication are more than a simple requirement for A-T base pairs. Switching the continuous eight adenines for thymidines, while preserving the overall A-T content of the region, drastically reduced replication efficiency (Dean et al., 1987).

The yeast ARS1 element contains a 50 bp sequence that shows evidence of curvature (Snyder et al., 1986). This sequence contains six poly(A) stretches of three to five nucleotides in length phased ten base pairs apart. Mutations in this sequence that disrupt DNA bending also impair replication efficiency of ARS1. Nearly wild type replication activity could be restored in such mutants by the addition of a synthetic bent DNA insert (Williams et al., 1988). The synthetic bent DNA was constructed by annealing and ligating 11-base A-T rich

oligomers that had limited homology to the natural bent sequence. These results suggest that the bent structure per se is crucial for ARS function.

(iii) DNA cruciform structures

It has long been recognized that the properties of DNA molecules, the bases and the bonds that hold them together allow for a wide variety of non-B-form conformations. Indeed, on a purely structural level, DNA molecules can assume an impressive number of conformations (i.e. B-DNA, Z-DNA, DNA triplexes). On an informational level, the double helical nature of DNA molecules offers the possibility to organize the informational content of the base sequences in a variety of ways. DNA cruciforms are peculiar structures in that they contain an inherent duplication of a base sequence with a drastic change in DNA structure. Cruciform formation is possible when there is an inverted and complementary tract of bases along a strand of DNA. Unwinding of such a region of DNA, followed by intrastrand base pairing can generate a structure that contains a junction from which emerges four double stranded DNA helices. Inverted repeats are fundamentally different than direct repeats or a simple palindromes, in that although in all three there is a duplication of a short nucleotide sequence, in the latter two there is no intrastrand complementarity.

The dynamic nature of cruciforms make them attractive structural features that may play a role signalling the initiation of DNA replication. Cruciform formation has been implicated in the initiation of DNA replication from the ColE1 plasmid (Masukata and Tomizawa, 1986), Herpes simplex virus (Weller et al., 1985), human

mitochondrial DNA (Hixon et al., 1986), and autonomously replicating sequences derived from mouse cells (Iguchi-Ariga et al., 1987). In the early seventies, it was observed that supercoiled covalently closed circular DNA molecules could be cleaved by single-strand-specific nucleases (Campbell and Jolly, 1973; Beard et al., 1973). The nature of this cleavage reaction was investigated by mapping the recognition sites of the single-strand-specific endonuclease S1 on supercoiled plasmid preparations (Lilley, 1980; Panayotatos and Wells, 1981)). The single stranded regions cut by this enzyme were highly specific and mapped to regions of short inverted repeats. The nucleotide sequence surrounding these cleavage sites consisted of inverted repeats of nine to thirteen base pairs separated by two to six non-repetitious base pairs. It is thought that the negative supercoiling of the plasmid favours the formation of cruciform structures and the un-base-paired loops of each hairpin are susceptible to cleavage.

The extrusion of cruciform structures is influenced by a number of factors in addition to the superhelical density of the DNA. An analysis of cruciforms from different plasmids, one from ColE1 (C-type) and one from pIRbke8 (S-type) revealed two distinct classes of kinetic behaviour (Lilley, 1985). The C-type cruciform extrudes at low ionic strength and has a large activation energy (180 kcal/mole), while the S-type cruciform extrudes at an optimal salt concentration of 100 uM $MgCl_2$ and has a lower activation energy (42 kcal/mole). The difference between these two cruciforms was found not to be in the

sequences of the inverted repeats themselves, but rather in the composition of the DNA flanking regions (Sullivan and Lilley, 1986). The C-Type cruciform was flanked by sequences with an unusually high A-T content (80%). The dominant effect of the flanking sequence context was demonstrated by switching S and C type cruciforms into the opposite type of DNA flanking sequences. The flanking sequences from the ColE1 cruciform could convert the pIRbke cruciform to C-type kinetics and vice versa. A high A-T content as short as 100 bp is required on only one side of the cruciform to confer C-type kinetics and polarity is unimportant.

Cruciform structures are energetically unfavourable and it has been postulated that they form transiently within a DNA molecule reverting rapidly to linear duplexes (Hand, 1978). Stable cruciform structures can be created by annealing two single strands of DNA that contain different inverted repeats flanked by complementary sequences. Since these structures are completely base paired only in the cruciform configuration, they do not revert to linear molecules. Such a stable cruciform was created by the denaturing and reannealing of a fragment from the SV40 origin (SphI to HindIII), to a similar fragment from a mutant SV40 origin that contained an unrelated substitution of the central inverted repeat (Nobile and Martin, 1986). The heteroduplex molecule was purified away from the homoduplex molecule by polyacrylamide gel electrophoresis and used to immunize mice. Hybridomas were obtained that produced monoclonal antibodies which recognized conformational determinants specific to stem loop

structures (Frappier et al., 1987). One of these hybridomas (2D3) produced IgG1 antibodies, the other (4B4) IgM antibodies. These monoclonals bound to T structures and cruciforms of different base sequences and through interactions with the base of the stem loop or "elbow".

These two monoclonals were used to study the involvement of cruciforms in the initiation of DNA replication. CV-1 cells were synchronized to the onset of S phase, permeabilized, and the progression of DNA replication was monitored in the presence or absence of these antibodies (Zannis-Hadjopoulos et al., 1988). Exposure of the cells to either 2D3 or 4B4 antibodies increased the incorporation of labelled precursor nucleotide by two to six fold over controls either lacking antibodies or with control antibody preparations against B or Z-DNA. Roughly 50% of this enhanced synthesis was sensitive to aphidicolin and dot blot hybridizations of DNA isolated from the treated cells was consistent with the hypothesis that the antibodies stimulated multiple initiations at replication origins and the amplification of DNA sequences.

Monoclonal antibodies specific for cruciform structures were used to determine the quantity and distribution of cruciforms in CV-1 nuclei (Ward et al., 1990). Exponentially growing and cells synchronized to the G1/S boundary by serum starvation and aphidicolin block, were assayed for cruciforms by immunoflourescence analysis. Anti-cruciform staining revealed a non-uniform granular pattern of

fluorescence compared to a nondeductible staining of a control antibody. Quantification of the antibody binding by fluorescence activated cell sorter analysis revealed a bimodal distribution with peaks around 0.6×10^5 and 3×10^5 bound antibodies per nuclei.

V. Outline of the Research Presented in this Thesis

The aim of this work is to isolate and characterize those sequences that play a role in the initiation of DNA replication in human cells. The isolation of these human origin enriched sequences (hors) is presented in chapter 3. Chapter 4 deals with the structural and functional characterization of these sequences in terms of properties that have been associated with other origins of DNA replication. Chapter 5 addresses DNA/protein interactions involving one of these sequences (hors106). Chapter 6 comprises a discussion of the results obtained and suggestions for further research.

CHAPTER TWO

MATERIALS AND METHODS

Unless otherwise indicated, standard techniques of molecular biology were used throughout (ie. for the preparation of plasmid DNA, restriction digests, hybridizations etc.).

I. Synchronization of CV-1 cells

Between 5-7 x 10^5 CV-1 cells were plated per 100 mm tissue culture plate (Corning, N.Y.) in Dulbecco's minimal essential medium (DMEM) supplemented with 5% (v/v) fetal calf serum (FCS, Flow Lab., McLean, VA) and penicillin - streptomycin (50 U/ml and 50 ug/ml respectively, Gibco, Grand Island, N.Y.) and cultured for 24 hours under standard conditions. The plates were then washed with 5 ml phosphate buffered saline (PBS) and the cells were covered with 8 ml DMEM containing penicillin and streptomycin. After 24 hours of serum starvation, the cells were synchronized at the G1/S boundary by the addition of FCS and aphidicolin (Sigma, St. Louis, MO) to a final concentration of 20% (v/v) and 2 ug/ml, respectively. 16-18 hours later, the cells were released into S phase synchrony by removing the aphidicolin-containing medium, washing with PBS and adding DMEM with 5% FCS (v/v) and antibiotics prewarmed to 37°C.

II. ^{3}H-thymidine pulse labelling

Two sets of plates of synchronized CV-1 cells were grown in parallel. One set was sampled at hourly intervals following release

from aphidicolin block and the DNA content profiles were determined by staining with propidium iodide and microfluorometric analysis. The other set of plates was pulse-labelled with ^{3}H-thymidine (5 uCi/ml; 60 Ci/mmol; ICN, Irvine, CA.) for the last 15 min of each hourly sampling. This set allowed for determination of the rate of DNA synthesis based on precursor incorporation at hourly intervals following release from the aphidicolin block. After the 15-minute pulse, plates were washed with 5 ml ice cold PBS and the cells were lysed with 2 ml lysis buffer (10 mM Tris, pH 7.5, 1mM EDTA, 0.5% SDS). The plates were scraped and rinsed with 2 ml PBS and the DNA was precipitated in 5% trichloroacetic acid (TCA) on ice for at least 30 minutes. The acid precipitable material was collected on GF/C filters washed with 5% TCA, 1% sodium pyrophosphate and ethanol, dried and counted in 5 ml scintillation fluid.

III. Microfluorometric analysis of DNA content

At hourly intervals after removal of aphidicolin, the cells were trypsinized (Bacto trypsin Difco laboratories, Detroit, MI) off the plates; trypsinization was stopped by adding 2 ml of FCS and the cells were collected by centrifugation. The resulting pellet was resuspended and fixed in 70% ethanol and stored at 4°C for subsequent analysis. On the day of analysis, the cells were washed 3 times with PBS with centrifugation at 600 x G. The rehydrated cells were resuspended at 10^{6} cells/ml in a propidium iodide solution containing

50 mg/ml propidium iodide (Sigma, St.Louis, MO), 0.01 M tris (pH 7.0), 5 mM $MgCl_2$ and 20ug/ml DNase-free RNase A (Boehringer-Mannheim, West Germany). The samples were incubated at 37°C for 1 hour, placed on ice for 30 minutes and passed through a fluorescence activated cell sorter (FACS-III, Becton-Dickenson, Sunnyvale, CA). Prior to each experiment, the FACS-III was calibrated for florescence and small angle light scatter as described previously (Ward et al. 1986). The FACS-III was calibrated for DNA content by determining the channel number of the G_0 G_1 population of cells in a sample of propidium iodide-labelled HL60 cells (DNA content is approximately 72 pg). Periodically, over the course of running the CV-1 cells, HL60 cells were passed through the FACS-III to ensure consistent calibration. Approximately 20,000 cells were collected for each CV-1 DNA profile and the normalized frequency distribution of DNA content was calculated.

IV. Calculation of DNA Content

The diploid content of CV-1 cells was calculated by determining the weighted mean of the zero hour DNA profile. In order to determine the mean DNA content of cells released from synchrony, the normalized zero hour DNA profile was subtracted from the normalized hourly DNA profiles and an hourly weighted mean DNA content was calculated from the difference. The mean DNA content determined in this manner was normalized against the CV-1 diploid DNA content calculated above.

V. Synchronization of WI-38 cells

Cells were propagated in Dulbecco's minimal essential medium supplemented with glutamine and 10% fetal calf serum (DV10). WI-38 cells at passage 30 were plated at a density of $5x10^5$ cells per 100 x 20 mm plate and incubated in DV10 supplemented with glutamine at 37°C for 48 hrs. The media was changed to DV0 (no serum) and incubated for 28 hours. Fetal calf serum and aphidicolin were added to a final concentration of 20% and 2ug/ml respectively and incubation was continued for a further 18 hours. The plates were then washed with PBS to remove the aphidicolin and warm DV10 was added. At this point the plates were separated into two groups and allowed to enter S phase for 0 to 2 hours. One set of plates was scraped, fixed with ethanol, treated with RNAse A and stained with 50 ug/ml propidium iodide. These cells were run through a fluorescence activated cell sorter (FACS) and DNA content histograms based on over 15,000 cells were produced. The other set of plates was pulsed for 15 minutes with warm DV10 containing 5 uCi/ml of ^3H thymidine (60 uCi/mmol) at hourly intervals following release from aphidicolin block. One ml of lysis buffer was added to each plate, and 10% trichloroacetic acid (TCA) was added to a final concentration of 5%. Following a 30 minute incubation on ice, the lysates were scraped and TCA precipitable material was trapped on GF/C filters. The filters were dried and counted in 5 ml liquefluor in a scintillation counter.

VI. Extrusion and isolation of nascent strands of DNA

The extrusion procedure used was a slight modification of the method previously described (Kaufmann et al., 1985). Briefly, WI-38 cells synchronized to the G1/S border as above were washed free of aphidicolin and floated on a water bath at 37°c in DV10 for 90 seconds. The plates were then immediately washed with cold hypotonic buffer, scraped and the cells were broken open in a dounce homogenizer with seven strokes of pestle B. The nuclei were pelleted and then resuspended in a 50 mM KCL 250 mM sucrose buffer. Elongation of the initiated replication origins was achieved by incubating the nuclei in a replication mixture containing a WI-38 cytosol fraction, free nucleotides, 50 uM ddTTP, Hg- dUTP (Sigma) and 32p-dCTP (3100 Ci/mmol) (ICN) for 10 minutes at 32°C. Following the in-vitro elongation reaction, the nuclei were pelleted, digested with RNAse A, proteinase K and then extracted with phenol and chloroform/isoamyl alcohol. The sample was gently pipetted a few times to shear mildly the highly viscous high molecular weight DNA and then incubated at 50°C for 5 hours in order to extrude the nascent DNA strands. After extrusion, the nascent strands were sedimented on a 32 ml 5-30% neutral sucrose gradient spun in a SW-27 rotor at 25 krpm for 26 hours. Twenty three fractions were collected from 5 ml above the bottom of the tube. A 50 ul aliquot from each fraction was precipitated with TCA and counted in 5 ml of liqueflor. Fractions 11-17, corresponding to DNA ranging from several hundred to several thousand bp in length were pooled and concentrated by ethanol

precipitation. This sample was applied to a 10 ml N-acetyl-D,L-homocysteine column (Pierce Chemical Company, Rockford Illinois) equilibrated in TN buffer (50 mM Tris HCl pH 7.5, 50 mM NaCl). The column was washed with 50 ml TN buffer and then the bound DNA was eluted with 60 ml TN buffer containing 0.1 M beta-mercaptoethanol. Aliquots of 50 ul from each fraction were counted in 5 ml aquasol. Fractions 36-54 containing the material that was selectively bound to the column were pooled and concentrated. This DNA was then digested to give blunt ends by mung bean nuclease and cloned into the Nrul site of pBR322.

VII. Replication Assays

Transient semiconservative replication assays were performed as previously described (Frappier and Zannis-Hadjopoulos, 1987). In brief, 5 ug of purified plasmid DNA was transfected into HeLa cells by the calcium phosphate method (Graham and van der Eb, 1973). The cells were incubated at 37°C overnight in DV5, then washed and incubated in DV5 containing 12.5 ug/ml bromodeoxyuridine (BrdU) for 24 hours. Plasmid DNA was recovered by Hirt's lysis (Hirt, 1967) and centrifuged on a CsCl gradient (initial refractive index of 1.4150) at 70 krpm in a Beckman vertical Vti80 rotor for 16 hours. Fractions of 0.25 ml were collected from the bottom of the tube and the refractive indices of every second fraction were recorded. The relative DNA content of each fraction was determined by densitometry scanning of dot blot

hybridizations probed with nick translated pBR322.

VIII. Unwinding Assays

Mung bean nuclease reactions were performed as described by Umek and Kowalski (1987). Briefly, negatively supercoiled plasmid DNA was nicked by incubation with limiting amounts of mung bean nuclease. The nicked plasmids were then linearized at either a unique Sall or EcoRl restriction site, and the fragments produced were end labelled by the T4 kinase exchange reaction. These fragments were run on a denaturing glyoxal 1.4% agarose gel along with radiolabelled restriction fragments of lambda digested with HindIII. The gel was dried and exposed to Kodak XAR film.

IX. Bent DNA Assays

The presence of anomalously migrating fragments within the hors clones was tested by two methods. Hors plasmids were first digested with restriction enzymes and separated on a 2% agarose gel at room temperature. The lanes of interest were excised, orientated at 90° and separated in a second dimension on a 7% acrylamide gel at 9°C. A small amount (1 ug) of a 123 bp ladder marker (BRL) was included with each sample as a control for the migration of unbent DNA. The arc of DNA fragments was visualised by staining with ethidium bromide.

Digested plasmids were also loaded onto parallel 5% acrylamide gels, run at 9°C and 23°C, respectively and migration distances were compared to those of the 123 base pair marker.

X. Sequencing

Five of the shorter hors clones were sequenced by the Sequenase (United States Biochemical Corporation, Cleveland, Ohio) reaction. Oligonucleotide primers of 15 bases were made complimentary to both DNA strands flanking the Nrul site of pBR322. Sequencing was performed on 2 ug of denatured plasmid and run on an 8.3M urea- 6% acrylamide gel. Sequences were analyzed with the LKB 2020 DNASIS and GCG programs as previously described (Rao et al., 1990).

XI. DNA Bandshift Assays

Labelled DNA subfragments were prepared by first end-labelling 15-mer oligonucleotide primers by the T4 polynucleotide kinase exchange reaction. These primers were then used in a polymerase chain reaction with the desired template for 30 cycles of amplification. The amplified product was digested with the appropriate restriction endonucleases and the desired radiolabelled sub-fragment was purified by acrylamide gel electrophoresis. The labelled DNA fragment was incubated with or without nuclear extract and competitors depending on

the experiment. A typical 20 ul band shift reaction contained: 10 ng of end-labelled DNA (1,000 c.p.m.), 1.5 ug of poly-dIC, 4 ul of binding buffer (60 mM Hepes, 50 mM Tris pH 7.9, 300 mM KCl, 31 mM $MgCl_2$, 3 mM EDTA), and 5 ug of protein extract. After incubation for 20 minutes at room temperature, loading dye was added and the reaction was electrophoresed on a 4% acrylamide gel for 1.5 hours at 180 volts. Gels were then dried and autoradiographed.

XII. Preparation of HeLa Cell Extracts

Hela cell extracts were prepared for both the bandshift and <u>in-vitro</u> replication experiments as described in Pearson et al., (1991). Hela S3 cells, adapted for suspension culture were grown to $5x10^5$ cells/ml and harvested by centrifugation (600 x G for 15 min). Cells were washed twice with isotonic buffer (20 mM Tris-HCl (pH 7.4), 137 mM NaCl, 5 mM KCl, 1 mM $CaCl_2$, 0.5 mM $MgCl_2$) plus 250 mM sucrose then with hypotonic buffer (20 mM Hepes (pH 7.8), 5 mM potassium acetate, 0.5 mM $MgCl_2$, 0.5 mM DTT) at 4°C. Cells at 7 x 10^7 cells/ml were lysed in a Dounce homogenizer (four passes with pestle B) and nuclei were pelleted by centrifugation at 1200 x g for 5 minutes. The nuclei were resuspended in 2.5 x vol. of hypotonic buffer plus 500 mM KoAc and extracted for 90 minutes on ice with occasional vortexing. This extract was then spun in a Beckman SW50.1 rotor at 300,000 x g for 1 hr, and the resulting supernatant (nuclear extract) was frozen in liquid nitrogen and stored at -70°C. The supernatant from the 1200 x

g spin was centrifuged at 100,000 x g for 1 hour in a Beckman Ti 50 rotor and the supernatant (cytoplasmic extract) was frozen in liquid nitrogen and stored at -70°C.

XIII. Protein Purification

Sixty ml of crude nuclear extract (3.06 mg/ml) were applied on a 40 ml DEAE-sephadex A-50 column (Pharmacia) and the column was washed with 250 ml of buffer (20 mM Hepes pH 7.9, 20% v/v glycerol, 0.2 mM EDTA, 10 mM $Na_2S_2O_4$, 1 mM PMSF, 0.5 mM DTT, 0.1 M KCl), and the flow through fractions of 7.5 ml each were collected. Bound protein was then eluted with a linear gradient using the same buffer containing from 0.1 to 1.5 M KCl. Every second fraction recovered from the column was assayed for bandshift activity and protein concentration estimated by spectrophotometry. Fractions 3-11 of the flow through contained the bulk of the bandshift activity and were concentrated to a volume of 30 ml (1.11 mg/ml) with polyethylene glycol. This material was applied to a 40 ml PC-11 phosphocellulose column (Whatman) and protein was washed and eluted with the same buffers as above. Again every second fraction was assayed for protein content and bandshift activity. The peak bandshift activity was found in fractions 61-65 from the KCL gradient. Every fraction between 59-67 was taken for analysis on SDS PAGE by silver staining and for bandshift assays. The peak bandshift activity was found in fractions 63-65, and these fractions were pooled and concentrated (0.14 mg/ml)

to yield the enriched DNA binding fraction from the PC-11 column.

XIV. Southwestern Analysis

Two parallel 4/12.5% SDS polyacrylamide gels were run with the enriched protein extract from the PC-11 column and a low molecular weight marker (BRL). Both gels were pre-run at 150 v for 0.5 hours and after the samples were loaded for 1 hour at 150 volts. One gel was stained with Commasie blue, and the other was electroblotted onto a nitrocellulose filter with 25 mM Tris, 192 mM glycine, and 20% v/v methanol, for 17 hours at 4°C, 23 volts and 40 mA, as described in the Mini Trans Blot protocol (Biorad). The nitrocellulose membrane was washed with 50 ml of binding buffer containing 10 mM Tris HCl (pH 7.2), 50 mM NaCl, 1 mM EDTA, 0.2% polyvinylpyrolidine, 0.2% ficoll, 0.2% bovine serum albumin, and 2.5% NP40 at room temperature for 1 hour. The membrane was then washed twice with 50 ml binding buffer without NP40 for 30 minutes at room temperature. The membrane was probed with 40,000 cpm of end-labelled HinfI fragment of hors106 in 5 ml of binding buffer for 2.5 hours at room temperature. The filter was rinsed briefly in water and autoradiographed.

XV. <u>In-vitro</u> replication assays

<u>in-vitro</u> replication assays (Pearson et al., 1991) were carried

out with 250 ng of supercoiled plasmid DNA in a 50 ul reaction volume containing: 15 ul of cytoplasmic extract, 8 ul of nuclear extract, and a reaction mix containing a final concentration of 45 mM Hepes (pH 7.8), 5 mM $MgCl_2$, 1 mM EGTA, 60 mM sucrose, 240 mM ethylene glycol, 5% poly(ethylene glycol) (Mr 12 000, Fluka), 6 mM phosphoenolpyruvate, 0.3 U pyruvate kinase (Boehringer-Mannheim), 2 mM ATP, 100 mM each CTP, GTP UTP, dGTP, dATP, and dTTP, 10 uM dCTP and 10 uCi of alpha^{32}P-dCTP. Purified protein fraction were added as indicated and the reaction was incubated for up to an hour at 30°C, after which they were stopped by the addition of 1 vol. of 1% SDS, 30 mM EDTA, 10 ug tRNA and 10 ug proteinase K. The reaction was incubated for 1 hour at 37°C and the DNA was purified and concentrated by phenol extraction and ethanol precipitation. The labelled DNA was run on a 1% agarose gel (20 cm x 20 cm) for 16 hours at 60 volts, the gel was dried and then autoradiographed.

CHAPTER THREE

SYNCHRONIZATION OF CELLS TO EARLY S PHASE AND THE ISOLATION OF NASCENT DNA

I. Synchronization of CV-1 cells to early S phase

The objective of this thesis is to research at the molecular level the mechanisms involved in the initiation of eukaryotic DNA replication. To this end, it was decided to make use of techniques that would permit the isolation of those DNA sequences that are replicated at the beginning of S phase. The cell line CV-1 is an immortalised, non-transformed line derived from African green monkey kidney. These cells are easily cultured under standard conditions, may be synchronized to various points of the cell cycle, and are amenable to studies on the progression of DNA replication (Kaufmann et al., 1985; McAlear et al., 1989). This cell line was used in the initial synchronization studies in order to develop a synchronization protocol and determine the effects of cell cycle arrest on the progression of DNA replication. It was of interest to determine not only the efficiency of cell cycle arrest to the G1/S border, but also the kinetics of recovery from aphidicolin block and the initiation of DNA replication. Logarithmically growing CV-1 cells were synchronized to the G1/S border by serum starvation and aphidicolin block as described in the materials and methods. Briefly, cells are first arrested in the G1 phase by serum starvation and then blocked at the G1/S interface by serum stimulation in the presence of aphidicolin. Aphidicolin is a potent inhibitor of the replicative DNA polymerases alpha and delta that effectively precludes DNA synthesis (Pedrali-Noy et al., 1980). Apidicolin does not inhibit the enzyme DNA primase, thus permitting the early initiating events to occur in its presence

(Dinter-Gotleib and Kaufmann, 1982). Once aphidicolin is removed from the media, cells immediately progress into S phase. The efficiency of the synchronization protocol and the ease with which cells may be released from this arrest was determined by microfluorometric analysis.

A logarithmically growing culture of CV-1 cells was harvested, stained with propidium iodide, run through a fluorescence activated cell sorter and analyzed with respect to DNA content. Under the condition used, a G1 content of DNA (C) corresponded to a fluorescent intensity centred around channel 22 and a G2/M DNA content (2C) centred around channel 45. DNA content histograms detailing the percentage of cells within each channel of fluorescence were constructed based on over 15,000 cells per analysis. In the logarithmically growing culture (fig. 1a), one can readily identify cells in each phase of the cell cycle, having DNA contents ranging from C (G1) to 2C (G2/M). In contrast, DNA content histograms from CV-1 cells 15 minutes after release from the aphidicolin block reveal a tight population of cells entering S phase (fig. 1b). In the ensuing hours, the CV-1 cells continued to synthesize DNA and progressed as a synchronous wave through S phase (fig. 1 c-f).

The progression of CV-1 cells passing through S phase was monitored for a total of nine hours following release from aphidicolin arrest. The average DNA content of the synchronized cell populations was calculated at hourly intervals and used to position each

population of cells within S phase (fig. 2). The relative DNA contents of the synchronized populations were seen to increase with time following entry into S phase, approaching completion of DNA synthesis at nine hours. It is clear from this analysis that the synchronization protocol can effectively synchronize cells to the G1/S interface and that DNA synthesis occurs soon after removal of aphidicolin. In fact, after only one hour, almost one quarter of the parental DNA has been replicated.

Figure 1. DNA content histograms of synchronized CV-1 cells. DNA content histograms of CV-1 cells growing logarithmically or following release from aphidicolin block were constructed by FACS analysis of over 15,000 cells per time point. Fluorescent intensities corresponding to the G1 DNA content (C), and G2/M DNA content (2C) are centred on channels 22 and 45 respectively. (A) logarithmically growing cells; (B-F) cells released from block for (B), 15 minutes; (C), 1 hour; (D), 2 hour; (E), 4 hours; (F), 6 hours.

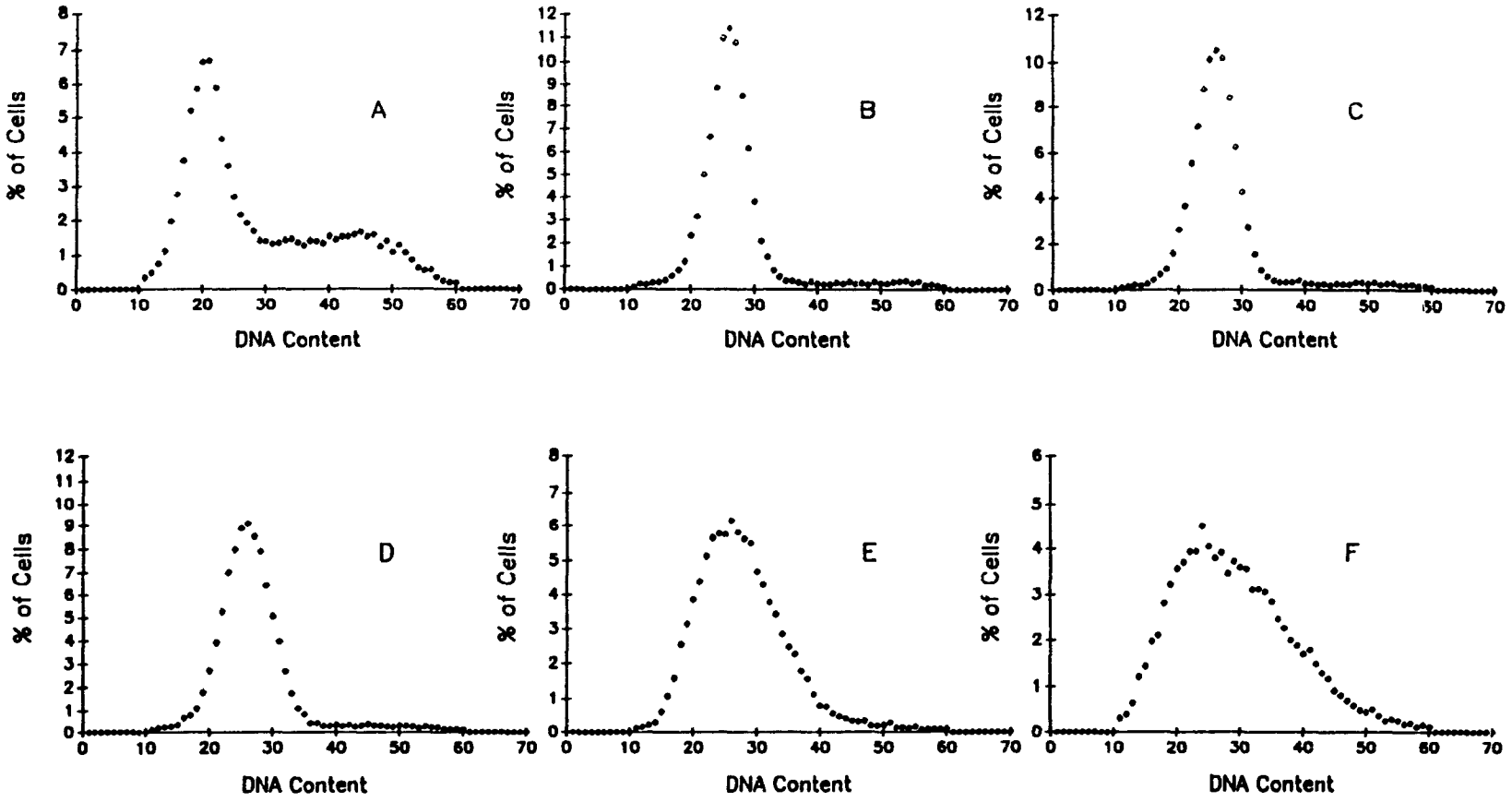
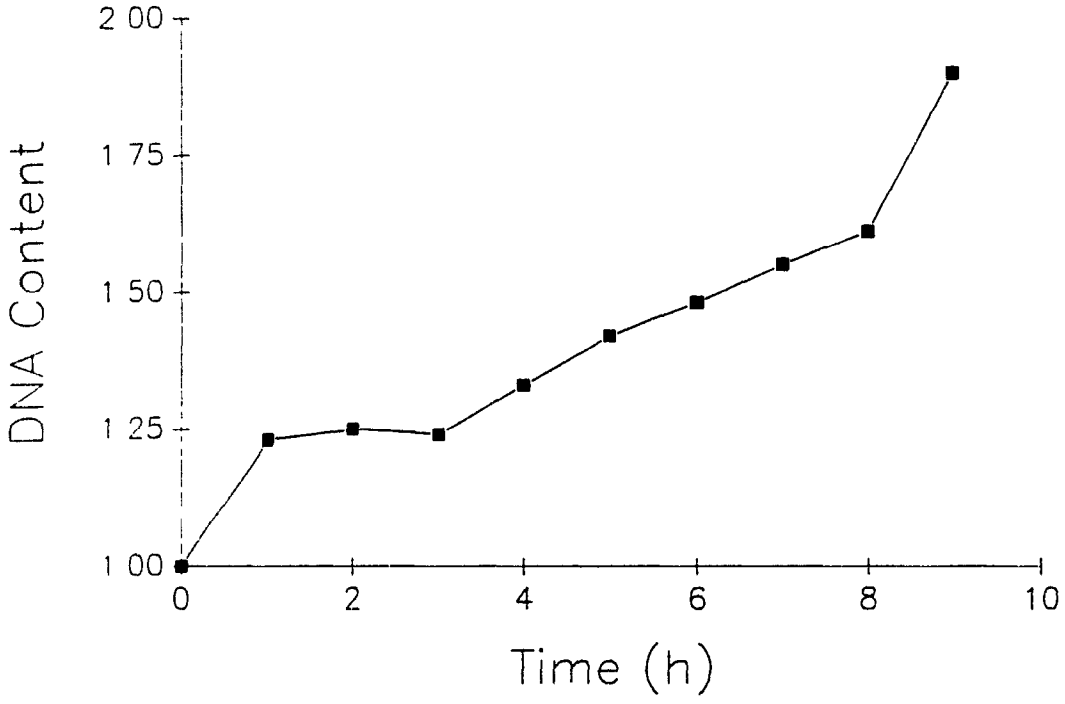

A
B
C
D
E
F
% of Cells
DNA Content

Figure 2. The progression of S phase in synchronized CV-1 cells. The relative position in S phase of the synchronized cell populations was calculated from the hourly DNA content histograms up to nine hours following release from aphidicolin arrest. The G1 content of DNA is defined as 1.00 and the G2/M DNA content as 2.00.


2 00
1 75
1 50
1 25
1 00
DNA Content
0
2
4
6
8
10
Time (h)

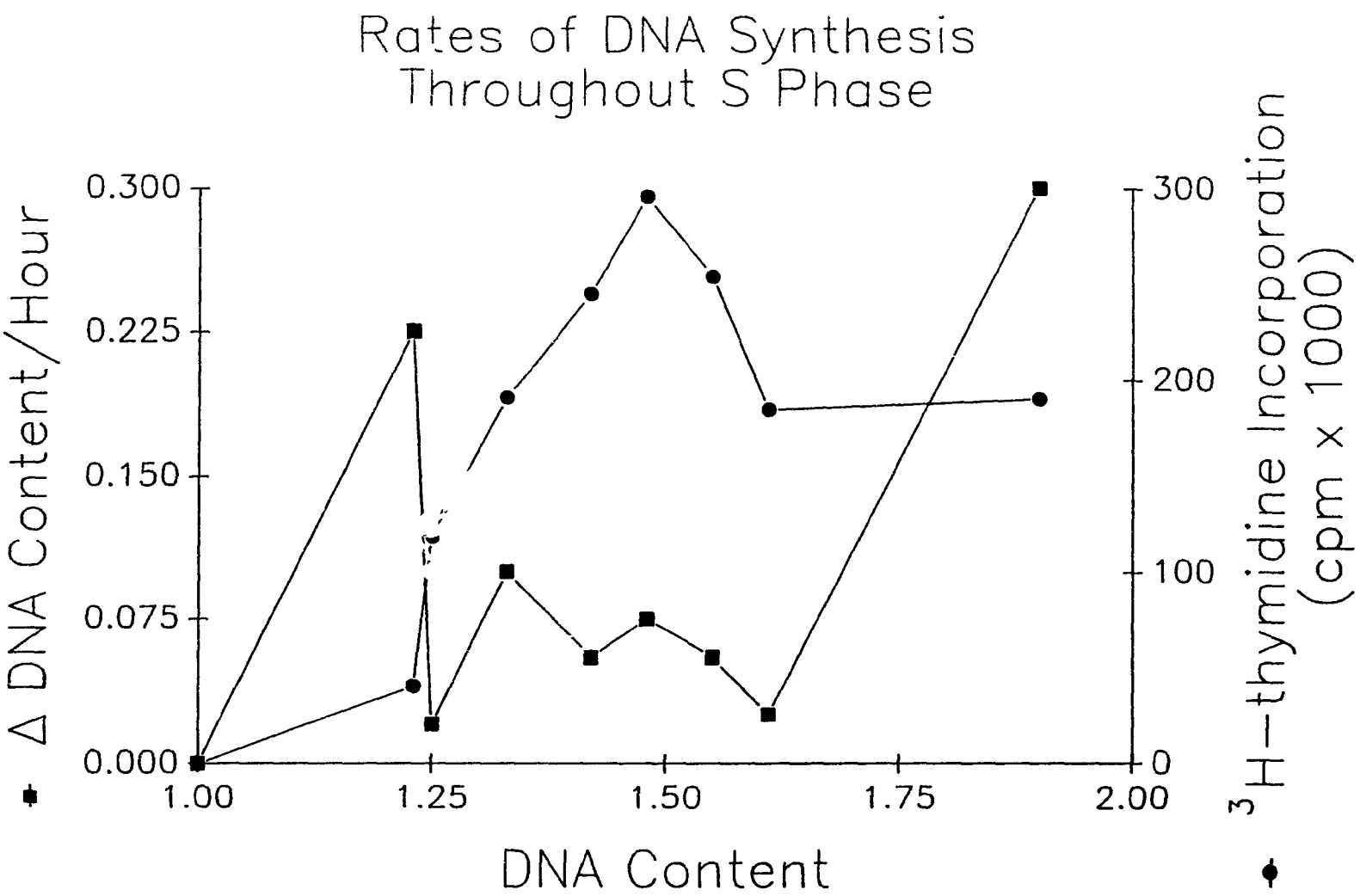
II. The rate of DNA synthesis in CV-1 cells throughout S phase

The rate at which DNA is synthesized is a fundamental property of DNA replication. One can calculate a mean rate of DNA replication by dividing the DNA content by the length of S phase, but this calculation does not provide insight into how the rate of DNA replication might vary throughout S phase. At any one time in S phase, the rate of DNA replication reflects the number of active replication forks and the rate of polymerization of nucleotides at each fork. If the rate of nucleotide polymerization is assumed to be relatively constant among replication forks, then the rate of DNA replication can be used to measure the number of active replication forks, which in turn reflects the number of activated origins of replication. For instance, the increased rate of doubling in embryonal cells is thought to result from an increased activation of origins of replication, rather than an increased rate of nucleotide polymerization by the replicative machinery.

The synchronization procedure and the DNA content histograms were used to measure the rate of DNA synthesis as CV-1 cells progress through S phase (Collins et al., 1980). The average DNA content of the cell populations calculated from each DNA histogram was used to place each population of cells at a point within S phase. The rate of DNA synthesis was determined following removal of aphidicolin in two ways: a) as a change in the DNA content over time; and b) by measuring the incorporation of ^{3}H-thymidine into acid precipitable material

during a 15 minute pulse (fig. 3). The synchronized cells were followed for 9 hours after release into S phase. When the rate of DNA synthesis was calculated based on the change in DNA content, it was observed that there was a high rate of DNA synthesis in the first hour of S (DNA content of 1.0 to 1.25), with subsequent lower rates fluctuating until approximately 60% of the genome was replicated (DNA content of 1.25 to 1.6) and, finally, a major late increase between 6 to 8 hours in S phase until complete duplication of the DNA (DNA content of 1.6 to 2.0) was achieved. This profile did not mirror the replication profile obtained by measuring the incorporation of ^{3}H-thymidine at hourly intervals through S phase. These labelling studies revealed an initial low rate of incorporation, followed by a sharp increase extending up to 6 hours and then dropping to a lower level for the remainder of S phase. These results are similar to ones we have reported previously with regard to the rate of precursor incorporation in aphidicolin-synchronized CV-1 cells (Kaufmann et al. 1985). Although the two methods for measuring the rate of DNA synthesis through S phase yield different profiles, they nonetheless correlate very well with the data obtained from similar experiments performed with HeLa cells that were synchronized by a double thymidine block (Collins, 1978). In that study, the rates of DNA synthesis obtained through microfuorimetric analyses were considered more representative of the true rates of DNA synthesis, as those measurements were not susceptible to the complexities of thymidine metabolism and intracellular nucleotide pools.

Figure 3. The rate of DNA synthesis in synchronized CV-1 cells. The rate of DNA replication throughout S phase in synchronized CV-1 cells was determined by microfluorometric analysis and by ^{3}H-thymidine incorporation. The rate of DNA replication was estimated by (A) the increase in DNA content per hour (-●-), and (B) by measuring the incorporation of ^{3}H-thymidine during a 15 minute pulse (-▲-).

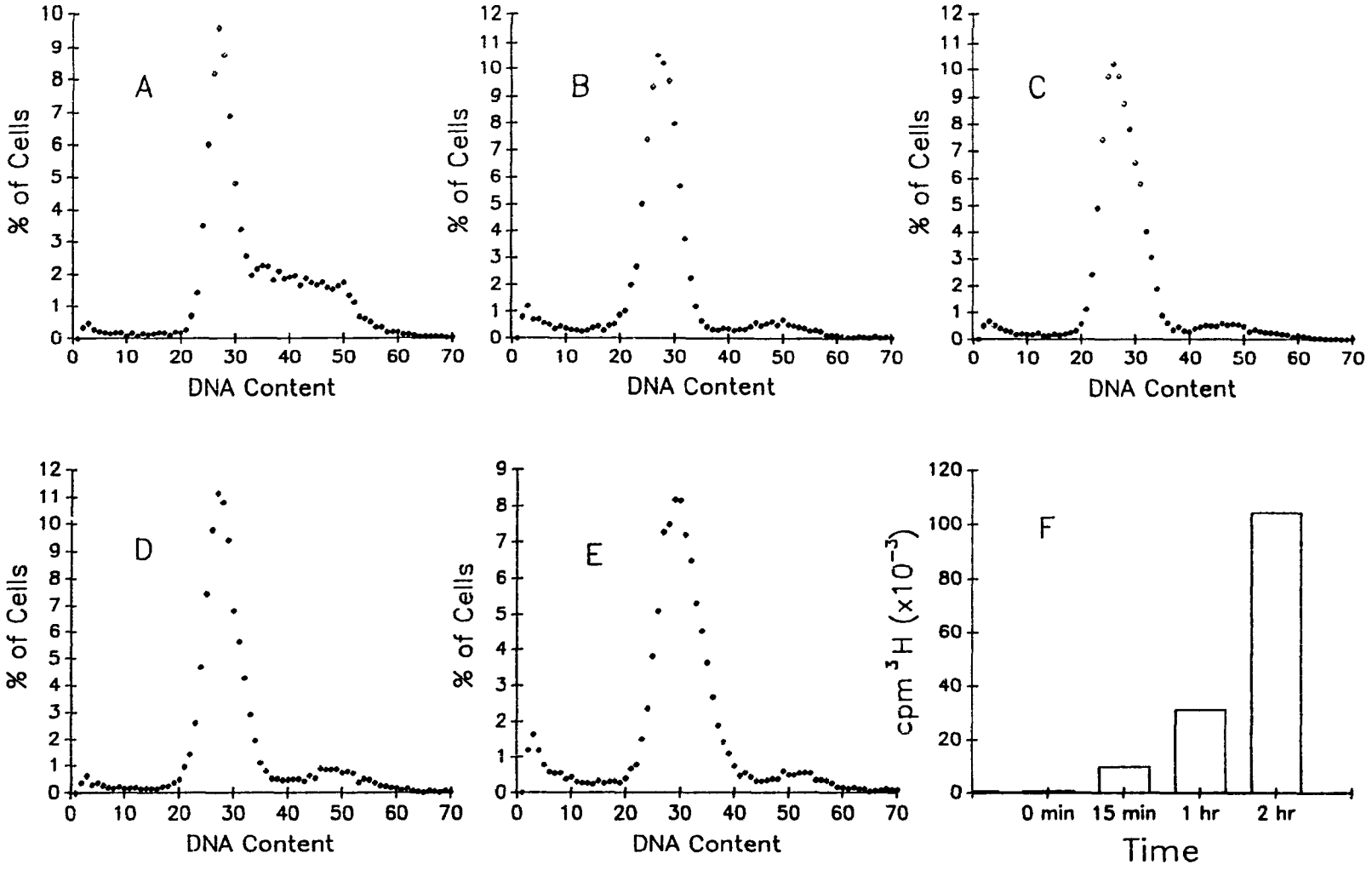

Rates of DNA Synthesis
Throughout S Phase
0.300
0.225
0.150
0.075
0.000
Δ DNA Content/Hour
300
200
100
0
^{3}H—thymidine Incorporation
(cpm x 1000)
1.00
1.25
1.50
1.75
2.00
DNA Content

III. Synchronization of human (WI-38) Cells

The cell strain WI-38 is a normal human fetal lung fibroblast strain that has a limited number of population doublings. These cells were chosen for the subsequent studies on the isolation of early replicating DNA sequences because they represent a non-immortalized population of cells whose progression through S phase may be more normal than that of immortalized or transformed cell lines. WI-38 cells at passage 30 were cultured and synchronized to the G1/S border by serum starvation and aphidicolin block. DNA content histograms were prepared by FACS analysis of non-synchronized cells as well as synchronized cell populations at various times following release from aphidicolin block in the same way as was done for the CV-1 cells. In a proliferating cell culture (fig. 4 A), the cells are distributed throughout the phases of the cell cycle, having DNA contents ranging from diploid (G1) to tetraploid (G2/M). Under the defined parameters, the G1 and G2/M DNA contents correspond to the fluorescent intensities centred around channels 27 and 50, respectively. The efficiency of the synchrony method is evident in that the DNA content of over 70% of the cells measured at 0 h (fig 4 B) fell within channels 23-33. After release from aphidicolin block, cells progressed as a synchronous wave into S phase (fig 4 C, D and E). Quantitative measurements of the rates of DNA replication by ^{3}H-thymidine pulse labelling in cells blocked in G1/S up to 2 h after their release into S indicate that DNA synthesis in the blocked population is essentially zero but increases progressively over the next two hours (fig 4 F). This technique

allowed us to focus on that fraction of DNA replication that occurrs at the very onset of S phase in WI-38 cells.

Fig. 4. The synchronization of WI-38 cells. WI-38 cells were synchronized to the G1/S border, released from aphidicolin block and followed into S phase with microfluorometric analysis and ^{3}H-thymidine incorporation. The progression of DNA synthesis was monitored by both the DNA content histograms (A), log culture; (B), 0 hr; (C), 15 min; (D), 1 hr, (E), 2 hrs; and (F) by the incorporation of ^{3}H thymidine during a 15 minute pulse.


A
B
C
D
E
F
% of Cells
DNA Content
cpm ^{3}H (x10^{-3})
0 min
15 min
1 hr
2 hr
Time

IV. Extrusion and purification of nascent WI-38 DNA

A synchronized cell population was the first step towards the isolation of those sequences that initiate DNA replication. The approach taken was to purify short nascent strands of DNA, as they should contain origins of replication at or near their centre (Zannis-Hadjopoulos et al., 1983). The extrusion method was used to isolate those sequences of DNA that are replicated at the very onset of S phase (Kaufmann et al., 1985) (fig. 5). WI-38 cells were synchronized as above and allowed to enter S phase for 90 seconds. In this brief interval of release from aphidicolin block, only a limited amount of DNA replication takes place initiated from the origins that are activated at the onset of S phase. Each of the short nascent DNA strands should, therefore, contain an origin of DNA replication. Nuclei were prepared from these cells and the nascent strands were elongated <u>in-vitro</u> in the presence of mercurated and radiolabelled nucleotides. This labelling reaction served to tag the replicating molecules and facilitate their purification. After the labelling reaction, the DNA was sheared slightly and the nascent strands were extruded from the replication bubbles as linear double stranded DNA. The extruded nascent DNA was separated from the high molecular weight parental DNA by sedimentation on a neutral 5-30% sucrose gradient (fig 6 A). The bulk of the labelled material sedimented in the top half of the gradient, corresponding to DNA of 100 to 5000 base pairs in length. Fractions 18-23 contained DNA of the size associated with Okazaki fragments (data not shown) and were discarded. Fractions 11-

17 were pooled and applied to a mercury affinity homocysteine column. This chromatography step was used to separate the Hg-dUTP labelled nascent DNA from any short strands of sheared parental DNA that could have been produced during the experimental manipulations. More than 80% of the alpha-^{32}-P-dCTP labelled material was selectively bound to the column and could be eluted with buffer containing beta-mercaptoethanol (fig. 6 B). In contrast, a non-mercurated sample of DNA was applied to a control column, and found to elute within the first several fractions (data not shown). Fractions 36-54, consisting of short DNA sequences that replicate at the onset of S phase, were pooled and concentrated. This DNA was digested with mung bean nuclease to yield blunt ends and cloned into the Nrul site of pBR322 (fig. 7 A). Recombinant clones were isolated and the size of the insert DNA was determined by running Bgll digests on acrylamide gels. The resulting library of human origin enriched sequences (hors) consisted of 149 A^rT^s clones with inserts ranging from 200 to 4000 base pairs in length. A random sample of ten recombinant clones was chosen for further analysis concerning the properties of these origin enriched sequences (fig. 7 B)

Figure 5. The extrusion protocol. The method for isolating origin enriched sequences required (A) the synchronization of cells to the G1/S border, (B) the initiation of replication in-vivo for 90 seconds, (C) elongation in-vitro with Hg-dUTP and $dCTP^{32}$, (D) extrusion of nascent strands from replication bubbles, (E) purification of nascent strands.

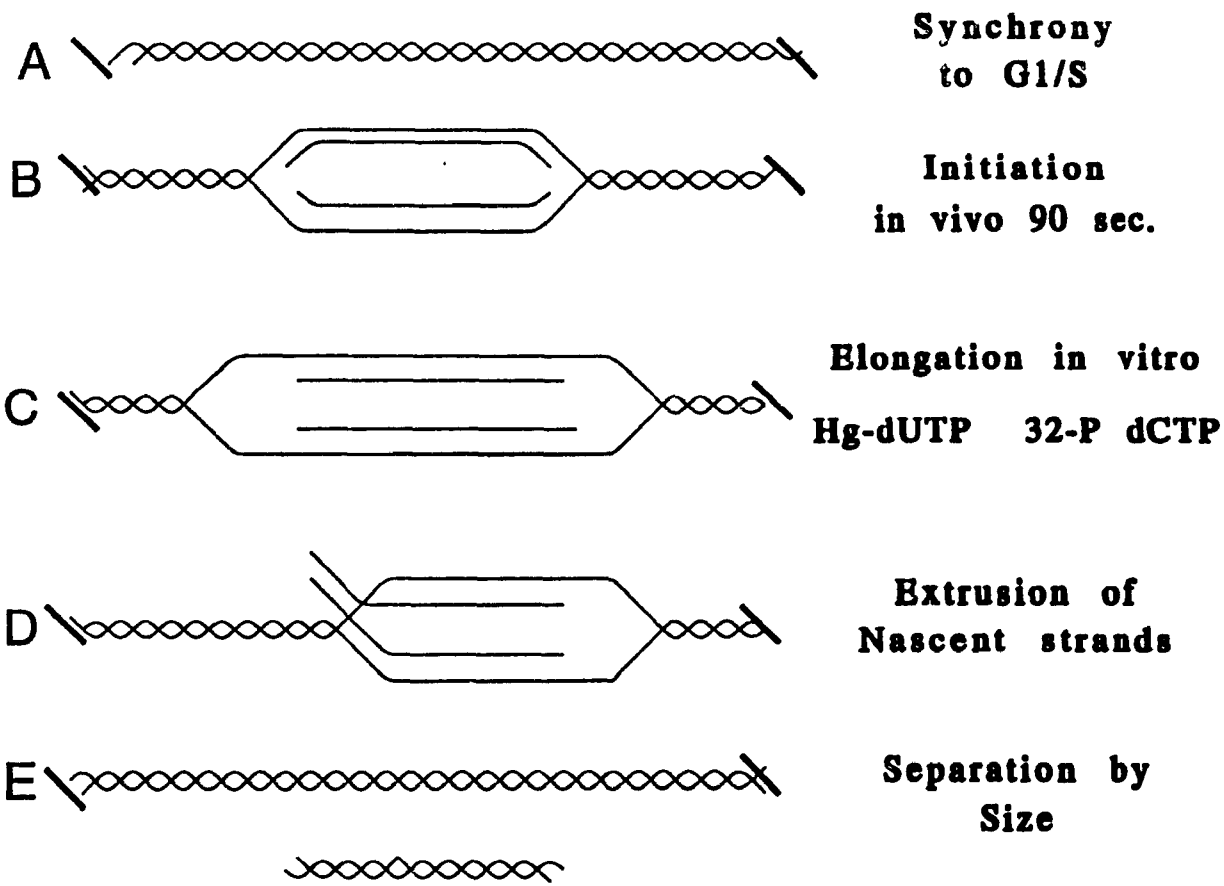

A
Synchrony
to G1/S
B
Initiation
in vivo 90 sec.
C
Elongation in vitro
Hg-dUTP 32-P dCTP
D
Extrusion of
Nascent strands
E
Separation by
Size

Fig. 6. The purification of nascent DNA. Early replicating WI-38 DNA was labelled in-vitro with $dCTP^{32}$ and Hg-dUTP and extruded from replication bubbles as double stranded DNA. This short, nascent DNA was separated from high molecular weight parental DNA by (A) centrifugation on a 5-30% neutral sucrose gradient, (B) by mercury affinity homocysteine chromatography. Elution of the bound DNA from the homocysteine column began at fraction 30 with buffer containing beta-mercaptoethanol.

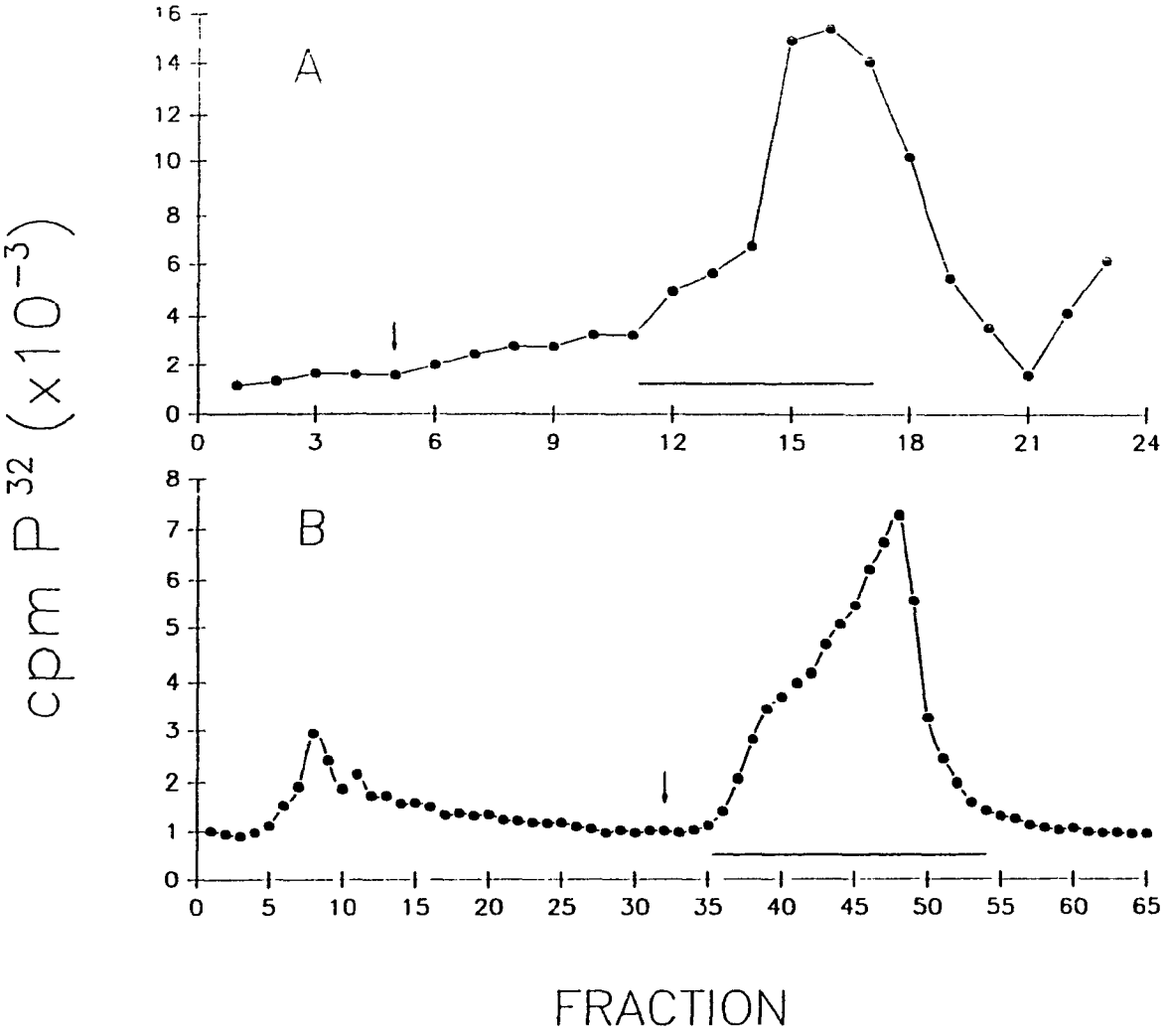
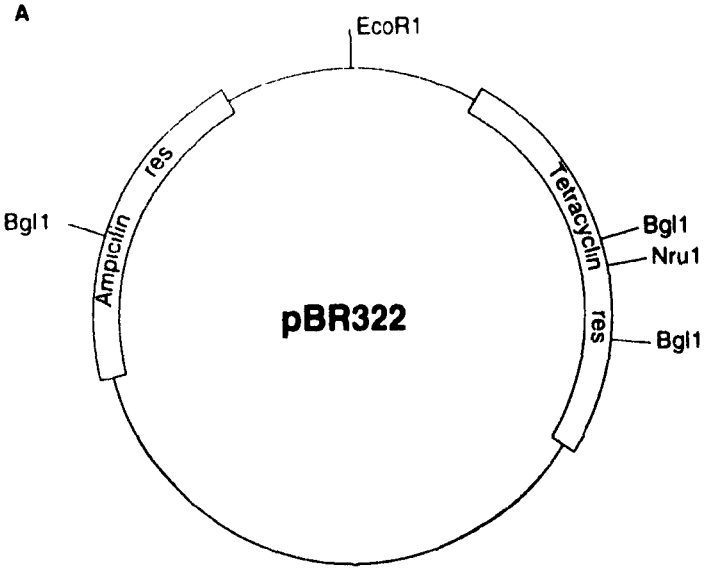

A
B
cpm P 32 (x10−3)
FRACTION
0
2
4
6
8
10
12
14
16
0
1
2
3
4
5
6
7
8
0 3 6 9 12 15 18 21 24
0 5 10 15 20 25 30 35 40 45 50 55 60 65

Figure 7. Cloning of human origin enriched sequences. The origin enriched WI-38 DNA was digested with mung bean nuclease to give blunt ends and cloned into the NruI site of pBR322 (A). Ten clones were chosen at random, digested with BglI and run on a 4% acrylamide gel to reveal the DNA inserts at the Nrul site (B).

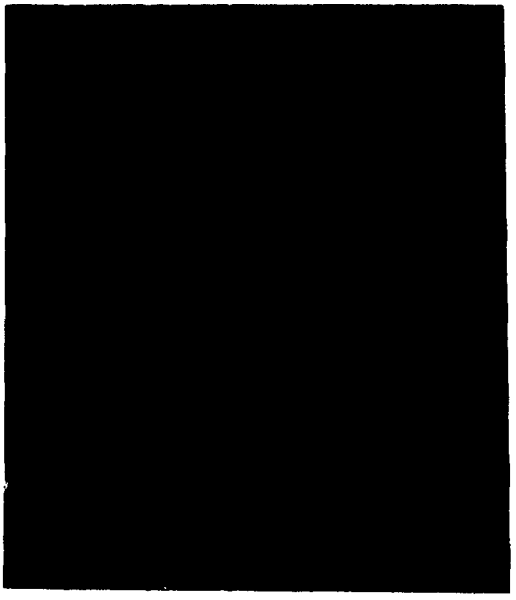
A



B

HORS Bgl1

M 67 74 76 98 103 106 110 112 129 133



CHAPTER FOUR

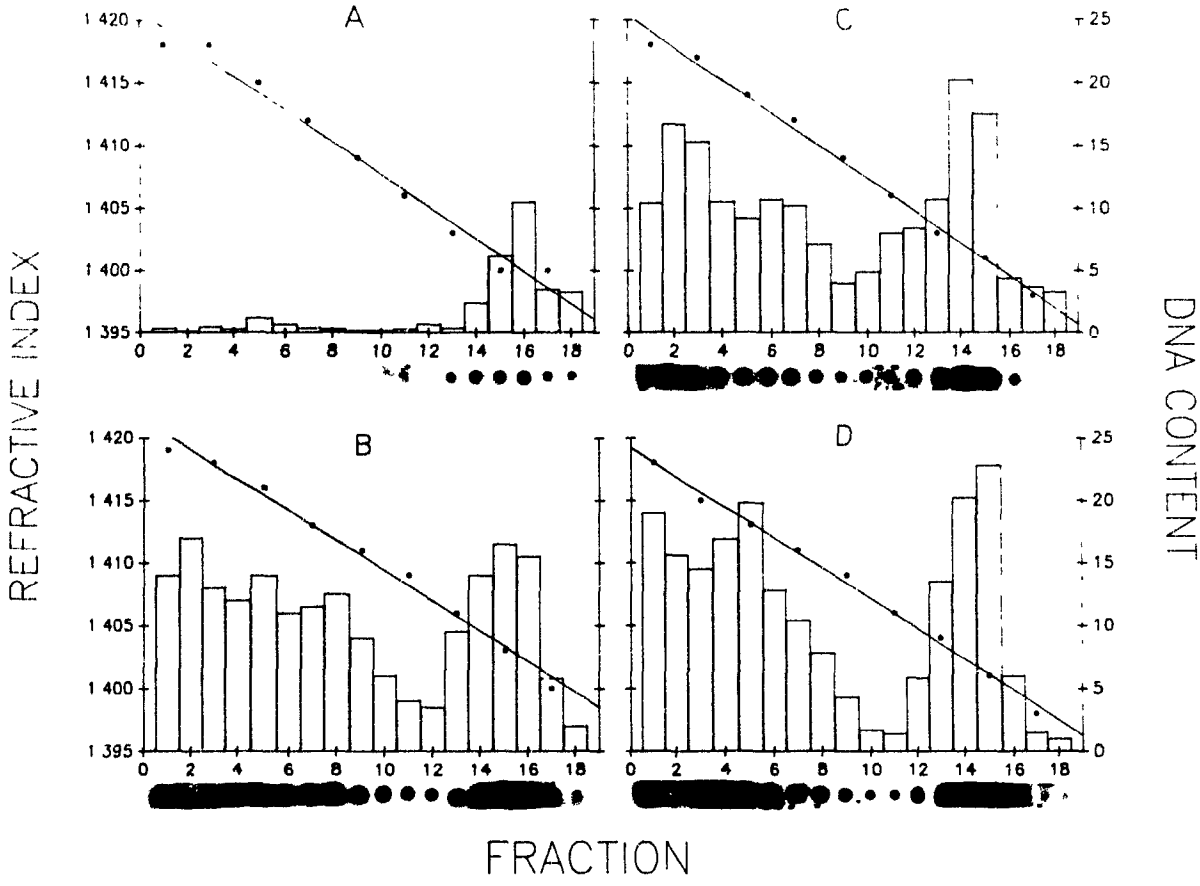
PROPERTIES OF ORIGIN ENRICHED SEQUENCES (HORS)

I. Autonomous Replication of Hors Plasmids

The random sample of ten hors clones were tested for the ability to replicate autonomously as plasmids in HeLa cells by the bromodeoxyuridine (BrdU) substitution assay. It was of interest to determine whether the sequences isolated could confer autonomous replication activity on the pBR322 vector. Plasmids were transfected into HeLa cells by the calcium phosphate method and allowed to replicate in the presence of BrdU for 24 hours. BrdU is a heavy analog of thymidine and any replicating molecules incorporating it into DNA will have a greater density than unreplicated plasmids. Hirt's lysates were prepared from the transfected cells and replicated molecules were separated from unreplicated ones on isopycnic CsCl gradients. Fractions were collected and the integrity of each gradient was verified by reading the refractive index of every second fraction. The relative DNA content of each fraction was determined by dot blot hybridizations. DNA recovered from cells transfected with the vector pBR322 was of light-light (LL) density, identical to that of the input DNA (fig 8 A). In contrast, replication profiles for hors clones 98, 106, and 112 (fig 8 B, C and D) in addition to LL DNA reveal considerable quantities of heavy-heavy (HH) and heavy-light (HL) DNA, indicating that some molecules underwent one or more than one rounds of replication. In all, 5 out of the 10 clones (hors 76, 98, 106, 112 and 129; see table 2) had autonomous replication activity as judged by this assay . Southern blot analyses of the fractionated gradients confirmed that the replicated DNA was neither rearranged nor

integrated into the genome (data not shown).

Fig. 8. Autonomous replication of hors clones. Clones were checked for transient autonomous replication by the BrdU substitution assay. Plasmids were transfected into HeLa cells, allowed to replicate in the presence of BrdU for 24 hrs, recovered and separated by density on CsCl gradients. The refractive indices and DNA contents of the fractionated gradients are shown for (A), pBR322; (B), hors98; (C), hors106; (D), hors112.


A
B
C
D
1 420
1 415
1 410
1 405
1 400
1 395
0 2 4 6 8 10 12 14 16 18
25
20
15
10
5
0
REFRACTIVE INDEX
DNA CONTENT
FRACTION

II. Bent DNA assays

DNA of abnormal electrophoretic mobility has been associated with origins of replication from several viruses, bacteria and yeast plasmids. This retarded mobility is more pronounced at low temperatures and is thought to result from a bend in the double helix Restriction digested DNAs of the hors clones were checked for anomalously migrating fragments by a 2-dimensional gel assay; digests were first run on a 2% agarose gel at 23°C, then excised and run at 9°C on a 7% acrylamide gel at a 90° angle. DNA fragments of random composition run as a smooth arc in the second dimension whereas abnormal DNA structures are associated with fragments that run above or below the arc. The digested hors samples were "spiked" with DNA from the 123 bp ladder in order to help in defining the arc of normal mobility. Two of the hors clones (hors76 and hors110) analyzed showed evidence of bent DNA. A two-dimensional gel of an AvaII digest of hors76, reveals two fragments of altered mobility (fig 9 A) One fragment migrates above the arc, indicating that it has a decreased mobility at 9°C. This fragment contains the NruI site and the insert DNA. The other smaller fragment that migrates below the arc is a fragment of pBR322 that has previously been shown to have an increased electrophoretic mobility. A more detailed analysis of the bent clones was performed by running parallel acrylamide gels at 9°C and 23°C. Bent DNA structures are more stable at lower temperatures and altered mobilities are most pronounced when the bend is located at the centre of the DNA fragment. Three different digests of hors76 were analyzed

in order to characterize the anomalously migrating fragments. The relative mobilities were calculated by referring to the 123 bp ladder marker. In each digest one can detect bands derived from the hors76 insert that migrate significantly slower at 9°C (fig. 10). The vector pBR322 contains bent DNA sequences distinct from the hors inserts. A similar analysis of hors110 revealed a bent DNA fragment and a summary of the bent fragments from hors76 and hors110 and their relative mobilities are listed in table 1.

Fig. 9. Two dimensional gel bent DNA assay. An AvaII digest of hors76 was spiked with a 123 bp ladder marker and run on a two-dimensional gel assay; first dimension 2 % agarose, 23°C, second dimension 7% acrylamide, 9°C a.d visualised with ethidium bromide staining. The arrow indicates a retarded DNA fragment.

Hors 76 Avall

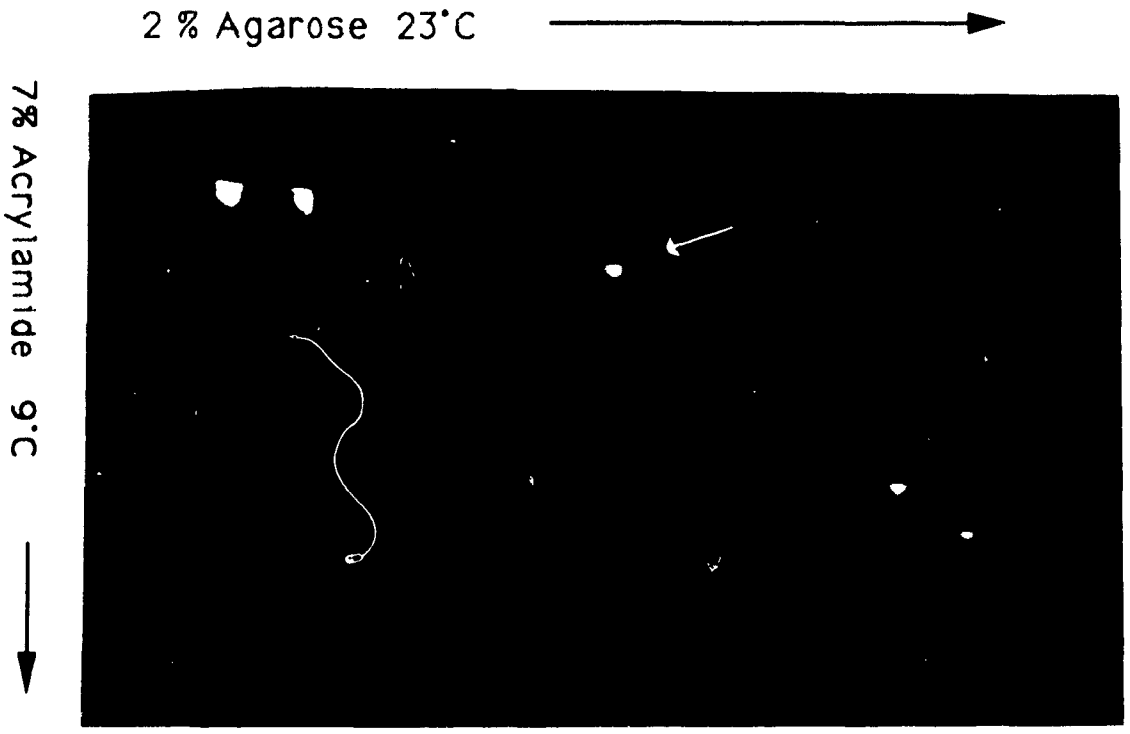


Fig. 10. Parallel gel bent DNA assay. Restriction digests of hors76 were run on parallel 5% acrylamide gels at 9°C and 23°C and were visualized with ethidium bromide staining. B, BglI; N, NaeI; A, AvaII; M, 123 bp ladder marker,

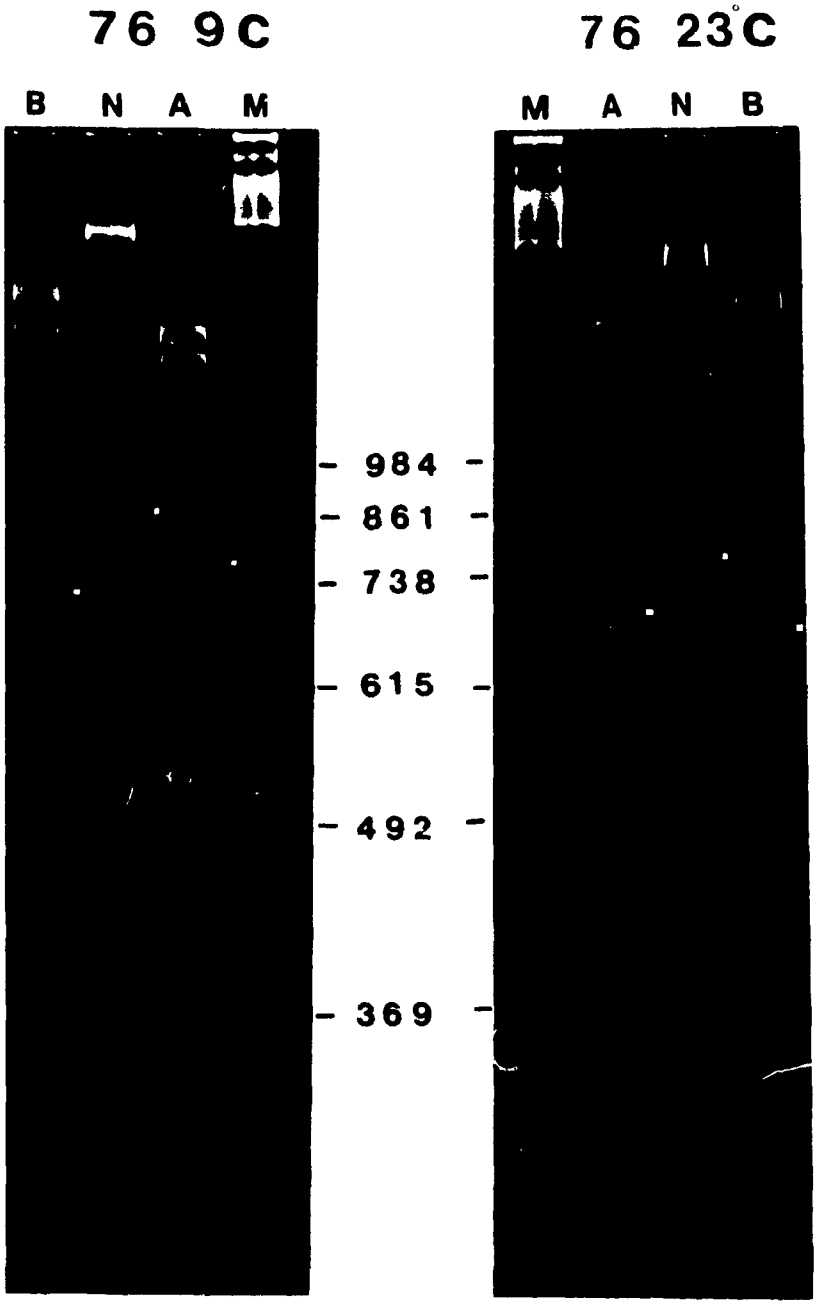

76 9C
B N A M
76 23°C
M A N B
984
861
738
615
492
369

Table 1. Bent DNA fragments of hors76 and hors110 The relative mobilities of restriction digest fragments are summarized for the insert bands of hors76 and hors110 under normal and cold conditions. Fragment sizes are reported in bp as determined by comparison with the 123 bp ladder marker.

Table 1 Bent DNA fragments of hors76 and hors110

Digest		Relative Mobility (bp) 9°C	23°C	9°C/23°C
Hors76	BglI	720	680	1.06
	NaeI	840	785	1.07
	AvaII	760	700	1.08
Hors110	DdeI	910	870	1.05
	TaqI	425	390	1.09
	AvaII	800	760	1.05

III. Unwinding assays

One of the first steps involved in the initiation of DNA replication involves the melting out of the double helix. This local unwinding, often at A/T rich regions, gives the replicative machinery access to the DNA template. Specific DNA sequences of low unwinding energy have been found in association with origins of replication from bacteria and yeast and it has been suggested that they promote the initiation of DNA replication by destabilizing the local DNA duplex. The panel of hors plasmids were surveyed for DNA unwinding elements (DUE) by the method of Umek and Kowalski (1987). Any unwound region of DNA would be sensitive to nicking by the single strand specific enzyme mung bean nuclease. Once a single nick is introduced into a supercoiled molecule, it becomes relaxed thereby relieving the torsional strain that favours local unwinding. Therefore this assay will detect the sequence element in a plasmid that has the least energy requirement for unwinding. Supercoiled plasmids were nicked with mung bean nuclease and the nicking site was mapped by either Sal1 or EcoR1 restriction digestion, end labelling and denaturing electrophoresis (fig. 11). The fragments produced by this assay include a strand of DNA equal to the full plasmid length as well as a pair of strands adding up to full length that are defined by the nicking and Sal1 or EcoR1 sites. Mung bean hypersensitive sites can be mapped by measuring the size of the fragments produced and comparing them to the site of linearization. The vector pBR322 has a native mung bean nuclease hypersensitive site within the Amp^r

terminator. Any DUE within the hors clone would have to be more prone to unwind than this sequence in order to be detected. In the case of each clone tested, the strongest mung bean nuclease hypersensitive site mapped to a location of 1.8 kb base pairs away from the SalI site and 1.1 kb away from the EcoRI site. This maps to the native hypersensitive site in pBR322. In clone 76 we could detect a faint set of bands at 0.9 and 4.0 kbp from the SalI digest that suggest a weak DUE within the hors clone.

A summary of the properties of the ten hors clones is outlined in table 2. Half of the ten hors clones were capable of autonomous replication as measured by the BrdU substitution assay. Two of these clones contained bent DNA sequences and one of these contained a weak DNA unwinding element. The presence of bent DNA or DNA unwinding elements was not required for replication activity nor was there an increased likelihood of replication potential with larger insert sizes.

Fig. 11. DNA unwinding element assay. Clones were surveyed for DNA unwinding elements by the procedure of Umek and Kowalski (1987) as described in the materials and methods. Plasmids were linearized with either EcoRI (A-E) or SalI (F-J). Hors76, A and F; hors106, B and G; hors112, C and H; hors129, D and I; hors133, E and J; lambda hindIII and phiX174 HaeIII markers, M.


A
B
C
D
E
M
F
G
H
I
J
23
9.4
6.6
4.4
2.3
2.0
1.3
1.1
0.9
0.6

Table 2. Summary of hors properties. The properties of the ten hors clones are tabulated with respect to: their size in bp, ability to replicate autonomously (BUDR), the presence of bent DNA (BENT), and DNA unwinding elements (UNWOUND).

TABLE 2 **Properties of HORS**

HORS	SIZE(b.p.)	BUDR	BENT	UNWOUND
67	288	-	-	-
74	3000	-	-	-
76	426	+	+	+
98	452	+	-	-
103	4000	-	-	-
106	852	+	-	-
110	553	-	+	-
112	571	+	-	-
129	395	+	-	-
133	227	-	-	-

IV. DNA sequence analysis

In addition to the functional and structural analysis of the hors DNA, it was of interest to determine the primary DNA sequences of the hors clones. This type of investigation was required to determine what if any sequence characteristics were shared among the hors clones. It was also of interest to determine if the sequences contained elements that have been identified in other origins of DNA replication. Five of the clones were sequenced by the dideoxy sequenase reaction and are reported in figure 12. On first inspection, the clones did not reveal any remarkable features with the exception that they were slightly G/C rich. They had an average G/C content of 52% compared to the estimated 40% G/C content of total human DNA (Kariya et al., 1987). There were no obvious stretches of unusual base sequences or repeats.

As discussed above, there are a number of sequence elements and structures that have been associated with various origins of DNA replication. One of the features often associated with origins of replication is the presence of repeats of short sequence motifs. Origins from E. coli, SV40 and yeast all contain repeats of short sequences that are required for replicative function, presumably through interaction with their initiator proteins (ie. dnaA protein, T antigen). The five sequences were scanned for a variety of these elements by the DNASIS and GCG DNA sequence analysis programs. The clones were searched for the presence of direct repeats, true

Figure 12. DNA sequence of the hors clones 98, 106, 112, 129, and 133.

```
HORS 98          452 BP
   1 CTGGGCCGAT CTGTGGAAGC CAGAGTACAA AGGCAGCCTG CTGTTGACCG
  51 ACGATGCCCG TGAAGTGTTC CAGATGGCGC TGCGTAACGT GGGCTACTCC
 101 GGTAACACCA CCGATCCGAA AGAGATTGAA GCTGCATATA ACGAGCTGAA
 151 AAACTGATGC CAAACGTGGC AGCCTTTAAC TCCGATAACC CGGCGAACCC
 201 GTACATGGAA GGCGAAGTTA ACCTCGGCAT GATCTGGAAC GGTTCTGCTT
 251 TCGTTGCACG CCAGCGGGTA CGCCAATTGA CGTGGTGTGG CCGAAAGAAG
 301 GCGGCATTTT CTGGATGGAC AGCCTGGCGA TCCCGCAAAT GCCAAAAACA
 351 AAGAAGGCGC GCTGAAATTG ATCAACTTCC TGTTGCGCCC GGATGTGGCA
 401 AAACAGGTTG CTGAAACTAT CGGTTATCCA ACGCCAAACT TGCGGTGTAA
 451 GC

HORS 106         852 BP
   1 CTGCCACATT GCGTCCAGGC CCAGGCATCC CGGCATAACC GGATCGCCAA
  51 TAAAGTGGCA TCCGAAGAAC CACAGATCCG GATTGATATC CAGTTCTGCT
 101 TCAACATACC CTTTGTCGAA GTTACCACCC GTTTCGGTCA TTTTGACCAC
 151 ACGGTCCATC ATCAGCATGT TCGGTGCTGG CAATTGCGGG CCTTTAGCGC
 201 AAACAGTTCA CCGCGACCAG AGGCAAGAAG GTCTTCTTTT GTATAGGATT
 251 CGCGTTTATC TACCATGTTC TCTGTAAGCC TTATTTTATT GAAGCACGCA
 301 GGATAGCTAA CACGTGTACG CTGAACAAGT CCGATCAGTT CGGAATAAAC
 351 CAGTTCAGCC AACGTAATGG CCATGGAAAA CGGTGACGTC CTTCCTGTTG
 401 CGATCGTTGC GCGATACGTT CCTGGATGGT TTGCATCAGC GTCGTTTGCC
 451 TTCGCCATCC CACACCAGAT TAATAATAAC GGCAGTGCGT CAGTCACATC
 501 GTCTACTGCC CAAGATGGTG AATTTGCCTT CTTCTACAGC TTTCACCAGT
 551 TCACTGTGAA GACTTAAATG GCGAACGTTA GCTGTGGGAT AATGACACCT
 601 TGTTTCCCGG TTAACTCACG TTGCTGGCAA ATAGCAAAGA AGCCTTCGAT
 651 TTTCTCATTT AAACCACCGA CCGGCTGGGC GCGACCGAAC TGATCGACTG
 701 AACCTGTGAT AGCGATACTC TGATTCACCG GCACATCGGC GAGGGCGCTT
 751 ATCAGGGCGC AGAGTTCAGC CATCGAGGCA CTATCTCCAT CAACTTCACT
 801 GTATGACTGC TCAAATGTCA GCGATGCTGA GAAGGGGATC TGTTGCTCAA
 851 GC

HORS 112         571 BP
   1 CTGGCAGAGG CCATCCGCGC AAGAATGGAT GGCCATTTCG ATAAAGTTTT
  51 AAAACAGGCG GTTTAAACCG TTTAACGCAG CTACCCGATA GCTTCCGCCA
 101 TCGTCGGGTA GTTAAAGGTG GTGTTGACGA AGTACTCAAT AGTGTTGCCG
 151 CCACCTTTCT GTTCCATAAT CGCCTGACCG ATATGAATAA TTTCGGCAGC
 201 GCGCTCGCCA AAGCAGTGAA TACCCAGAAT CTCTTTTGTT TCCCGATGGA
 251 ACAAAAATTT TCAGCGTGCC ACGTTCATGC CGACGATTTG TGCGCGTGCC
 301 AGATGTTTAA CTGGGCGCGG CCCACTTCAT ATGGCACTTT CATTGCGGTC
 351 AGCTGCTGTT CGGTTTTGCC CACAGAGCTG ATTTCCGGGA TGGTGTAAAT
 401 ACCGGTAGGG ATATCTTCAA TCAGATGTGC GGTCGTCGCC TTTTACCAGG
 451 CCTGCGCGGC AATGCGCCCC TGGTCATAGG CCGCCGACGC CAGGCTCGGA
 501 TAACCAATCA CGTCGCCCAC CGCGTAAACG TGTGGCTGTG CGGTCTGATA
 551 CATGCTGTTG ACCTTCAGCG A

HORS 129         395 BP
   1 CTGTCTGTTC CGGAAAAAAC ACCGGAAAAA AAGATTATGG TGAGAAAAAA
  51 GGCCAGATAC CCTTTAATGC CCACTTTTTC TGTCGCCAGC TCCCCGGCTA
 101 CAGCATCCCC TTGTTGTGTC ATAACAATCC CCATTGTTTT TTTGTGTTTG
 151 TGTGTTGCTT GAGCCAGGAA AATTCCTGGA ACTGAAAAAC CTTATCAGCC
 201 ATGTGGGTAT GGGGAAAATG CAAGTATGTT TCTGGCTATG CGTTTTTTGC
 251 ATAGCGGAAA AATTACGCTG TGTGGAAAAG GCTCTGCGCA GGATTGAGTT
 301 GCAGAAAATT CATAAACCAA GTGACAATGA CATCAGTAAA TAAAAACAGA
 351 AAATCCAGGG ACGAAAACGA GATGGCAAAG TACAACGAAA AAGAG

HORS 133         227 BP
   1 CTGGCACCTG CGGAGAACCC GCTGCAGAAG AGCGGCCACG GCTCGCTGCT
  51 GCAGACGCAT ACGGGTGAAT GGTACATGGC CTACCTCACC AGCCGCCCGC
 101 TGCGCCTGCC CGGCGTGCGC TGCTGGCCTC CGCGGACGCG GCTACTGCCC
 151 GCTGGGGCGC GAGACGGGCA TCGCCGCATT GAATGGCGCG ACGGCTGGCC
 201 GTACGTGGAA GGCGGCAAGC AGCGCAG
```

palindromes, potential cruciform structures, and A-T rich regions. It was of interest to tabulate the occurrence of these sequence elements and determine what, if any, correlation there was between these elements and the structural and functional properties of the hors clones.

The hors clones were scanned for the presence of direct repeats by the algorithm of the DNASIS program. The search criteria were such that a direct repeat of 8 bp or more constituted a match. The repeats had to be perfect and the pair of elements could be located anywhere within the sequence. The size and number of such direct repeats are recorded in table 3. To a large extent, the clones were not internally repetitive in nature. Although each clone contained direct repeats of at least 8 bp, the largest directly repetitive sequence detected was only 10 bp in length. For example hors98 contained 3 pairs of repeats of 8 bp in length and one pair of 10 bp. Where repeats were found, they were scattered throughout the clones.

True palindromes consist of a contiguous stretch of nucleotides that contains a central axis of symmetry. The popular linguistic palindrome "A man, a plan, a canal, Panama" illustrates this property well. When considered in terms of a nucleotide sequence, it is important to note that the symmetrical order of the nucleotides is limited to the same strand of the double helix. That is, although there is an inherent repeat of a stretch of adjacent nucleotides, in one arm it runs 5'-3', and in the other, it runs 3'-5'. It is also

Table 3. Hors sequence elements. The five hors sequences were scanned for a variety of sequence elements by the DNASIS and GCG programs. The length and number of exact direct repeats greater than 7 bp are tabulated for each clone (ie. hors98 had 3 sets of exact direct repeats 8 bp in length and one set of repeats 10 bp in length). True palindromes of at least 80% match over 6 or more bp are indicated (ie. hors98 had two palindromes, each half having a match of 5 out of 6 bp). The three largest potential cruciforms are indicated by their relative score from the STEMLOOP algorithm and the position of the first nucleotide in the stem. The total number of potential cruciforms scoring over 15 are also indicated for each clone. A/T rich regions of at least 70% A/T over 20 or more bp are indicated as the number of A/T bp over the total region length.

Table 3. Hors Sequence Analysis

	98	106	112	129	133
Size (b.p.)	452	852	571	395	227
Direct Repeats	8 (3), 10	8 (3), 9 (3), 10	8, 9	8 (2), 10	8 (4)
True Palindromes	5/6 (2)	5/6 (7)	5/6 (3)	10/12, 9/11, 6/7	5/6
Pot. Cruciforms	24 (nt. 265)	23 (nt. 73)	33 (nt. 432)	32 (nt. 243)	37 (nt. 92)
	18 (nt. 356)	21 (nt. 600)	32 (nt. 114)	24 (nt. 78)	32 (nt. 123)
	17 (nt. 301)	20 (nt. 583)	26 (nt. 6)	19 (nt. 268)	27 (nt. 89)
total # > 15	11	17	28	10	30
A/T Rich	23/35, 14/20	26/36, 14/20	17/20, 14/20	27/37, 22/31	-
		15/20	20/27	14/21, 14/20	
				17/23, 38/50	

important to note that there are no intervening bases separating the two arms of a true palindrome. The hors clones were scanned by the DNASIS program for the occurrence of true palindromes by an algorithm that required a match of over 80% in palindrome arms of at least 6 bases. True palindromes fitting these criteria were found in all of the five clones (table 3). Only hors129 contained large palindromes, the remainder of the clones contained true palindromes having arms of only 6 bases.

Potential cruciforms are unusual elements of DNA that can form a unique structural configuration based on an inherent duplication of a short nucleotide sequence. They combine the repetitive nature of direct repeats (with the order of nucleotides following the same polarity) with the central axis of symmetry of true palindromes. It is this peculiar arrangement of the bases that permits such inverted repeats to form cruciforms. The hors clones were scanned for potential cruciforms by the STEMLOOP algorithm of the GCG program. This program calculates a score for each inverted repeat depending on the quality and length of base pairing within the stem of the cruciform. A bonding value is given for the paired bases in the stem (G:C=3, A:T=2, G:T=1). The cut-off score was set at a level of 16 or greater. Each clone contained a number of cruciforms and the nucleotide position of the first base of the three largest cruciforms in each sequence are indicated in table 3. The total number of potential cruciforms scoring over 15 is also indicated for each clone. These values represent a relative grading of the energetics of the

different cruciform structures. Each of the clones contains cruciforms that score over 20, with the largest score being 37. The significance of these values may be addressed by comparing these scores with the values of other potential cruciforms that have been detected in other systems. When subjected to the same scoring algorithm, the ColE1 cruciform ranks as a score of 33, and the PhiX174 inverted repeat scores at 29. Both of these cruciforms have been shown to form transiently in supercoiled DNA (Lilley, 1980).

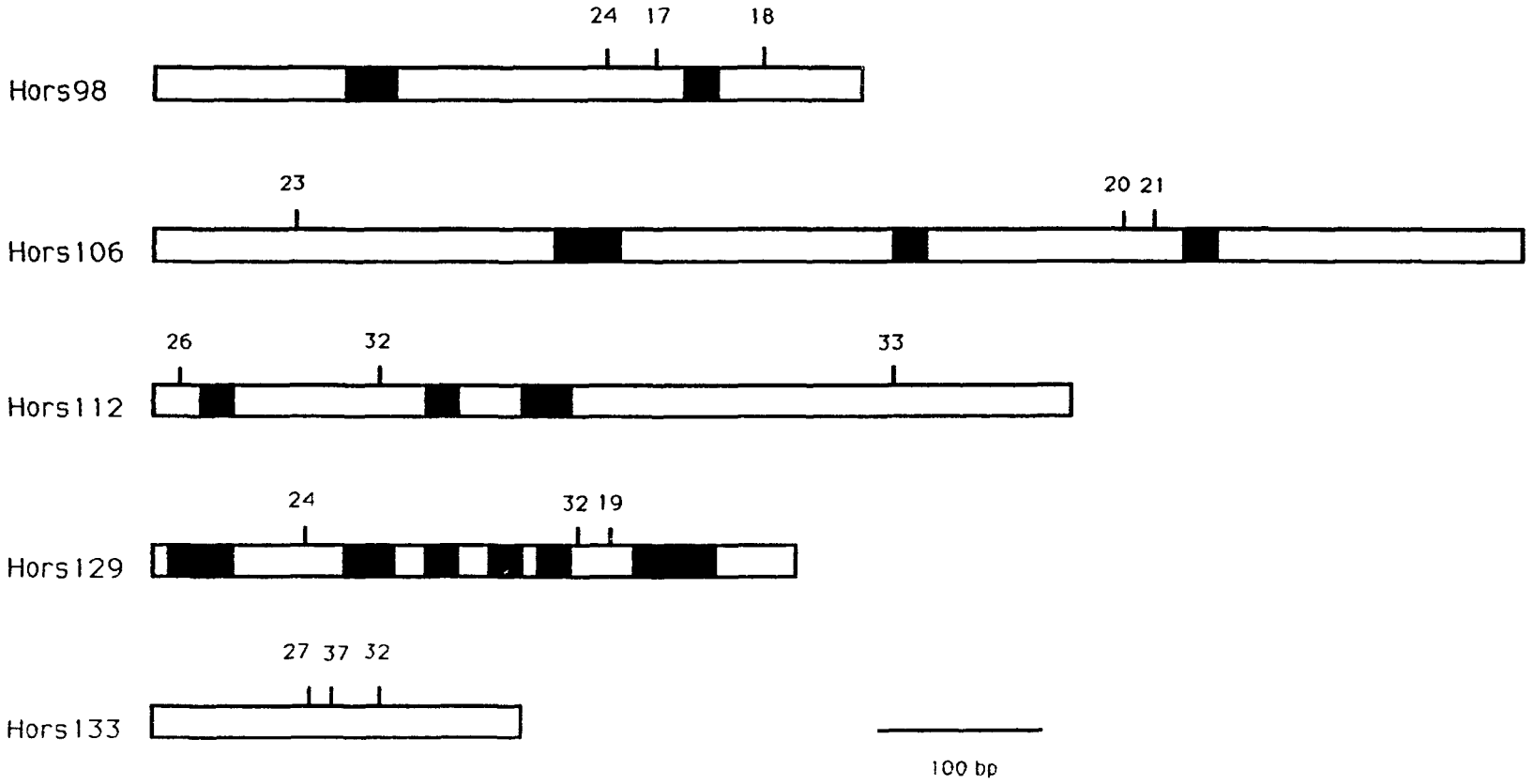
Perhaps one of the most common sequence elements associated with origins of DNA replication is A/T rich regions. The energetics of the base pairing of A/T rich sequences makes them the preferred sequences at which to unwind the double helix. As it has been shown with the OriC origin from E. coli, short A/T rich sequences can function in the initiation of DNA replication (Bramhill and Kornberg, 1988). The five hors clones were scanned for A/T rich sequences by the DNASIS program. Regions of at least 70% A/T over 20 or more bp were identified and recorded in table 3. While none of the clones were A/T rich, four of the clones did contain more than one A/T rich region that met these criteria. It is interesting to note that hors133 did not have any A/T rich sequences and this clone was the only one of the five clones sequenced that failed to replicate autonomously.

The presence of large potential cruciforms and A/T rich sequences was analyzed further. It is known that cruciform formation

is affected by the nature of the flanking DNA regions and in particular, A/T rich flanking sequences have been shown to favour cruciform formation (Sullivan and Lilley, 1986). The position of these two sequence elements was mapped on each of the clones (fig. 13). Both the potential cruciform structures and the A/T rich sequences (where they were found) were distributed throughout the clones. The A/T rich sequences are represented as black boxes, and the centre of the largest cruciforms is indicated by a line. In all of the clones except hors133, we could detect large potential cruciform structures lying close to an A/T rich region. In fact, out of the twelve largest potential cruciforms in the replicating clones, ten of them were located within 50 bp of an A/T rich sequence.

There are a number of short DNA sequence motifs that have been associated with a range of biological properties and activities. A large and growing number of transcription factors have been identified that interact specifically with defined nucleotide sequences. These DNA binding proteins regulate the complex process of transcription through their interaction with their recognition sequences of DNA and the transcriptional enzymatic machinery. The obvious similarity between this process and that of DNA replication has led some researchers to investigate a possible link between transcription and replication (reviewed in DePamphilis, 1988). In addition to the transcriptional sequence motifs, there are short nucleotide sequences that have been associated with other biological properties such as ARS replication in yeast (ie. ARS consensus) and scaffold attachement sequences. The five hors clones were scanned for a variety of DNA

Fig. 13 Map of hors clones. The relative positions and sizes of potential cruciform structures and A/T rich regions are mapped for the five hors clones. A/T rich regions are represented as black boxes and the potential cruciforms are indicated by a line positioned at the centre of the loop below the relative score as calculated by the STEMLOOP algorithm.


24 17 18
Hors98
23 20 21
Hors106
26 32 33
Hors112
24 32 19
Hors129
27 37 32
Hors133
100 bp

consensus sequences associated with a range of factors and functions. The consensus sequences analyzed are indicated in table 4 and include those sequences that were previously searched for within the monkey ors clones. Each clone was scanned for an exact match to the different consensus sequences, and where a match occurred, the first nucleotide of the sequence is indicated under each clone heading (when the consensus was found in the reverse orientation, the nucleotide number is followed by an r). The probability of finding a match to a given consensus sequence at random within the clones was considered. The chances of finding a short consensus sequence within a clone of a given length is more likely than that of finding a larger sequence. The expected number of perfect matches to a given sequence was calculated for the total number of bp (2497x2) scanned. For example, if one screened a random sequence of 4994 bp of DNA, one would expect to find the TGGCA sequence 4.8 times and the GCTGCAG sequence 0.3times. A standard deviation was calculated from these data as reported previously (Rao et al., 1990). For simplicity, expected values of less than 0.1 are indicated by a 0. The presence of the consensus sequences were tabulated and it was determined if the hors clones were enriched for any particular sequences. Several of the consensus sequences were found within the clones at an occurrence of greater than one standard deviation from the expected frequency. The NF-1 related sequence (Leegwater et al., 1986) was found 14 times within the hors clones and it would be expected to occur by chance only 4.8 times. The AP2-A consensus sequence (Imagawa et al., 1987) occurs 7 times and is expected only 2.4 times. The iron response

Table 4. Hors consensus sequence search. Both strands of the hors clones were searched for the presence of a variety of short consensus sequence elements previously identified in other systems. The sequence code was: R = G or A; K = G or T; S = G or C; W = A or T; M = A or C; Y = T or C; D = G or A or T; V = G or A or C; B = G or T or C; and N = any nt. Where a match was found to the consensus sequence, the first nucleotide of the matching sequence within the hors clone is indicated (an r indicates that the match is on the complementary strand). The mean number of times the given consensus sequence was expected to occur in the hors clones by chance was calculated by multiplying the total number of comparisons by the probability of finding a match in a single comparison (Exp.). Expected = (4994 - 5 x (length of consensus -1)) x (match in a single comparison) = nP, where n is the total number of comparisons and P is the probability of a match in a single comparison. The probability of obtaining a match (P) was calculated based on a random distribution of nucleotides of A/T content of 50%. The expected distribution's standard deviation (S.D.) is approximated from the binomial frequency distribution S.D. = nP(1 - P). For simplicity, where the expected number of matches was found to be lower than 0.1, it has been tabulated as a 0.

Table 4.

Hors Consensus Sequence Search

Element	Sequence	Reference	98	106	112	129	133	Exp	SD
Eryf	WGATAR	Evans et al. (1988)		256r, 707, 749r, 781r		192r		4 8	2 2
Nf-E1	MYWATCWY	Wall et al. (1988)		780	501	31r		2.4	1 6
(i)	TGACKHW	Mermod et al. (1988)		490r	125			3 6	1 9
AP2. A	VSSCCMNSSS	Imagawa et al. (1987)		452, 672r	510, 511		33, 150r, 192r	2 4	1.5
AP2. B	CCCCAGGC	Mitchell et al. (1987)						0	
AP4 (i)	HCAGCKSBDG	Mermod et al. (1988)			348			0 5	0 7
(s)	YCAGCTCYGG							0	
B1 intron (s)	AAARRGGAARY	Gaison and Housman (1988)						0	
(i)	AARRGGAA							0 3	0 5
ß-promoter. A	TCYWAAGYAAKTGMS	Antonion et al. (1988)						0	
ß-promoter. B	KCYWAMGYMAKTGHS							0	
related seq.	TGG(N)$_7$DCCAR	Gronostajski (1986)						0	
NF1	TGGCA	Leegwater et al. (1986)	158r, 167, 340r, 396	2r, 56, 178, 625	2, 266r, 296r, 331	373	2	4 8	2 2
	GCCAAT	Rosenfeld et al. (1987)	272	46				1 2	1 1
	CCAAT		273	47	501			4 8	2 2
GCS and related	CACCC	Dierks et al. (1983)		126			63r	4 8	2.2
seq.	CYHCACCC	Mitchell et al. (1987)						0 45	0 7
	RCYHCACCC	Myers et al. (1986)						0 23	0 5
NF-kB	GGACTTTCC	Sen and Baltimore (1986)						0	
BGP1	GGGGGGG	Lewis et al. (1988)						0.30	0.54
Pit-1 and related	AWWTATTCAT	Nelson et al. (1988)			183r			0	
seq.	WHWHAWBCAH			658	183r			3 0	1.7
Estrogen rec.	WKGTMANNNTGWSYK	Evans (1988)		136r				0	
GRE-related	AGAWCAGW	Yamamoto (1985)						0.30	0.55
sequences	BKNHYHNNNWGTYCT	Evans (1988)		83				1.0	1.0
HAP1 (cycl1)	GCNNTTANNGCNNTTAC	Pfeifer et al. (1987)						0	
(cycl7)	GGGNTTTANGGACNANGA							0	
HSE-related seq.	CNGAANTTCNG	Pelham (1985)						0	
Interferon response	RGTTTCNNWTYY	Levy et al. (1988)						0	
Iron response	CNNNNNCAGNG	Casey et al. (1988)	360r	212, 319r	208	282r	20r, 44r, 97r	4 8	2.2
Metal response	TGCRCHC	Stuart et al. (1985)	255r	754r	460			1 8	1.3
cAMP element	ACGTCA	Sassone-Corsi (1988)	278r	384r				1.2	1.1
	KCGTCA		487					2 4	1.6
Octamer	ATTTGCAT	Singh et al. (1986)						0	
Poly(A)	AATAAA	Fitzgerald and Shenk (1981)		49, 285r, 344		339		1 2	1.1
ß-globin enhancer Region A, chick p-globin upstream footprint	GCTGCAG	Rao et al. (1990)					21, 48	0 3	0 5
Region B, chick p-globin upstream footprint	CATTTG	Rao et al (1990)	336	812r				1.2	1.1
Chick alpha-globin promoter footprint	CCGTGCGG	Kemper et al. (1987)						0	
ARS (S. cer.)	WTTTATRTTTW	Palzkill and Newton (1988)						0	
SAR A(s)	AATAAAYAAA	Gasser and Laemmli (1986)						0	
(i)	WADAWAYAWW					133r		0.45	0 7
SAR T(s)	TTWTWTTWTT			281				0	
(i)	TWWTDTTWWW			281, 469r				0.45	0.7
Topo 2	GTNWAYATTNATNNR	Sander and Hsieh (1985)						0	
(chicken)	RNYNNCNNGYNGKTNYNY	Spitzner and Muller (1988)						1.2	1.1
SP1. A	KRGGCGKR	Briggs et al. (1986)						1.2	1.1
SP1. B	GGCGGG							1.2	1.1

element (Casey et al., 1988), the poly(A) element (Fitzgerald and Shenk, 1981), the chick globin upstream region A (Rao et al, 1990), and the low stringency scaffold attachement region (Gasser and Laemmli, 1986) are also slightly enriched for in these clones.

The hors sequences where searched for potential open reading frames and for homology to each other, the previously isolated monkey ors sequences and the sequences in the GENBANK database. The hors sequences did not show considerable homology to each other or to the monkey ors sequences. Only short stretches of homology could be found between any pair of sequences. Unlike some of the monkey ors sequences, the hors clones did not show homology to repetitive sequences (O family, Alu etc.). They did however contain open reading frames. The entire 227 bp of hors133 consisted of an open reading frame (ORF) corresponding to a potential 75 amino acids. Hors129 had an ORF of 60 amino acids that was open on one end. Hors98 had an internal open reading frame of 85 amino acids. Hors106 had an ORF of 95 amino acids that was open on one end of the clone and hors112 had an ORF of 65 amino acids. The relevance of these observations and any putative transcripts arising from these sequences on the ability of these clones to initiate DNA replication is not known.

CHAPTER FIVE

DNA-PROTEIN INTERACTIONS AT HORS

I. DNA binding activity to HORS fragment

It is clear that the initiation of DNA replication will require the interaction of the trans-acting proteins of the replication machinery with the cis-acting DNA sequences that constitute an origin of DNA replication. One approach towards characterizing the mechanisms of the initiation of DNA replication is to identify protein/DNA interactions between the hors sequences and proteins isolated from human nuclei. For this reason, I chose to look for DNA/protein interactions between HeLa cell extracts and the hors sequences by the DNA bandshift assay (Garner and Revzin, 1981).

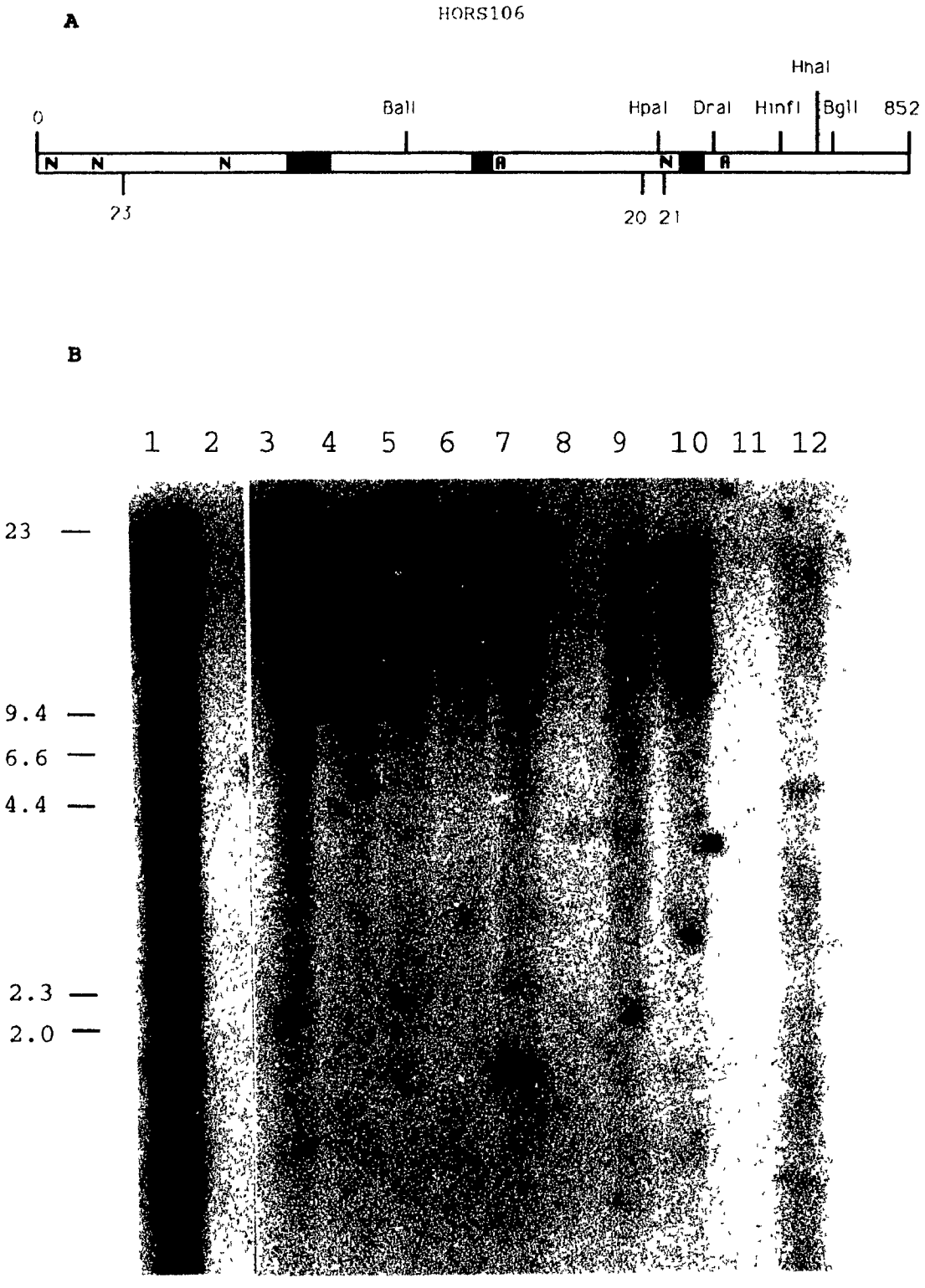
Several sub-fragments from a range of hors clones were assayed for DNA/protein interactions with crude Hela cell nuclear extracts. A range of bandshift activities was observed among these varied fragments, the strongest occurring with a HpaI fragment from the 3' end of hors106. This DNA fragment was found to bind to a factor (or factors) from HeLa nuclear extracts as determined by the bandshift assay. Hors106 was able to replicate autonomously in HeLa cells and the 3' end of this clone contained the sequences of two of the enriched consensus elements (NF-1 related, and AP2-A). This region of hors106 also contained two of the three largest potential cruciforms in the clone and they were in close association with an A/T rich sequence (fig. 14 A). For these reasons, it seemed reasonable to focus in on this region of this autonomously replicating sequence in search of proteins that may play a role in the initiation of DNA

replication.

Quantitative dot blot analysis of human DNA indicated that hors106 was a low copy number (<20) sequence (data not shown). Recently, mapping protocols have been developed to determine whether or not a given sequence functions as an origin of replication in-vivo in its chromosomal context (Burhans et al., 1990). These techniques require that the sequence of interest be of a single copy, and require the cloning of adjacent DNA fragments. It was of interest therefore, to determine if the hors106 was a single copy sequence and if so, on what chromosome was it located. A southern blot from a panel of somatic cell hybrids was probed with the entire hors106 sequence. Hybridization with the control lane of total human DNA and with a number of the somatic cell hybrids revealed a range of bands indicating that the hors106 sequence was present in some 10-12 copies (fig. 14 B). Due to this number of bands, it was not possible to map these bands to particular chromosomes.

One of the first experiments in the investigation of the DNA/protein interactions that occur at hors106, was to characterize the nature of the bandshift complex. An end-labelled HpaI fragment of hors 106 was incubated with HeLa nuclear extracts and then treated with proteinase K, RNase A, or phenol (fig. 15 A). The bandshift reaction was sensitive to phenol and proteinase K, but not RNase A suggesting that the factors involved in the bandshift reaction were in fact protein. The specificity of this DNA/protein interaction was

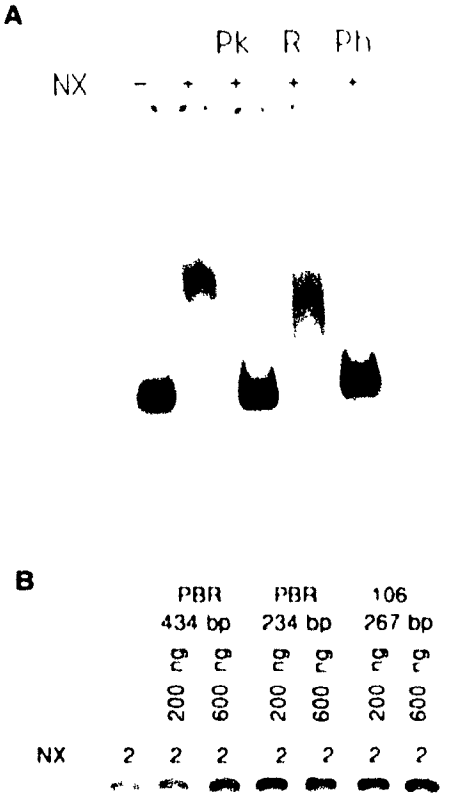
Fig. 14. Characteristics of hors106. (A) map of 852 bp hors106: relevant restriction sites are listed above the clone, A/T rich sequences are designated as black boxes, NF-1 related and AP2-A consensus sequences are indicated as N and A respectively. The three largest potential palindromes are mapped below the clone with their calculated scores. (B) Southern blot of somatic cell human/hamster hybrid DNA digested with EcorI and probed with hors106. 1, total human control DNA; 2, hamster control DNA; 3-12 hybrids 8017a, 8014, 8005d, 7905, 7633, 7631, 7614, 4543, 4501, 4106 (see appendix).


A
HORS106
Hhal
0
BalI
HpaI
DraI
HinfI
BglI
852
N
N
N
A
N
A
23
20
21
B
1 2 3 4 5 6 7 8 9 10 11 12
23
9.4
6.6
4.4
2.3
2.0

tested by a competition assay (Singh et al , 1986; Carthew et al , 1985) (fig 15 B). The bandshift activity in the crude nuclear extracts could be competed by the addition of cold excess fragments of the hors106 DNA, but not by the addition of excess similar and larger sized DNA fragments from pBR322.

Having identified an activity in HeLa cell nuclei that binds to this sequence, it was of interest to partially purify it and further characterize the nature of this DNA/protein interaction. The sequence information analysis suggested that this region of hors106 may contain determinants that play a role in the initiation of DNA replication. As yet, the identity of the proteins that serve to regulate the initiation of eukaryotic DNA replication are unknown. It is possible that proteins that bind to this region of hors106, even if they themselves are not initiator proteins, may play a role in the molecular mechanisms that serve to initiate and regulate DNA replication.

Fig. 15 DNA bandshift assay of DNA/protein interactions at hors106. An end-labelled HpaI fragment of hors106 was incubated with crude nuclear extract (N.X.) from HeLa cells, run on a 4% acrylamide gel, dried and autoradiographed. (A) The DNA fragment was incubated without extract (-), or with 3 ul of crude extract (+) for 20 minutes. The reactions were then treated with either, 10 ug proteinase K (Pk), 10 ug RNase A (R), or 5 ul phenol (Ph) at 37°C for 30 minutes. (B) Bandshift reactions were carried out in the presence of excess cold competitor DNA (200 or 600 ng) from 434 bp or 234 fragments of pBR322 or the 267 bp HpaI fragment of hors106.


A
Pk R Ph
NX - + + + +
B
PBR 434 bp
PBR 234 bp
106 267 bp
ng
200
600
NX 2 2 2 2 2 2 2

II. Purification of DNA binding activity

The DNA binding activity from the crude nuclear HeLa cell extracts was purified for further analysis. The crude extract was subjected to anion exchange chromatography on a DEAE column as described in the materials and methods. The column was washed and eluted with a linear 0.1-1.5 M KCl gradient starting at fraction 33. Every second fraction was analyzed for protein content and subjected to a bandshift assay with the hors106 DraI fragment. The bulk of the DNA binding activity eluted in the first several fractions of the DEAE column (fig 16 A). Fractions 3-11 also contained the bulk of the protein that came off in the flow through (fig 16 B) and were pooled, concentrated and then applied to a PC-11 column. Once again the column was washed and bound protein was eluted with a 0.1-1.5 M KCl gradient starting at fraction 30. Both protein concentration and DNA binding activity were monitored for every second fraction. DNA binding activity was seen to peak at around fraction 65 (fig 17 A). The smearing of the bandshift complex was likely due to the high salt content in these fractions. The proteins contained in these fractions were retained on the column through the initial wash and were released only during the elution with the salt gradient at an estimated concentration of 0.8 M KCl (fig 17 B). The PC-11 (phosphocellulose) column has properties similar to DNA and as such it is useful for the purification of DNA binding proteins. It can be seen that the bandshift activity was limited to only a few of the fractions from the eluate of this column.

Fig 16. DEAE column purification. HeLa crude nuclear extracts were applied to a DEAE column, washed and eluted with a 0.1-1.5 M KCl gradient starting at fraction 33 as described in materials and methods. (A) DNA bandshift assays were performed with 10 ul from every second fraction. (B) The optical density of every second fraction was used to estimate protein concentration.

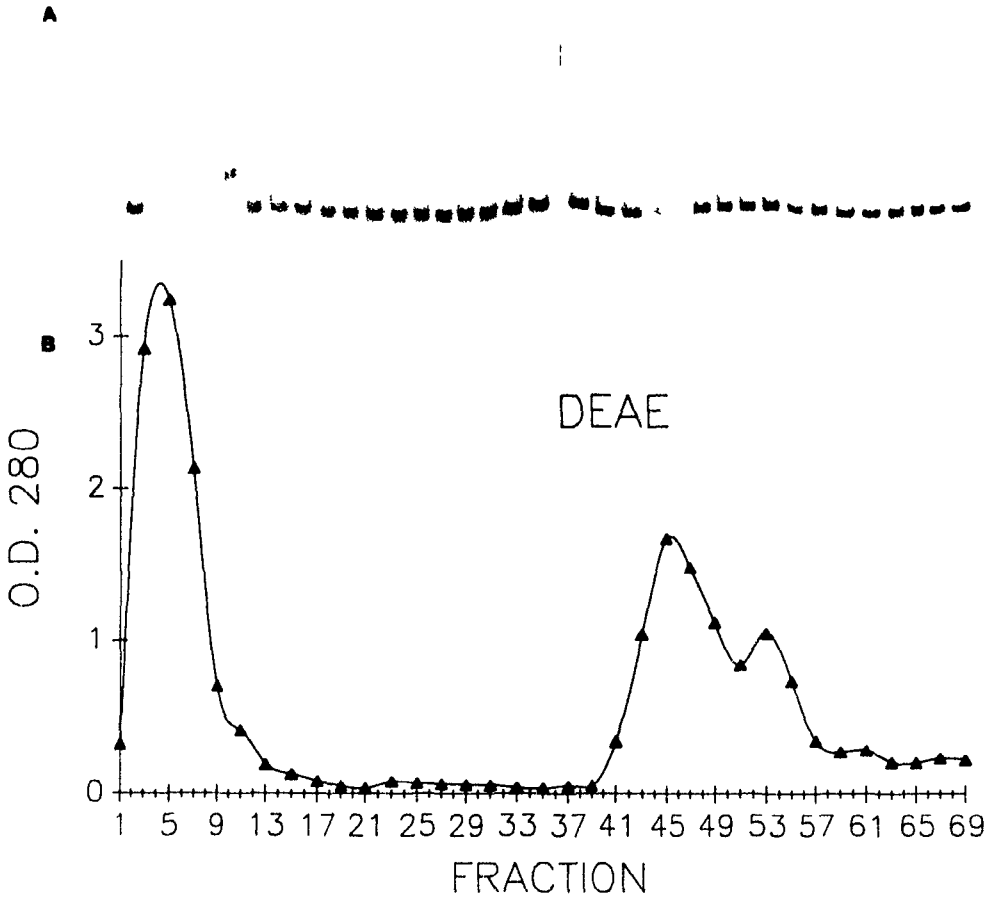
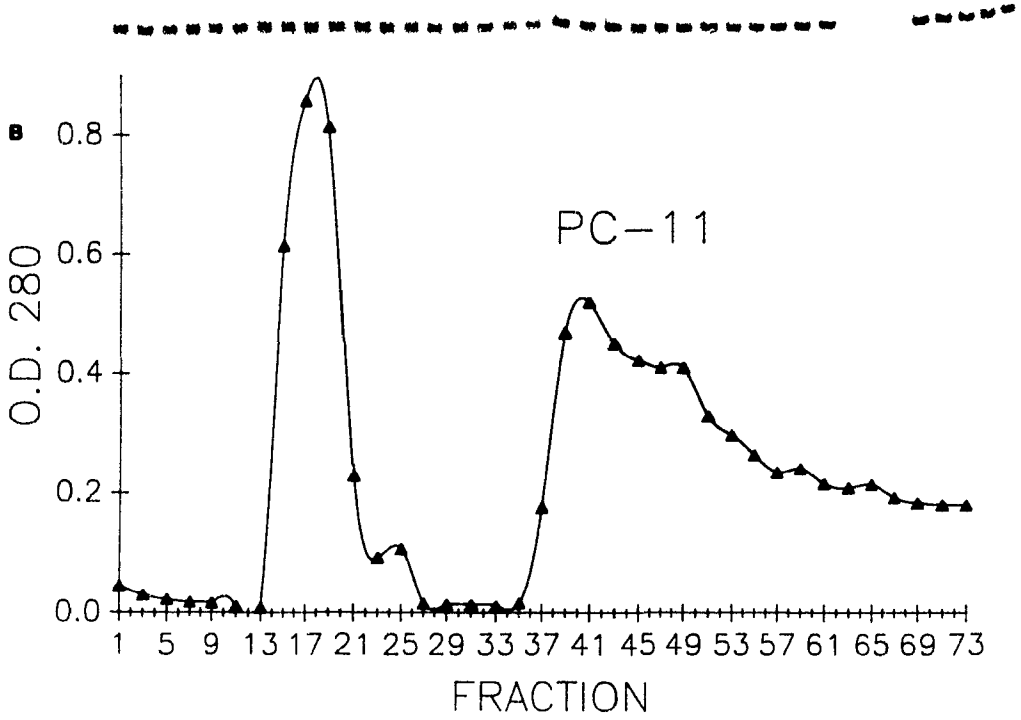

A
B
DEAE
O.D. 280
3
2
1
0
1 5 9 13 17 21 25 29 33 37 41 45 49 53 57 61 65 69
FRACTION

Fig 17. PC-11 column purification. Protein from fractions 3-11 of the DEAE column flow through were concentrated and applied to a PC-11 column as described in materials and methods. The column was washed and bound proteins were eluted with a 0.1-1.5 M KCl gradient starting at fraction 30. (A) DNA bandshift reactions were performed with 10 ul from every second fraction. (B) the optical density of every second fraction was used to estimate protein concentration.

A


B
0.8
0.6
0.4
0.2
0.0
O.D. 280
PC–11
1 5 9 13 17 21 25 29 33 37 41 45 49 53 57 61 65 69 73
FRACTION

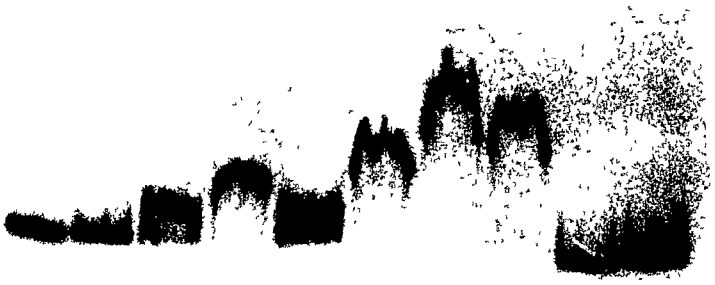
Having narrowed down the bandshift activity to several fractions coming from the gradient of the PC-11 column, it was of interest make a more detailed analysis of these fractions. Fractions 59-67 were chosen for a determination of the protein constituents within each fraction that correlated with the bandshift activity (fig 18). DNA binding profiles using 15 ul of each fraction were run parallel to a silver stained protein gel (SDS PAGE) (Oakley et al., 1980) as a characterization of the DNA binding proteins within these fractions. It was noticed that the bandshift activity of the fractions correlated to a doublet of protein bands at around 21-23 kDa. The waviness of the bandshift reactions was again likely due to the high KCl content of these fractions. These gels were deliberately stained with silver, so as to reveal the full spectrum of minor proteins within these fractions. Although the silver stained gel indicates that there were other proteins within these peak fractions, the 21-23 kDa doublet constituted the majority of the protein and was consistant with the bandshift activity. Under Coomassie staining, these minor proteins were barely detected.

The correlation of bandshift activity with the protein doublet is not a proof that these proteins were in fact involved in the DNA binding. To address this point, a southwestern analysis was performed on these partially purified fractions (fig. 19). The purified fractions from the PC-11 column were electrophoresed on a 12% SDS gel alongside a molecular weight marker and electroblotted to a nitrocellulose filter. Coomasie staining was done both before and

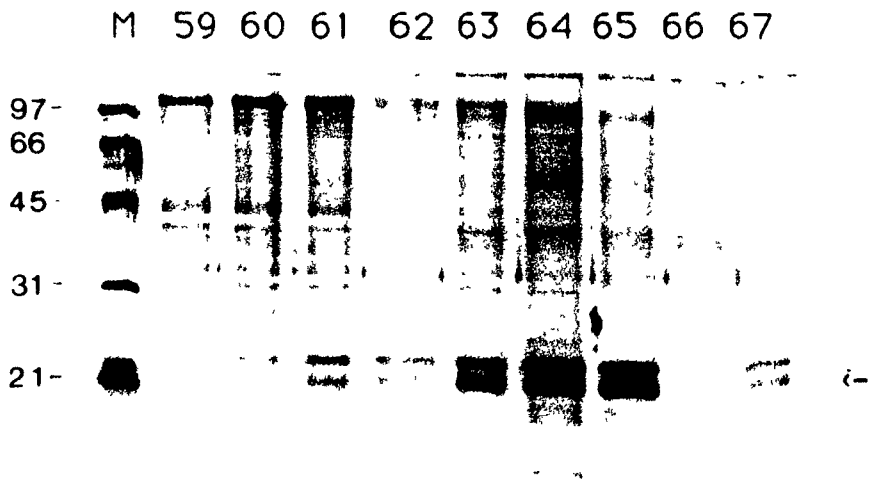
Figure 18. DNA bandshift and SDS PAGE analysis. 15 ul aliquots from fractions 59-67 off the PC-11 column were (A) subjected to bandshift analysis and (B) run on a 4/10% SDS PAGE and visualised by silver staining.

A

B - S

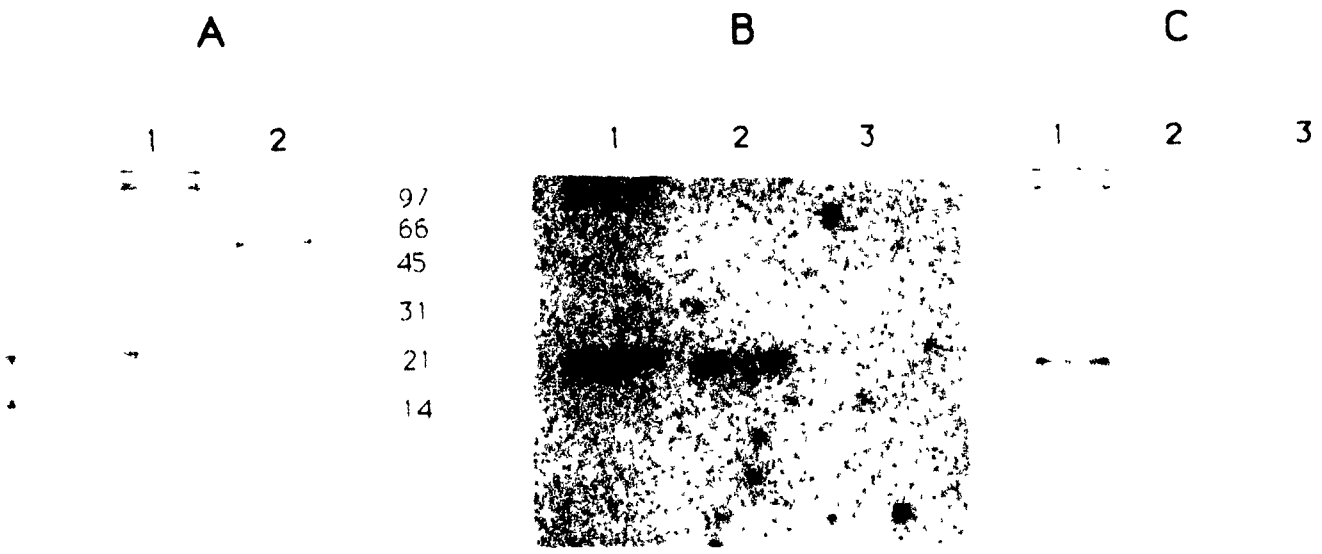


B



SDS

Fig 19. Southwestern analysis. Samples from the pooled fractions 63-65 of the PC-11 column were run on a 4/12.5% acrylamide gel and either (A) stained with commasie blue (1; 15 ul extract, 2; low molecular weight marker) (B) electroblotted onto nitrocellulose, probed with end-labelled HinfI fragment of hors106 and autoradiographed (1; 15 ul fractions 63-65, 2; 25 ul fractions 63-65, 3; low molecular weight marker), or (C) stained with commassie blue after electroblotting (1; 15 ul fraction 63-65, 2; 25 ul fraction 63-65, 3; low molecular weight marker).


A
B
C
1
2
1
2
3
1
2
3
97
66
45
31
21
14

after electroblotting, so as to follow the proteins in these fractions closely. The filter was treated as described in the materials and methods and probed with a radiolabelled HinfI fragment of hors106. In fig 19 A, one can see the protein doublet at around 21-kDa and a background of minor polypeptides. Two samples of this partially purified fraction were run on a parallel gel, electroblotted, and hybridized (fig 19 B). After electroblotting, the gel was stained again to check that transfer had occurred (fig 19 C). One can see that the DNA hybridized only to the 21-23 kDa region and not to the minor background proteins or the markers.

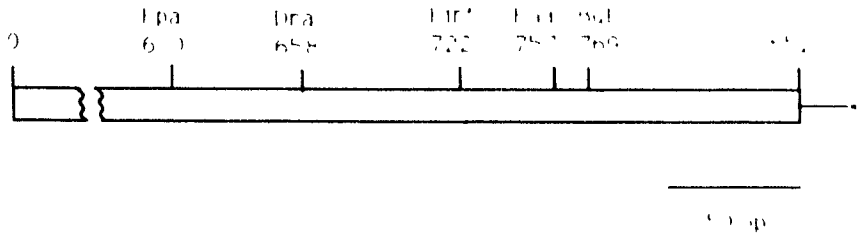
III. Properties of DNA protein interaction

It was of interest to better define the region of hors 106 that was recognized by the DNA binding fraction. The HpaI fragment was digested further with a series of restriction enzymes and these smaller fragments were subjected to bandshift analysis with the purified fractions from the PC-11 column (fig 20). One can observe a significant drop in bandshift activity when the fragment is digested at the BglI site. Lowered bandshift activity was found for the fragment digested with HhaI. Thus the region of the clone required for bandshift activity mapped between the HinfI and BglI sites.

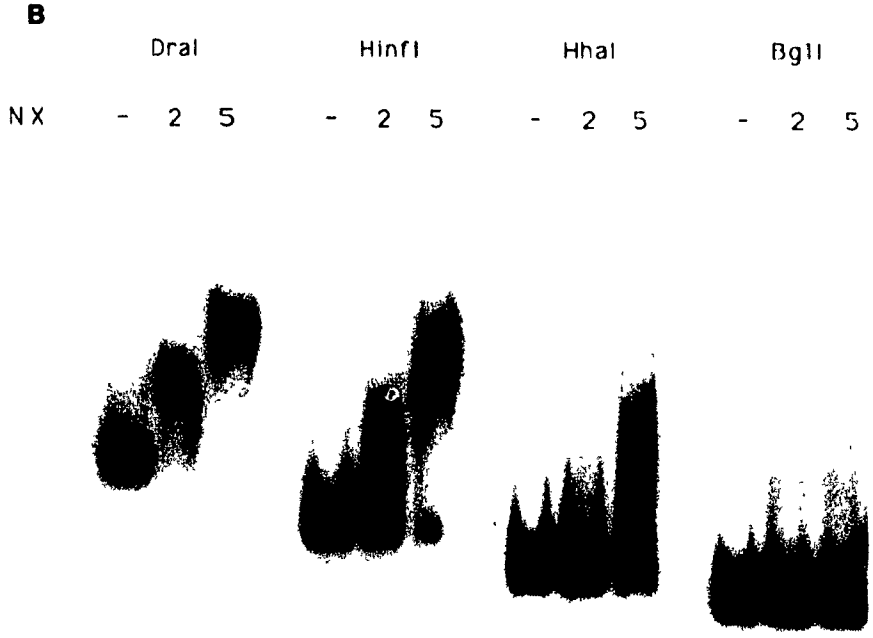
Figure 20. Restriction bandshift assays. End-labelled fragments of hors106 were created by the polymerase chain reaction and digests of this clone were subjected to bandshift analysis with the fraction 63-65 of the PC-11 column. (A) restriction map of the 5' end of hors106. (B) the bandshift assay with 0, 2 or 5 ul of fraction 63-65 on the fragments of hors106.

A

HORS106



B



IV. In-vitro replication profiles

Recently an in-vitro replication system has been developed that can follow the replication of supercoiled plasmids using HeLa cell extracts (Pearson et al., 1991). This system combines a reaction mixture of nucleotides, crude nuclear and cytoplasmic extracts with 250 ng of supercoiled plasmid DNA. The replication of the plasmid molecules are followed by incorporation of labelled nucleotides followed by agarose gel electrophoresis. Although this system is similar to that defined for SV40 (Li and Kelly, 1984), it is considerably less efficient as it does not have the benefits of a purified initiator protein (T antigen). It was of interest to determine if the partially purified DNA binding activity from the PC-11 column had an effect on the replication profile of hors106 plasmid in this system (fig 21). 250 ng of supercoiled hors106 plasmid was incubated with the standard replication mixture, or with the addition of 5 or 15 ul of the partially purified PC-11 DNA binding fraction. After 20 and 60 minutes of incubation, one could detect labelled replication products corresponding in size to both linear and nicked plasmid molecules, as well as a continous smear of replication intermediates. The addition of the partially purified PC-11 fraction did not alter the incorporation profile considerably. There was a slight decrease in the reaction supplemented with 15 ul of the partially purified fraction, but this change was not dramatic.

Fig. 21 In-vitro replication assay. 250 ng of hors106 was incubated in the standard replication mixture with or without the addition of the partially purified hors106 DNA binding fraction from the PC-11 column. (A) incubation for 20 minutes, (B) incubation for 60 minutes. (1: standard reaction, 2: with 5 ul PC-11 fraction, 3: with 15 ul PC-11 fraction).


A
B
1 2 3
1 2 3

CHAPTER SIX

DISCUSSION

The initiation of DNA replication is a complex and highly ordered biological process. The progression of eukaryotic cells into S phase constitutes a major transition in the cycle and as such it is subjected to a variety of controls (Murray, 1987). The complexity of replicative machinery that catalyses the synthesis of DNA and the fidelity with which DNA is replicated are staggering. The DNA sequences at which DNA replication is initiated are important in that they may suggest to us some of the mechanisms that serve to regulate and orchestrate DNA replication. One important subset of DNA replication origins are those sequences that are initiated at the beginning of S phase. Activation of these origins signals a turning point in the eukaryotic cell cycle that ends ultimately in cell division. Unravelling the molecular mechanisms that control this process will help our understanding of issues related to cell proliferation.

I. The Synchronization of Cells with Aphidicolin and the Extrusion Method for Isolating Nascent DNA

The ability to synchronize cell populations offers a great advantage to researchers of the cell cycle (reviewed in Ashihara and Baserga, 1979). There are a number of questions concerning cell division that cannot be easily addressed in logarithmically growing cultures. Aphidicolin is a powerful drug that is particularly suited for the synchronization of cells to the G1/S border (Pedrali-Noy et al., 1980; Kaufmann et al., 1985). This is due to the fact that

aphidicolin is a reversible inhibitor of the replicative DNA polymerases alpha and delta (Hammond et al., 1987; Marracino et al., 1987), and has a lower toxicity than 5-fluorodeoxyuridine, or hydroxyurea (D'Anna et al., 1985). Cells first arrested in G1 phase by serum starvation can be pushed to the G1/S border by the addition of serum in combination with aphidicolin. These cells will have all of the growth factors and mitogens required for entry into S phase (Phillips and Cristofalo, 1981), except that the replication machinery is stalled by aphidicolin. These arrested cells can be easily maintained for several hours in this arrested state and, when washed free of aphidicolin, quickly commence DNA replication. Indeed, DNA primase activity is not inhibited by aphidicolin (Dinter-Gotleib and Kaufmann, 1982), and as such, activated origin sequences may be primed and ready for elongation by a competent DNA polymerase. This observation is supported by the fact that DNA replication proceeds immediately following the removal of aphidicolin in both CV-1 and WI-38 cells (figures 1 and 4). This property makes aphidicolin synchronization, as opposed to synchronization by other methods such as double thymidine block or hydroxyurea arrest, particularly useful for the analysis of events at the onset of S phase. It is interesting to note that not all cells can be effectively synchronized by this method. We were unable to synchronize HeLa cells using the same technique (data not shown). The decreased requirement for growth factors exhibited by HeLa cells may prevent them from accumulating in G1 phase when deprived of serum. If cells are exposed to aphidicolin during S phase, interruption of DNA replication leads to cell death.

CV-1 cells synchronized by aphidicolin arrest may be released and followed as they progress through S phase (figure 2). This technique allows for a detailed analysis of the progression of DNA replication. Measurements of the rate of DNA synthesis by microfluorometric analysis indicate that there is a biphasic pattern of DNA replication (figure 3). The rate of DNA synthesis is highest at the beginning and end of S phase. In contrast, if the rate of DNA replication is estimated by the incorporation of ^{3}H-thymidine, one observes a gradual increase in the rate of DNA replication, peaking in mid S phase and falling towards the end of S. One problem concerning this type of analysis is that the amount of labelled nucleotide incorporated into DNA over a given pulse period changes with the size of the free nucleotide pools within the cell (Richter and Hand, 1979). Regardless of the method used to measure replication, it is clear that there is little detectable DNA replication in the arrested cell populations, and that replication commences immediately following release from aphidicolin block.

The unique properties of this synchronization protocol make it a useful technique with which to isolate early replicating DNA. If one could isolate short fragments of DNA that are replicated at the onset of S phase, one would expect that they should contain origins of DNA replication. A crucial property of replication bubbles is that they are unstable at elevated temperatures. The short nascent DNA can be extruded from these replication bubbles as double stranded DNA. This extrusion process likely occurs by a mechanism involving DNA

breathing, renaturation of the nascent strands, and branch migration (Zannis-Hadjopoulos et al., 1981). This DNA should contain an origin of DNA replication at or near its centre and can be purified away from the high molecular weight parental DNA by virtue of its size and incorporated Hg-dUTP during an the in-vitro elongation step. Theoretically, the only DNA that would have incorporated the tagged nucleotides would be the short nascent fragments.

When the post extrusion DNA sample was layered on a neutral sucrose gradient and sedimented, it was noticed that the labelled material was in fact of low molecular weight (figure 6 A). Part of this material should contain Okazaki fragments that were not yet ligated to the growing DNA strand. As an indication of the size range of the labelled DNA, fractions 19-23 were run on an acrylamide gel and visualised by autoradiography. These fractions contained DNA of up to several hundred base pairs in length, and were discarded. In a parallel sucrose gradient, supercoiled pBR322 DNA (18s) was recovered in fraction 5. Therefore, fractions 11-17 were estimated to contain DNA ranging from several hundred to several thousand bp in length and they were pooled and concentrated for further purification.

It is likely that during the manipulation of the DNA sample in the extrusion process that some of the parental DNA would have been sheared down to the size of the nascent DNA. This would result in the origin-enriched DNA fraction containing random non-origin sequences. To compensate for this potential problem, the nascent DNA was further

purified by chromatography on a homocysteine column (figure 6 B). Mercurated DNA will bind selectively to a homocysteine column and can be eluted off with buffers containing beta-mercaptoethanol. The labelled nascent DNA was loaded on and bound to the homocysteine column. It was found that over 80% of the labelled DNA applied to the column could be eluted off of the column with buffer containing beta-mercaptoethanol. Fractions 36-54 were pooled and concentrated for cloning into the NruI site of pBR322. The resulting library of recombinant plasmids contained inserts ranging in size from several hundred to several thousand bp in length.

II. Properties of the Origin-Enriched Sequences

The BrdU substitution assay (Frappier and Zannis-Hadjopoulos, 1987) was used to determine if the hors clones could replicate autonomously as plasmids in HeLa cells. Detection of semiconservative replication is assayed by determining the extent of incorporation of a heavy nucleotide analog (BrdU). An alternative method to measure DNA replication in plasmids is to use restriction endonucleases (ie. DpnI, MboI) to assess the methylation status of the plasmid DNA before and after replication. Although the DpnI resistance assay has been used to monitor DNA replication successfully in viral systems, it is not always capable of detecting the low levels of replication in ors clones (Landry and Zannis-Hadjopoulos, 1991). Out of the ten clones tested in this study, five showed evidence of autonomous replication as determined by the incorporation of BrdU into the plasmid DNA. Both

HH and HL plasmid molecules were recovered after 24 hours of replication indicating that some plasmids underwent more than one round of replication. It is possible that the HH DNA came from either reinitiation of replication in a population of plasmids or that some cells passed through two S phases within the 24 hour labelling period.

This percentage of replication-positive clones (50%) is similar to the frequency observed in the monkey ors clones (similarly isolated by the extrusion process), and is considerably higher than that observed for a random selection of similar sized human DNA (G. Price personal communication). This high rate of replication-positive clones is likely due to the manner in which these sequences were isolated. The extrusion method is designed to isolate sequences arising from functional origins of replication that are activated at the onset of S phase. In contrast, Krysan et al., (1989) screened a library of random human DNA fragments and isolated a heterogenous collection of human genomic sequences which, when cloned into an Epstein-Barr virus vector replacing the viral region of dyad symmetry replicate auotonomously in human cells. The average size of the human DNA insert in that study was found to be 12 kb and replication potential was found to correlate strongly with the size of the human DNA insert (Heinzel et al., 1991). The mechanism by which these clones promote replication is unclear. Without evidence of origin activity leading to semi-conservative replication it is unclear whether these sequences represent chromosomal replication origins or

fortuitously complement the activity of the dyad symmetry region of the Epstein-Barr virus origin. It was observed that although the two-dimensional gel analysis suggests that replication initiates at many sites within the clone, sub-clones of the larger sequences have a lowered replication efficiency (Krysan and Calos, 1991).

The screening of autonomously replicating sequences has been used to identify DNA replication origin sequences in yeast (Broach et al., 1982), although interestingly, not all sequences that function as ARS elements function as origins of replication in their chromosomal context (Reynolds et al., 1989). The ability to replicate autonomously as plasmids does not prove that a sequence is in fact a bona fide chromosomal origin. Methods are currently being developed to test for origin activity of sequences on chromosomal DNA (Burhans et al., 1990). These techniques work best on single copy sequences and require the cloning of unique sequences flanking the potential origin of interest.

Bent DNA is one of several non-B form DNA structures that has been identified in recent years (Anderson, 1986). Considerable research has been carried out to characterize both the occurrence and function of this unusual DNA structure. Bent DNA has been associated with origins of replication from bacteriophage lambda (Zahn and Blattner, 1987), SV40 (Gerard and Gluzman, 1986), and yeast (Snyder et al., 1986). Several proteins have been purified that enhance DNA bending in these systems, but they have yet to be shown to have a role

in replication from these origins (Baur and Knippers, 1988). Bent DNA can be identified by its altered electrophoretic mobility, when run under cold conditions. The two dimensional gel technique (first dimension 2% agarose, second dimension 7% acrylamide) offers a visual assay for the identification of bent DNA fragments (figure 9). The parallel acrylamide gel assay (one at room temperature, one run under cold conditions) permits a quantification of the extent of altered fragment mobility (figure 10). DNA fragments that migrate anomalously on cold acrylamide gels were found in 2 hors clones. One of these clones had ARS activity (hors76), one did not (hors110). Bent DNA was therefore neither present in the majority of the hors clones, nor was its occurrence correlated with the ability of the hors clones to replicate autonomously. Bent DNA was previously identified within an early labelled fragment of the DHFR amplicon, but more recent data suggests that it may lie several hundred bp away from the origin of bidirectional replication associated with this replicon (Burhans et al., 1990).

Another DNA structure that has been associated with origins of replication is the DNA unwinding element. As a prerequisite to the initiation of DNA replication, there must be an unwinding of the double helix, permitting access of the replicative machinery to the DNA templates. DNA unwinding elements are A/T rich sequences of DNA that preferentially unwind under supercoiled conditions. This local unwinding can be detected with the single strand specific enzyme mung bean nuclease. One of the 10 clones tested, hors76 contained a DNA

unwinding element comparable (albeit somewhat weaker) in energy to that of the A/T rich native pBR322 mung bean nuclease hypersensitive site. The unwinding element in pBR322 is located in the Amp^r terminator region (Umek and Kowalski, 1988) and although it does not play a role in the replication of pBR322, it can serve as a control for the relative tendency to unwind. That only one DUE was found in the hors clones does not proscribe that these clones may possess DNA unwinding elements, which through interaction with cellular factors, may unwind in-vivo. It merely implies that each plasmid, as naked supercoiled DNA, is most likely to unwind within the vector sequence. This result is not surprising, as the pBR322 site is considered to be a strong unwinding element (D. Kowalski personal communication). In E. coli oriC, opening up of the 13 mer A/T rich repeats is facilitated by the binding of the dnaA protein to its recognition sites (Bramhill and Kornberg, 1988). Like DNA bending, DNA unwinding can be affected by local protein/DNA interactions.

Five of the hors clones were sequenced and each was analysed for the presence of a variety of short sequence elements and potential structures. Short repetitive elements are required in a number of prokaryotic, viral, and eukaryotic origins. E. coli OriC contains repeats of the dnaA binding sequence (Bramhill and Kornberg, 1988), SV40 Ori has multiple binding sites for T antigen (Shortle et al., 1979), and yeast origins have at least two ARS consensus sequences (Palzkill and Newlon, 1988). Disruption of the number or spacing of these closely linked short repeats is often enough to disrupt the

ability of an origin to initiatee replication. It was of interest to characterize the internally repetitive nature of the hors clones, as this analysis may point towards unusual sequence motifs that may function in the initiation of DNA replication. Although each clone contained at least one direct repeat of at least 8 bp of length, as a whole, the hors clones were not internally repetitive to any large degree. When direct repeat sequences were found, they were scattered throughout the clones, and no striking array of repeats was found.

We also searched for inverted repeats consistent with the formation of cruciforms. Although cruciform formation in relaxed DNA is energetically unfavourable, it has been shown that cruciforms may form transiently in negatively supercoiled DNA (Panayotatos and Wells, 1981). Cruciforms have been postulated to play important roles in a number of replication systems (Weller et al., 1985). The hors clones were screened for potential cruciform structures by the STEMLOOP algorithm of the GCG sequence analysis program. This program grades potential cruciforms on the basis of the base pairing in the stem. Base pairs of G:C, A:T, and G:T are given values of 3, 2, and 1, respectively. The summation of the base pairing values in the stem is used to calculate the overall score of the potential cruciform. Each hors clone was found to contain a number of potential cruciforms having a score greater than 15, and several contained potential cruciforms scoring 25 or greater. The significance of these values can be addressed with comparison to the values of cruciform structures that have been shown to function in other systems. Both the ColE1 and

PhiX174 cruciforms have been shown to form in-vitro in negatively supercoiled DNA, and when graded by the same algorithm, score at 33 and 29 respectively.

It has been reported that one of the factors involved in the kinetics of cruciform formation is the composition of the flanking DNA sequences. A/T rich regions favour cruciform formation and can have effects at distances of at least 100 bp (Lilley, 1985). We found many hairpin structures within the hors clones and each of the replication positive clones has a large potential cruciform structure in close association with an A/T rich region. If cruciforms play a role in the initiation of DNA replication, then the presence of these closely associated A/T rich sequences may be an important determinant in promoting cruciform extrusion.

The search for previously reported DNA consensus sequences within the hors clones revealed that the hors clones were enriched for the NF-1 related, AP2-A, iron response and poly(A) sequence elements. The calculations of the expected frequency of occurrence of these consensus sequences was estimated for the total length of bp searched. These calculations were based on a G/C content of 50% and a random distribution of bases, and as such are therefore only estimates. In fact the G/C content of the hors clones was 52% and the distribution of nucleotides in human DNA is not random.

Nuclear factor I is a DNA binding protein that is required for

the replication of adenovirus (Nagata et al., 1983). Purification of NF-1 from HeLa cells has revealed a number of proteins ranging in size from 52 to 66 kDa (Rosenfeld and Kelly, 1986). Members of this family of polypeptides have been shown to stimulate the transcription of both viral and cellular genes through interactions at promoter sequences (Gil et al., 1988; Henninghausen and Fleckstein, 1986). NF-1 and activator protein-2 (AP2) bind in a mutually exclusive way to overlapping promoter sequences and trans-activate the human growth hormone gene (Courtois et al., 1990). Although these sequences were found to be enriched for in the hors clones, it is not known if this relates to the issue of DNA replication. When the monkey ors clones were subjected to the same sort of analysis, they were also slightly enriched for the AP2-A and NF-1 related sequences as well as others including the control sequence described by Dierks et al (1983) and the scaffold attachement regions (Gasser and Laemmli, 1986).

III. Proteins that interact with hors106

If replication in human cells follows a similar process to that of E. coli OriC and the SV40 ori, then one would expect that the initiation of DNA replication will involve the interaction of an initiator protein with the DNA sequence that serves as an origin of replication. Indeed, all models of replication involve DNA/protein interactions. The question that remains to be answered is "What proteins are interacting with what sequences, and what happens when they do?". Unfortunately, when it comes to DNA replication in human

cells, this is still a very big question, as little is known about either the protein or DNA constituents of this process. Apart from the enzymatic machinery that catalyses DNA synthesis, little is known concerning the proteins that initiate and regulate DNA replication. Answering this question will require the purification and characterization of proteins that interact with origins of DNA replication.

I have approached this issue by partially purifying an activity from HeLa cells that binds to a region of hors106. While it is likely that the identification of a eukaryotic initiator protein will require extensive characterization of the DNA/protein interactions at an origin sequence, it is possible that proteins binding close to an origin of replication may have an affect on the initiation of replication. The binding of proteins to a DNA sequence (ie. transcription factors) has the potential to alter the local DNA conformation and alter the highly regulated processes of transcription and replication. Indeed, the SV40 origin is flanked by transcriptional control sequences and these are required for efficient initiation of replication (DePamphilis, 1988). The identification of the DNA/protein interactions at an origin sequence is a necessary step in characterizing the molecular mechanisms that execute the initiation of DNA replication.

Having decided to investigate DNA/protein interactions within the hors clones, it was necessary to choose a fragment within one of

the hors clones to study in detail. Several fragments from hors106 and hors112 were assayed for bandshift activity with crude nuclear extracts from HeLa cells. It was noticed that among these fragments, the strongest bandshift activity was with a fragment of DNA from the 3′ end of hors106. This region of hors106 contained the DNA sequences and structures that were identified as possible determinants in origin function. The 3′ end of hors106 contained several large potential cruciforms in association with an A/T rich sequence as well as the AP2-A and NF-1 related consensus sequences. This fragment therefore seemed to be a reasonable choice for the characterization of DNA/protein interactions.

One obvious step in the characterization of this DNA/protein interaction is the purification of the bandshift activity. Crude nuclear extracts from HeLa cells were fractionated on a DEAE and then a PC-11 column and assayed for bandshift activity. It was observed that the bandshift activity was found in the flow through fractions of the DEAE column, but was retained on the PC-11 column under low salt conditions. The fractions with measurable bandshift activity were eluted off of the PC-11 column at a KCl concentration of about 0.6 M. Protein analysis by silver staining of SDS PAGE of the partially purified fractions containing bandshift activity revealed that the bulk of the protein consisted of a doublet at around 21-23 kDa. The confirmation that it was not any of the minor background proteins that bound the DNA was achieved by southwestern analysis. The partially purified fraction was subjected to SDS PAGE, electroblotted, and

probed with radiolabelled hors106 HinfI fragment. Unfortunately, due to the size of the electroblotting apparatus, the closeness of the protein doublet, and the diffusion and imprecision inherent in blotting and autoradiography, it was not possible to determine wether one or both bands of the protein doublet bound the DNA probe. It would be interesting to investigate whether or not the two bands constituted related proteins (one may be a slight modification of the other ie. phosphorylation etc.), and if so, how this correlates with DNA binding activity.

The region of the hors106 sequence that is required for DNA binding was investigated in order to further characterize the nature of this DNA/protein interaction. DNase I footprinting analysis of the DNA/protein interaction was tried but was unsuccessful. Although hints of DNase I protection were observed in a region some 100 bp from the 3 end of the clone, due to either technical difficulties or the nature of the DNA/protein interaction or both, no clearly protected regions were identified. It was, however, possible to delineate the sequences necessary for DNA binding by restriction analysis of the hors106 clone. It was found that the 40 bp sequence 5' to the BglI restriction site was required for bandshift activity.

An <u>in-vitro</u> replication system has been developed in our lab to follow the replication of DNA in autonomously replicating plasmids (Pearson et al., 1991). This system was used to test the activity of the partially purified DNA binding fraction from the PC-11 column. It

was of interest to determine whether or not the addition of this protein fraction would influence the replication of hors106 in-vitro. It was observed that incorporation of the labelled nucleotides increased from 20 minutes to 60 minutes under the standard reaction conditions. Addition of 5 or 15 ul of the partially purified DNA binding activity did not greatly alter the replication profile of this clone. There was a slight decrease in the amount of incorporation in the 20 and 60 minute reactions supplemented with 15 ul of partially purified fraction.

There are a number of reasons why there might not have been a considerable difference between the standard and supplemented reactions. First of all, the DNA binding activity simply may not play a role in the initiation of replication of this plasmid Although this binding activity maps to a region of the hors clone that contains elements enriched for in the hors clones in general, that does not guarantee that binding in this region affects the initiation of DNA replication. Even if this region of hors106 was important in the regulation of initiation of DNA replication in-vivo, it is not necessary that this relationship be maintained when the hors106 sequence is used to replicate as a plasmid. The binding protein may function with a structural component in the nucleus not maintained in the in-vitro reaction. Alternatively, this binding may be important in the replication of hors106, but under the conditions of the in-vitro reaction, addition of the purified fraction may have no effect. The protein may not be in limiting concentrations in the standard

reaction, or the fraction may not have retained activity.

The nature of these systems makes the identification of relevant replication proteins difficult. Diffley and Stillman (1988) have reported the purification of a protein that binds to a yeast origin of replication, but have not ascribed it any functional activity. Interestingly, this protein has a MWr of 21 kDa. Dailey et al, (1990) have used a similar strategy to investigate the protein/DNA interactions that occur at the DHFR origin region. They took a fragment of the DHFR replicon that contained elements characteristic of origins of replication (this region also contained the consensus binding sequences to AP1 and OTF1/NFIII) and characterized the proteins that bound to this region. One of the proteins that they identified enhanced DNA bending within this sequence (Caddle et al., 1990), and was associated with an ATP dependant helicase (Dailey et al., 1990). Although these results are suggestive of a role for these proteins in replication, lack of a reliable assay has precluded this demonstration.

IV Conclusions

The hors clones do not show considerable sequence homology to each other or to the previously identified monkey ors sequences. They were however enriched for the NF-1 related, AP2-A and iron response consensus sequences. One half of these clones were able to replicate autonomously as plasmids in HeLa cells as judged by the BrdU

substitution assay, but this property was not linked to either the presence of bent DNA or DNA unwinding elements. The replication positive clones all contained large potential cruciform structures within 50 bp of an A-T rich region. Dissecting the mechanisms that regulate the initiation of DNA replication will require the characterization of proteins interacting with origin sequences. DNA bandshift assays identified a DNA-binding activity from HeLa cells that interacts in a specific manner with a sub-fragment of hors106. Further purification steps are under way that should enable us to characterize the protein and assess its possible role in replication.

CLAIMS FOR ORIGINAL RESEARCH

The following results presented in this thesis are original:

1) Measurement of the rate of DNA replication in CV-1 cells after release from serum starvation and aphidicolin block.

2) Application of the extrusion method to human (WI-38) cells and the isolation of human origin enriched sequences (hors).

3) Demonstration that 5 of the 10 hors clones can replicate autonomously as plasmids in HeLa cells.

4) Demonstration that two of the hors clones contain bent DNA and one contains a weak DNA unwinding element.

5) Determination and analysis of the primary DNA sequence of five of the hors clones.

6) Purification and characterization of a DNA binding activity from HeLa cell nuclei that interacts in a specific manner with a sub-fragment of hors106.

REFERENCES

Aguinaga, M. de P., Sharan, C., Singh, D.N., Kegler, D., Siddiqui, K.A., and Valenzuela, M.S. (1989). Chromosomal mapping and nucleotide sequence of a human DNA autonomously replicating sequence. Genomics 5:605-611.

Anachkova, B., and Hamlin, J.L., (1989) Replication in the amplified dihydrofolate reductase domain in CHO cells may initiate at two distinct sites, one of which is a repetitive sequence element. Mol. Cell. Biol. 9:532-540.

Anderson, J.N., (1986) Detection, squence patterns and function of unusual DNA structures. Nucleic Acids Res. 14:8513-8533.

Antoniou, M., deBoer, E., Habets, G., and Grosveld, F., (1988) The human B-globin gene contains multiple regulatory regions: identification of one promoter and two downstream enhancers. EMBO J.7:377-384.

Ashihara, T., Baserga, R., (1979) Specialized Techniques. Cell Synchronization. Methods in Enzym. 58:248-262.

Baur, C.P., and Knippers, (1988) Protein-induced bending of the simian virus 40 origin of replication. J. Mol. Biol. 203:1009-1019.

Bayev, A A., Georgiev, O.I., Hadjiolov, A.A., Kermekchiev, M.B., Nikolaev, N., Skryabin, K.G., and Zaharyev, V.M. (1980) The structure of the yeast ribosomal RNA genes. 2. The nucleotide sequence of the initiation site for ribosomal RNA transcription. Nuc. Acids Res. 8:4919-4926.

Beard, P., Morrow, J.F., and Berg, P. (1973) Cleavage of circular superhelical simian virus 40 DNA to a linear duplex by S1 nuclease. J. Virol. 12:1303-1313.

Bell, G.I., DeGennaro, L.J., Gelfand, D.H., Bishop, R.J., Valenzuela, P., and Rutter, W.J. (1977) Ribosomal RNA genes of saccharomyces cerevisiae. I. Physical map of the reapeating unit and locations of the regions coding for 5S, 5.8S, 18S, and 25S ribosoma 'NAs J.Biol. Chem. 252:8118-8125.

Berezney, R., and Coffey, D.S., (1977) Isolation and characterization of a framework structure rat liver nuclei. J. Cell. Biol. 73:616-637.

Berezney, R., and Buchholtz, L.A. (1981) Dynamic association of replicating DNA fragments with the nuclear matrix of regenerating liver. Exp. Cell Res. 132:1-14.

Berrios, M., Osheroff, N., and Fisher, P.A. (1985) In situ localization odf DNA topoisomerase II, a major polypeptide component of the drosophila nuclear matrix fraction. Proc. Natl. Acad. Sci. USA

82:4142-4146.

Blumenthal, A.B., Kriegstein, H.J., and Hogness, D.S., (1974) The units of DNA replication in Drosophila melanogaster chromosomes. Cold Spring Harbor Symp. Quant. Biol. 38:205-223.

Borsa, J., and Whitmore, B.F., (1969) Cell killing studies on the mode of action of methotrexate on L-cells in-vitro. Cancer Res. 29:737-744.

Bramhill, D. and Kornberg, A. (1988) Duplex opening by DNAa protein at novel sequences in initiation of replication at the origin of the E.coli chromosome. Cell 52:743-755.

Bramhill, D., and Kornberg, A. (1988) A model for initiation at origins of DNA replication. Cell 54:915-918.

Braunstein, J.D., Schulze, D., DelGiudice, T., Furst, A. and Schildkraut, C.L. (1982) The temporal order of replication of murine immunoglobin heavy chain constant region sequences corresponds to their linear order in the genome. Nucl. Acids Res. 10:6887-6902.

Bravo, R., Fey, S.J., Bellatin, J., Mose Laron, P., Arevalo, J., and Celis, J.E. (1981) Identification of a nuclear and of a cytoplasmic polypeptide whose relative proportions are sensitive to changes in the rate of cell proliferation. Exp. Cell Res. 136:311-320.

Bravo, R., Frank, R., Blundell, P.A., and MacDonald-Bravo H., (1987) Cyclin/PCNA is the auxiliary protein of DNA polymerase delta. Nature 326:515-520.

Brewer, B.J. and Fangman, W.L. (1987) The Localization of Replication Origins on ARS Plasmids in S. cerevisiae. Cell 51:463-471.

Briggs, M.R., Kadonaga, J.T., Bell, S.P., and Tijan, R. (1986) Purification and biochemical characterization of the promoter-specific transcription factor, Spl. Science 234:47-52.

Broach, J.R., Li, Y.Y., Feldman, J., Jayaram, M., Nasmyth, K.A. and Hicks, J.B. (1982) Localization and sequence analysis of yeast origins of DNA replication. Cold Spring Harbor Symp. Quant. Biol. 42:1165-1173.

Brown, E. H., Iqbal, M.A., Stuart, S., Hatton, K.S., Valinski, J. and Schildkraut, C.L. (1987) Rate of replication of the murine immunoglobin heavy-chain locus: evidence that the region is part of a single replicon. Mol. Cell. Biol. 7:450-457.

Burgers, P.M.J. (1989) Eukaryotic DNA polymerases alpha and delta: conserved properties and interactions from yeast to mammalian cells. Prog. Nucleic Acids Res. Mol. Biol. 16:6297-6306.

Burhans, W.C., Vassilev, L., Caddle, M.S., Heintz, N.H., and

Depamphilis, M.L. (1990). Identification of an Origin of Bidirectional DNA Replication in Mammalian Chromosomes. Cell **62**:955-965.

Caddle, M.S., Dailey, L., and Heintz, N.H. (1990). RIP60, a Mammalian Origin-Binding Protein, Enhances DNA Bending near the Dihydrofolate Reductase Origin of Replication. Mol. Cell. Biol. **10**:6236-6243.

Caddle, M.S., Lussier, R.H., and Heintz, N.H. (1990) Intra molecular DNA Triplexes, Bent DNA and DNA Unwinding Elements in the Initiation Region of an Amplified Dihydrofolate Reductase Replicon. J. Mol. Biol. **211**:19-33.

Cairns, J. (1966). Autoradiography of HeLa cell DNA. J. Mol. Biol. **15**:372-373.

Callan, H.G. (1972) Replication of DNA in the chromosomes of eukaryotes. Proc. R. Soc. London, Ser B. **181**:19-41.

Callan, H. G., (1973) DNA replication in the chromosomes of eukaryotes. Cold Spring Harbor Symp. Quant. Biol. **38**:195-204.

Campbell, A.M., and Jolly, D.J. (1973) Light-scattering studies on superhelical unwinding. Biochem. J. **133**:209-226.

Campbell, J.L. (1986) Eukaryotic DNA Replication. Ann. Rev. Biochem. **55**:733-771.

Carthew, R.W., Chodosh, L.A., and Sharp, P.A. (1985) An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. Cell **43**:439-448.

Casey, J.L., Hentze, M.W., Koeller, D.M., Caughman, S.W., Rouault, T.A., Klausner, R.D., and Harford, J.B., (1988) Iron-responsive elements: regulatory RNA sequences that control mRNA levels and translation. Science **240**:924-928.

Celis, J.E., and Madsen, P., (1986) Proliferation sensitive nuclear phosphoprotein "dividin" is synthesized almost exclusively during the S phase of the cell cycle in AMA cells. Proc. Natl. Acad. Sci. USA **83**:8187-8193.

Celis, J. E., Peterson Ratz, G., and Celis, A. (1987) Progressin: A novel proliferation sensitive and cell cycle-regulated human protein whose rate of synthesis increases at o near the G1/S transition border of the cell cycle. FEBS Lett. 223 -24?.

Chan, C.S.M., and Tye, B. (1980) Autonomously replicating sequences in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA **77**:6329-6333.

Collins, J.M., (1978) Rates of DNA synthesis during the S phase of Hela cells. J. Biol. Chem. **253**:8570-8577.

Collins, S., and Groudine, M. (1982) Amplification of endogenous myc-

related DNA sequences in a human myeloid leukaemia cell line. Nature **298**:679-684.

Collins, J.M., Berry, O.E., and Bagwell, C.B., (1980) Different rates of DNA synthesis during the S phase of log phase HeLa S3, WI-38 and 2RA cells. J. Biol. Chem **255**:585-3590.

Conaway, R.C., and Lehman, I.R. (1982) A DNA primase activity associated with DNA polymerase alpha from Drosophila melanogaster embryos. Proc. Natl. Acad. Sci. USA **79**:2523-2527.

Cook, P.R. (1991) The nucleoskeleton and the topology of replication. Cell **66**:627-635.

Courtois, S.J., Lafontaine, D.A., Lemaigre, F P., Durviaux, S.M , and Rousseau, G.G. (1990) Nuclear factor I and activator protein-2 bind in a mutually exclusive way to overlapping promoter sequences and trans-activate the human growth hormone gene. Nucleic Acids Res **18**:57-64.

D'Anna, J.A., Crissman, H.A., Jackson, R.J., and Tobey, R (1985) Time dependant changes in H1 content, H1 turnover, DNA elongation, and the survival of cells blocked in early S phase by hydroxyurea, aphidicolin or 5-fluorodeoxyuridine. Biochemistry **24**:5020-5026.

Dailey, L., Caddle, N.H., Heintz, N., and Heintz, N.H. (1990) Purification of RIP60 and RIP100, Mammalian Proteins with Origin-

Specific DNA-Binding and ATP-Dependant DNA Helicase Activities. Mol. Cell. Biol. **10**:6225-6235.

Dana, K.J., and Nathans D., (1971) Specific cleavage of Simian Virus 40 DNA by restriction endonuclease of Hemophilus Influenzae. Proc. Natl. Acad. Sci. USA **68**:2913-2917.

Dean, F.B., Borowiec, J.A., Ishimi, Y., Deb, S., and Tegtmyer, P. (1987) Simian virus 40 large tumour antigen requires three core replication origin domains for DNA unwinding and replication in-vitro. Proc. Natl. Acad. Sci. USA **84**:8267-8271.

Deb, S., DeLucia, A.L., Baurr, C. Koff, A., and Tegtmeyer, P. (1986a) Domain structure of the simian virus 40 core origin of replication. Mol. Cell. Biol. **6**:1663-1667.

Deb, S., Delucia, A.L., Koff, A., Shanli, T. and Tegtmeyer, P (1986b) The adenine-thymidine domain of the simian virus 40 core origin directs DNA bending and coordinately regulates DNA replication. Mol. Cell. Biol. **6**:4578-4584.

DePamphilis., M.L., (1988) Transcriptional elements as components of eukaryotic origins of DNA replication. Cell **52**:635-638.

Der, C.J., 1989. The ras family of oncogenes. In Oncogenes (ed. C. Benzand E. Liu), 74-119. Kluwer Academic Publishers, Amsterdam.

Deusberg, P., Bister, K., and Vogt, P.K. (1977) The RNA of avian acute leukemia virus MC29. Proc. Natl. Acad. Sci. USA 74:4320-4329

Dhar, V., Mager, D., Iqbal, A., and Schildkraut, C L. (1988) The coordinate replication of the human B-globin gene domain reflects its transcriptional activity and nuclease hypersensitivity Mol. Cell Biol. 8:4958-4965.

Dierks, P., Van Ooyen, A., Cochran, M.D , Dobkin, C., Reiser, J , and Weissmann, C., (1983) Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit B globin gene in mouse 3T3 cells. Cell 32:695-706

Diffley, J.F.X., and Stillman, B. (1988) Purification of a yeast protein that binds to origins of DNA replication and a transcriptional silencer. Proc. Natl. Acad. Sci. USA 85:2120-2124.

Dijkwel, P.A., Wenink, P.W., and Poddighe, J. (1986) Permanent attachment of replication origins to the nuclear matrix in BHK cells Nuc. Acids Res. 14:3241-3250.

Dingman, C.W. (1974) Bidirectional chromosome replication: some topological considerations. J. Theor. Biol. 43:187-195.

Dinter-Gotleib, G., and Kaufmann, G. (1982) Nucleic Acids Res. 10: 763-773.

Dodson, M., Roberts, J., McMacken, and Echols, H. (1985) Specialized nucleoprotein structures at the origin of replication of bacteriophage lambda: complexes with lambda O protein and with lambda P, and Escherichia coli DNA B proteins. Proc. Natl. Acad. Sci. USA **82**:4678-4782.

Earnshaw, W.C., and Laemmli, U.K. (1983) Aechitecture of metaphase chromosomes and chromosome scaffolds. J. Cell Biol **96**:84-93.

Earnshaw, W.C., Halligan B., Cooke, C.A., Heck, M.M.S., and Lui, L.F. (1985) Topoisomerase II is a structural component of mitotic chronosome scaffolds. J. Cell Biol. **100**:1706-1715.

Eisenman, R.N., Tachibana, C.H., Abrams, H.D., and Hann, S.R. (1985) v-myc and c-myc-Encoded proteins are associated with the nuclear matrix. Mol. Cell. Biol. **5**:114-126.

Evans, T., Reitman, M., and Felsenfeld, G., (1988) An erythrocyte-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes. Proc. Natl. Acad. Sci. **85**:5967-5980.

Fairman, M.P., and Stillman, B., (1988) Cellular proteins required for multiple stages of SV40 DNA replication in-vitro. EMBO J. **7**:1211-1218.

Falaschi, A., Biamonti, G., Cobianchi, F., Csordas-Toth, E., Faulkner, G., Giacca, M., Pedacchia, D., Perini, G., Riva, S., and Tribioli,

Presence of transcriptional signals in two putative DNA replication origins of human cells. Biochem. Biophysica Acta. **951**: 430-442

Fischer, P.A., Berrios, M., and Blobel, G. (1982) Isolation and characterization of a proteinacious subnuclear fraction composed of nuclear matrix, peripheral lamina, and nuclear pore complexes from embryos of drosophila melanogaster. J. Cell Biol. **92**:674-686

Fitzgerald, M., and Shenk, T. (1981) The sequence 5'-AAUAAA-3' forms part of the reconition site for polyadenylation of late SV40 mRNAs Cell **24**:251-260.

Frappier, L. and Zannis-Hadjopoulos, M. (1987) Autonomous replication of plasmids bearing monkey DNA origin-enriched sequences Proc. Natl Acad. Sci. USA **84**:6668-6672.

Frappier, L., Price, G.B., Martin, R.G., and Zannis-Hadjopoulos, M. (1987) Monoclonal antibodies to cruciform DNA structures J Mol. Biol. **193**:751-758.

Galson, D.L., and Housman, D.E., (1988) Detection of two tissue specific DNA-binding proteins with affinity for sites in the mouse B-globin intervening sequence 2. Mol Cell. Biol. **8**:381-392.

Garner, M.M., and Revzin, A. (1981) A gel electrophoresis method for quantifying the binding of proteins to specific regions: application

to components of the E. coli lactose operon regulatory system Nucl. Acids Res. **9**:3047-3060.

Gasser, S.M., and Laemmli, U.K., (1986) Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of D. melanogaster. Cell **46**:521-530.

Gellert, M , Fisher, L.M., Ohmori, H., ODea, M.H , Mizuuchi, K. (1981) DNA gyrase: site specific interactions and transient double-stranded breakage of DNA. C.S.H. Symp. Quant. Biol; 45:391-398.

Gerace, L., and Blobel, G. (1980) The nuclear envelope lamina is reversibly depolymerized during mitosis. Cell **19**:277-287.

Gerace, L., Blum, A., and Blobel, G. (1978) Immunocytochemical localization of the major polypeptides of the nuclear pore complex-lamina fraction. Interphase and mitotic distribution. J. Cell. Biol. **79**:546-566.

Gerard, R., and Gluzman, Y. (1986) Functional analysis of the role of the A/T rich region and upstream flanking sequences in simian virus 40 DNA replication Mol. Cell. Biol. **6**:4570-4577.

Giacca, M., Biamonti, G., Cobianchi, F., Colonna, M., Tribioli, C., Riva, S., and Falaschi, A. (1988) Sequence-Specific DNA binding Protein(s) that Bind(s) to a Putative Human DNA Replication Origin.

37: 1807-1808.

Gil, G., Smith, R.J., Goldstein, C.A., Slaughter, K., Orth, M.S., Brown, M., and Osborne, T.F., (1988) Multiple genes encode nuclear-factor I like proteins that bind to the promoter for 3' hydroxy-3-methylglutaryl-coenzyme A reductase. Proc. Natl. Acad. Sci. USA 85:8963-8967.

Glazer, R.I., and Knode, M.C., (1984) 1-beta-D-arabinosyl-5-azocytosine. Cytocidal activity and effects on the synthesis and methylation of DNA in human colon carcinoma cells. Mol. Pharmacol 26:381-387.

Goscin, L.P., and Byrnes, J.J. (1982) DNA polymerase delta· One polypeptide, two activities. Biochemistry 21:2513-2518.

Graham, F.L. and Van der Eb, A. J. (1973) A new technique for the assay of infectivity of Human Adenovirus 5 DNA. Virology 52:456-467

Gronostajski, R.M., (1986) Analysis of nuclear factor 1 binding to DNA using degeneraten oligonucleotides. Nuc. Acids Res. 14:9117-9132.

Guggenheimer, R.A., Nagata, K., Kenny, M., Hurwitz, J. (1983) Protein primed replication of plasmids containing the terminus of the adenovirus genome II. Purification and characterization of a host protein required for the replication of DNA templates devoid of the

terminal protein. J. Biol. Chem. 259:7815-7825.

Hamlin, J.L., Leu, T.H. Foreman, P.K., Ma, C.A., Vaughn, J.P., Dijkwel, P.A., Creeper, L. and Willard, H. (1988) Cancer Cells, 6:329-333.

Hammond, R.A., Byrnes, J.J., and Miller, M.R. (1987) Biochemistry 26: 6817-6824.

Hancock, R., and Boulikas, T., (1982) Functional organization in the nucleus. Int. Rev. Cytol. 79:165-214.

Hand, R. (1978). Eukaryotic DNA: Organization of the genome for replication. Cell 15:317-325.

Hand, R., (1975) Regulation of DNA replication on subchromosomal units of mammalian cells. J. Cell Biol. 64:89-97.

Hand, R., (1979) Eukaryotic chromosome replication and its regulation. Cell Biol. 2:389-437.

Hand, R., and Tamm, I. (1974) Initiation of DNA replication in mammalian cells and its inhibition by reovirus infection. J. Mol. Biol. 82:175-183.

Harland, R.M., and Laskey, R.A., (1980) Regulated replication of DNA

microinjected into eggs of Xenopus laevis. Cell 21:761-771.

Hartwell, L.H., Culotti, J., Pringle, J.R., and Reid, B J. (1974) Genetic control of the cell cycle in yeast. Science 183·46-51

Hay, R.T., and DePamphilis, M.L., (1982) Initiation of SV40 DNA replication in-vivo: location and structure of 5' ends of DNA synthesized in the ori region. Cell 28:767-779.

Hayflick, L. and Moorhead, P.S. (1965) The serial cultivation of human diploid cell strains. Exp. Cell Res. 25·585-621

Heintz, N.H., and Hamlin, J.L. (1982) An amplified chromosomal sequence that includes the gene for dihydrofolate reductase initiates replication within specific restriction fragments. Proc Natl. Acad. Sci. USA 79: 4083-4087.

Heintz, N.H., Milbrandt, J.D., Greisen, K.S., and Hamlin, J L. (1983) Cloning of the initiation region of a mammalian chromosomal replicon. Nature 302: 439-441.

Heinzel, S. S., Krysan, P.J., Tran, C.T., and Calos, M.P. (1991). Autonomous DNA Replication in Human Cells is Affected by the Size and the Source of the DNA. Mol. Cell. Biol. 11: 2263-2272.

Henninghaysen, L., and Fleckstein, B. (1986) Nuclear factor I

interacts with 5 DNA elements in the promoter region of the human cytomegalovirus major immediate early gene. EMBO J **5**:1367-1371.

Hirt, B. (1967) Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **26**: 365-369.

Hixon, J., Wong, T. and Clayton, D.J. (1986) Both the conserved stem-loop and divergent 5'-flanking sequences are required for initiation at the human mitochondrial origin off light strand DNA replication. J.Biol. Chem., **261**: 2384-2390.

Holley, R.W., (1975) Growth control of mammalian cells in cell culture. Nature **258**:487-493.

Howard, A., and Pelc, S.R., (1951) Nuclear incorporation of P^{32} as demonstrated by autoradiographs. Exp. Cell Res. 2:178-183.

Huberman, J.A. and Riggs, A.D. (1968) On the mechanism of DNA replication in mammalian chromosomes. J. Mol. Biol. 32: 327-341.

Huberman, J.A., Spotila, L.D., Nawotka, K.A., El-Assouli, S.M. and Davis, L.R. (1987) The <u>in-vivo</u> replication origin of the yeast 2um plasmid. Cell **51**: 473-481.

Iguchi-Ariga, S.M.M., Itani, T., Yamaguchi, M., and Ariga, H. (1987a) C-myc protein can be sustituted for SV40 T antigen in SV40 DNA

replication. Nuc. Acids Res. 15:4889-4899.

Iguchi-Ariga, S.M.M., Itani, T., Kiji, Y. and Ariga, H. (1987b) Possible function of the c-myc product: promotion of cellular DNA replication. EMBO 6:2365-2371.

Iguchi-Ariga, S.M.M., Okazaki, T.I., Ogata, M., Sato, Y., and Ariga, H. (1988a) An initiation site of DNA replication with trancriptional enhancer activity present upstream of the c-myc gene. EMBO J. 7:3135-3142.

Iguchi-Ariga, S.M.M., Okazaki, T., Itani, T., and Ariga, H. (1988b) Cloning of the p53-dependent origin of DNA replication. Oncogene 3:509-515.

Imagawa, M., Chiu, R., and Karin, M. (1987) Transcription factor AP-2 mediates induction by two different signal-transduction pathways: protein kinase C and cAMP. Cell 51:251-260.

Jacob, F., Brenner, J., and Cuzin, F., (1963) On the regulation of DNA replication in bacteria. Cold Spring Harbor Symp. Quant. Biol 28:329-348.

Jackson, D.A., and Cook, P.R. (1986) Replication occurs at a nucleoskeleton. EMBO J. 5:1403-1410.

Kariya, Y., Kato, K., Hayashizaki, Y., Himeno, S., Tarui, S., and Matsubara, K. (1987) Revision of consensus sequence of human Alu repeats. Gene **53**:1-10.

Kaufmann, G., Zannis-Hadjopoulos, M. and Martin, R.G. (1985) Cloning of nascent monkey DNA synthesized early in the cell cycle. Mol. Cell. Biol. 5:721-727.

Kearsey, S. (1984) Structural requirements for the function of a yeast chromosomal replicator. Cell **37**:299-307.

Kelly, K., Cochran, B.H., Stiles, C.D., and Leder, P. (1983) Cell-specific regulation of the c-myc gene by lymphoocyte mitogens and platelet-derived growth factor. Cell 35:603-611.

Kemper, B., Jackson, P.D., and Felsenfeld, G. (1987) Protein-binding sites within the 5' DNaseI-hypersensitive region of the chicken a-globin gene. Mol. Cell. Biol. **7**:2059-2069.

Krysan, P.J., Haase, S.B., and Calos, M.P. (1989). Isolation of Human Sequences that Replicate Autonomously in Human Cells. Mol. Cell. Biol. **9**: 1026-1033.

Krysan, P.J., and Calos, M.P. (1991). Replication Initiates at Multiple Locations on an Autonomously Replicating Plasmid in Human Cells. Mol. Cell. Biol. **11**: 1464-1472.

Lacatena, R.M. and Cesareni, G., (1981) Base pairing of RNAI and its complementary sequence in the primer precursor inhibits ColE1 replication. Nature **294**:623-626.

Landry, S., and Zannis-Hadjopoulos, M. (1991) Classes of autono ously replicating sequences are found among early-replicating monkey DNA. Biochem. Bio. Acta. **108**:234-244.

Lebkowski, J. A., and Laemmli, U.K. (1982) Non-histone proteins and long range organization of HeLa interphase DNA. J. Mol. Biol. **156**:103-119.

Leegwater, P.A.J., Van der Vilet, P.C., Rupp, R.A.W., Nowock, J., and Sippel, A.E. (1986) Functional homology between the sequence-specific DNA-binding proteins nuclear factor I from HeLa cells and the TGGCA protein from chicken liver. EMBO J **5**:381-386.

Leu, T.H., and Hamlin, J.L., (1989) High resolution mapping of replication fork movement through the amplified dihydrofolate reductase domain in CHO cells by in-gel renaturation analysis. Mol. Cell. Biol. **9**:523-531.

Levy, D.E., Kessler, D.S., Pine, R., Reich, N., and Darnell, J.E. (1988) Interferon-induced nuclear factors that bind a shared promoter element correlate with positive and negative transcriptional control. Genes Develop. **2**:383-393.

Lewis, C.D., Clark, S.P., Felsenfeld, G., and Gould, H. (1988) An erythrocytre-specific protein that binds to the poly(dG) region of the chicken B-globin gene promoter. Genes Develop. 2:863-873.

Li, J.J., and Kelly, T., (1984) Simian virus 40 DNA replication in-vitro Proc. Natl. Acad. Sci USA **81**:6973-6982.

Li, J.J., Peden, K.W.C., Dixon, R.A.F., and Kelly, T. (1986) Functional organization of the simian virus 40 origin of replication. Mol. Cell. Biol. **6**:1117-1128.

Lilley, D.M.J. (1980) The inverted repeat as a recognizable structural feature in supercoiled DNA molecules. Proc. Natl. Acad. Sci. USA **77**:6468-6472.

Lilley, D.M.J. (1985) The kinetic properties of cruciform extrusion are determined by DNA base sequence. Nucl. Acids Res. **13**:1443-1465.

Linskens, M.H.K., and Huberman, J.A. (1988) Organization of replication of ribosomal DNA in Saccharomyces cerevisiae. Mol. Cell. Biol. **8**:4927-4935.

Looney, J.E., and Hamlin, J.L. (1988) Isolation of the amplified dihydrofolate reductase domain from methotrexate-resistant Chinese hamster ovary cells. Mol. Cell. Biol. 7:569-577.

Maniatis, T., Jeffrey, A., and van de Sande, H. (1975) Chain length determination of small double stranded and single stranded DNA molecules by polyacrylamide gel electrophoresis. Biochemistry 14:3787-3794.

Marini, J.C., Weisberg, R., and Landy, A. (1977) The isolation of restriction fragments containing the primary and secondary (galT) bacterial att sites of phage lambda. Virology 83:254-270.

Marini, J.C., Levenue, S.D., Crothers, D.M., and Englund, P.T. (1982) Proc. Natl. Acad. Sci. USA 79:7664-7668.

Marraccino, K.L., Wahl, A.F., Keng, P.C., Lord, M.E., and Bambera, R.A., (1987) Cell cycle dependant activities of DNA polymerases alpha and delta in chinese hamster ovary cells. Biochemistry 26: 7864-7870.

McAlear, M., Ward, G.K., McKenzie,R., Price, G.B. and Zannis-Hadjopoulos, M. (1989) Biphasic activation of DNA replication. Mol. Gen. (life Sci. Adv.) 8:11-14.

McConaughy, B.L., Young, L.S., and Champoux, J.J., (1981) The effect of salt on the binding of the eukaryotic DNA nicking-closing enzyme to DNA and chromatin. Biochem. Biophys. Acta 655:1-9.

McCready, S.J., Godwin, J., Mason, D.W., Brazell, I.A., and Cook, P.R. (1980) DNA is replicated at the nuclear cage. J. Cell Sci. 46:365-386.

McWhinney, C., and Leffak, M. (1990) Autonomous replication of a DNA fragment containing the chromosomal replication origin of the c-myc gene. Nuc. Acids Res. **18**:1233-1242.

Meijer, M., Beck. E. Hansen, F.G., Bergman, H.E.N., Messer, W., von Meyenberg, K., and Schaller, H. (1979) Nucleotide sequence of the origin of replication of the Escherichia coli K-12 chromosome. Proc. Natl. Acad. Sci. USA **76**:580-584.

Mermod, N., Williams, T.J., and Tijan, R., (1988) Enhancer binding factors AP-4 and AP-1 act in concert to activate SV40 late transcription in-vitro. Nature 322:557-561.

Mertz, J.E., and Berg, P., (1974) Viable deletion mutants of simian virus 40: selective isolation by means of a restriction endonuclease from Hemophilus parainfluenzae. Proc. Natl. Acad. Sci. USA 12:4879-4883.

Milbrandt, J.D., Heintz, N.H., White, W.C., Rothman, S.M., and Hamlin, J.L. (1981) Methotrexate-resistant Chinese hamster ovary cells have amplified a 135-kilobase-pair region that includes the dihydrofolate reductase gene. Proc. Natl. Acad. Sci. USA. **78**: 6043-6047

Milbrandt, J.D., Azizkhan, J.C., Greisen, K.S., and Hamlin, J.L. (1983) Organization of a Chinese Hamster Ovary Dihydrofolate Reductase Gene Identified by Phenotypic Rescue. Mol. Cell. Biol. 3:1266-1273.

Mitchell, P.J., Wang, C., and Tijan, R., (1987) Positive and negative regulation of transcription in-vitro: enhancer-binding protein AP-2 is inhibited by SV40 T antigen. Cell 50:847-861.

Murray, A.W., (1987) A cycle is a cycle, is a cycle. Nature 327: 14-15.

Musukata, H. and Tomizawa, J. (1986) Control of primer formation for ColE1 plasmid replication: conformational change to the primer transcript. Cell 44:125-136.

Muzard, G., Theveny, B., and Revet, B. (1990) Electron microscopy mapping of pBR322 DNA curvature, comparison with theoretical models. EMBO J 9:1289-1298.

Myers, R.M., Tilly, K., and Maniatis, T. (1986) Fine structure genetic analysis of the beta globin promoter. Science 232:613-618.

Nagata, K. Guggenheimer, R.A., Enomoto, T., Lichy, J.H., Hurwitz, J. (1983) Adenovirus DNA replication in-vitro: Identification of a host factor that stimulates synthesis of the pre-terminal protein-dCMP complex. Proc. Natl. Acad. Sci. USA 79:6438-6442.

Nakamura, H., Morita, T., and Sato, C., (1986) Structural organizations of replicon domains during DNA synthetic phase in the mammalian nucleus Exp. Cell Res. 165:291-297.

Nelson, W.G., Pienta, K.J., Barrack, E.R., and Coffey, D.S. (1986) The role of the nuclear matrix in the organization and function of DNA. Ann. Rev. Biophys. Chem. **15**:457-475.

Nobile, C., and Martin, R.G. (1986) Stable stem-loop and cruciform DNA structures: isolation of mutants with rearrangements of the palindromic sequence at the simian virus 40 replication origin. Intervirology **25**:158-171.

Oakley, B.R., Kirsch, D.R., and Morris, N.R. (1980) A simplified ultra sensitive silver stain for detecting proteins in acrylamide gels Anal. Biochem. **105**:361-363.

O'Neil, E.A., and Kelly, T.J. (1988) Purification and characterization of nuclear factor III (origin recognition protein C), a sequence-specific DNA binding protein required for efficient initiation of adenovirus DNA replication. J. Biol. Chem. **263**:931-937.

Painter, R.B., and Schaeffer, A.W., (1969) Rate of synthesis along replicons of different kinds of mammalian cells. J. Mol. Biol. **45**:467-479.

Palzkill, T.G., and Newlon, C.S. (1988) A yeast replication origin consists of multiple copies of a small conserved sequence. Cell **53**:441-450.

Panayotatos, N., and Wells, R.D. (1981) Cruciform structures in supercoiled DNA. Nature **289**:466-470.

Paulson J.R., and Laemmli, U.K., (1977) The structure of histone depleted metaphase chromosomes. Cell **12**:817-828.

Peden, K.W., Pipas, J.M., Pearson-White, S., and Nathans, D. (1980) Isolation of mutants of an animal virus in bacteria. Science **209**:1392-1396.

Pearson, C.E., Frappier, L., and Zannis-Hadjopoulos, M. (1991) Plasmids bearing mammalian DNA replication origin enriched (ors) fragments initiate semiconservative replication in a cell free system. Biochim. Biophys. Acta **1090**:156-166.

Pedrali-Noy, G., Spadari, S., Miller-Faures, A., Miller, A.O., Kruppa, J., and Koch G., (1980) Synchronization of HeLa cell cultures by inhibition of DNA polymerase alpha with aphidicolin. Nuc. Acids Res. **8**:377-387.

Pelham, H., (1985) Activation of heat shock genes in eukaryotes. Trends Genet. **1**:31-35.

Pfeifer, K., Prezant T., and Guarente, L. (1987) Yeast HAP1 activator binds to two upstream activation sites of different sequence. Cell **49**:19-27.

Phillippsen, P., Thomas, M.J., Kramer, R.A., and Davis, R.W. (1978) Unique arrangement of coding sequences for 5S, 5.8S, 18S, and 25S ribosomal RNA in Saccharomyces cerevisiae as determined by R-loop and hybridization analysis. J. Mol. Biol. **123**:387-404.

Phillips, P.D., and Cristofalo, V.J. (1981) Growth regulation of WI-38 cells in a serum-free medium. Exp. Cell Res. **134**:297-302

Prelich, G., Kostura, M., Marshak, D.R., Mathews, M.B., and Stillman, B. (1987) The cell cycle regulated proliferating cell nuclear antigen is required for SV40 DNA replication in-vitro. Nature **326**:471-476.

Pruijin, G.J.M., van Miltenburg, R.T., Classens, J.A.J., and van der Vilet, P.C. (1988) Interaction between the octamer-binding protein nuclear factor III and the adenovirus origin of DNA replication. J. Virol. **62**:3092-3102.

Rao, B.S., Zannis-Hadjopoulos, M., Price, G.B., Reitman, M. and Martin, R.G. (1990) Sequence similarities among monkey ori-enriched (ors) fragments. Gene **87**:233-242.

Razin, S.V., Kekelidze, M.G., Lukanidin, E.M., Scherrer, K. and Georgiev, G.P. (1986) Replication origins are attached to the nuclear skeleton. Nucl. Acids Res. 14:8189-8207.

Rekosh, D.M.K., Russel, W.C., Bellet, A.J.D., and Robinson, A.J.

(1977) Identification of a protein linked to the ends of adenovirus DNA. Cell **11**:283-295.

Richter, A., and Hand, R. (1979) Eukaryotic DNA replication. J. Cell Physiol. **101**:407-417.

Reynods, A.E., McCarroll, R.M., Newlon, C.S., and Fangman, W.L. (1989) Time of Replication of ARS Elements along Yeast Chromosome III. Mol. Cell. Biol. **9**:4488-4494.

Robinson, A.J., Younghusband, H.B., and Bellett, A.J.D. (1973) A circular DNA-protein complex from adenoviruses. Virology **56**:54-59.

Rosenfeld, P.J., and Kelly, T.J., (1986) Purification of nuclear factor I by DNA recognition site affinity chromatography. J. Biol. Chem. **261**:1398-1408.

Rosenfeld, P.J., O'Neil, E.A., Wides, R.J., and Kelly, T.J. (1987) Sequence-specific interactions between cellular DNA-binding proteins and the adenovirus origin of DNA replication. Mol. Cell. Biol. **7**:875-886.

Ryder, K., Silver, S., DeLucia, A.L., Fanning, E. and Tegtmeyer, P. (1986) An altered DNA conformation in origin region I is a determinant for the binding of SV40 large T antigen. Cell **44**:719-725.

Sander, M., and Hsieh, T. (1985) Drosophila topoisomerase II double strand DNA cleavage: analysis of DNA sequence homology at the cleavage site. Nuc. Acids Res. **13**:1057-1072.

Sassone-Corsi, P. (1988) Cyclic AMP induction of early adenovirus promoters involves sequences required for E1A trans activation. Proc. Natl. Acad. Sci. USA **85**:7192-7196.

Sen, R., and Baltimore, D. (1986) Multiple nuclear factors interact with the immunoglobin enhancer sequences. Cell **46**:705-716.

Shelfin, L.G. and Kowalski,D. (1984) Mung bean nuclease cleavage of a dA+dT-rich sequence or an inverted repeat sequence in supercoiled PM2 DNA depends on ionic environment. Nucl. Acids Res. **12**: 7087-7104.

Sheflin, L.G. and Kowalski, D. (1985) Altered DNA conformations detected by mung bean nuclease occur in promoter and terminator regions of supercoiled pBR322 DNA. Nuc. Acids Res. **13**:6137-6154.

Shortle, R., Margolskee, R.F., and Nathans, D. (1979) Proc. Natl. Acad. Sci USA **76**:6128-6131.

Simpson, L., (1979)Isolation of maxicircle component of kinetoplast DNA from hemoflagellate protozoa. Proc. Natl. Acad. Sci. USA **76**:1585-1588.

Singh, H., Sen, R., Baltimore, D., and Sharp, P.A. (1986) A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobin genes. Nature 319:154-158.

Skryabin, K.G., Eldarov, M.A., Larionov, V.L., Bayev, A.A., Klootwijk, J., de Regt, V.C.H.F., Veldman, G.M., PLanta, R.J., Georgiev, O.I. and Hadjiolov, A.A (1984) Structure and function of the nontranscribed spacer regions of yeast rDNA. Nuc. Acids Res. 12:2955-2968.

Smith, H.C., Puvion, E., Buchholtz, L.A., and Berezney, R. (1984) Spatial distribution of DNA loop attachement and replication sites in the nuclear matrix. J. Cell Biol. 99:1794-1802.

Snyder, M., Buchman, A.R. and Davis, R.W., (1986) Bent DNA at a yeast autonomously replicating sequence. Nature 324:87-89.

Spadari, S., Sala, F. and Pedrali-Noy, G. (1982) Aphidicolin: a specific inhibitor of nuclear DNA replication in eukaryotes. Trends Biochem. Sci. 7: 29-32.

Spitzner, J.R., and Muller, M.T. (1988) A consensus sequence for cleavage by vertebrate DNA topoisomerase II. Nuc. Acids Res. 16:5533-5556.

Spotila, L.D., and Huberman, J.A., (1985) Method for mapping DNA relication Origins. Mol. Cell Biol. 5:85-92.

Stuart, G.W., Searle, P.F., and Palmiter, R.D. (1985) Identification of multiple metal regulatory elements in mouse metallotionein-I promoter by assaying synthetic sequences. Nature 317:828-831.

Stillman, B. (1989) Initiation of eukaryotic DNA replication in-vitro. Ann. Rev. Cell Biol. 5:197-245.

Stinchcomb, D.T., Struhl, K., and Davis, R.W., (1979). Isolation and characterization of a yeast chromosome replicator. Nature 282:39-43.

Struhl, K., Stinchcomb, D.T., Scherer, S., and Davis, R.W. (1979) High-frequency transformation of yeast: Autonomous replication of hybrid DNA molecules. Proc. Natl. Acad. Sci. USA 76: 1035-1039.

Sullivan, K.M. and Lilley, M.J. (1986) A dominant influence of flanking sequences on a local structural transition in DNA Cell 47: 817-827.

Tan, C.K., Castillo, C., So, A.G., and Downey, K.M. (1986) An auxiliary protein for DNA polymerase delta from fetal calf thymus. J. Biol. Chem. 261:12310-12316.

Trempe, J.P., Lindstrom, Y. I., and Leffak, M., (1988) Opposite replication polarities of transcribed and nontranscribed histone H5 genes. Mol. Cell. Biol. 8:1657-1663.

Tribioli, C., Biamonti, G., Giacca, M., Colonna, M., Riva, S. and Falaschi, A. (1987). Characterization of human sequences synthesized at the onset of S-phase. Nuc. Acids Res. 15: 10211-10232.

Trifonov, E.N., and Sussman, J.L. (1980) The pitch of chromatin DNA is reflected in its nucleotide sequence. Proc. Natl. Acad. Sci. USA 77:3816-3820.

Tsurimoto, T., and Stillman, B. (1989) Purification of a cellular replication factor, RF-C, that is required for coordinated synthesis of leading and lagging strands during SV40 DNA replication in-vitro. Mol. Cell. Biol. 9:609-619.

Umek, R.M. and Kowalski, D., (1988) The ease of DNA unwinding as a determinant of initiation at yeast replication origins. Cell 52: 559-567

Umek, R.M. and Kowalski,D. (1987) Yeast regulatory sequences preferentially adopt a non-B conformation in supercoiled DNA. Nucleic Acids Res., 15:4467-4480.

Vogelstein, B., Pardoll, D.M., and Coffey, D.S., (1980) Supercoiled loops and DNA replication. Cell 22:79-87.

Wall, L., De Boer, E., and Grosveld, F., (1988) The human beta-globin gene 3' enhancer contains multiple binding sites for an erythroid-

specific protein. Genes Develop. 2:1089-1100.

Wang, J.C., Peck, L.J. and Becherer, K. (1983) DNA supercoiling and its effects on DNA structure and function. Cold Spring Harbor Symp. Quant. Biol. 47:85-91.

Ward, G.K., McKenzie, R. Zannis-Hadjopoulos, M., and Price, G.B. (1990) The dynamic distribution and quantification of DNA cruciforms in eukaryotic nuclei. Exp. Cell Res. **188**:235-246.

Watson, J.D., and Crick, F. (1953) Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. Nature. **171**:737-738.

Weller, S.K., Spadaro,A., Schaffer,J.E., Murray,A.W., Maxam,A.M. and Schaffer,P.A. (1985) Cloning, sequencing, and functional analysis of oriL, a herpes simplex virus typel origin of DNA synthesis. Mol. Cell. Biol., 5:930-942.

Wellinger, R.J., and Zakian, V.A. (1989) Lack of positional requirements for autonomously replicating sequence elements on artificial yeast chromosomes. Proc. Natl. Acad. Sci. USA **86**:973-977.

Williams, J.S., Eckdahl, T.T., and Anderson, J.N. (1988) Bent DNA functions as a replication enhancer in Saccharomyces cerevisiae. Mol. Cell. Biol. **8**:2763-2769.

Wirak, D.O., Chalifour, L.E., Wasserman, P.M., Muller, W.J., Hassel, J.A., and DePamphilis, M.L. (1985) Sequence-dependant DNA replication in preimplantation mouse embryos. Mol. Cell. Biol. 5:2924-2935.

Wilson, D.A., and Thomas, C.A. Jr. (1973) Biochim. Biophys. Acta. 331:331-340.

Wobbe, C.R., Weissbach, L., Borowiec, J., Dean, F.B., Murakami, Y., Bullock, P., and Hurwitz, J. (1987) Replication of simian virus 40 origin containing DNA in-vitro with purified proteins. Proc. Natl. Acad. Sci. USA 84:1834-1838.

Wold, M. S., and Kelly, T. (1988) Purification and characterization of replication protein A, a cellular protein required for SV40 DNA replication in-vitro. Proc. Natl. Acad. Sci. USA 85:

Wood, S.H., and Collins, J.M. (1986) Preferential binding of DNA primase to the nuclear matrix in HeLa cells. J. Biol. Chem. 261:7119-7122.

Yamamoto, K.R. (1985) Steroid receptor regulated transcription of specific genes and gene networks. Annu. Rev. Genet. 19:209-252.

Yates, J.L., Warren, N., and Sugden, B. (1985) Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. Nature 313:812-815.

Zahn, K. and Blattner, F.R. (1985) Sequence-induced DNA curvature at the bacteriophage lambda origin of replication. Nature **317**:451-453.

Zahn,K. and Blattner, F.R. (1987) Direct evidence for DNA bending at the lambda replication origin. Science **236**:416-422.

Zannis-Hadjopoulos, M., Persico, M., and Martin, R.G. (1981) The remarkable instability of replication loops provides a general method for the isolation of origins of DNA replication. Cell **27**:155-163.

Zannis-Hadjopoulos, M., Chepelinsky, B., and Martin, R.G., (1983) Mapping of the 3'-end positions of simian virus 40 nascent strands. J. Mol. Biol. **165**:599-607.

Zannis-Hadjopoulos, M., Kaufmann, G. and Martin, R.G. (1984) Mammalian DNA enriched for replication origins is enriched for snap-back sequences. J. Mol. Biol. **179**: 577-586.

Zannis-Hadjopoulos, M., Kaufmann, G Wang, S.S., Lechner, R.L., Karawya, E., Hesse, J., and Martin, R.G. (1985) Properties of some monkey DNA sequences obtained by a procedure that enriches for DNA replication origins Mol. Cell. Biol. 5:1621-1629.

Zannis-Hadjopoulos, M., Frappier, L., Khoury, M., and Price, G.B. (1988) Effect of anti-cruciform DNA monoclonal antibodies on DNA replication EMBO J **7**:1837-1844.

Zyskind, J.W., and Smith, D.W. (1986) The bacterial origin of replication, oriC. Cell 46:479-490.

APPENDIX

List of human-hamster somatic cell hybrid lines.

CELL LINE	HUMAN CHROMOSOME NUMBER																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	Y
4106	-	-	+	-	+	-	-	-	-	+	-	-	+	+	+	+	+	+	-	+	-	-	+	-
4501	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-
4543	+	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	+	+	+	+	+	-	+	-
7614	-	-	+	-	+	+	-	-	-	-	+	-	+	+	-	-	+	+	-	+	+	-	+	+
7631	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
7633	+	-	+	+	+	-	-	+	-	+	-	-	+	-	+	+	+	+	+	-	+	-	+	-
7905	-	-	+	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	-	+	-	+	-
8005[d]	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	+	-	-	+	-	-	+	+
8014[c]	-	-	-	+	+	+	+	-	-	+	-	+	-	-	-	-	-	-	-	+	+	-	+	-
8017[a]	-	-	-	+	+	-	-	-	+	-	+	-	-	-	+	+	+	+	+	+	+	+	+	-
8282[a]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-
190.03	-	-	-	+	+	+	-	+	-	+	-	+	-	-	-	-	-	-	+	+	-	+	+	+
8516[a]	-	+	+	+	+	-	-	+	-	+	+	-	+	-	+	+	-	-	-	-	+	-	+	-
8927	-	-	-	-	+	-	-	-	+	-	-	+	-	-	+	+	+	+	-	+	-	+	-	-