Characterisation of *Xenopus* Origin Recognition Complex

A thesis submitted to the University of London
for the degree of Doctor of Philosophy

by

Tamara Tugal

Imperial Cancer Research Fund
Clare Hall Laboratories
South Mimms

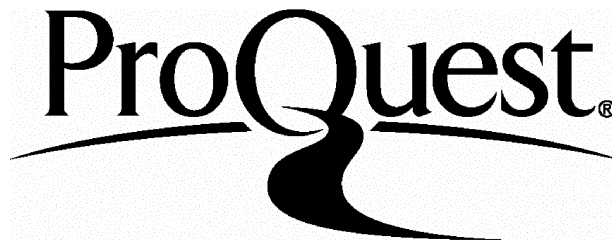University College London
Gower Street
London

2000

ProQuest Number: 10013328

ProQuest 10013328

With special thanks to Sir Kenneth Murray

Abstract

The Origin Recognition Complex (ORC) of *S. cerevisiae* plays a crucial role in recognition of origins of DNA replication within the yeast genome. In *Xenopus* egg extracts DNA replication initiates from random DNA sequences suggesting that origins of DNA replication are not required. Despite this apparent lack of requirement for DNA sequences serving as origins of DNA replication, a protein homologous to the Orc1p subunit of yeast ORC has been identified in *Xenopus*. This discovery has indicated that the mechanism which recognises replication origins in higher eukaryotes could be similar to that in yeast. Work described in this thesis started at the time when *Xenopus* Orc1p was identified. In the quest to understand *Xenopus* ORC and it's function in origin recognition I set out to purify this protein complex.

I produced monoclonal antibodies which recognised *Xenopus* Orc1p subunit of ORC with high affinity and specificity. Using these antibodies, I purified a protein complex which contained *Xenopus* Orc1p in association with at least four other polypeptides. I showed that one of these polypeptides was *Xenopus* Orc2p which was meanwhile identified by an independent research. Microsequencing of another Orc1p associated protein led to identification of mouse Orc4 gene from the EST database. I used the identified mouse Orc4 sequence to screen *Xenopus* cDNA library and I cloned *Xenopus* Orc4 gene. Microsequencing of another two proteins co-purified with Orc1p confirmed that they were *Xenopus* Orc3p and Orc5p and helped identification of their corresponding genes from frogs and humans.

I described a mobility shift of *Xenopus* Orc1p occurring upon the exit from mitosis and showed that it correlated with dephosphorylation of the protein. I determined that the amount of ORC present in a single frog egg is 100 000 fold higher than in a single somatic cell.

Finally, I showed that ORC purified from *Xenopus* egg extract binds to *S. cerevisiae* ARS1 sequence, as well as to DNA sequences derived from a bacterial plasmid. I observed that this ORC-DNA interaction was dependent on ATP. However, the way by which ORC determines the sites of initiation of DNA replication in higher eukaryotes remains unknown.

Acknowledgements

I would like to thank my PhD supervisor, Tim Hunt, whose enthusiasm for science is an unforgettable experience. Thank you for having me in your lab.

I would like to thank all the people from the Hunt laboratory who supported me throughout my PhD. I would like to thank Angel Nebreda for his advice about the Orc1p "shift". I would like to thank Julian Gannon for helping me to generate the anti-Orc1p monoclonal antibodies. I would like to thank Ralph Graeser for his advice and help with screening phage library in search for the Orc4 gene. I would like to thank Jane Kirk for making sure that I always had everything I needed. I would like to thank Birke Bartosh, Michael Brandeis, Patrick Descombes, Tod Duncan, Claudia Ellenrieder, Stephan Geley, Dolores Harrison, Helfrid Hochegger, Andrea Klotzbücher, Hiro Mahbubani, Steve Mayall, Jon Moore, Yoshimi Tanaka, Chizuko Tsurumi and Hiro Yamano for their support, suggestions, encouragement and advice. I especially would like to thank Jane Kirk, Tod Duncan and Jon Moore for their patience and care proof-reading parts of these thesis.

I would like to thank John Diffley for giving me the opportunity to attend his group's lab meetings and for his advice about my PhD work. I would like to thank all the people from the Diffley laboratory for their suggestions. In particular, I would like to thank Miguel Ferreira for the endless discussions about replication origins.

I would like to thank Bruce Stillman, Ryuji Kobayashi and Helena Zou-Yang from Cold Spring Harbour Laboratories for the interesting collaboration we had. I would like to thank Piotr Romanowski from Ron Laskey's lab, Alison Rowles from Julian Blow's lab and Phil Carpenter from Bill Dunphy's lab for sharing their reagents with me. I would like to thank all people form ICRF services for the excellent work they did for me.

Finally, I would like to thank my husband Bülent for encouraging me and advising me throughout my PhD.

4

Contents

List of Figures

List of Tables

Abbreviations

ABF1	ARS binding factor 1
ACE	Amplification control element
ACS	ARS consensus sequence
AER	Amplification enhancing regions
ARS	Autonomously replicating sequence
ATP	Adenosine 5'-triphosphate
BSA	Bovine serum albumin
CDC	Cell division cycle
CSF	Cytostatic factor
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
DHFR	Dihydrofolate reductase
ECL	Enhanced chemiluminescence
ED	Edman degradation
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glyco-bis(β-aminoethyl ether)N,N,N'N' tetraacetic acid
FCS	Foetal calf serum
HEPES	N-[2-Hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid]
HRP	Horseradish peroxidase
ICRF	Imperial Cancer Research Fund
IPTG	Isopropyl-β-D-thiogalactopyranoside
kDa	KiloDaltons
LCR	Locus control region
MS	Mass spectrometry
MAb	Monoclonal antibody
MBT	Mid-blastula transition
neo^r	Neomycin phosphotransferase gene
ORC	Origin Recognition Complex
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
SSB	Single strand DNA binding protein
SV 40	Simian virus 40
TBSA	Tris-buffered saline, version A
Tris	Tris[hydroxymethyl]aminomethane
2D	Two-dimensional

Chapter 1

Control of initiation of DNA replication

Living cells must duplicate their chromosomes before cell division so that each daughter cell has a complete and accurate copy of the genome. The fidelity of transmission of the genetic information stored in DNA depends in part on accurate replication, proofreading and repair, and also on the specialised mechanisms that are often called cell cycle checkpoints. Thus, replication does not begin unless certain conditions are met, such as adequate reserves of energy and DNA precursors (carbon, nitrogen and phosphorus). In multicellular organisms, signals from neighbouring cells can advance or delay the onset of S phase. There is normally a delay between the last division and the onset of DNA replication-the period known as G1-which allows cells to reach a critical size threshold before embarking on a program that leads to cell division. Once replication is initiated, it proceeds to completion barring accidents, and entry into mitosis is tightly coupled to that completion, so that eukaryotic cells do not divide with incompletely replicated chromosomes. At the same time, an extremely tight control mechanism exists to prevent the re-replication of already–replicated stretches of DNA. Chromosomes are replicated once, and only once, in each cell cycle.

This thesis is about the protein complex which is required for initiation of DNA synthesis in eukaryotes and is responsible for determination of the initiation site. Similar principles apply to DNA replication in bacteria and viruses where DNA replication is initiated from specific sites on the chromosome which are recognised by specialised origin recognition proteins.

Initiation of DNA replication in bacteria and viruses

Initiation of DNA replication in λ-phage occurs at oriλ

The process of initiation of DNA replication is perhaps best understood in the case of bacteriophage lambda. The origin of DNA replication in the genome of lambda phage (oriλ) was originally defined by electron microscopy of replicating lambda molecules and by analysis of defective prophages incapable of autonomous replication (Schnos and Inman, 1970). The 164 bp DNA fragment containing a functional lambda origin was identified by its ability to support the replication of a chimerical phage in *E. coli*. In this experiment, all λ gene products were supplied by a helper phage (Furth *et al.*, 1977). The sequence analysis of the 164 bp fragment containing oriλ revealed four 18 bp tandemly repeated, closely spaced DNA sequences and an adjacent AT rich region (Figure 1/1; Denniston-Thompson *et al.*, 1977; Grosschedl and G, 1979; Hobom *et al.*, 1979; Moore *et al.*, 1979).

Structure of the replication origin of the phage lambda
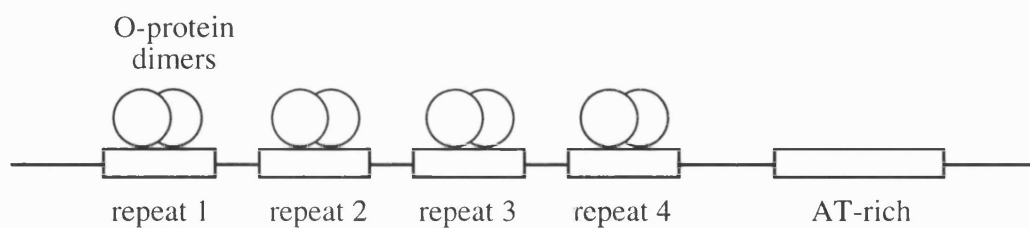
Figure 1/1. The 164 bp fragment carrying the full origin of DNA replication of bacteriophage lambda contains four tandem imperfect repeated sequences and an AT-rich region. An O-protein dimer is thought to bind to the each repeat forming a typical structure called the O-some in which the DNA is folded or wrapped around the O-protein. Structural changes caused by O-protein binding to the lambda origin region are thought to cause strand melting in the AT rich region.

The initiation of DNA replication from oriλ requires two phage encoded proteins, the O and the P protein and several *E. coli* replication proteins. Interestingly, the oriλ is located within the coding sequence of the O protein (Furth *et al.*, 1977). The O and P proteins are transcribed in the same transcriptional unit and their transcription activates oriλ. The timing of initiation of DNA replication is determined by the cascade of transcriptional regulation of the lambda phage life cycle (Mensa-Wilmot *et al.*, 1989a).

In vitro experiments using purified O protein showed that it binds to each of the four 18 bp tandemly arranged repeats (Figure 1/1) within oriλ. O protein binds as a dimer and four such dimers might bind to the origin sequence (Tsurimoto and Matsubara, 1981). Electron microscopy revealed that the O protein bound origin DNA forms a typical structure termed the "O-some" consisting of multiple O protein molecules with DNA being wrapped or folded around it (Dodson *et al.*, 1985). ATP or other molecular sources of energy are not required for assembly of the O-some. The formation of the O-some represents the first stage of the initiation of λ genome replication and serves to localise subsequent protein-protein and protein-DNA interactions.

The second stage of the initiation reaction is the recruitment of *E. coli* DnaB helicase facilitated by the phage encoded P-protein. The P-protein forms a tight complex with the bacterial DnaB (Wickner, 1979) and the P-DnaB complex interacts with the O-some to form a larger nucleoprotein structure at oriλ (Dodson *et al.*, 1985).

In the next stage of the reaction, *E. coli* heat shock proteins DnaJ and DnaK bind to the nucleoprotein structures assembled at oriλ (Alfano and McMacken, 1989b; Dodson *et al.*, 1989). DnaJ and DnaK act in an ATP-dependent manner to free DnaB

from it's tight inhibitory association with P-protein (Dodson *et al.*, 1986), and to remove P protein from oriλ (Alfano and McMacken, 1989a). The activation of the DnaB helicase results in localised unwinding, and the unwound DNA helix is stabilised by the *E. coli* single strand DNA binding protein (SSB). Adjacent to the O-some is the AT rich region (Figure 1/1) which is proposed to be the initiation site for DNA unwinding because of its high AT content (Denniston-Thompson *et al.*, 1977; Hobom *et al.*, 1979).

The unwound complex which is formed in this third stage "prepriming reaction" is competent to interact with primase (DnaG) to enable synthesis of primers which can be elongated by DNA polymerase III holoenzyme. Nine purified proteins comprising of phage encoded O and P proteins and *E. coli* DnaB helicase, DnaJ and DnaK heat shock proteins, single-stranded DNA binding protein, DnaG primase, DNA polymerase III holoenzyme and DNA gyrase are sufficient to replicate DNA molecules *in vitro* from oriλ. Although DNA gyrase (type II topoisomerase) is not required for the initiation of DNA synthesis, it is required for extensive propagation of the replication fork (Mensa-Wilmot *et al.*, 1989b).

A cell-free replication system composed of the nine purified proteins did not require transcriptional activation of oriλ which is required for DNA replication to occur from oriλ *in vivo* (Mensa-Wilmot *et al.*, 1989b). However, when the histone-like HU protein was added to the replication reaction, the dependence on transcription was restored (Mensa-Wilmot *et al.*, 1989a). HU protein does not inhibit the propagation of replication forks but apparently interferes with the assembly or function of the nucleoprotein structures assembled at oriλ. The chromatin structure of the DNA in the region surrounding oriλ may therefore play a central role in the negative

regulation of initiation of DNA replication from oriλ *in vivo* (Learn *et al.*, 1993;

Mensa-Wilmot *et al.*, 1989a).

Initiation of DNA replication at oriλ provides a simple model which is repeated in a number of biological systems. To successfully initiate DNA replication, replication origins on the DNA molecule are first identified. Specialised origin recognition proteins bind to the replication origin to mark the initiation site. The decision to initiate DNA replication is linked to the life cycle of the virus or cell. In the case of bacteriophage λ, this is facilitated by the transcriptional activation of the replication origin. Additional proteins recruited to the initiation site ensure localised DNA unwinding, usually at AT-rich sequences, followed by the actual initiation of bidirectional DNA synthesis.

Replication of the *E. coli* chromosome initiates at *OriC*

The origin of DNA replication on the *E. coli* chromosome (*OriC*) was identified as a fragment of the chromosomal DNA that is able to promote *E. coli* specific initiation of DNA replication (Meijer *et al.*, 1979; Oka *et al.*, 1980; Yasuda and Hirota, 1977). Functional motifs of the 245 bp minimal origin include two types of AT-rich sequences: four 9 bp repeated sequences and three 13 bp repeats (Figure 1/2).

The DnaA protein specifically recognises and binds to the 9 bp repeated sequence within *OriC*. Binding of DnaA is highly cooperative with 20-30 monomers of DnaA interacting at a single *OriC*, covering 200-250 bp of DNA (Fuller *et al.*, 1984). DNase I footprinting experiments suggest that the origin DNA is wrapped on the outside of the oligomerised DnaA protein core. Interestingly, one or more copies of the 9 bp sequence are found in a number of bacterial plasmids (Fuller *et al.*, 1984).

The DnaA protein binds ATP with very high affinity and hydrolyses it slowly to ADP in the presence of DNA. Only the ATP bound DnaA is active in replication (Sekimizu *et al.*, 1987). The ADP bound form of DnaA is inactive and does not exchange ADP for ATP. Remarkably, cell membrane phospholipids can reactivate the inert ADP form of DnaA. This interaction represents a potential link between the regulation of bacterial cell cycle and initiation of chromosomal replication (Sekimizu and Kornberg, 1988). Binding of DnaA to *OriC* DNA and consequent wrapping of the DNA around the protein structure induces melting of DNA strands in the 13 mers region and formation of an "open" complex (Bramhill and Kornberg, 1988).

Structure of *OriC*
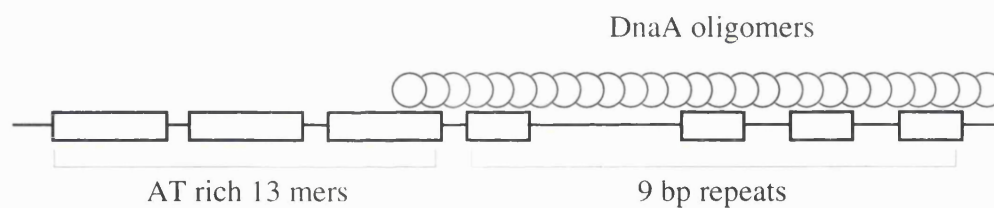
Figure 1/2. The 245 bp fragment carrying the full *OriC* contains four 9 bp repeats and three 13 bp AT rich repeats. The DnaA protein binds to the 9 bp repeats and oligomerises. *OriC* DNA wraps around the oligomerised DnaA resulting in melting of the DNA strands in the 13 mers region.

Following the melting of the AT-rich region, DnaB helicase is delivered to the opened complex by DnaC. DnaC interacts with DnaB to form a complex which is stabilised by ATP and when DnaB is complexed to DnaC it is inactive. Upon delivery of DnaB, DnaC is released from the protein-DNA complex. This release is ATP dependent and activates the DnaB helicase activity (Wahle *et al.*, 1989a; Wahle *et al.*, 1989b). Following the delivery to the melted region of *OriC*, DnaB acts as a helicase at the replication forks. Single strands of DNA are coated by the single

strand DNA binding protein (SSB) and DNA gyrase (type II topoisomerase) acts to relief torsional stress. DnaG primase then binds to DnaB complexes and synthesises primers on both strands. This is followed by replication of the template by DNA polymerase III holoenzyme.

Despite the detailed understanding of proteins assembled at *OriC* which leads to the initiation of DNA replication, the regulation of the timing of the initiation event is not well understood. The study of minichromosomes which replicate their DNA from *OriC* demonstrated that the timing of initiation from *OriC* is regulated. The minichromosomes replicated in a specific interval in the cell division cycle which corresponded to the timing of initiation of chromosome replication (Leonard and Helmstetter, 1986). However, the nature of this regulation remains unclear. The only connection between DNA replication from *OriC* and the bacterial cell cycle is the above mentioned association of DnaA with the plasma membrane.

Initiation at *OriC* resembles the initiation at $ori\lambda$ in some basic principles. The site of initiation on the DNA molecule is determined by a specific origin sequence. DnaA serves as an origin recognition protein and binds to the origin sequence. The analogous function in the phage λ is accomplished by the phage encoded O-protein. However, while recognition of ori λ by O-protein is ATP independent, only ATP-bound DnaA is competent to bind *OriC*. Although the timing of the initiation at *OriC* is regulated by the bacterial cell cycle, not much is known about the nature of this regulation. Finally, as in the case of λ, localised DNA unwinding and beginning of DNA synthesis is accomplished by additional proteins recruited to the initiation site.

SV40 encoded T-antigen binds the viral replication origin

Replication of the Simian virus 40 (SV40) genome has been studied extensively as a model for eukaryotic DNA replication. SV40 DNA replicates in the nucleus of human or monkey cells where the viral genome is complexed with histones to form a nucleoprotein structure indistinguishable from host chromatin. Only a single replication protein, the T-antigen, is encoded by the virus. SV40, like phage λ, highjacks the host replication machinery to replicate its own DNA, although both SV40 and phage λ encode their own origin recognition proteins to re-direct the host replication machinery to their own genomes. The virally encoded origin recognition proteins also overcome the checkpoints which prevent re-replication, permitting multiple rounds of replication of the viral genomes.

The SV40 origin of DNA replication consists of a 64 bp core region and several ancillary elements. The ancillary elements are not essential for the initiation of DNA replication, but they stimulate it 5 to 25 fold. The core sequence itself is tripartite (Figure 1/3). At the centre is 27 bp Site II which contains four GAGGC pentamers flanked by a 17 bp region containing exclusively A and T residues and a 15 bp imperfect palindromic element (Figure 1/3).

Only one viral protein, SV 40 large T-antigen, is required for the replication of SV40 genome (Myers and Tjian, 1980). DNase I protection studies showed that T-antigen binds SV40 DNA within a specific region and recognises the GAGGC pentamers within Site II of the core origin (Tjian, 1979). The interaction of origin DNA with T-antigen was further confirmed by specific co-immunoprecipitation of a fragment of DNA carrying the origin sequence using anti-SV40 T-antigen antibodies (McKay, 1983).

SV40 core origin
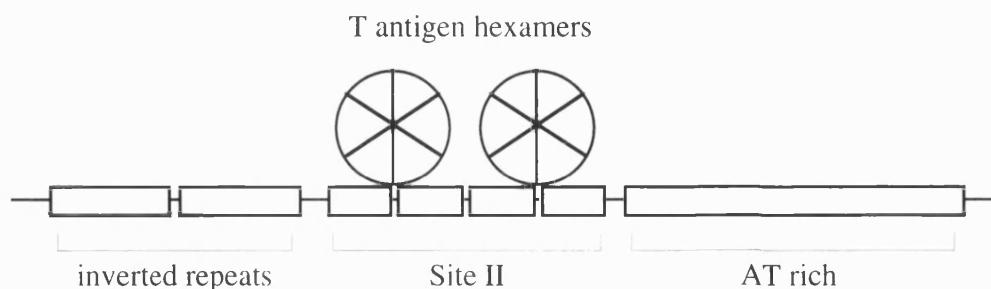
Figure 1/3. The 64 bp core sequence of SV40 replication origin consists of Site II which contains four GAGGC pentamers and a 17 bp AT rich region and 15 bp imperfect palindromic element. Two T–antigen hexamers bind to Site II.

In the presence of ATP, T-antigen oligomerises into double hexamer, centered on the core origin and extending beyond it by 12 bp in each direction (Dean *et al.*, 1987; Mastrangelo *et al.*, 1989). Despite the requirement for ATP, the hydrolysis of ATP is not required for the assembly of the T-antigen double hexamer. The formation of functional hexamer complexes is dependent on the presence of DNA (Dean *et al.*, 1992). The role of ATP may be to stabilise a conformation of the T-antigen molecule that is competent for multimerisation and DNA binding. The assembly of the T-antigen double hexamer on the origin DNA is critical for local unwinding of the duplex during initiation of DNA replication (Borowiec *et al.*, 1990). In addition to its function in origin recognition, T-antigen also possesses a helicase activity. Once bound to the origin, T-antigen is capable of entering the duplex and catalysing ATP dependent unwinding of the two DNA strands.

The development of SV 40 cell free replication system significantly advanced the understanding of the eukaryotic cellular replication machinery. Ten host replication proteins were identified which, together with T-antigen could carry out the complete process of replicating the SV 40 genome. Following DNA strand separation by T-

antigen, replication protein A (RPA) binds to unwound single strands preventing re-annealing of the DNA strands. Topoisomerase I relieves the superhelical torsion created by unwinding of supercoiled DNA. DNA polymerase alpha (Polα) catalyses the synthesis of RNA primers and extends the RNA primers into short DNA chains, with elongation of the newly synthesised DNA chains being catalysed by DNA polymerase delta (Polδ). Both RPA and PCNA are required for high efficiency elongation. Finally Rnase HI and FEN-1 (MF1) are required to remove the RNA primers. DNA strands are joined by DNA ligase I and untangled by topoisomerase I.

SV40 DNA replication is effectively controlled by T-antigen. The activity of T-antigen is regulated both positively and negatively by phosphorylation (Fanning, 1994; McVey *et al.*, 1989). T-antigen was found to interact with cyclinA/cdk2 kinase (Adamczewski *et al.*, 1993) which is the best physiological candidate to activate replication incompetent unphosphorylated T-antigen (Fanning, 1994). However, hyperphosphorylated T-antigen is also inactive in the SV 40 cell free replication system, suggesting that negative regulation of T-antigen function by phosphorylation also exists (Fanning, 1994). The function of regulation of T-antigen activity by phosphorylation remains however, unclear.

Initiation of DNA replication in SV40, likewise that in *E. coli* and λ, requires a specific origin sequence, an origin recognition protein and additional proteins to be recruited to the initiation site. The SV40 T-antigen not only recognises the replication origin, but also acts as a helicase, this function is in other systems accomplished by separate proteins.

20

Replication origins in yeast

Replication in *S. cerevisiae* initiates at the ARS sequence

In eukaryotic cells DNA replications starts at multiple sites on chromosomes to enable the replication of the larger eukaryotic genomes. Origins of DNA replication in *S. cerevisiae* were originally identified by their presence in fragments of yeast genomic DNA which conferred autonomous replication of circular plasmids in yeast cells without requiring integration into a yeast chromosome (Hsiao and Carbon, 1979; Stinchcomb *et al.*, 1979; Struhl *et al.*, 1979; Tschumper and Carbon, 1980). A number of such sequences were isolated from the yeast genome and the frequency of their occurrence in randomly cloned yeast DNA indicated that they are present on average every 30-40 kb (Beach *et al.*, 1980; Chan and Tye, 1980). This number was in a good agreement with an inter-origin distance obtained by electron microscopy analysis. The DNA sequences which facilitated the autonomous replication became known as autonomously replicating sequences (ARS). Curiously, almost half of autonomously replicating sequences isolated from a yeast genomic library showed homology to repetitive sequences (Chan and Tye, 1980).

Development of two-dimensional gel electrophoresis techniques allowed analysis of replication intermediates by identifying whether replication initiated within a particular DNA sequence (Brewer and Fangman, 1987; Huberman *et al.*, 1987). The neutral/neutral 2D gel method identifies either replication bubbles or replication forks. DNA fragments containing intact replication bubbles or forks are separated from each other and from nonreplicating linear DNA by their unique electrophoretic properties due to distinct shapes. This method is able to identify whether replication initiated within a particular DNA fragment. The neutral/alkaline 2D gel method identifies newly sythesized strands derived from replication bubbles. By identifying

the shortest newly synthesized strands, this method specifies, which DNA sequence was closest to the initiation site. Using these techniques it was shown that DNA replication initiated within the ARS sequence carried on a plasmid, confirming the long standing assumption that ARS elements functioned as origins of DNA replication in yeast (Brewer and Fangman, 1987; Huberman *et al.*, 1987). Many ARSs were subsequently shown to act as replication origins in their normal chromosomal context (Ferguson *et al.*, 1991; Huberman *et al.*, 1988; Linskens and Huberman, 1988; Zhu *et al.*, 1992).

Comparison of nucleotide sequences of seven independed ARS sequences identified an 11 bp consensus sequence (A/T)TTTA(T/C)(A/G)TTT(A/T). No other extended consensus sequence has been found, however, a 100 bp region in which the consensus is located has high A/T content (Broach *et al.*, 1983).

Mutational analysis of the 11 bp ARS consensus sequence (ACS) showed it to be essential for origin function (Van Houten and Newlon, 1990). However, the region required for full ARS activity was shown to extend beyond the ACS sequence (Van Houten and Newlon, 1990) and deletion studies of the ARS1 sequence identified three functional domains. Domain A contains the ACS and is necessary, but not sufficient to support efficient DNA replication (Celniker *et al.*, 1984). Domain B is required for efficient replication in addition to domain A. There is no obvious sequence homology between domain B of ARS1 and any other ARS sequence. Finally, domain C has weak effect on the origin function, which is only realised in the absence of domain B (Celniker *et al.*, 1984).

S.cerevisiae replication origin ARS1
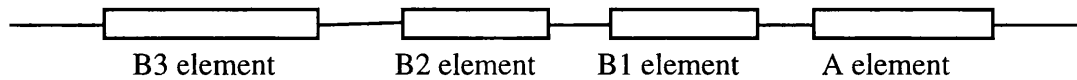
Figure 1/4. The 193 bp fragment of ARS1. The functional elements were identified as 15 bp A element which contains the 11 bp ARS consensus sequence and three distinct B elements.

Linker substitution mutations across an 193 bp fragment of ARS1 confirmed the presence of the 15 bp essential A element which contained an 11 bp match to the ARS consensus sequence and revealed the existence of three additional elements; a 13 bp B1, a 12 bp B2 and a 18 bp B3 which collectively are also essential for origin function (Marahrens and Stillman, 1992). Mutation of a single B element effected the origin function only slightly, but abolition of any two B sequences was deleterious for replication. Elements A, B1, B2 and B3 were sufficient to function as a replication origin when transferred to an unrelated sequence provided that the spacing between the elements resembled the spacing of the elements in the wild type origin (Marahrens and Stillman, 1992).

Linker substitution mutagenesis was also used to analyse the yeast replication origin ARS 307. These studies identified three DNA sequence elements in ARS 307, comprising of an A element containing the ACS, and two elements in similar positions to B1 and B2, that were important for efficient origin function (Rao *et al.*, 1994; Theis and Newlon, 1994). Although these two elements show little sequence similarity to the B1 and B2 elements of ARS1, they can be functionally substituted by their respective analogs (Rao *et al.*, 1994).

Origins in *S.pombe* contain AT rich sequence blocks

Origins of DNA replication in *S. pombe* were isolated as chromosomal sequences capable of supporting autonomous replication of plasmids in *S. pombe* (Maundrell *et al.*, 1988; Wohlgemuth *et al.*, 1994; Maundrell *et al.*, 1985). The isolation of ARSs from *S. pombe* was complicated by the fact that almost any plasmid was able to replicate in *S. pombe*, possibly due to rearrangements during which they were thought to obtain a chromosomal ARS (Maundrell *et al.*, 1988; Maundrell *et al.*, 1985). Despite these complication, ARS elements were successfully isolated from *S. pombe* and were estimated to occurs in the *S. pombe* genome approximately every 20 kb (Maundrell *et al.*, 1985). A number of *S. pombe* ARSs have been shown to function as origins of DNA replication in their normal chomosomal location (Dubey *et al.*, 1994; Wohlgemuth *et al.*, 1994).

Fission yeast ARS elements are AT rich and have minimal size of 500-1 000 bp which is considerably larger than the approximately 200 bp replication origins in budding yeast (Maundrell *et al.*, 1988). Detailed analysis of *S. pombe* ARS sequences revealed that each contains one or more 20 to 50 bp elements that are important for the ARS function. These sequences are generally extremely AT rich, and display tendency toward clustering of A and T residues. They do not however, share a common consensus sequence comparable to the 11 bp ARS consensus sequence of *S. cerevisiae* replication origins (Clyne and Kelly, 1995; Dubey *et al.*, 1996; Kim and Huberman, 1998; Maundrell *et al.*, 1988). *S. pombe* ARS elements are also characterised by a high degree of functional redundancy. In many cases elements that are important for function appear to be composed of smaller AT-rich sequence elements that can be deleted individually without significantly affecting origin activity. Chromosomal replication origins appear to be clustered into larger

initiation zones in *S. pombe*, perhaps to ensure that initiation occurs at regular intervals on the chromosome (Dubey *et al.*, 1994; Wohlgemuth *et al.*, 1994).

Replication origins in higher eukaryotes

In early *Xenopus* embryos initiation occurs at random sites

The identification of defined DNA sequences that serve as replication origins in viruses, bacteria and yeast supports the classical "replicon model". This model postulates that DNA replication would be controlled by a specific DNA sequence acting as a replicator and by a diffusible protein, an initiatior, that acts on the replicator to allow beginning of replication of the neighbouring DNA (Jacob *et al.*, 1963). Experiments in early *Xenopus* embryos showed that DNA replication can initiate from apparently random sequences however, with no need for a specific replicator and challenged the "replicon model" (Harland and Laskey, 1980; Laskey and Gurdon, 1973). In a landmark experiment, SV40 DNA injected into unfertilised *Xenopus* eggs was shown to undergo semiconservative DNA replication under cell cycle control (Harland and Laskey, 1980). Double-stranded circular DNA of bacterial viruses, bacterial plasmid DNA, or SV40 DNA lacking the viral origin of DNA replication also replicated semiconservatively when injected into the unfertilised eggs (Harland and Laskey, 1980). These findings provided strong evidence that no specialised DNA sequence is required to serve as an origin of DNA replication in *Xenopus* eggs. This apparent lack of specific sequence requirement for DNA replication in *Xenopus* eggs was confirmed by experiments testing DNA replication of variety of prokaryotic and eukaryotic substrates that included fragments of *Xenopus* genomic DNA. The efficiency of DNA replication in *Xenopus* eggs was mainly determined by the size of the DNA molecule tested, rather than any specific sequence (Mechali and Kearsey, 1984). Cell free extracts prepared from activated

Xenopus eggs behaved in a similar manner. Extracts that were able to initiate semiconservative DNA replication of sperm nuclei were also able to initiate and complete semiconservative DNA replication of DNA from a number of sources including DNA from bacteria or *Xenopus* (Blow and Laskey, 1986).

The neutral/neutral 2D gel electrophoresis technique was used to investigate whether replication initiated within a particular sequence on a given DNA molecule (Brewer and Fangman, 1987). One of the plasmids studied carried the *Xenopus* ribosomal repeat unit. The *Xenopus laevis* genes coding for the 40S ribosomal RNA precursor are arranged as a single unit of 11–15 kb repeated 400 times. Each unit contains the rRNA transcription unit and an intragenic spacer. This stretch of DNA is probably too long to be replicated from origins located outside of it. If, therefore, a specific origin exists, each repeat unit should contain such an origin. In fact, clusters of replication microbubbles were observed in the spacer region of the consecutive tandem units by electron microscopy of purified ribosomal DNA from *Xenopus* larvae (Bozzoni *et al.*, 1981). Plasmids carrying a single repeat unit were replicated either in the *Xenopus* cell free extracts or microinjected into *Xenopus* eggs. Analysis of the replication intermediates within the repeat unit identified initiation bubbles, single forks passing all the way through the repeat unit and repeat units replicated by two converging forks (Hyrien and Mechali, 1992; Mahbubani *et al.*, 1992). A similar pattern was observed for any fragment of the plasmid tested, whether it contained the ribosomal transcription unit, intragenic spacer or pBR322 plasmid sequences (Mahbubani *et al.*, 1992). These results indicated that initiation occurred at random places throughout the plasmid molecule. Similar results were observed when plasmids carrying the yeast 2 μm plasmid replication origin or a fragment of *Xenopus* genomic DNA were tested (Mahbubani *et al.*, 1992). Despite the abundance of potential origins on the tested plasmids however, multiple initiation events on a single

plasmid were not observed and the time required for plasmid replication was consistent with the time expected for replication starting from a single origin on the DNA molecule (Hyrien and Mechali, 1992; Mahbubani *et al.*, 1992).

To investigate whether replication of *Xenopus* chromosomal DNA initiates from specific sequences, the neutral/neutral 2D gel electrophoresis technique was used to study replication intermediates of the chromosomal rDNA cluster in early blastula (1 000-2 000) cell stage embryos. Initiation bubbles, single forks passing through, and two converging forks were observed in both the intragenic spacer and in the rRNA transcription unit, indicating that initiation occurred at multiple, apparently random positions on the chromosomal rDNA repeat (Hyrien and Mechali, 1993). These experiments also estimated that initiation of DNA replication within the ribosomal repeat unit occurred on average every 9-12 kb at this stage in *Xenopus* development (Hyrien and Mechali, 1993).

In later stage *Xenopus* embryos initiation becomes localised

In *Xenopus* early embryos, essentially all transcription, including that of the 40S ribosomal RNA precursor, is inhibited. After the mid-blastula transition, transcription of rRNA genes resumes. Analysis of the replication intermediates at the early gastrula stage of development showed strong inhibition of initiation from the rRNA transcription unit and increased incidence of initiation events in the intragenic spacer, although a low residual level of initiation could still be detected within the transcription unit (Hyrien *et al.*, 1995). The nature of this change in origin specificity during the *Xenopus* mid-blastula transition is far from clear and will be further discussed later. One possible explanation however, is that the intragenic spacer contains a complex battery of repeated promoter and enhancer elements that can bind the UBF transcription factor. This represents a potential link to the regulation of

some *S. cerevisiae* replication origins by the transcription factor ABF1 (Diffley and Stillman, 1988; Sweder *et al.*, 1988).

Attempts to clone mammalian origins have failed

Attempts to clone mammalian replication origins by plasmid rescue procedures similar to those employed in yeast have given controversial results. Some results showed that none of the human genomic fragments tested could support autonomous replication in human cells (Biamonti *et al.*, 1985), while other results suggested that any human sequences could support autonomous replication, with certain sequences promoting replication more efficiently than others, and that longer DNA fragments replicated better (Krysan *et al.*, 1993; Masukata *et al.*, 1993).

Initiation localises to the intragenic region at the DHFR locus

The identity of mammalian origins of DNA replication has also been studied by mapping the actual sites of initiation on chromosomes. These experiments identified several chromosomal locations from which DNA replication initiates. Perhaps the most extensively studied mammalian replication origin lies in the dihydrofolate reductase (DHFR) locus of Chinese hamster ovary (CHO) cells. In the CHOC400 cell line, the 250 kb DHFR locus has been amplified to give about 1 000 tandemly repeated units (Milbrandt *et al.*, 1981). The amplification of the DHFR locus meant that on an ethidiun bromide stained agarose gel of a total genomic digest, fragments derived from the DHFR locus were clearly visible against the total cellular DNA background. Radiolabelling of DNA during early S phase showed that out of the total 50 restriction fragments of the DHFR locus only three became labelled, indicating that replication initiated within a defined region of the DHFR locus (Heintz and Hamlin, 1982).

Each DHFR unit contains at least three transcription units: 25 kb DHFR gene, 65 kb REP3 gene, 34 kb 2BE2121 gene and a number of intragenic sequences (Figure 5/1; Foreman and Hamlin, 1989; Milbrandt *et al.*, 1983; Mitchell *et al.*, 1986). Cloning and mapping of the three DNA fragments from the DHFR locus which replicated early localised them to a single 28 kb locus within the intragenic region between the DHFR and 2BE2121 genes, further confirming localised initiation of DNA replication (Heintz *et al.*, 1983).

Further analysis of DNA fragments from the DHFR locus labelled early in S phase localised initiation into two short regions, termed oriβ and oriγ that are separated by approximately 22 kb (Figure 5/1; Leu and Hamlin, 1989). These findings were confirmed by the study of Okazaki fragments, which are found on different DNA strands on the opposite side of the replication origin. This study localised at least 80% of the initiation events into 450 bp region which contained oriβ (Burhans *et al.*, 1990), suggesting that oriβ represents the major initiation site.

The neutral/neutral 2D gel electrophoresis technique detected however, replication bubbles as well as forks travelling in both directions in every restriction fragment spanning the 55 kb of the DNA that lies between the DHFR and 2BE2121 genes (Dijkwel and Hamlin, 1992; Dijkwel *et al.*, 1994; Vaughn *et al.*, 1990), suggesting that initiation occurred at multiple sites within the 55 kb intragenic region. To resolve this apparent paradox initiation in the DHFR locus was carefully re-examined by a number of origin mapping methods. The consensus of these experiments is that somewhat preferred subzones at oriβ and oriγ reside in a larger zone of multiple potential initiation sites in the intragenic region (Wang *et al.*, 1998).

DHFR locus
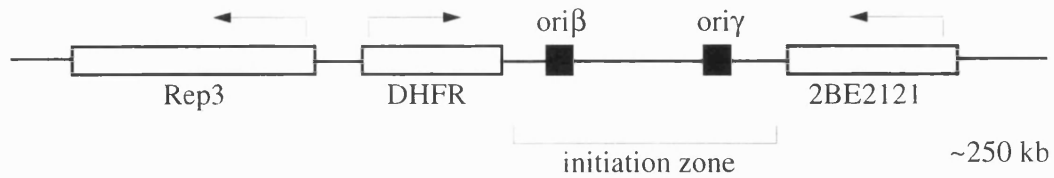
Figure 1/5. Organization of the Chinese hamster DHFR locus. Arrows show direction of transcription of the Rep3, DHFR and 2BE2121 genes.

The ability of oriβ to function as an origin on a plasmid molecule has been investigated further, based on the conclusion that it constitutes a major site of initiation of DNA replication in this locus (Burhans *et al.*, 1990). Plasmids containing a 4.5 kb fragment which included oriβ did not replicate in human cells. Plasmid carrying a 13.3 kb fragment, which in addition to oriβ contained surrounding sequences could replicate, but the activity was no higher than that of randomly chosen human fragments of similar size. Furthermore, two-dimensional gel analysis indicated that initiation did not preferentially occur within oriβ (Caddle and Calos, 1992).

To study the sequence requirement for initiation of DNA replication on the chromosome, a 4.3 kb fragment that contains oriβ was replaced with the *neo*^r cassette, but this had no dramatic effect on origin activity as fragments both upstream and downstream from the *neo*^r gene were able to initiate efficiently. Interestingly, replication bubbles were not detected within the *neo*^r gene which replaced the oriβ. Knocking out oriβ completely generated initiation pattern indistinguishable from the wild type. Surprisingly, a 13.5 kb deletion of sequences upstream of oriβ which

30

spanned the 3' end of the DHFR gene and a part of the intragenic spacer completely eliminated origin activity. Both oriβ and oriγ were devoid of any replication intermediates (Kalejta *et al.*, 1998). It will be interesting to establish whether essential regulatory DNA sequence exists within this region, or whether the loss of 3' processing signals of the DHFR gene allows the transcription complex to travel well into the intragenic region and interfere with the initiation process.

A fragment from the β-globin locus functions as an origin

Another mammalian origin of DNA replication has been identified by studying the direction of DNA replication covering more than 200 kb across the human β-globin locus which encodes six β-like globin genes. A single bidirectional origin of DNA replication has been located upstream of the β-globin gene using this method. This site was narrowed down to a defined 2 kb DNA fragment (Kitsberg *et al.*, 1993). This locus is used to direct DNA replication in both erythropoietic cells which actively transcribe the globin genes and in cells which do not (Kitsberg *et al.*, 1993). In cells which do not actively transcribe the globin genes, replication of the locus occurs late, while transcribing cells replicate the locus early (Epner *et al.*, 1988; Hatton *et al.*, 1988).

Patients suffering from the haemoglobin Lepore syndrome carry an approximately 8 kb deletion within the β-globin locus that includes the identified 2 kb replication origin (Mears *et al.*, 1978). In cells derived from these patients, the direction of DNA replication is reversed upstream of the origin locus, indicating that replication initiated from a site outside the studied locus (Kitsberg *et al.*, 1993).

50 kb upstream of the human β-globin replication origin lies the β-globin "locus control region" (LCR), which controls transcription, chromatin structure and

replication timing of the entire locus (Forrester *et al.*, 1990). A targeted insertion into the LCR locus which inactivated β-globin gene transcription and the chromatin structure of the locus did not affect initiation from the β-globin origin (Aladjem *et al.*, 1995; Kim *et al.*, 1992). In patients suffering from Hispanic thalassemia, a 35 kb region of DNA encompassing the LCR has been deleted (Driscoll *et al.*, 1989). This mutation inactivates the transcription of β-globin genes and causes a major alteration in chromatin structure (Driscoll *et al.*, 1989; Forrester *et al.*, 1990). Despite being about 50 kb away from the β-globin replication origin, this deletion completely abolished initiation from the origin. Instead, the locus apparently replicated from a downstream initiation site providing evidence for requirement of distantly separated sequences for initiation of DNA replication (Aladjem *et al.*, 1995).

To study the ability of the β-globin replication origin to function in novel chromosomal locations, an 8 kb fragment carrying the origin was transferred to ectopic sites in the simian genome. DNA synthesis indeed initiated from β-globin replication origin sequences located within the monkey chromosome, suggesting that specific sequences within this DNA fragment can function as a classic replicator. In these contexts, initiation from the β-globin replication origin was independent of the LCR which may act to alleviate a repressive effect of chromatin structure. In an ectopic location, this function may be served by other cis-acting elements.

The study of the 8 kb β-globin replication origin deletion mutants in ectopic location showed that deletions of 1.5 kb 5' or 4 kb 3' did not prevent initiation. However, a construct carrying both deletions; or a deletion of a 2.6 kb "core" region failed to initiate. The shortest fragment that was able to support initiation was 4.2 kb long and included the core region (Aladjem *et al.*, 1998). These results suggest a modular

structure of mammalian replication origins comprising of "core" sequences and auxiliary elements that affect replication efficiency.

Initiation in *Drosophila* chorion cluster localises to the ACE

The chorion genes of *Drosophila* provide an unique opportunity to study the replication of specific chromosome regions in a higher eukaryote. Chorion genes encode structural proteins of the eggshell and are arranged in a tandem cluster of four on chromosome three and a cluster of six on the X chromosome (Spradling, 1981; Spradling and Mahowald, 1980). During oogenesis, ovarian follicle cells over-replicate the chorion genes, during the same cell cycle, producing high levels of eggshell proteins. The cluster on chromosome X over-replicates approximately 15-fold, while the cluster on the third chromosome amplifies about 60-fold; equivalent to four and six rounds DNA replication, respectively. Gene amplification on the third chromosome occurs by initiating multiple rounds of DNA replication in the vicinity of the four tandemly arranged genes (Figure 1/6). Each newly initiated fork progresses along the chromosome, duplicating both the gene region and up to 50 kb of flanking chromosomal DNA, producing a bell-shaped gradient of amplification centred at the site of the gene cluster (Spradling, 1981).

To study the ability of DNA sequences from the third chromosome chorion locus to initiate the additional rounds of DNA replication, DNA segments derived from the cluster were inserted into novel chromosomal sites using P-element mediated transformation. Only a transposon containing a specific 3.8 kb segment derived from the cluster underwent amplification during oogenesis (Figure 1/6). However, the ability of this transposon to support initiation was subject to position effects. Attempts to induce amplification with subfragments of the 3.8 kb segment were unsuccessful, suggesting that most of this fragment is required for amplification (de

Cicco and Spradling, 1984). Deletion analysis of this 3.8 kb fragment identified a

510 bp region upstream of the s18 gene, which is essential for its function. This

Drosophila chorion cluster on chromosome three
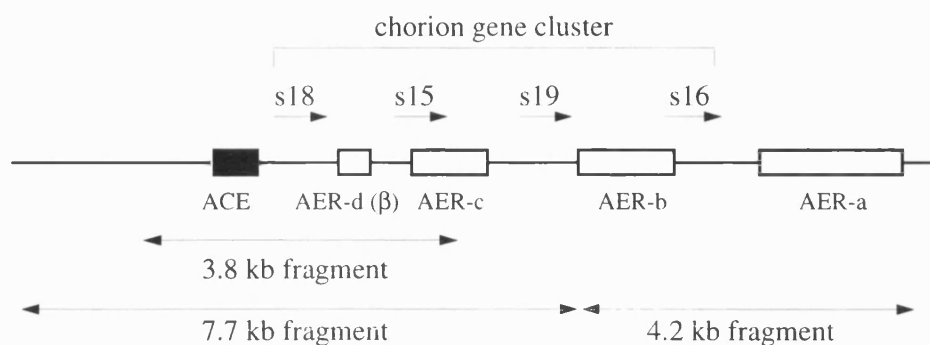
Figure 1/6. Organization of the *Drosophila* chorion gene cluster on chromosome three. Arrows labelled s18, s15, s19 and s16 represent the tandemly arranged chorion genes. ACE is the amplification control element and AER a-d represent the amplification enhancing regions. The AER-d region is also known as β. The 3.8 kb fragment was initially identified by its ability to confer replication competence at a novel chromosomal location. The majority of replication bubbles occured in the 7.7 kb fragment with smaller number of bubbles found in the 4.2 kb fragment.

region became known as the "amplification control element" (ACE) (Figure 1/6; Orr-

Weaver and Spradling, 1986). However additional sequences outside of the ACE

were found to enhance amplification. These "amplification enhancing regions"

(AER) could support a low level of amplification even in the absence of ACE,

suggesting functional redundancy within the chorion replicon (Delidakis and Kafatos,

1987; Swimmer *et al.*, 1989).

The neutral/neutral 2D gel electrophoresis analysis of the chorion cluster revealed

that the majority of the initiation events occurred within a 7.7 kb region with small

number of bubbles found in the adjacent 4.2 kb fragment (Figure 1/6). However,

multiple start sites within the 7.7 kb region were observed (Brewer and Fangman,

1987; Heck and Spradling, 1990). To clarify which DNA sequences were important

for the initiation of replication, the 3.8 kb fragment was studied by two-dimensional electrophoresis in a novel chromosomal location. In this setting, the majority of the initiation events localised to the ACE region and to the previously identified AER-d region which is also known as β (Heck and Spradling, 1990).

In summary, the emerging picture of a replication origin in higher eukaryotes involves interplay between DNA sequences, which appear to have a modular structure and a high degree of redundancy, and regulation by chromatin structure. In the case of *Xenopus* eggs however, DNA replication initiaties from random sequences and the mechanism responsible for this phenomena remains unclear.

The Origin Recognition Complex (ORC)

S. cerevisiae ARS sequences are recognised by ORC

The identification of a yeast replication origin sequence prompted a search for proteins that bind to this sequence to regulate initiation of DNA replication. The first protein to be identified interacting with *S. cerevisiae* ARS1 sequence was the 135 kDa ARS binding factor 1 (ABF1) (Diffley and Stillman, 1988; Sweder *et al.*, 1988). ABF1 was purified by its ability to effect a gel shift in ARS1 DNA (Diffley and Stillman, 1988; Sweder *et al.*, 1988). DNA footprinting defined the region within ARS1 recognised by ABF1 (Diffley and Stillman, 1988) and this region was shown to be within the B3 element of the ARS1 sequence (Marahrens and Stillman, 1992). ABF1 mutants show DNA replication defects similar to those observed in ARS1 mutants (Rhode *et al.*, 1992).

Interestingly, ABF1 is a transcription factor with a recognition motif found at many sites in the yeast genome. In an experiment where ABF1 binding site was replaced

by recognition sequence for another active transcription factor such as RAP1, the RAP1 protein could substitute for ABF1 function in origin firing (Marahrens and Stillman, 1992). Analysis of seven different ARSs revealed that ABF1 is not a universal ARS binding factor (Diffley and Stillman, 1988). The precise role of ABF1 in origin function is unclear, although in may be similar to the transcriptional activation of the replication origin of the phage λ.

A breakthrough in understanding of proteins controlling the regulation of the initiation of DNA replication in yeast came from purification of a multisubunit protein complex which specifically protected both the A and B1 elements of the replication origin ARS1 (Bell and Stillman, 1992). This protein complex, composed of six subunits, is called "origin recognition complex" (ORC).

Dnase I footprinting experiments of the ARS1 sequence using purified ORC identified 10 bp periodicity of cleavage on both strands, suggesting that DNA may wrap around ORC during binding (Bell and Stillman, 1992). The same 10 bp periodic cleavage was also observed in the ARS1 footprint *in vivo* (Diffley and Cocker, 1992). Increasing the amount of ORC in the binding reactions resulted in an increased region of protection, suggesting that perhaps more that one ORC molecule may interact with the replication origin (Bell and Stillman, 1992).

Interaction between ARS1 and ORC was also studied by analysing the binding of ORC to linker scanning mutants of ARS1. These experiments showed that mutations in the ARS consensus sequence (ACS) diminished ORC binding, implying that ACS is essential for recognition of ARS1 by ORC. ARS1 mutants which failed to efficiently bind ORC were diminished in their ability to replicate (Bell and Stillman, 1992). ORC was shown to recognise ACS in three other ARS sequences (ARS121,

ARS307 and H4ARS) suggesting a general role of ORC in recognition of the conserved sequence in replication origins of yeast (Bell and Stillman, 1992).

In addition to ORC recognition of ACS, the B1 element of the ARS1 was shown to significantly enhance ORC binding (Rao and Stillman, 1995; Rowley *et al.*, 1995). However, neither B2 or B3 elements seem to be required for ORC binding to ARS1 *in vivo* (Rowley *et al.*, 1995).

The six subunits constituting the *S. cerevisiae* ORC are called Orc1p (120 kDa), Orc2p (72 kDa), Orc3p (62 kDa), Orc4p (56 kDa), Orc5p (53 kDa) and Orc6p (50 kDa) (Bell and Stillman, 1992). The purified ORC subunits were microsequenced and the resulting peptide sequences were used to identify cDNAs encoding the individual subunits (Bell *et al.*, 1993; Bell *et al.*, 1995; Loo *et al.*, 1995). Orc2p and Orc6p were first isolated in independent genetic screens as genes interacting with a yeast origin of DNA replication (Foss *et al.*, 1993; Li and Herskowitz, 1993).

The function of ORC in DNA replication was also investigated by studying temperature sensitive mutations of the ORC subunits. Mutations in genes encoding Orc2p and Orc5p show cell cycle aberrations consistent with a defect in DNA replication (Bell *et al.*, 1993; Foss *et al.*, 1993; Fox *et al.*, 1995a; Loo *et al.*, 1995; Micklem *et al.*, 1993). Analysis of replication intermediates in Orc1p, Orc2p and Orc5p mutants, using two–dimensional gel electrophoresis indicated a reduced initiation frequency from chromosomal origins (Hori *et al.*, 1996; Liang *et al.*, 1995).

ORC is thus good candidate for the initiator protein of yeast (Jacob *et al.*, 1963). It specifically binds to yeast replication origin sequences, and recognises the ARS

consensus sequence. Like O protein of the λ phage, DnaA protein of *E. coli* and T-antigen of SV40, ORC in *S. cerevisiae* is likely to be the protein complex which recognises the origins of DNA replications and marks the point on the chromosome for other replication proteins to act upon.

To understand how the multisubunit ORC recognises the replication origin, the structure of the ORC–ARS complexes has been studied (Lee and Bell, 1997). ORC complexes lacking a single ORC subunit (either Orc1p, Orc2p, Orc3p Orc4p or Orc5p) were unable to bind to the ARS sequence. However, ORC complexes lacking Orc6p were capable of binding origin DNA in a sequence specific and DNA dependent manner. Cross-linking studies of ORC at either ARS1 or ARS305 origins have defined a working model for ORC-origin core interactions. In this model, Orc1p, Orc2p and Orc4p are associated with the ACS and Orc5p recognises the B1 element (Lee and Bell, 1997).

Importantly, ORC recognition of ARS1 DNA is strongly dependent on the presence of ATP. None of the remaining three nucleoside triphosphates stimulated ORC DNA binding at a level comparable to ATP (Bell and Stillman, 1992). Both Orc1p and Orc5p subunits of ORC contain putative nucleotide binding sites (Bell *et al.*, 1995; Loo *et al.*, 1995). To determine if ATP binding was mediated by these sites, ORC complexes that contained mutation in either the Orc1p ATP binding site or in the Orc5p ATP binding site were tested. ORC with the mutation in Orc5p bound little ATP in the absence of ARS1 DNA, but retained normal ARS1 DNA stimulated ATP binding activity. In contrast, ORC with a mutation in Orc1p bound ATP to the same extend as wild type in the absence of DNA but exhibited little or no stimulation by ARS1 DNA. ORC complexes containing Orc1p with a mutation in its ATP binding site were unable to bind origin DNA efficiently, while binding of ORC complexes

with mutation in Orc5p ATP binding site bound DNA at the wild type level. Together, these results implicates a strong link between ATP metabolism by Orc1p and sequence specific DNA binding by ORC (Klemm *et al.*, 1997).

Crosslinking studies with radiolabelled ATP identified three ORC subunits that were crosslinked with ATP. In addition to Orc1p and Orc5p, Orc4p was also labelled with the radioactive ATP. Complexes with mutations in the Orc1p ATP binding site eliminated ATP cross-linking to Orc1p and Orc4p, but not to Orc5p. Mutations in Orc5p ATP binding site only eliminated ATP cross-linking to Orc5p. Orc4p does contain a weak match to the ATP binding site consensus, but ORC complexes containing Orc4p mutated at this site showed no defect in ATP binding and cross-linking. These results suggest that Orc1p and Orc5p interact directly with ATP and that Orc4p is situated very near the Orc1p binding site (Klemm *et al.*, 1997).

ORC has weak ATPase activity, and addition of ARS1 DNA reduced this ATPase activity 8-fold. The inhibition was dependent on the presence of an intact ORC binding site. Although ATP is hydrolysed by Orc5p during the ORC-DNA binding reactions, this hydrolysis does not appear to be required for DNA binding. Instead, ATP bound to Orc1p seems to be a part of stable tripartite complex with ORC and the origin. The hydrolysis of the bound ATP molecule may play important role in the subsequent steps of the initiation reaction (Klemm *et al.*, 1997).

S.pombe ORC contains an "AT-hook" DNA recognition motif

Homologues of *S. cerevisiae* Orc1p, Orc2p, Orc4p and Orc5p have been identified in *S. pombe* (Chuang and Kelly, 1999; Grallert and Nurse, 1996; Ishiai *et al.*, 1997; Leatherwood *et al.*, 1996; Muzi-Falconi and Kelly, 1995), implying the presence of

an ORC complex in *S. pombe*. *S. pombe* Orc1p is 30% identical to its *S. cerevisiae* homologue and was shown to be essential for DNA replication (Grallert and Nurse, 1996; Muzi-Falconi and Kelly, 1995). Similarly, *S. pombe* Orc2p which is 22% identical to *S. cerevisiae* Orc2p is required for DNA replication. Surprisingly, *S. pombe* Orc4p consists of two similarly sized domains; a C-terminal domain showing high sequence similarity to the *S. cerevisiae* Orc4p, and an N-terminal domain with nine repeats of a so-called "AT-hook" motif which is absent in *S. cerevisiae* Orc4p. The AT-hook motif has been shown to bind in the minor groove of AT-tracts of double-stranded DNA (Geierstanger *et al.*, 1994; Solomon *et al.*, 1986). The presence of nine AT-hook motifs in *S. pombe* Orc4p suggests it may recognise AT-rich sequences in *S. pombe* replication origins. *In vitro* experiments using Orc4p or its N–terminal domain showed that it specifically recognises *S. pombe* ARS1 DNA suggesting how *S. pombe* ORC is targeted to origin DNA.

ORC is required for DNA replication in higher eukaryotes

Despite the differences in the DNA sequences that determine the position of replication origins in yeast and higher eukaryotes, the proteins determining the initiation site are well conserved. Proteins homologous to the subunits of ORC have been identified in a number of higher eukaryotes. First identified were *Drosophila* Orc2p and Orc5p proteins (Gossen *et al.*, 1995). The genes encoding these proteins were initially identified by analysis of the *Drosophila* genome project. The sequence conservation between the *Drosophila* and *S. cerevisiae* ORC protein is relatively low, with 21% identity between the Orc2p proteins from the two species and *Drosophila* Orc5p has 22% identity with the *S. cerevisiae* Orc5p (Gossen *et al.*, 1995). It was the low degree of conservation between ORC subunits from different species, which made their identification in higher eukaryotes difficult.

Drosophila Orc2p and Orc5p proteins were found to be complexed with four other proteins, suggesting the presence of an Origin Recognition Complex similar to that from yeast (Chesnokov *et al.*, 1999; Gossen *et al.*, 1995). Microsequencing of three co-purified proteins led to the identification of *Drosophila* Orc4p which is 24% identical to *S. cerevisiae* Orc4p, Orc3p with 21% identity to the yeast counterpart and Orc6p which is 19% identical to the *S. cerevisiae* Orc6p (Chesnokov *et al.*, 1999). The *Drosophila* homologue of Orc1p was also found to be part of the *Drosophila* ORC (Chesnokov *et al.*, 1999; Pak *et al.*, 1997). Interestingly, *Drosophila* Orc2 was independently identified as a gene whose mutation reduced chorion amplification in ovarian follicle cell, implying a function for ORC in the initiation of DNA replication in *Drosophila* (Landis *et al.*, 1997).

A fragment of human Orc1p was identified in two-hybrid screen for proteins interacting with the herpes simplex virus γICP34.5 protein and consequently used to isolate full length *Xenopus* Orc1p (Rowles *et al.*, 1996). *Xenopus* Orc2p was identified in a screen for inhibitors of cdc2 kinase activity that prevented mitotic catastrophe (Carpenter *et al.*, 1996). The purification and microsequencing of proteins associated with *Xenopus* Orc1p and Orc2p led to the identification of *Xenopus* Orc3p and Orc4p (Carpenter and Dunphy, 1998; Tugal *et al.*, 1998). Aminoacid identities between *S. cerevisiae* and *Xenopus* Orc1p, Orc2p, Orc3p and Orc4p are 20%, 23%, 14% and 22% respectively (Carpenter and Dunphy, 1998; Carpenter *et al.*, 1996; Rowles *et al.*, 1996; Tugal *et al.*, 1998).

Human Orc1p and Orc2p were identified by degenerate PCR (Gavin *et al.*, 1995). Human Orc1p is 27% identical to the *S. cerevisiae* protein. Human Orc4p was identified by searching EST databases and was found to be 29% identical to the Orc4p from *S. cerevisiae* (Quintana *et al.*, 1997; Tugal *et al.*, 1998). To identifiy

human Orc5p, databases were searched with peptide sequences derived from a polypeptide co-purifying with *Xenopus* Orc1p (Tugal *et al.*, 1998). These searches identified human Orc5p which has 29% identity to the yeast version of the gene (Ishiai *et al.*, 1997; Quintana *et al.*, 1998; Tugal *et al.*, 1998).

The identification of ORC subunit homologues in several higher eukaryotes suggests the presence of a complex with a function in origin recognition similar to that of *S. cerevisiae* ORC. ORC subunits in *Drosophila* and *Xenopus* are found in high molecular weight complexes (Carpenter and Dunphy, 1998; Carpenter *et al.*, 1996; Chesnokov *et al.*, 1999; Gossen *et al.*, 1995; Romanowski *et al.*, 1996b; Rowles *et al.*, 1996; Tugal *et al.*, 1998). Individual ORC subunits were also found to co-imunoprecipitate from human extracts (Quintana *et al.*, 1997; Quintana *et al.*, 1998). Moreover, purified ORC complexes from *Xenopus* and *Drosophila* were able to restore DNA replication in ORC depleted *Xenopus* extracts establishing the crucial role of ORC in DNA replication in higher eukaryotes (Chesnokov *et al.*, 1999; Romanowski *et al.*, 1996b; Rowles *et al.*, 1996). A function for ORC in DNA replication in higher eukaryotes is also supported by the fact that certain Orc2p mutants of *Drosophila* are defective in chorion locus amplification (Landis *et al.*, 1997). Interestingly, one copy of human Orc5 gene has been found to be deleted in uterine leiomyomas, which are benign tumours that rarely progress to malignancy (Quintana *et al.*, 1998). It is possible that deletion of one copy of Orc5 slows the cell cycle, preventing abundant proliferation and genomic instability leading to malignancy.

The identification of ORC complexes from higher eukaryotes provided a new prospective on the process of origin specification. Despite no clear DNA sequence serving as an origin of DNA replication being identified in higher eukaryotes,

complexes similar to those in *S. cerevisiae* are clearly required for DNA replication. This allowed studying origin specification in higher eukaryotes by investigating the sequence elements recognised by ORC. In *Drosophila, in vivo* and *in vitro* approaches have been used to demonstrate that *Drosophila* ORC binds to replication elements that direct initiation in the chorion gene cluster (Figure 1/6, Austin *et al.*, 1999). The ACE element, previously identified to be required for the ability of the DNA fragment from the chorion gene cluster to confer initiation competence to novel chromosomal location, was shown to direct immunolocalisation of ORC in follicle cells. *In vivo* cross-linking and chromatin immunoprecipitation also demonstrated association of ORC with ACE and with AER-d (Figure 1/6). Moreover, purified *Drosophila* ORC binds to ACE and AER-d in vitro, and like its *S. cerevisiae* counterpart, this binding is dependent on ATP. These experiments for the first time demonstrated that ORC from higher eukaryotes binds specifically to the DNA sequence which serve as an origin of DNA replication. In the view of these results, it is likely that ORC in *Xenopus* somatic cells recognises DNA sequence elements within the replication origins of *Xenopus*. The random initiation of DNA replication in *Xenopus* egg extracts remains however unexplained. The possible explanations for this apparent paradox include embryonic ORC being different to the somatic ORC or ORC levels in the egg being extremely high, driving ORC towards non-specific ORC-DNA interactions.

ORC plays a role in transcriptional silencing

Position effects are influences on the expression of a gene as a function of its location within the genome. Gene silencing is a common phenomenon in eukaryotes during which the silenced DNA is refractory to transcription. Silencing requires a combination of DNA sequences, called silencers, and multiple proteins, at least some

of which act at silencers. It has been suggested that silencing results from altered chromatin structure at the silenced locus.

Some silencers were found to also function as chromosomal origins of replication (McNally and Rine, 1991; Rivier and Rine, 1992). In addition, mutations in ORC not only reduced replication initiation, but also caused concomitant loss of silencing (Fox *et al.*, 1995b). These results led to the hypothesis of some link between replication and silencing. However, ORC alleles have been identified, which were defective in silencing, but not replication, suggesting that the replication and silencing functions of ORC are separable (Fox *et al.*, 1995b). This idea is also supported by the finding that *Drosophila* ORC2 could complement the silencing, but not the replication defect of the *S. cerevisiae* ORC2 mutant (Ehrenhofer-Murray *et al.*, 1995). Consistent with its role in silencing, ORC has been found to localise preferentially to heterochromatin in *Drosophila* and to interact with a heterochromatin localised protein, HP1, in both *Drosophila* and in *Xenopus* (Pak *et al.*, 1997).

Linking initiation to the cell cycle

Pre-replicative complexes licence origins for replication

The initiation of DNA replication in eukaryotes is an important event in the cell cycle. Cells must ensure that DNA replication has been completed before the start of mitosis and that the new round of DNA replication does not initiate before the end of mitosis. The use of genomic footprinting in *S. cerevisiae* allowed *in vivo* study of the chromatin changes occurring at yeast replication origins during the cell cycle (Diffley *et al.*, 1994). These experiments revealed that yeast replication origins exist in two states during the cell cycle. Immediately after DNA replication, post-replicative complexes are found on the replication origins. The post-replicative footprint

resembles closely a footprint generated by purified ORC and ABF1 proteins, implying that these are the main proteins responsible for the post-replicative footprinting pattern (Diffley *et al.*, 1994). Confirmation that the post-replicative genomic footprint across the ARS consensus sequence is generated by ORC came from an experiment showing that a temperature sensitive Orc2p mutant could cause disappearance of the post-replicative footprint (Santocanale and Diffley, 1996). Finally, chromatin binding and cross-linking assays showed that ORC is part of the post-replicative complexes and is also associated with the origin DNA throughout the cell cycle (Aparicio *et al.*, 1997; Donovan *et al.*, 1997; Liang and Stillman, 1997; Tanaka *et al.*, 1997).

Chromatin at origins is also found in a pre-replicative state. In the pre-replicative footprint, ORC induced hypersensitive sites are suppressed and the region of protection over the B domain is extended (Diffley *et al.*, 1994). These features of the pre-replicative complexes suggest that additional proteins, besides ORC and ABF1 participate in pre-replicative complex formation (Diffley *et al.*, 1994). The pre-replicative complexes appear near the end of mitosis and persist during G1, until the origins are either fired or passively replicated by a passing replication fork.

Replication origins do not fire simultaneously during S phase, (Fangman and Brewer, 1992) and cells arrested during S phase possess early origins that have already initiated replication while late origins remain unfired (Santocanale and Diffley, 1998). Genomic footprinting analysis of early and late origins in these arrested cells showed that late origins were maintained in the pre-replicative state while early origins already converted to the post-replicative state (Bousset and Diffley, 1998). Furthermore, once the cells are released from the arrest, the pre-replicative complexes on the late origins reverted into the post-replicative complexes as the late origins fire.

These experiments indicate that the pre-replicative complexes are indeed required for conferring competence for initiating DNA replication, and implicate that the proteins which compose the pre-replicative complexes are essential for the initiation process.

Cdc6 protein participates in pre-replicative complex formation

One of the components of pre-replicative complexes in *S. cerevisiae* is the Cdc6 protein. Temperature sensitive mutants of Cdc6 arrest with unreplicated DNA (Hartwell, 1976) and both genetic and biochemical evidence suggests that Cdc6 interacts with ORC (Liang *et al.*, 1995). Cdc6 is an unstable protein which appears in late mitosis and disappears as cells enter S phase (Drury *et al.*, 1997; Piatti *et al.*, 1995). This pattern of expression coincides with the time when pre-replicative complexes are present on the yeast replication origins. Cdc6 was directly shown to be a component of the pre-replicative complexes by releasing yeast cells from a mitotic block with the expression of the Cdc6 gene being turned off. When cells exit mitosis and enter G1 in the absence of Cdc6, complexes formed on replication origins, monitored by genomic footprinting, closely resembled the pre-replicative complexes suggesting that Cdc6 is required for the pre-replicative complex assembly (Cocker *et al.*, 1996). Cdc6 was also shown to bind to chromatin during G1 (Donovan *et al.*, 1997) and to interact with yeast origin DNA (Tanaka and Nasmyth, 1998). Taken together, these result implicate Cdc6p in the formation of pre-replicative complexes and that it plays essential role in the regulation of origin firing in *S. cerevisiae*.

The essential role of Cdc6 in regulation of DNA replication was confirmed in a number of other higher eukaryotes. A *S. pombe* Cdc6 homologue is known as cdc18 and is 28% identical to its budding yeast counterpart (Kelly *et al.*, 1993). The overexpression of cdc18 drives *S. pombe* cells into S phase, while deletion causes an

arrest with unreplicated DNA (Kelly *et al.*, 1993; Muzi-Falconi *et al.*, 1996; Nishitani and Nurse, 1995). The cell cycle profile of cdc18 expression is also consistent with its role at G1/S boundary (Kelly *et al.*, 1993; Muzi-Falconi *et al.*, 1996; Nishitani and Nurse, 1995).

A *Xenopus* Cdc6 gene was identified by degenerate PCR based on the homology between the *S. cerevisiae* and *S. pombe* genes and was found to be 39% identical to the *S. cerevisiae* gene (Coleman *et al.*, 1996). Immunodepletion of Cdc6 from *Xenopus* egg extract inhibited DNA replication of added sperm chomatin, but not the replication of single stranded DNA, implying a role in the initiation process. Unlike yeast, the protein levels of Cdc6 in *Xenopus* do not oscillate during the embryonic cell cycle. Interestingly, Cdc6 associates with chromatin upon entry into interphase in *Xenopus* extract and this association is dependent on ORC (Coleman *et al.*, 1996). Thereafter, Cdc6 is eliminated from chromatin during the replication process and is found associated with the nuclear envelope, a feature similar to *E. coli* DnaA association with the lipids of the bacterial plasma membrane (Coleman *et al.*, 1996; Sekimizu and Kornberg, 1988). Human Cdc6 was cloned by PCR based on the sequence of the *Xenopus* gene (Williams *et al.*, 1997). The Cdc6 proteins levels are relatively constant throughout the HeLa cell cycle (Fujita *et al.*, 1999). Instead, the subcellular localization of cdc6 is regulated. In G1, cdc6 is nuclear and re-localises to the cytoplasm during S phase (Fujita *et al.*, 1999; Petersen *et al.*, 1999).

MCM proteins interact with origins prior to initiation

The MCM proteins are a family of related proteins required for initiation of DNA replication. The MCM proteins were first identified in *S. cerevisiae* in genetic screens for mutations defective for plasmid maintenance (MCM-minichromosome

maintenance proteins) or cell cycle progression (Maine *et al.*, 1984; Moir *et al.*, 1982). MCM genes have been identified in a wide range of eukaryotes, including yeast, frogs and humans, and based on sequence homology are clustered into six groups called MCM2-7 (Chong *et al.*, 1996). The 240-aminoacid region of high conservation between MCM proteins from the six groups include a modified match to the Walker A motif of the nucleotide binding site (Koonin, 1993; Walker *et al.*, 1982). However, different members of MCM family are quite diverse outside this 240 amino acid region.

Studies of a variety of eukaryotes suggests that MCM proteins function together in a large multi-protein complex. In *S. cerevisiae,* individual MCM proteins were shown to interact with each other both genetically and biochemically (Hennessy *et al.*, 1991; Lei *et al.*, 1996). Interaction between individual MCM subunits was also demonstrated in *S. pombe, Drosophila, Xenopus* and mouse (Chong *et al.*, 1995; Hendrickson *et al.*, 1996; Kimura *et al.*, 1995; Kubota *et al.*, 1995; Madine *et al.*, 1995a; Madine *et al.*, 1995b; Okishio *et al.*, 1996; Romanowski *et al.*, 1996a; Su *et al.*, 1996).

The levels of MCM proteins are relatively constant throughout the cell cycle. However, the localization of MCM proteins in *S. cerevisiae* is significantly altered as cells replicate their DNA. MCM proteins enter the nucleus at the end of mitosis, persist there throughout G1 and are rapidly exported into cytoplasm soon after initiation of DNA replication (Dalton and Whitbread, 1995; Hennessy *et al.*, 1990; Yan *et al.*, 1993). Once inside the nucleus, a fraction of MCM proteins becomes tightly associated with DNA (Yan *et al.*, 1993). Chromatin binding experiments clearly demonstrated that less that 5% of MCM5 (Cdc46p) and MCM7 (Cdc47p) are associated with chromatin at the G2/M transition, while about half of the total cellular

MCM5 and MCM7 is associated with chromatin in G1 (Donovan *et al.*, 1997). Similar results were obtained for the *S. cerevisiae* MCM3 protein, which was shown to be bound to chromatin in G1, becoming displaced during S phase (Liang and Stillman, 1997).

In vertebrate and fission yeast cells the nuclear localisation of MCM proteins does not seem to be cell cycle regulated. However, MCM association with the insoluble chromatin is under cell cycle control. MCM proteins associate with sperm chromatin added to *Xenopus* egg extract and are gradually released from chromatin as replication proceeds (Chong *et al.*, 1995; Hendrickson *et al.*, 1996; Kubota *et al.*, 1995; Madine *et al.*, 1995a; Madine *et al.*, 1995b; Romanowski *et al.*, 1996a).

The timing of the association of the MCM proteins with chromatin implies their presence in pre-replicative complexes. Indeed, *in vivo* cross-linking studies in *S. cerevisiae* showed MCM4 (Cdc54p) and MCM7 (Cdc47p) associated specifically with origin DNA during G1, supporting the idea that they are a part of pre-replicative complexes (Aparicio *et al.*, 1997; Tanaka *et al.*, 1997). During S phase, MCM proteins associate with non-origin DNA, suggesting that after initiation they move with the replication forks (Aparicio *et al.*, 1997). The participation of MCM proteins in the pre-replicative complexes was further supported by experiments showing that MCM association with chromatin was prevented in cells which carried a mutation in the Orc1 or Orc5 gene (Aparicio *et al.*, 1997; Tanaka *et al.*, 1997).

In addition to functional ORC, association of MCM proteins with chromatin requires the presence of Cdc6 protein. In the absence of Cdc6 expression in *S. cerevisiae*, MCM proteins are unable to bind chromatin (Donovan *et al.*, 1997). Similarly, in Cdc6 mutant cells association of MCM proteins with the origin DNA is reduced

(Aparicio *et al.*, 1997; Tanaka *et al.*, 1997). Despite the requirement for functional ORC and Cdc6 for loading of the MCM proteins onto chromatin, these proteins are not required for the maintenance of the MCM-chromatin interaction, as MCM proteins remain chromatin bound following salt extraction of ORC and Cdc6 (Donovan *et al.*, 1997).

Similar results were obtained in the *Xenopus* system, where binding of Cdc6 to chromatin was shown to require ORC to bind first (Coleman *et al.*, 1996). Loading of MCM3 required both ORC and Cdc6 proteins to be pre-bound (Coleman *et al.*, 1996; Romanowski *et al.*, 1996b; Rowles *et al.*, 1996). Thus, the assembly of replication competent chromatin in *Xenopus* involves sequential binding of ORC, Cdc6 and MCMs to DNA. Interestingly, salt extraction of ORC and Cdc6 from chromatin, following the loading of the MCM proteins does not inhibit DNA replication, suggesting that loading of MCMs might be the essential function of ORC and Cdc6 (Rowles *et al.*, 1999).

Despite their importance in the process of DNA replication, the precise function of MCM proteins remains unclear. MCM complexes isolated from HeLa cells had both ATPase and helicase activities, suggesting a role as a replicative helicase (Ishimi, 1997; Koonin, 1993). The finding that MCMs, although initially associated with replication origin sequences in *S. cerevisiae*, associate with chromatin distal to replication origins as replication proceeds supports the idea that MCM proteins are associated with replication forks (Aparicio *et al.*, 1997).

Interestingly, MCM proteins were identified as a part of a licencing factor, which is required for initiation, but is inactivated or destroyed as DNA replication progresses (Blow and Laskey, 1988; Chong *et al.*, 1995). This suggests a possible role for

MCMs in distinguishing the replicated and unreplicated DNA and restricting DNA replication to once per cell cycle. In agreement with this model, in replicating HeLa nuclei, localisation of MCM proteins coincides with subnuclear sites of unreplicated chromatin. MCM proteins are displaced from their chromatin sites at the time when these sites are replicated (Krude *et al.*, 1996).

CDC45 protein is also found in the pre-replicative complex

The *S. cerevisiae* CDC45 gene was isolated by complementation of the cold sensitive *cdc45-1* mutant which showed a phenotype typical for a gene functioning in DNA replication. CDC45 protein levels are constant throughout the cell cycle and unlike MCM proteins, CDC45 is always in the nucleus and does not change localisation as a function of the cell cycle (Hopwood and Dalton, 1996; Owens *et al.*, 1997).

CDC45 was shown to interact genetically with the MCM and ORC proteins, confirming its role in DNA replication (Hardy, 1997; Hennessy *et al.*, 1991; Zou *et al.*, 1997). The interaction between CDC45 and MCM proteins was also demonstrated by immunoprecipitation, yeast two hybrid assays and gel filtration (Hopwood and Dalton, 1996; Dalton and Hopwood, 1997). 2D gel electrophoresis analysis showed that initiation of DNA replication was defective in cells carrying a CDC45 mutation (Zou *et al.*, 1997). The transition from pre-replicative complexes to post-replicative complexes was also shown to be impaired in CDC45 mutants implying a problem in initiation (Owens *et al.*, 1997).

In vivo cross-linking studies demonstrated that like MCM proteins, CDC45 is also a component of the pre-replicative complexes; is associated with origins in G1 and moves with the replication forks following the initiation of DNA replication

(Aparicio *et al.*, 1997). However, the precise function of CDC45 in the initiation process remains unknown.

CDC7/DBF4 protein kinase is required for origin firing

The CDC7 gene was identified as a gene required for DNA replication in *S. cerevisiae* likely to be involved in the initiation process (Hartwell, 1973; Hartwell, 1978). Comparison of the amino acid sequence of the CDC7 protein with the sequence of *S. cerevisiae* CDC28 protein suggested that CDC7 is a protein kinase (Patterson *et al.*, 1986). Indeed, CDC7 immune complexes were able to phosphorylate histone H1 and the more likely physiological target MCM2 (Hollingsworth and Sclafani, 1990; Oshiro *et al.*, 1999; Yoon and Campbell, 1991). Although CDC7 protein levels are constant throughout the cell cycle, the associated kinase activity is periodic, reaching the maximum at the G1/S boundary (Jackson *et al.*, 1993; Yoon *et al.*, 1993).

A possible explanation for the periodic kinase activity of CDC7 is an association with a cyclin like regulatory subunit. A candidate regulatory subunit of CDC7 is DBF4 protein. DBF stands for dumbbell former which reflects its terminal phenotype of a cell with a single bud. This dumbbell phenotype is characteristic of *S. cerevisiae* DNA replication mutants and is identical to that found for temperature sensitive CDC7 mutants (Jackson *et al.*, 1993). DBF4 was originally identified during a search for replication mutants in *S. cerevisiae* and, like CDC7, has been shown to be involved in the initiation of DNA replication (Johnston and Thomas, 1982).

Genetic and biochemical studies demonstrated a clear link between CDC7 and DBF4. Over-expression of DBF4 suppresses mutations of CDC7 and CDC7 on a multi-copy plasmid suppresses DBF4 mutation (Kitada *et al.*, 1992). CDC7 and DBF4 interact

in a two-hybrid assay and DBF4 function is needed for CDC7 kinase activity on histone H1 (Dowell *et al.*, 1994; Jackson *et al.*, 1993). In contrast to CDC7, whose protein levels remain constant in the cell cycle, DBF4 protein levels are low during G1, rise as cells enter S phase and decline during transition into the next cell cycle (Chapman and Johnston, 1989; Oshiro *et al.*, 1999). DBF4 protein levels correlate with the activity of CDC7 kinase, implicating DBF4 as a potential CDC7 regulatory factor. However, DBF4 levels persist for longer then CDC7 kinase activity, arguing for at least one further level of control. This could be phosphorylation of CDC7 which correlates with CDC7 kinase activity (Yoon and Campbell, 1991; Yoon *et al.*, 1993). Together, these data suggest that active CDC7 kinase is a heterodimeric CDC7-DBF4 complex. In one-hybrid interaction assay, DBF4 was found to interact with the ARS1 origin sequence in *S. cerevisiae* suggesting that DBF4 may direct CDC7 kinase to the origin (Dowell *et al.*, 1994). Origins of DNA replication fire asynchronously throughout S phase. Interestingly, CDC7-DBF4 protein kinase is required for firing of individual replication origins throughout S phase, rather than for the S phase entry (Bousset and Diffley, 1998; Donaldson *et al.*, 1998). Presumably, at the origin, the CDC7-DBF4 complex phosphorylates key targets assembled in the pre-replication complex to trigger initiation of DNA replication. The most likely targets of the CDC7/DBF4 protein kinase are the members of the MCM protein family (Hardy *et al.*, 1997; Lei *et al.*, 1997).

Homologues of CDC7 and DBF4 were identified in *S. pombe*, mouse, *Xenopus* and human confirming the universal role of this protein kinase in initiation of DNA replication (Brown and Kelly, 1998; Jiang and Hunter, 1997; Kim *et al.*, 1998; Kumagai *et al.*, 1999; Landis and Tower, 1999; Masai *et al.*, 1995; Sato *et al.*, 1997).

CDKs function to promote S phase and prevent re-replication

In eukaryotes, DNA replication initiates from multiple origins which fire at different times during S phase (Friedman *et al.*, 1995; Huberman and Riggs, 1968). As it is essential to ensure that all DNA sequences are replicated precisely once in each cell cycle, none of the large number of the replication origins must fire more than once per cell cycle.

The mechanism ensuring that origins fire no more than once in a cell cycle was extensively studied in a number of species. In the classic set of experiments HeLa cells in different stages of the cell cycle were fused. Fusion of cells in G1 with cells in S phase induced DNA replication in the G1 nucleus, while fusion of G2 cells with cells in S phase did not induce DNA synthesis in the G2 nucleus (Rao and Johnson, 1970). These experiments indicated that a diffusible S phase activator can only act on G1 nuclei which are competent to replicate and not on nuclei in G2.

The recently described pre-replicative complexes assembled on the replication origins in G1 in *S. cerevisiae* can explain the replication potential of G1, but not G2 nuclei (Diffley *et al.*, 1994). Only in G1, pre-replicative complexes are assembled on the replication origins, and only those origins are competent to fire. Following initiation of DNA replication from a particular origin, the pre-replicative complex is disassembled and the origin is incompetent to initiate another round of DNA replication. New round of DNA replication can only initiate from this origin following the re-setting of pre-replicative complex.

Attention therefore turned towards understanding the mechanism which allows pre-replicative complexes to assemble on replication origins in G1 phase and prevents their re-setting within the same cell cycle. To ensure that DNA replication occurs

precisely once per cell cycle a mechanism must link the process of origin licencing to the control of the cell cycle.

In *S. cerevisiae*, CDC28 protein kinase is the major regulator of the cell cycle. The CDC28 protein kinase in association with Clb1-6 B-type cyclins is absolutely required for the initiation of DNA replication as cells in which the Clbs are not functional fail to enter S phase (Schwob *et al.*, 1994).

Replication origins in cells arrested with nocodazole at the G2/M boundary are normally in a post-replicative state. However, expression of SIC1 protein, which is a specific inhibitor of CDC28/Clb kinases caused formation of pre-replicative complexes in these cells. This suggests that the kinases responsible for origin firing, function to prevent re-setting of the pre-replicative complexes later on in the cell cycle (Dahmann *et al.*, 1995; Mendenhall, 1993; Schwob *et al.*, 1994). Indeed, during G1–phase, when pre-replicative complexes are formed, SIC1 accumulates and only disappears shortly before S phase (Schwob *et al.*, 1994). These results provide a simple model for restricting origin firing to once per cell cycle. The pre-replicative complexes assemble during G1, when CDC28/Clb activity is low. The activation of CDC28/Clb kinase drives cells into S phase, however, the high CDC28/Clb kinase prevents pre-replicative complexes reforming after origin firing. Pre-replicative complexes can only be reset when CDC28/Clb activity drops following the exit from mitosis.

S. cerevisiae CDC6 protein is required for the formation of pre-replicative complexes and for initiation of DNA replication (Cocker *et al.*, 1996). CDC6 protein can only promote DNA replication in a restricted window in the cell cycle between the destruction of Clbs after anaphase and before activation of Clb5/CDC28 and

Clb6/CDC28 in late G1 (Piatti *et al.*, 1996). When CDC6 is expressed only later in the cell cycle, following the activation of the Clb/CDC28 kinases, it is no longer able to induce pre-replicative complex formation (Piatti *et al.*, 1996). Conversely, when Clb2 is expressed in G1, prior CDC6 expression, formation of pre-replicative complexes is also inhibited (Detweiler and Li, 1998). Presumably, Clb/CDC28 kinases prevent resseting of the replication origins by preventing CDC6 forming pre-replicative complexes.

Experiments in *S. pombe* provided further evidence that mitotic kinases serve to prevent re-replication. In *S. pombe*, CDC2 is the major cell cycle kinase and cdc13 encodes the single B-type cyclin. Inactivation of CDC2 in cells arrested in G2 and consequent re-activation of the kinase led to re-entry in to S phase rather mitosis (Broek *et al.*, 1991). Presumably, pre-replicative complexes re-assembled at the replication origin during the low kinase activity and subsequent elevation of the kinase triggered the origin firing. Consistent with this interpretation, deletion of the CDC13 B-type cyclin also causes re-replication (Hayles *et al.*, 1994). Over-expression of CDC2/CDC13 protein kinase in cells arrested in G1 induced entry into mitosis, without going through S phase (Hayles *et al.*, 1994). Finally, overexpression of RUM1, a specific inhibitor of CDC2/CDC13 protein kinase, also induced re-replication (Correa-Bordes and Nurse, 1995; Moreno and Nurse, 1994).

In *Xenopus* egg extracts, entry into mitosis is controlled by activation of CDC2 protein kinase which associates with mitotic cyclins. However, depletion of CDC2 kinase did not inhibit DNA replication (Fang and Newport, 1991). Instead, a related protein kinase CDK2, which is 66% identical to CDC2 is normally required for the initiation of S phase. The *Xenopus* egg extracts from which CDK2 kinase was depleted failed to initiate replication, but were able to elongate previously initiated

replication forks. Re-addition of the CDK2 kinase restored the extracts' ability to initiate DNA replication (Blow and Nurse, 1990; Fang and Newport, 1991).

In *Xenopus* egg extracts, cyclin E is the major partner of CDK2 (Jackson *et al.*, 1995). Immunodepletion of cyclin E from *Xenopus* egg extract led to the inhibition of DNA synthesis and this inhibition was overcome by addition of cyclin E protein (Jackson *et al.*, 1995). Interestingly, addition of cyclin A could also rescue DNA replication, suggesting that although CDK2/cyclin E is responsible for triggering the initiation of DNA replication, CDK2/cyclin A can substitute. However, addition of cyclin B was unable to rescue the replication defect (Jackson *et al.*, 1995).

The CIP1 protein is a specific inhibitor of cyclin dependent kinases, with preference for CDK2/cyclin E and CDK2/cyclin A (Jackson *et al.*, 1995). CIP1 added to *Xenopus* egg extract strongly inhibited chromosomal replication, presumably due to binding and inactivation of CDK2/cyclin E (Chen *et al.*, 1995; Jackson *et al.*, 1995; Strausfeld *et al.*, 1994). This inhibition could again be relieved by addition of cyclins A or E, but not cyclin B (Strausfeld *et al.*, 1994).

The requirement for cyclin dependent kinases for initiation of DNA replication was further examined in interphase *Xenopus* egg extracts from which cdk's and associated cyclins were depleted. In this assay CDK2/cyclin E, CDK2/cyclin A and CDC2/cyclin A complexes were all able to promote initiation (Strausfeld *et al.*, 1996). In a similar set of experiments, both CDC2 and CDK2 were able to rescue replication (Chevalier *et al.*, 1995). Experiments in cell free system derived from HeLa cells confirmed the results obtained using *Xenopus* egg extracts. G1 nuclei in this system could be induced to enter S phase by CDK2/cyclin E and CDK2/cyclin A complexes, but not by CDC2/cyclin B (Krude *et al.*, 1997).

The emerging view is that CDK2/cyclin E kinase is responsible for initiating DNA replication in higher eukaryotes. Consistent with this idea, cyclin E was found to be associated with replicating S phase nuclei in the *Xenopus* system. Under limiting cyclin E conditions, only the replicating nuclei showed cyclin E staining, suggesting that cyclin E was essential for replication (Chevalier *et al.*, 1996). While cyclin A associated kinases can substitute for the CDK2/cyclin E function in promoting S phase, CDK2/cyclin E normally performs this function.

Similarly to the mechanism preventing re-replication in yeast, CDK2/cyclin E and CDK2/cyclin A kinases are implicated in having dual function. In addition to their function in S phase promotion, these kinases prevent replication if they are elevated prior proteins essential for DNA replication bind chromatin. In particular, MCM3 protein failed to bind chromatin in the presence of high levels of CDK2/cyclin E (Hua *et al.*, 1997). The mechanism which restricts initiation of DNA replication to only once per cell cycle is therefore likely to be conserved among all eukaryotic systems.

The issues addressed in this thesis

The work described in this thesis started at the time of identification of *Xenopus* Orc1p. The existence of ORC homologues in higher eukaryotes had not been reported at that time. The finding that an ORC, similar to that recognising origins of DNA replication in yeast, is required for initiation in higher eukaryotes challenged how we view initiation process in higher eukaryotes. The identification of ORC in frogs, where no replication origins as such had been identified, poses the question: what does the *Xenopus* ORC recognise? This thesis aims to address this question by studying the ORC itself.

Chapter 2

Materials and Methods

Plasmids

pET21b-Orc1

The 2661 bp *Xho*I/*Nde*I DNA fragment carrying the full length *Xenopus* Orc1 cDNA obtained from Alison Rowles from Julian Blow's lab was subcloned into *Xho*I/*Nde*I digested pET21b DNA (Novagen; Figure 2/1). This was to generate a hexahistidine tag at the C-terminus of the protein. The advantage of a C-terminal tag is that polypeptides produced by a partial translation of the fusion protein do not have the tag and therefore do not co-purify with the full-length protein in the nickel agarose affinity step.

pET21b-Orc1 plasmid map
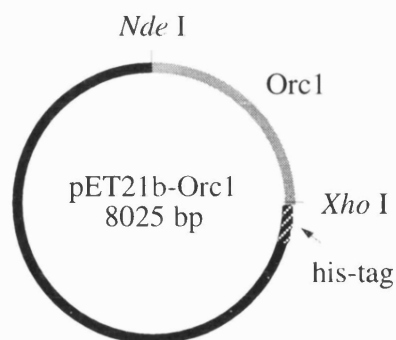
Figure 2/1. The Orc1 cDNA subcloned into *Nde* I and *Xho* I sites of the pET21b plasmid generates a C-terminal hexahistidine fusion.

pET21b-Orc4

The 1296 bp DNA fragment coding for *Xenopus* Orc4p gene was amplified from the Bluescript plasmid rescued from λ-ZAP library clones (see Chapter 4) by PCR. Primers used (see below) added an *Xho* I restriction site to both ends of the amplified fragment. The amplified fragment was then subcloned into the *Xho* I cloning site of the pET21b expression vector (NOVAGEN), with the resulting sequence coding for C-terminally hexahistidine tagged Orc4p.

3' primer: GTC TGT CTC GAG ATG AGC AAA CGA AAG T

5'primer: CCT TCC CTC GAG GCT GAG TGA GGA CAT C

pET16b-Orc1ΔN280 and pET16b-Orc1ΔN423

The pET16b-Orc1ΔN280 and pET16b-Orc1ΔN423 plasmids were kindly provided by Alison Rowles from Julian Blow's lab.

pET16b-Orc1N-280, pET16b-Orc1N-270 and pET16b-Orc1N-260

To make pET16b-Orc1N-280, pET16b-Orc1N-270 and pET16b-Orc1N-260 plasmids, the appropriate N-terminal fragments of Orc1 coding sequence were produced by PCR and then subcloned into the *Nde* I–*Xho* I cloning sites of pET 16b vector. The 5' primer used was identical for all three truncations produced, while the 3' primers were specific for the individual truncations (see below). The 5' primer provided *Nde* I restriction site and the 3' primers added an *Xho* I site.

5' primer: ATC GAA GGT CGT CAT ATG ACC

3' primers: GGA TTC CTC GAG TTT GAA TTC CTC ATC C

                  TGT GGT CTC GAG TTG CTC AAG GCT AGA G

                  TTC AGA CTC GAG TGG ACT ACT TAG TTG C

pET24d-Orc2

The pET24d-Orc2 plasmid was provided by Phil Carpenter from Bill Dunphy's lab.

pARS1-1.4.1

The pARS1-1.4.1 plasmid was kindly provided by Lucy Drury from John Diffley's lab.

Bacterial strains

NovaBlue (DE3)	*endA1 hsdR17*(r_{K12}^{-} m_{K12}^{+}) *supE44 thi-1 recA1 gyrA96 relA1 lac* [F' *proA*$^+$*B*$^+$ *lacI*^{q} *ZΔM15* ::Tn *10*] (DE3)
XL1-Blue	*supE44 hsdR17 recA1 endA1 gyrA46 thi relA1 lac*$^-$ F' [*proAB*$^+$ *lacI*^{q} *lacZΔM15* Tn *10(tet*^{r}*)*]
SURE	e14$^-$(McrA$^-$) Δ(*mcrCB-hsdSMR-mrr*)*171 endA1 supE44 thi-1 gyrA96 relA1 lac recB recJ sbcC umuC*::Tn*5* (Kan^r) *uvrC* [F' *proAB lacI*^{q}*ZΔM15* Tn*10* (Tet^r)]
LE 392	e14$^-$(McrA$^-$) *hsdR514 supE44 supF58 lacY1 galK2 galT22 metB1 trpR55*

Media

L-broth	10 g Bactotryptone; 5 g Yeast Extract; 10 g NaCl per litre
SOC	20 g Bactotryptone; 5 g Yeast Extract; 0.6 g NaCl; 0.2 g KCl; 2 g MgCl$_2$; 2.5 g MgSO$_4$; 3.6 g D-glucose per litre
2xTY	10 g Bactotryptone; 10 g Yeast Extract; 5 g NaCl per litre
NZY	5 g Yeast Extract; 5 g NaCl; 2 g MgCl$_2$; 10 g NZ amine per litre

Standard molecular biology

Plasmid minipreparation

For plasmid minipreparation, a QIAGEN miniprep kit was used, following the manufacturer's instruction.

Purification of DNA fragments from agarose gels

To extract DNA fragments from agarose gels, QIAGEN gel extraction kit was used following the manufacturer's instruction.

Transformation of bacteria by calcium chloride method

Transformation competent cells were prepared first. A single colony of an appropriate *E. coli* strain was picked from a fresh plate and was used to inoculate 5 ml of 2xTY media to produce an overnight culture. 500 ml of L-broth was inoculated with 1 ml of the overnight culture and incubated at 37°C, ensuring vigorous aeration, until an optical density reading at A_{600} of 0.5 was reached. The culture was then chilled in an ice bath and centrifuged for 10 min at 5 000 rpm in Sorvall GSA rotor. The obtained pellet was resuspended in 30 ml of ice cold 100 mM $CaCl_2$ and kept on ice for 30 min. Cells were then centrifuged at 5 500 rpm for 7 min in Sorvall SS 34 rotor and the pellet resuspended gently in 5 ml of ice-cold 100 mM $CaCl_2$; 15% glycerol and aliquots frozen at -70°C.

To transform prepared cells 2-5 µl of transforming DNA (0.01-1 µg) was added to 22.5 µl of 5 mM $MgCl_2$; 5 mM Tris-Cl, pH 7.5, followed by 50 µl of the competent cells. After vortexing, tubes were incubated on ice for 1 hour, followed by 5 min incubation in 37°C waterbath and brief chilling on ice. 200 µl of L-broth was then

added and tubes were incubated at 37°C for one hour. Finally, aliquots of the transformed cells were spread on appropriate selection plates.

Electroporation of bacteria

To make electro-competent cells, a single bacterial colony from a freshly streaked agar plate was picked and inoculated into L-broth to grow an overnight culture. On the next day, 1 litre of L-broth was inoculated with 10 ml of the overnight culture and grown at 37°C with good aeration until an A_{600} between 0.5 and 1 was reached. The culture was then placed on ice for 30 min and all subsequent manipulations were done at 4°C. Cultures were then centrifuged at 4000g for 6 min and the supernatant was discarded. The cell pellet was resuspended in 1 litre of ice cold 0.1 mM HEPES and the centrifugation step was repeated. The obtained pellet was resuspended in 500 ml of ice cold 0.1 mM HEPES and cells were harvested once again and resuspended in 20 ml of ice cold 10% glycerol in water. Cells were then harvested for the last time, resuspended in 3 ml of 10% glycerol, aliquoted, frozen on dry ice and stored at -70°C.

Before electroporation, the 0.2 cm electroporation cuvettes were chilled on ice and the electroporation apparatus was set at 25 μF capacitance and 2.5 kV. The pulse controller unit was set to 200 Ω. An aliquot of previously prepared electrocompetent cells was thawed on ice and 45 μl of cells were mixed with 1-5 μl of DNA and incubated on ice for 1 minute. The mixture was then transferred into the cuvette and a pulse was applied at the above settings. 1 ml of SOC media was immediately added and the mixture was then incubated at 37°C for one hour. Appropriate dilutions were then plated on selective agar plates.

Manual DNA sequencing

For manual DNA sequencing reactions 1.5 µg of plasmid DNA was incubated with

4.5 pmol of primer in the presence of 100 mM of NaOH in a total volume of 10 µl for

10 min at 68°C. Then 4 µl of TDMN (280 mM TES; 120 mM HCl; 80 mM MgCl$_2$;

200 mM NaCl; 50 mM DTT) was added and the reaction was incubated for 10 min at

37°C. Tubes were then placed on ice and 5.5 µl of pre-mixed 1 µl of 0.1 M DTT;

0.4 µl of 5x Sequenase Reaction Buffer (200 mM Tris pH 7.5; 100 mM MgCl$_2$;

250 mM NaCl); 0.4 µl of 5x Labeling Mix (7.5 µM dGTP; 7.5 µM dCTP; 7.5 µM

dTTP); 0.5 µl [^{35}S] dATP (10 mCi/ml); 1.2 µl water, and 2 µl Sequenase (UBS),

diluted 1:8 in Enzyme Dilution Buffer (10 mM Tris pH 7.5; 5 mM DTT; 0.5 mg/ml

BSA) was added followed by 3 min incubation at room temperature. For each

sequencing reaction, four tubes were labelled A, C, G and T and 2.5 µl of appropriate

Termination Mix (80 µM dGTP; 80 µM dATP; 80 µM dCTP; 80 µM dTTP and 8 µM

of either ddGTP, ddATP, ddCTP, ddTTP) was added to each tube. Then, 3.5 µl of

the sequencing reaction described above was added to each tube and samples were

incubated at 37°C for 3 min. The reactions were terminated by addition of 4 µl of

Stop Solution (95% Formamide; 20 mM EDTA; 0.05% Bromophenol Blue; 0.05%

Xylene Cyanol FF) into each tube and samples were then stored on ice. Sequencing

reactions were run on 6% denaturating polyacrylamide gel, with denaturation of the

samples at 90°C for three minutes before loading.

Automatic DNA sequencing

Automatic DNA sequencing is based on repeated cycles of linear PCR in conjunction

with dye terminators. For this protocol, DNA Sequencing Kit Big dye terminator-

ABI PRISM (PE Applied Biosystems) was used. Briefly, 500 ng of plasmid DNA,

3.2 pmole of primer and 8 µl of Terminator Ready Reaction Mix in total volume of

20 μl were placed into PTC-200 Peltier Thermal Cycler and cycled 25 times as

follows: 96°C for 30 seconds

               50°C for 15 seconds

               60°C for 4 minutes.

Samples were then mixed with 30 μl of water, 5 μl of 3 M Sodium acetate, pH 4.6

and 137.5 μl of 100% ethanol, incubated on ice for 10 min and centrifuged in a bench

top centrifuge at 4°C for 30 min. The obtained pellet was washed with 70% ethanol,

dried and sent to Graham Clark in the ICRF Equipment Park for sequence analysis.

Polymerase Chain Reaction (PCR)

The standard PCR reaction used during these thesis was performed in 100 μl

volumes. The amount of DNA used for PCR reactions varied depending on the

experiment. To "PCR-out" a DNA fragment of interest from a plasmid (see

production of pET21b-Orc4, pET21b-Orc1N-280, pET21b-Orc1N-270 and pET21b-

Orc1N-260 described earlier in this chapter) 100-200 ng of DNA substrate was used.

On the other hand, in experiments where ORC bound DNA from the WAK cells was

recovered by immunoprecipitation (see Chapter 8), the amount of DNA available as a

PCR template was extremely low. Primer concentration was usually, 100–200 nM of

each oligonucleotide. All PCR reactions contained 20 μM of each dNTP and 1 μl of

Taq DNA polymerase produced in the lab and were done in the presence of 10 mM

Tris, pH 8.3; 50 mM KCl; 1.5 mM $MgCl_2$ and 0.001% gelatin. PCR reactions were

placed into PTC-200 Peltier Thermal Cycler, and the machine was programmed in a

following way: denaturation step: 94°C for 2 min

                  cycling: 94°C for 30 sec

                  44-60°C (depending on the annealing temperature of individual

                  primers) for 30 sec

                  72°C for 1 min

In the case of PCR from a plasmid, 15 PCR cycles were done to minimise chances of introducing a mutation. However, PCR reactions in which extremely low amounts of DNA were used as a template, 30 PCR cycles were usually done to maximise the amount of product.

Cloning of *Xenopus* Orc4 gene

Making the probe

Xenopus Orc4 gene was identified on the bases of its homology to the mouse counterpart. The DNA probe for screening *Xenopus* cDNA library was produced by PCR amplification of a 473 bp fragment of the mouse Orc4 gene which was obtained in the form of EST clone AA110784 from the gene bank. The primers used for the PCR reaction are shown below and the PCR reaction was performed as described earlier in this chapter.

5' primer: CAT GCT TTG AAA GAA CTC AT

3' primer: GGT AGA GAT AAC TGT TCT TT

To generate a labelled probe the produced DNA fragment was then radiolabelled using eight reactions of Megaprime DNA labelling kit (Amersham) following the manufacturer's instructions and the unincorporated radionucleotides were removed using a G50 Sephadex spin column.

Library

For the cloning of the *Xenopus* Orc4 gene, a *Xenopus laevis* ovary cDNA library in λ-ZAP, kindly provided by Eva Dworkin from Vienna Biocentre was used. Before plating the library, the number of viable phages per ml was estimated to be

approximately 3.5x10^8 by titrating sequential dilutions of the library using the plating strain LE 392.

To plate the library for screening, a single colony of *E. coli* strain LE 392 from a fresh plate was inoculated into 50 ml of NZY media supplemented with 10 mM MgSO$_4$ and 0.2% maltose and incubated overnight at 30°C, shaking vigorously. Cells were then spun in SORVALL RT7 centrifuge for 10 min at 2 000 rpm, resuspended in 15 ml of 10 mM MgSO$_4$, diluted in MgSO$_4$ to A$_{600}$ 0.5 and kept on ice.

Six 50 ml Falcon tubes were used to perform the infection of the bacteria by phage. To each tube 1.7 ml cells and 1.5 µl of the phage library, previously diluted 1:10 in SM buffer (100 mM NaCl; 14.4 mM MgSO$_4$.7H$_2$O; 50 mM Tris, pH 7.5; 0.1 % gelatin) were added and incubated for 15 min at 37°C, shaking slowly.

50 ml of NZY media, supplemented with 0.7 % agarose and kept melted at 48°C, was added to each tube and the mixture was rapidly poured onto large pre-warmed and pre-dried (22 cm x 22 cm) NZY-agar plates and incubated at 37°C for 8 hours. Approximately 60 000 plaques (10 000 per plate) were obtained.

The plated phage library was cooled at 4°C for several hours. "Library lifts" were taken by placing dry Hybond-N (Amersham) membrane onto plates and allowing the phage to transfer to the membrane for 2 min, marking 5-10 spots for subsequent orientation. The membrane was then taken off the plate, turned "phage side up", and incubated for 5 min in 500 mM NaOH; 1.5 M NaCl, then for 5 min in 1 M Tris pH7.5; 1.5 M NaCl and another 5 min in fresh 1 M Tris pH7.5; 1.5 M NaCl and 5 min in 2xSSC (300 mM NaCl; 30 mM Na$_3$citrate.2H$_2$O pH 7.0). The filters were then dried for 15 min on 3 mm Whatman paper and baked at 80°C for 2 hours.

Hybridisation

Filters carrying the DNA from the library phages were first wetted in 2xSSC (300 mM NaCl; 30 mM Na$_3$citrate.2H$_2$O pH 7.0) and then incubated in prehybridisation solution 900 mM NaCl; 90 mM Na$_3$citrate.2H$_2$O pH 7.0; 0.1% SDS; 1 mg/ml Ficoll 400, 1 mg/ml polyvinylpyrrolidone, 1 mg/ml BSA; 5% Dextran sulphate and 100 µg/ml of sonicated salmon sperm DNA for 2 hours at 50°C. The previously prepared labelled probe was then denatured by 5 min boiling and added to the warm hybridisation solution. Hybridisation was carried on overnight and the filters were washed twice the next day with 2xSSC (300 mM NaCl; 30 mM Na$_3$citrate.2H$_2$O pH 7.0)/ 0.1% SDS for 30 min at 37°C. Washed filters were then sealed into plastic bags and exposed to film.

Second round screening

Six positive phages were identified in the initial library screen. These were cut out of the agarose layer and incubated in 1 ml of SM buffer (100 mM NaCl; 14.4 mM MgSO$_4$.7H$_2$O; 50 mM Tris, pH 7.5; 0.1% gelatin) for 2 hours at 4°C. Aliquots of eluted phages were then adsorbed to competent LE 392 cells, plated out on small bacterial plates and screened in essentially the same way as described above. Clones number 5 and 6 of the six initially identified phages were positive in this secondary screen and these two phage clones were isolated from the plates.

In vivo excision

To excise the Bluescript phagemid from the λ-ZAP vector, 200 µl of competent XL1-Blue cells were incubated with 200 µl of the isolated λ phage clones and with 2.5 µl of the σ408 helper phage for 15 min at 37°C in 50 ml Falcon tubes, shaking slowly. Then 5 ml of LB media was added and samples were incubated for 3 hours at 37°C,

heated for 20 min at 70°C and centrifuged in the SS 34 rotor for 10 min. The resulting supernatant was stored as phagemid stock. 200 µl of competent XL1-Blue cells were incubated with 10 µl of the phagemid stock for 15 min at 37°C and plated onto ampicilin agar plates to generate colonies of Bluescript plasmids carrying the cloned genes. DNA produced from these plasmids was digested with the *Eco*RI restriction enzyme, which was originally used to insert cDNA clones into the λ vector during the library preparation. The insert in the clone 5 was approximately 1.6 kb and the insert in clone 6 was about 1.55 kb. DNA sequencing of the obtained clones revealed that they encoded identical proteins but in the two isolated clones the cDNA's were inserted into the library vector in opposite directions. The insert in clone 5 was slightly longer due to extra untranslated sequence.

Separation and detection of proteins

SDS-polycrylamide gel electrophoresis

To prepare gels for protein sample separation either 15 % (15% polyacrylamide; 0.09% bis-acrylamide; 375 mM Tris pH 8.0) or 12.5% polyacrylamide gels (12.5% polyacrylamide; 0.1% bis-acrylamide; 375 mM Tris pH 8.0) were poured and topped with stacking gel (5% polyacrylamide; 0.12% bis-acrylamide; 120 mM Tris pH 6.8). Ammonium persulphate and Temed were used to drive polyacrylamide polymerization. Protein samples were mixed with sample buffer to final concentration 2% SDS; 10% glycerol; 80 mM Tris pH 6.8; 5% mercaptoethanol; 0.001% bromophenol blue, loaded on prepared polyacrylamide gels and run using SDS-running buffer (384 mM glycine; 50 mM Tris; 0.1% SDS) at 30 mA, 160 V until the front dye reached the bottom edge of the gel.

Coomassie blue staining

After separation of loaded protein samples, polyacrylamide gels were incubated in Coomassie blue stain (0.5% Coomassie blue; 45% methanol; 10% acetic acid) for 3 min and destained in Destain solution (25% methanol; 7% acetic acid) for 1 h.

Silver staining

Silver staining of polyacrylamide gels was done using Rapid-Ag-Stain kit (ICN Radiochemicals).

Immunoblotting

After separation of loaded protein samples, polyacrylamide gels were briefly soaked in transfer buffer (24 mM Tris, 193 mM glycine, 10% methanol). Six sheets of Whatman 3 mm paper soaked in transfer buffer were put into Hoeffer semidry transfer apparatus followed by nitrocellulose also soaked in transfer buffer. The protein gel was put on top of the nitrocellulose followed by another six sheets of soaked Whatman paper. Transfer was done at approximately 1 mA/cm^2 of the nitrocellulose membrane for 1 hour. After transfer, the nitrocellulose was removed from the transfer apparatus and soaked in Ponceau S stain (0.2% Ponceau S; 3% TCA) for 3 min to assess the efficiency of the transfer and mark position of the molecular weight markers. The stain was washed off in PBS (136.87 mM NaCl; 2.68 mM KCl; 10.14 mM Na$_2$HPO$_4$; 1.76 mM KH$_2$PO$_4$ pH 7.4). Non-specific binding sites on the nitrocellulose were blocked by 45 min incubation in either 4% milk (Marvel)/TBSA (10 mM Tris pH 8.0; 150 mM NaCl; 0.5 % Tween-20) or 10% milk/2% yeast extract (Difco)/TBSA (for anti-Orc4p antibodies). The nitrocellulose was then placed into the appropriate dilution of the primary antibody in either 1% milk/TBSA (10 mM Tris pH 8.0; 150 mM NaCl; 0.5 % Tween-20) or 5% milk/2%

yeast extract/TBSA (10 mM Tris pH 8.0; 150 mM NaCl; 0.5 % Tween-20) for 1 hour at room temperature. Unbound primary antibody was removed by four washes with TBSA (10 mM Tris pH 8.0; 150 mM NaCl; 0.5 % Tween-20). Horse radish peroxidase (HRP) conjugated secondary antibody was applied in the appropriate dilution in 1% milk/TBSA for 1 hour at room temperature. Unbound antibody was removed by six washes with TBSA (10 mM Tris pH 8.0; 150 mM NaCl; 0.5 % Tween-20). ECL reagents were applied onto the nitrcellulose for 1 min and the chemiluminiscence was detected by exposure to the light sensitive film (Fuji).

Antibodies

Production of histidine-tagged recombinant proteins

Histidine tagged fusion proteins (full length Orc1p, Orc2p and Orc4p) cloned into either pET16b, pET21b, or pET24d vectors were produced in *E. coli* strain Nova Blue DE3 (Novagen). Cells containing the expression plasmid were grown at 37°C to A_{600} of about 0.7. Synthesis of recombinant proteins was then induced with 1 mM IPTG for 2h at 37°C. Expressed proteins were purified on nickel agarose beads (Qiagen) under either native or denaturing conditions.

To isolate the soluble fraction of the expressed recombinant protein, cells were lysed in presence of 20 mM Tris pH 8.0; 500 mM NaCl; 5 mM imidazole pH 8.0; 5mg/ml lysosyme; 2.5% NP-40 and briefly sonicated. After centrifugation at 25 000 g for 30 minutes at 4°C, the cell lysate supernatant was batch incubated with nickel agarose beads for 2 hours at 4°C. Beads were then washed in a column with the same buffer containing increasing concentration of imidazole (5 mM, 25 mM, 50 mM). Purified protein was eluted by raising the imidazole concentration to 150 mM.

To isolate the insoluble fraction of the expressed recombinant protein, inclusion bodies were dissolved in 6M guanidine hydrochloride; 20 mM Tris, pH 8.0; 500 mM NaCl; 5 mM imidazole, pH 8.0 and batch bound to nickel agarose beads. Beads were then washed with the same buffer containing increasing concentration of imidazole (5 mM, 25 mM, 50 mM) followed by a wash with 1 x PBS (136.87 mM NaCl; 2.68 mM KCl; 10.14 mM Na_2HPO_4; 1.76 mM KH_2PO_4 pH 7.4). Purified protein was released by boiling in 4% SDS; 40 mM EDTA; 80 mM Tris-Cl pH 6.8; 5% mercaptoethanol buffer and than dialysed against 0.2% SDS; 1xPBS.

Production of polyclonal antibodies

Anti-Orc1p, 2 and 4 rabbit polyclonal antibodies were raised against the histidine-tagged full length *Xenopus* Orc1p, Orc2p and Orc4p proteins purified under denaturing conditions as described above. Purified antigens were given to Del Watling in the Biological Resource Unit of ICRF to be injected into New Zealand white rabbits as described by Harlow and Lane (Harlow and Lane, 1988).

Generation of anti-Orc1p monoclonal antibodies

To raise monoclonal antibodies against *Xenopus* Orc1p, the Orc1p protein, purified under either native or denaturationg conditions was given to Del Watling in the Biological Resource Unit of ICRF to be injected into fifteen young adult female BALB/c mice as described by Harlow and Lane (Harlow and Lane, 1988). The three best-responding mice were boosted by intraperitoneal injection of 1.5 µg of native Orc1p protein 5 days before sacrifice. Spleenocytes were extracted from the spleens of these mice as described in Antibodies (Harlow and Lane, 1988), page 209 and used for the production of immortalised antibody producing hybridomas.

Sp2 myeloma cells were thawed from the liquid nitrogen stock six days before fusion and grown in Dulbecco's Modified Eagles Medium supplemented with 20 % of FCS (foetal calf serum) to obtain 10^7-10^8 of rapidly dividing cells at the day of fusion.

Fusion of the spleen cells with the Sp2 myeloma cells was done by stirring method as described in Antibodies (Harlow and Lane, 1988), page 211. After fusion the cells were resuspended in 200 ml of Dulbecco's Modified Eagles Medium supplemented with 20 % of FCS (foetal calf serum) and with 1:100 dilution of hypoxanthine (1.36 mg/ml; Sigma) and of azaserine (0.1 mg/ml; Sigma) and with 1:5000 dilution of interleukine-6 (4 µg/ml) and dispensed in 100 µl aliquots into wells of 96-well microtiter plates and allowed to grow at 37°C in a 7% CO_2 incubator.

Growing hybridomas were tested for the presence of anti-Orc1p antibodies by aseptically removing 50 µl of tissue culture supernatant from each well of the 96-well microtiter plates and using them for ELISA assay as described in Antibodies (Harlow and Lane, 1988), page 182-183. Positives were re-tested by a second round of ELISA screening and 40 hybridomas which produced antibodies recognising most strongly the recombinant Orc1p on the ELISA plates were selected for further characterisation.

To test these antibodies for their ability to immunoprecipitate Orc1p from *Xenopus* egg extract, 5 µl of extract was incubated with 165 µl of IP buffer (50 mM Tris, pH 7.5; 150 mM NaCl; 1% NP-40; 5 mM EDTA; 5 mM EGTA) plus 250 µl of tissue culture supernatant from each of the ELISA-positive hybridomas and 10 µl (packed volume) of protein G beads (Biorad) for 3 hours at 4°C. The beads were washed three times with 500 µl of IP buffer and bound proteins released by boiling in 10 µl of 2x sample buffer (4% SDS; 20% glycerol; 160 mM Tris, pH 6.8;

10% mercaptoethanol; 0.002% bromophenol blue). Obtained samples were analysed by SDS-polyacrylamide gel electrophoresis, and Orc1p protein was detected by immunoblotting using a polyclonal anti-Orc1p antibody. Cell lines producing antibodies which immunoprecipitated Orc1p were cloned as single cells by limiting dilution method similar to that described in Antibodies (Harlow and Lane, 1988), page 223, with the difference that no feeder cells were used. This procedure established fourteen cell lines producing antibodies capable of immunoprecipitating Orc1p from *Xenopus* egg extract.

Mapping of anti-Orc1p monoclonal antibodies

In order to map the antibody binding sites on Orc1p protein, full length Orc1p protein and two deletion mutants, $\Delta N\,280$ and $\Delta N\,423$ were separated by SDS-PAGE, transferred to nitrocellulose and cut into strips which were immunoblotted with each of the fourteen of the anti-Orc1p monoclonal antibodies produced. Antibodies TK1 and TK15 reacted exclusively with the full length protein and therefore must recognise epitopes within the first 280 residues of Orc1p. Antibodies TK22 and TK23 reacted with both full length protein and the $\Delta N\,280$ construct, but not with the $\Delta N\,423$ deletion mutant and therefore bind to the region of Orc1p protein between the residues 280 and 423. Finally, all ten remaining antibodies reacted with all three polypeptides and therefore recognise a protein sequence between residue 423 and the C-terminus of Orc1p.

Purification of ORC

Antibody coupling to beads

To couple monoclonal antibodies which were of IgG2 subclass to protein A beads (Biorad), 100 µl of protein A beads were mixed with 50 ml antibody tissue culture

supernatant and rotated at room temperature for 1 hour. Beads were then washed three times with PBS (136.87 mM NaCl; 2.68 mM KCl; 10.14 mM Na_2HPO_4; 1.76 mM KH_2PO_4 pH 7.4), two times with coupling buffer (27 mM Sodium tetraborate; 73 mM Boric acid, pH 9.0) once with 1% Dimethyl Pimelimidate Dihydrochloride (Sigma)/0.1M Sodium tetraborate, pH 8.5-9.0 and incubated in 1% Dimethyl Pimelimidate Dihydrochloride (Sigma)/0.1M Sodium tetraborate, pH 8.5-9.0 overnight at 4°C. The next day, beads were washed three times with coupling buffer, incubated with 1 ml of 1M Tris, pH 9.0 for 10 minutes, washed three times with storage buffer (3.2 mM Sodium tetraborate; 45.5 mM Boric acid, pH8.0) and stored in storage buffer at 4°C.

Immunoaffinity purification of *Xenopus* ORC

To perform immunoaffinity purification, anti Orc1p monoclonal antibodies TK1, TK15, TK37 and TK47 as well as control control anti-SV40 T-antigen monoclonal antibody 423 were separately covalently coupled to protein A beads. 12 µl of antibody-coupled beads were incubated with 400 µl of *Xenopus* egg extract and 240 µl of purification buffer (50 mM Tris pH 7.5; 150 mM NaCl; 0.5 % NP-40; 2.5 mM EDTA; 2.5 mM EGTA; 0.5 mM DTT), rotating at 4°C for 2 hours. The beads were then washed seven times with 1 ml of ice-cold purification buffer by careful pipeting. The purified proteins were released by boiling in 15 µl of 2x sample buffer (4% SDS; 20% glycerol; 160 mM Tris pH 6.8; 10% mercaptoethanol; 0.002% bromophenol blue)r. This procedure yielded 0.1-0.2 µg of each of the ORC subunits.

Protein Microsequencing

Protein microsequencing by Edman degradation

To prepare proteins for microsequencing by the Edman degradation method, samples

were run on the polyacrylamide gel using only freshly made electrophoresis buffers and well cleaned plates, gel aparatus and all containers. Gels were then stained with 0.05% Coomassie brilliant blue G* in 5% AcOH; 10% methanol for 15 min and destained with 5% AcOH; 10% MeOH. Finally, gels were soaked in water for 10 min, sealed into a plastic bag and sent to Ryuji Kobayashi in Cold Spring Harbur Laboratories for further analysis.

Protein microsequencing by Mass spectrometry

To prepare samples for microsequencing by the Mass spectrometry, polyacrylamide gels were run under clean conditions as described above and the separated protein were transferred onto Imobilon–P (Millipore) transfer membrane, pre-wetted in methanol. After transfer, membrane was washed twice for 15 min in water and dried. Dry membrane was stained in 30% methanol; 0.2% acetic acid; 0.005% sulforhodamine B for few seconds, washed in water for few seconds, dried and given to Darryl Pappin in the ICRF protein microsequencing department for analysis.

Frog techniques

Preparation of interphase *Xenopus* egg extract

The *Xenopus* egg extract used for ORC purification and for DNA replication assay was essentially prepared following the protocol described by Julian Blow (Blow, 1993).

Briefly, freshly laid *Xenopus* eggs were collected and dejellied for 5 min in a Cysteine buffer (2% cysteine HCl, pH 7.8; 1 mM EGTA) at room temperature. During this time, the bad eggs were removed. Eggs were then washed three times with Barth buffer (88 mM NaCl; 2 mM KCl; 1 mM $MgCl_2$; 0.5 mM $CaCl_2$; 15 mM

Tris pH 7.5) at room temperature and activated with 2 µg/ml calcium ionophore A23187 in the Barth buffer for 5 min. Eggs were then washed three with Barth buffer at room temperature and five times with Extraction buffer (50 mM KCl; 50 mM HEPES-KOH, pH 7.6; 5 mM $MgCl_2$; 2 mM DTT) at 4°C. Eggs were transferred to 15 ml tubes and centrifuged in the HB-4 swing-out rotor at 3000 rpm for 30 sec at 4°C to pack the eggs. The buffer was then removed, and so were bad eggs which floated on top. Packed eggs were then crushed by centrifugation in the same rotor at 12 000 rpm for 10 min at 4°C. The cytoplasm (the middle layer) was then transferred into fresh tubes and supplemented with 10 µg/ml cytochalasin B and Dilution buffer (50 mM KCl; 50 mM HEPES-KOH pH 7.6; 0.4 mM $MgCl_2$; 2 mM DTT; 0.4 mM EDTA; 1 µg/ml pepstatin; 1 µg/ml leupeptin; 1 µg/ml aprotinin; 10% sucrose) of a volume equal to 15% of the original volume of the sample was added. Samples were centrifuged in an SW55 rotor for 15 min at 20 000 rpm at 4°C. The clear, golden layer was removed and passed through 25 µm Nylon membrane. Obtained extract was suplemented with 1% glycerol and stored at -70°C.

Preparation of sperm chromatin

Demembranated sperm chromatin for replication assays was prepared from testes of male frogs primed with 50U Folligon seven days in advance. Testes were rinsed three times in cold MMR buffer (100 mM NaCl; 2 mM KCl; 1 mM $MgCl_2$; 2 mM $CaCL_2$; 0.1 mM EDTA; 5 mM Hepes pH7.8), twice in cold NPB buffer (250 mM sucrose; 15 mM Hepes 7.4; 1 mM EDTA; 0.5 mM spermidine trichloride; 0.2 mM spermine tetrachloride; 1 mM DDT) and finely chopped with a razor. The obtained material was homogenised in a homogeniser, filtered through 25 µm Nylon membrane and centrifuged at 3000 rpm for 10 min at 4°C in HB-4 swing-out rotor. The pellet was resuspended in 1 ml of NPB buffer (250 mM sucrose; 15 mM Hepes 7.4; 1 mM EDTA; 0.5 mM spermidine trichloride; 0.2 mM spermine tetrachloride;

77

1 mM DDT) at room temperature and 50 µl of 10 mg/ml lysolecithin was added and samples were incubated at room temperature for 5 min. The demembranation of the sperm was tested by mixing 1 µl of the sample with 1 µl of Hoescht stain # (1 µl/ml). Following demembranation greater then 95%, 10 ml of cold NPB buffer (250 mM sucrose; 15 mM Hepes 7.4; 1 mM EDTA; 0.5 mM spermidine trichloride; 0.2 mM spermine tetrachloride; 1 mM DDT) supplemented with 3% BSA was added to the 1 ml sample and centrifuged at 3000 rpm for 10 min at 4°C in HB-4 swing-out rotor. Obtained pellet was resuspended in 500 µl of cold NPB buffer (250 mM sucrose; 15 mM Hepes 7.4; 1 mM EDTA; 0.5 mM spermidine trichloride; 0.2 mM spermine tetrachloride; 1 mM DDT) supplemented with 0.3% BSA and 30% glycerol. The sperm density was then counted, adjusted to 10^7 sperms/ml and aliquots were quickly frozen in liquid nitrogen.

The replication assay

The replication assay was done as described by Julian Blow (Blow, 1993). For a single replication reaction, 20 µl of interphase extract was supplemented with 25 mM phosphocreatine, 10 µg/ml creatine phosphokinase, 0.1 µl of [α^{32}P] CTP (10 mCi/ml) and 0.1 µl of prepared sperm chromatin. Replication reactions were incubated at 23°C for 3 hours and terminated by addition of 160 µl of Stop C solution (20 mM Tris pH 7.5; 5 mM EDTA; 0.5% SDS; 0.2 mg/ml proteinase K), followed by incubation at 37°C for 30 min. Then the reactions were added to 4 ml of cold 10% TCA, 2% Na-pyrophosphate solution, mixed and incubated at 4°C for 30 min. The high molecular weight precipitates were trapped on glass fibre filters and the filters were washed twice with 10% TCA, 2% Na-pyrophosphate solution and once with methanol, dried and the bound radioactivity was measured in the scintillation counter.

Histone HI kinase assay

In the Histone HI assay described in Chapter 6 (Figure 6/2) 15 µl CSF extract (arrested in meiosis II), prepared by Andrea Klotzbücher, was incubated with 1.8 µl of 4 mM $CaCl_2$ at 23°C. In a parallel sample, no $CaCl_2$ was added. At 0, 15, 30, 45 and 60 min, 1 µl aliquots were taken from both tubes and added to 12 µl of HIK buffer (80 mM Sodium β-glycerophosphate; 20 mM EGTA; 15 mM $MgCl_2$; 1mM DTT; 1mM PMSF pH 7.4), mixed and spun in the bench top centrifuge for 30 sec. To perform the histone HI kinase assay, 4 µl of the supernatant was incubated with 3 µg of histone HI, 2 µCi of $[\gamma^{32}P]$ ATP and 33 µM of ATP in the total volume of 8 µl in the presence of HIK buffer (80 mM Sodium β-glycerophosphate; 20 mM EGTA; 15 mM $MgCl_2$; 1mM DTT; 1mM PMSF pH 7.4) for 15 min at room temperature. Obtained samples were run on polyacrylamide gel, which was then dried and exposed to film.

Protein dephosphorylation

In the phosphatase assay described in Chapter 6 (Figure 6/3) 10 µl of CSF extract was incubated with 1 µl of 4 mM $CaCl_2$ for 30 min at 23°C essentially to produce a sample of an interphase extract. Then 10 µl of untreated CSF extract and 10 µl of $CaCl_2$ treated (interphase) extract were each mixed with 40 µl of PAP/SDS buffer (20 mM MES, pH 6.5; 100 mM NaCl; 1 mM PMSF; 1% SDS) and boiled for 3 min. 5 µl samples of either CSF or interphase extracts were then incubated with either PAP buffer (20 mM MES, pH 6.5; 100 mM NaCl; 1 mM PMSF) alone or with 0.2 U of potato acid phosphatase in the presence of PAP buffer or with 0.2 U of potato acid phosphatase and 50 mM NaF in the presence of the PAP buffer at 30°C for 1 hour. Orc1p mobility shift was then analysed by SDS-PAGE.

ORC-DNA interaction

Cross-linking assay

The cross-linking assay described in Chapter 8 was based on cross-linking protocols previously published by Orlando (Orlando and Paro, 1993) and Tanaka (Tanaka *et al.*, 1997). In this experiment six plates (15 cm diameter) of WAK cells were washed 3 times with ice-cold PBS (136.87 mM NaCl; 2.68 mM KCl; 10.14 mM Na_2HPO_4; 1.76 mM KH_2PO_4 pH 7.4), scraped and collected into 1.5 ml eppendorf tubes and incubated with 1 ml of cross-linking buffer (50 mM Tris pH 8.0; 100 mM NaCl; 1 mM EDTA; 0.5 mM EGTA; 1% formaldehyde) rotating at 4°C for 15 min. The cross-linking buffer was then removed and cells were incubated in a neutralising buffer (50 mM Tris pH 8.0; 100 mM NaCl; 1 mM EDTA; 0.5 mM EGTA; 125 mM glycine) rotating at 4°C for 5 min. Cells were then washed 4 times in ice-cold TBS (20 mM Tris 7.5; 150 mM NaCl) and gently sonicated in 500 µl of lysis buffer (50 mM Tris pH 7.5; 500 mM NaCl; 5 mM EDTA; 5 mM EGTA; 1% Triton X-100). Obtained samples were centrifuged at 10 000 rpm in a bench top centrifuge for 5 min at 4°C and the supernatants were used for immunoprecipitation using either anti-Orc1p antibody TK15 or a control anti-T antigen antibody. The immunoprecipitations were performed as described in "Immunoaffinity purification of *Xenopus* ORC" section of this chapter. The beads bound proteins were analysed by incubating beads in 60 µl of 2x sample buffer (4% SDS; 20% glycerol; 160 mM Tris, pH 6.8; 10% mercaptoethanol; 0.002% bromophenol blue) at 95°C for 30 min, followed by SDS-PAGE. The bead-bound DNA was recovered by incubation of the beads in 200 µl of TE (10 mM Tris pH 8.0; 1 mM EDTA)/1% SDS at 65°C overnight. Then 140 µl of TE (10 mM Tris pH 8.0; 1 mM EDTA) was added together with 5 µl of proteinase K (20 mg/ml) and incubated for 2 hours at 37°C. Then 3 µl of glycogen (20 mg/ml, Boehringer Mannheim) and 15 µl of 5 M NaCl were added and the

samples were extracted with phenol and chloroform, and precipitated with ethanol.

As the amount of DNA recovered from this assay was extremely low, it had to be amplified by PCR. Firstly, the samples were incubated with 0.3 µl of mung bean nuclease (114 000 U/ml, Pharmacia Biotech) in a total volume of 60 µl at 30°C for 5 min to create blund ends. Samples were then precipitated with ethanol in the presence of 5 µl of glycogen and the pellet was resuspended in 3 µl of TE (10 mM Tris pH 8.0; 1 mM EDTA). 0.4 µg of linkers (produced by annealing oligonucleotides AAGAATTCCTCGAGGCGGCCGC and GCGGCCGCCTCGAGG AATTC) were ligated to the blunt-ended DNA. The ligation product was then used as a PCR substrate using oligonucleotide AAGAATTCCTCGAGGCGGCCGC as a single primer, as the linkers had a 5' double AA overhang on one strand and therefore only ligated to the substrate in one direction. The PCR was otherwise performed as described in the Polymerase Chain Reaction section of this chapter.

ORC-DNA binding in extract

In the experiment studying ORC-DNA interaction in *Xenopus* egg extract, described in Chapter 8, Figure 8/4, 500 µl of interphase extract was suplemented with 25 mM phosphocreatine, 10 µg/ml creatine phosphokinase and 2 mM $MgCl_2$. 12.5 µg of pARS1-1.4.1 plasmid was then added and the sample was incubated for 15 min at 23°C and spun at 14 000 rpm for 8 min at 4°C in a bench top centrifuge. 200 µl of obtained supernatant was then incubated in the presence of 150 µl of binding buffer (50 mM Tris, pH 7.5; 150 mM NaCl; 0.1% NP-40; 1 mM DDT) with 15 µl of protein A beads coupled to either anti-Orc1p antibody TK15 or to a control, anti-SV40 T-antigen antibody, and rotated for 2 hours at 4°C. The beads were then washed three times with the binding buffer (50 mM Tris pH 7.5; 150 mM NaCl; 0.1% NP-40; 1 mM DDT) and 10% of the sample was analysed by SDS-PAGE to

examine whether ORC was bound to the beads. The remaining beads were incubated with 150 μl of 50 mM Tris pH 8.0/10 mM EDTA/1% SDS at 65°C for 20 min. Then 140 μl of TE (10 mM Tris pH 8.0; 1 mM EDTA), 2 μl of glycogen (20 mg/ml, Boehringer Mannheim) and 7.5 μl of proteinase K (20 mg/ml) were added and the samples were incubated for 2 hours at 37°C. Samples were then extracted with phenol and chloroform, precipitated with ethanol, resuspended in TE (10 mM Tris pH8.0; 1 mM EDTA), treated with RNAse and analysed by agarose gel. electrophoresis.

ORC-DNA binding *in vitro*

Figure 8/5 in Chapter 8 shows an experiment in which 48 μl of protein A beads coupled to anti-Orc1p antibody TK15 were incubated with 800 μl of interphase extract in the presence of 400 μl of binding buffer (36 mM Hepes, pH 7.6; 0.72 mM EDTA; 0.72 mM EGTA; 3.6 mM Magnesium acetate; 7.2% glycerol; 120 mM KCl; 0.2% NP-40) for 1 hour rotating at 4°C. Beads were then washed three times with the binding buffer (36 mM Hepes, pH 7.6; 0.72 mM EDTA; 0.72 mM EGTA; 3.6 mM Magnesium acetate; 7.2% glycerol; 120 mM KCl; 0.2% NP-40) and 6 μl of beads were used per binding reaction. Binding reactions were carried out in 50 μl final volume in the binding buffer (36 mM Hepes, pH 7.6; 0.72 mM EDTA; 0.72 mM EGTA; 3.6 mM Magnesium acetate; 7.2% glycerol; 120 mM KCl; 0.2% NP-40). To each reaction, 2 μg of DNA fragments produced by a restriction digest of pARS1-1.4.1 plasmid, and end labelled with [γ-^{32}P] ATP using T4 polynucleotide kinase and subsequent purification on G 50 Sepharose coloum, were added. Reactions were performed either in the presence or in the absence of 1 mM ATP, 10 mM phosphocreatine and 4 μg/ml creatine phosphokinase. Increasing amounts of DNA competitor poly (dI-dC) was added to the reactions to final concentrations of 0, 2 and

6 μg/ml. Binding reactions were carried out by rotating for 15 min at room temperature and samples were washed three times with binding buffer (36 mM Hepes, pH 7.6; 0.72 mM EDTA; 0.72 mM EGTA; 3.6 mM Magnesium acetate; 7.2% glycerol; 120 mM KCl; 0.2% NP-40) following the binding. Bound DNA was extracted by incubating the beads with 150 μl of 50 mM Tris pH 8.0; 10 mM EDTA; 1% SDS at 65°C for 20 min. Then 140 μl of TE (10 mM Tris pH8.0; 1 mM EDTA), 2 μl of glycogen (20 mg/ml, Boehringer Mannheim) and 7.5 μl of proteinase K (20 mg/ml) were added and the samples were incubated for 2 hours at 37°C. Samples were then extracted with phenol and chloroform, precipitated with ethanol, resuspended in TE (10 mM Tris, pH 8.0; 1 mM EDTA) and analysed on a 5% non-denaturing acrylamide gel.

Chapter 3

Immunopurification of *Xenopus* ORC

I started my PhD shortly after the *Xenopus* Orc1p was identified as a result of collaboration of three ICRF labs (Rowles *et al.*, 1996). The idea behind my project was to re-address the controversial question of specificity of replication origins in higher eukaryotes (see Introduction). The strategy was to attempt the identification of DNA sequenced *Xenopus* ORC recognises. I used the Orc1p subunit of the presumed *Xenopus* ORC to get a handle on the whole complex. This chapter describes raising a number of monoclonal antibodies against the Orc1p protein, development of *Xenopus* ORC immunopurification scheme and isolation of pure and presumably complete *Xenopus* ORC.

Polyclonal antisera to *Xenopus* Orc1p an Orc2p

I first generated polyclonal antibodies against *Xenopus* Orc1p in rabbits. These were useful to assess the results of immunoaffinity purification of ORC using anti-Orc1p monoclonal antibodies raised in mice. Mouse antibodies used for immunoaffinity purification can be present in the sample of the eluted protein. When such sample is analysed by immunobloting with anti-mouse secondary antibody, the mouse antibody on the blot is recognised, giving strong interfering signal. Using anti-rabbit secondary antibodies is a good way to get around this problem.

To generate polyclonal antibodies against the *Xenopus* Orc1p the full length open reading frame was subcloned into pET21b vector so as to add a hexahistidine tag to the C-terminus (see Methods). This fusion protein was expressed very efficiently in bacteria (see Methods). The majority of the produced protein was insoluble and was solubilised in guanidine hydrochloride prior to the purification on nickel agarose beads (see Methods). Figure 3/1 shows that the purification yielded a preparation of relatively pure polypeptide of molecular weight of approximately 110 kDa as expected for full length *Xenopus* Orc1p protein. Two rabbits (TK2 and TK3) were injected with 7 x 500 μg doses of the purified Orc1p protein. Figure 3/2 shows that the preimmune sera did not recognise any protein in the *Xenopus* egg extract. The immune respose increased with the increasing number of immunisations. Antisera (final bleed) produced by both rabbits recognised a band of the molecular weight corresponding to Orc1p on an immunoblot of frog egg extracts. The cross-reacting band was of very different molecular weight than Orc1p and therefore did not interfere with the use of these antibodies for immunoblotting. The TK3 antibody was used for all subsequent experiments, as it produced cleaner immunoblot.

The cloning of *Xenopus* Orc2p by Phil Carpenter allowed generation of antibodies against another subunit of frog ORC complex (Carpenter *et al.*, 1996). These antibodies became very important for determining whether the Orc2p protein co-purified with the Orc1p. This, in turn, served as an indicator, whether the whole ORC complex remained intact during the ORC purification procedure.

The clone encoding *Xenopus* Orc2p was kindly given to me by Phil Carpenter (see Methods). A histidine-tagged version of the full length gene was expressed in

Production of anti-Orc1p polyclonal antibodies

(1) Purification of recombinant Orc1p
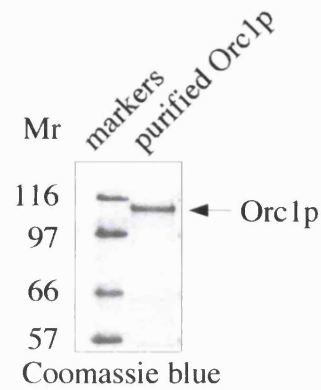
Figure 3/1. Recombinant Orc1p protein was produced in bacteria and purified on nickel agarose beads. The purified protein was run on polyacrylamide gel along side with molecular weight markers and stained with Coomassie blue to assess its size purity and protein concentration.

(2) Immunoblot of frog egg extract with anti-Orc1p antisera
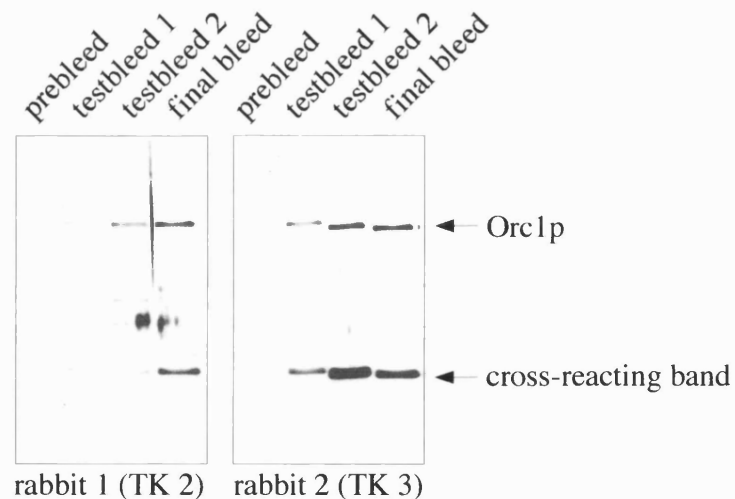
Figure 3/2. A total of 0.3 μl of *Xenopus* egg extract was run per lane and immunoblotted to test for the presence of anti-Orc1p antibodies in the sera of rabbit 2 and 3. The pre-immune sera (prebleed) did not recognise any protein from *Xenopus* egg extract. After immunisation (testbleed 1, testbleed 2 and final bleed), there was an increasing immune response to the Orc1p protein.

Production of anti-Orc2p polyclonal antibodies

(3) Purification of recombinant Orc2p
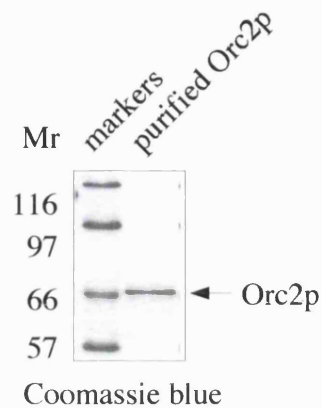
Coomassie blue

Figure 3/3. Recombinant Orc2p was produced in bacteria and purified on nickel agarose beads. Sample of the purified Orc2p protein was run on gel and stained with Coomassie blue. Size of purified Orc2p protein was compared to that of molecular weight markers.

(4) Immunoblot of frog egg extract with anti-Orc2p antisera

rabbit 1 (TK 5)   rabbit 2 (TK 6)
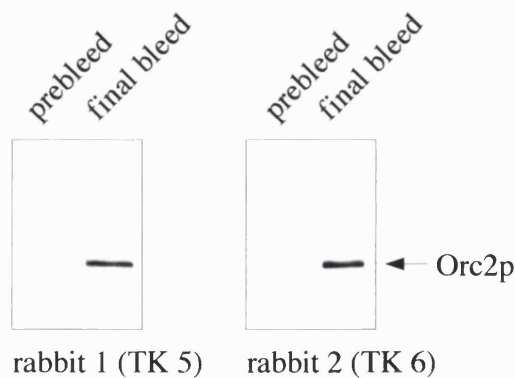
Figure 3/4. Total of 0.3 μl of *Xenopus* egg extract was run per lane and immunoblotted to test for the presence of anti-Orc2p antibodies in the sera of rabbits 5 and 6. The pre-immune sera (prebleed) did not recognise any protein from *Xenopus* egg extract. Antisera produced by both rabbits recognised a single band of a molecular weight corresponding to that of Orc2p on the immunoblot of frog egg extracts.

bacteria, producing large quantities of insoluble protein (see Methods). The purification on nickel agarose beads was therefore done under denaturing conditions (see Methods). Purified 67 kDa protein showed in Figure 3/3 is of slightly higher molecular weight than the 62 kDa Orc2p in frog extracts. This is likely to be due to the histidine tag changing the mobility of Orc2p protein on polyacrylamide gels. Two rabbits (TK5 and TK6) were injected with 5x100 μg of the purified Orc2p protein. Figure 3/4 shows that the pre-immune sera (prebleed) did not recognise any protein in frog egg extracts, while the antisera (final bleed) produced by both rabbits (TK5 and TK6) recognised a single band of the molecular weight of the Orc2p protein on an immunoblot of *Xenopus* egg extract. Antibody TK5 was used in subsequent experiments.

Raising anti-Orc1p monoclonal antibodies

The crucial role of yeast ORC in control over initiation of DNA replication focused the interest of a number of laboratories on purification of complexes homologues to yeast ORC from frog, flies and humans. However, purification of ORC from higher eukaryotes using conventional chromatography proved to be difficult and extremely inefficient (Gossen *et al.*, 1995; Rowles *et al.*, 1996). This chapter describes ORC purification by immunoaffinity chromatography. The success of this method relies on monoclonal antibodies, which provide a molecular handle on the whole ORC complex. A considerable effort was put into production of anti-Orc1p monoclonal antibodies which would be highly specific in recognising and binding the Orc1p in frog egg extracts.

To prepare anti-Orc1p monoclonal antibodies I wanted to immunise mice with Orc1p purified under native purification conditions (without the use of denaturing agent, such as guanidine hydrochloride) with the idea that the antibodies produced would

recognise the native Orc1p in frog egg extracts and would therefore be suitable for imunopurification of the native ORC complex. However, the Orc1p produced in bacteria was largely insoluble, with only very small fraction of the protein found in the supernatant. It took several rounds of purification therefore to produce enough soluble Orc1p protein to immunise ten mice (mice 1-10, data not shown). The animals were injected with 5 x 5 μg doses of recombinant Orc1p, which is the amount at the lowest end of the recommended range (Harlow and Lane, 1988). To prevent the potential failure of the monoclonal antibodies production due to the levels of injected protein being too low, five additional mice (mice 11-15) were injected with 5 x 20 μg doses of recombinant Orc1p protein purified under denaturing purification conditions, such as the one described in Figure 3/1. Figure 3/5, lanes 7, 10, 14 and 15 show that a number of the injected mice responded well to the recombinant Orc1p protein.

Mice 10 and 15 responded most strongly to the injected antigen and were therefore selected for production of anti-Orc1p monoclonal antibodies (Figure 3/5, lanes 10 and 15). Mouse 2 (Figure 3/5, lane 2) was chosen for a "practice round" of the relatively complicated protocol for making monoclonal antibodies. Spleen cells from the selected mice were fused to Sp2 myeloma cells to generate immortal antibody producing hybridomas (see Methods). The next three paragraphs describe the screening the antibodies produced by these hybridomas to identify antibodies capable of immunoprecipitating Orc1p. The selection is described individually for each of the three mice used. Following the three round of selection, the hybridomas producing identified antibodies were cloned to produce immortal antibody producing cell lines.

Response of mice to the injected Orc1p
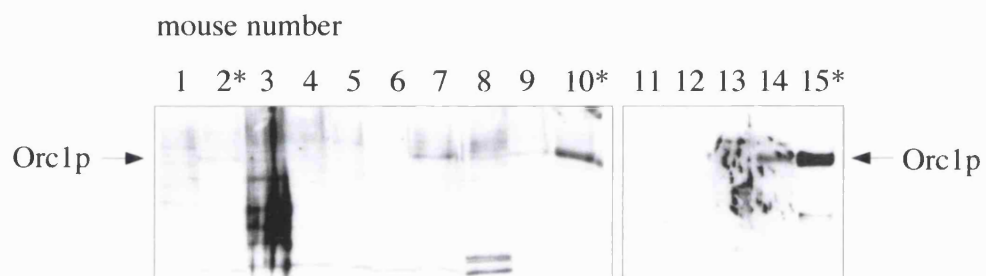
Figure 3/5. Sera from 15 mice taken after fifth immunisation with Orc1p protein were tested for their ability to recognise Orc1p protein in *Xenopus* egg extract. Total of 0.3 µl of *Xenopus* egg extract was run per lane and immunoblotted to test for the presence of anti-Orc1p antibodies. Mice number 7, 10, 12 and 15 responded well to the injected Orc1p protein. Asterisks indicate which mice were actually used for anti-Orc1p monoclonal antibody production.

Mouse 2

This mouse was initially intended for a "practice" round of the monoclonal antibody production. Therefore only a sample of 1/20 of the fused hybridoma cells were taken further for testing. These cells were plated in a single 96 well microtitre plate and the antibodies produced were tested for their ability to recognise recombinant Orc1p protein in ELISA assay. Antibodies from four wells out of the 96 tested gave positive signal in the first round of ELISA. Cells in these wells were allowed to grow and produce more antibodies which were tested again for their ability to recognise Orc1p by second round of ELISA assay. Out of these four antibodies tested one continued to show strong reaction to Orc1p. This antibody was named TK1 and was tested for the ability to immunoprecipitate Orc1p from *Xenopus* egg extract. Briefly, 5 μl of the interphase extract was incubated with 250 μl of antibody containing tissue culture supernatant in the presence of 130 μl of IP buffer (see Methods). After one hour, 10 μl of protein G beads were added to harvest the immunoprecipitated proteins. Figure 3/6 shows that antibody TK1 immunoprecipitated Orc1p from frog egg extracts, while the preimmune sera of the mouse 2 did not.

Mouse 15

Hybridoma cells generated from the spleen of mouse 15 were plated onto twenty microtitre plates. Out of approximately 2 000 hybridomas tested, 80 were selected as positive in the first round of ELISA assay, out of these, 22 were positive in the second round of ELISA assay. These antibodies were named TK2-23 and were again tested by immunoprecipitation. Figure 3/7 shows that antibodies secreted by five out of the twenty two selected hybridomas were able to immunoprecipitate Orc1p from frog egg extract.

<u>Mouse 10</u>

The final round of the anti-Orc1p antibody production was done using the spleen cells from mouse number 10. Out of almost 2 000 hybridomas tested, 90 were positive in the first round of ELISA assay, out of these, 35 re-tested positive in the second round. These antibodies were named TK31-65 and were re-tested by immunoprecipitation. Figure 3/8 shows that out of the 35 tested hybridomas seventeen secreted antibodies, which could recognise Orc1p in frog egg extracts, and be therefore potentially useful for immunopurification of frog ORC.

In total, twenty three hybridomas producing antibodies capable of immunoprecipitation of Orc1p from frog egg extracts were selected from the three fusions performed. To establish stable cell lines, these were then cloned by the limited dilution procedure. Five clones were lost during this procedure because they failed to divide at a sufficient rate. Seventeen stable cell lines were eventually generated and re-tested for their ability to immunoprecipitate Orc1p from frog egg extracts. Figure 3/9 shows that although three of these cell lines lost their ability to make immunoprecipitating antibodies, fourteen cell lines producing monoclonal antibodies capable of immunoprecipitating Orc1p from *Xenopus* egg extracts were successfully established.

Intrestingly, the proportion of antibodies capable of immunoprecipitating Orc1p from egg extracts was higher among the antibodies generated using mouse 10, which was injected with the native Orc1p. Curiously, a number of antibody secreting hybridomas derived from this mouse grew extremely slowly and failed to be cloned.

Screening anti-Orc1p antibodies by immunoprecipitation (first round)

(6) MAb TK1 derived from mouse 2 immunoprecipitated Orc1p
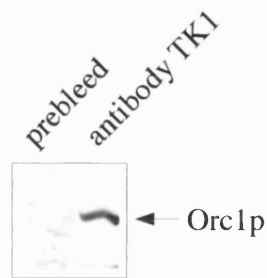
Figure 3/6. Antibody TK1 was generated using the spleen cells from mouse 2. To test the ability of this antibody to immunoprecipitate Orc1p from *Xenopus* egg extracts, both the TK1 antibody and the preimmune sera taken from the mice 2 were incubated with frog egg extracts and harvested on protein G beads. Immunoprecipitated proteins were separated by SDS-page and blotted with anti-Orc1p polyclonal antibody TK 3. Antibody TK1 immunoprecipitated the Orc1p from frog egg extract, while the preimmune sera did not.

(7) Five MAbs derived from mouse 15 immunoprecipitated Orc1p

antibody number
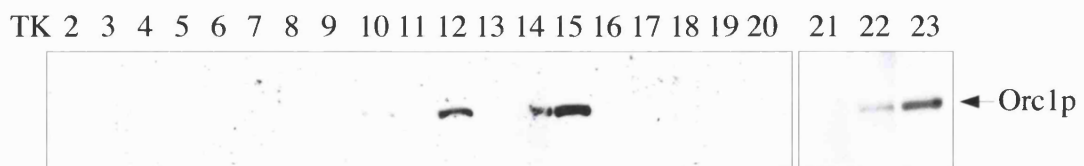
Figure 3/7. Antibodies TK2-23 were generated using the spleen cells from mouse 15. To test the ability of these antibodies to immunoprecipitate Orc1p from *Xenopus* egg extracts, the antibodies were incubated with frog egg extracts and harvested on protein G beads. Immunoprecipitated proteins were separated by SDS-page and blotted with anti-Orc1p polyclonal antibody TK 3. Antibodies TK10, 12, 15, 22 and 23 immunoprecipitated the Orc1p from frog egg extract, while the rest of the produced antibodies did not.

93

Screening anti-Orc1p antibodies by immunoprecipitation (first round) continued

(8) Seventeen MAbs derived from mouse 10 immunoprecipitated Orc1p

antibody number

TK 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50

Orc1p ►

antibody number

TK 51 52 53 54 55 56 57 58 59 50 61 62 63 64 65
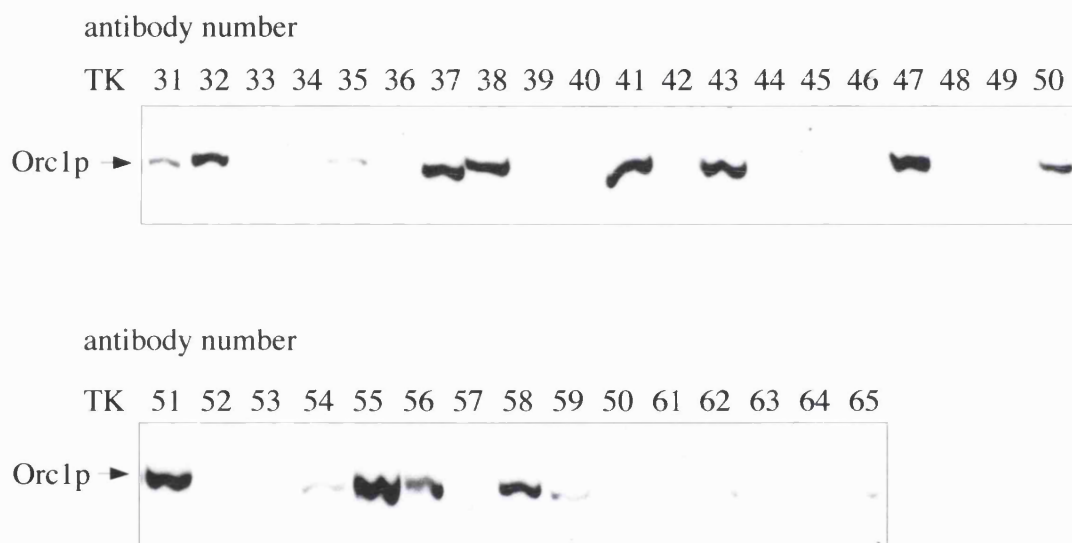
Orc1p ►

Figure 3/8. Antibodies TK31-65 were generated using the spleen cells from mouse 10. To test the ability of these antibodies to immunoprecipitate Orc1p from *Xenopus* egg extracts, the antibodies were incubated with frog egg extracts and harvested on protein G beads. Immunoprecipitated proteins were separated by SDS-page and blotted with anti-Orc1p polyclonal antibody TK 3. Seventeen antibodies from this group immunoprecipitated the Orc1p from frog egg extract, while the remaining eighteen antibodies did not.

Screening anti-Orc1p monoclonal antibodies by immunoprecipitation (second round)

monoclonal antibody number

TK   1 10 12 15 22 23 31 37 41 43 47 50 51 55 56 58 59
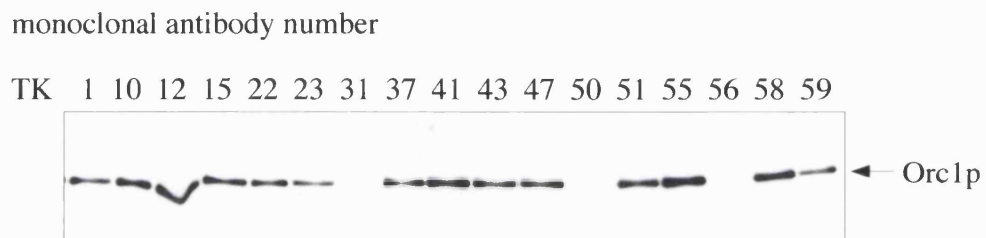
◄— Orc1p

Figure 3/9. Monoclonal antibodies generated against Orc1p protein were re-tested for their ability to immunoprecipitate Orc1p from *Xenopus* egg extracts following the single cell cloning procedure. The antibodies were incubated with frog egg extract and harvested on protein G beads. Immunoprecipitated proteins were separated by SDS-page and blotted with anti-Orc1p polyclonal antibody TK3. Fourteen anti-Orc1p monoclonal antibodies were able to specifically immunoprecipitate Orc1p protein from *Xenopus* egg extracts.

Characterising anti-Orc1p monoclonal antibodies

Monoclonal antibodies produced by the fourteen established cell lines were next tested for their ability to recognise Orc1p protein on immunoblot of *Xenopus* egg extract. All the antibodies recognised Orc1p protein, although with very different affinities (Figure 3/10).

The immunoglobulin subclass of the produced anti-Orc1p monoclonal antibodies was kindly determined by Maggie Stubbs in the Hybridoma Development Unit of ICRF (Table 3/1).

Table 3/1

Determination of immunoglobulin subclass of anti-Orc1p antibodies

Antibody name	Subclass	Antibody name	Subclass
TK1	IgG2a	TK41	IgG2a
TK10	IgG3	TK43	IgG1 & IgG2a
TK12	IgG1	TK47	IgG2a
TK15	IgG2a	TK51	IgG2a (weak)
TK22	IgG1	TK55	IgG2a & IgM
TK23	IgG1	TK58	IgG2b
TK37	IgG2a	TK59	ND

Immunoblot with anti-Orc1p monoclonal antibodies

monoclonal antibody number
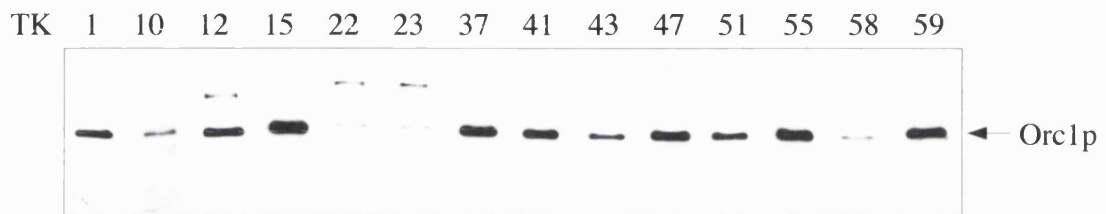
Figure 3/10. Total of 0.3 µl of *Xenopus* egg extract was run per lane. This was used to test all fourteen anti-Orc1p monoclonal antibodies for their ability to recognise Orc1p protein on immunoblot.

Next, the epitopes recognised by the monoclonal antibodies were coarsely mapped to three regions of Orc1p protein by testing their ability to recognise either the full length Orc1p or the two N-terminal deletion mutants, Δ N 280 and Δ N 423 produced in bacteria (see Methods). Antibodies TK1 and TK15 bind to the N-terminus, between residues 1-280 of Orc1p protein (Figure 3/11A, lanes 1 and 4 and Figure 3/11B), antibodies TK22 and TK23 recognise the region between residues 280 and 423 (Figure 3/11A, lanes 5 and 6 and Figure 3/11B) and all the rest of the antibodies recognise C-terminus of Orc1p beyond residue 423 (Figure 3/11A, lanes 2-3 and 7-14 and Figure 3/11B).

Epitopes recognised by antibodies TK1 and TK 15, already shown to be in the first 280 residues, were analysed further by testing their ability to recognise three N–terminal fragments of Orc1p protein; N-280, N-270 and N-260. Antibody TK1 bound between residues 1-260 of Orc1p protein (Figure 3/12A, lanes 1-3 and Figure 3/12B) and antibody TK15 recognised the region between residues 270 and 280 (Figure 3/12A, lanes 4-6 and Figure 3/12B).

The identified nucleotide sequence of the epitope recognised by the monoclonal antibody TK15 is 268–LEQEDTTLDEEFKMLE-283 (considering possible epitope overlap of the deletion boundaries). The identity of the epitope recognised by this antibody was further studied by Julian Gannon using the phage-display method. The consensus epitope identified by the phage-display method was DEEFKI/LI. There is very good agreement between the peptide sequences identified by the two mapping methods with the slight sequence discrepancy being likely to be due to the similar structure of the amino acids methionine, leucine and isoleucine, as the antibody may recognise oxidised methionine.

Mapping of anti-Orc1p monoclonal antibodies

A

monoclonal antibody number

B

TK 1 and 15

TK22 and 23

TK 10, 12, 37, 41, 43,
47, 51, 55, 58 and 59
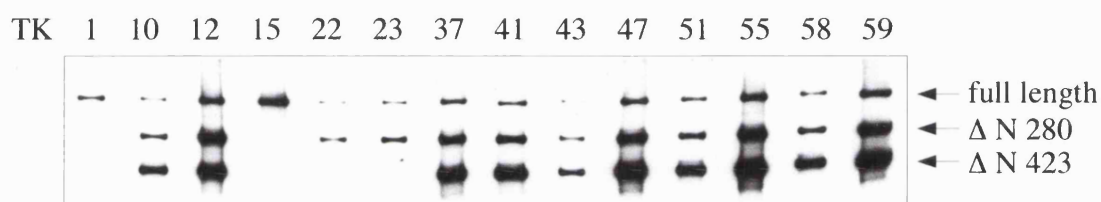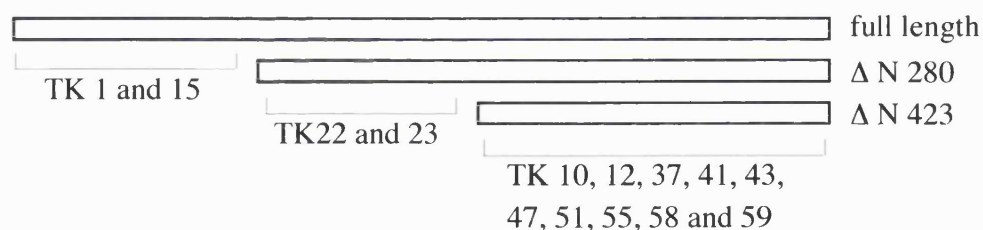
Figure 3/11. (A) A mixture of full length Orc1p protein with and two N-terminal deletion mutants, Δ N 280 and Δ N 423 were expressed in bacteria, separated by SDS-PAGE and immunobloted with all fourteen anti-Orc1p monoclonal antibodies. (B) A diagram of the Orc1p deletion mutants showing that antibodies which react only with the full length protein must recognise the region of Orc1p protein within the first 280 aminoacids (TK1 and TK15). Those reacting with both full length protein and the Δ N 280 deletion mutant, but not with the Δ N 423 deletion mutant bind to the region of Orc1p protein between the residues 280 and 423 (TK 23 and TK23). Finally, the antibodies which react with all three polypeptides must recognise a protein sequence between the residue 423 and the C-terminus of the Orc1p protein (TK10, 12, 37, 41, 43, 47, 51, 55, 58 and 59).

Epitope maping of anti-Orc1p monoclonal antibodies TK1 and TK15
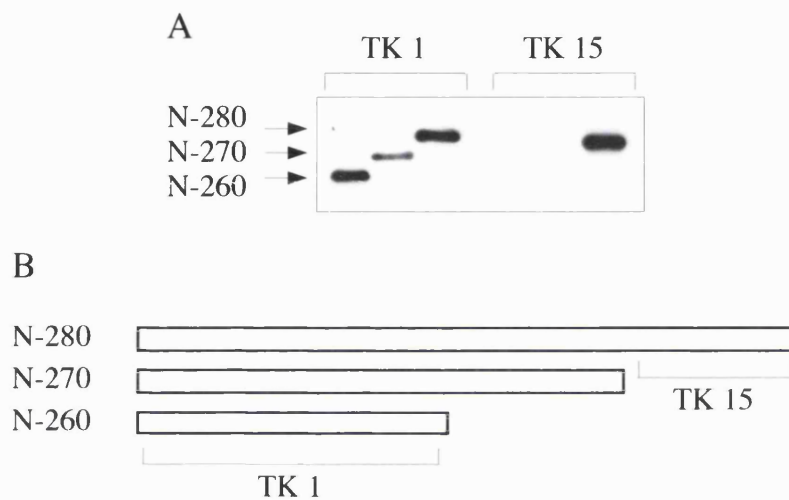
Figure 3/12. (A) Three N-terminal fragments, N-280, N-270 and N-260 of Orc1p protein were expressed in bacteria, separated by SDS-PAGE and immunobloted with anti-Orc1p monoclonal antibodies TK1 and TK15. (B) A diagram of the Orc1p N-terminal fragments showing that antibody TK15 previously shown to recognise the N-terminal 280 residues of Orc1p reacts only with the N-280 fragment of Orc1p protein and not with the N-270 and N-260 fragments. Antibody TK15 therefore binds between residue 270 and 280 of of Orc1p protein. Antibody TK1 which react with all N-terminal fragments must recognise a protein sequence between the N-terminus of the Orc1p protein and the residue 260.

To confirm that the identified epitope was indeed the one recognised by the monoclonal antibody TK 15, four peptides overlapping the identified site were synthesised by the Peptide Synthesis Unit of ICRF (see Table 3/2).

Table 3/2

Peptides tested as recognition epitopes of Mab TK15

Peptide identified by deletion mapping:		LEQEDTTLDEEFKMLE
Peptide identified by phage-display;		DEEFK(I/L)I
Peptides synthesised:	peptide 0	EDTTLDEEFK
	peptide 1	EQEDTTLD
	peptide 2	DTTLDEEF
	peptide 3	DEEFKMLE

The synthetic peptides were tested for their ability to act as epitopes of the monoclonal antibody TK 15. Peptides 0 and 3 were recognised by the monoclonal antibody TK15 in an ELISA assay. However, none of the peptides were able to completely inhibit the binding of the antibody TK15 to Orc1p in frog egg extracts under the conditions tested (data not shown). In the experiment shown in Figure 3/13, the interference of the peptides with the ability of the antibody to deplete Orc1p from the extract is shown. Monoclonal antibody TK15 depletes all Orc1p from the extract in the presence of PBS only, or in the presence of peptides 1, 2 and 3. Only peptide 0 could reduced the ability of antibody TK15 to deplete Orc1p from the extract, leaving about half of the Orc1p protein behind. The nucleotide sequence EDTTLDEEFK is therefore the best representation of the *in vivo* epitope of the anti-Orc1p monoclonal antibody TK15.

Inhibition of monoclonal antibody TK 15 by peptides

Figure 3/13. Anti-Orc1p antibody TK15 was covalently coupled to protein A beads and incubated in *Xenopus* egg extract in the presence of the indicated peptides (lanes 3-6), or in the absence of any peptide (lane 2). Monoclonal antibody TK15 completely depleted the Orc1p from frog egg extract both in the absence of any peptide (lane 2) or in the presence of peptides 1-3 (lanes 4-6). Only peptide 0 (lane 3) was able to prevent TK15 from depleting Orc1p from the extract. Lane 1 shows total undepleted extract.

Immunoaffinity purification of *Xenopus* ORC

I next tested if the monoclonal antibodies capable of immunoprecipitating Orc1p from frog egg extracts could be used for immunopurification of the whole ORC complex from *Xenopus* extracts. The initial immunopurification experiments were very promising. Following a simple immunoprecipitation, a strong band corresponding to the molecular weight of Orc1p was clearly visible by Coomassie blue staining. This band was only present in the immunoprecipitates done with anti-Orc1p monoclonal antibodies, and not in the immunoprecipitates done with a control anti-SV40 T-antigen antibody. The stained band was confirmed to be the *Xenopus* Orc1p by immunoblotting (data not shown). Identification of the other ORC subunits co-purified with Orc1p protein was more difficult. In retrospect, this was probably due to the unexpected instability of *Xenopus* ORC in increasing amounts of salt. A compromise had to be found to identify conditions in which contaminating proteins were removed, but the ORC complex remained intact. In the course of these experiments, the cDNA encoding *Xenopus* Orc2p was identified (Carpenter *et al.*, 1996) and we were thus able to raise anti-Orc2p antibodies (Figure3/4). These antibodies became a very valuable tool in assessing the integrity of the ORC complex in the immunopurification experiments. The final conditions for purification of *Xenopus* ORC used low amounts of antibody coupled beads (enough to deplete about half of Orc1p from the extract) which reduced the amount of binding space available for non-specific binding, and mild washing buffer (150 mM KCl, 0.5 % NP-40) to preserve the integrity of the ORC complex.

Briefly, purification of *Xenopus* ORC was done in the following way: antibodies TK1, TK15, TK37 and TK47 were used as they are from the IgG2a subclass which binds efficiently to protein A beads (Affiprep protein A beads gave much lower

background binding than the Sepharose based protein G beads). The four selected anti-Orc1p monoclonal antibodies as well as an anti-SV40 T-antigen monoclonal antibody to serve as negative control, were covalently coupled to protein A beads (see Methods). 12 μl of coupled protein A beads were incubated with 400 μl of *Xenopus* egg extract in the presence of 240 μl of buffer. Beads were then washed in the presence of 150 mM of KCl and 0.5 % of NP-40 (see Methods). Bound proteins were extracted and analysed by SDS-PAGE. Figure 3/14 shows that all four anti-Orc1p monoclonal antibodies selectively immunoprecipitated a protein corresponding to the size of Orc1p, easily visible by Coomassie blue staining while a control anti-SV40 T-antigen antibody did not. Five other polypeptides that specifically associated with all of the anti-Orc1p antibodies and not with the control antibody were also visible on Coomassie blue stained gel (Figure 3/14A, see arrows). Immunoblotting with anti-Orc1p polyclonal antibody confirmed that the 110 kDa polypeptide is indeed Orc1p protein and immunoblotting with anti-Orc2p polyclonal antibody showed that the 64 kDa polypeptide is Orc2p (Figure 3/14B). The 110 and 64 kDa polypeptides were also reconfirmed to be previously identified *Xenopus* Orc1p and Orc2p respectively by Dr Darryl Pappin and his colleagues in the Protein Sequencing Unit of ICRF by Mass spectrometry. The four other proteins (~68, 48, 43 and 27 kDa) co-purified with the *Xenopus* Orc1p and Orc2p are therefore good candidates for the missing subunits of *Xenopus* ORC.

Attempt to elute the purified ORC from the antibody TK15

Purification using antibody TK15 yielded the cleanest preparation of frog ORC complex. It would have been useful to be able to elute the purified complex from this antibody for further biochemical studies. The peptides previously identified as epitopes of the monoclonal antibody TK15 were therefore used for elution

Immunoaffinity purification of *Xenopus* ORC
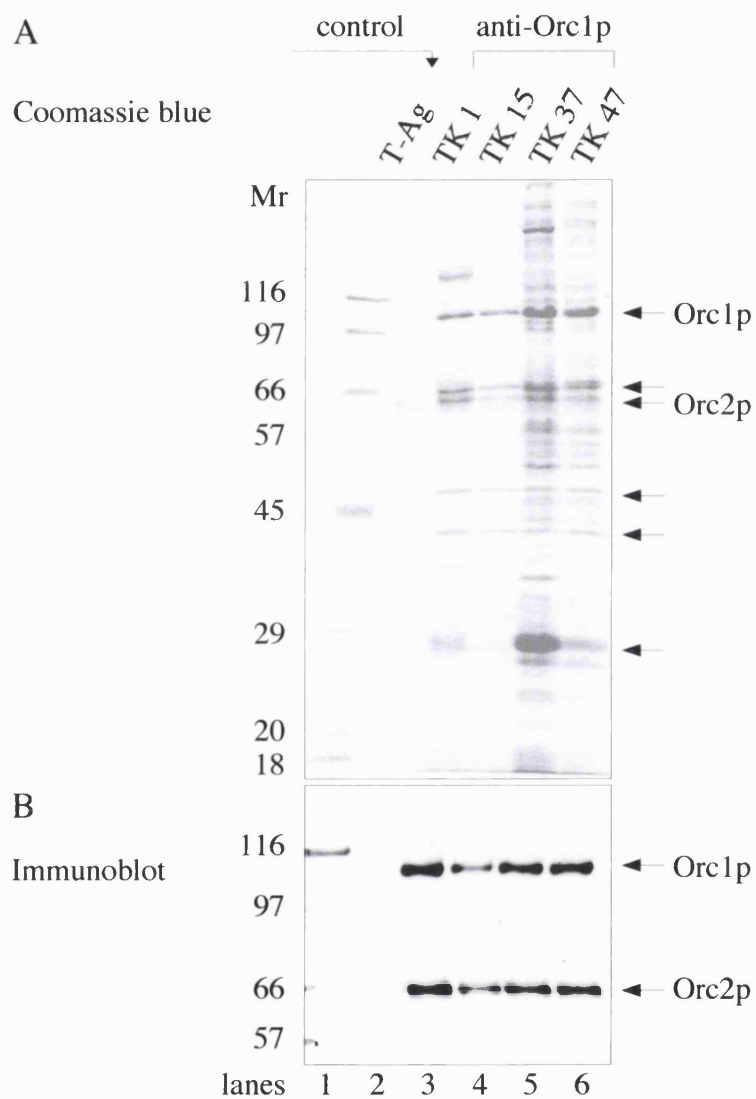
Figure 3/14. (A) Interphase extract from *Xenopus* eggs was incubated with protein A beads coupled to either control anti-SV40 T-antigen mAb 423 (lane 2) or anti-Orc1p mAbs TK1, TK15, TK37 and TK47 (lanes 3-6). The bound proteins were analysed by SDS-PAGE and detected by Coomassie blue staining. Polypeptides found in all four samples obtained with anti-Orc1p antibodies (lanes 3-6), but not in the control sample (lane 2) are marked with arrows. (B) A parallel gel was immunoblotted with anti-Orc1p and Orc2p polyclonal antibodies. Lanes 1, marker proteins.

experiments (see Table 3/2). However, elution of the purified ORC of the monoclonal antibody TK15 proved to be extremely difficult. Figure 3/15 shows that buffer as mild as PBS causes the Orc2p and Orc4p (see Chapter 4) subunits of ORC to be lost from the immunoprecipitates while all of the Orc1p remains tightly bound (see Methods). Addition of the peptides, that had been identified as epitopes of the monoclonal antibody TK15 (see Table 3/2) did not help to elute Orc1p from the antibody. A number of buffers commonly used for elution of proteins from antibodies (Harlow and Lane, 1988) were used in an attempt to elute ORC from the monoclonal antibody TK15 in combination with the epitope peptides. Most of these buffer-peptide combinations tested were unsuccessful in releasing the intact ORC. In fact, releasing of any Orc1p at all from the TK15 antibody proved difficult. It seems that high affinity of the antibody TK15 is due to a very low off rate of the antibody-Orc1p interaction, causing very slow and poor elution of Orc1p from the antibody.

The only successful attempt to elute ORC from the monoclonal antibody TK15 is shown in Figure 3/16, where a combination of 1 mg/ml of peptide 0 (EDTTLDEEFK) in an EB buffer plus 100 mM $MgCl_2$, 1% NP-40 and 10% ethylene glycol were used. However, under these conditions the ORC complex tends to dissociate and would have to be re-assembled for use in further biochemical studies.

Immunoaffinity as a method of choice for ORC purification

We have purified the Origin Recognition Complex from frog egg extracts using highly specific monoclonal antibodies directed against the Orc1p subunit of this complex. The high affinity of these antibodies for the Orc1p protein was demonstrated by the great difficulty we encountered when we attempted to elute purified ORC from the monoclonal antibody TK15. We have shown that at least three out of the four antibodies (TK1, TK15, TK37 and TK47) used for the

Attempt to elute purified ORC from the monoclonal antibody TK15

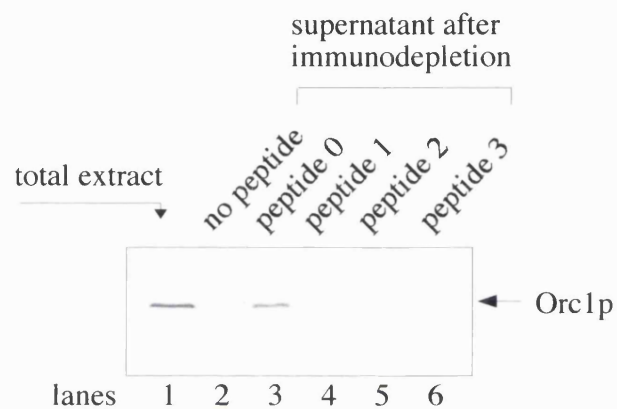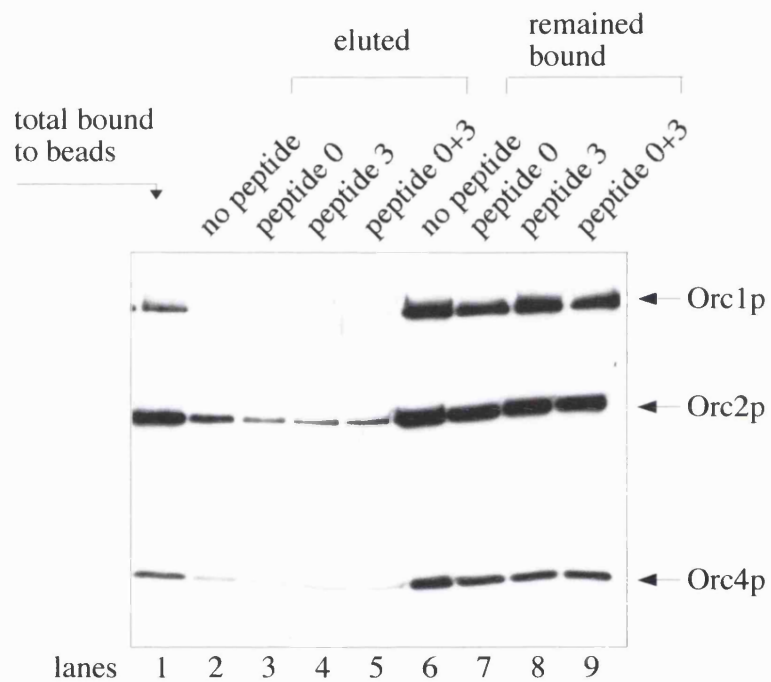**Figure 3/15.** Anti-Orc1p antibody TK 15 was covalently coupled to protein A beads and incubated in *Xenopus* egg extract. Bound proteins are shown in lane 1. To elute the bound ORC, beads were incubated with either PBS only (no peptide) or with PBS in the presence of either peptide 0, peptide 3 or both peptides 0 and 3. Lanes 2-5 show protein eluted, while lanes 6-9 show proteins which remained bound to the beads. (Please note that in lanes 6-9 double amount of sample was loaded by mistake.)

Elution of purified ORC from the monoclonal antibody TK15
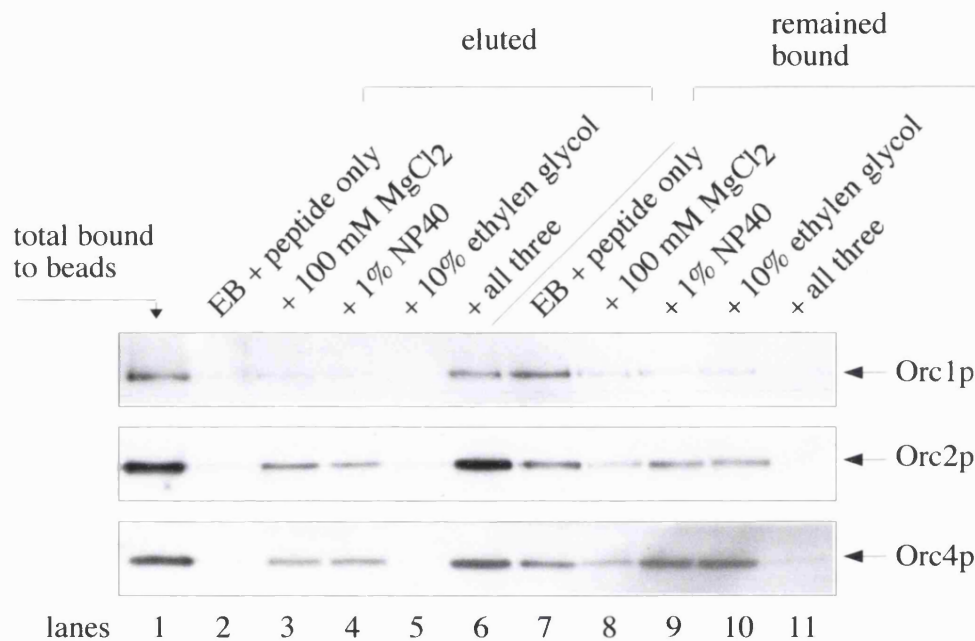
Figure 3/16. Anti-Orc1p antibody TK15 was covalently coupled to protein A beads and incubated in *Xenopus* egg extract. Bound proteins are shown in lane 1. To elute the bound ORC, beads were incubated for 2 hours at room temperature in the presence or 1mg/ml of peptide 0 (EDTTLDEEFK) in EB buffer only (lane 2) or in the presence of 100 mM $MgCl_2$ (lane 3), or 1% NP-40 (lane 4) or 10% ethylene glycol (lane 5) in addition to the peptide. Finally, the elution was done in the presence of EB, peptide, 100 mM $MgCl_2$, 1% NP-40 and 10% ethylene glycol (lane 6). Lanes 2-5 show the eluted proteins, while lanes 7-11 show proteins left behind on the beads.

immunoaffinity purification of *Xenopus* ORC recognised different epitopes on the Orc1p molecule. This gave us confidence that the proteins co-immunoprecipitated with all four anti-Orc1p antibodies are those which truly interact with Orc1p. It is obviously possible that some of the proteins that co-immunoprecipitated with only some of the anti-Orc1p monoclonal antibodies also interact with Orc1p rather then being spurious contaminants. This opens a whole range of opportunities to investigate proteins interacting with *Xenopus* Orc1p and their potential function in DNA replication or in chromatin structure. The following chapters of this thesis will deal in some detail with proteins that co-immunoprecipitated with all four anti-Orc1p monoclonal antibodies.

As mentioned at the beginning of this chapter, a number of laboratories attempted purification of complexes homologues to yeast ORC from frog, flies and humans. It became, however, gradually apparent that conventional chromatography was unsuitable for the purification of ORC from higher eukaryotes (Gossen *et al.*, 1995, Stillman-personal communication; Rowles *et al.*, 1996). It appears that the ORC complex from higher eukaryotes, at least in the case of frogs, is very labile in the presence of moderately increased levels of salt such as one used in conventional ion-exchange chromatography. The majority of purification schemes which succeeded were based, at least partially, on the immunoaffinity approach (Carpenter and Dunphy, 1998; Gossen *et al.*, 1995; Romanowski *et al.*, 1996).

In addition to ORC purification described in this chapter, isolation of ORC from frog egg extracts was attempted by Alison Rowles, Piotr Romanowski and Phil Carpenter (Carpenter and Dunphy, 1998; Romanowski *et al.*, 1996; Rowles *et al.*, 1996). Comparison of the protein bands on the pictures of gels in their publications suggests that they all contain the 110 kDa Orc1p, 64 kDa Orc2p and the 68, 48, and 43 kDa

proteins presumably identical to those described in this chapter (Figure 3/14). They do not seem to contain the 27 kDa polypeptide. The ORC purified by Alison Rowles contains several additional proteins (Rowles *et al.*, 1996).

It is striking that the molecular masses of the ORC subunits purified from *Xenopus* eggs (~110, 68, 64, 48, 43 and 27 kDa) corresponded closely to those of the ORC complex purified from *Drosophila* embryos (~115, 82, 79, 47, 42 and 30 kDa; Gossen *et al.*, 1995), whereas the ORC identified in yeast was somewhat different (120, 72, 62, 56, 53 and 50 kDa; Bell and Stillman, 1992; Figure 3/17).

Isolation of relatively large amounts of *Xenopus* ORC described in this chapter opened new oportunities to study new ORC subunits, proteins interacting with *Xenopus* ORC and the DNA sequences recognised by this protein complex.

ORC complexes purified from yeast, flies and frogs

A   B   C

S. cerevisiae   *D. melanogaster*   *X. laevis*
ORC   ORC   ORC
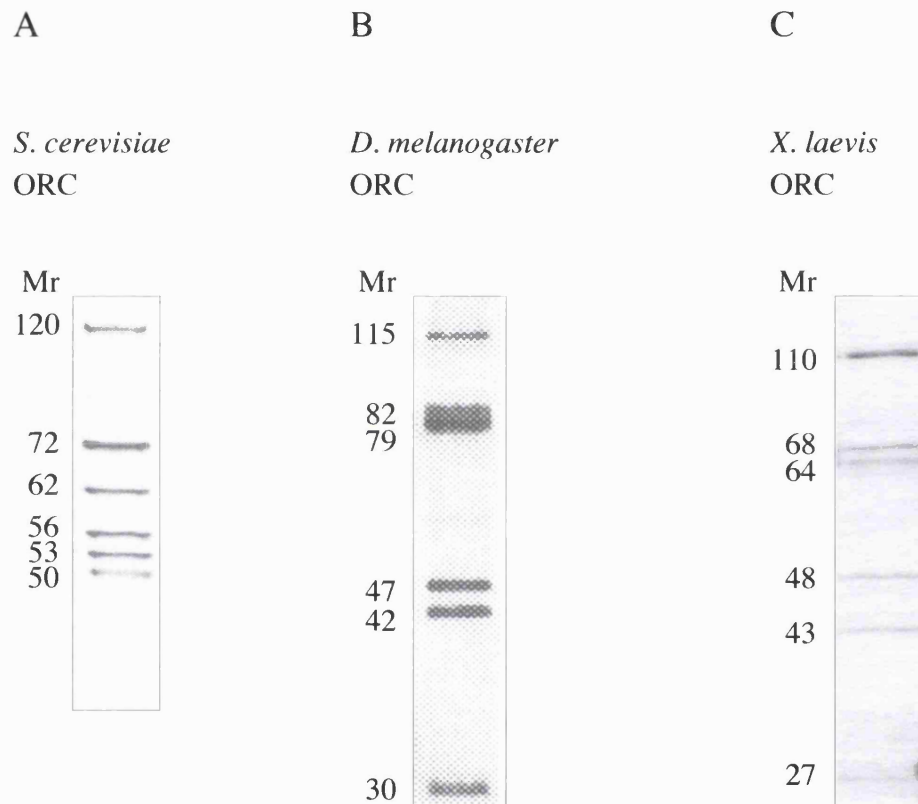
Figure 3/17. (A) ORC complex purified from *S. cerevisiae* by Bell and Stillman 1992, figure taken from Bell and Stillman 1995. (B) ORC complex purified from *D. melanogaster* by Gossen et al., 1995. (C) ORC complex purified from *X. laevis*, purification described in this thesis and in Tugal et al., 1998.

Chapter 4

Identification of *Xenopus* Orc4p

Purification of *Xenopus* ORC described in the previous chapter provided an opportunity to identify the individual subunits of frog ORC. Orc1p and Orc2p subunits of *Xenopus* ORC had previously been cloned, both displaying relatively low levels of homology to their yeast counterparts (Carpenter *et al.*, 1996; Rowles *et al.*, 1996). Despite extensive searches of sequence databases, homologues of the remaining four subunits of yeast ORC were not found in higher eukaryotes at the time of this work. This chapter describes microsequencing of the 43 kDa protein that copurified with *Xenopus* Orc1p and Orc2p and subsequent identification of Orc4 genes from both mouse and frog.

Identification of mouse Orc4 EST

To establish the identity of the four polypeptides specifically co-purifying with the *Xenopus* Orc1p and Orc2p, purified samples of *Xenopus* ORC (Chapter 3, Figure 3/9) were prepared for protein microsequencing (see Methods) and given to Darryl Pappin and his colleagues in the Protein Sequencing Unit of ICRF to be sequenced by mass spectrometry. Initially, the most successful was sequencing of the 43 kDa co-purified protein, which generated two peptide sequences: (K/R)TV(L/I)HGESNS A(L/I)(L/I)(L/I)GPR and (K/R)Q(L/I)ED(L/I)(F/M)G(F/M)K. A sequence database search with first of the two peptides identified a mouse clone AA110785 from the EST database (Figure 4/1). When in turn the AA110785 mouse EST clone was used

112

to search protein databases it was found to be homologous to the *S. cerevisiae* Orc4p. EST database contains short, approximately 400 bp sequences of the actual clones in the gene bank. To find out the full length sequence of the identified candidate mouse Orc4 gene we searched the EST databases for other sequences with overlapping homology to the original mouse clone. In this way, a full length hypothetical mouse-human hybrid Orc4 gene was reconstituted (Figure 4/2).

The mouse clone AA110785 was obtained from the EST gene bank, and DNA sequencing revealed the full length mouse Orc4 gene. The deduced protein sequence is homologous to the *S. cerevisiae* Orc4p throughout the whole sequence, although the overall homology is relatively low, with only 24% identities (Figure 4/3). The sequence of the mouse Orc4 gene also confirmed that the initial computer-generated hybrid mouse-human Orc4 gene was correctly assembled.

Cloning of *Xenopus* Orc4 gene

The mouse Orc4 sequence was used as a probe to identify the *Xenopus* Orc4 cDNA. A 473 bp fragment of the mouse ORC4 cDNA was produced by PCR (see Methods) and labelled with [^{32}P] dCTP using the Megaprime DNA labelling system (see Methods). Approximately 60,000 plaques of a *Xenopus* ovary cDNA library in λ-ZAP, which was kindly provided by Eva Dworkin from Vienna Biocentre, were screened (see Methods). Initially, six positives were detected by autoradiography and were isolated and re-tested in the second round of screening (see Methods). Two of the initially isolated clones tested positive in the Secondary screen (Figure 4/4). Plasmid DNA from both of the clones was obtained by *in vivo* excision (see Methods). Sequencing of the two clones revealed that they were derived from the same cDNA, cloned into the library vector in the opposite directions. The deduced

Identification of mouse EST clone which encodes the Orc4p

<pre>
Xenopus peptide sequence: R T V I H G E S N S A L I L G P R
                             R T   I + G E S N S   L I + G P R
   mouse AA110785 clone: R T A I Y G E S N S V L I V G P R
</pre>

Figure 4/1. Homology between a peptide sequence generated from the 43 kDa polypeptide which specifically co-purified with the *Xenopus* Orc1p and Orc2p proteins and a mouse EST clone number AA110785.

Assembly of the full length Orc4 gene
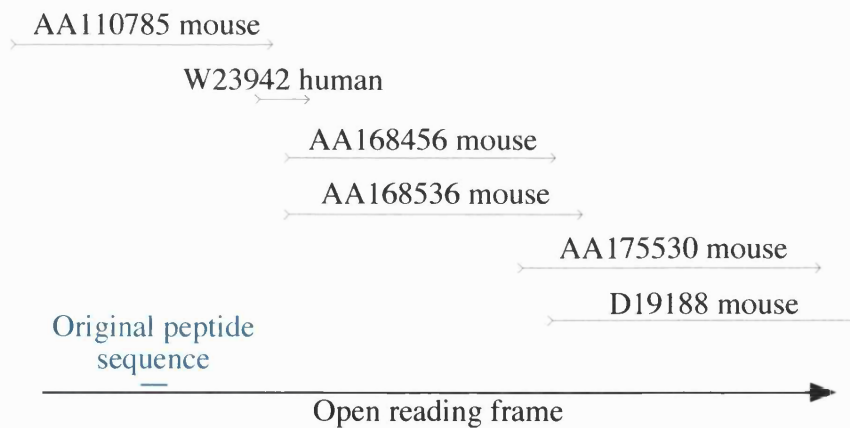
Figure 4/2. The EST database was searched with sequences from the ends of the AA110785 EST clone. Clones with overlapping homologous sequences were fused and used to search for further overlapping clones. In this way the full-length mouse-human hybrid Orc4 gene has been assembled.

Mouse EST clone AA110785 encodes a homologue of *S. cerevisiae* Orc4p

```
yeast Orc4p    MTISEARLSP QVNILPIKRH SNEEVEETAA IIKKRTIDNE KCKDSDPGFG
mouse Orc4p    MSSRKTKSNA HAEQLS---- ---QVQR--- ILRERF---- -CHHSPHS--

yeast Orc4p    SLQRRLLQQL YGTLPTDEKI IFTYLQDCQQ HIDRIIKQSI IQKESHSVIL
mouse Orc4p    --------NL FGV-----QV QYKHL----- -IE-LIKRTA IYGESNSVLI

yeast Orc4p    VGPRQSMKTY LLDYELSLLQ QSYKEQFITI RLNGFIHSEQ TAINGIATQL
mouse Orc4p    VGPRGSGKT- ------TLLN HAIKE-LMFI EVSE--NVIQ VHINGL---L

yeast Orc4p    EQQLCKIHGS HEKIDDTSLE TISSGSLTEV FEKILLILDS TIKTRNEDSG
mouse Orc4p    QTN-HKI-AL KEITRQLNLD NVVEDKVFGS FAENLSFLLE AIK-----RG

yeast Orc4p    EVDRESITKI TVMFIHDEID IFAGPVRQTL LYNLFDMVEH SRVPVCIFGC
mouse Orc4p    --DRIS--SC PVIFIIDEFD IFAHQKNQTL LYNLFDISQS AQTPVAVIGL

yeast Orc4p    TTKLNILEML EKRVKSRFSQ RMIYMPQIQN LDDMVDAVRN ILTVRSEISP
mouse Orc4p    TCRLDILEIL EKRVKSRFSH RQIHLMNSFD FPQYLKIFKE QLSLPAEFPD

yeast Orc4p    W--VSQWNET LEKFLSDPRS NINRHIRMNF ETFRSLPTIK NSIIPLVATS
mouse Orc4p    KAFAERWNEN VH-QLSEDST VLE-VLQKHF SVNKNLQSLH --MLLMLALN

yeast Orc4p    KNFGSLCTAI KSCSFLDIYN KNQLSNNLTG RLQSLSDLEL AIILSAARVA
mouse Orc4p    RVTVSH-PFM TSADLMEAQH MCSLDSK-AN IVHGLSMLEI CLIIAMKHLN

yeast Orc4p    LRAKDGSFNF NLAYAEYEKM IKAINSRIPT VAPTTNVGTG QSTFSIDNTI
mouse Orc4p    DIYEEEFFNF QMVYNEFQKF IQ----R--- --------KA HSVYNFEKPV

yeast Orc4p    KLWLKKDVKN VWENLMQLDF FTEKSAVGLR DNATAAFYAS NYQFQGTMIP
mouse Orc4p    -------VMK AFEHLQQLEL IKPVERTSV- -NSQREYQLV KLLLDNTQIM

yeast Orc4p    FDLRSYQMQI ILQELRRIIP KSNMYYSWTQ L
mouse Orc4p    NALQKY--SN CPTDVRQ-WA TSSL--SWLZ -
```

Figure 4/3. Mouse EST clone AA110785 encodes a full lenght mouse Orc4p which has 24% identities with the *S. cerevisiae* Orc4p.

Cloning of *Xenopus* Orc4 gene

A   Primary Screen
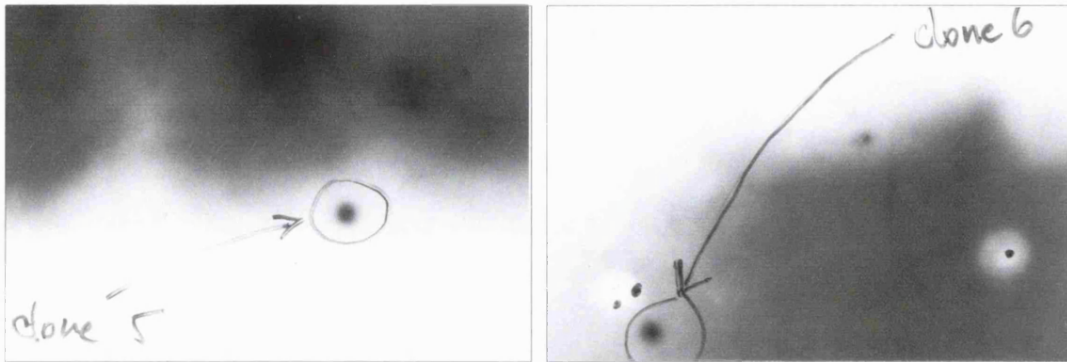
B   Secondary Screen

clone 5   clone 6
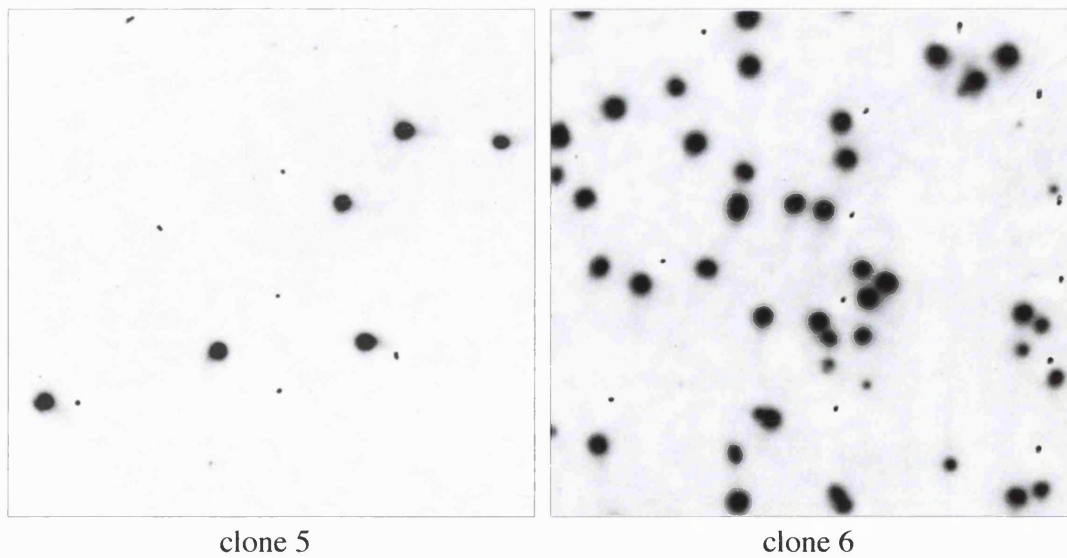
Figure 4/4. (A) 60 000 plaques of *Xenopus* total ovary cDNA library in λ-ZAP were screened with a radiolabelled probe derived from the mouse Orc4 gene. Primary screen identified six positives. (B) Clones 5 and 6 tested positive in the secondary screen.

amino acid sequence of *Xenopus* Orc4p is 69% identical to the mouse Orc4p with

homology through the length of the protein (Figure 4/5). However, as is the case for

the mouse counterpart, the overall homology between the *Xenopus* Orc4p and the

yeast protein is low, with only 22% identities. The homology between the mouse

Orc4p and the human Orc4p, which was meanwhile independently cloned in two

different labs, is even higher than that between frog and mouse with 89% identities

(Quintana *et al.*, 1997; Tugal *et al.*, 1998; Figure 4/5).

The amino-acid sequence of the identified Orc4p contains both peptides originally

microsequenced, which led to the identification of this gene (Figure 4/5). In addition,

the 43kDa polypeptide was microsequenced by Ryuji Kobayashi in Cold Spring

Harbour by Edman degradation method. This procedure identified additional four

peptides from this protein. All peptide sequences obtained by microsequencing of

Table 4/1

Microsequencing of *Xenopus* Orc4p

Peptide	Peptide sequence	Orc4p sequence
MS-4/1	(K/R)TV(L/I)HGESNSA(L/I)(L/I)(L/I)GPR	RTVIHGESNSALIIGPR
MS-4/2	(K/R)Q(L/I)ED(L/I)(F/M)G(F/M)K	KGALEDIFGMK
ED-4/1	LSLPASFPDsqFfEk	LSLPASFPDSQFAEK
ED-4/2	LQDIYxGexf	LQDIYDGEPF
ED-4/3	xLVELxK	HLVELLK
ED-4/4	aLEDIFG	ALEDIFG

MS, mass spectrometry; ED, Edman degradation
Lower case letters indicate amino acids identified with lower certainty

Orc4p is highly conserved between frog, mouse and human

```
frog Orc4p    MSKRKSKDLC LPVGECISKV HGILRQRLFQ HHGKP--FGV DSQHKHIMEL
mouse Orc4p   MSSRKIKSNA --HAECISCV QRILRERFCH HSPHSNIFGV QVQMKHLIEL
human Orc4    MSSRKSKSNS LIHTECISCV QRILRERFCR QSPHSNIFGV QVQMKHLSEL
                                                          peptide ED-4/3

frog Orc4p    LKRIVIHGES NSALIIGPRG SGKSMLLKGA LEDIFGMKQM KETAIQVNLN
mouse Orc4p   LKRIAIYGES NSVLIMGPRG SGKTTLLNHA LKELMEIE-V SENVIQVHLN
human Orc4    LKRIALHGES NSVLIIGPRG SGKTMLISHA LKELMEIEEV SENVIQVHLN
                    peptide MS-4/1       Motif A      peptides MS-4/2 & ED-4/4

frog Orc4p    GLLQTTLKIA LKEITRQLHL HNVVGDRVFG SFAENLSFLL EALKIGDRKS
mouse Orc4p   GLLQTNEKIA LKEITRQLNL INVVEDKVFG SFAENLSFLL EALKKGDRIS
human Orc4    GLLQINLKIA LKEITRQLNL HNVVGDKVFG SFAENLSFLL EALKKGDRIS

frog Orc4p    SCPVLFMLDE FDIFAHHKNQ TLLYNLFDIA QSAQTPVAVI GLTCRLDVME
mouse Orc4p   SCPVIFILDE FDIFAHQKNQ TLLYNLFDIS QSAQTPVAVI GLTCRLDILE
human Orc4    SCPVIFILDE FDIFAHHKNQ TLLYNLFDIS QSAQTPIAVI GLTCRLDILE
                   Motif B

frog Orc4p    LLEKRVKSRF SHRQIHLLNS FSFSQYLCIF QEKLSLPASF PDSQFAEKWN
mouse Orc4p   LLEKRVKSRF SHRQIHLMNS FDFHQYLKIF KEQLSLPAHF PDKAFAEHWN
human Orc4    LLEKRVKSRF SHRQIHLMNS FGFHQYVKIF KEQLSLPAHF PDKVFAEKWN
                                                        peptide ED-4/1

frog Orc4p    ESIKSLMESK IVELVLQKQY NASKDVHSLH MIMILAVGRV NVSHPHITAA
mouse Orc4p   ENVHQLSEDS TVLEVLQKHF SVNKNLQSLH MILMLALNRV TVSHPFMTSA
human Orc4    ENVQYLSEDR SVQFVLQKHF NISKNLFSLH MILMLALNRV TASHPFMTAV

frog Orc4p    DFIEVFRLRN QDSKANILHG MSVLELCLII AMKHILQDIYD GEPFNFQMVH
mouse Orc4p   DLMEAQHMCS IDSKANIVHG ISVLEICLII AMKHILNDIYE HEPFNFQMVY
human Orc4    DLMEASQLCS MDSKANIVHG ISVLEICLII AMKHILNDIYE HEPFNFQMVY
                                                        peptide ED-4/2

frog Orc4p    NEFQKFIQRK AHSVYNFEKA VVIKAFEHLH QLELIKPMEG ISVRTQKEYR
mouse Orc4p   NEFQKFIQRK AHSVYNFEKP VVMKAFEHLQ QLELIKPVER TSVNSQREYQ
human Orc4    NEFQKFVQRK AHSVYNFEKP VVMKAFEHLQ QLELIKPMER TSGNSQREYQ

frog Orc4p    LMKLLLDNTQ IVEALQKYEN CPTDVKQWAM SSLSZ--
mouse Orc4p   IMKLLLDNTQ IMNALQKYSN CPTDVRQWAT SSLSWLZ
human Orc4    IMKLLLDNTQ IMNALQKYEN CPTDVRQWAT SSLSWL-
```

Figure 4/5. Frog sequence is 69% identical to the mouse sequence. Human and mouse sequences are 89% identical. Peptide sequences identified by microsequencing are underlined. Motifs A and B of the putative nucleotide binding site are labelled.

the 43 kDa protein, which co-purified with *Xenopus* Orc1p and Orc2p are given in Table 4/1. All identified peptide sequences can be found in the aminoacid sequence of the *Xenopus* Orc4p, confirming that the gene cloned indeed encodes the 43 kDa protein originally microsequenced (Figure 4/5).

Cloned Orc4p is part of frog ORC

For further characterisation of newly identified *Xenopus* Orc4p protein, I generated polyclonal antibodies directed against this protein. The full length Orc4 gene was amplified by PCR from the original library vector and was subcloned into the *Xho* I site in pET21b (see Methods) vector to create a C-terminal hexahistidine tag. The fusion protein was produced in bacteria (see Methods) and purified on nickel agarose beads under denaturing conditions (see Methods). Figure 4/6 shows that the molecular weight of the expressed and purified polypeptide was approximately 45 kDa, slightly higher then expected for the full length Orc4p protein (Figure 3/9), presumably due to the fused histidine tag. A high molecular weight contaminating polypeptide is also present in the purified Orc4p sample. Two rabbits were injected with the purified Orc4p protein and the preimmune sera were compared to the immune sera. Because these antibodies had the tendency to recognise a number of proteins non-specifically, the immunoblotting with these antibodies was always done in the presence of 10% milk and 2% yeast extract. Figure 4/7 shows that the antisera produced by both rabbits (named as TK8 and TK9) recognised Orc4p protein on the immunoblot of *Xenopus* egg extract.

The obtained anti-Orc4p antibodies were used to determine whether the cloned Orc4p is indeed component of *Xenopus* ORC. The ORC complex was purified in essentially the same way as described in the previous chapter (Figure 3/9) using the anti-Orc1p monoclonal antibody TK15. Figure 4/8 shows that the newly cloned *Xenopus* Orc4p

Production of anti-Orc4p polyclonal antibodies

(6) Purification of recombinant Orc4p
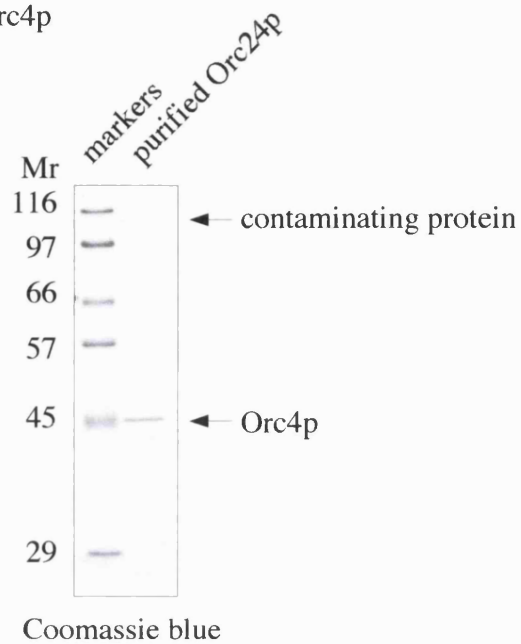
Coomassie blue

Figure 4/6. Recombinant Orc4p was produced in bacteria and purified on nickel agarose beads. Sample of the purified Orc4p protein was run on gel and stained with Coomassie blue. Size of purified Orc4p protein was compared to that of molecular weight markers.

(7) Immunoblot of frog egg extract with anti-Orc4p antisera

rabbit 1 (TK 8)   rabbit 2 (TK 9)
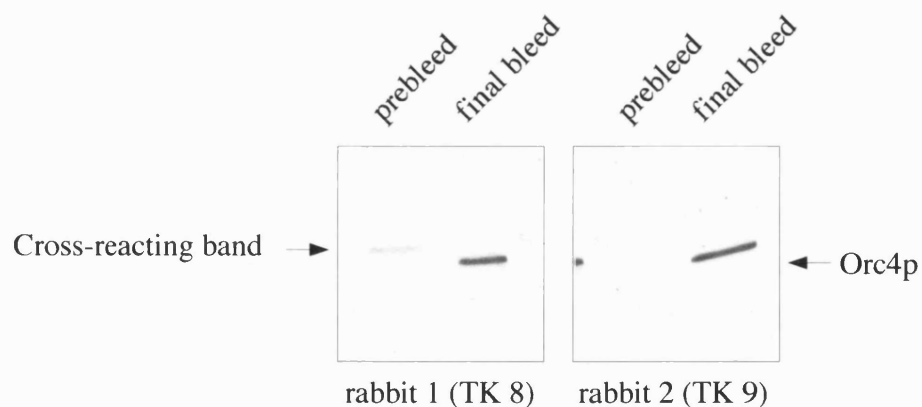
Figure 4/7. Total of 0.3 μl of *Xenopus* egg extract was run per lane and immunoblotted to test for the presence of anti-Orc4p antibodies in the sera of rabbits 5 and 6. Antisera produced by both rabbits recognised a single band of a molecular weight corresponding to that of Orc4p on the immunoblot of frog egg extracts.

Cloned Orc4p is part of the ORC complex
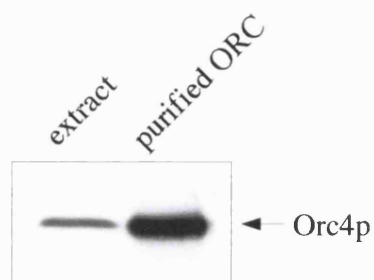
Figure 4/8. Interphase extract from *Xenopus* egg was incubated with protein A beads coupled to anti-Orc1p mAb TK15 and the bound proteins were resolved by SDS-PAGE and immunoblotted with anti-Orc4p polyclonal antibodies. Lane 1; total interphase extract.

co–immunorecipitated with anti-Orc1p monoclonal antibody confirming that it is indeed part of the ORC complex.

Intriguing Orc4p features

This chapter described an identification of a novel subunit of *Xenopus* ORC which is homologous to Orc4p of *S. cerevisiae* (Bell *et al.*, 1995; Tugal *et al.*, 1998). In parallel with this work, two groups identified a human homologue of the *S. cerevisiae* Orc4p (Quintana *et al.*, 1997; Tugal *et al.*, 1998). Soon after, sequences encoding the Orc4p in *S. pombe* and in *Drosophila* have been identified (Chesnokov *et al.*, 1999; Chuang and Kelly, 1999).

The deduced nucleotide sequence of Orc4p identified from frog, mouse, humans and from *S. pombe* revealed several interesting features of this ORC subunit. Surprisingly, a nucleotide binding site consensus sequence is conserved between frog, human, mouse, *Drosophila* and *S. pombe,* whereas the match found in the A motif of the nucleotide binding site in *S. cerevisiae* Orc4p is much weaker (Koonin, 1993; Walker *et al.*, 1982). *S. cerevisiae* Orc4p has been shown to cross-link ATP, when part of the *S. cerevisiae* ORC complex interacting with ARS1 DNA. However, mutations which removed this motif showed no gross phenotypic defect (Klemm *et al.*, 1997; see Introduction). Furthermore, the cross-linking to ATP was completely abrogated when the A motif of the ATP binding site in Orc1p was mutated, indicating the Orc4p-ATP binding might have been indirect or that it at least requires a functional Orc1p. It is not clear at present whether the ATP-binding site of Orc4p is functional in frogs, and is so, what role does it play. Further biochemical analysis is necessary to address this question.

Analysis of Orc4p sequence revealed that is strikingly similar to Orc1p and Cdc6p, both involved in control of DNA replication. The similarities between Orc1p and Cdc6p have been described before and now Orc4p was found to be related to the Orc1p and Cdc6p protein families (Bell *et al.*, 1995). Figure 4/9 shows that the conserved region includes the Walker A and the Walker B motifs of the putative nucleotide binding site as well as additional amino acid sequence blocks. The sequence similarities between Orc1p, Cdc6p and Orc4p suggest that the three proteins might have similar structure or that they all evolved form a common ancestral protein.

The identification of the Orc4p from *S.pombe* uncovered yet another peculiarity of this protein (Chuang and Kelly, 1999; see Introduction). The *S. pombe* Orc4p protein consists of two distinct functional domains. The C-terminal domain shows strong sequence similarity to human, frog and yeast Orc4p proteins. The N-terminal domain contains nine copies of the AT-hook motif found in a number of DNA-binding proteins. AT-hook motifs are known to mediate binding to the minor groove of AT-tracks in DNA. This special feature of Orc4p of *S. pombe* offers an alternative explanation for origin sequence recognition by ORC in this organism. Similar motif was however not found in Orc4p from other organisms.

Chapter 5

Other proteins associated with frog Orc1p

Following the successful identification of *Xenopus* Orc4p achieved through microsequencing of the 43 kDa subunit of the ORC complex, purification of which is described in Chapter 3 of this thesis (see Figure 3/14), the remaining proteins co–purified with *Xenopus* Orc1p were microsequenced in search for homologues of *S. cerevisiae* Orc3, Orc5 and Orc6 genes from higher eukaryotes. Peptide sequences obtained from the 48 kDa polypeptide led to identification of the human Orc5 gene. Microsequencing of the 68 kDa protein helped to confirm the identity of the *Xenopus* Orc3p. However, the attempts to microsequence the 27 kDa co-purified protein did not lead to identification of any new gene. In addition, a new 50 kDa polypeptide weakly associated with *Xenopus* Orc1p was discovered. However, neither cdc6, cyclin A, cyclin E or cdc2 were found to be associated with ORC immunoprecipitated from frog interphase extract.

Identification of human Orc5p

The 48 kDa polypeptide co-purified with Orc1p, Orc2p and Orc4p in the immunoaffinity purification described in Chapter 3, Figure 3/14 was microsequenced by both mass spectrometry in Protein Sequencing Unit of ICRF and by Edman degradation done by Ryuji Kobayashi in Cold Spring Harbour Laboratory. The identified peptide sequences were then used by Helena Zou-Yang in Bruce Stillman's lab to identify, in conjunction with EST database searching, the human homologue of

S. cerevisiae Orc5 gene. The peptide sequences initially identified by microsequencing of the 48 kDa frog protein associated with Orc1p are shown in Table 5/1 alongside with the actual protein sequence of the newly identified human

Table 5/1

Microsequencing of *Xenopus* Orc5p

Peptide	Frog 48 kDa peptide sequence	Human Orc5p sequence
MS-5/1	(K/R)ESQVST(L/I)(L/I)S(L/I)(F/M)GER	RESQVSILQSLFGER
MS-5/2	(K/R)TY(L/I)(L/I)QTV(L/I)R	KTYVTQTLLK
MS-5/3	(K/R)TV(L/I)(F/M)D(L/I)VR	RTVNFDIIK
MS-5/4	(K/R)GHTVS(L/I)N(F/M)(L/I)R	KCTVSLDFIR
ED-5/1	NIEPhLK	NIEPHLK
ED-5/2	rFFLK	RFFLK
ED-5/3	LSAHAHVELPYYSK	LSAHTHVELPYYSK

MS, mass spectrometry; ED, Edman degradation
Lowercase letters indicate amino acids identified with lower certainty

Orc5p. Figure 5/1 shows the sequence alignment of the identified human Orc5p sequence with *D. melanogaster and S. cerevisiae* Orc5p (Gossen *et al.*, 1995; Loo *et al.*, 1995). The human Orc5p is 39% identical to the *Drosophila* Orc5p and 29% identical to the *S. cerevisiae* Orc5p. Although the cDNA encoding *Xenopus* Orc5p has not been identified, its protein sequence is presumably highly related to the human sequence as the peptides identified by the microsequencing of the frog protein are almost identical to the human protein sequence and the calculated molecular

Alignment of Orc5p sequences from yeast, drosophila and human

```
drosophila  MEAICSSLEP LFPQREAAIE TLGELIGDSS E-TYPSAIYL FGHSGIGKTA
human       M----PHLEN VVLQRESQVS ILQSLFGERH HFSFPS-IFI YGHTASGKTY
yeast       M----NVTTP EVAEREYQTN CLASYISADP -DITPSNLIL QGYSGIGKTY
                            peptide MS-5/1                        Motif A

drosophila  LTRAFLKECG KRQNVRTAHL NAIECYTTKI MLEIL----- --LDSL----
human       VTQTLLKTL- ELPHV---FV NCVECFTLRL ILEQI----- --LNKLNHLS
yeast       ---TLKKYFN ANPNLHAVWL EPVELVSWKP LLQAIARTVQ YKLKTLYPNI
            peptide MS-5/2

drosophila  APDQGDALKV --DNMLDFVE QLRRQAAT-- RVEDQG-FLI AVDNAERLRD
human       SSEDGCSTEI TCETFNDFV- RLFKQVTTAE NLKDQT-VYI VLDKAEYLRD
yeast       PTTDYDPLQV E-EPFL-LVK TLHNIFVQYE SLQEKTCLFL ILDGFDSIQD
                                                         Motif B

drosophila  MDANVLPVLL RLHEI--TNL NLCVILLSQL PFEKFYNKTG LSEIVCLHLA
human       MLANLLPGFL RLHEI--ADR NVTVLFLSEI VWEKFRPNTG CFEPFVLYFP
yeast       LLAALFNKYI KLNEILPKDS KINIKFIYTM LETSFLQRYS THCIPTVMFP

drosophila  QYNKAETQRI IGSDFQQVRN QLLEQFAQDK KRLEICQEAV TED-FYN---
human       DYSIGNLQKI ISHDHPP--- ---EYSA--- ---------- --D-FYA---
yeast       RYNVDEVSTI I---VMSRCG ELMEDSCLRK RIIEEQITDC TDDQFQNVAA

drosophila  NYLNLFLSVF YKAC-RDVPE LQLTARKCLS TYLEPVLDGT VDATDISRLW
human       AYINILLGVF YTVC-RDLKE IRHLAVLNFP KYCEPVVKGE ASERDTRKLW
yeast       NFIHLIVQAF HSYTGNDIFA LNDLIDFKWP KYVSRITK-- ------ENIF

drosophila  RHIAGHIRSA LTQIYMR-IE KPA-EEVE-D FTAIEDQSVR KL-AQS-LEL
human       RNIEPHIKKA MQTVYLREIS SSQWEKLQKD DT--DPGQLK GLSAHTHVEL
yeast       EPLA--LIYKS AIKLFL-STD DNLSENGQGE -SAITTNRDD LENSQTY-DL
            peptide ED-5/1                                peptide ED-5/3

drosophila  PYYAKFILIA AFLASHNAAK QDKRLFVKHH GKQRKRMQTV NARAKTTEKM
human       PYYSKFILIA AYLASYNPAR TDKRFFLKHH GKIKK----T NF-LKKHEKT
yeast       SIISKYILIA SYIQSYLEPR YDASIFSR-- -KTRIIQGRA AYGRRKKKEV
                                          peptide ED-5/2

drosophila  ST-TLGPKSF SIDRLLAIFY AI-------- ------LEEK VGLTCNI--L
human       SNHLLGPKFF PLDRLLAILY SI-------- ------VDSR VAPTANI--F
yeast       NPRYLQPSIF AIFRLLAIFQ AIFPIQGKAE SGSLSALREE SLMKANIEVF

drosophila  SQISTLVHLN ILSFVSGEQ- NIMEGSARLQ CTIGLEFVLQ IGKVVGFNVR
human       SQITSLVTLQ LLTLV-GHD- DQLDG-PKYK CTVSLDFIRA IARTVNFDII
yeast       QNLSELHTLK IIATTMNKNI DYLSPKVRWK VNVPWEIIKE ISESVHFNIS
                                                 peptide MS-5/4    peptide MS-5/3

drosophila  QYLQDFM   460
human       KYLYDFL   435
yeast       DYFSDIHE  479
```

Figure 5/1. The human Orc5p is 39% identical to the *Drosophila* Orc5p and 29% identical to the *S. cerevisiae* protein. Underlined the original peptides microsequenced from the frog protein. Motif A and Motif B of the nucleotide binding site are marked.

weight of the human protein is very similar to the molecular weight of the frog protein microsequenced (Tugal *et al.*, 1998). Interestingly, Orc5p shows weak homology to the entire Orc4p and to the C-terminus of Orc1p and Cdc6p (Tugal *et al.*, 1998). It is therefore likely that all four proteins evolved from an ancestral protein and that the retained sequence similarity is required for a common function such as oligomerisation or an interaction with the DNA.

Identification of *Xenopus* Orc3p

During the course of experiments described in these thesis, the *Xenopus* ORC complex was purified by Phil Carpenter using anti-Orc2p polyclonal antibodies (Carpenter and Dunphy, 1998). Microsequencing of co-purified p81 protein led to identification of a novel component of *Xenopus* ORC. Initially, it was not clear, whether this protein was identical to the 68 kDa protein that co-purified with Orc1p in immunopurification described in Chapter 3, Figure 3/14, because of the discrepancy in the estimated molecular weight between the two proteins. The 68 kDa

Table 5/2

Microsequencing of *Xenopus* Orc3p

Peptide	The 68 kDa protein	Sequence of p81
ED-3/1	tDLYQLQK	TDLYQLQK
ED-3/2	LEEFLINFQSLEETTQNEED	LEEFLINFQSLEETTQNEED
ED-3/3	LFENLIxFLRE	LFENLIGFLRK

ED, Edman degradation
Lowercase letters indicate amino acids identified with lower certainty

Orc1p associated protein described in Figure 3/14 was microsequenced by Edman

degradation done by Ryuji Kobayashi. The obtained peptides shown in Table 5/2

matched extremly well the deduced nucleotide sequence of the Orc2p associated p81

protein identified by Phil Carpenter. To determine whether the two proteins were

identical, *Xenopus* ORC was immunopurified as described in Chapter 3, resolved by

SDS-page and analysed by immunoblotting with anti-Orc1p, anti-Orc2p, anti-Orc4p

and anti-p81 antibodies (the anti-p81 antibodies were kindly provided by Phil

Carpenter). Figure 5/2 shows the anti-p81 antibody recognised the Orc1p-associated

68 kDa protein described in Chapter 3, Figure 3/14. We therefore concluded that the

68 kDa Orc1p associated protein described in Chapter 3 is in fact identical to the

Orc2p-associated p81 identified by Phil Carpenter, with the discrepancy in the

molecular weight being attributed to the difference in the composition of the

polyacrylamide gels used by the two labs.

The p81 is a likely candidate for the missing homologue of *S. cerevisiae* Orc3p.

Figure 5/3 shows the sequence alignment of p81 with the *S. cerevisiae* Orc3p and the

recently identified Orc3p from *Drosophila*. The homology between the yeast Orc3p

and frog Orc3p candidate p81 is extremely low with only 14% identity. However,

there is 30% identity between the *Xenopus* and *Drosophila* Orc3p. Despite the

sequence conservation of Orc3p being extremely low, based on the fact that this

protein co-purifies with Orc1p, Orc2p, Orc4p and Orc5p in both frogs and flies, we

can conclude with a reasonable degree of confidence that the Orc2p associated p81

identified by Phil Carpenter (identical to 68 kDa protein identified in this work) is the

Xenopus homologue of the *S. cerevisiae* Orc3p.

The anti-p81 antibody recognised the Orc1p-associated 68 kDa protein
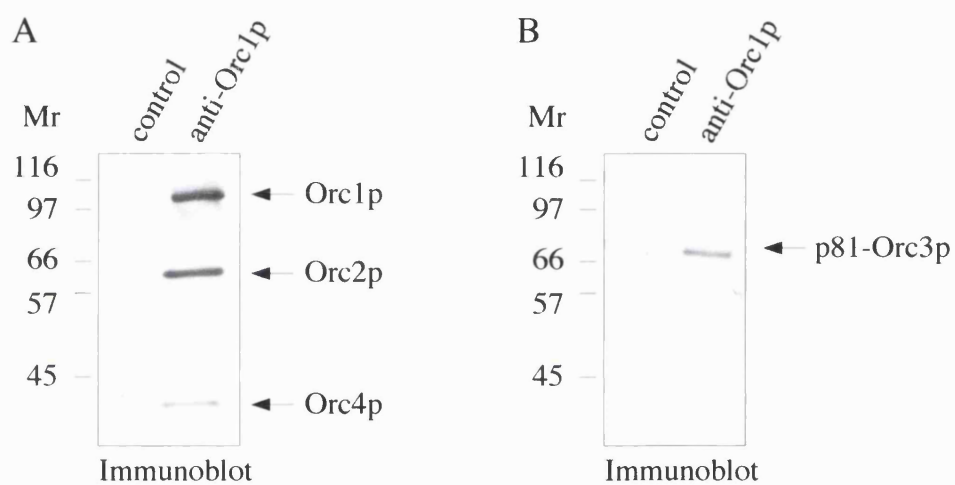
Figure 5/2. Interphase extract from *Xenopus* eggs was incubated with protein A beads coupled to either control anti-SV40 T-antigen mAb 423 or anti-Orc1p mAb TK15. The bound proteins were analysed by SDS-PAGE and detected by immunoblotting with anti-Orc1p, anti Orc2p, anti-Orc4p (panel A) and anti-p81 (panel B) polyclonal antibodies.

Alignment of Orc3p from *S. cerevisiae*, *Drosophila* and *Xenopus*

```
yeast         M-SDLNQSKK MNVSEFADAQ RSHYTVYPSL PQSNKNDKHI PFVKLLSGKE SEVNVEKRWE
drosophila    MDPTISVSKG CFVYKNGATR AGKKAASKRK RPAAESSSLL GKEVVQ--QP FYEEYRKAWN
frog          M-TTSSVSKG CFVFKPSAKK KKTSLTVADY FNEGLRDSED SKKRFESCQL LWQQMKSQTE


LYH-QLHSHF HDQVDHIIDN IEADLKAEIS DLL--YSETT QKRRCFNTIF LLGSDSTTKI ELKD--ESSR
QIN-DHIADL QHRSYARTLE QLVDFVVGQA ERD--TPDEV LPTAALLTGI NQPDHLSQFT ALTQRLHAQR
QTSRGNEKRL FENLIGFLRK SHADFHNKKD DWSCRMRASE IPTAALVLGV NVTDHDLTFN SLSDILHETI
           peptide ED-3/3

YNVLIELTPK ESPNVRMMLR RSMYKLYSAA DAEEHPTIKY EDINDEDGDF T----EQ--N NDVSYDLSLV
AAMVCVLQSR DCATLKAAVE TLVFGLVEDN AEVEQMEDED EDEDGAERDR K----RL--R RSQCTMKQLK
TPFVVLLQSK ECTGIKQLLQ KLLTQLMGNT VDID-LEEEE EQVTISQRKM NCTLASLSGW YKRATKKSAS


ENFKRLFG-- --KDLAMVFN FKDVDSINFN TLDNFIILLK S---AFKYDH V-KI-SLIFN INTNLS-NIE
SWYTNNFDSE Q-KRRQLVVI LPDFECFNAS VLQDLILILS AHCGSLPFVL VLGVATAMTA VHGTLPYHVS
PKKKRSLTST HWESPPVVVI FKDLESFTAS VLQEFIVISS GYVQDLPLVL VFGIATSPMI IHRILSHSVS


KNLR------ QSTIRLLKRN YHKLDVSSNK GFKYGNQIFQ ----SFL--- ----DTVLG- KLNLSDRFVE
SKIRLRVFQT QAAPTGLNEV LDKVLLSPKY AFHLSGKTFK FLTHIFLYYD FSIHGFIQGF KYCLMEHFFG
SRLCIELFQS MSCTEHLATV VDQLLLTNHF PIKLSGRVMQ VLITIFLYHD FSVQNFIKGL QLSVVEHFYT


FILSKMANNT NHNLQLLTKM LDYSLMSYFF QNAFSVFIDP VN-----VDF L-NDDYLK-- ---ILSR---
GNAFALCTDY SKALGRIKQL THEDMETIRR LPSFRPYVEQ INDCKRIIAV LTDDDYLKKK LPQLL-R--D
QPLSVLCCSL SESRKRIKNL SHAQCENIRH LSSFMSYVE- SQTPENQVNL LTNDRFLKEM TQEFLERLNS


CP-TF--MF- --------FV EGLIK--QHA PADFILSL-L TNKNRGLEEF -----FVEFL VRENPINGHA
CLLHFL-LFR CSLEFLTELV GDLPRCPLGK LRRELYVNCL NRAIISTPEY KECLQMLSFL SKDEFVAKVN
YHENFTPILR C----LHHFT CILPKYPLGK QIREIYCACL EKKVWETEDY NSALPLLRML AKDEIVATLQ


KFVAR---FL -EE----EL- ---------N ITNF------ -------NLI ELYHN---LL IGKLDSYLDR
RALERTEQFL VEEIAPLELG EACTAVLRPK LEAIRLAVDE VVKATMATIT TTSPNETRQA TDHLTPVASR
KCVAVLKPYF -EK----KLG SA----L-EK LEEF------ LIN--FQSLE ETTQNEEDED TSPQKSLQKK
                                                       peptide ED-3/2

-------WSA CKEYKDRLHF E-F---IDTI FQELFTLDN- ---RSGLLT- QSIFP----- SYK--SNIED
QELKDQLLQR SKEDKMRHQL NTPTTQFGRA LQKTLQLIET QIVQDHLRAL QDAPPIHELF VFSDIATVRR
TDL-YQLQKK LLEMKETRRT KKP-SRFELL RQDVVDFIDG -LVREYLLP- PEMIPLHEVV YFSAASTLRR
peptide ED-3/1

NLL--SWEQV LPSLDK-ENY ---DTLS-GD LDKIMA--PV LGQLFKLYRE ANMTINIYDF YIAFRETLPK
NIIGAPRAAL HTALNNPHFY MQCKCCELQD QSLLVGTLPD LSVVYKLHLE CGRMINLFCW LQAFRSVVSG
HLNAAPRVAL HTALNNPASY L--KCLE-NE GGSISNAAPD ICIAYKLHLE CGRLINLYCW LEAFATVVHA


EEILNFIRKD PSNTKLLELA ETPDAFDKVA LILFMQAIFA FENMGLIKFQ STKSYDLVEK CVWRGI  616
SD-------- --NE-----E VAQEQIDPQI QARFTRAVAE LQFLGYIKMS KRKT-LHATR LTW---  721
AE-------- -GSE-----S DSAQQVDDVT HARFIRAVSE LELLGFVKPT KQKT-DHVAR LTWGGC  709
```

Figure 5/3. Yeast Orc3p protein sequence is 14% identical to the frog Orc3p (p81). The *Drosophila* Orc3p protein sequence is 30% identical to the frog protein. The peptides microsequenced are underlined.

131

Microsequencing of the 27 kDa Orc1p associated protein

The 27 kDa polypeptide co-purified with *Xenopus* Orc1p, Orc2p, Orc3p, Orc4p and Orc5p in the immunoaffinity purification described in Chapter 3, Figure 3/14 was microsequenced by Ryuji Kobayashi in Cold Spring Harbour Laboratory by the Edman degradation method. Two peptide sequences were identified from this protein: LLVGDGHMGYPEEXPY and PGGRLILPVGPAGGNQMXE. Both these peptides matched extremely well a nucleotide sequence of an enzyme called protein-L-isoaspartate (D-aspartate) O–methyltransferase, strongly implying that the 27 kDa protein co-purified with Orc1p was a contaminant. Protein L-isoaspartate (D-aspartate) O-methyltransferase repairs damage to proteins caused primarily by spontaneous deamination of protein asparaginyl residues, by methylating the deaminated proteins (Szymanska *et al.*, 1998). It is therefore possible that this kind of protein repair protein might be associated with proteins, which are stored in the frog oocyte for substantial amount of time (up to 6 months) before the mature egg is laid.

The identification of the 27 kDa protein co-purified with *Xenopus* Orc1p as contaminant revived the question of the identity of the frog homologue of the yeast Orc6p. The *Drosophila* Orc6p protein is 30 kDa, supporting the idea that the *Xenopus* Orc6p will be found in similar molecular weight region. In addition, there is no other candidate for Orc6p in the sample purified with the monoclonal antibody TK15. It is possible that both frog Orc6p and the contaminating protein-L-isoaspartate (D-aspartate) O–methyltransferase run with a similar size on a gel, and could perhaps be separated by further chromatography. It is also possible that *Xenopus* Orc6p is phosphorylated on multiple sites and therefore does not run as a single band on a polyacrylamide gel as it is in the case of *S. cerevisiae* Orc6p (Stillman-personal communication). Alternatively, the Orc6p could have been lost

from the complex during the purification procedure, or there might be no Orc6p in the *Xenopus* egg extract.

A novel 50 kDa protein loosely associated with ORC

During the development of ORC purification scheme described in Chapter 3, a different way of purifying ORC has been tested. The SW buffer (50 mM Hepes 7.2, 100 mM KCl, 250 mM sucrose and 1 mM $MgCl_2$) used for this purification procedure was even milder that that used for the experiments in Chapter 3. Anti-Orc1p monoclonal antibodies TK1 and TK47 and a control, anti-SV40 T-antigen monoclonal antibody were covalently coupled to protein A beads as described in Methods. The immunopurification was performed with extract which was diluted 1:1 with SW buffer and centrifuged in the TL 100 rotor at 50 000 rpm for an increasing amount of time. Figure 5/4 shows that immunoprecipitates which were washed with the mild SW buffer contain a lot of contaminating proteins in comparison to the immunoprecipitates washed with more stringent buffer described in Chapter 3. However, when the extract is pre-cleared by high speed centrifugation prior to the immunoprecipitation, a number of contaminating proteins, especially in the lower molecular weight range are removed. When the immunoprecipitation is done using extract which was pre-cleared by high speed centrifugation for two hours, Orc1p, Orc4p and Orc5p are clearly recognisable on Coomassie blue stained gel as proteins which are found in both immunoprecipitates obtained by using anti-Orc1p monoclonal antibodies and not in the control immunoprecipitate. There also is a new band of an approximate molecular weight of a 50 kDa associated with anti-Orc1p antibodies, but not with the control antibodies. This band is not found in the immunopurified ORC described in Chapter 3, which were prepared under more stringent conditions and it is therefore likely to be associated with ORC only very

A novel 50 kDa band associated with *Xenopus* ORC
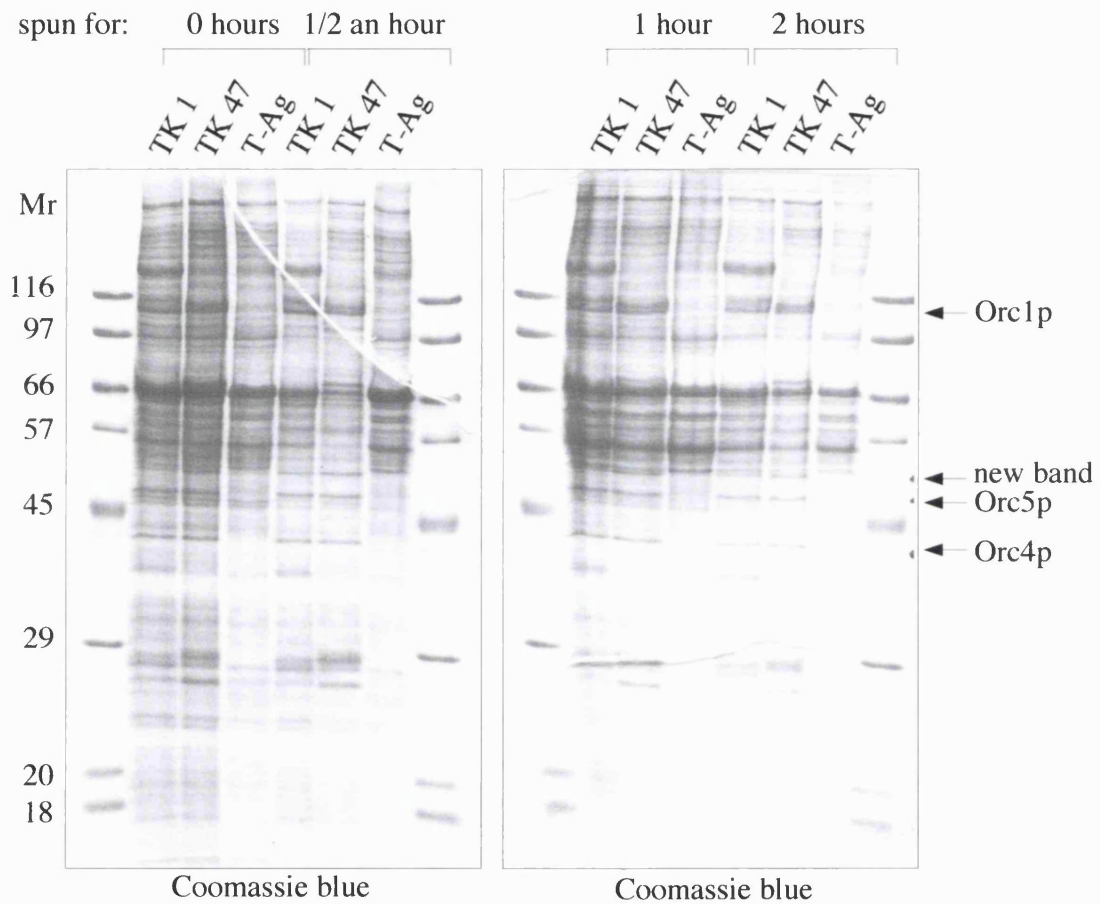
Figure 5/4. Interphase extract from *Xenopus* eggs was diluted with SW buffer and centrifuged in the TL 100 rotor at 50 000 rpm for an increasing amount of time. Following the centrifugation, the supernatant was incubated with protein A beads coupled to anti-Orc1p mAb TK1, anti-Orc1p monoclonal antibody TK47 or control anti-SV40 T-antigen mAb 423. The bound proteins were analysed by SDS-PAGE and detected by Coomassie blue staining.

The 50 kDa ORC associated protein
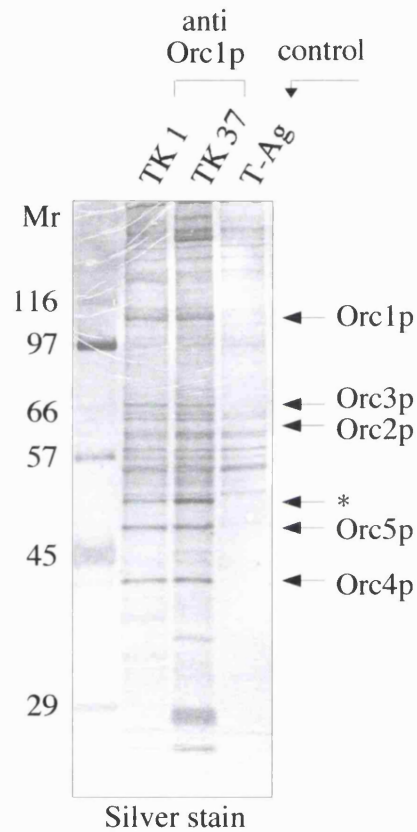
Figure 5/5. Interphase extract from *Xenopus* eggs was diluted in SW buffer and centrifuged in TL 100 rotor for 2 hours and the supernatant was incubated with protein A beads coupled to either anti-Orc1p mAbs TK1 and TK37 or to a control anti-SV40 T-antigen mAb 423. The bound proteins were analysed by SDS-PAGE and detected by silver staining. Polypeptides found in samples obtained with anti-Orc1p antibodies, but not in the control sample are marked with arrows.

weakly. This protein could either be the missing Orc6p subunit of *Xenopus* ORC or a novel protein associated with ORC. It would require microsequencing of this polypeptide to discover its identity.

To confirm the association of the novel 50 kDa protein with ORC, the purification was repeated several times using different anti-Orc1p antibodies. Figure 5/5 shows a silver-stained gel of such experiment.

ORC interaction with cdc6, cdc2, cyclin A and cyclin B.

In the purification of *Xenopus* ORC described in Chapter 3, Figure 3/14, number of proteins other than ORCs are associated with Orc1p immunoprecipitates, in particular in the case of anti-Orc1p monoclonal antibodies TK37 and TK47. We were interested whether any of the proteins implicated in regulation of DNA replication could be found in these immunoprecipitates. The immunopurification of ORC was therefore performed as described in Chapter 3, Figure 14 using the anti-Orc1p antibodies TK15 and TK37 and a control anti-SV40 T-antigen antibody. Figure 5/6 shows that all the tested proteins are detected in the sample of the interphase extract. None of them were immunoprecipitated with the control anti-Tag monoclonal antibody. Both anti-Orc1p monoclonal antibodies co-immunoprecipitate Orc1p, Orc2p and Orc4p, but not the Cdc6, cyclin E, cyclin A1 and Cdc2 proteins. We therefore conclude that Cdc6, cyclin A, cyclin E or Cdc2 are not associated with ORC in frog interphase extract, or at least not strongly enough to survive immunopurification, which is carried out under relative mild conditions.

The cdc6, cyclins E and A1 and cdc2 do not interact with ORC in *Xenopus* egg extract
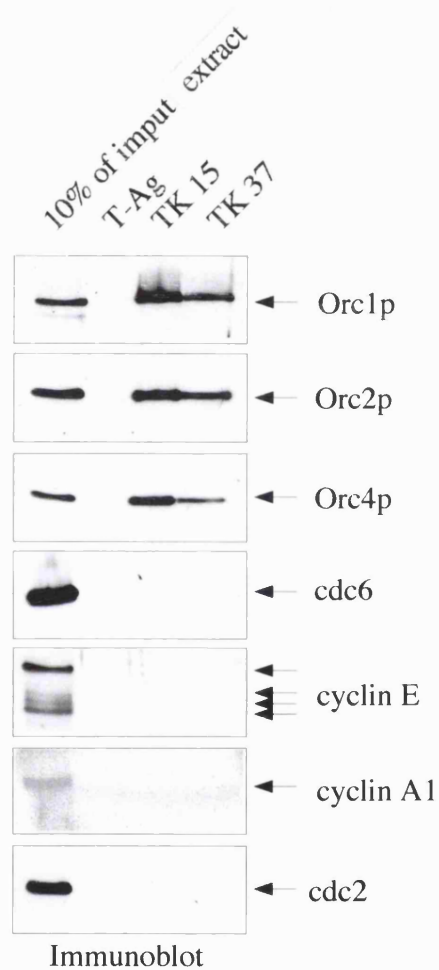
Figure 5/6. Interphase extract from *Xenopus* eggs was incubated with protein A beads coupled to either control anti-SV40 T-antigen mAb 423 or anti-Orc1p mAbs TK15 and TK37. The bound proteins were analysed by SDS-PAGE and detected by immunoblotting with anti-Orc1p, anti-Orc2p, anti-Orc4p, anti- cdc6, anti-cyclin E, anti-cyclin A1 and anti-cdc2 polyclonal antibodies.

Chapter 6

Orc1p is phosphorylated during cell cycle

In the course of experiments investigating Orc1p in progesterone treated oocytes we noticed that the Orc1p protein mobility on polyacrylamide gel changed as the oocytes progressed through the process known as oocyte maturation. This chapter describes studies of the nature of this change of Orc1p mobility and the timing of this event.

Mobility shift of Orc1p during oocyte maturation

To study the Orc1p during the early stages of frog development, oocytes were isolated from ovaries by manual dissection and were treated with progesterone in pools of ten for an increasing amount of time to observe the transition into meiosis II (see Methods). Figure 6/1 shows, that following 2 hours treatment with progesterone, a fraction of the Orc1p appears on the polyacrylamide gels as a slower migrating band. Timing of the appearance of the modified Orc1p corresponds to the time when the activation of mitotic kinesis is expected. The continuous presence of a fraction of Orc1p protein with the faster mobitily on the gel after 6 hours treatment may be due to a fact that a pool of the oocytes progress through maturation in an asynchronous manner.

Orc1p modification during oocyte maturation

hours after progesterone addition

0   1   2   3   4   5   6

◄— Orc1p
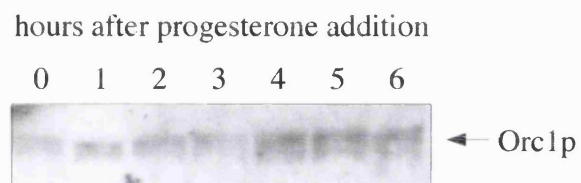
Figure 6/1. Oocytes were isolated from frog ovaries by manual dissection and incubated with progesterone for an increasing amount of time. Protein samples from these oocytes were prepared by extraction in HIK buffer. Pools of ten oocytes were used for each sample.

Orc1 p shows altered mobility in M-phase extracts

To investigate further the original observation that Orc1p mobility changes upon entry into second meiotic metaphase, Orc1p protein from CSF-egg extract (arrested in meiosis II) was compared to the extract which was induced by added $CaCl_2$ to enter interphase. The slow migrating form of the protein was identified in metaphase extracts (Figure 6/2 A, lanes 1 and 6-10) and the fast migrating form appeared upon Ca^{2+}-induced transition to interphase (Figure 6/2 A, lanes 2-5). To find out how well the Orc1p mobility shift correlated with the activity of meiotic kinases, their activity was measured by their ability to phosphorylate histone HI. The histone HI kinase assay was performed on both the CSF-arrested and the calcium induced samples. Figure 6/2 shows that in this time-course experiment, the shift to a faster migrating form occured concomitantly with the inactivation of mitotic kinases.

The mobility shift of Orc1p is due to phosphorylation

In order to find out whether phosphorylation was responsible for the Orc1p modification, we treated both, the CSF arrested extract as well as the calcium induced interphase extract with with potato acid phosphatase. Figure 6/3 shows that the slow migrating form of Orc1p, found in the CSF extract, was converted to the faster migrating one upon this treatment. Furthermore, this conversion was blocked by 50 mM NaF, a potent phosphatase inhibitor (Figure 6/3, lane5).

Cell cycle dependent Orc1p modification
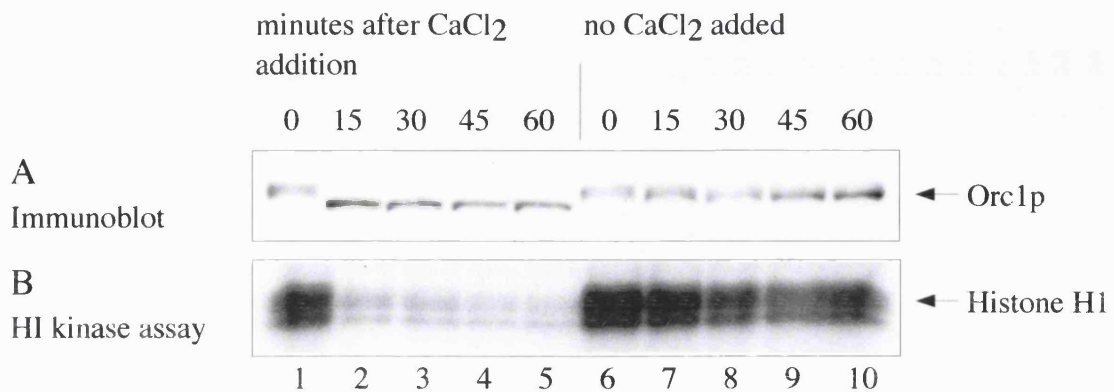
Figure 6/2. (A) CSF (meiotic) egg extract was incubated with CaCl2 for increasing lengths of time to trigger the transition to interphase (lanes 1-5). After 15 minutes incubation, Orc1p appeared on polyacrylamide gels as a faster migrating band (lane 2). There was, however, no change in the Orc1p mobility in samples where no Ca2+ was added and therefore remained in the meiotic phase (lanes 6-10). (B) The activity of mitotic kinases in samples incubated with Ca2+ as well as in the control samples was tested by their ability to phosphorylate histone HI. There was a tight correlation of the slow migrating form of Orc1p and the meiotic extract (lanes 1, 6-10). Conversely, the faster migrating form of Orc1p was found only in the meiotically inactive (interphase) samples (lanes 2-5).

Phosphorylation of *Xenopus* Orc1p
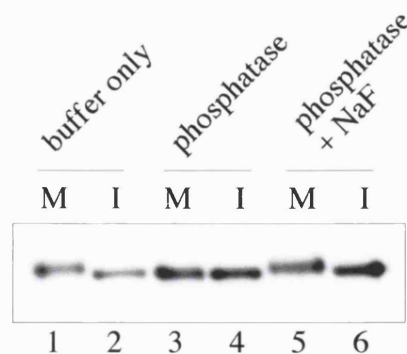
Figure 6/3. Both meiotic CSF extract labelled M (lanes 1, 3 and 5) and Ca2+ treated interphase extract labelled I (lanes 2, 4 and 6) were treated with buffer alone (lanes 1 and 2), or with potato acid phosphatase (lanes 3 and 4), or with phosphatase and 50 mM NaF (lanes 5 and 6).

Orc1p phosphorylation could regulate ORC function

This chapter shows that *Xenopus* Orc1p is phosphorylated in cell cycle dependent manner with the hyperphosphorylated form appearing in correlation with the activity of mitotic kinases. The phosphorylation of *Xenopus* Orc1p and Orc2p was also described by Phil Carpenter (Carpenter and Dunphy, 1998; Carpenter *et al.*, 1996) who has demonstrated that both *Xenopus* Orc1p and Orc2p are phosphorylated on both serine and threonine residue during M-phase, whereas the interphase forms appeared to contain mainly phosphothreonine, indicating the existence of at least three phosphorylation forms of both Orc1p and Orc2p. In Figure 6/1 of this chapter, there are possibly all three forms of Orc1p resolved, however gels shown in Figures 6/2 and 6/3 only resolve the mitotic from the interphase one.

Xenopus ORC was shown to be temporarily removed from condensed chromosomes during M-phase (Romanowski *et al.*, 1996). It has been suggested that the phosphorylation of ORC subunits may regulate such dissociation (Carpenter and Dunphy, 1998). Alternatively, ORC phosphorylation may serve to prevent the assembly of the pre-replicative complexes during mitosis (see Introduction). However, further experiments are needed to determine what is the function, if any, of this modification.

Chapter 7

Xenopus ORC in development

DNA replication during the cleavage cell cycles of early frog development proceeds much faster than in somatic cells because replication origins are spaced ten times more frequently. The initiation of DNA replication may be therefore regulated differently in embryos and in somatic cell cycles of frogs and other higher eukaryotes. This chapter attempts to identified the similarities and differences between the ORC complex present in frog eggs and somatic cells. No differences in Orc1p, Orc2p and Orc4p size and immunoreactivity have been detected between the embryonic and somatic ORC. All three proteins are found in a complex both in embryonic and somatic extracts. However, the ratio of ORC to the number of origins of DNA replication varies dramatically between the embryonic and somatic cell. The regulation of Orc1p expression during oogenesis is also described.

Quantification of Orc1p, Orc2p and Orc4p

To estimate the number of ORC molecules present in the frog egg extract, Orc1p, Orc 2p and Orc4p were produced in bacteria and purified on nickel agarose beads as described in Methods and in Chapters 3 and 4. The protein concentrations of the prepared purified protein samples were estimated by comparison to two different protein standards. Figure 7/1 show an immunoblot comparing the amounts of Orc1p, Orc 2p and Orc4p in the frog extract to that of the prepared protein standards.

Estimate of Orc1p, Orc2p and Orc4p levels in egg
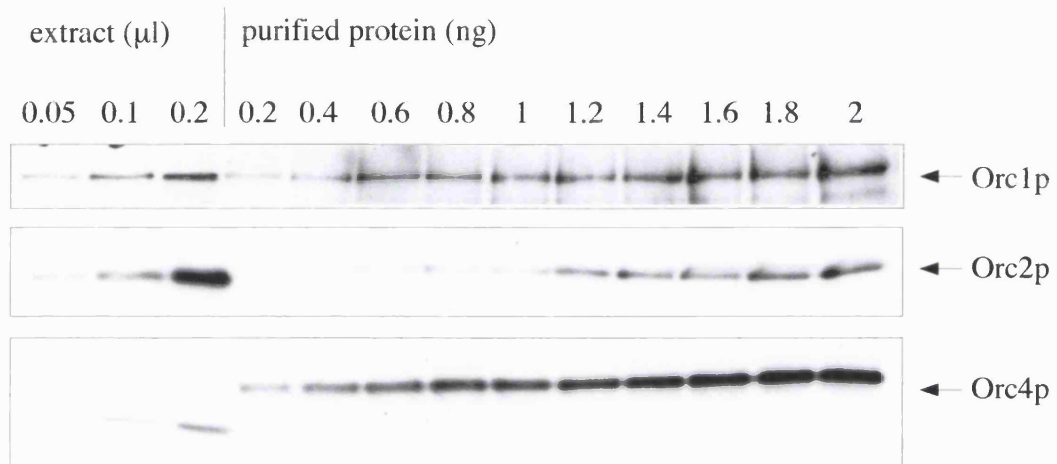
Figure 7/1. Orc1p, Orc2p and Orc4p were expressed in bacteria and purified on nickel agarose beads. The protein concentration of the purified proteins was determined by comparison to BSA and phosphorylase A protein samples of a known protein concentration on a Coomassie blue stainde gel. Samples of egg extract and of the purified proteins were analysed by SDS-PAGE and immunoblotted with anti-Orc1p, anti-Orc2p and anti-Orc4p antibodies.

The amount of Orc1p in the egg extract was determined by quantification of Orc1p immunoblot in Figure 7/1. Frog egg extract contains approximately 4 ng of Orc1p in 1 μl of extract. The molecular weight of Orc1p is 110 kDa and therefore the concentration of Orc1p in frog egg is approximately 36 nM. The amount of Orc2p and Orc4p present in the egg extract was determined in the same way to be 16 ng of Orc2p in 1 μl of extract and 1.5 ng of Orc4p in 1 μl of extract. The concentration of 64 kDa Orc2p frog egg extract is therefore 250 nM and the concentration of 43 kDa Orc4p is 34 nM.

According to the protein estimation described above, there is seven fold excess of Orc2p over the Orc1p and Orc4p. Although this type of protein estimation is not very accurate, it is unlikely for the mistake to be as high as seven fold. The possible significance of the excess of Orc2p is at present unclear. However, when looking of the Coomassie blue stained gel of the purified ORC (Chapter 3, Figure 14), there does not seem to be an access of Orc2p in comparison to the other ORC subunits in the purified complex. For the calculations latter on in this chapter, the concentration of ORC complex in frog egg extract will be assumed to be 35 nM.

Orc1p, Orc2p and Orc4p levels in frog egg and somatic cell

To establish the relative levels of individual ORC proteins present in *Xenopus* eggs compared to those in somatic cells, a sample of egg extract was run on the polyacrylamide gel alongside with a sample of extract from somatic (WAK) cells. Figure 7/2 shows that there is an equal amount of Orc1p, Orc2p, and Orc4p in 0.1 μl of egg extract (equals approximately 0.1 egg) as there in 10 000 somatic cells. In other words, there is the same amount of ORC proteins in one egg as there is in 100 000 somatic cells.

There is 100 000 times more ORC in an egg then in an somatic cell
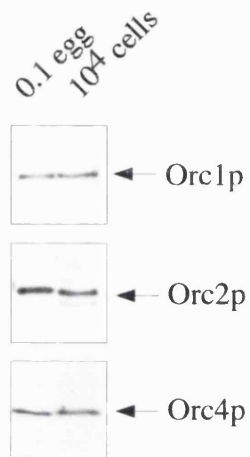
Figure 7/2. 0.1 μl of *Xenopus* interphase egg extract (equals approximatelly 0.1 egg) and an extract from 10^4 *Xenopus* somatic cells (WAK) were analysed by SDS-PAGE and immunoblotted with either anti-Orc1p, anti-Orc2p or anti-Orc4p antibodies.

In *Xenopus* genome, there are approximately 60 000 origins of DNA replication (the size of *Xenopus* diploid genome size is 6.14×10^9 bp, origin frequency is estimated to be one each 100 kb). As described earlier in this chapter, the concentration of ORC in frog egg extract is approximately 35 nM. This means that one microlitre of egg extract contains approximately 35 fmol of ORC (2.1×10^{10} molecules of ORC per microlitre of egg extract). One somatic cell contains 100 000 times less ORC then an egg. There is therefore 2.1×10^5 molecules of ORC in an somatic cell, making approximately 3.5 ORC complexes available per each replication origin. In an egg, however, there are 350 000 ORC complexes present per replication origin at the first S phase after fertilization, reducing to approximately 35 ORCs per origin at the MBT.

It is possible that this huge access of ORC complex to the replication origin in early frog embryo is responsible for the initiation of DNA replication from random sequences.

Orc1p, Orc2p and Orc4p are in complex in frog somatic cells

To determine whether Orc1p, Orc2p and Orc4p form a complex in frog somatic cell, as they do in egg extract, ORC was extracted from WAK cells in the presence of the same buffer as the one described in ORC purification from egg extract in Chapter 3 (also see Methods) and incubated with protein A beads which were coupled to either control anti-SV40 T-antigen antibody or to an anti-Orc1p antibody TK15. Figure 7/3 shows that Orc1p, Orc2p and Orc4p co-immunoprecipitated with the anti-Orc1p antibody, but not with the control antibody. This result confirms that not only Orc1p, Orc2p and Orc4p have the same size and immunoreactivity in both the embryonic and somatic extracts, but also they are found together in an complex in both types of extracts.

Orc1p, Orc2p and Orc4p are in a complex in frog somatic cells
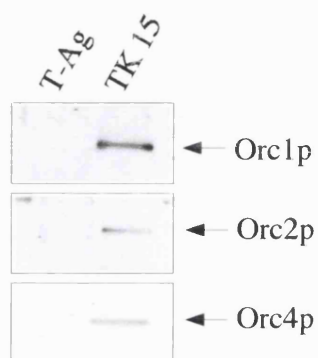
Figure 7/3. WAK cell lysate was incubated with protein A beads coupled to either control anti-SV40 T-antigen Mab 423 or anti-Orc1p Mab TK15. The bound proteins were analysed by SDS-PAGE and immunoblotted with anti-Orc1p, anti-Orc2p and anti-Orc4p polyclonal antibodies.

Orc1p expression is regulated in oogenesis

To investigate the regulation of Orc1p expression during the process of *Xenopus* egg development, oocytes of different developmental stages were compared. Figure 7/4 shows that there was no detectable Orc1p in early stage *Xenopus* oocytes. In the latter stages of the oocyte development (stages V and VI), a small amount of Orc1p was detected. It is important to realise that the majority of the oocyte proteins are expressed during this time, as this is the time when oocyte undergoes a significant increase in size. However, it can be seen in the Figure 7/4, the expression of Orc1p is largely restricted to the the period known as oocyte maturation, which is normally induced by treatment with progesterone. A possible function of this regulation could be restriction of premature DNA replication. Regulation of expression of other ORC subunits during egg development was not tested.

Orc1p expression during *Xenopus* egg development

oocyte stage | hours after progesterone addition

I II III V VI | 0 1 2 3 4 5 6 14

◄— Orc1p
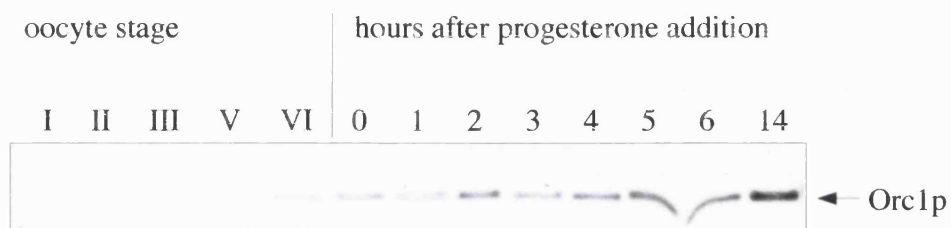
Figure 7/4. (A) Frog ovaries were treated with collagenase and oocytes of various developmental stages were isolated by manual dissection and sorted according to their developmental stage. Proteins present in the oocytes of different developmental stages were analysed by SDS-PAGE and immunoblotted with anti-Orc1p monoclonal antibody TK1. (B) Oocyte maturation experiment was done with stage VI oocytes manually dissected from the ovaries and treated with progesterone for 0-14 hours. Oocyte proteins were again analysed by SDS-PAGE and immunoblotted with anti-Orc1p monoclonal antibody TK1.

Chapter 8

ATP-dependent ORC-DNA interaction

Following the thorough characterisation of *Xenopus* ORC described in previous chapters of these thesis, I focused on investigation of DNA binding properties of this complex. I actually spent most of my time during my PhD trying to understand how are origins of DNA replication determined in the frog system. The experiments described in this chapter were built upon presumption that although DNA replication in *Xenopus* egg extract initiates from random sequences in ORC dependent manner, ORC itself would show some sequence specificity as does the yeast ORC, with the non-specific initiation in the egg system being attributed to other factors, such as the high concentration of ORC in the egg.

My initial idea was to isolate the DNA sequences to which ORC is actually bound to in frog somatic cells. This was because unlike in eggs, in somatic cells of higher eukaryotes initiation of DNA replication is known to be localised to certain DNA sequences (see Introduction). The first problem was the isolation of ORC-DNA complexes from the cells, as the chromatin is inherently insoluble and requires high amounts of salt to be solubilised. On the other hand, the ORC-DNA interaction is salt sensitive and is lost in an increased amounts of salt. To overcome this problem, I crosslinked ORC to the DNA *in vivo* to preserve the interaction during the purification procedure. Although initially promising, this approach did not lead to isolation of DNA actually bound to ORC in frog somatic cells due to technical

problems such as extremely low DNA recovery and high level of contaminating cellular DNA.

Following the failure of the experiments studying the ORC-DNA interaction in frog somatic cell, I decided to investigate DNA binding properties of ORC isolated from frog eggs. The great advantage of this approach is the relative ease of obtaining high amounts of purified ORC from egg extract, in comparison to very low amounts recovered from somatic cell. I tested the binding of ORC to plasmid DNA added to the *Xenopus* egg extract, as well as *in vitro* binding of DNA fragments to ORC purified by immunoaffinity on protein A beads. However, I detected no sequence specific DNA recognition by ORC in these assays, despite a number of attempts to isolate DNA fragments which would bind *Xenopus* ORC with higher affinity.

Perhaps not surprisingly, I found that the detected ORC-DNA interaction was stimulated by the addition of ATP, as it is the case for the *S. cerevisiae* ORC.

Finally, ORC immunoprecipitated from frog somatic cells did not show sequence preference on the tested DNA in the *in vitro* binding assay either.

Monoclonal antibody TK15 does not inhibit DNA replication

The ORC-DNA binding studies described in the following experiments were done using ORC purified by immunoprecipitation and bound to the beads via an anti-Orc1p monoclonal antibody. It was therefore very important to determine whether the antibody used, interfered with ORC-DNA interaction. The monoclonal antibody most frequently used for these experiments was TK15, because of its ability to immunoprecipitate a pure preparation of ORC (see Chapter 3, Figure 14). To test whether Mab TK15 interfered with ORC function, the antibody was first purified on

protein A beads. The purified antibody was tested for interference with ORC function in the replication assay. Briefly, 100 μl of the interphase egg extract was incubated with either 5 μl of PBS (control) or with 5 μl of the purified sample of the monoclonal antibody TK15 in PBS. 20 μl aliquots of untreated interphase extract, PBS treated extract and antibody treated extract were tested in the replication assay (see Methods) in both the presence and the absence of chromatin. Figure 8/1A shows that the amount of radioactive phosphate incorporated into chromatin was not reduced in samples where anti-Orc1p antibody was added. In fact, there appeared to be a slight increase in the replication efficiency after the addition of the antibody. To determine what proportion of Orc1p molecules was actually bound to the Mab TK15 in this experiment, 50 μl aliquots of both control (PBS treated) and antibody treated extract were incubated with protein A beads, and after discarding the beads, proteins that remained in the supernatants were analysed by SDS-PAGE. Figure 8/1B shows that addition of protein A beads depleted most of the Orc1p from Mab TK15 treated extract and not from the control extract. The majority of Orc1p in the experiment was therefore bound to Mab TK15. Interestingly, at least half of Orc2p remained in the extract. I concluded that Mab TK15 did not interfere with the DNA-ORC interaction as antibody bound ORC could function in the replication assay. However, it is difficult to exclude the possibility that the fraction of Orc1p, which remained unbound to Mab TK15 in this experiment, could be responsible for the observed DNA replication.

Attempt to isolate the DNA sequences bound to somatic ORC

To isolate DNA sequences somatic ORC is actually bound to, a frog cell line (WAK) was used. However, the prepared chromatin samples were insoluble despite extensive DNase treatment until a sufficient amount of salt was added. This is

Mab TK15 does not inhibit DNA replication

A   Replication assay

chromatin   -   +   -   +   -   +

extract   extract   extract
           + PBS   + Mab TK15
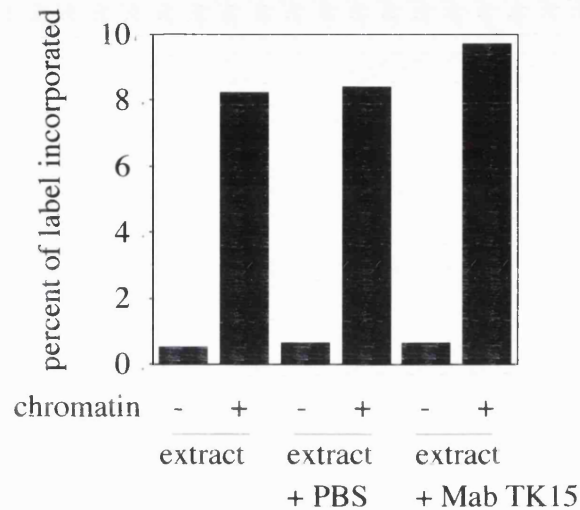
B   Added antibody bound to majority of Orc1p in the extract

◄— Orc1p

◄— Orc2p
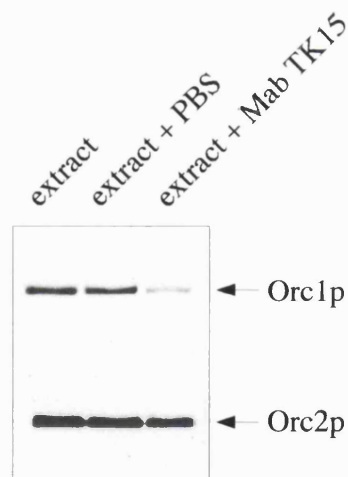
Figure 8/1. (A) Aliquotes of an intephase extract, extract incubated with PBS only and extract incubated with purified monoclonal antibody TK15 were used in replication assay in either the presence or the absence of chromatin. The graph shows the amount of rabiolabel incorporated into DNA as a proportion of total radiolabel added. (B) To determine the propotrion of Orc1p bound to the Mab TK15 in the replication experiment, aliquotes of an intephase extract, extract incubated with PBS only and extract treated with monoclonal antibody TK15 were incubated with protein A beads. After the removal of the beads, proteins which remained in the supernatant were analysed by SDS-PAGE.

illustrated in an experiment described in Figure 8/2 where cells were lysed in a buffer containing increasing amount of salt both in the presence or in the absence of DNase. When the amount of KCl in the sample is relatively low (e.g. 70 mM), ORC proteins are found in the pellet fraction, both in the absence and in presence of DNase I. Figure 8/2 B shows that the DNase digested DNA is found in the pellet fraction under these conditions. When however, the KCl concentration was increased into 100 mM, a fraction of the ORC proteins was found in the supernatant, even in the absence of DNase treatment, presumably due to the ORC complex being stripped off the DNA by the increased salt concentration. It is therefore very hard to find salt concentration conditions under which ORC-DNA interactions would be preserved and the complexes would be soluble.

To overcome these problems, ORC proteins were cross-linked to their substrate DNA by *in vivo* cross-linking with formaldehyde (Orlando and Paro, 1993; Tanaka *et al.*, 1997; see Methods). Briefly, WAK cells were washed with ice-cold PBS and *in vivo* cross-linked in the presence of 1% formaldehyde at 4°C for 15 min. The free cross-linker was then neutralised by 125 mM glycine and cells were washed, lysed in the presence of 1% TritonX-100 and 500 mM NaCl and gently sonicated to shear the DNA into fragments between 200-1000 bp. These samples were then used for immunoprecipitation with anti-Orc1p monoclonal antibody TK15. Figure 8/3A shows that anti-Orc1p antibody co-immunoprecipitated Orc1p, Orc2p and Orc4p from these samples, while tubulin did not co-precipitate, indicating that the cross-linker did not link proteins together non-specifically. In the sample, where the cross-linker was omitted, only Orc1p immunoprecipitated with anti-Orc1p antibody, as 500 mM NaCl completely disrupts the ORC complex.

Salt levels required to solubilise chromatin strip ORC from DNA

A   Titration of an amount of KCl required to solubilise ORC-DNA complexes
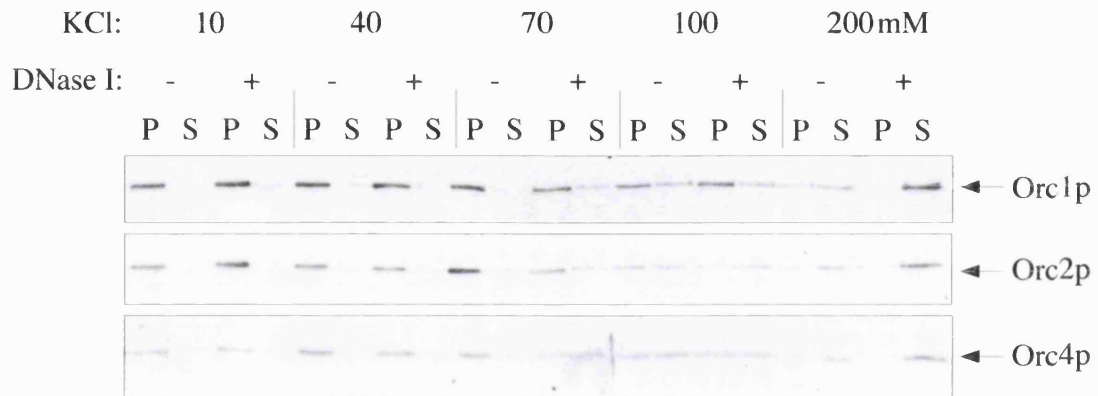
B   DNA control
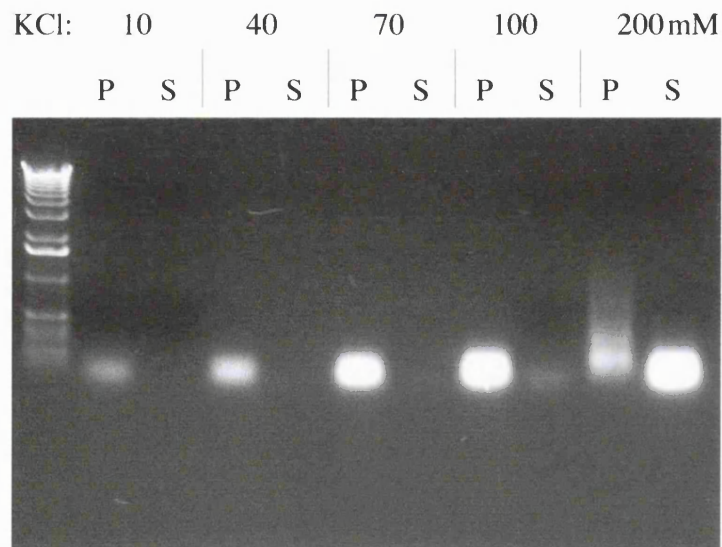
Figure 8/2. (A) Frog somatic cells (WAK) were lysed in buffer containing increasing amounts of KCl in the presence or the absence of DNase I. Following 15 min incubation on ice, the pellet and supernatant were separated by centrifugation through a sucrose cushion, analysed by SDS-PAGE and immunoblotted with anti-Orc1p, anti-Orc2p and anti-Orc4p antibodies. (B) DNA was extracted from the DNase + samples to determine, whether it was soluble or insoluble.

156

Crosslinking ORC to the DNA

A Immunoprecipitation of ORC in high salt conditions in the absence or the presence of cross-linking
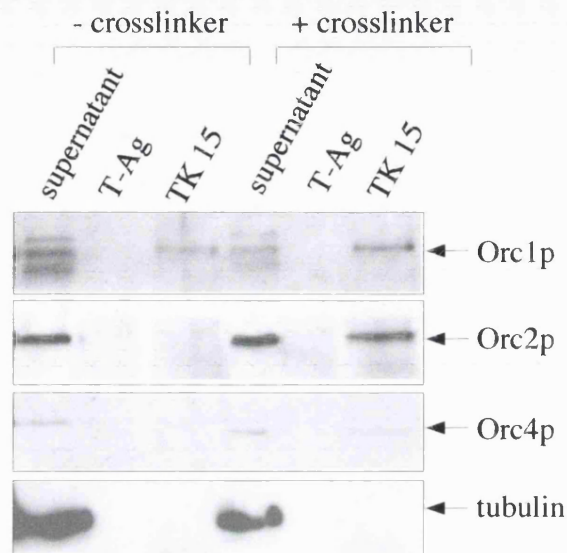
B PCR amplification of the immunoprecipitated DNA
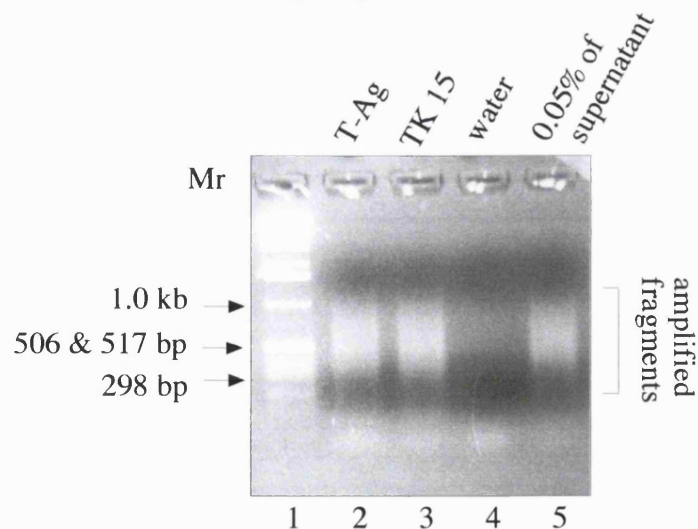
Figure 8/3. (A) Frog somatic cells (WAK) were split into two samples and one of them was *in vivo* crosslinked by formaldehyde. DNA in both samples was sheared by sonication into approximately 200-1000 bp fragments and an immunoprecipitation was performed with either anti-Orc1p antibody TK15 or with control anti-Tag antibody. The immunoprecipitated proteins were analysed by SDS–PAGE and immunoblotted with anti-Orc1p, anti-Orc2p anti-Orc4p and anti-tubulin antibodies. (B) DNA was extracted from the formaldehyde cross-linked samples, ligated to linkers and amplified by PCR. Gels shows PCR products of samples using anti-SV40 T-antigen immunoprecipitate (lane 2), anti-Orc1p immunoprecipitate (lane 3), water (lane 4) and 0.05% of DNA recovered from the supernatant sample (lane 5) as a substrate.

The *in vivo* cross-linking looked very promising at this stage as it specifically cross-linked Orc1p with Orc2p and Orc4p and not with tubulin. The DNA fragments to which ORC was actually bound to should have been therefore also part of this complex. The DNA from the cross-linked immunoprecipitates was recovered by incubating samples in 1% SDS overnight at 65°C, followed by proteinase K treatment and phenol-chloroform extraction. The amount of DNA recovered by this procedure was extremely low and was not detectable by ethidium bromide staining. To amplify the extracted DNA, the samples were treated with Mung bean nuclease to create blunt ends, ligated to linkers and amplified by PCR (see Methods). This process was not very efficient and required a lot of optimisation to finally produce a PCR product. However, as shown in Figure 8/3 , the same amount of PCR product was obtained from both the anti-Orc1p immunoprecipitates as from the control anti-T antigen immunoprecipitates. The origin of this DNA was presumably contaminating cellular DNA which was carried through the immunoprecipitation procedure. The control sample, where water was used as a substrate, was treated in the same way as the DNA extracted from the immunoprecipitates. This sample did not produce any PCR product and therefore the product seen in the immunoprecipitation samples did not originate from any step following the immunoprecipitation (e.g. nuclease treatment, linker ligation). In the positive control sample 0.05% of the equivalent of the supernatant sample used for each immunoprecipitation was used. This sample gave similar amount of PCR product as the immunoprecipitate samples indicating that 0.05% of contaminating cellular DNA failed to be removed from the immunoprecipitates.

Further attempts to remove the contaminating DNA or to improve the background to signal ration were unsuccessful with both the positive (anti-Orc1p immunoprecipitate) and the negative (anti-SV40 T-antigen immunoprecipitate)

signals always disappeared at the same point. In conclusion, the cross-linking experiments, although initially very promising, failed to select the DNA fragments ORC is actually bound to due to low level of a specific signal and a high background level.

ORC binds to plasmid DNA added to the egg extract

An alternative approach to studying the ORC-DNA interaction is to investigate the binding of ORC found in the frog egg extract to externally added DNA. It has been demonstrated by several laboratories that any DNA added to the *Xenopus* interphase egg extract can replicate (see Introduction). To test whether ORC bound to plasmid DNA added to the extract, pUC19-ARS1 plasmid (a pUC based plasmid containing the *S. cerevisiae* ARS1 sequence; see Methods) was added to this extract and following 15 min incubation at 23°C, immunoprecipitation was performed with either the control anti-SV40 T-antigen antibody or with the anti-Orc1p antibody TK15. Figure 8/4 shows that the control antibody did not immunoprecipitate any DNA from the extract and did not bring down any Orc1p and Orc2p proteins. In contrast, the anti-Orc1p monoclonal antibody TK15 immunoprecipitated a small fraction of the added plasmid as well as both Orc1p and Orc2p proteins, suggesting that at least a fraction of the plasmid added to the frog egg extract was recognised and bound by ORC. The additional high molecular weight DNA immunoprecipitated from the extract by the anti-Orc1p antibody is presumably frog mitochondrial DNA.

The fact that the majority of plasmid was not immunoprecipitated in this experiment is not due to lack of ORC in the binding reaction as approximately 18 pmol (in 500 µl of interphase extract) of ORC and 6 pmol of plasmid were present in the binding reaction. It is therefore not clear, whether only small proportion of ORC bound to the plasmid DNA, or whether the majority of the binding was lost during the extraction.

ORC binds to plasmid added to the interphase extract
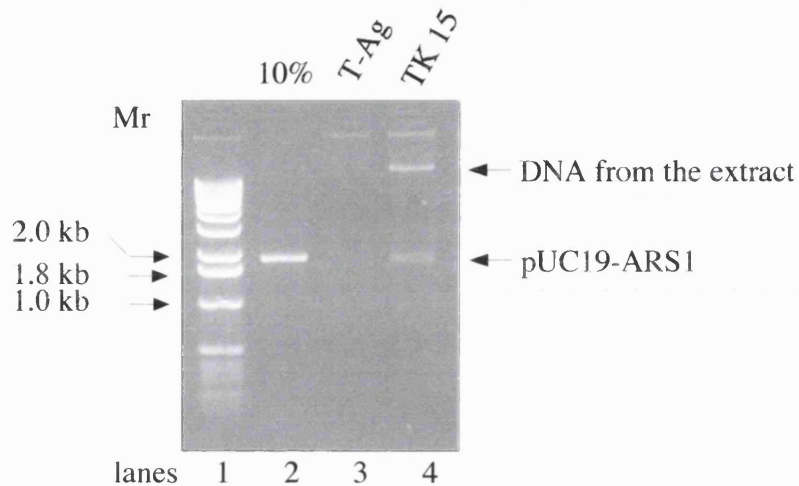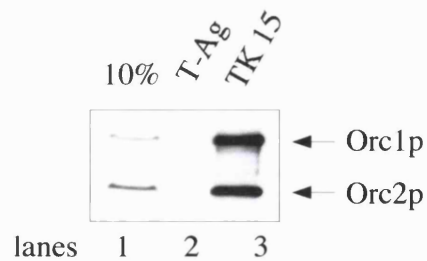
Figure 8/4. (A) Plasmid pUC19-ARS was incubated in the interphase egg extract for 15 min at 23°C. This was followed by immunoprecipitation with either anti-Orc1p monoclonal antibody TK15 (lane 4) or with a control anti-SV40 T-antigen antibody (lane 3). Lane 2; 10% of the imput DNA, lane 1; markers. The DNA was extracted from the immunoprecipitated and resolved on an agarose gel. (B) Immunoblot of Orc1p and Orc2p extracted from the same samples. Lane 1; 10% of the imput protein.

Embryonic ORC binds DNA in ATP-dependent manner

Having established that ORC binds to pUC19-ARS1 DNA added to frog egg extract, an interesting question was whether there is any sequence in the pUC19-ARS preferentially recognised by ORC. An *in vitro* assay, measuring ORC-DNA interaction, based on DNA binding to ORC which was immunoprecipitated and bound to the beads was therefore developed. The pUC19-ARS DNA was digested with several restriction enzymes and end labeled with polynucleotide kinase to give the various DNA fragments and equal intensity radio-label. ORC was immunoprecipitated from interphase egg extract with monoclonal antibody TK15 and incubated (while still bound to beads) with the DNA fragments generated from the pUC19 plasmid either in the presence or in the absence of ATP. Increasing amounts of poly-dI-poly-dC competitor were added to the binding reactions. Figure 8/5 shows that although DNA fragments bound to ORC both in the presence and the absence of ATP, the binding in the presence of ATP was more efficient. This result would be expected considering that the DNA binding of *S. cerevisiae* ORC is ATP dependent. The residual binding of DNA to ORC in the absence of ATP might be due to some ATP being carried through from the egg extract.

Binding of ORC to the substrate DNA is easily competed out by the poly-dI-poly-dC competitor however, suggesting that the interaction is not of very high affinity, possibly not very specific. No single fragment shows a preferential binding to ORC (including the yeast ORC recognition site). It therefore seems obvious, that if there is a DNA sequence which interacts with ORC purified from *Xenopus* eggs at a higher frequency, such sequence is probably not present in pUC19 plasmid, nor in the *S. cerevisiae* origin sequence.

Binding of digested pUC19-ARS1 plasmid to ORC IP
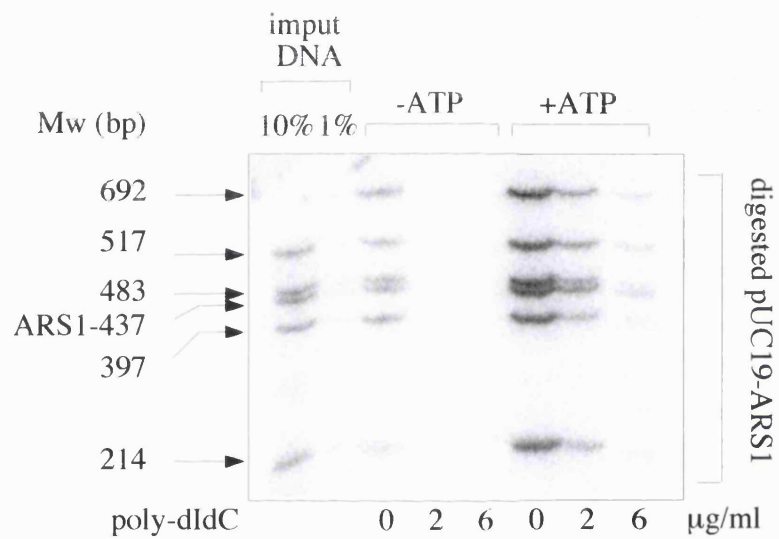
Figure 8/5. Plasmid pUC19-ARS1 was digested with restriction enzymes to produce a number of DNA fragments. The 437 bp fragment is containing the sequence of ARS1. The obtained restriction fragments were end labeled with ^{32}P and used in binding to ORC immunoprecipitated with anti-Orc1p monoclonal antibody TK15. The DNA binding was performed either the presence or the absence of 1 mM ATP. An increasing amount of the poly dIdC was used in these reactions.

Embryonic ORC interacts with any DNA sequence tested

The *in vitro* ORC-DNA interaction assay was very reproducible and I therefore decided to use it for selecting DNA sequences which would bind to *Xenopus* ORC with higher affinity. The idea was to search for such DNA sequences in the frog genome. In the first round of experiments, I generated a frog genomic library in the pUC19 plasmid. I then performed the *in vitro* ORC-DNA binding assay using the library plasmids as substrate. I looked for conditions under which I could recover plasmids bound to ORC in the presence, but not in the absence of ATP, assuming that the ATP independent binding might be non-specific. I recovered a number of plasmids in this way, but when I re-tested these plasmids themselves, or the insert cloned within them in a competition assay, their ability to bind to ORC was exactly the same as that of the pUC19-ARS.

Next, I thought that perhaps the plasmid itself was interfering with the binding assay, as it has the ability to bind to ORC. I was also concerned that my genomic library was perhaps not large enough. I therefore partially digested the total frog genomic DNA with *Sau* IIIA and used these fragments for the selection of sequences binding to ORC with higher affinity. DNA fragments recovered as those which bind to ORC were then subcloned into a *Bam* HI site of Bluescript vector, amplified and re-isolated. The 1.2 kb DNA fragment identified in this experiment was then re-tested for its ability to bind ORC. Figure 8/6, however, shows that this fragment bound to ORC with the same affinity as the previously tested ARS1 fragment.

In conclusion, all of the screens attempted isolated DNA fragments which bound to ORC with exactly same affinity as any other fragments. While it is quite possible that the selection systems used were not sufficiently stringent, or not exhaustive

Competition assay
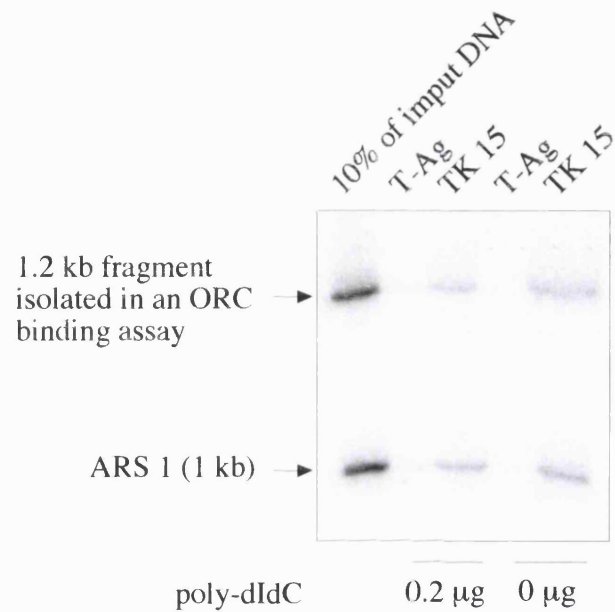
Figure 8/6. The 1.2 kb DNA fragment was isolated from the total frog genomic digest as DNA fragment interacting with ORC. The 1.2 kb DNA fragment and the 1 kb fragment containing the ARS1 sequence were end labeled with ^{32}P and used in binding to either control immunoprecipitate (using anti-SV40 T-antigen monoclonal antibody) or to immunoprecipitated ORC (using anti-Orc1p monoclonal antibody TK15). The DNA binding was performed in the presence of 1 mM ATP and in either the absence or in the presence of 0.2 μg of poly dIdC per binding reaction.

enough, it seems likely that the embryonic ORC used indeed does not have any strong preferences for a particular DNA sequence.

In vitro interaction of somatic ORC with DNA

In the lack of apparent sequence specificity of ORC isolated from *Xenopus* egg extract, the obvious question whether this would be the same for ORC isolated from frog somatic cells arises. Although it is much harder to obtain sufficient amount of ORC for ORC-DNA *in vitro* binding assay, I decided to try and see whether I could detect any sequence preferences in binding of the somatic ORC to DNA.

ORC was immunoprecipitated from WAK cell lysate in using either anti-Orc1p monoclonal antibody TK15 or a control anti-SV40 T-antigen antibody. Plasmid pUC19-ARS1 was digested and the obtained fragments were end labelled and used for binding to the ORC immunoprecipitated for the frog somatic cells. Figure 8/7 shows a weak binding of the DNA fragment to the ORC immunoprecipitate and not to the control immunoprecipitate. However, all fragments bound ORC with the same affinity, suggesting that somatic ORC does not bind preferentially any sequences in pUC 19 nor in the *S. cerevisiae* ARS1 sequence.

Binding of digested pUC19-ARS1 plasmid to somatic ORC IP
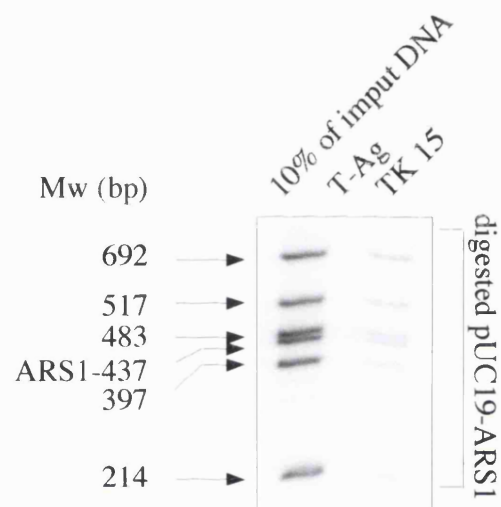
Figure 8/7. Plasmid pUC19-ARS1 was digested with restriction enzymes to produce a number of DNA fragments. The 437 bp fragment is containing the sequence of ARS1. The obtained restriction fragments were end labeled with ^{32}P and used in binding to either control immunoprecipitate (using anti-SV40-antigen monoclonal antibody) or to ORC immunoprecipitated from frog somatic cells (WAK) (using anti-Orc1p monoclonal antibody TK15). The DNA binding was performed in the presence of 1 mM ATP. No poly dIdC was used in these binding reactions.

Chapter 9

Discussion

The overriding aim of the work described in this thesis was to understand the specification of the DNA replication origin sites in the frog chromosomes. There was an obvious discrepancy between the finding that essentially any DNA sequence can be replicated in the frog egg extract and the requirement for ORC, which is a functional homologue of the yeast protein complex that recognises origins of replication in yeast. I therefore set off to investigate how *Xenopus* ORC recognises its target sequences.

To be able to study the DNA binding properties of ORC, I isolated ORC from *Xenopus* egg extract, which is a rich source of this protein complex present in a free (DNA unbound) form. The approach I chose was based on immunoaffinity purification, taking advantage of the Orc1p subunit of the complex, which was the only subunit of ORC known in higher eukaryotes at the time. Immunoaffinity purification proved to be successful and allowed a simple way to isolate the complex. This success was largely due to high affinity and specificity of anti-Orc1p monoclonal antibodies, which I prepared with this purpose in mind.

Microsequencing of the purified complex led to the isolation of Orc4p from *Xenopus*, helped to identify *Xenopus* Orc3p, and was used for isolation of human Orc5p. Curiously, Orc6p was not identified from my purified ORC sample, and it remains to

be determined whether this was due to technical problems, or whether ORC in *Xenopus* eggs lacks an Orc6p subunit. If the latter proves to be the case, the possibility of lack of Orc6p being responsible for non-specific ORC-DNA interactions in the frog egg extract will have to be considered.

Isolation of Orc4 and Orc5 cDNAs from frog and human led to the identification of previously unexpected homologies between the two proteins encoded by these two genes and the previously identified human and frog Orc1p and Cdc6p proteins. These homologies imply descent from a common ancestral protein, or a common functional feature, such as oligomerisation, or both. Clearly, much remains to be learned about the functioning of ORC on the molecular level.

Some of the anti-Orc1p monoclonal antibodies used in these studies bound a number of additional proteins besides the identified ORC subunits. These may represent interesting ORC interacting proteins, although the fact that Cdc6p protein was not found among them makes this possibility less likely. The difficulty associated with large-scale protein microsequencing limited any further investigation of these polypeptides.

In the course of this work, I noticed that *Xenopus* Orc1p becomes phosphorylated when the oocytes enter M phase. This modification may have important regulatory function in the regulation of the pre-replicative complex assembly, for example. However, this theory is not easily tested in the frog system and remains to be investigated in the future.

To identify differences between embryonic and somatic ORC, which might explain the presumed differential behaviour of ORC in eggs and somatic cells as regards the

recognition of specific origin-binding sites, I compared ORC obtained from the two sources. I found no differences in immunoreactivity and the molecular weight between Orc1p, Orc2p and Orc4p from eggs and somatic cells. In support for this finding is the fact that in *Drosophila* all ORC subunits were found to be encoded by single genes (Chesnokov *et al.*, 1999).

The only difference I found between the embryonic and somatic ORC, besides the possible lack of Orc6p, was the ratio between ORC and the number of replication origins. I estimated that there are approximately 350 000 ORC molecules per replication origin present in the frog egg during the first S phase while in the somatic cell there are only 3.5 ORCs per origin. This finding offers an alternative explanation for the observed difference in the initiation specificity between eggs and somatic cells.

I put significant efforts into attempts to isolate the DNA sequences to which ORC is actually bound in *Xenopus* somatic cells (WAK). The rationale behind these experiments was that although replication initiates at apparently random sites in *Xenopus* eggs, it becomes localised to specific sites later on in the development. Moreover, the study of human, *Drosophila* and Chinese hamster somatic cells provided further support for localised initiation in higher eukaryotes. However, I encountered significant technical difficulties in this approach and I was unable to isolate ORC bound DNA from somatic cells.

To learn more about ORC-DNA interaction I decided to study the DNA binding properties of embryonic ORC. I showed that ORC can bind plasmid DNA added to the egg extract, which is in an agreement with the previous reports which demonstrated that plasmids from a variety of sources can replicate in the egg extract

system and that replication in this system is ORC dependent (Harland and Laskey, 1980; Laskey and Gurdon, 1973; Rowles *et al.*, 1996). In an *in vitro* binding assay, any DNA tested bound ORC purified from *Xenopus* eggs with the same kinetics and I could find no evidence for the existence of DNA fragments that bound to the embryonic ORC with higher affinity. These results suggest that *Xenopus* embryonic ORC is indeed able to bind any DNA sequence, explaining why any DNA sequence can replicate in the egg extract system. The binding of embryonic ORC to DNA was not very strong however, as it was easily competed with a non-specific competitor (poly-dI-dC).

Interestingly, the observed ORC-DNA interaction was stimulated by the presence of ATP. This finding is in slight disagreement with the results from the *S. cerevisiae* and the *Drosophila* systems, which claim that only sequence specific origin recognition by ORC is ATP dependent, and that background levels of non-specific ORC-DNA binding is ATP independent (Austin *et al.*, 1999; Klemm *et al.*, 1997).

In the light of recently published data studying the interaction of ORC with the ACE element of the *Drosophila* chorion cluster, it appears that somatic ORC of higher eukaryotes binds to specific sequences within replication origins (Austin *et al.*, 1999). To further our understanding of the identity of *Xenopus* replication origins, DNA fragments which *Xenopus* ORC binds in somatic cells will have to be identified. However, the questions of the nature of the random initiation in the *Xenopus* eggs and the mechanism that sets the origin spacing in this system remain still far from understood.

References

Adamczewski, J., Gannon, J. and Hunt, T. (1993) Simian virus 40 large T-antigen associates with cyclin A and p33cdk2. *J Virol,* **67,** 6551-6557.

Aladjem, M., Groudine, M., Brody, L., Dieken, E., Fournier, R., Wahl, G. and Epner, E. (1995) Participation of the human beta-globin locus control region in initiation of DNA replication. *Science,* **270,** 815-819.

Aladjem, M., Rodewald, L., Kolman, J. and Wahl, G. (1998) Genetic dissection of a mammalian replicator in the human beta-globin locus. *Science,* **281,** 1005-1009.

Alfano, C. and McMacken, R. (1989a) Heat shock protein-mediated disassembly of nucleoprotein structures is required for the initiation of bacteriophage lambda DNA replication. *J Biol Chem,* **264,** 10709-10718.

Alfano, C. and McMacken, R. (1989b) Ordered assembly of nucleoprotein structures at the bacteriophage lambda replication origin during the initiation of DNA replication. *J Biol Chem,* **264,** 10699-10708.

Aparicio, O., Weinstein, D. and Bell, S. (1997) Components and dynamics of DNA replication complexes in *S. cerevisiae*: redistribution of MCM proteins and Cdc45p during S phase. *Cell,* **91,** 59-69.

Austin, R., Orr-Weaver, T. and Bell, S. (1999) *Drosophila* ORC specifically binds to ACE3, an origin of DNA replication control element. *Genes Dev,* **13,** 2639-2649.

Beach, D., Piper, M. and Shall, S. (1980) Isolation of chromosomal origins of replication in yeast. *Nature,* **284,** 185-187.

Bell, S., Kobayashi, R. and Stillman, B. (1993) Yeast origin recognition complex functions in transcription silencing and DNA replication. *Science,* **262,** 1844-1849.

Bell, S., Mitchell, J., Leber, J., Kobayashi, R. and Stillman, B. (1995) The multidomain structure of Orc1p reveals similarity to regulators of DNA replication and transcriptional silencing. *Cell,* **83,** 563-568.

Bell, S.P. and Stillman, B. (1992) Nucleotide dependent recognition of chromosomal origins of DNA replication by a multi-protein complex. *Nature,* **357,** 128-134.

Biamonti, G., Della Valle, G., Talarico, D., Cobianchi, F., Riva, S. and Falaschi, A. (1985) Fate of exogenous recombinant plasmids introduced into mouse and human cells. *Nucleic Acids Res.,* **13,** 5545-5561.

Blow, J. (1993) Preventing re-replication of DNA in a single cell cycle: evidence for a replication licensing factor. *J Cell Biol,* **122,** 993-1002.

Blow, J. and Laskey, R. (1988) A role for the nuclear envelope in controlling DNA replication within the cell cycle. *Nature*, **332**, 546-548.

Blow, J. and Nurse, P. (1990) A cdc2-like protein is involved in the initiation of DNA replication in Xenopus egg extracts. *Cell*, **62**, 855-862.

Blow, J.J. and Laskey, R.A. (1986) Initiation of DNA replication in nuclei and purified DNA by a cell-free extract of *Xenopus* eggs. *Cell*, **47**, 577-587.

Borowiec, J., Dean, F., Bullock, P. and Hurwitz, J. (1990) Binding and unwinding--how T antigen engages the SV40 origin of DNA replication. *Cell*, **60**, 181-184.

Bousset, K. and Diffley, J. (1998) The Cdc7 protein kinase is required for origin firing during S phase. *Genes Dev*, **12**, 480-490.

Bozzoni, I., Baldari, C., Amaldi, F. and Buongiorno-Nardelli, M. (1981) Replication of ribosomal DNA in *Xenopus laevis. Eur J Biochem*, **118**, 585-590.

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.

Brewer, B. and Fangman, W. (1987) The localization of replication origins on ARS plasmids in *S. cerevisiae. Cell*, **51**, 463-471.

Broach, J.R., Li, Y.-Y., Feldman, J., Jayaram, M., Abraham, J., Nasmyth, K.A. and Hicks, J.B. (1983) Localization and sequence analysis of yeast origins of DNA replication. *Cold Spring Harbour Symp. Quant. Biol.*, **47**, 1165-1173.

Broek, D., Bartlett, R., Crawford, K. and Nurse, P. (1991) Involvement of p34cdc2 in establishing the dependency of S phase on mitosis. *Nature*, **349**, 388-393.

Brown, G. and Kelly, T. (1998) Purification of Hsk1, a minichromosome maintenance protein kinase from fission yeast. *J Biol Chem*, **273**, 22083-22090.

Burhans, W., Vassilev, L., Caddle, M., Heintz, N. and DePamphilis, M. (1990) Identification of an origin of bidirectional DNA replication in mammalian chromosomes. *Cell*, **62**, 955-965.

Caddle, M. and Calos, M. (1992) Analysis of the autonomous replication behavior in human cells of the dihydrofolate reductase putative chromosomal origin of replication. *Nucleic Acids Res*, **20**, 5971-5978.

Carpenter, P.B. and Dunphy, W.G. (1998) Identification of a novel 81-kDa component of the *Xenopus* Origin Recognition Complex. *Journal of Biological Chemistry*, **273**, 24891-24897.

Carpenter, P.B., Muller, P.R. and Dunphy, W.G. (1996) Role for a *Xenopus* ORC 2-related protein in controlling DNA replication. *Nature*, **379**, 357-360.

Celniker, S.E., Sweder, K., Srienc, F., Bailey, J.E. and Campbell, J.L. (1984) Deletion mutations affecting autonomously replicating sequence ARS1 of *Saccharomyces cerevisiae. Mol. Cell. Biol.,* 2455-2466.

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

Chapman, J. and Johnston, L. (1989) The yeast gene, DBF4, essential for entry into S phase is cell cycle regulated. *Exp Cell Res,* **180,** 419-428.

Chen, J., Jackson, P., Kirschner, M. and Dutta, A. (1995) Separate domains of p21 involved in the inhibition of Cdk kinase and PCNA. *Nature,* **374,** 386-388.

Chesnokov, I., Gossen, M., Remus, D. and Botchan, M. (1999) Assembly of functionally active Drosophila origin recognition complex from recombinant proteins. *Genes Dev,* **13,** 1289-1296.

Chevalier, S., Couturier, A., Chartrain, I., Le Guellec, R., Beckhelling, C., Le Guellec, K., Philippe, M. and Ford, C. (1996) *Xenopus* cyclin E, a nuclear phosphoprotein, accumulates when oocytes gain the ability to initiate DNA replication. *J Cell Sci,* **109,** 1173-1184.

Chevalier, S., Tassan, J., Cox, R., Philippe, M. and Ford, C. (1995) Both cdc2 and cdk2 promote S phase initiation in Xenopus egg extracts. *J Cell Sci,* **108,** 1831-1841.

Chong, J., Mahbubani, H., Khoo, C. and Blow, J. (1995) Purification of an MCM-containing complex as a component of the DNA replication licensing system. *Nature,* **375,** 418-421.

Chong, J., Thommes, P. and Blow, J. (1996) The role of MCM/P1 proteins in the licensing of DNA replication. *Trends Biochem Sci,* **21,** 102-106.

Chuang, R. and Kelly, T. (1999) The fission yeast homologue of Orc4p binds to replication origin DNA via multiple AT-hooks. *Proc Natl Acad Sci U S A,* **96,** 2656-2661.

Clyne, R. and Kelly, T. (1995) Genetic analysis of an ARS element from the fission yeast Schizosaccharomyces pombe. *EMBO J,* **14,** 6348-6357.

Cocker, J., Piatti, S., Santocanale, C., Nasmyth, K. and Diffley, J. (1996) An essential role for the Cdc6 protein in forming the pre-replicative complexes of budding yeast. *Nature,* **379,** 180-182.

Coleman, T., Carpenter, P. and Dunphy, W. (1996) The *Xenopus* Cdc6 protein is essential for the initiation of a single round of DNA replication in cell-free extracts. *Cell,* **4,** 53-63.

Correa-Bordes, J. and Nurse, P. (1995) p25rum1 orders S phase and mitosis by acting as an inhibitor of the p34cdc2 mitotic kinase. *Cell,* **83,** 1001-1009.

Dahmann, C., Diffley, J. and Nasmyth, K. (1995) S-phase-promoting cyclin-dependent kinases prevent re-replication by inhibiting the transition of replication origins to a pre-replicative state. *Curr Biol,* **5,** 1257-1269.

Dalton, S. and Hopwood, B. (1997) Characterization of Cdc47p-minichromosome maintenance complexes in Saccharomyces cerevisiae: identification of Cdc45p as a subunit. *Mol Cell Biol*, **17**, 5867-5875.

Dalton, S. and Whitbread, L. (1995) Cell cycle-regulated nuclear import and export of Cdc47, a protein essential for initiation of DNA replication in budding yeast. *Proc Natl Acad Sci U S A*, **92**, 2514-2518.

de Cicco, D. and Spradling, A. (1984) Localization of a cis-acting element responsible for the developmentally regulated amplification of Drosophila chorion genes. *Cell*, **38**, 45-54.

Dean, F., Borowiec, J., Eki, T. and Hurwitz, J. (1992) The simian virus 40 T-antigen double hexamer assembles around the DNA at the replication origin. *J Biol Chem*, **267**, 14129-14137.

Dean, F., Dodson, M., Echols, H. and Hurwitz, J. (1987) ATP-dependent formation of a specialized nucleoprotein structure by simian virus 40 (SV40) large tumor antigen at the SV40 replication origin. *Proc Natl Acad Sci U S A*, **84**, 8981-8985.

Delidakis, C. and Kafatos, F. (1987) Amplification of a chorion gene cluster in Drosophila is subject to multiple cis-regulatory elements and to long-range position effects. *J Mol Biol*, **197**, 11-26.

Denniston-Thompson, K., Moore, D., Kruger, K., Furth, M. and Blattner, F. (1977) Physical structure of the replication origin of bacteriophage lambda. *Science*, **198**, 1051-1056.

Detweiler, C. and Li, J. (1998) Ectopic induction of Clb2 in early G1 phase is sufficient to block prereplicative complex formation in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A*, **95**, 2384-2389.

Diffley, J. and Cocker, J. (1992) Protein-DNA interactions at a yeast replication origin. *Nature*, **357**, 169-172.

Diffley, J., Cocker, J., Dowell, S. and Rowley, A. (1994) Two steps in the assembly of complexes at yeast replication origins *in vivo*. *Cell*, **78**, 303-316.

Diffley, J. and Stillman, B. (1988) Purification of a yeast protein that binds to origins of DNA replication and a transcriptional silencer. *Proc Natl Acad Sci U S A*, **85**, 2120-2124.

Dijkwel, P. and Hamlin, J. (1992) Initiation of DNA replication in the dihydrofolate reductase locus is confined to the early S period in CHO cells synchronized with the plant amino acid mimosine. *Mol Cell Biol*, **12**, 3715-3722.

Dijkwel, P., Vaughn, J. and Hamlin, J. (1994) Replication initiation sites are distributed widely in the amplified CHO dihydrofolate reductase domain. *Nucleic Acids Res*, **22**, 4989-4996.

Dodson, M., Echols, H., Wickner, S., Alfano, C., Mensa-Wilmot, K., Gomes, B., LeBowitz, J., Roberts, J. and McMacken, R. (1986) Specialized nucleoprotein structures at the origin of replication of bacteriophage lambda: localized unwinding of duplex DNA by a six-protein reaction. *Proc Natl Acad Sci U S A*, **83**, 7638-7642.

Dodson, M., McMacken, R. and Echols, H. (1989) Specialized nucleoprotein structures at the origin of replication of bacteriophage lambda. Protein association and disassociation reactions responsible for localized initiation of replication. *J Biol Chem*, **264**, 10719-10725.

Dodson, M., Roberts, J., McMacken, R. and Echols, H. (1985) Specialized nucleoprotein structures at the origin of replication of bacteriophage lambda: complexes with lambda O protein and with lambda O, lambda P, and Escherichia coli DnaB proteins. *Proc Natl Acad Sci U S A*, **82**, 4678-4682.

Donaldson, A., Fangman, W. and Brewer, B. (1998) Cdc7 is required throughout the yeast S phase to activate replication origins. *Genes Dev*, **12**, 491-501.

Donovan, S., Harwood, J., Drury, L. and Diffley, J. (1997) Cdc6p-dependent loading of Mcm proteins onto pre-replicative chromatin in budding yeast. *Proc Natl Acad Sci U S A*, **94**, 5611-5616.

Dowell, S., Romanowski, P. and Diffley, J. (1994) Interaction of Dbf4, the Cdc7 protein kinase regulatory subunit, with yeast replication origins in vivo. *Science*, **265**, 1243-1246.

Driscoll, M., Dobkin, C. and Alter, B. (1989) Gamma delta beta-thalassemia due to a de novo mutation deleting the 5' beta-globin gene activation-region hypersensitive sites. *Proc Natl Acad Sci U S A*, **86**, 7470-7474.

Drury, L., Perkins, G. and Diffley, J. (1997) The Cdc4/34/53 pathway targets Cdc6p for proteolysis in budding yeast. *EMBO J*, **16**, 5966-5976.

Dubey, D., Kim, S., Todorov, I. and Huberman, J. (1996) Large, complex modular structure of a fission yeast DNA replication origin. *Curr Biol*, **6**, 467-473.

Dubey, D., Zhu, J., Carlson, D., Sharma, K. and Huberman, J. (1994) Three ARS elements contribute to the ura4 replication origin region in the fission yeast, Schizosaccharomyces pombe. *EMBO J*, **13**, 3638-3647.

Ehrenhofer-Murray, A., Gossen, M., Pak, D., Botchan, M. and Rine, J. (1995) Separation of origin recognition complex functions by cross-species complementation. *Science*, **270**, 1671-1674.

Epner, E., Forrester, W. and Groudine, M. (1988) Asynchronous DNA replication within the human beta-globin gene locus. *Proc Natl Acad Sci U S A*, **85**, 8081-8085.

Fang, F. and Newport, J. (1991) Evidence that the G1-S and G2-M transitions are controlled by different cdc2 proteins in higher eukaryotes. *Cell*, **66**, 731-742.

Fangman, W. and Brewer, B. (1992) A question of time: replication origins of eukaryotic chromosomes. *Cell*, **71**, 363-366.

Fanning, E. (1994) Control of SV40 DNA replication by protein phosphorylation: a model for cellular DNA replication? *Trends Cell Biol*, **4**, 250-255.

Ferguson, B.M., Brewer, B.J., Reynolds, A.E. and Fangman, W.L. (1991) A yeast origin of replication is activated late in S-phase. *Cell*, **65**, 507-515.

Foreman, P. and Hamlin, J. (1989) Identification and characterization of a gene that is coamplified with dihydrofolate reductase in a methotrexate-resistant CHO cell line. *Mol Cell Biol*, **9**, 1137-1147.

Forrester, W., Epner, E., Driscoll, M., Enver, T., Brice, M., Papayannopoulou, T. and Groudine, M. (1990) A deletion of the human beta-globin locus activation region causes a major alteration in chromatin structure and replication across the entire beta-globin locus. *Genes Dev*, **4**, 1637-1649.

Foss, M., McNally, F., Laurenson, P. and Rine, J. (1993) Origin recognition complex (ORC) in transcriptional silencing and DNA replication in *S. cerevisiae*. *Science*, **262**, 1838-1844.

Fox, C., Loo, S., Dillin, A. and Rina, J. (1995a) The origin recognition complex has essential functions in transcriptional scilencing and chromosomal replication. *Genes Dev*, **9**, 911-924.

Fox, C., Loo, S., Dillin, A. and Rine, J. (1995b) The origin recognition complex has essential functions in transcriptional silencing and chromosomal replication. *Genes Dev*, **9**, 911-924.

Friedman, K., Raghuraman, M., Fangman, W. and Brewer, B. (1995) Analysis of the temporal program of replication initiation in yeast chromosomes. *J Cell Sci Suppl*, **19**.

Fujita, M., Yamada, C., Goto, H., Yokoyama, N., Kuzushima, K., Inagaki, M. and Tsurumi, T. (1999) Cell Cycle Regulation of Human CDC6 Protein. *J Biol Chem*, **274**, 25927-25932.

Fuller, R., Funnell, B. and Kornberg, A. (1984) The dnaA protein complex with the *E. coli* chromosomal replication origin (oriC) and other DNA sites. *Cell*, **38**, 889-900.

Furth, M., Blattner, F., McLeester, C. and Dove, W. (1977) Genetic structure of the replication origin of bacteriophage lambda. *Science*, **198**, 1046-1051.

Gavin, K.A., Hidaka, M. and Stilman, B. (1995) Conserved Initiation Proteins in Eukaryotes. *Science*, **270**, 1667-1671.

Geierstanger, B., Volkman, B., Kremer, W. and Wemmer, D. (1994) Short peptide fragments derived from HMG-I/Y proteins bind specifically to the minor groove of DNA. *Biochemistry*, **33**, 5347-5355.

Gossen, M., Pak, D.T.S., Hansen, S.K., Acharya, J.K. and Botchan, M.R. (1995) A *Drosophila* Homologue of the Yeast Origin Recognition Complex. *Science*, **270**, 1674-1677.

Grallert, B. and Nurse, P. (1996) The ORC1 homolog orp1 in fission yeast plays a key role in regulating onset of S phase. *Genes Dev*, **10**, 2644-2654.

Grosschedl, R. and G, H. (1979) DNA sequences and structural homologies of the replication origins of lambdoid bacteriophages. *Nature*, **277**, 621-627.

Hardy, C. (1997) Identification of Cdc45p, an essential factor required for DNA replication. *Gene*, **187**, 239-246.

Hardy, C., Dryga, O., Seematter, S., Pahl, P. and Sclafani, R. (1997) mcm5/cdc46-bob1 bypasses the requirement for the S phase activator Cdc7p. *Proc Natl Acad Sci U S A*, **94**, 3151-3155.

176

Harland, R.M. and Laskey, R.A. (1980) Regulated replication of DNA microinjected into eggs of *Xenopus laevis. Cell*, **21**, 761-771.

Harlow, E. and Lane, D. (1988) Antibodies. A laboratory manual.

Hartwell, L. (1973) Three additional genes required for deoxyribonucleic acid synthesis in Saccharomyces cerevisiae. *J Bacteriol*, **115**, 966-974.

Hartwell, L. (1976) Sequential function of gene products relative to DNA synthesis in the yeast cell cycle. *J Mol Biol*, **104**, 803-817.

Hartwell, L. (1978) Cell division from a genetic perspective. *J Cell Biol*, **77**, 627-637.

Hatton, K., Dhar, V., Brown, E., Iqbal, M., Stuart, S., Didamo, V. and Schildkraut, C. (1988) Replication program of active and inactive multigene families in mammalian cells. *Mol Cell Biol*, **8**, 2149-2158.

Hayles, J., Fisher, D., Woollard, A. and Nurse, P. (1994) Temporal order of S phase and mitosis in fission yeast is determined by the state of the p34cdc2-mitotic B cyclin complex. *Cell*, **78**, 813-822.

Heck, M. and Spradling, A. (1990) Multiple replication origins are used during Drosophila chorion gene amplification. *J Cell Biol*, **110**, 903-914.

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

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

Hendrickson, M., Madine, M., Dalton, S. and Gautier, J. (1996) Phosphorylation of MCM4 by cdc2 protein kinase inhibits the activity of the minichromosome maintenance complex. *Proc Natl Acad Sci U S A*, **93**, 12223-12228.

Hennessy, K., Clark, C. and Botstein, D. (1990) Subcellular localization of yeast CDC46 varies with the cell cycle. *Genes Dev*, **4**, 2252-2263.

Hennessy, K., Lee, A., Chen, E. and Botstein, D. (1991) A group of interacting yeast DNA replication genes. *Genes Dev*, **5**, 958-969.

Hobom, G., Grosschedl, R., Lusky, M., Scherer, G., Schwarz, E. and Kossel, H. (1979) Functional analysis of the replicator structure of lambdoid bacteriophage DNAs. *Cold Spring Harb Symp Quant Biol*, **43**, 165-178.

Hollingsworth, R.J. and Sclafani, R. (1990) DNA metabolism gene CDC7 from yeast encodes a serine (threonine) protein kinase. *Proc Natl Acad Sci U S A*, **87**, 6272-6276.

Hopwood, B. and Dalton, S. (1996) Cdc45p assembles into a complex with Cdc46p/Mcm5p, is required for minichromosome maintenance, and is essential for chromosomal DNA replication. *Proc Natl Acad Sci U S A*, **93**, 12309-12314.

Hori, Y., Shirahige, K., Obuse, C., Tsurimoto, T. and Yoshikawa, H. (1996) Characterization of a novel CDC gene (ORC1) partly homologous to CDC6 of Saccharomyces cerevisiae. *Mol Biol Cell*, **7**, 409-418.

Hsiao, C.-L. and Carbon, J. (1979) High frequency transformation of yeast by plasmids containing the clonad yeast ARG4 gene. *Proc Natl AcadSsci USA*, **76**, 3829-3833.

Hua, X., Yan, H. and Newport, J. (1997) A role for Cdk2 kinase in negatively regulating DNA replication during S phase of the cell cycle. *J Cell Biol*, **137**, 183-192.

Huberman, J. and Riggs, A. (1968) On the mechanism of DNA replication in mammalian chromosomes. *32*, **2**.

Huberman, J., Spotila, L., Nawotka, K., el-Assouli, S. and Davis, L. (1987) The in vivo replication origin of the yeast 2 microns plasmid. *Cell 1987 Nov 6;51(3):473-81*, **51**, 473-481.

Huberman, J.A., Zhu, J.G., Davis, L.R. and Newlon, C.S. (1988) Close assotiation of a DNA replication origin and an ARS element on chromosome III of the yeast, *Saccharomyces cerevisiae* . *Nucleic Acids Res.*, **16**, 6373-6384.

Hyrien, O., Maric, C. and Mechali, M. (1995) Transition in Specification of Embryonic Metazoan DNA Replication Origins. *Science*, **270**, 994-997.

Hyrien, O. and Mechali, M. (1992) Plasmid replication in *Xenopus* eggs and egg extracts: a 2D gel electrophoretic analysis. *Nucleic Acids Res*, **20**, 1463-1469.

Hyrien, O. and Mechali, M. (1993) Chromosomal replication initiates and terminates at random sequences but at regular intervals in the ribosomal DNA of *Xenopus* early embryo. *EMBO J.*, **12**, 4511-4520.

Ishiai, M., Dean, F., Okumura, K., Abe, M., Moon, K., Amin, A., Kagotani, K., Taguchi, H., Murakami, Y., Hanaoka, F., O'Donnell, M., Hurwitz, J. and Eki, T. (1997) Isolation of human and fission yeast homologues of the budding yeast origin recognition complex subunit ORC5: human homologue (ORC5L) maps to 7q22. *Genomics*, **46**, 294-298.

Ishimi, Y. (1997) A DNA helicase activity is associated with an MCM4, -6, and -7 protein complex. *J Biol Chem*, **272**, 24508-24513.

Jackson, A., Pahl, P., Harrison, K., Rosamond, J. and Sclafani, R. (1993) Cell cycle regulation of the yeast Cdc7 protein kinase by association with the Dbf4 protein. *Mol Cell Biol*, **13**, 2899-2908.

Jackson, P., Chevalier, S., Philippe, M. and Kirschner, M. (1995) Early events in DNA replication require cyclin E and are blocked by p21CIP1. *J Cell Biol*, **130**, 755-769.

Jacob, F., Brenner, S. and Cuzin, F. (1963) On the regulation of DNA replication in bacteria. *Cold Spring Harbor Symp. Quant. Biol.*, **58**, 383-387.

Jiang, W. and Hunter, T. (1997) Identification and characterization of a human protein kinase related to budding yeast Cdc7p. *Proc Natl Acad Sci U S A*, **94**, 14320-14325.

Johnston, L. and Thomas, A. (1982) A further two mutants defective in initiation of the S phase in the yeast Saccharomyces cerevisiae. *Mol Gen Genet*, **186**, 445-448.

Kalejta, R., Li, X., Mesner, L., Dijkwel, P., Lin, H. and Hamlin, J. (1998) Distal sequences, but not ori-beta/OBR-1, are essential for initiation of DNA replication in the Chinese hamster DHFR origin. *Mol Cell*, **2**, 797-806.

Kelly, T., Martin, G., Forsburg, S., Stephen, R., Russo, A. and Nurse, P. (1993) The fission yeast cdc18+ gene product couples S phase to START and mitosis. *Cell*, **74**, 371-382.

Kim, C., Epner, E., Forrester, W. and Groudine, M. (1992) Inactivation of the human beta-globin gene by targeted insertion into the beta-globin locus control region. *Genes Dev*, **6**, 928-938.

Kim, J., Sato, N., Yamada, M., Arai, K. and Masai, H. (1998) Growth regulation of the expression of mouse cDNA and gene encoding a serine/threonine kinase related to *Saccharomyces cerevisiae* CDC7 essential for G1/S transition. *J Biol Chem*, **273**, 23248-23257.

Kim, S. and Huberman, J. (1998) Multiple orientation-dependent, synergistically interacting, similar domains in the ribosomal DNA replication origin of the fission yeast, *Schizosaccharomyces pombe*. *Mol Cell Biol*, **18**, 7294-7303.

Kimura, H., Takizawa, N., Nozaki, N. and Sugimoto, K. (1995) Molecular cloning of cDNA encoding mouse Cdc21 and CDC46 homologs and characterization of the products: physical interaction between P1(MCM3) and CDC46 proteins. *Nucleic Acids Res*, **23**, 2097-2104.

Kitada, K., Johnston, L., Sugino, T. and Sugino, A. (1992) Temperature-sensitive cdc7 mutations of Saccharomyces cerevisiae are suppressed by the DBF4 gene, which is required for the G1/S cell cycle transition. *Genetics*, **131**, 21-29.

Kitsberg, D., Selig, S., Keshet, I. and Cedar, H. (1993) Replication structure of the human beta-globin gene domain. *Nature*, **366**, 588-590.

Klemm, R., Austin, R. and Bell, S. (1997) Coordinate binding of ATP and origin DNA regulates the ATPase activity of the origin recognition complex. *Cell*, **88**, 493-502.

Koonin, E. (1993) A common set of conserved motifs in a vast variety of putative nucleic acid-dependent ATPases including MCM proteins involved in eukaryotic DNA replication. *Nucleic Acids Res*, **21**, 2541-2547.

Krude, T., Jackman, M., Pines, J. and Laskey, R. (1997) Cyclin/Cdk-dependent initiation of DNA replication in a human cell-free system. *Cell*, **88**, 109-119.

Krude, T., Musahl, C., Laskey, R. and Knippers, R. (1996) Human replication proteins hCdc21, hCdc46 and P1Mcm3 bind chromatin uniformly before S-phase and are displaced locally during DNA replication. *J Cell Sci*, **109**, 309-318.

Krysan, P., Smith, J. and Calos, M. (1993) Autonomous replication in human cells of multimers of specific human and bacterial DNA sequences. *Mol Cell Biol*, **13**, 2688-2696.

Kubota, Y., Mimura, S., Nishimoto, S., Takisawa, H. and Nojima, H. (1995) Identification of the yeast MCM3-related protein as a component of *Xenopus* DNA replication licensing factor. *Cell*, **81**, 601-609.

Kumagai, H., Sato, N., Yamada, M., Mahony, D., Seghezzi, W., Lees, E., Arai, K. and Masai, H. (1999) A novel growth- and cell cycle-regulated protein, ASK, activates human Cdc7-related kinase and is essential for G1/S transition in mammalian cells. *Mol Cell Biol*, **19**, 5083-5095.

Landis, G., Kelley, R., Spradling, A. and Tower, J. (1997) The k43 gene, required for chorion gene amplification and diploid cell chromosome replication, encodes the *Drosophila* homolog of yeast origin recognition complex subunit 2. *Proc Natl Acad Sci U S A*, **94**, 3888-3892.

Landis, G. and Tower, J. (1999) The *Drosophila* chiffon gene is required for chorion gene amplification, and is related to the yeast Dbf4 regulator of DNA replication and cell cycle. *Development*, **126**, 4281-4293.

Laskey, R. and Gurdon, J. (1973) Induction of polyoma DNA synthesis by injection into frog-egg cytoplasm. *Eur J Biochem*, **37**, 467-471.

Learn, B., Karzai, A. and McMacken, R. (1993) Transcription stimulates the establishment of bidirectional lambda DNA replication in vitro. *Cold Spring Harb Symp Quant Biol*, **58**, 389-402.

Leatherwood, J., Lopez-Girona, A. and Russell, P. (1996) Interaction of Cdc2 and Cdc18 with a fission yeast ORC2-like protein. *Nature*, **379**, 360-363.

Lee, D. and Bell, S. (1997) Architecture of the yeast origin recognition complex bound to origins of DNA replication. *Mol Cell Biol*, **17**, 7159-7168.

Lei, M., Kawasaki, Y. and Tye, B. (1996) Physical interactions among Mcm proteins and effects of Mcm dosage on DNA replication in Saccharomyces cerevisiae. *Mol Cell Biol*, **16**, 5081-5090.

Lei, M., Kawasaki, Y., Young, M., Kihara, M., Sugino, A. and Tye, B. (1997) Mcm2 is a target of regulation by Cdc7-Dbf4 during the initiation of DNA synthesis. *Genes Dev*, **11**, 3365-3374.

Leonard, A. and Helmstetter, C. (1986) Cell cycle specific replication of Escherichia coli minichromosome. *Proc Natl Acad Sci USA*, **83**, 5101-5105.

Leu, T. and Hamlin, J. (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.

Li, J. and Herskowitz, I. (1993) Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system. *Science 1993 Dec 17;262(5141):1870-4*, **262**, 1870-1874.

Liang, C. and Stillman, B. (1997) Persistent initiation of DNA replication and chromatin-bound MCM proteins during the cell cycle in cdc6 mutants. *Genes Dev*, **11**, 3375-3386.

Liang, C., Weinreich, M. and Stillman, B. (1995) ORC and Cdc6p interact and determine the frequency of initiation of DNA replication in the genome. *Cell*, **81**, 667-676.

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

Loo, S., Fox, C., Rine, J., Kobayashi, R., Stillman, B. and Bell, S. (1995) The origin recognition complex in silencing, cell cycle progression, and DNA replication. *Mol Biol Cell*, **6**, 741-756.

Madine, M., Khoo, C., Mills, A. and Laskey, R. (1995a) MCM3 complex required for cell cycle regulation of DNA replication in vertebrate cells. *Nature*, **375**, 421-424.

Madine, M., Khoo, C., Mills, A., Mushal, C. and Laskey, R. (1995b) The nuclear envelope prevents reinitiation of replication by regulating the binding of MCM3 to chromatin in *Xenopus* egg extracts. *Curr Biol*, **5**, 1270-1279.

Mahbubani, H.M., Paull, T., Elder, J.K. and Blow, J.J. (1992) DNA replication initiates at multiple sites on plasmid DNA in *Xenopus* egg extract. *Nucleic Acids Res.*, **20**, 1457-1462.

Maine, G., Sinha, P. and Tye, B. (1984) Mutants of *S. cerevisiae* defective in the maintenance of minichromosomes. *Genetics*, **106**, 365-385.

Marahrens, Y. and Stillman, B. (1992) A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science*, **255**, 817-823.

Masai, H., Miyake, T. and Arai, K. (1995) hsk1+, a *Schizosaccharomyces pombe* gene related to Saccharomyces cerevisiae CDC7, is required for chromosomal replication. *EMBO J*, **14**, 3094-3104.

Mastrangelo, I., Hough, P., Wall, J., Dodson, M., Dean, F. and Hurwitz, J. (1989) ATP-dependent assembly of double hexamers of SV40 T-antigen at the viral origin of DNA replication. *Nature*, **338**, 658-662.

Masukata, H., Satoh, H., Obuse, C. and Okazaki, T. (1993) Autonomous replication of human chromosomal DNA fragments in human cells. *Mol Biol Cell*, **4**, 1121-1132.

Maundrell, K., Hutchison, A. and Shall, S. (1988) Sequence analysis of ARS elements in fission yeast. *EMBO J*, **7**, 2203-2209.

Maundrell, K., Wright, A., Piper, M. and Shall, S. (1985) Evaluation of heterologous ARS activity in S. cerevisiae using cloned DNA from *S. pombe. Nucleic Acids Res*, **13**, 3711-3722.

McKay, R. (1983) Immunoassay for sequence-specific DNA-protein interaction. *Methods Enzymol*, **92**, 138-416.

McNally, F. and Rine, J. (1991) A synthetic silencer mediates SIR-dependent functions in *Saccharomyces cerevisiae. Mol Cell Biol*, **11**, 5648-565'9.

McVey, D., Brizuela, L., Mohr, I., Marshak, D., Gluzman, Y. and Beach, D. (1989) Phosphorylation of large tumour antigen by cdc2 stimulates SV40 DNA replication. *Nature*, **341**, 503-507.

Mears, J., Ramirez, F., Leibowitz, D. and Bank, A. (1978) Organization of human delta--and beta-globin genes in cellular DNA and the presence of intragenic inserts. *Cell*, **15**, 15-23.

Mechali, M. and Kearsey, S. (1984) Lack of specific sequence requirement for DNA replication in *Xenopus* eggs compared with high sequence specificity in yeast. *Cell*, **38**, 55-64.

Meijer, M., Beck, E., Hansen, F., Bergmans, H., Messer, W., von Meyenburg, K. and Schaller, H. (1979) Nucleotide sequence of the origin of replication of the Escherichia coli K-12 chromosome. *Proc Natl Acad Sci U S A*, **76**, 580-584.

Mendenhall, M. (1993) An inhibitor of p34CDC28 protein kinase activity from Saccharomyces cerevisiae. *Science*, **259**, 216-219.

Mensa-Wilmot, K., Carroll, K. and McMacken, R. (1989a) Transcriptional activation of bacteriophage lambda DNA replication in vitro: regulatory role of histone-like protein HU of Escherichia coli. *EMBO J*, **8**, 2393-2402.

Mensa-Wilmot, K., Seaby, R., Alfano, C., Wold, M., Gomes, B. and McMacken, R. (1989b) Reconstitution of a nine-protein system that initiates bacteriophage lambda DNA replication. *J Biol Chem*, **264**, 2853-2861.

Micklem, G., Rowley, A., Harwood, J., Nasmyth, K. and Diffley, J. (1993) Yeast origin recognition complex is involved in DNA replication and transcriptional silencing. *Nature*, **366**, 87-89.

Milbrandt, J., Azizkhan, J., Greisen, K. and Hamlin, J. (1983) Organization of a Chinese hamster ovary dihydrofolate reductase gene identified by phenotypic rescue. *Mol Cell Biol*, **3**, 1266-1273.

Milbrandt, J., Heintz, N., White, W., Rothman, S. and Hamlin, J. (1981) Methotrexate-resistant Chinese hamster ovary cells have amplified a135-kilobase-pair region that includes the dihydrofolate reductase gene. *Proc Natl Acad Sci U S A*, **78**, 6043-6047.

Mitchell, P., Carothers, A., Han, J., Harding, J., Kas, E., Venolia, L. and Chasin, L. (1986) Multiple transcription start sites, DNase I-hypersensitive sites, and an opposite-strand exon in the 5' region of the CHO dhfr gene. *Mol Cell Biol*, **6**, 425-440.

Moir, D., Stewart, S., Osmond, B. and Botstein, D. (1982) Cold-sensitive cell-division-cycle mutants of yeast: isolation, properties, and pseudoreversion studies. *Genetics*, **100**, 547-563.

Moore, D., Denniston-Thompson, K., Kruger, K., Furth, M., Williams, B., Daniels, D. and Blattner, F. (1979) Dissection and comparative anatomy of the origins of replication of lambdoid phages. *Cold Spring Harb Symp Quant Biol*, **43**.

182

Moreno, S. and Nurse, P. (1994) Regulation of progression through the G1 phase of the cell cycle by the rum1+ gene. *Nature*, **367**, 236-242.

Muzi-Falconi, M., Brown, G. and Kelly, T. (1996) cdc18+ regulates initiation of DNA replication in Schizosaccharomyces pombe. *Proc Natl Acad Sci U S A*, **93**, 1566-1570.

Muzi-Falconi, M. and Kelly, T. (1995) Orp1, a member of the Cdc18/Cdc6 family of S-phase regulators, is homologous to a component of the origin recognition complex. *Proc Natl Acad Sci U S A*, **92**, 12475-12479.

Myers, R. and Tjian, R. (1980) Construction and analysis of simian virus 40 origins defective in tumor antigen binding and DNA replication. *Proc Natl Acad Sci U S A*, **77**, 6491-6495.

Nishitani, H. and Nurse, P. (1995) p65cdc18 plays a major role controlling the initiation of DNA replication in fission yeast. *Cell*, **83**, 397-405.

Oka, A., Sugimoto, K., Takanami, M. and Hirota, Y. (1980) Replication Origin of *Escherichia coli* K-12 Chromosome: The Size and Structure of the Minimum DNA segment Carrying the Informationfor Autonomous Replication. *Molec Gen Genet*, **178**, 9-20.

Okishio, N., Adachi, Y. and Yanagida, M. (1996) Fission yeast Nda1 and Nda4, MCM homologs required for DNA replication, are constitutive nuclear proteins. *J Cell Sci*, **109** (Pt 2), 319-326.

Orlando, V. and Paro, R. (1993) Mapping Polycomb-repressed domains in the bithorax complex using in vivo formaldehyde cross-linked chromatin. *Cell*, **75**, 1187-1198.

Orr-Weaver, T. and Spradling, A. (1986) *Drosophila* chorion gene amplification requires an upstream region regulating s18 transcription. *Mol Cell Biol*, **6**, 4624-4633.

Oshiro, G., Owens, J., Shellman, Y., Sclafani, R. and Li, J. (1999) Cell cycle control of Cdc7p kinase activity through regulation of Dbf4p stability. *Mol Cell Biol*, **19**, 4888-4896.

Owens, J., Detweiler, C. and Li, J. (1997) CDC45 is required in conjunction with CDC7/DBF4 to trigger the initiation of DNA replication. *Proc Natl Acad Sci U S A*, **94**, 12521-12526.

Pak, D., Pflumm, M., Chesnokov, I., Huang, D., Kellum, R., Marr, J., Romanowski, P. and Botchan, M. (1997) Association of the origin recognition complex with heterochromatin and HP1 in higher eukaryotes. *Cell*, **91**, 311-323.

Patterson, M., Sclafani, R., Fangman, W. and Rosamond, J. (1986) Molecular characterization of cell cycle gene CDC7 from Saccharomyces cerevisiae. *Mol Cell Biol*, **6**, 1590-1598.

Petersen, B., Lukas, J., Sorensen, C., Bartek, J. and Helin, K. (1999) Phosphorylation of mammalian CDC6 by cyclin A/CDK2 regulates its subcellular localization. *EMBO*, **18**, 396-410.

Piatti, S., Bohm, T., Cocker, J., Diffley, J. and Nasmyth, K. (1996) Activation of S-phase-promoting CDKs in late G1 defines a "point of no return" after which Cdc6 synthesis cannot promote DNA replication in yeast. *Genes Dev*, **10**, 1516-1531.

Piatti, S., Lengauer, C. and Nasmyth, K. (1995) Cdc6 is an unstable protein whose de novo synthesis in G1 is important for the onset of S phase and for preventing a 'reductional' anaphase in the budding yeast *Saccharomyces cerevisiae. EMBO J*, **14**, 3788-3799.

Quintana, D., Hou, Z., Thome, K., Hendricks, M., Saha, P. and Dutta, A. (1997) Identification of HsORC4, a member of the human origin of replication recognition complex. *J Biol Chem*, **272**, 28247-28251.

Quintana, D., Thome, K., Hou, Z., Ligon, A., Morton, C. and Dutta, A. (1998) ORC5L, a new member of the human origin recognition complex, is deleted in uterine leiomyomas and malignant myeloid diseases. *J Biol Chem*, **273**, 27137-27145.

Rao, H., Marahrens, Y. and Stillman, B. (1994) Functional conservation of multiple elements in yeast chromosomal replicators. *Mol Cell Biol*, **14**, 7643-7651.

Rao, H. and Stillman, B. (1995) The origin recognition complex interacts with a bipartite DNA binding site within yeast replicators. *Proc Natl Acad Sci U S A*, **92**, 2224-2228.

Rao, P. and Johnson, R. (1970) Mammalian cell fusion: studies on the regulation of DNA synthesis and mitosis. *Nature*, **225**, 159-164.

Rhode, P., Elsasser, S. and Campbell, J. (1992) Role of multifunctional autonomously replicating sequence binding factor 1 in the initiation of DNA replication and transcriptional control in Saccharomyces cerevisiae. *Mol Cell Biol*, **12**, 1064-1067.

Rivier, D. and Rine, J. (1992) An origin of DNA replication and a transcription silencer require a common element. *Science*, **256**, 659-663.

Romanowski, P., Madine, M. and Laskey, R. (1996a) XMCM7, a novel member of the *Xenopus* MCM family, interacts with XMCM3 and colocalizes with it throughout replication. *Proc Natl Acad Sci U S A*, **93**, 10189-10194.

Romanowski, P., Madine, M., Rowles, A., Blow, J.J. and Laskey, R.A. (1996b) The *Xenopus* origin recognition complex is essential for DNA replication and MCM binding to chromatin. *Current Biology*, **6**, 1416-1425.

Rowles, A., Chong, J.P.J., Brown, L., Howell, M., Evan, G.I. and Blow, J.J. (1996) Interaction between the Origin Recognition Complex and the Replication Licencing System in *Xenopus*. *Cell*, **87**, 287-296.

Rowles, A., Tada, S. and Blow, J. (1999) Changes in association of the *Xenopus* origin recognition complex with chromatin on licensing of replication origins. *J Cell Sci*, **112**, 2011-2018.

Rowley, A., Cocker, J., Harwood, J. and Diffley, J. (1995) Initiation complex assembly at budding yeast replication origins begins with the recognition of a bipartite sequence by limiting amounts of the initiator, ORC. *EMBO*, **14**, 2631-2641.

Santocanale, C. and Diffley, J. (1996) ORC- and Cdc6-dependent complexes at active and inactive chromosomal replication origins in *Saccharomyces cerevisiae. EMBO J*, **15**, 6671-6679.

184

Santocanale, C. and Diffley, J. (1998) A Mec1- and Rad53-dependent checkpoint controls late-firing origins of DNA replication. *Nature*, **395**, 615-618.

Sato, N., Arai, K. and Masai, H. (1997) Human and *Xenopus* cDNAs encoding budding yeast Cdc7-related kinases: in vitro phosphorylation of MCM subunits by a putative human homologue of Cdc7. *EMBO J*, **16**, 4340-4351.

Schnos, M. and Inman, R. (1970) Position of branch points in replicating lambda DNA. *J Mol Biol*, **51**, 61-73.

Schwob, E., Bohm, T., Mendenhall, M. and Nasmyth, K. (1994) The B-type cyclin kinase inhibitor p40SIC1 controls the G1 to S transition in *S. cerevisiae*. *Cell*, **79**, 233-244.

Sekimizu, K., Bramhill, D. and Kornberg, A. (1987) ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the *E. coli* chromosome. *Cell*, **50**, 259-265.

Sekimizu, K. and Kornberg, A. (1988) Cardiolipin activation of dnaA protein, the initiation protein of replication in *Escherichia coli*. *J Biol Chem*, **263**, 7131-7135.

Solomon, M., Strauss, F. and Varshavsky, A. (1986) A mammalian high mobility group protein recognizes any stretch of six A.T base pairs in duplex DNA. *83*, 5.

Spradling, A. (1981) The organization and amplification of two chromosomal domains containing Drosophila chorion genes. *Cell*, **27**, 193-201.

Spradling, A. and Mahowald, A. (1980) Amplification of genes for chorion proteins during oogenesis in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A*, **77**, 1096-1100.

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

Strausfeld, U., Howell, M., Descombes, P., Chevalier, S., Rempel, R., Adamczewski, J., Maller, J., Hunt, T. and Blow, J. (1996) Both cyclin A and cyclin E have S-phase promoting (SPF) activity in *Xenopus* egg extracts. *J Cell Sci*, **109**, 1555-1563.

Strausfeld, U., Howell, M., Rempel, R., Maller, J., Hunt, T. and Blow, J. (1994) Cip1 blocks the initiation of DNA replication in *Xenopus* extracts by inhibition of cyclin-dependent kinases. *Curr Biol*, **4**, 876-883.

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

Su, T., Feger, G. and O'Farrell, P. (1996) *Drosophila* MCM protein complexes. *Mol Biol Cell*, **7**, 319-329.

Sweder, K., Rhode, P. and Campbell, J. (1988) Purification and characterization of proteins that bind to yeast ARSs. *J Biol Chem*, **263**, 17270-17277.

185

Swimmer, C., Delidakis, C. and Kafatos, F. (1989) Amplification-control element ACE-3 is important but not essential for autosomal chorion gene amplification. *Proc Natl Acad Sci U S A*, **86**, 8823-8827.

Szymanska, G., Leszyk, J. and O'Connor, C. (1998) Carboxyl methylation of deamidated calmodulin increases its stability in *Xenopus* oocyte cytoplasm. Implications for protein repair. *J Biol Chem*, **273**, 28516-28523.

Tanaka, T., Knapp, D. and Nasmyth, K. (1997) Loading of an Mcm protein onto DNA replication origins is regulated by Cdc6p and CDKs. *Cell*, **90**, 649-660.

Tanaka, T. and Nasmyth, K. (1998) Association of RPA with chromosomal replication origins requires an Mcm protein, and is regulated by Rad53, and cyclin- and Dbf4-dependent kinases. *EMBO J*, **17**, 5182-5191.

Theis, J. and Newlon, C. (1994) Domain B of ARS307 contains two functional elements and contributes to chromosomal replication origin function. *Mol Cell Biol*, **14**, 7652-7659.

Tjian, R. (1979) Protein-DNA interactions at the origin of simian virus 40 DNA replication. *Cold Spring Harb Symp Quant Biol*, **43**, 655-661.

Tschumper, G. and Carbon, J. (1980) Sequence of a yeast DNA fragment containing a chromosomal replicator and the TRP1 gene. *Gene*, **10**, 157-166.

Tsurimoto, T. and Matsubara, K. (1981) Purified bacteriophage lambda O protein binds to four repeating sequences at the lambda replication origin. *Nucleic Acids Res*, **9**, 1789-1799.

Tugal, T., Zou-Yang, X., Gavin, K., Pappin, D., Canas, B., Kobayashi, R., Hunt, T. and Stillman, B. (1998) The Orc4p and Orc5p subunits of the *Xenopus* and human origin recognition complex are related to Orc1p and Cdc6p. *J Biol Chem*, **273**, 32421-32429.

Van Houten, J.V. and Newlon, C.S. (1990) Mutational analysis of the consensus sequence of a replication origin from yeast chromosome III. *Mol. Cell. Biol.*, **10**, 3917-3925.

Vaughn, J., Dijkwel, P. and Hamlin, J. (1990) Replication initiates in a broad zone in the amplified CHO dihydrofolate reductase domain. *Cell*, **61**, 1075-1087.

Wahle, E., Lasken, R. and Kornberg, A. (1989a) The dnaB-dnaC replication protein complex of *Escherichia coli*. I. Formation and properties. *Biol Chem 1989 Feb 15;264(5):2463-8*, **264**, 2463-2468.

Wahle, E., Lasken, R. and Kornberg, A. (1989b) The dnaB-dnaC replication protein complex of *Escherichia coli*. II. Role of the complex in mobilizing dnaB functions. *J Biol Chem*, **264**, 2469-2475.

Walker, J., Saraste, M., Runswick, M. and Gay, N. (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J*, **1**, 945-951.

Wang, S., Dijkwel, P. and Hamlin, J. (1998) Lagging-strand, early-labelling, and two-dimensional gel assays suggest multiple potential initiation sites in the Chinese hamster dihydrofolate reductase origin. *Mol Cell Biol*, **18**, 39-50.

Wickner, S. (1979) DNA replication proteins of Escherichia coli and phage lambda. *Cold Spring Harb Symp Quant Biol*, **43**, 303-310.

Williams, R., Shohet, R. and Stillman, B. (1997) A human protein related to yeast Cdc6p. *Proc Natl Acad Sci U S A*, **94**, 142-147.

Wohlgemuth, J., Bulboaca, G., Moghadam, M., Caddle, M. and Calos, M. (1994) Physical mapping of origins of replication in the fission yeast *Schizosaccharomyces pombe*. *Mol Biol Cell*, **5**, 839-849.

Yan, H., Merchant, A. and Tye, B. (1993) Cell cycle-regulated nuclear localization of MCM2 and MCM3, which are required for the initiation of DNA synthesis at chromosomal replication origins in yeast. *Genes Dev*, **7**, 2149-2160.

Yasuda, S. and Hirota, Y. (1977) Cloning and mapping of the replication origin of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA*, **74**, 5458-5462.

Yoon, H. and Campbell, J. (1991) The CDC7 protein of Saccharomyces cerevisiae is a phosphoprotein that contains protein kinase activity. *Proc Natl Acad Sci U S A*, **88**, 3574-3578.

Yoon, H., Loo, S. and Campbell, J. (1993) Regulation of Saccharomyces cerevisiae CDC7 function during the cell cycle. *Mol Biol Cell*, **4**, 195-208.

Zhu, J.G., Newlon, C.S. and Huberman, J.A. (1992) Localization of a DNA replication origin and termination zone on chromosome III of *Saccharomyces cerevisiae*. *Mol. Cell Biol.*, **12**, 4733-4741.

Zou, L., Mitchell, J. and Stillman, B. (1997) CDC45, a novel yeast gene that functions with the origin recognition complex and Mcm proteins in initiation of DNA replication. *Mol Cell Biol*, **17**, 553-563.

Appendix

Xenopus Orc1-full length sequence

```
1/1                                  31/11
ATG ACC AAA TAT TTT ACC CGT GTG AAA AAG CAC AGA GTC TAT GCC TGG AAA GGG AGA CCT
M   T   K   Y   F   T   R   V   K   K   H   R   V   Y   A   W   K   G   R   P

61/21                                91/31
TCC ATA AAA GAC AGA AAA TTA CTG AAT CTG CAT TAT GGA GGG CTC GAT ATA AAA TCT GAA
S   I   K   D   R   K   L   L   N   L   H   Y   G   G   L   D   I   K   S   E

121/41                               151/51
GGC ATT GGG TGT ATA ACG GTA ACT CCT GGA GAC TTT GTT CTG ATA GAA GGG GAA AAT GAG
G   I   G   C   I   T   V   T   P   G   D   F   V   L   I   E   G   E   N   E

181/61                               211/71
GAG AGA CCA TTT GTT GCA AAG CTC CAA GAA TTA TAC GAT GAT GGA AAT GAA AAA CAC ACA
E   R   P   F   V   A   K   L   Q   E   L   Y   D   D   G   N   E   K   H   T

241/81                               271/91
AGC AAG CAT GCC ATT GTA CAA TGG TTT TTA AGG TAT GAA GAA GTG CCA TAC AAG AAG CGA
S   K   H   A   I   V   Q   W   F   L   R   Y   E   E   V   P   Y   K   K   R

301/101                              331/111
GCC CTA CTT GGG CGG GAG CCT CAT CAA GAG GAA ATA TTC TTG TAT GAC GTC CCT TCC TGC
A   L   L   G   R   E   P   H   Q   E   E   I   F   L   Y   D   V   P   S   C

361/121                              391/131
GAA AAT GAC ATT GAT GCA GAG ACA ATT ATT GGA AGT GTT AAG GTA ACA CAG TTG AAA CCA
E   N   D   I   D   A   E   T   I   I   G   S   V   K   V   T   Q   L   K   P

421/141                              451/151
GGA GAG AAT TTC CCT GTA GAA AAA GAT GCT GGA ATA CTA TTT GTG AAG TTG GCA TGG AAC
G   E   N   F   P   V   E   K   D   A   G   I   L   F   V   K   L   A   W   N

481/161                              511/171
GGG AAA GCC TTC AAG CCA TTA ACT TCC AGT ATT TGT CAT AAA TAT GGA GAG ATT GCA GCA
G   K   A   F   K   P   L   T   S   S   I   C   H   K   Y   G   E   I   A   A

541/181                              571/191
AAC AAT CAC CTA GAG GAT CAT AAT TCT TTG CCA AAG CCC CTC TTC TAT CCG TTG GAG GGA
N   N   H   L   E   D   H   N   S   L   P   K   P   L   F   Y   P   L   E   G

601/201                              631/211
GAT ACT GCC GAA ATC AAG CAA ATC ACT CGC AGT GCA ACC AGG GGA CAA AGA AAC GAG GCA
D   T   A   E   I   K   Q   I   T   R   S   A   T   R   G   Q   R   N   E   A

661/221                              691/231
GAA GTA GAG TCA AGA CAT TCG GCT TCC AAG AGC TCC CAT TGT AAG GAG AGA GGG ATT CAG
E   V   E   S   R   H   S   A   S   K   S   S   H   C   K   E   R   G   I   Q

721/241                              751/251
AGA GTC TCT AGT ACA ACC AGC ACT CCT GGA GCT AGA AAG AAG CTG CAA CTA AGT AGT CCA
R   V   S   S   T   T   S   T   P   G   A   R   K   K   L   Q   L   S   S   P

781/261                              811/271
ACA AAA TCT GAA GTC TCT AGC CTT GAG CAA GAG GAC ACC ACA TTG GAT GAG GAA TTC AAA
T   K   S   E   V   S   S   L   E   Q   E   D   T   T   L   D   E   E   F   K

841/281                              871/291
ATG CTG GAA TCC ATA GCT CCC AAG CGG AAA GTG GCA TTT AGC GGA ATC CCT GAT TCA CCA
M   L   E   S   I   A   P   K   R   K   V   A   F   S   G   I   P   D   S   P

901/301                              931/311
CCA AAG AAA GTC CTC AGC GAC ATC CCT AAT AGT GCT AGT ATC TTG CCA TTA AAG GGA AAC
P   K   K   V   L   S   D   I   P   N   S   A   S   I   L   P   L   K   G   N

961/321                              991/331
GCA AAT AGT CAG ACA AGA CTC CGT TTA AGA AAA AAC TGC ATA GAA GTG AAT GGA AAT AAA
A   N   S   Q   T   R   L   R   L   R   L   T   P   L   R   L   P   L   E   Q

1021/341                             1051/351
CTG ACT CCA TTG AGA CTT CCG CTG GAA CAA TCA TGC AGC AGT CTG CCA GAG AAC CAG AGT
S   C   S   S   L   P   E   N   Q   S   K   N   C   I   E   V   N   G   N   K
```

```
1081/361                                1111/371
GCT GTT GAT AAT GGA AGA CCT GTT AAG TCT AAA ACA CCC CAA AAA CTG GAA CCC ACA CCA
A   V   D   N   G   R   P   V   K   S   K   T   P   Q   K   L   E   P   T   P

1141/381                                1171/391
AAG AAA GGC ACG TCT CGA CCT TTA AAA AGG TCT GAG GAA GAC TTT GAG GGA GAC CCA CAG
K   K   G   T   S   R   P   L   K   R   S   E   E   D   F   E   G   D   P   Q

1201/401                                1231/411
ACA CCC CGT TCC CGC AGG ACA TCT GCT CGG AAG ACT GCA AAC AAA GTC CGT GCT CAA CTG
T   P   R   S   R   R   T   S   A   R   K   T   A   N   K   V   R   A   Q   L

1261/421                                1291/431
AAG GCT TTA ATG GAC AGT GAT GAC GAT GTT GAT GAG GAC AAA GAC TTT GTT CCA ATA AAA
K   A   L   M   D   S   D   D   D   V   D   E   D   K   D   F   V   P   I   K

1321/441                                1351/451
GAG ACT TGT ACA CCC AGT GAC AGT GAA GAA GAG AAC GAA CAA GAG GAA GTG CCT TTA AAA
E   T   C   T   P   S   D   S   E   E   E   N   E   Q   E   E   V   P   L   K

1381/461                                1411/471
AGG AAT GTA AAG AGA GGA TGT GCT ACA CCT AGA ACA CCA ACC TCC TCC AGA AAA ACC TCT
R   N   V   K   R   G   C   A   T   P   R   T   P   T   S   S   R   K   T   S

1441/481                                1471/491
TCA AGA ACA CCC AAT AAA ACA CCC ACC TCC AAA ACG CCA CAT TCT GCA ACT CCT AGA ATT
S   R   T   P   N   K   T   P   T   S   K   T   P   H   S   A   T   P   R   I

1501/501                                1531/511
CCA GAG CGG AAC CAG CCA GTT AAA AAA CCG AGC AAT ATG CTG GAG GAA GCC AGA ATT AGA
P   E   R   N   Q   P   V   K   K   P   S   N   M   L   E   E   A   R   I   R

1561/521                                1591/531
CTT CAT GTC TCT GCT GTG CCT GAA TCA CTG CCC TGT CGG GAG CAA GAA TAC CAA GAT GTT
L   H   V   S   A   V   P   E   S   L   P   C   R   E   Q   E   Y   Q   D   V

1621/541                                1651/551
TAT AAC TTT GTG GAA AGC AAG CTT TTA GAT GGC ACT GGG GGA TGC ATG TAC ATC TCT GGG
Y   N   F   V   E   S   K   L   L   D   G   T   G   G   C   M   Y   I   S   G

1681/561                                1711/571
GTA CCA GGT ACA GGA AAA ACA GCC ACT GTT CAT GAG GTG ATT CGC AGT TTA CAG GAA AGT
V   P   G   T   G   K   T   A   T   V   H   E   V   I   R   S   L   Q   E   S

1741/581                                1771/591
GCA GAG GAA GAA GAG CTG CCT ATG TTC CAC TAC ATT GAA ATA AAT GGC ATG AAA CTG ACT
A   E   E   E   E   L   P   M   F   H   Y   I   E   I   N   G   M   K   L   T

1801/601                                1831/611
GAT CCT CAT CAG GCC TAT GTG CAG ATC CTG AAG CTC TTA ACT GGC CAA AAA GCC ACA GCA
D   P   H   Q   A   Y   V   Q   I   L   K   L   L   T   G   Q   K   A   T   A

1861/621                                1891/631
GAT CAC GCT GCA GCA CTC CTG GAG AAG CGC TTT AGC ACC CCA GCA TCC AAA AAA GAA ACT
D   H   A   A   A   L   L   E   K   R   F   S   T   P   A   S   K   K   E   T

1921/641                                1951/651
ACA GTG CTC TTG GTG GAT GAG CTG GAC TTA TTG TGG ACA AGA AAA CAG AAT GTC ATG TAT
T   V   L   L   V   D   E   L   D   L   L   W   T   R   K   Q   N   V   M   Y

1981/661                                2011/671
AGT TTG TTT GAC TGG CCG ACC AGA AAA CAA GCA AAA TTG ATT GTC CTG GCA ATA GCA AAT
S   L   F   D   W   P   T   R   K   Q   A   K   L   I   V   L   A   I   A   N

2041/681                                2071/691
ACT ATG GAT CTG CCA GAG AGG ATC ATG ATG AAT AGA GTG GCC AGC AGA CTG GGT CTC ACC
T   M   D   L   P   E   R   I   M   M   N   R   V   A   S   R   L   G   L   T

2101/701                                2131/711
AGG ATG TCC TTC CAG CCT TAC ACT CAC AAA CAG CTG CAG CAG ATA ATT ACT TCT AGG CTA
R   M   S   F   Q   P   Y   T   H   K   Q   L   Q   Q   I   I   T   S   R   L

2161/721                                2191/731
AAT CAC ATC AAA GCT TTT GAT GAT GAT GCT ATT CAG CTT GTT GCA AGA AAG GTG GCT GCA
N   H   I   K   A   F   D   D   D   A   I   Q   L   V   A   R   K   V   A   A

2221/741                                2251/751
CTG TCG GGA GAT GCT CGC AGA TGC TTG GAT ATA TGT CGT AGA GCC ACA GAA ATA TGT GAG
L   S   G   D   A   R   R   C   L   D   I   C   R   R   A   T   E   I   C   E

2281/761                                2311/771
TTT TCT CAT TCG ATG GCT GAC TCA TCT CTG GTT AAG ATG TCC CAT GTG ATT GAA GCA TTG
F   S   H   S   M   A   D   S   S   L   V   K   M   S   H   V   I   E   A   L
```

189

```
2341/781                                   2371/791
GAG GAA ATG TTT TCT TCA CCT TAT GTA ACT   GCA ATC CGA AAC TCT TCG TTA ATG GAA CAG
E   E   M   F   S   S   P   Y   V   T     A   I   R   N   S   S   L   M   E   Q

2401/801                                   2431/811
ACA TTC CTA CGG GCT GTA ATT GCT GAA TTT   CGC AGA TCT GGA CTT GAG GAG GCT ACT TTT
T   F   L   R   A   V   I   A   E   F     R   R   S   G   L   E   E   A   T   F

2461/821                                   2491/831
CAA CAG ATA TAT CGA CAG CAT GTT GCA CTG   TGC AGG ATA GAG GGA TTG CAA CCT CCA CAG
Q   Q   I   Y   R   Q   H   V   A   L     C   R   I   E   G   L   Q   P   P   Q

2521/841                                   2551/851
ATG TCT GAG ACC ATG GCT GTG TGC CAT AGA   CTT GGA GCC TCT CGG CTA CTC CTG GTG GAA
M   S   E   T   M   A   V   C   H   R     L   G   A   S   R   L   L   L   V   E

2581/861                                   2611/871
TCA AGC AGA AAT GAC CTT CAC CTA AGA GTG   AGG ATA AAC GTC AGT CAG GAT GAT ATT ATG
S   S   R   N   D   L   H   L   R   V     R   I   N   V   S   Q   D   D   I   M

2641/881
TAT GCC CTG AAG GAG GAA c
Y   A   L   K   E   E
```

Xenopus Orc4-full length sequence

```
1/1                                        31/11
ATG AGC AAA CGA AAG TCT AAG GAT CTT TGT   CTC CCG GTA GGA GAA TGC ATC TCA AAG GTT
M   S   K   R   K   S   K   D   L   C     L   P   V   G   E   C   I   S   K   V

61/21                                      91/31
CAT GGA ATT TTA AGG CAA AGG CTT TTT CAG   CAT CAC GGC AAA CCG TTT GGA GTA GAT TCT
H   G   I   L   R   Q   R   L   F   Q     H   H   G   K   P   F   G   V   D   S

121/41                                     151/51
CAG CAC AAG CAT TTG GTT GAA CTC CTG AAA   AGA ACC GTG ATT CAT GGT GAA AGT AAC TCG
Q   H   K   H   L   V   E   L   L   K     R   T   V   I   H   G   E   S   N   S

181/61                                     211/71
GCC CTG ATT ATT GGC CCT AGG GGA TCA GGA   AAG AGC ATG TTG CTA AAA GGA GCT TTA GAA
A   L   I   I   G   P   R   G   S   G     K   S   M   L   L   K   G   A   L   E

241/81                                     271/91
GAC ATA TTT GGA ATG AAG CAG ATG AAG GAG   ACT GCA CTG CAA GTT AAT CTA AAT GGT CTG
D   I   F   G   M   K   Q   M   K   E     T   A   L   Q   V   N   L   N   G   L

301/101                                    331/111
TTA CAA ACT ACT GAT AAG ATT GCT CTC AAA   GAG ATT ACC AGA CAG CTA CAT TTA GAA AAT
L   Q   T   T   D   K   I   A   L   K     E   I   T   R   Q   L   H   L   E   N

361/121                                    391/131
GCT GAA AAC TTG TCT TTC CTT CTT GAA GCA   TTA AAA ACA GGT GAT CGG AAA AGT AGC TGT
V   V   G   D   R   V   F   G   S   F     A   E   N   L   S   F   L   L   E   A

421/141                                    451/151
GTC GTT GGG GAC AGA GTT TTT GGA AGT TTT   CCA GTG CTG TTT GTT TTG GAT GAG TTT GAC
L   K   T   G   D   R   K   S   S   C     P   V   L   F   V   L   D   E   F   D

481/161                                    511/171
TTA TTT GCC CAT CAC AAG AAT CAG ACT CTG   CTT TAT AAT CTT TTC GAT ATT GCT CAG TCG
L   F   A   H   H   K   N   Q   T   L     L   Y   N   L   F   D   I   A   Q   S

541/181                                    571/191
GCC CAA ACA CCT GTT GCA GTA ATT GGA CTA   ACA TGC CGC CTG GAC GTT ATG GAA CTT CTA
A   Q   T   P   V   A   V   I   G   L     T   C   R   L   D   V   M   E   L   L

601/201                                    631/211
GAG AAA AGG GTG AAA TCT CGG TTC TCC CAC   AGA CAG ATC CAT CTG CTC AAT TCC TTC AGC
E   K   R   V   K   S   R   F   S   H     R   Q   I   H   L   L   N   S   F   S

661/221                                    691/231
TTC TCC CAG TAC CTG CAA ATA TTT CAA GAG   AAA CTT TCT CTG CCT GCG TCT TTC CCA GAC
F   S   Q   Y   L   Q   I   F   Q   E     K   L   S   L   P   A   S   F   P   D

721/241                                    751/251
ATT AAG AGT CTT GTA GAA AGT AAG CTG GTG   GAA GAT GTT CTT CAA AAA CAG TAC AAT GCC
S   Q   F   A   E   K   W   N   E   S     I   K   S   L   V   E   S   K   L   V

781/261                                    811/271
TCC CAG TTT GCT GAG AAA TGG AAC GAG AGC   AGC AAA GAT GTG CGC TCC TTA CAT ATG CTC
E   D   V   L   Q   K   Q   Y   N   A     S   K   D   V   R   S   L   H   M   L
```

190

```
841/281                              871/291
ATG CTC CTG GCT GTC TGC CGT GTC AAT GTG TCT CAC CCC CAT ATA ACT GCC GCG GAT TTC
 M   L   L   A   V   C   R   V   N   V   S   H   P   H   I   T   A   A   D   F

901/301                              931/311
TCA AAA GCA AAT ATA CTG CAC GGT GTG TCT GTT CTA GAG CTC TGT CTA ATC ATT GCC ATG
 L   E   V   F   R   L   R   N   Q   D   S   K   A   N   I   L   H   G   V   S

961/321                              991/331
TTG GAA GTC TTT CGC CTC CGA AAC CAG GAC AAG CAC TTA CAG GAC ATT TAT GAT GGG GAA
 V   L   E   L   C   L   I   I   A   M   K   H   L   Q   D   I   Y   D   G   E

1021/341                             1051/351
CCG TTT AAT TTC CAG ATG GTG CAC AAT GAA TTC CAG AAG TTT ATT CAG AGG AAA GCT CAC
 P   F   N   F   Q   M   V   H   N   E   F   Q   K   F   I   Q   R   K   A   H

1081/361                             1111/371
TCT GTC TAT AAC TTT GAA AAG GCC GTA GTA ATT AAG GCG TTT GAG CAC TTG CAC CAG TTG
 S   V   Y   N   F   E   K   A   V   V   I   K   A   F   E   H   L   H   Q   L

1141/381                             1171/391
GAG CTG ATC AAG CCC ATG GAA GGT CTG TCC GTG CGC ACG CAG AAA GAA TAC CGA CTC ATG
 E   L   I   K   P   M   E   G   L   S   V   R   T   Q   K   E   Y   R   L   M

1201/401                             1231/411
AAG CTC CTC CTG GAT AAC ACT CAG ATT GTA GAA GCA CTT CAG AAG TAC CCC AAC TGC CCT
 K   L   L   L   D   N   T   Q   I   V   E   A   L   Q   K   Y   P   N   C   P

1261/421                             1291/431
ACA GAT GTG AAG CAG TGG GCA ATG TCC TCA CTC AGC TGA
 T   D   V   K   Q   W   A   M   S   S   L   S   *
```

Mouse Orc4-full length sequence

```
1/1                                  31/11
ATG AGC AGT CGT AAA ACC AAG AGT AAT GCA CAT GCA GAG TGT CTT TCA CAG GTA CAA AGA
 M   S   S   R   K   T   K   S   N   A   H   A   E   C   L   S   Q   V   Q   R

61/21                                91/31
ATT TTA CGT GAA AGA TTT TGT CAT CAT AGT CCA CAT AGC AAC CTC TTT GGA GTG CAA GTA
 I   L   R   E   R   F   C   H   H   S   P   H   S   N   L   F   G   V   Q   V

121/41                               151/51
CAA TAC AAG CAT TTG ATT GAG CTA CTC AAA AGA ACT GCT ATC TAT GGA GAA AGC AAT TCT
 Q   Y   K   H   L   I   E   L   L   K   R   T   A   I   Y   G   E   S   N   S

181/61                               211/71
GTA CTC ATT GTT GGA CCA CGA GGA TCA GGA AAG ACC ACA TTA TTA AAT CAT GCT TTG AAA
 V   L   I   V   G   P   R   G   S   G   K   T   T   L   L   N   H   A   L   K

241/81                               271/91
GAA CTC ATG GAA ATA GAA GTG AGT GAA AAT GTG ATA CAA GTT CAC TTA AAT GGA CTC CTG
 E   L   M   E   I   E   V   S   E   N   V   I   Q   V   H   L   N   G   L   L

301/101                              331/111
CAG ACA AAT GAA AAA ATT GCT CTG AAA GAG ATC ACA AGA CAA TTA AAT CTA GAC AAT GTA
 Q   T   N   E   K   I   A   L   K   E   I   T   R   Q   L   N   L   D   N   V

361/121                              391/131
GTT GAA GAT AAA GTT TTT GGA AGC TTT GCT GAA AAC CTT TCA TTT CTT CTG GAA GCT TTA
 V   E   D   K   V   F   G   S   F   A   E   N   L   S   F   L   L   E   A   L

421/141                              451/151
AAA AAA GGT GAC CGG ACT AGC AGT TGC CCA GTG ATC TTC ATA CTG GAT GAA TTT GAT ATT
 K   K   G   D   R   T   S   S   C   P   V   I   F   I   L   D   E   F   D   I

481/161                              511/171
TTT GCT CAT CAG AAA AAT CAA ACA CTC CTT TAT AAT CTT TTT GAC ATT TCT CAA TCT GCA
 F   A   H   Q   K   N   Q   T   L   L   Y   N   L   F   D   I   S   Q   S   A

541/181                              571/191
CAG ACA CCA GTA GCA GTT ATT GGA CTT ACA TGT AGA TTG GAT ATT TTG GAA CTC TTA GAA
 Q   T   P   V   A   V   I   G   L   T   C   R   L   D   I   L   E   L   L   E

601/201                              631/211
AAG AGA GTG AAG TCA CGA TTT TCT CAC CGG CAG ATA CAT TTA ATG AAT TCA TTT GAT TTT
 K   R   V   K   S   R   F   S   H   R   Q   I   H   L   M   N   S   F   D   F
```

191

```
661/221                                   691/231
CCA CAA TAT TTG AAA ATA TTT AAA GAA CAG TTA TCT CTA CCT GCA GAA TTC CCA GAT AAG
P   Q   Y   L   K   I   F   K   E   Q   L   S   L   P   A   E   F   P   D   K

721/241                                   751/251
GCT TTT GCT GAG AGA TGG AAT GAG AAT GTT CAC TGT CTC TCT GAA GAT TCA ACT GTG CTT
A   F   A   E   R   W   N   E   N   V   H   C   L   S   E   D   S   T   V   L

781/261                                   811/271
GAA GTC CTA CAG AAA CAT TTC AGT GTC AAC AAA AAC TTG CAG TCA TTA CAC ATG CTA TTG
E   V   L   Q   K   H   F   S   V   N   K   N   L   Q   S   L   H   M   L   L

841/281                                   871/291
ATG CTT GCT TTA AAT CGA GTA ACC GTA TCA CAC CCA TTT ATG ACT TCA GCA GAT CTG ATG
M   L   A   L   N   R   V   T   V   S   H   P   F   M   T   S   A   D   L   M

901/301                                   931/311
GAG GCA CAG CAT ATG TGT AGC TTG GAT TCT AAG GCG AAT ATT GTA CAT GGT CTG TCA GTC
E   A   Q   H   M   C   S   L   D   S   K   A   N   I   V   H   G   L   S   V

961/321                                   991/331
TTG GAA ATC TGT CTT ATA ATA GCA ATG AAA CAT TTA AAT GAC ATA TAT GAA GAG GAG CCC
L   E   I   C   L   I   I   A   M   K   H   L   N   D   I   Y   E   E   E   P

1021/341                                  1051/351
TTT AAT TTT CAA ATG GTG TAT AAT GAA TTT CAG AAA TTC ATT CAA AGA AAG GCC CAT TCT
F   N   F   Q   M   V   Y   N   E   F   Q   K   F   I   Q   R   K   A   H   S

1111/371                                  1141/381
GTT TAT AAC TTT GAG AAA CCT GTG GTC ATG AAG GCA TTT GAG CAC TTA CAA CAG TTG GAA
V   Y   N   F   E   K   P   V   V   M   K   A   F   E   H   L   Q   Q   L   E

1081/361                                  1171/391
TTA ATA AAA CCC GTG GAA AGA ACT TCA GTA AAT TCT CAG AGA GAA TAC CAG CTA GTG AAA
L   I   K   P   V   E   R   T   S   V   N   S   Q   R   E   Y   Q   L   V   K

1201/401                                  1231/411
CTA CTT TTG GAT AAT ACT CAA ATT ATG AAT GCT CTA CAG AAA TAC TCC AAC TGC CCT ACA
L   L   L   D   N   T   Q   I   M   N   A   L   Q   K   Y   S   N   C   P   T

1261/421                                  1291/431
GAT GTT AGG CAG TGG GCA ACA TCC TCA CTA AGC TGG CTG TGA
D   V   R   Q   W   A   T   S   S   L   S   W   L   *
```

The Orc4p and Orc5p Subunits of the *Xenopus* and Human Origin Recognition Complex Are Related to Orc1p and Cdc6p*

Tamara Tugal‡**, X. Helena Zou-Yang§**, Kimberley Gavin§, Darryl Pappin¶, Benito Canas¶,
Ryuji Kobayashi§, Tim Hunt‡, and Bruce Stillman§‖

From the ‡Imperial Cancer Research Fund, Clare Hall Laboratories, South Mimms, Potters Bar, Herts EN6 3LD,
United Kingdom, the ¶Imperial Cancer Research Fund, 44 Lincoln's Inn Fields, London WC2A 3PX, United Kingdom,
and the §Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724

The location of origins of DNA replication within the *Saccharomyces cerevisiae* genome is primarily determined by the origin recognition complex (ORC) interacting with specific DNA sequences. The analogous situation in vertebrate cells is far less clear, although ORC subunits have been identified in several vertebrate organisms including *Xenopus laevis*. Monoclonal antibodies were raised against *Xenopus* Orc1p and used for single-step immunoaffinity purification of the entire ORC from an egg extract. Six polypeptides (~110, 68, 64, 48, 43, and 27 kDa) copurified with *Xenopus* Orc1p. Protein sequencing also showed the 64-kDa protein to be the previously identified *Xenopus* Orc2p. Microsequencing of the 43- and 48-kDa proteins that copurified with Orc1p and Orc2p led to their identification as the Orc4p and Orc5p subunits, respectively. Peptide sequences from the 43-kDa protein also allowed the isolation of cDNAs encoding the *Xenopus*, mouse, and human *ORC4* subunits. Human *ORC5* was also cloned; its sequence displayed extensive homology to both *Drosophila* and yeast *ORC5*. Surprisingly, comparison of the amino acid sequences of Orc1p, Orc4p, and Orc5p suggests that they are structurally related to each other and to the replication initiation protein, Cdc6p. Finally, we present the sequence of the putative *Xenopus* and human Orc3p.

The initiation of DNA replication in eukaryotes is tightly regulated such that all the DNA is replicated precisely once per cell cycle from a number of discrete replication origins (Refs. 1–3; for reviews, see Refs. 4 and 5). The nature of this regulation, however, is not well understood as only a small fraction of proteins involved in this process have been identified. Moreover, with the exception of budding yeast, the DNA sequences that define the origin of DNA replication on chromosomes remain an almost complete mystery in higher eukaryotes.

Present knowledge of the control over initiation of DNA replication in an eukaryotic cell is largely based on experiments done in the yeast *Saccharomyces cerevisiae*. In this yeast, DNA replication starts from well defined origins, known as the autonomously replicating sequences (6). An essential, bipartite DNA sequence in all *S. cerevisiae* replication origins is recognized by the six-subunit origin recognition complex (ORC),[1] which is itself essential for initiation of DNA replication (7–20). The ORC also functions in the control of gene silencing, a function that is, at least partially, separable from the functions in initiation of DNA replication (9, 10, 15, 21–23). Proteins homologous to individual subunits of the *S. cerevisiae* ORC have been identified recently in other yeast species and from animal and plant cells. Orc1p and Orc2p were found in *Schizosaccharomyces pombe*, *Drosophila*, *Xenopus*, and humans, and Orc5p has been identified in *Drosophila* (24–30). More recently, the human Orc4p has been identified and shown to co-immunoprecipitate with human Orc2p (31). Immunodepletion of ORC inhibited initiation of DNA replication in a *Xenopus* egg extract, suggesting that ORC is essential for initiation of DNA replication in higher eukaryotes (28, 29, 32).

The nature of the DNA sequence specificity for origin function has not been resolved in species other than the budding yeast, but we imagine that characterization of ORC and its interaction with DNA will help clarify this situation. In pursuance of this goal, we sought to identify the protein subunits of ORC from *Xenopus* and human cells. Monoclonal antibodies against *Xenopus* Orc1p were made and used for immunoaffinity purification of the presumed, complete ORC protein complex from *Xenopus* egg extract. The purified *Xenopus* ORC was composed of six subunits with molecular masses of ~110, 68, 64, 48, 43, and 27 kDa. The 110- and 64-kDa proteins were identified previously as the *Xenopus* Orc1p and Orc2p proteins, respectively. We identified the 43-kDa protein as the *Xenopus* homologue of the yeast Orc4p, and the 48-kDa polypeptide as the homologue of Orc5p. Based on this protein sequence information, cDNAs encoding the human Orc4p and Orc5p were also isolated. The sequences reveal a previously unrecognized structural similarity among the ORC subunits.

EXPERIMENTAL PROCEDURES

Raising of Antibodies—Mice were injected with bacterially expressed full-length *Xenopus* Orc1p as described (33). Spleen cells from three best-responding mice were used to generate antibody-producing hybridomas, and antibodies were screened by enzyme-linked immunosorbent assay and immunoprecipitation (33). Two deletional mutants of *Xenopus* Orc1p, ΔN 280 and ΔN 423, were use to map the epitopes of selected antibodies. Antibodies TK1 and TK15 were found to recognize epitopes within the first 280 residues, whereas TK37 and TK47 recognize a protein sequence between residue 423 and the C terminus of *Xenopus* Orc1p.

Polyclonal antibodies against bacterially expressed full-length Xeno-

[1] The abbreviations used are: ORC, origin recognition complex; MES, 4-morpholineethanesulfonic acid; PAGE, polyacrylamide gel electrophoresis; EST, expressed sequence tag; ORF, open reading frame; mAb, monoclonal antibody; CSF, cytostatic factor.

pus Orc1p and Orc2p were raised as described in (33).

Immunoaffinity Purification of Xenopus ORC—Anti-Orc1p mono-clonal antibodies TK1, TK15, TK37, and TK47 were covalently coupled to protein A beads (Bio-Rad) using dimethylpimelimidate dihydrochloride (Sigma) as a coupling agent, essentially as described in Ref. 33. To perform immunoaffinity purification, antibody-coupled beads were incubated with *Xenopus* egg extract. The beads were then washed several times with ice-cold purification buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 0.5% Nonidet P-40, 2.5 mM EDTA, 2.5 mM EGTA, and 0.5 mM dithiothreitol). The purified proteins were released by boiling in Laemmli sample buffer.

Protein Microsequencing—Two procedures, Edman degradation method and mass spectrometry were used to microsequence the isolated *Xenopus* peptides. The first procedure was performed essentially as described in (9). Very briefly, proteins were separated by SDS-PAGE and in-gel digested with lysyl-endopeptidase (*Achromobacter* protease I; Wako Chemical). Peptide fragments were subsequently extracted and separated on a Hewlett-Packard HPLC model 1090 using a Vydac C18 column. Separated peptides were sequenced with an automated protein sequencer (Applied Biosystems model 494).

In the second procedure, the proteins were electroblotted and digested with trypsin (Boehringer Mannheim), and the peptides were extracted with 1:1 v/v formic acid:ethanol (34). Aliquots were analyzed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry using a LaserMat 2000 mass spectrometer (Thermo Bioanalysis) (34). Additional aliquots were quantitatively esterified using 1% v/v thionyl chloride in methanol and also analyzed by matrix-assisted laser desorption ionization to provide acidic residue composition (35). Native and esterified peptide masses were then screened against the MOWSE peptide mass fingerprint data base (36). The remaining digested peptides (90% of total digest) were also reacted with *N*-succinimidyl-2(3-pyridyl) acetate to enhance b-ion abundance and facilitate sequence analysis by tandem mass spectrometry (37). Dried peptide fractions were treated with 1% w/v *N*-succinimidyl-2(3-pyridyl) acetate in 0.5 M HEPES (pH 7.8 with NaOH) containing 15% v/v acetonitrile on ice. The reaction was terminated by the addition of hepta-fluorobutyric acid and an equal volume of water. The solution was then injected onto a capillary reverse-phase column (300 μm × 15 cm) packed with POROS R2/H material (Perseptive Biosystems) equilibrated with 2% v/v acetonitrile/0.05% v/v trifluoroacetic acid running at 3 μl/min. After washing with 10% v/v acetonitrile/0.05% v/v trifluoroacetic acid, peptides were eluted with a single-step gradient to 75% v/v acetonitrile/0.1% v/v formic acid and collected in a single 3–4-μl fraction. The derivatized peptides were sequenced by low-energy collision-activated dissociation using a Finnigan MAT TSQ7000 fitted with a nanoelectrospray source (38, 39). Collision-activated dissociation was performed using 2.5 mTorr argon with collisional offset voltages between −18V and −28V. The production spectra were collected with Q3 scanned at 500 amu/sec.

Isolation of cDNA Clones Encoding Mouse, Human, and Xenopus ORC4—A search of the dbEST data base using the BLAST or PSI-BLAST programs (40) with either the *Drosophila* Orc4p sequence (DmOrc4p)[2] or the *S. cerevisiae* Orc4p sequence (ScOrc4p) (14), a partial human cDNA clone (accession no. W23942). W23942 was obtained from ATCC, and sequence analysis performed using an ABI automated sequencer (Applied Biosystems model 373). Nested polymerase chain reactions to map the 5′-end were carried out on a human teratocarcinoma cDNA library kindly provided by Dr. Jacek Skowronski at the Cold Spring Harbor Laboratory. A full-length clone was assembled from the polymerase chain reaction fragment and the EST cDNA clone. Two mouse ESTs, AA168456 and AA110785, were also identified in data base searches using ScOrc4p, DmOrc4p, or peptide MS4–1 from the 43-kDa band as queries. Sequencing analyses revealed a single complete open reading frame (ORF) with significant homology to *ORC4* was shared by these two EST cDNA clones, henceforth called *MmORC4*. A 473-base pair fragment to *MmORC4* was also used to identify *Xenopus* ORC4 using a *Xenopus* ovary cDNA library (Strategene), a generous gift from Dr. Eva Dworkin at Vienna University). Screening of approximately 60,000 phage plaques was done essentially as described (41). Restriction endonuclease analysis of the two isolated clones revealed that they were derived from the same cDNA. The larger insert (1.6 kilobases) was then sequenced on both strands with Sequenase, version 2.0 (Amersham Pharmacia Biotech).

Isolation of cDNA Clones Encoding Human ORC5—The dbEST nucleotide sequence data base was searched using the BLAST or PSI-BLAST (40) programs for sequences related to the peptide sequences

---

[2] I. Chesnokov and M. Botchan, personal communication.
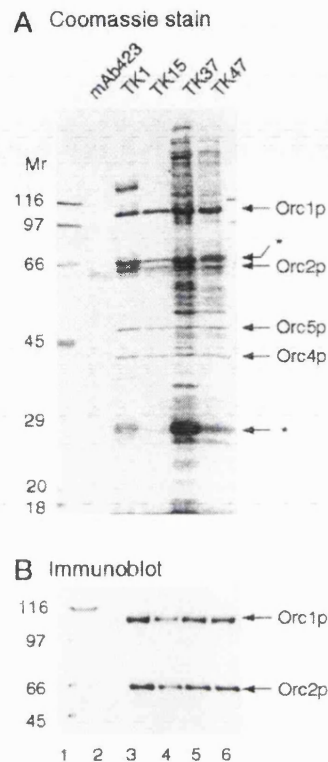
FIG. 1. **Immunoaffinity purification of Xenopus ORC.** *A,* interphase extract from *Xenopus* eggs was incubated with protein A beads coupled to either control anti-SV40 T-antigen mAb 423 (*lane 2*) or anti-Orc1p mAbs TK1, TK15, TK37, and TK47 (*lanes 3–6*). The bound proteins were analyzed by SDS-PAGE and detected by Coomassie Brilliant Blue staining. Polypeptides found in all four samples obtained with anti-Orc1p antibodies (*lanes 3–6*) but not in the control sample (*lane 2*) are marked with *arrows*. *B,* a parallel gel was immunoblotted with anti-Orc1p and Orc2p polyclonal antibodies. *Lane 1* in both *A* and *B,* marker proteins.

obtained from the *Xenopus* Orc5p (48-kDa band) or to the *S. cerevisiae* Orc5p (15). One partial human EST cDNA clone from Jurkat cells (accession no. AA353934; ATCC) was identified and then sequenced completely. A single nucleotide discrepancy (A162G, numbering from the start of AA353934) was observed between AA353934 and the polymerase chain reaction products derived from amplifying related cDNAs from a teratocarcinoma cDNA library using cloned *Pfu* polymerase (Strategene). The most 5′-terminal ATG codon was tentatively assigned as the initiation codon, the context of which matched with the Kozak consensus for eukaryotic translation initiation (42).

Xenopus ORC1 Phosphorylation—CSF-arrested extracts were activated by the addition of $CaCl_2$ to a final concentration of 0.4 mM. Histone HI-kinase assay was done as described (43). Phosphatase sensitivity was tested by adding 0.4 units of potato acid phosphatase (Boehringer Mannheim) in the presence of phosphatase buffer (20 mM MES, pH 6.5, 100 mM NaCl, 1 mM $MgCl_2$ and 1 mM phenylmethylsulfonyl fluoride). In the control sample, phosphatase was inhibited by addition of 50 mM NaF.

Data Base Accession Numbers—The DNA sequences reported in this paper were submitted to either the EMBL or GenBank™ data bases, and the accession numbers are as follows: human Orc4p, AF047598; mouse Orc4p, Y16386; *Xenopus* Orc4p, Y16385; human Orc5p, AF047599.

RESULTS

Immunopurification of the Xenopus Origin Recognition Complex—To purify *Xenopus* ORC, mouse monoclonal antibodies were raised against the recombinant Orc1p subunit (29). Antibodies were selected on the basis of their ability to immunoprecipitate Orc1p from a *Xenopus* egg extract under native conditions. We mapped the epitopes of antibodies TK1 and

TABLE I

Microsequencing of the Xenopus ORC

Peptide	Method	Peptide sequence[a]	Corresponding gene products
K28	Edman	LTDPHQAYVQILK	X1Orc1p, LTDPHQAYVQILK (aa 599–611)[b]
K16	Edman	NEFISTT	X1Orc2p, NEFISTT (aa 151–157)
K32	Edman	VQFTPDEEVLEHI	X1Orc2p, VQFTPDEEVLEHI (aa 14–26)
K35	Edman	LLDQTPSAFADELHRISK	X1Orc2p, LLDQTPSAFADELHRISK (aa 256–273)
K37	Edman	NTASNLVEEYFEAHSSSK	X1Orc2p, NTASNLVEEYFEAHSSSK (aa 215–232)
K32–5	Edman	LSLPASFPDsqFfEk	X1Orc4p, LSLPASFPDSQFAEK (aa 232–246)
K35–5	Edman	LQDIYxGexf	X1Orc4p, LQDIYDGEPF (aa 333–342)
K24	Edman	xLVELxK	X1Orc4p, HLVELLK (aa 44–50)
K33	Edman	aLEDIFG	X1Orc4p, ALEDIFG (aa 78–84)
MS4–1	MS	(K/R)TV(L/I)HGESNSA(L/I)(L/I)(L/I)GPR	X1Orc4p, RTVIHGESNSALIIGPR (aa 51–67)
MS4–2	MS	(K/R)Q(L/I)ED(L/I)(F/M)G(F/M)K	X1Orc4p, KGALEDIFGMK (aa 76–86)
K12	Edman	NIEPhLK	Putative X1Orc5p, Hs: NIEPHLK (aa 256–262)
K21	Edman	rFFLK	Putative X1Orc5p, Hs: RFFLK (aa 326–330)
K24–4	Edman	LSAHAHVELPYYSK	Putative X1Orc5p, Hs: LSAHTHVELPYYSK (aa 294–307)
MS5–1	MS	(K/R)ESQVST(L/I)(L/I)S(L/I)(F/M)GER	Putative X1Orc5p, Hs: RESQVSILQSLFGER (aa 11–25)
MS5–2	MS	(K/R)TY(L/I)(L/I)QTV(L/I)R	Putative X1Orc5p, Hs: KTYVTQTLLK (aa 43–52)
MS5–3	MS	(K/R)TVL(F/M)D(L/I)VR	Putative X1Orc5p, Hs: RTVNFDIIK (aa 421–429)
MS5–4	MS	(K/R)GHTVS(L/I)N(F/M)(L/I)R	Putative X1Orc5p, Hs: KCTVSLDFIR (aa 408–417)

[a] Lowercase letters indicate amino acids identified with lower certainty.

[b] aa, amino acids; MS, mass spectrometry.

TK15 to the N terminus of Orc1p (residues 1–280) and those of antibodies TK37 and TK47 to the C-terminal half of Orc1p (residues 423–886). These antibodies, together with a control antibody against SV40 large T-antigen (mAb423), were individually coupled to protein A-Sepharose beads. Beads were then incubated with an interphase *Xenopus* egg extract (44), and the bound proteins were analyzed by SDS-PAGE. Fig. 1 shows that all four anti-Orc1p mAbs selectively immunoprecipitated a protein corresponding to the size of Orc1p, easily visible by Coomassie Brilliant Blue staining, whereas the control antibody did not (Fig. 1A).

Five additional polypeptides also reproducibly associated with all four anti-Orc1p antibodies but not with the control antibody (Fig. 1A, *arrows*). Immunoblotting the same samples with an anti-Orc1p polyclonal antibody confirmed that the 110-kDa polypeptide was indeed *Xenopus* Orc1p and an anti-Orc2p polyclonal antibody identified the 64-kDa polypeptide as *Xenopus* Orc2p (Fig. 1B, *lanes 3–6*). It seemed likely, therefore, that the four proteins that copurified with the *Xenopus* Orc1p and Orc2p were components of *Xenopus* ORC. It is striking that the molecular masses of the ORC subunits purified from *Xenopus* eggs (~110, 68, 64, 48, 43, and 27 kDa) corresponded closely to those of the ORC purified from *Drosophila* embryos (~115, 82, 79, 47, 42, and 30 kDa), whereas the ORC identified in yeast was somewhat different (120, 72, 62, 56, 53, and 50 kDa) (7, 14, 25).

Identification of the Xenopus, Mouse and Human Orc4p Subunit—Polypeptides that bound specifically to anti-ORC1 antibody TK15 were analyzed by mass spectrometry and the Edman degradation method (Table I). Both methods further confirmed that the 110- and 64-kDa proteins were the previously identified *Xenopus* Orc1p and Orc2p, respectively (29, 32). In addition, a tryptic peptide derived from the 43-kDa band, (R/K)TV(I/L)HGESNSA(I/L)(I/L)(I/L)GPR, was similar to the sequence RTAIYGESNSVLIVGPR translated from a *Mus musculus* EST, AA110785. Interestingly, an independent search of the dbEST data base using both ScOrc4p (14) and the *Drosophila* Orc4p[2] as queries identified another mouse EST, AA168456, as well as a human EST, W23942. Both AA110785 and AA168456 were sequenced completely and were found to contain an identical ORF. This ORF is capable of encoding 433 amino acids closely related to the other *ORC4* sequences and is therefore likely to be the mouse homologue. W23942 was also sequenced completely and was found to be lacking the 5'-end. Isolation of the 5'-end using nested polymerase chain reaction

led to the assembly of a full-length cDNA, the coding sequence of which was highly related to the yeast, fly, frog (see below and Fig. 2), and mouse Orc4p sequences.

A 473-base pair fragment from the mouse ORC4 cDNA clone was also used to screen a λ-ZAP *Xenopus* ovary cDNA library. Two identical clones containing a 1.3-kilobase ORF were identified. Importantly, more microsequencing of the 43-kDa polypeptide by both Edman degradation and mass spectrometry subsequently identified peptides (Table I) that showed close matches to the predicted sequence from the cDNA (Fig. 2).

Several observations can be made upon aligning multiple Orc4p sequences from human, mouse, frog, and yeast (Fig. 2). First of all, all these sequences share similar lengths and are highly related to each other. A BLAST search comparing HsOrc4p with other sequences in GenBank showed that HsOrc4p is most related to MmOrc4p and XlOrc4p (P = 3.1e-251 for MmOrc4p, and P = 2.0e-204 for XlOrc4p). Not surprisingly, Orc4p from these intimately related organisms (human, mouse, and frog) share higher levels of homology (compare with P = 3.1e-24 for ScOrc4p). Second, multiple blocks of conserved residues can be readily identified. One intriguing feature is a strong match to the nucleotide binding site consisting of both the Walker A and B motifs (45). This bipartite putative nucleotide binding site appears to be well conserved among human, mouse, and frog, whereas only a weak match to the Walker A motif, but not to the B motif, is found in ScOrc4p (Fig. 2). It has been postulated that the Walker A motif is involved in ATP/GTP binding, whereas the Walker B motif is involved in ATP hydrolysis (45). In agreement with this, ScOrc4p has been shown to cross-link ATP in the presence of autonomously replicating sequence 1 and ScOrc1p (46). However, mutations that removed this motif in ScOrc4p showed no gross phenotypic defect. Furthermore, the cross-linking to ATP was completely abrogated when the Walker A motif in Orc1p was mutated, indicating that the Orc4p-ATP binding might have been indirect or that it at least requires a functional ScOrc1p. It is likely that the nucleotide binding motif is nonfunctional in yeast but fully functional in higher eukaryotes due to selection pressure. Further biochemical analysis is necessary to address this question.

Cloning of Human ORC5—A data base search identified a human EST clone, AA353934, as a potential homologue of the *Drosophila* Orc5p (DmOrc5p). This cDNA was sequenced entirely on both strands, and a conceptual translation with the most 5'-terminal ATG as tentative initiation codon revealed an

ORC Protein Subunits in Vertebrates

```
HsOrc4p                                                         MSSR KSKSNSEIHT
MmOrc4p                                                         MSSR KTKSNA..HA
XlOrc4p                                                         MSKR KSKDLCEPVG
ScOrc4p    MTISEARLSP QVNLLPIKRH SNEEVEETAA ILKKRTIDNE KCKDSDPGFG

HsOrc4p    ECLSQVQRIL RERRCRQSPH SNLFGVQVQ. .YKHLSELLK RTALHGESNS
MmOrc4p    ECLSQVQRIL RERRCHHSPH SNLFGVQVQ. .YKHLIELLK RTALYGESNS
XlOrc4p    ECISKVHGIL RQRLFQH..H GKPFGVDSQ. .HKHLVELLK RTVIHGESNS
ScOrc4p    ...SLQRRLL QQLYGTLPTD EKIIFTYLQD CQQEIDRIIK QSTIQKESHS
                        Walker A

HsOrc4p    VLLIGPRGSG KTMLISHALK ELMELEEVSE NVLQVHLNGL LQINDKIALK
MmOrc4p    VLIVGPRGSG KTTLLNHALK ELMEL.EVSE NVIQVHLNGL LQTNEKIALK
XlOrc4p    ALLIGPRGSG KSMLLKGALE DIFGMKQMKE TALQVNLNGL LQTTDKIALK
ScOrc4p    MILVGPRQSY KTYLLDYEL. SLLQ.QSYKE QFITIRLNGF IH.SEQTAIN

HsOrc4p    EITRQL..NL ENVVG..DKV .........E GSFA...ENL SFLLEALKK.
MmOrc4p    EITRQL..NL DNVVE..DKV .........E GSFA...ENL SFLLEALKK.
XlOrc4p    EITRQL..HL ENVVG..DRV .........E GSFA...ENL SFLLEALKT.
ScOrc4p    GKATQLEQQL QKIHGSEEKI DDTSLETISS GSLTEVFEKI LLELDSTTKT
                                       Walker B

HsOrc4p    .....G..DR TS..SCPVIF LLDEFDLFAH HKNQTLLYNL FDISQSAQTP
MmOrc4p    .....G..DR TS..SCPVIF LLDEFDIFAH QKNQTLLYNL FDISQSAQTP
XlOrc4p    .....G..DR KS..SCPVLF VLDEFDLFAH HKNQTLLYNL FDIACSAQTP
ScOrc4p    PNEDSGEVDR ESITKITVVF IFDEIDTEAG PVRQTLLYNL EDMVEHSRVP

HsOrc4p    IAVIGLTCRL DILELLEKRV KSRFSHRQIH LMNSFGFPQY VKIEKEQLSL
MmOrc4p    VAVIGLTCRL DILELLEKRV KSRFSHRQIH LMNSFDFPQY LKFFKEQLSL
XlOrc4p    VAVIGLTCRL DVMELLEKRV KSRFSHRQIH LLNSESESOY LQLLQEKLSL
ScOrc4p    MCIFGCFTKL NILEYLEKRV KSRFSQRVLY MPQIQNLDDM VDAVRNLETV

HsOrc4p    PAEFPDKVFA EKWNENV.QY LSEDRSVQEV LQKHFNIS.K NLRSLHMLLM
MmOrc4p    PAEFPDKAFA ERWNENV.HC LSEDSTVLEV LQKHFSVN.K NLQSLHMLLM
XlOrc4p    PASFPDSQFA EKWNESI.KS LVESKLVEDV LQKQYNAS.K DVRSLHMLML
ScOrc4p    RSEI..SPWV SQWNETLEKE LSDPESS...N LNRHIRMNFE TFRSLPTKKN

HsOrc4p    LALNRVTASH PF...MTAVD LMEASQLCSM DSKANIVHG. ...LSVLEIC
MmOrc4p    LALNRVIVSH PF...MTSAD LMEAQHMCSL DSKANIVHG. ...LSVLEIC
XlOrc4p    LAVCRVNVSH PH...ITAAD FLEVFRIRNQ DSKANILHG. ...VSVLELC
ScOrc4p    SIIPLVATSK NEGSLCTAIK SCSFLDIYNK NQLSNNLTGR LQSLSDLELA

HsOrc4p    LIIAMKHLND IYBEEPFNFQ MVYNEFQKFV Q...RKAHSV ...YNFEKPV
MmOrc4p    LIIAMKHLND IYBEEPFNFQ MVYNEFQKFI Q...RKAHSV ...YNFEKPV
XlOrc4p    LIIAMKHLQD IYDGEPFNFQ MVHNEFQKFI Q...RKAHSV ...YNFEKAV
ScOrc4p    ILISAARVAL RAKDGSFNEN LAYAEYEKML KAINSRIPTV APTTNVGTGQ

HsOrc4p    VMKAFEHLQQ LELIKPMERT SGN.SQREYQ LMKL...LLI NTQIMNALQK
MmOrc4p    VMKAFEHLQQ LELIKPVERT SVN.SQREYQ LVKL...LLI NTQIMNALQK
XlOrc4p    VTKAFEHLHQ LELIKPMEGL SVR.TQKEVR LMKL...LLI NTQIVEALQK
ScOrc4p    STFSIDNTIK LWLKKDVKNV WENLVQLDFF TEKSAVGLRD NATAAFYASN

HsOrc4p    Y....PNCPT DVRQWATS.S LSWL*
MmOrc4p    Y....SNCPT DVRQWATS.S LSWL*
XlOrc4p    Y....PNCPT DVKQWAMS.S LS*
ScOrc4p    YQFQGTMIPF DLRSYQMQII SQELRRIIPK SNMYYSWTQL*
```

FIG. 2. **Sequence comparison among *H. Sapiens*, *M. musculus*, *X. laevis*, and *S. cerevisiae* Orc4p performed with the GCG Pileup program.** Identical residues are indicated by *dark boxes*, and similar residues are indicated by *light boxes*. The putative P-loop and A-loop are marked (*Walker A* and *Walker B*).

ORF strikingly similar to both DmOrc5p (25) and the *S. cerevisiae* Orc5p (ScOrc5p; 15); this ORF will henceforth be referred to as HsOrc5p. Others have also identified this same EST clone encoding HsOrc5p (47). The predicted protein consists of 435 amino acids with a calculated molecular mass of ~50 kDa. Overall, HsOrc5p is approximately 39% identical and 50% similar to DmOrc5p (Fig. 3). It is also 29% identical and 41% similar to ScOrc5p, with the sequence similarity most evident in the C-terminal half of the proteins. Nonetheless, a search of the PROSITE data base failed to identify any recognizable motif in this region. In the N terminus, all three proteins contain a putative Walker A motif, implicating a role of Orc5p in ATP binding. Indeed, ScOrc5p, when complexed with the other ORC subunits, displayed a DNA-independent ATP binding activity that was sensitive to mutations within the consensus (46).

Additional evidence supporting a role for this gene product

as the human homologue of Orc5p was obtained from microsequencing the 48-kDa band that copurified with XlOrc1p and XlOrc2p. Several peptides were generated from this band (Table I; Fig. 3), all of which displayed extensive similarities to the deduced sequence of HsOrc5p, as well as DmOrc5p and ScOrc5p. Thus, the 48-kDa polypeptide that co-immunoprecipitated with XlOrc1p was most likely to be XlOrc5p.

In a survey of primary human tissues using HsOrc5 mRNA as a probe, a single 1.9-kilobase transcript, was detected by Northern blotting in most tissues (data not shown). The size of the HsOrc5 transcript detected in this experiment suggested that the isolated cDNA was likely to be full-length. Intriguingly, *HsORC5* mRNA was particularly abundant in testis (data not shown).

A bacterial artificial chromosome (BAC) clone (RG126M09) that contained the last 700 base pairs of HsOrc5 was recently deposited in the data base. This information suggested that the

FIG. 3. **Sequence alignment among H. Sapiens, D. melanogaster, and S. cerevisiae Orc5p using the GCG Pileup program.** Identical residues are indicated by *dark boxes*, and similar residues are indicated by *light boxes*. The putative P-loop and A-loop are underlined by a *black line* and *two arrows*. HsOrc5p and DmOrc5p possess only imperfect matches to the A-loop consensus. Peptides that are highly related between the *Xenopus* 48-kDa band and HsOrc5p are indicated by *black lines*.

```
                1                                                    50
HsOrc5p     MPHLEN VVLCRESQVS ILQSLFGERH HFSFPS.IFI YGHTASGKTY
DmOrc5p  MEAICSSLEP LFPCREAAIE TLGELIGDSS E.TYPSAIYL FGHSGTGKTA
ScOrc5p     MNVTTP EVAFREYQTN CLASYISADP DIT.PSNLIL QGYSGTGKTY
            51                                                   100
HsOrc5p  VTQTLLKTL. ELPHV...FV NCVECFTLRL LLEQI..... ..LNKLNHLS
DmOrc5p  LTRAFLKECG KRQNVRTAHL NAIECYTTKI MDEIL..... ..LDSL....
ScOrc5p  ...TLKKYFN ANPNLHAVWL EPVELVSWKP LLQAIARTVQ YKLKTLYPNI
            101                                                  150
HsOrc5p  SSEDGCSTEI TCETFNDFV. RLFKQVTTAE NLKDQT.VYI VLDKAEYLRD
DmOrc5p  APDQGDALKV ..DNMLDFVE QLRRQAAT.. RVEDQG.FLI AVDNAERLRD
ScOrc5p  PTTDYDPLQV E.EPFL.LVK TLHNIFVQYE SLQEKTCLFL ILDGFDSLQD
            151                                                  200
HsOrc5p  MEANLLPGFL RLQEL..ADR NVIVLFLSEI VWEKFRPNTG CFEPFVLYFP
DmOrc5p  MDANVLPVLL RLQEL..TNL NLCVILLSQL PFEKFYNKTG LSEIVCLHLA
ScOrc5p  LDAALFNKYI KLNELLPKDS KINIKFIYTM LETSFLQRYS THCIPTVMFP
            201                                                  250
HsOrc5p  DYSIGNLQKI LSHDHPP... ...EYSA... .......... ..D.FYA...
DmOrc5p  QYNKAETQRI LGSDFQQVRN QLLEQFAQDK KRLEICQEAV TED.FYN...
ScOrc5p  RYNVDEVSTI L...VMSRCG ELMEDSCLRK RIIEEQITDC TDDQFQNVAA
            251                                                  300
HsOrc5p  AYINILLGVF YTVC.RDLKE LRHLAVLNFP KYCEPVVKGE ASERDTRKLW
DmOrc5p  NYLNLFLSVF YKAC.RDVPE LQLTARKCLS TYLEPVLDGT VDATDISRLW
ScOrc5p  NFIHLIVQAF HSYTGNDIFA LNDLIDFKWP KYVSRITK.. ......ENIF
            301                                                  350
HsOrc5p  RNIEPHLKKA MQTVYLREIS SSQWEKLQKD DT..DPGQLK GLSAHTHVEL
DmOrc5p  RHIAGPLRSA LTQIYMR.IE KPA.EEVE.D FTAIEDQSVR KL.AQS.LEL
ScOrc5p  EPLA..LYKS AIKLFL.STD DNLSENGQGE .SAITTNRDD LENSQTY.DL
            351                                                  400
HsOrc5p  PYYSKFILIA AYLASYNPAR TDKRFFLKHH GKIKK....T NF.LKKHEKT
DmOrc5p  PYYAKFLLIA AFLASHNAAK QDKRLFVKHH GKQRKRMQTV NARAKTTEKM
ScOrc5p  SIISKYLLIA SYICSYLEPR YDASIFSR.. .KTRIIQGRA AYGRRKKKEV
            401                                                  450
HsOrc5p  SNHLLGPKPF PLDRLLAILY SI........ ......VDSR VAPTANI..F
DmOrc5p  ST.TLGPKSF SIDRLLAIFY AI........ ......LEEK VGLTCNL..L
ScOrc5p  NPRYLQPSLF AIFRLLAIFQ AIFPIQGKAE SGSLSALREE SLMKANIEVF
            451                                                  500
HsOrc5p  SQITSLVTLQ LLTLV.GHD. DQLDG.PKYK CTVSLDFIRA IARTVNFDII
DmOrc5p  SQISTLVHLN LLSFVSGEQ. NIMEGSARLQ CTIGLEFVLQ IGKVVGFNVR
ScOrc5p  QNLSELHTLK LIATTMNKNI DYLSPKVRWK VNVPWEIIKE ISESVHFNIS
            501
HsOrc5p  KYLYDFL*
DmOrc5p  QYLCDFM*
ScOrc5p  DYFSDIHE*
```

genomic copy of HsORC5 is located on chromosome 7, between 7q21 and 7q22 and direct mapping has confirmed this location (47). Intriguingly, genes encoding human Mcm7p, as well as histone H2A and H2B are also found near this locus.

Characterization of Xenopus Orc4p in Xenopus Eggs and Cultured Cells—Full-length *Xenopus* Orc4p was expressed in bacteria and used to produce a rabbit polyclonal antibody that reacted with a 43-kDa protein in both crude *Xenopus* egg extract and the immunopurified *Xenopus* ORC (Fig. 4A, lanes 1 and 2). Quantitative immunoblotting was employed to estimate the concentration of Orc4p in the *Xenopus* egg extract, which gave a value of approximately 35 nM (data not shown). This is in reasonable agreement with previous published estimates for Orc1p and Orc2p, and it suggested that *Xenopus* Orc1p, Orc2p, and Orc4p were present in the egg extract in roughly equimolar amounts. The protein levels of ORC subunits in the egg extract was also compared with those present in WAK cells, an established *Xenopus* cell line. Approximately the same concentrations and ratio of Orc1p, Orc2p, and Orc4p were present in 10^5 exponentially growing cells as in a single egg ($\sim$0.5 μl of extract) (Fig. 4B). To determine whether Orc1p, Orc2p, and Orc4p also formed a complex in these somatic cells, we performed the same immunoprecipitation using anti-Orc1p mAb TK15 and anti-T antigen mAb423 as a control. TK15 specifically immunoprecipitated Orc1p, Orc2p, and Orc4p from the somatic cell extract (Fig. 4C), arguing that a very similar ORC is also formed in *Xenopus* somatic cells.

XlOrc1p and XlOrc2p Show Altered Mobility in M-phase Extracts—Although we did not detect any differences in the molecular masses of either Orc1p, Orc2p, or Orc4p between

FIG. 4. **ORC in Xenopus egg and somatic cell extracts.** A, anti-Orc4p antibody. Interphase egg extract (*lane 1*) and an aliquot of purified ORC (*lane 2*) were separated by SDS-PAGE and immunoblotted with a rabbit polyclonal anti-Orc4p antibody that recognized the 43-kDa band in both samples. B, relative amounts of ORC subunits in *Xenopus*. 0.05 μl of CSF-extract (*lane 3*), 0.05 μl of interphase extract (*lane 4*), and a sample of 10^4 WAK cells (*lane 5*) were separated by SDS-PAGE and immunoblotted with either anti-Orc1p, Orc2p, or Orc4p antibodies. C, immunoprecipitation of ORC from WAK cells. WAK cell lysate was incubated with protein A beads coupled to either control anti-SV40 T-antigen mAb 423 (*lane 6*) or anti-Orc1p mAb TK15 (*lane 7*). The bound proteins were analyzed by SDS-PAGE and immunoblotted with anti-Orc1p, anti-Orc2p, and anti-Orc4p polyclonal antibodies.

A

+ CaCl$_2$ | · CaCl$_2$

0 15 30 45 60 | 0 15 30 45 60 Minutes

← Orc1p

B

← Histone
H1

1 2 3 4 5 6 7 8 9 10

C

Buffer only   Phosphatase   Phosphatase + NaF
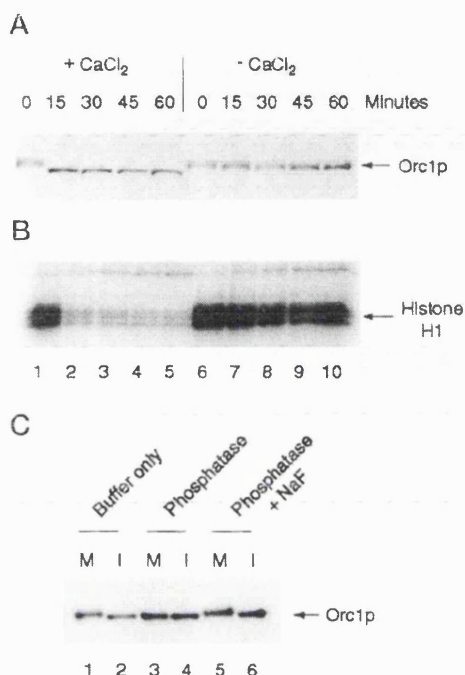
M I   M I   M I

← Orc1p

1 2 3 4 5 6

FIG. 5. **Cell cycle-dependent Orc1p modification.** *A*, CSF (meiotic) egg extract was incubated with CaCl$_2$ for increasing lengths of time to trigger the transition to interphase *(lanes 1–5)*. After a 15-min incubation, Orc1p appeared on polyacrylamide gels as a faster migrating band *(lane 2)*. There was, however, no change in the Orc1p mobility in samples to which no Ca^{2+} was added, and therefore, the cells remained in the meiotic phase *(lanes 6–10)*. *B*, the activity of mitotic kinases in samples incubated with Ca^{2+} as well as in the control samples was tested by their ability to phosphorylate histone HI. There was a tight correlation of the slow migrating form of Orc1p and the meiotic extract *(lanes 1 and 6–10)*. Conversely, the faster migrating form of Orc1p was found only in the meiotically inactive (interphase) samples *(lanes 2–5)*. *C*, both meiotic CSF extract (labeled *M*; *lanes 1, 3, and 5*) and Ca^{2+}-treated interphase extract (labeled *I*; *lanes 2, 4, and 6*) were treated with buffer alone *(lanes 1 and 2)*, with potato acid phosphatase *(lanes 3 and 4)*, or with phosphatase and 50 mM NaF *(lanes 5 and 6)*.

eggs and somatic cells (Fig. 4*B*), there was a slight mobility shift of Orc1p and Orc2p, but not of Orc4p, when the interphase extract was compared with the CSF-arrested metaphase extract (Figs. 4*B*). A similar modification of Orc2p mobility was observed previously (32). Analysis of the Orc1p modification revealed that the slow migrating form of the protein present in CSF-arrested extracts (Fig. 5*A, lanes 1* and *6–10*) was converted to a faster migrating form upon Ca^{2+}-induced transition to interphase (Fig. 5, *lanes 2–5*). The mobility shift of XlOrc1p is likely due to phosphorylation/dephosphorylation for the following reasons: in a time-course experiment, the shift to a faster migrating form occurred concomitantly with the inactivation of mitotic kinases (Fig. 5*B*). Furthermore, treatment of CSF-arrested extracts with potato acid phosphatase could also convert the slow migrating form of Orc1p to the faster one, and this conversion was blocked by 50 mM NaF, a potent phosphotase inhibitor (Fig. 5*C*).

DISCUSSION

The origin recognition complex was first discovered as a six-polypeptide assembly that binds *S. cerevisiae* origins of DNA replication and governs initiation of DNA replication at these sites (4, 5, 7). Since then, similar multisubunit complexes have been purified in *Drosophila* and *Xenopus*. The *Drosophila* ORC contains six polypeptides homologous to the six subunits

of ScORC (25, 30). An ORC-resembling complex was also purified by biochemical means, which contained at least XlOrc1p and XlOrc2p (29). These and other biochemical and genetic studies in *Drosophila* and *Xenopus* demonstrate that the ORC functions in chromosomal DNA replication in multicellular eukaryotes, just as it does in yeast (25, 28–30, 48, 49). In this report, we describe an immunopurification scheme of *Xenopus* ORC from interphase egg extract and the cloning and characterization of some of the genes encoding *Xenopus* and human ORC subunits. Six polypeptides, with apparent molecular masses ranging from 27 to 110 kDa, were retained reproducibly by a panel of monoclonal antibodies against XlOrc1p. The 110- and 64-kDa polypeptides in the purified complex were XlOrc1p and XlOrc2p (Fig. 1), whereas the 48-kDa subunit was most likely to be the XlOrc5p subunit, as it contained peptide sequences highly related to both DmOrc5p and the newly identified HsOrc5p (Table I; Fig. 3). We further showed that the 43-kDa subunit was Orc4p, the fifth largest subunit of *Xenopus* ORC. XlOrc4p associated stoichiometrically with XlOrc1p and XlOrc2p in both egg and somatic cell extracts.

Although ORC has been purified in only a handful of species as mentioned previously, individual *ORC1, ORC2, ORC4,* and *ORC5* subunits have been identified from a variety of organisms including *S. pombe, Kluyveromyces lactis, Arabidopsis thaliana,* and human cells (24, 26, 27, 31, 47, 50–52, and this report). This argues strongly that virtually all ORC subunits are conserved throughout evolution. It also raises the question of whether these subunits assemble into a similar higher-order complex as they do in yeast, fly, and frog. In the case of human cells, immunoprecipitation experiments strongly implicated that at least HsOrc1p, HsOrc2p, and HsOrc4p were among one complex (24, 31).[3]

In this report, we also described the comparison of ORC from frog egg extract with that from somatic cells in an attempt to explain why previous studies came up with conflicting conclusions about the identity of origins of DNA replication in vertebrates (4). We found no difference in the immunoreactivity or protein mobility ofXlOrc1p, XlOrc2p, and XlOrc4p between those present in eggs and those in somatic cells (Fig. 4). In addition, all three subunits co-immunoprecipitated from somatic cell extract as well as from egg extract, implying the presence of the same complex in both situations. The main difference appears to lie in the concentration of ORC, which is 10-fold higher in eggs than in somatic cells. Given the large size of a frog egg, there are approximately 100,000 times more ORC molecules in one egg than in one somatic cell. The ORC-to-origin ratio is therefore much higher in *Xenopus* eggs than it is in somatic cells, at least in the early cleavage cell cycles. Whether this high ORC-to-origin ratio promotes the apparently random initiation of DNA replication in the early *Xenopus* embryo is at present unclear.

The polyacrylamide gel mobility of XlOrc1p changed upon Ca^{2+}-induced transition of CSF-arrested extracts to interphase. The timing of this shift correlated with the inactivation of mitotic kinases monitored by the histone H1 kinase assay, and could be mimicked by phosphatase treatment (Fig. 5). It therefore seems that XlOrc1p is phosphorylated by protein kinase(s) present in M-phase cell extracts, similar to what was previously reported for Orc2p (32). The function of these modifications remains to be elucidated.

It is of interest to notice that a subset of the monoclonal antibodies also immunoprecipitated other cellular proteins. We currently cannot rule out the possibility of cross-reactivity by

---

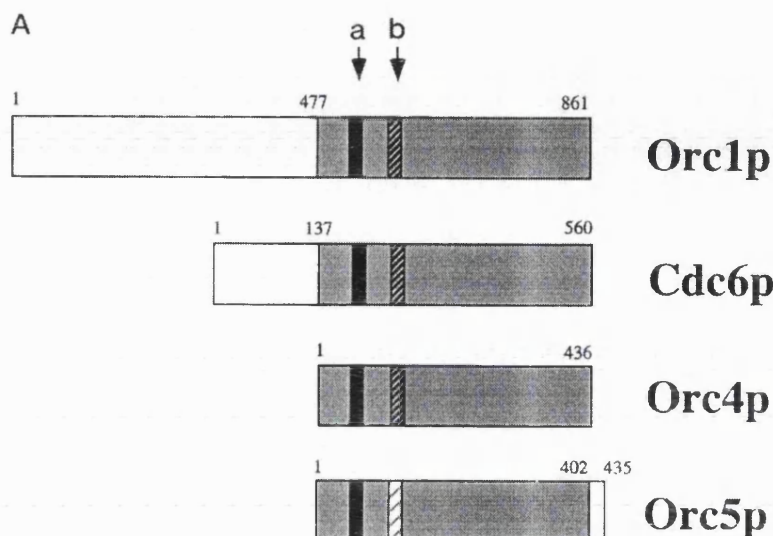[3] H. Yang, K. Simpson, M. Hidaka, and B. Stillman, unpublished results.

FIG. 6. *A*, comparison of Orc5p from *H. sapiens*, Orc4p from *H. sapiens* and *X. laevis*, Cdc6p from *H. sapiens* and *X. laevis*, and Orc1p from *H. sapiens* and *X. laevis*. The amino acid numbers shown *above* each box refer to the sequences from the four human protein sequences. The conserved amino acid blocks are indicated. *B*, multiple sequence alignment of the above sequences using the PIMA (60, 68) program. At any given position, only residues conserved among at least four sequences are *shaded*. Numbers in *parentheses* indicate number of residues in each spacer region.

antibodies. It is equally likely, however, that some of the cellular proteins might represent genuine ORC binding proteins *in vivo*. For example, several high molecular weight proteins have been consistently brought down in the immunoprecipitates using mAbs TK37 and TK47, reminiscent of the 168-kDa protein that copurified with *Xenopus* ORC (29). Originally identified as an origin binding protein required for initiation of DNA replication, ORC has since been implicated in a number of other cellular processes, including the mating-type and telomeric silencing in budding yeast (9–11, 15, 21–23, 53, 54) and heterochromatin assembly in *Drosophila* (30). At this moment,

it is not clear whether ORC is indeed pleiotropic or its impact on other cellular processes is secondary to its primary role in initiation of DNA replication. Identifying cellular proteins that copurified with *Xenopus* ORC, therefore, will undoubtedly help us to answer this question.

The isolation of cDNAs encoding ORC from multiple species has allowed a re-examination of the primary structure of the ORC subunits. It was known that the Orc1p and the Cdc6p proteins shared extensive sequence similarity in a broad region of the proteins, including the ATP binding sites (14). Cdc6p from *S. cerevisiae* and *Xenopus*, and the related *S. pombe*

TABLE II
Microsequencing of the Xenopus 68-kDa protein

Footnotes are the same as in Table I.

Peptide	Peptide sequence	Corresponding amino acids in 81-kDa
K16	tDLYQLQK	TDLYQLQK (aa 526-533)
K40	LEEFLINFQSLEETTQNEED	LEEFLINFQSLEETTQNEED (aa 494-513)
K43	EeFLINFQgle	EEFLINFQSLE (aa 495-505)
K46	LFENLIlFLRE	LFENLIGFLRK (aa 69-79)

FIG. 7. **Alignment of the protein sequences from *S. cerevisiae* Orc3p (Bell et al., 1995), the *Xenopus* 81-kDa protein reported by Dunphy and Carpenter, and the putative human Orc3p.**

cdc18⁺ protein, function with ORC in the initiation of DNA replication (26, 55–59, 61–65). Furthermore, Cdc6p/cdc18⁺ plays a critical role in the control of initiation of DNA replication during the cell cycle, ensuring that replication occurs only once during each cell division (57, 63, 64, 66). Analyses of the new Orc4p and Orc5p sequences using the PSI-BLAST (40) and PROBE (67) algorithms have revealed unexpected similarity between these two proteins and the members of the Orc1p and Cdc6p protein families. For convenience, only the human and *Xenopus* sequences are shown in Fig. 6, but the sequence similarity extends to protein sequences from other species (data not shown).[4]

The C terminus of the Orc1p and Cdc6p shows sequence similarity to the entire protein sequence of Orc4p and nearly entire Orc5p (Fig. 6, *A* and *B*). The Walker A and B motifs of the putative nucleotide binding site are located toward the N terminus of the region of similarity and are conserved in all four

proteins, as discussed above. In addition to these, the conserved region contains other amino acid sequence blocks that are found in each of the three ORC subunits and Cdc6p (Fig. 6), some of which are also present in other ATP binding proteins involved in DNA replication, such as the replication factor C and the δ′ subunit of *Escherichia coli* DNA polymerase III (Fig. 6, first 10 sequence blocks, 67–69). We suggest that these four proteins share significant tertiary structure and that they all evolved from an ancestral protein sequence. Interestingly, these sequences are also related to single amino acid sequences in the genomes of the two archea species *Methanobacterium thermoformicicum* and *Archaeoglobus fulgidus* (data not shown), suggesting that there may have been an ancestral initiator protein that diverged to form the multisubunit ORC.

In modern eukaryotes, three of the six ORC subunits and the Cdc6p are related to each other. Because ORC and Cdc6p exist in a prereplication complex, a cluster of related ATP binding proteins forms at eukaryote origins of DNA replication. In prokaryotes, a single initiator protein, DnaA in *E. coli,* forms a multimeric protein complex at the origin of DNA replication,

---

[4] A. F. Neuwald, L. Aravind, J. L. Spouge, and E. V. Koonin, submitted for publication.

thereby assembling a number of ATP binding proteins that are identical (71). Thus, the similarities between DNA replication in prokaryotes and eukaryotes (72) might extend beyond the proteins that function at the replication fork to the structure of the initiation complexes. It is striking that many ATP-interacting DNA replication proteins, such as replication factor C and the minichromosome maintenance proteins also form complexes containing multiple sequence-related ATP-interacting proteins (5, 72), and now ORC is known to have similar features.

Note Added in Proof—Carpenter and Dunphy recently identified an 81-kDa protein that was associated with the *Xenopus* ORC and reported its sequence. (Carpenter, P. B., and Dunphy, W. G. (1998) *J. Biol. Chem.* **273**, 24891–24897). As part of the studies reported here, we obtained four peptide sequences by the Edman degradation technique from the 68-kDa protein in our *Xenopus* ORC (see *asterisk* in Fig. 1 at 68 kDa). These peptides correspond to sequences within the 81-kDa open reading frame reported by Carpenter and Dunphy (see Table II). In addition, Carpenter and Dunphy kindly provided us with their anti-p81 antiserum, which recognized the polypeptide running at approximately 68 kDa on our gel marked with an *asterisk* in Fig. 1. We therefore conclude that the p81 identified by Carpenter and Dunphy is identical to our 68-kDa protein, the difference in apparent molecular mass being explained by differences in the composition of our polyacrylamide gels. Furthermore, we have identified human expressed sequence tag cDNAs that encode a protein with significant similarity to the *Xenopus* protein. To determine the sequence of the putative human homolog of *S. cerevisiae* Orc3p, three human EST cDNA clones (GenBank™ accession numbers U50950, H17704, and AA262375) were sequenced, and a composite was made (X. H. Zou-Yang and B. Stillman, unpublished results). Theoretical translation of the composite cDNA produced a 698-amino acid open reading frame (ORF) with the initiation codon being absent. This ORF displays significant sequence similarity to both p81 from *Xenopus* and ScOrc3p. An alignment of the human and *Xenopus* proteins with the *S. cerevisiae* Orc3p is shown in Fig. 7. Although the protein sequence alignments show only scattered regions of identity and similarity, we suggest that the human and *Xenopus* proteins are Orc3 subunits of the origin recognition complex.

REFERENCES

1. Cairns, J. (1966) *J. Mol. Biol.* **15**, 372–373
2. Huberman, J. A., and Riggs, A. D. (1966) *Proc. Natl. Acad. Sci. U. S. A.* **55**, 599–606
3. Rao, P. N., and Johnson, R. T. (1970) *Nature* **225**, 159–164
4. Diffley, J. F. X. (1996) *Genes Dev.* **10**, 2819–2830
5. Stillman, B. (1996) *Science* **274**, 1659–1664
6. Newlon, C. S., and Theis, J. F. (1994) *Curr. Opin. Genet. Dev.* **3**, 752–758
7. Bell, S. P., and Stillman, B. (1992) *Nature* **357**, 128–134
8. Marahrens, Y., and Stillman, B. (1992) *Science* **255**, 817–823
9. Bell, S. P., Kobayashi, R., and Stillman, B. (1993) *Science* **262**, 1844–1849
10. Foss, M., McNally, F. J., Laurenson, P., and Rine, J. (1993) *Science* **262**, 1838–1844
11. Micklem, G., Rowley, A., Harwood, J., Nasmyth, K., and Diffley, J. F. X. (1993) *Nature* **366**, 87–89
12. Diffley, J. F. X., Cocker, J. H., Dowell, S. J., and Rowley, A. (1994) *Cell* **78**, 303–316
13. Rao, H., Marahrens, Y., and Stillman, B. (1994) *Mol. Cell. Biol.* **14**, 7643–7651
14. Bell, S. P., Mitchell, J., Leber, J., Kobayashi, R., and Stillman, B. (1995) *Cell* **83**, 563–568
15. Loo, S., Fox, C. A., Rine, J., Kobayashi, R., Stillman, B., and Bell, S. (1995) *Mol. Biol. Cell* **6**, 741–756
16. Rao, H., and Stillman, B. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 2224–2228
17. Rowley, A., Cocker, J. H., Harwood, J., and Diffley, J. F. X. (1995) *EMBO J.* **14**, 2631–2641
18. Hardy, C. F. J. (1996) *Mol. Cell. Biol.* **16**, 1832–1841
19. Hori, Y., Shirahige, K., Obuse, C., Tsurimoto, T., and Yoshikawa, H. (1996) *Mol. Biol. Cell* **7**, 409–418
20. Lee, D. G., and Bell, S. P. (1997) *Mol. Cell. Biol.* **17**, 7159–7168
21. Ehrenhofer-Murray, A. E., Gossen, M., Pak, D. T. S., Botchan, M. R., and Rine, J. (1995) *Science* **270**, 1671–1674
22. Fox, C. A., Loo, S., Dillin, A., and Rine, J. (1995) *Genes Dev.* **9**, 911–924
23. Dillin, A., and Rine, J. (1997) *Genetics* **147**, 1053–1062
24. Gavin, K. A., Hidaka, M., and Stillman, B. (1995) *Science* **270**, 1667–1671
25. Gossen, M., Pak, D. T. S., Hansen, S. K., Acharya, J. K., and Botchan, M. R. (1995) *Science* **270**, 1674–1677
26. Muzi-Falconi, M., Brown, G. W., and Kelly, T. J. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 1566–1570
27. Leatherwood, J., Lopez-Girona, A., and Russell, P. (1996) *Nature* **379**, 360–363
28. Romanowski, P., Madine, M. A., Rowles, A., Blow, J. J., and Laskey, R. A. (1996) *Curr. Biol.* **6**, 1416–1425
29. Rowles, A., Chong, J. P. J., Brown, L., Howell, M., Evan, G. I., and Blow, J. J. (1996) *Cell* **87**, 287–296
30. Pak, D. T. S., Pflumm, M., Chesnokov, I., Huang, D. W., Kellum, R., Marr, J., Romanowski, P., and Botchan, M. R. (1997) *Cell* **91**, 311–323
31. Quintana, D. G., Hou, Z., Thome, K. C., Hendricks, M., Saha, P., and Dutta, A. (1997) *J. Biol. Chem.* **272**, 28247–28251
32. Carpenter, P. B., Meuller, P. R., and Dunphy, W. G. (1996) *Nature* **379**, 357–360
33. Harlow, E., and Lane, D. (1988) *Antibodies: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
34. Sutton, C. W., Pemberton, K. S., Cottrell, J. S., Corbett, J. M., Wheeler, C. H., Dunn, M. J., and Pappin, D. J. C. (1995) *Electrophoresis* **16**, 308–316
35. Mock, K. K., Sutton, C. W., and Cottrell, J. S. (1992) *Rapid Commun. Mass Spectrom.* **6**, 233–238
36. Pappin, D. J. C., Hojrup, P., and Bleasby, A. J. (1993) *Curr. Biol.* **3**, 327–332
37. Pappin, D. J. (1997) *Methods Mol. Biol.* **64**, 165–173
38. Hunt, D. F., Yates, J. R., Shabanowitz, J., Winston, S., and Hauer, C. R. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 6233–6237
39. Wilm, M., and Mann, M. (1996) *Anal. Chem.* **68**, 1–8
40. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25**, 3389–3402
41. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
42. Kozak, M. (1996) *Mamm. Genome* **7**, 563–574
43. Murray, A. W. (1991) *Methods Cell Biol.* **36**, 581–605
44. Blow, J. J. (1993) *J. Cell Biol.* **122**, 993–1002
45. Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982) *EMBO J.* **1**, 945–951
46. Klemm, R. D., Austin, R. J., and Bell, S. P. (1997) *Cell* **88**, 493–502
47. Ishiai, M., Dean, F. B., Okumura, K., Abe, M., Moon, K.-Y., Amin, A. A., Kagotani, K., Taguchi, H., Murakami, Y., Hanaoka, F., O'Donnell, M., Hurwitz, J., and Eki, T. (1997) *Genomics* **46**, 294–298
48. Coleman, T. R., Carpenter, P. B., and Dunphy, W. G. (1996) *Cell* **87**, 53–63
49. Landis, G., Kelley, R., Spradling, A. C., and Tower, J. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 3888–3892
50. Wolf, D. A., Wu, D., and McKeon, F. (1996) *J. Biol. Chem.* **271**, 32503–32506
51. Grallert, B., and Nurse, P. (1996) *Genes Dev.* **10**, 2644–2654
52. Ohtani, K., DeGregori, J., Leone, G., Herendeen, D. R., Kelly, T. J., and Nevins, J. R. (1996) *Mol. Cell. Biol.* **16**, 6977–6984
53. Triolo, T., and Sternglanz, R. (1996) *Nature* **381**, 251–253
54. Fox, C. A., Ehrenhofer-Murray, A. E., Loo, S., and Rine, J. (1997) *Science* **276**, 1547–1551
55. Kelly, T. J., Martin, G. S., Forsburg, S. L., Stephen, R. J., Russo, A., and Nurse, P. (1993) *Cell* **74**, 371–382
56. Liang, C., Weinreich, M., and Stillman, B. (1995) *Cell* **81**, 667–676
57. Nishitani, H., and Nurse, P. (1995) *Cell* **83**, 397–405
58. Piatti, S., Lengauer, C., and Nasmyth, K. (1995) *EMBO J.* **1141**, 3788–3799
59. Cocker, J. H., Piatti, S., Santocanale, C., Nasmyth, K., and Diffley, J. F. X. (1996) *Nature* **379**, 180–182
60. Smith, R. F., and Smith, T. F. (1992) *Protein Eng.* **5**, 35–41
61. Santocanale, C., and Diffley, J. F. X. (1996) *EMBO J.* **15**, 6671–6679
62. Donovan, S., Harwood, J., Drury, L. S., and Diffley, J. F. X. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 5611–5616
63. Jallepalli, P. V., Brown, G. W., Muzi-Falconi, M., Tien, D., and Kelly, T. J. (1997) *Genes Dev.* **11**, 2767–2779
64. Liang, C., and Stillman, B. (1997) *Genes Dev.* **11**, 3375–3386
65. Williams, R. S., Shohet, R. V., and Stillman, B. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 142–147
66. Tanaka, T., Knapp, D., and Nasmyth, K. (1997) *Cell* **90**, 649–660
67. Neuwald, A. F., Liu, J. S., Lipman, D. J., and Lawrence, C. E. (1997) *Nucleic Acids Res.* **25**, 1665–1677
68. Smith, R. F., and Smith, T. F. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 118–122
69. Cullman, G., Fien, K., Kobayashi, R., and Stillman, B. (1995) *Mol. Cell. Biol.* **15**, 4661–4671
70. Guenther, B., Onrust, R., Sali, A., O'Donnell, M., and Kuriyan, J. (1997) *Cell* **91**, 335–345
71. Kornberg, A., and Baker, T. A. (1992) *DNA Replication*, W. H. Freeman and Co., New York, NY
72. Stillman, B. (1994) *Cell* **78**, 725–728