AD_____

GRANT NUMBER DAMD17-94-J-4050

TITLE: General Methods for Identifying G1-Phase Substrates of CdK Protein Kinases

PRINCIPAL INVESTIGATOR: Dr. A. Bruce Futcher
Dr. Daniel R. Marshak

CONTRACTING ORGANIZATION: Cold Spring Harbor Laboratory
Cold Spring Harbor, New York 11724

REPORT DATE: July 1997

TYPE OF REPORT: Annual

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 1997	Annual (1 Jun 96 - 31 May 97)

4. TITLE AND SUBTITLE

General Methods for Identifying G1-Phase Substrates of CdK Protein Kinases

5. FUNDING NUMBERS

DAMD17-94-J-4050

6. AUTHOR(S)

Dr. A. Bruce Futcher
Dr. Daniel R. Marshak

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Cold Spring Harbor Laboratory
Cold Spring Harbor, New York 11724

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

We are using two-dimensional gels in combination with genetic and biochemical techniques to identify G1-phase substrates of cyclin-Cdk complexes. We have optimized protocols for labelling proteins with ^{32}P prior to two-dimensional gel electrophoresis. So far, we have successfully identified and characterized one important in vivo substrate, Sic1. Other workers have shown that an analogous substrate, p27, is important in human cells. More recently, we have identified and started to characterize two more substrates, Mcm3 and Cdc54. We have made peptide antigens for the purpose of making antibodies against phosphorylated Cdk substrates. Monoclonal antibodies possibly reactive against phospho-Ser-Pro are now being tested for their exact specificity.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer	9
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

BF In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

BF In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

BF In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Bruce Fuhrken Aug. 18, 1997

PI - Signature Date

TABLE OF CONTENTS

5. Introduction.

There is general agreement that in all eukaryotes, phosphorylation by various cyclin-Cdk complexes controls and orchestrates key cell cycle events. These events include commitment in G1 phase, initiation of DNA synthesis in S phase, and spindle formation and elongation in mitosis. However, despite knowing a great deal about the cyclin-Cdk complexes themselves, and despite years of investigation by many laboratories, we know only about half a dozen substrates of the cyclin-Cdk kinases, and none of these explain the control of critical cell cycle events. In particular, we do not know what substrates have to be phosphorylated for commitment to occur (although in mammalian cells, Rb is almost certainly one of the substrates).

The purpose of the present work is to develop methods for identifying substrates of the cyclin-Cdk complexes. In particular, we are interested in G1 substrates. To begin, experiments will be done in the yeast *S. cerevisiae*, and then the project will expand into mammalian cells.

We initially proposed two main approaches. The first approach uses two-dimensional gels to examine phosphoproteins. Various cyclins are expressed from a *GAL* promoter, and cells with the over-expressed cyclins are labelled with ^{32}P. The pattern of spots on a 2D gel is then compared between cells expressing and not expressing the cyclin. Extra spots in the cyclin-expressing cells may be substrates.

The second approach is to develop antibodies against phosphoserine followed by proline, and phosphothreonine followed by proline. Such antibodies would recognize proteins phosphorylated by Cdk complexes. Thus, such proteins could be immunoprecipitated and sequenced. This could also be combined with the 2D gels as an enrichment step.

6. Body of the Report.

A. Visualization of G1 Cdc28 substrates on 2D gels.

i. Experiments with 2D gels. We have done a large number of experiments examining yeast phosphoproteins on 2D gels; a manuscript detailing our experiences is almost ready to be submitted for publication. It has proven extremely difficult to find spots that are specifically phosphorylated by cyclin dependent kinases; the proteins that are phosphorylated in this way are non-abundant, and so the spots are relatively weak. This is not exactly a problem of sensitivity; rather, the problem is that the 2D gels contain numerous spots that are intensely labelled, and these obscure nearby weaker spots. At exposures that would allow us to see the weaker spots, the whole film has long since turned black. This problem, which we term "coverage" has limited the usefulness of this approach.

Nevertheless, by doing many experiments and many exposures, we were able to see (at least in some experiments) weak spots that may have been G1- or S-phase substrates of Cdc28 kinase. By examination of the molecular weight and isoelectric point of the spots, we guessed that they might be Sic1 (an inhibitor of Clb-Cdc28 kinases), Mcm3, and Cdc54 (Mcm proteins, which are important for DNA replication, and probably involved in movement of the replication fork). In last year's report, I described our work on Sic1. We were able to show that it almost certainly is an important substrate of Cdc28 at the end of G1, and this work has

been published (Schneider et al. 1996). Our work on Sic1 has been extended into mammalian cells by Roberts and co-workers (pers. comm.) who have shown that the mammalian CDK inhibitor p27 is likewise phosphorylated by G1 cyclin-CDK complexes, then degraded.

ii. Investigation of Potential Substrates involved in Replication. As described above, Mcm3 and Cdc54 are potential substrates identified in the 2D gel system. Mcm3 and Cdc54 are both "Mcm" (Mini-Chromosome Maintenance) genes. There are six different Mcm proteins in yeast, each of which is essential for viability. The six Mcms are structurally related to each other; and there are very closely related homologs in other eukaryotes, including humans. Eubacteria are not know to have Mcm proteins, but Mcm proteins are found in Archebacteria. The Mcm proteins are required for DNA replication (reviewed by Tye, 1994). Recent research suggests that the Mcms form a complex (a hexamer?) at origins of replication, and when replication begins, these complexes move away from the origins with the forks (possibly acting as helicases?) (S. Bell, pers. comm.).

Both Mcm3 and Cdc54 have clustered potential Cdc28 phosphorylation sites. The Xenopus Cdc54 homolog, called Mcm4, can be phosphorylated by Cdc2 (the Cdc28 homolog) (Hendrickson et al. 1996). We obtained soluble Mcm3 from baculovirus, mixed it with immunoprecipitated Cln-Cdc28 and Clb-Cdc28 complexes, and showed that Mcm3 could be phosphorylated by either form of Cdc28 in vitro. When the phosphorylated material was run on 2D gels, it migrated in about the same position as the Cdc28-dependent spots we had previously seen. However, mixing experiments were technically difficult, and we were not able to say whether the spots were precisely the same.

The N-terminal 120 amino acids of Cdc54 and its homologs in S. pombe (Cdc21), Xenupus (Mcm4) and humans (Mcm4) are shown below. There are two points of interest. First, there is little or no homology between the protein in this region; the homology (which is very high) begins at about amino acid 300. Second, all four proteins show multiple Cdc28 sites in this region. In particular, there is a consensus phosphorylation site starting at a.a. 7 in S. cerevisiae, Xenopus, and humans, and at a.a. 15 in S. pombe. (A Cdc28 consensus site is (S/T)-P-X-(K/R), and (S/T)-P is a minimal site.)

Cdc54 (Mcm4) N-terminus:

```
S.c.  MSQQSSSPTK EDNNSSSPVV PNPDSVPPQL SSPALFYSSS SSQGDIYGRN SNQNLSQGEG
S.p.  MSSSQQSGRA NELRTPGRAN SSSREAVDSS PLFFPASSPG STRLTTPRTT ARTPLASSPL
X.l.  MSSPTSTPSR RRVNVGEAVT LQLLVERKCS LLRHRKDGQK IPHQLVSFCP CQPHLQETFR
H.s.  MSSPASTPSR RGSRRGRATP AQTPRSEDAR SSPSQRRRGE DSTSTGELQP MPTSPGVDLQ


S.c.  NIRAAIGSSP LNFPSSSQRQ NSDVFQSGFR QGRIRSSASA SGRSRYHSDL RSDRALPTSS (23 aa) SSPRRIV
S.p.  LFESSSPGPN IPQSSRSHLL SQRNDLFLDS SSQRTPRSTR RGDIHSSVQM STPSRRREVD
X.l.  VPCFLALAPS RHSAHQSELD LSSPLTYGTP SSRVEGTPRS GIRGTPARQR ADLGSARKVK
H.s.  STAAQDVLFS SPPQMHSSAI PLDFDVSSPL TYGTPSSRVE GTPRSGVRGT PVRQRPDLGS
```

MCM3 has a notable cluster of conensus sites near the center of the protein:

Mcm3 SKRSPQKSPKKRQRVRQPASNSGSPIKSTPRRST

We mutated all the SP and TP sites in the N-terminal region of Cdc54 to AP, as well as a consenus site and near-consensus site later in the protein (seven mutations

total). That is, we made a multiple mutant which had essentially no Cdc28 sites left. We also mutated all the Cdc28 consensus sites in Mcm3. Unfortunately, neither multiply-mutant gene caused a significant phenotype on its own.

However, in the course of other work we had also made multiple site mutations in other replication proteins containing clusters of Cdc28 sites including Orc2, Orc6, and Cdc6. These also did not cause significant phenotypes on their own. We began combining multiple mutants, and now we began to get strong phenotypes. For instance, the *orc2* cdc54** double mutant is sick at 23°C, and inviable at 30°C, and shows a strong plasmid loss phenotype. The plasmid loss is seen with plasmids with a single ARS (origin of replication) and is suppressed by multiple ARSs, which shows that the defect has to do with initiation at an origin. The *mcm3* cdc6* orc2* orc6** quadruple mutant is likewise inviable at 30°C, with a strong plasmid loss phenotype at 23°C, whereas the *mcm3** single has no phenotype, and the *cdc6* orc2* orc6** triple has only a mild plasmid loss phenotype. We have found two sites in Orc2 such that reversion of just one of these sites reverts the entire phenotype.

An important question is whether the synthetic phenotypes we see are really due to a lack of phosphorylation, or whether they are due to the amino acid changes as such (that is, the substitution of alanine for serine or threonine). We are attempting to answer this by pseudoreversion. That is, we are placing novel Cdc28 consensus sites at novel locations in the N-terminus of Cdc54 and Orc2 (and will do the same for Mcm3), but leaving all the ser to ala mutations intact, to see if addition of novel sites can suppress the lack of the original one. If it can, it is a very strong argument that it is phosphorylation sites as such that are important.

If it turns out that the phosphorylation sites are important, then we will have a situation where (a) sites are redundant within a protein (since we need a multiple knockout to get a phenotype) and (b) sites are redundant between proteins. This is consistent with the idea that these proteins form complexes (as is strongly suspected on other grounds), and that complex formation is regulated by phosphorylation. Phosphates controlling association could be scattered at various places on relevant surfaces of several different proteins, thus accounting for the redundancy. It is important to know about this redundancy, since it then must be taken into account in any search for substrates. We already know that Clb-Cdc28 activity is essential for replication; the substrates and potential phosphorylation sites we are working on now may put us very close to the trigger for initiation of replication.

B. Develop antibodies against phosphoSer-Pro and phosphoThr-Pro.

Because of the difficulty in seeing spots phosphorylated by Cdc28 (not to mention the difficulty in purifying, identifying, and analysing them), we would like antibodies that could specifically recognize the phosphorylated forms of Cdc28 substrates. Since Cdc28 almost always phosphorylates a serine or threonine followed by a proline (i.e., SP or TP) we would like antibodies directed against phospho-S-P and phospho-T-P. It is likely that such antibodies can be generated; good antibodies have been made against phospho-tyrosine (Ross et al. 1981), and some antibodies have been made against phospho-threonine (Heffetz et al. 1989). The extra proline should make for a much better epitope than phospho-Ser or phospho-Thr alone.

We have synthesized a set of peptides to try and make the antibodies desired. The peptides are:

1. NH_2 C G G G **pS** P G G G K-OH COOH

2. NH_2 R A A **pS** P A A C-OH COOH

3. NH_2 C N N **pS** P N N H-OH COOH

4. NH_2 C G G **pT** P G G K-OH COOH

5. NH_2 R A A **pT** P A A C-OH COOH

6. NH_2 C N N **pT** P N N H-OH COOH

Each peptide has been conjugated (through the terminal cys residue) to three different carrier proteins, KLH, ovalbumin, and BSA. The immunization strategy to get (e.g.) anti-phospho-Ser-Pro antibodies has been to do a cycle of immunizations: peptide 1 coupled to KLH was the primary antigen; peptide 2 coupled to BSA was the first boost; peptide 3 coupled to KLH was the second boost; peptide 1 coupled to BSA was the third boost; peptide 2 coupled to KLH was the fourth boost; peptide 3 coupled to BSA was the fifth boost; and finally peptide 1 coupled to KLH was the sixth boost. Note that because of the different peptide sequences, the only thing in common between all the antigens was phospho-Ser-Pro. Additional immunizations can be done as necessary to get high titers. The peptide coupled to ovalbumin in not used for immunizations; it is used for the initial screens of fusion clones. Thus, when a positive is picked up, we know it is not reacting with one of the carrier proteins (KLH or BSA) used for the immunization.

At the time of last year's report, we were close to doing the fusions. Unfortunately, all six mice (three for phospho-SP, three for phospho-TP) died in the four weeks before the fusion was scheduled. Conceivably this may have had something to do with the antibodies they were producing. We have had to re-couple peptides to carrier, and immunize a new set of mice. We have settled for somewhat lower reactivities in the sera, in case the mice will be killed at higher titers. The fusion has been done for a phospho-SP mouse. We screened about 800 primary clones, and have 12 primary clones that pass the prelimary screens and may indeed be specific for phospho-SP. These are now being sub-cloned to give true monoclonals. At that point we will determine the exact specificity of these antibodies. We also have three mice showing reactivity to the phospho-TP antigens, and a fusion is scheduled in about a month.

C. Future plans.

We are working hard on the mutant replication proteins. We are concentrating on the pseudoreversion studies. If these studies indeed say that the mutant phenotypes are really due to the lack of phosphorylation, our next step will be the detailed biochemical measurements showing that the wild-type sites are phosphorylated in a cell-cycle dependent, Cdc28-dependent manner in vivo. Ultimately, we would like to know (a) exactly how these phosphorylations trigger replication; (b) extend studies to the homologous substrates in mammalian cells; and (c) generalize the studies to other substrates.

We are also working hard on the anti-phospho-SP and -TP antibodies. Our next two steps will be to characterize the putative anti-phosopho-SP antibodies more thoroughly. We still need to see if the phosphate and the proline are absolutely required. It is also possible that there are subtle context effects; that is, the antibodies might recognize phospho-SP in some amino acid contexts, but not others.

This will have to be investigated carefully with substrates carrying SP in different contexts. Of course, we are also anxious to do the fusions with the phospho-TP mice. If antibodies of the desired type are obtained, we would return to the 2D gels, but immunoprecipitate phosphate labelled protein with anti-phospho-SP and -TP antibodies before running the gels. This should largely solve the coverage problem which has bedeviled this approach.

7. Conclusions.

Without some enrichment step for substrates of the desired type, the 2D gel approach is probably not practical for non-abundant substrates. Nevertheless, it has helped us find one relevant substrate (Sic1) and probably two others (Mcm3 and Cdc54). Work on these substrates is going well. The anti-phospho-SP and -TP antibodies are potentially extremely useful; their characterization is continuing. These may allow the 2D approach to be used successfully.

8. References:

1. Heffetz, D., Fridkin, M., and Zick, Y. (1989) Eur. J. Biochem. *182*, 343.
2. Hendrickson, M., Madine, M., Dalton, S., and Gautier, J. (1996) PNAS *93*, 12223.
3. Ross, A. H., Baltimore, D., and Eisen, H. (1981) Nature *294*, 654.
4. Schneider, B., Yang, Q.-H., and Futcher, B. (1996) Science *272*, 560.
5. Tye, B.-K. (1994) Trends Cell Biol. *4*, 160.