

Substrate Specificity of CDC2 Kinase from Human HeLa Cells as Determined with Synthetic Peptides and Molecular Modeling

Jianwen Zhang, Roberto J. Sanchez, Shenglun Wang, Corrado Guarnaccia, Alessandro Tossi,* Sotir Zahariev, and Sándor Pongor¹

International Centre for Genetic Engineering and Biotechnology, Area Science Park, Padriciano 99, Trieste-34012, Italy; and
*Dipartimento di Biochimica Biofisica e Chimica delle Macromolecole, Università degli Studi di Trieste, 34127 Trieste, Italy

Received May 17, 1994, and in revised form July 13, 1994

A systematic study was undertaken in order to assess the substrate specificity of cyclin-B/cell division control protein kinase (CDC2) isolated from human HeLa cells, using 13–15 residue peptides with a central histone-like KKSPKK motif as a model. Replacement of the proline residue by any of the other 19 amino acids or D-proline drastically reduces or abolishes phosphorylation by CDC2. Changing the basic residues to Ala on either side of the -SP- structure differentially reduces phosphorylation. Molecular modeling and dynamics simulation indicated that the phosphorylation site of the peptide may have to adopt a turn-like conformation that will orientate the charged and hydrophobic residues so as to allow interaction with postulated binding surfaces within the CDC2 active site. It thus appears that, of the 20 coded amino acids, only proline can provide this conformation in short peptides. This is in agreement with the finding that sarcosine can replace proline in this respect (S. Ando *et al. Biochem. Biophys. Res. Commun.* 195, 837–843, 1993). © 1994 Academic Press, Inc.

Key Words: CDC2 kinase; phosphorylation; synthetic peptides.

The cyclin-dependent protein kinases are critically important regulators of the eukaryotic cell cycle (1).² Their transient activation is thought to trigger specific stages of the cell cycle, including mitosis and the initiation of a new cycle ("Start"). In yeast, both of these transitions are thought to be controlled by a single cyclin-dependent

kinase (CDC2³ in *Saccharomyces pombe* and CDC28 in *Saccharomyces cerevisiae*) (2). In human cells, a variety of cyclin-dependent kinases were described, two of which (CDC2 kinase and CDK2 kinase) have been studied extensively. Both CDC2 and CDK2 are structurally related to yeast CDC2/CDC28, but appear to carry out different functions. CDK2 is thought to be mainly involved in the G₁ to S phase transition (3, 4), whereas CDC2 has been implicated mainly in the control of mitosis (1). Even though a great many proteins have been identified as substrates of CDC2 kinase *in vivo* or *in vitro* (5, 6), until recently studies with synthetic peptides have been less common (7–9).

CDC2 is a serine/threonine kinase, composed of a catalytic subunit, p34^{cdc2}, and a regulatory cyclin component which is essential both for kinase activity and for cell-cycle dependent activation (6, 10, 11). In yeast, p34^{cdc2} is associated with a distinct group of G₁/S cyclins and plays a role in the G₁/S transition (10, 12, 13). The catalytic subunit of human CDC2 kinase, p34^{cdc2}, requires cyclin B for function, even though it has been shown to interact with a family of related cyclins (A,C,D1,D2,D3) (14, 15). CDK2 kinase of higher eukaryotes, on the other hand, has p34^{cdk2} as the catalytic and cyclin A as the regulatory subunit (3, 16, 17).

In vitro studies of CDC2 substrate specificity are complicated by the fact that CDC2 is regulated not only by cyclins but also by complex phosphorylation and dephosphorylation events. There is also evidence to indicate

³ Abbreviations used: P53, human anti-oncoprotein P53; RB, retinoblastoma tumor repressor gene product; T-antigen, simian virus large tumor antigen; myosin LC, myosin regulatory light chain; sucl, p13^{suc1} protein; cAPK, cAMP-dependent protein kinase; PKI, cAMP-dependent protein kinase inhibitor; ERK, extracellular signal-related kinase; CDC2, cell division control protein kinase; CDK2, cyclin-dependent kinase 2; Sar, sarcosine; TFA, trifluoroacetic acid.

¹ To whom correspondence should be addressed. Fax: +39-40-226-555. E-mail: pongor@icgeb.trieste.it.

² Portions of this paper are presented as a Miniprint Supplement at the end of this paper.

that some members of the Cdk family may have functions that are unrelated to the cell cycle (18). In spite of these complicating factors there is growing evidence that the sites phosphorylated by CDC2 kinase can be classified in two categories: (i) sites containing proline ("S/TP") and (ii) sites without proline ("SabR"). The sites in the first category are represented in many proteins and may be generalized with the consensus sequence *O-S/T-P-X-Z*, where *O* is polar or basic, *X* is polar, and *Z* is generally a basic amino acid (1). Proline and the basic amino acids are generally thought to be essential for phosphorylation to occur even though there are exceptions to this rule. For example, P53 (19), cyclin B (20), and RNA polymerase II (21) were shown to be substrates of cyclin-B/CDC2 kinase *in vivo* or *in vitro*, although their phosphorylation sites lack a basic residue in the second position after proline. With synthetic peptide substrates however, previous studies have demonstrated that the polar/basic residue *O* before *S/T* is not essential for the phosphorylation, but the second basic residue after proline is required (7). The second category is represented in the myosin light chain (22), in glial fibrillary acidic protein (23), in desmin (24), and in vimentin (9). The phosphorylation site for this category is less conserved than that of the previous group of substrates and can be written as *SabR*, where *a* is (often) serine, threonine, or alanine and *b* is lysine, glutamine, leucine, or alanine. The structural requirements for this site are, to our knowledge, not well documented. Ando *et al.* recently found that a synthetic peptide, LGSALRRR, corresponding to one of the CDC2 phosphorylation sites in vimentin was relatively weakly phosphorylated *in vitro* by CDC2, but that phosphorylation increased as alanine was replaced by proline, or, more interestingly, by sarcosine (9). This lead to the conclusion that the N-substituted structure, common to proline and sarcosine, may be necessary for the enzyme to recognise a substrate (9). Kitagawa and co-workers (8) recently described a selective, competitive inhibitor of CDC2/CDK2 kinases, butyrolactone I, a compound devoid of proline and of basic moieties. This compound is quite hydrophobic and apparently competes both with both ATP and the substrate.

One of the intriguing properties of human CDC2 kinase is the fact that it can specifically interact with a wide variety of protein substrates. *In vitro* specificity studies with recombinant proteins are greatly hampered by the fact that CDC2 may need a variety of factors in order to reach the activity seen *in vivo*. On the other hand, purification of sufficient amounts of active CDC2 kinase from cells is made difficult by the fact that the CDC2 and CDK2 families have catalytic and regulatory subunits that are similar in size and sequence. Recently, however, Pan *et al.* (25), as well as Kitagawa and associates (8), published purification methods that make it possible to isolate CDC2 kinase in sufficient amounts, and we have used this strategy to isolate CDC2 kinase from cultured HeLa cells.

In this paper we report on a systematic study in which specificity of cyclin-B/CDC2 kinase was probed with 34 synthetic peptides designed on the basis of the KKSPKK sequence motif of histone H1. In one set of peptides, the central proline was replaced with all the other amino acids, while, in another set, the lysines flanking -SP- were systematically replaced by alanines. Replacement of Pro by D-Pro and of Ser by Thr was also examined. A correlation of the specificity/inhibition data with 3D models of the peptide substrate made it possible to outline the approximate spatial arrangement of CDC2/substrate interactions which may allow us to design specific inhibitors to this enzyme in the future.

MATERIALS AND METHODS

Peptides synthesis. Two sets of peptides were designed and synthesized; one set was based on the sequence corresponding to residues 139–153 of trout histone H1 (26, Table IA), in which Pro 146 was replaced by all the other 19 amino acids. The other set had the central KKSPKK motif of the previous sequence flanked by GAGA sequences and the lysines were systematically replaced by alanines (Table IB). Peptides corresponding to the phosphorylation sites of the SV40 large T-antigen (27), P53 (19), RB (28), nucleolin (29) and myosin LC (22) were also synthesized and used as a control.

A total of 34 peptides was prepared by automated solid-phase peptide synthesis using the Fmoc chemistry (30) in dimethylformamide as a solvent on a type 9050 automated synthesizer from MilliGen/Biosearch. Fmoc-Ala substituted pepsyn-KA resin (Milligen, 0.1 mmol scale) was used for peptides p-1 and p-4. Fmoc-Ala substituted PEG-PS resin (Milligen, 0.02 mmol scale for each peptide) was used for the analogues of peptide p-1 and all peptides in Table IC. Fmoc-Phe-pepsyn-KA resin was used for peptide p-2, Fmoc-Leu-pepsyn-KA was used for peptide p-3, Fmoc-Met(*t*Bu)-Pepsyn-KA resin was used for peptide p-5, and Fmoc-Arg(Pmc)-Pepsyn-KA resin for peptide p-6. Peptides in Table IA were synthesized singly, whereas peptides in Table IC were synthesized in a multiple column mode and the 19 analogues of peptide p-1 (Table IB) were synthesized using a segmented column packing technique. A mixture of Fmoc-protected amino acid (4 M excess; Milligen or Novabiochem, Laufelfingen, Switzerland), *N*-hydroxybenzotriazole (4 M excess; Aldrich-Chemie, Steinheim, Germany), and 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate (4 M excess; Novabiochem) as well as diisopropylethylamine (8 M excess; Aldrich-Chemie) dissolved in dimethylformamide was used for each coupling step. Side-chain protection was as follows: trityl (Gln,Asn,Cys), *t*-butylester (Glu, Asp), *t*-butylether (Ser, Thr), *t*-butyloxycarbonyl (Lys,Trp), and 2,2,5,7,8-pentamethylchroman-6-sulfonyl (Arg). Cleavage from the resin and deprotection was carried out using a mixture of 82.5% trifluoroacetic acid (TFA), 5% each of phenol and ethanedithiol, and 2.5% each of water, thioanisole, and triisopropylsilane. Peptides were repeatedly extracted with ether and purified by RP-HPLC on a C18 column (19 × 300-mm Delta-Pak; Waters, Bedford, MA), using a 0–40% acetonitrile gradient in 0.05% TFA.

The homogeneity of the peptides was confirmed with analytical RP-HPLC on a C18 column (3.9 × 150-mm Delta-Pak) using the same conditions and the correct composition was confirmed by amino acid analysis using the PicoTag method (Waters). Peptides were quantified on the basis of the amino acid analysis, the optical densities at 205 or 280 nm, and from the molar concentration of the N-terminal groups, as determined by the ninhydrine reaction (31).

Phosphorylation assay. Phosphorylation assays were carried out as described by Kamijo *et al.* (7), except that the concentrations of the peptides and of the bovine histone H1 substrates (Boehringer, Mannheim, Germany) were adjusted to 300 and 6 μM, respectively. The reaction was stopped after 30 min and the samples were analyzed by acid-urea

polyacrylamide gel electrophoresis (32) combined with autoradiography (Fig. 1), followed by scintillation counting of the excised bands. The proline-substitution data in Table IB were determined by the P81 paper spotting assay (7). The substrate competition assays were carried out in the same manner as the histone H1 phosphorylation assays, except that the inhibitor was added to the reaction mixture before starting the reaction with ATP.

Kinetic constants were determined as described by Ando *et al.* (9). The reaction mixture with a final volume 50 μ l was incubated for 5 min. Each selected peptide was used in a 15-fold range of concentration around the apparent K_M value. Other assay conditions were identical to those described in the phosphorylation assay. The kinetic constant values were obtained from Lineweaver-Burk plots (9).

cdc2 purification. Cyclin-B/CDC2 kinase was purified from HeLa cells through a procedure similar to that used by Kitagawa *et al.* (8) and Pan *et al.* (25), and was followed by *suc1* affinity chromatography (33) and heparin chromatography. A synthetic peptide (p-0 in Table IC) was used to monitor the purification (the procedure is described in the Mini-print Supplement).

Modeling studies. Models representing the KKSPKK motif were built based on homology, using the Insight/Discover package (34) and the CVFF force field (35). The starting structures used were: (a) the theoretical model of Suzuki (36, 37), (b) an NMR-based preliminary model of Suzuki *et al.* (38), and (c) the ARSPDK motif in the histone-like domain of HNF3 (39). The side chains in the respective starting structures were first replaced by those in the KKSPKK motif. Subsequently the models were subjected to 100 steps of steepest descent energy minimization for the side chains, followed by 10-ps dynamics simulations at 300 K with a distance dependent dielectric constant and a final complete minimization until the maximum energy derivative was less than 0.01 kcal/Å. The VADAR program (version 0.9) from David Wishart and Leigh Willard (University of Alberta) was used to predict the H-bonds with the DEFINE-S algorithm (40).

RESULTS

The CDC2 kinase preparation used in this study was obtained through a series of chromatographic and affinity purification steps and was found to be the cyclin-B/p34^{cdc2}/suc1 complex. The *suc1* purification procedure was incorporated into the purification scheme in order to avoid a loss of enzyme activity experienced with repeated chromatographies (8, 25) and also because *suc1* does not influence the phosphorylation of the histone H1 substrate used in this study (41). The CDC2 preparation was electrophoretically homogeneous and devoid of p34^{cdc2} and other kinase activities. The two methods selected to quantitate phosphorylation (electrophoresis and P81 paper spotting) essentially gave identical results. We estimate that the relative phosphorylation values are reproducible within an experimental error of $\pm 5\%$.

Synthetic peptides corresponding to the phosphorylation sites of histone H1, SV40 T-antigen, P53, RB, and nucleolin were all found to be efficiently phosphorylated by the CDC2-kinase *in vitro*, with only exception of the myosin-LC-derived peptide. The relative phosphorylation was highest for the histone derived peptide; however, a 66% decrease was observed when the serine residue was replaced with a threonine in this peptide (Table IC, peptide p-0T).

The replacement of proline with any of the 19 natural amino acids or with D-proline (Table IB and peptide p-

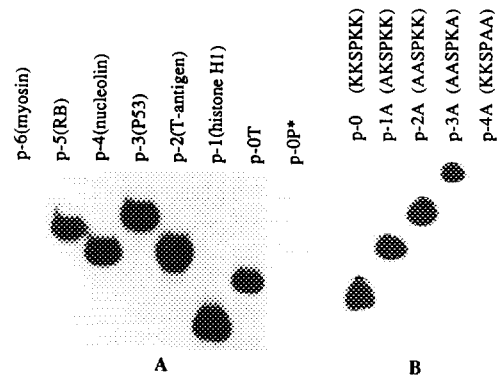


FIG. 1. Peptide phosphorylation assay: auto radiography of 20% acid-urea-SDS gels (the peptide sequences are shown in Table I).

0P* in Table IC) drastically decreased the phosphorylation of the synthetic peptides by CDC2 kinase. Low but detectable levels of phosphorylation were obtained with Met, Ile, Phe, Val, and Leu substitutions (Table IB). All other amino acids, except Cys, gave relative phosphorylation values below 5%. In view of the reproducibility of the assay, we consider these data as background. The phosphorylation of the cysteine containing peptide was found to increase with time. When the sample was oxidized with iodine, a total of 13% relative phosphorylation was observed, while the freshly reduced sample (CysH) showed relative phosphorylation of 9%. Acid-urea electrophoresis revealed two adjacent bands; the faster migrating one corresponding to the monomeric species (CysH) and the other corresponding to the phosphorylated dimer (Cys-SS), respectively. It would thus appear that both monomeric and dimeric Cys-containing peptides are substrates for CDC2, with a relative phosphorylation comparable to that of the Met- or Ile-substituted peptides.

Kamijo and co-workers (7) previously found that proline cannot be replaced by lysine in short, synthetic peptides. In this study we have found that proline cannot, in fact, be replaced by any other amino acid, not even by serine, which, according to Suzuki (36), can adopt a pseudo-proline conformation in the histone-like motifs. It is worth noticing that residues with long hydrophobic side chains, introduced in place of proline, gave detectable phosphorylation levels. This leads us to suppose that these residues may interact with the CDC2 active site, thereby promoting substrate binding. The size and polarity of this residue seems to be a crucial factor: short (Gly, Ser, Thr), charged (Glu, Asp, Lys, Arg), and polar side chains (Gln, Asn) all gave poor substrates. The largest of the residues that gave a detectable phosphorylation was Phe; the bulkier hydrophobic residues (Tyr, Trp) gave very low phosphorylation levels.

The effect of the basic amino acids was studied on model peptides that contained the KKSPKK motif flanked by GAGA sequences. The results in Table IC show that

TABLE I
Relative Phosphorylation Rate of Synthetic Peptides by Human CDC2 Kinase

A			
Peptide	Source	Sequence	Relative phosphorylation rate (%)
p-1	Histone H1	AVAAKKSPKKAKKPA ↓	100
p-2	SV40 T-antigen	DSQHS ^T TPPKKKR ^K VEDPKDF ↓	98
p-3	P53	PNNTSSSPQPKKKPL ↓	78
p-4	Nucleolin	AVTPAKKAVTPAKKA ↓	75
p-5	RB-P105	MYLSPLRS ^P PKKRTST ↓	54
p-6	Myosin CL	SSKRAKAKTTKKR ↓	2
B			
Peptide	Sequence		Relative phosphorylation rate (%)
p-1	AVAAKKSPKKAKKPA ↓		100
p-CC	(AVAAKKSC*KKAKKPA) ₂		13
p-M	AVAAKKSMKKAKKPA		12
p-V	AVAAKKSVKKAKKPA		12
p-I	AVAAKKSIKKAKKPA		10
p-C(SH)	AVAAKKSCKKAKKPA		9
p-L	AVAAKKSLKKAKKPA		8
p-F	AVAAKKSFKKAKKPA		8
p-T	AVAAKKSTKKAKKPA		4
p-Y	AVAAKKSYKKAKKPA		3
p-A	AVAAKKSAKKAKKPA		3
p-W	AVAAKKS ^W KKAKKPA		2
p-S	AVAAKKSSKKAKKPA		2
p-H	AVAAKKSHKKAKKPA		1
p-K	AVAAKKSKKKAKKPA		1
p-E	AVAAKKSEKKAKKPA		1
p-R	AVAAKKSRKKAKKPA		0.2
p-Q	AVAAKKSQKKAKKPA		0.2
p-G	AVAAKKSGKKAKKPA		0.2
p-N	AVAAKKSNKKAKKPA		0.1
p-D	AVAAKKSDKKAKKPA		0.0
C			
Peptide	Sequence		Relative phosphorylation rate (%)
p-0	GAGAKKSPKKAGAGA ↓		100 ^a
p-1A	GAGAAKSPKKAGAGA		85
p-2A	GAGAAASPKKAGAGA		51
p-3A	GAGAAASP ^K AGAGA		4
p-4A	GAGAKKSP ^A AGAGA		2
p-0P*	GAGAKKSP*KKAGAGA		3
p-0T	GGGAKKTPKKAGAGA		34

Note. (A) Synthetic peptides corresponding to proteins that are phosphorylated by CDC2 kinase at "↓." Sequential location of the peptides, given as position number in the respective entries of the Swiss-Prot database, release 26 (Bairoch and Boeckman, 1993): p1, H1_SALTR, 139-153; p2, TALA_SV40, 119-138; p3, P53_HUMAN, 309-323; p4, NUCL_CHICK, 75-90; p5, RB_HUMAN, 349-365; p6; MLRN_HUMAN, 1-13. (B, C) Peptides are designed to contain a histone-like KKSPKK motif flanked by GAGA sequences. The phosphorylated serine is marked by "↓," and the residues substituted as compared to p-0 are underlined. P*, D-proline. pCC and pC(SH) are the oxidized and reduced forms of the cystein-containing analogue, respectively, C*, cystine.

^a The relative phosphorylation of p-0 was measured to be 90% of that of p-1.

phosphorylation is markedly reduced by K to A replacements on either side of the serine residue. These data confirm the finding of Kamijo and associates on the importance of basic residue located in the second position, downstream from proline (7). In addition, we found that positive charges upstream from the serine residue contribute to both the efficiency of phosphorylation and the ability of the peptides to compete with the phosphorylation of histone H1. It appears that downstream lysines are more important, however. For example, peptide p-3 (SSSPQKKK) is a relatively good substrate, even though it contains no lysines upstream of the phosphorylation site.

The above tendencies are clearly illustrated by the kinetic constants of the phosphorylation reaction (Table II). The replacement of proline leads to an at least 50-fold increase in K_M and a 6-fold decrease of V_{max} . When the positive charges are replaced by alanine, the respective changes are clearly smaller (Table II).

None of the peptides were found to be efficient inhibitors when tested in a histone phosphorylation assay (Fig. 2). The peptides only suppressed histone phosphorylation at high concentrations and their relative inhibitory activity goes in parallel with the relative extent of their phosphorylation. Since peptide p-0, whose sequence is identical to that of the histone substrate, had the strongest effect, this phenomenon can thus be simply regarded as weak competitive inhibition.

As the currently available histone H1 structural coordinates do not yet include the KKSPKK motif (42), we used three different homologous structures for building three-dimensional models of the substrate peptides (Figs. 3a–3c). The theoretical model of Suzuki (model *a*) can be considered a mixture of a β -turn and an Asx-type turn

TABLE II
The Effect of Residue Substitutions in the Peptide Substrates on the Enzymatic Constants of cdc2 Phosphorylation, K_M , and V_{max} ^a

Peptide	Sequence	V_{max} ($\mu\text{mol}/\text{min}/\text{mg}$)	Apparent K_M (μM)
H1*		3.08	1.2
P-1	A V A A K K S P K K A K K P A	4.05	3.5
P-CC	A V A A K K S C K K A K K P A	0.58	48.2
P-M	A V A A K K S M K K A K K P A	0.65	88.1
P-E	A V A A K K S E K K A K K P A		ND
p-0	G A G A K K S P K K A G A G A	3.82	5.5
p-1A	G A G A <u>A</u> K K S P K K A G A G A	3.60	11.3
p-2A	G A G A <u>A A</u> S P K K A G A G A	1.94	26.4
p-3A	G A G A <u>A A A</u> S P K <u>A</u> A G A G A	0.31	54.0
p-4A	G A G A <u>K K</u> S P <u>A A</u> A G A G A		ND
p-0P*	G A G A K K S <u>P</u> * K K A G A G A		ND
p-0T	G A G A K K <u>T</u> P K K A G A G A	3.189	32.1

^a H1, histone H1; p-0P*, D-proline substitution; ND, phosphorylation with a rate less than $0.05 \mu\text{mol} \cdot \text{min}^{-1} \text{mg}^{-1}$ at a peptide concentration of 0.75 mM was considered not detectable.

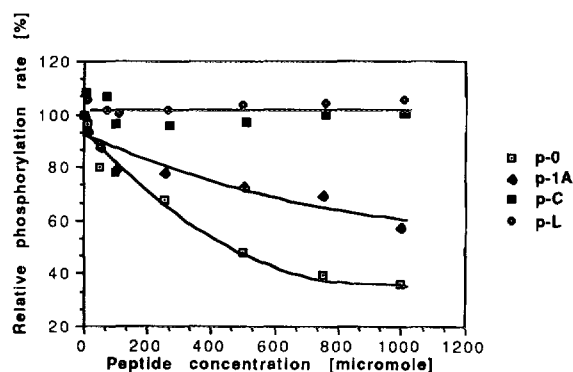


FIG. 2. Competitive phosphorylation inhibition assay. The reactions were carried out in the presence of $6 \mu\text{M}$ of bovine histone H1. The sequences are given in Table I.

(36, 37). The NMR structure used for model *b* is also a mixed turn type (type I β -turn and Asx turn; Ref. (38)). Finally, the loop motif, taken from the histone-like motif of HNF3 protein, is in an irregular region. Even though the resulting structures were all somewhat different, they all show a turn-like conformation in the SPKK segment (structure *b* is a type I turn). In all three structures an H-bond connects the carbonyl oxygen of serine with the α -amino group of Lys+3 (position with respect to serine). Lys-1 and Lys-2 are located on one face of the turn (Face A), while Lys-2 and Lys+3 are located on the opposite face (Face B). In all three models a triangular arrangement of Ser(γ -O) and the ϵ -N atoms of Lys-1 and Lys+3 is apparent (Fig. 3). The approximate distance between Ser(γ -O) and Lys-1(ϵ -N) is 6 \AA , between Ser(γ -O) and Lys-3(ϵ -N) is 7.5 \AA , and between Lys-1(ϵ -N) and Lys-3(ϵ -N) is around 7 \AA (Fig. 3, legend). All three structures were stable when subjected to molecular dynamics simulations at 300 K. In order to simulate the effect of proline replacements, L-proline in model *b* was replaced by D-proline or by sarcosine and the models were subjected to energy minimization and dynamics simulation as described under Materials and Methods. Both replacements extensively changed the conformation of the turn motif in such a way that the H-bond between Ser and Lys+3 was no longer possible (models not shown). The most conspicuous change was in the position of the Lys+3 chain, which is essential for phosphorylation in peptides. In the D-Pro-substituted model (which corresponds to a nonphosphorylated substrate), the Lys+3 side chain moved so that the ϵ -amino group was displaced by about 11 \AA from its original position in model *b*. In the Sar-substituted model, the Lys+3 side chain was in an intermediary position between the D-Pro variant and model *b* (the distance of ϵ -amino group was about 7 \AA from its original position). This result is in keeping with the experimental finding that both L-Pro or Sar replacement at position -*a* confers an increased level of phosphorylation to a *SabR*-type pep-

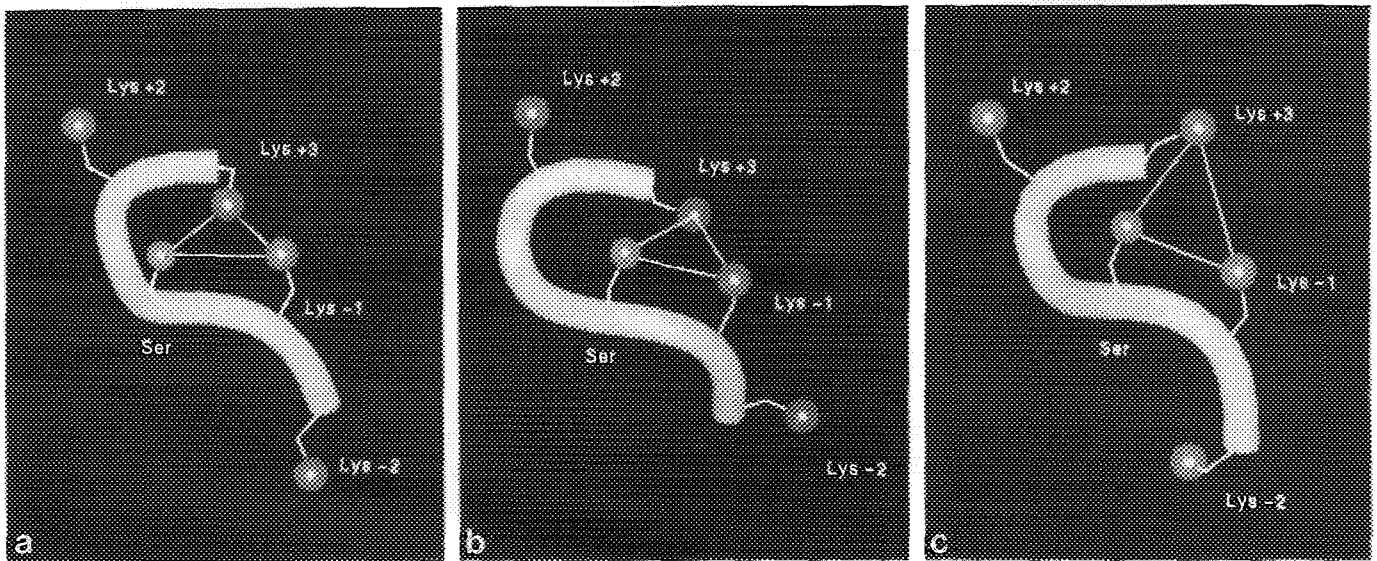


FIG. 3. Simplified molecular models of the KKSPKK substrate built from three different starting structures: (a) the theoretical model of Suzuki (36, 37), (b) an NMR-based preliminary model of Suzuki *et al.* (38), and (c) the ARSPDK motif in the histone-like domain of HNF3 (39). Lysine C, and serine O atoms are shown as spheres. The distances measured between the N atom of the Lys ϵ -amino groups and the Ser (γ -O) atom were as follows [numbers denote the average distance in the three models (\AA), followed in parentheses by the individual values for models a, b, and c, respectively]: Ser to Lys+3: 7.5 (6.8, 7.1, 8.7); Ser to Lys-1: 6.0 (6.8, 5.3, 6.0); Lys-1 to Lys+3: 7.2 (7.4, 8.8, 5.5). (The distances are approximate as Ser (γ -O) is presumed to rotate freely.)

tide, albeit that the effect is considerably greater for L-Pro (9).

DISCUSSION

Protein substrates such as those phosphorylated by CDC2 *in vivo* can interact with the catalytic and regulatory subunits in a variety of ways which may influence their phosphorylation. Short peptides, on the other hand, supposedly interact predominantly with the active site and its surroundings. In the following discussion we seek to interpret our data in terms of three major factors, i.e., the conformation, the size, and the polarity of the substrate.

Generally speaking, a peptide or protein substrate may need to adopt a "suitable conformation" in order to bind to the active site of CDC2. In principle, this conformation may either be *a priori* present in the substrate molecule or may be induced upon binding to the enzyme. Short peptides, like the ones used in this study, have little regular conformation in solution, so the conformation necessary for phosphorylation may have to be stabilized on binding to the active site. Three-dimensional models of the peptide substrate (Fig. 3) indicate that the substrate molecule may have a bent conformation that brings the lysine residues flanking the phosphorylation site near to each other. Even though stable turns are known to be present in short peptides, it is improbable that such a conformation could be stable in the highly charged KKSPKK motif without help from the enzyme; i.e., the

structure suitable for phosphorylation of this substrate may be formed upon binding to the active site. A turn-promoting residue, such as proline, may thus, in fact, be crucial for the suitable conformation to occur. On the other hand, intact protein molecules can provide additional forces in order to create a turn-like conformation at the CDC2 phosphorylation site. This is in keeping with the fact that the *SabR*-type motifs, which are phosphorylated in myosin, vimentin, and some other proteins, are not, or are only weakly, phosphorylated when presented to the enzyme in short peptide substrates. Also, when *a* was replaced by proline in the *SabR*-type motif of vimentin, the otherwise weakly phosphorylated peptide became a good substrate to CDC2 (9). In addition, sarcosine, which is known for its ability to replace proline in turn-like structures (43), was able to increase phosphorylation of the same peptide (9). The experimental data thus suggest that a β -turn may be a general structural feature of the CDC2 phosphorylation site, in a manner similar to the way in which it is supposed to be important in other kinase phosphorylation sites (44, 57). This is corroborated by the fact that all of the well-phosphorylated substrate sequences used in this study have predicted β -turns, when analyzed with the method of Garnier *et al.* (45) (data not shown). Pines and Hunter (6) suggested that the phosphorylation site of myosin has a predicted β -turn which might substitute the absent proline. The data presented here and those published by Ando *et al.* (9) indirectly confirm this suggestion. On the other hand, it is interesting to note that sarcosine and proline also share a ten-

dency to *cis/trans*-isomerise (46); therefore, one may thus speculate whether, in addition to turn formation, this flexibility plays a role in creating the conformation necessary for phosphorylation.

The importance of positive side chains in the sequential vicinity of the phosphorylation site is apparent both from this and from previous studies. Side chain +3 is especially important; its replacement by alanine abolishes phosphorylation. The replacement of the other lysines reduces phosphorylation by 15–50% but does not abolish it. We suppose that side chain +3 may provide a primary interaction with the enzyme while the other positive charges may either be involved in less important enzyme/substrate interactions or simply provide a favorable ionic milieu for the reaction. One of the mechanisms can be the stabilization of the phosphorylated product. It is well known that phosphoserine (or phosphothreonine)-containing peptides are chemically more stable if there are basic moieties in the sequential vicinity of the phosphorylated residue (47).

Threonine was phosphorylated at a relative rate of 34% compared with serine within the same sequential environment. The phosphorylation data of the nucleolin peptide, p-4, (75%), are in reasonable quantitative agreement with the known fact that two threonine residues are phosphorylated in the peptide. It appears that we can safely conclude that threonine is less efficiently phosphorylated by CDC2 than serine (7).

In the light of the foregoing study, the specificity of CDC2 for synthetic peptides can be summarized as follows: (i) Serine is a better phosphate recipient than threonine. (ii) The phosphorylation site may have to adopt a turn-like conformation in order to fit the active site. In short peptides, this conformation is best ensured by a turn-promoting residue after the phosphorylated side chain. (iii) The third residue after the phosphorylated side chain should bear a positive charge. (iv) Positively charged residues on either side of the phosphorylated serine increase phosphorylation, those downstream being more important.

The approximate 3D models of the peptide substrate suggest that the substrate binding site of the enzyme may be a pocket-like structure (Fig. 4) that can accommodate the protruding serine residue and may have binding sites capable of interacting with the positive charges (Fig. 4, I). The fact that some hydrophobic residues, introduced in place of proline confer a well-measurable phosphorylation to the p-1 peptide substrate prompts us to suggest that a hydrophobic surface may be present in the active site opposite the proline residue of the substrate (Fig. 4, II). It is conspicuous that only residues of a certain size, namely Ile, Met, and Leu, are acceptable, whereas smaller (Gly, Ala, Val) or bulkier ones (Tyr, Trp) are not.

It is interesting to compare our schematic active site model with the known enzyme structures cAPK(48), cdk2 (49), and ERK2 (50). The cdk2 active site is expected to be the most similar to cdc2, however, the cdk2 structure

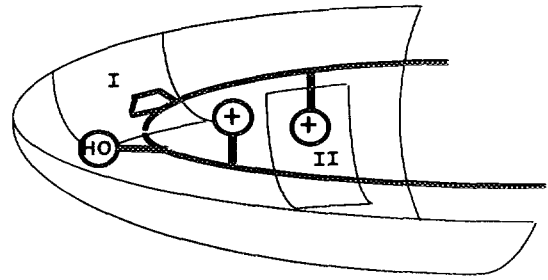


FIG. 4. Schematic representation of the CDC2 binding site interacting with the KKSPKK. Lys-1 and Lys+2 are denoted with a + sign. "I" indicates the approximate location a putative hydrophobic surface. "II" indicates the approximate location of a surface interacting with the positive charge in position +3.

is that of the inactive monomer and the role of cycline binding and Thr60 phosphorylation may have an effect on substrate binding which is difficult to assess. The +1 residue of PKI is bound to a deep hydrophobic pocket in PKA (51). This hydrophobic pocket is partially blocked (i.e., shallower) both in cdk2 (49) and in ERK2 (50). So, one could argue that the role of proline in the +1 position comes from the fact that, among all the natural amino acids, it is the only one small enough yet sufficiently hydrophobic to fit into this shallower hydrophobic pocket. The fact that sarcosine can replace proline in this respect (9) confirms this view. In other words, "proline-directed phosphorylation" seems to be a simple consequence of the physicochemical and conformational properties of the +1 residue. The fact that sarcosine is a good replacement for proline (9) supports this view. This model is necessarily approximate, however, we feel that it may provide a basis for the rational design of CDC2 inhibitors, such as correctly cyclized peptides.

ACKNOWLEDGMENTS

The authors thank Professors Arturo Falaschi (ICGEB, Trieste) and Tim Hunt (MRC, Cambridge) for suggestions on the experimental design. p13^{src1} protein was a generous gift of Drs. Dénes Dudits and Ferenc Felföldi (Biological Research Centre, Szeged, Hungary). Cultured HeLa cells were kindly provided by Dr. Georgine Faulkner and Ms. Maria Elena Lopez (ICGEB, Trieste). Professor J. E. Allende (Universidad de Chile, Santiago de Chile), Dr. Lawrence Banks (ICGEB, Trieste), and Dr. Bence Asbóth (ABC, Gödöllő) gave useful advice regarding the manuscript. J.Z., R.S., and S.W. are recipients of an ICGEB fellowship. The authors thank Ms. Suzanne Kerbavcic for her help in preparing the manuscript.

REFERENCES

1. Norbury, C., and Nurse, P. (1992) *Annu. Rev. Biochem.* **61**, 441–470.
2. Forsburg, S. L., and Nurse, P. A. (1991) *Rev. Cell Biol.* **7**, 227–256.
3. Fang, F., and Newport, J. W. (1991) *Cell* **66**, 731–742.
4. Pagano, M., Pepperkok, R., Lukas, J., Baldin, V., Ansorge, W., Bartek, J., and Draetta, G. (1993) *J. Cell Biol.* **121**, 101–111.
5. Moreno, S., and Nurse, P. (1990) *Cell* **61**, 549–551.

6. Pines, J., and Hunter, T. (1990) *New Biol.* **2**, 389–401.
7. Kamijo, M., Yasuda, H., Yau, P. M., Yamashita, M., Nagahama, Y., and Ohba, Y. (1992) *Peptide Res.* **5**, 281–285.
8. Kitagawa, M., Okabe, T., Ogino, H., Matsumoto, H., Suzuki-Takahashi, I., Kokubo, T., Higashi, H., Saitoh, S., Taya, Y., Yasuda, H., Ohba, Y., Nishimura, S., Tanaka, N., and Okuyama, A. (1993) *Oncogene* **8**, 2425–2432.
9. Ando, S., Tsujimura, K., Matsuoka, Y., Tokui, T., Hisanaga, S., Okumura, E., Uchiyama, M., Kishimoto, T., Yasuda, H., and Inagaki, M. (1993) *Biochem. Biophys. Res. Commun.* **195**, 837–843.
10. Nurse, P. (1990) *Nature* **344**, 503–508.
11. Draetta, G. (1990) *Trends Biochem. Sci.* **15**, 378–383.
12. Nash, R., Tokiwa, G., Anand, S., Erikson, K., and Fletcher, A. B. (1988) *EMBO J.* **7**, 4335–4346.
13. Hadwiger, J. A., Wittenberg, C., Richardson, H. E., de Barros Lopes, M., and Reed, S. I. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 6255–6259.
14. Wang, J., Chenivess, X., Henglein, B., and Brechot, C. (1990) *Nature* **343**, 555–557.
15. Lew, J., Beaudette, K., Litwin, C., and Wang, J. H. (1992) *J. Biol. Chem.* **267**, 13383–13390.
16. Pagano, M., Pepperkok, R., Verde, F., Ansorge, W., and Draetta, G. (1992) *EMBO J.* **11**, 961–971.
17. Roosenblatt, J., Gu, Y., and Morgan, D. O. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 2824–2828.
18. Beaudette, K. N., Lew, J., and Wang, J. H. (1993) *J. Biol. Chem.* **268**, 20825–20830.
19. Bischoff, J. R., Friedman, P. N., Marshak, D. R., Preives, C., and Beach, D. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4766–4770.
20. Pines, J., and Hunter, T. (1989) *Cell* **58**, 833–846.
21. Cisek, L. J., and Corden, J. L. (1989) *Nature* **339**, 679–684.
22. Satterwhite, L., Cisek, L., Corden, J., and Pollard, T. (1990) *Ann. N.Y. Acad. Sci.* **582**, 307.
23. Matsuoka, Y., Nishizawa, K., Yano, T., Shibata, M., Ando, S., Takahashi, T., and Inagaki, M. (1992) *EMBO J.* **11**, 2895–2902.
24. Kusubata, M., Matsuoka, Y., Tsujimura, K., Ito, H., Ando, S., Kamijo, M., Yasuda, H., Ohba, Y., Okumura, E., Kishimoto, T., and Inagaki, M. (1993) *Biochem. Biophys. Res. Commun.* **190**, 927–934.
25. Pan, Z. Q., Amin, A., and Hurwitz, J. (1993) *J. Biol. Chem.* **268**, 20433–20442.
26. Mcleod, A. R., Wong, N. C. W., and Dixon, G. H. (1977) *Eur. J. Biochem.* **78**, 281–291.
27. McVey, D., Brizuela, L., Mohr, I., Marshak, D. R., Gluzuman, Y., and Beach, D. (1989) *Nature* **341**, 503–507.
28. Lin, B. T. Y., Gruenwald, S., Moral, A. O., Lee, W. H., and Wang, J. Y. J. (1991) *EMBO J.* **10**, 857–864.
29. Peter, M., Nakagawa, J., Doree, M., Labbe, J. C., and Nigg, E. A. (1990) *Cell* **60**, 791–801.
30. Fields, C. C., Lloyd, D. H., McDonald, R. L., Otteson, K. M., and Noble, R. L. (1991) *Peptide Res.* **4**, 95–101.
31. Sarin, V. K., Kent, S. B. H., Tam, J. P., and Merrifield, R. B. (1981) *Anal. Biochem.* **117**, 147–157.
32. Hurley, C. K. (1977) *Anal. Biochem.* **80**, 624–626.
33. Labbe, J. C., Cavadore, J. C., and Doree, M. (1991) *Methods Enzymol.* **200**, 291–301.
34. Biosym (1993) Discover User Guide, version 2.9/3.1. Biosym Technologies, San Diego.
35. Dauber-Osguthorpe, P., Roberts, V. A., Osguthorpe, D. J., Wolff, J., Genest, M., and Hagler, A. T. (1988) *Proteins* **4**, 31–47.
36. Suzuki, M. (1989) *J. Mol. Biol.* **207**, 61–84.
37. Suzuki, M. (1989) *EMBO J.* **8**, 797–804.
38. Suzuki, M., Gerstein, M., and Johnson, T. (1993) *Prot. Eng.* **6**, 565–574.
39. Clark, K. L., Halay, E. D., Lai, E., and Burley, S. K. (1993) *Nature* **364**, 412–420.
40. Richards, F., and Kundrot, C. (1988) *Proteins* **3**, 71–84.
41. Brizuela, L., Draetta, G., and Beach, D. (1987) *EMBO J.* **6**, 3507–3514.
42. Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L., and Sweet, R. M. (1993) *Nature* **362**, 219–223.
43. Freidinger, R. M., Perlow, D. S., Randall, W. C., Saperstein, R., Arison, B. H., and Veber, D. F. (1984) *Int. J. Pept. Protein Res.* **23**, 142–50.
44. Mariano, O., Donella-Deana, A., Brunati, A. M., Fischer, S., and Pinna, L. A. (1991) *J. Biol. Chem.* **266**, 17798–17803.
45. Garnier, J., Osguthorpe, D. J., and Robson, B. (1978) *J. Mol. Biol.* **120**, 97–120.
46. Matsoukas, J. M., and Moore, G. J. (1986) *Arch. Biochem. Biophys.* **248**, 419–423.
47. Martensen, T. M. (1984) *Methods Enzymol.* **107**, 3–23.
48. Knighton, D. R., Zheng, J., Ten Eyck, L. F., Ashford, V. A., Xuong, N.-H., Taylor, S. S., and Sowadski, J. M. (1991) *Science* **253**, 407–414.
49. De Bondt, H. L., Rosenblatt, J., Jancarik, J., Jones, H. D., Morgan, D. O., and Kim, S.-H. (1993) *Nature* **363**, 595–602.
50. Zhang, F., Strand, A., Robbins, D., Cobb, M., and Goldsmith, J. (1993) *Nature* **367**, 704–711.
51. Knighton, D. R., Zheng, J., Ten Eyck, L. F., Xuong, N.-H., Taylor, S. S., and Sowadski, J. M. (1991) *Science* **253**, 414–420.
52. Tuteja, N., Tuteja, R., Rahman, K., Kang, L. Y., and Falaschi, A. (1990) *Nucleic Acids Res.* **18**, 6785–6792.
53. Maller, J. E., Kemp, B. E., and Krebs, E. G. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 248–251.
54. Erikson, E., and Maller, J. L. (1991) *J. Biol. Chem.* **266**, 5294–5255.
55. Hathaway, G. M., and Traugh, J. A. (1983) *Methods Enzymol.* **99**, 317–331.
56. Cheng, H. C., Kemp, B. E., Pearson, R. B., Smith, A. J., Misconi, L., Van Patten, S. M., and Walsh, D. A. (1986) *J. Biol. Chem.* **261**, 989–992.
57. Small, D., Chou, P. Y., and Fassman, G. D. (1977) *Biochem. Biophys. Res. Commun.* **79**, 341–346.

CDC2 kinase preparation

Cyclin-B/CDC2 kinase was purified from HeLa cells by a procedure similar to that used by Kitagawa et al. (8) and Pan et al. (25) and was followed by *suc1* affinity chromatography (33) and heparin chromatography. A synthetic peptide (p-0 in Table 1C) was used to monitor the purification.

The steps of purification (summarised in Table S1) were as follows: Step 1: Crude extracts of HeLa cells were obtained as described previously (52). 33 g of frozen cell material (corresponding to about 10^{10} cells) were thawed and lysed by homogenising in 100 ml of buffer A (25mM sodium phosphate, pH 7.0, 5mM NaF, 2.5mM MgCl₂, 0.5mM EDTA, 5mM β-glycerophosphate, 35mM NaCl, 0.5mM EGTA, 25mg/L of PMSF). The homogenate was clarified through centrifugation at 30,000g for 45 min, and the supernatant was passed through a glass fibre column. Step 2: 110 ml of the crude extract was loaded to a DEAE Sepharose CL-6B (Pharmacia) column (1.6x25 cm) equilibrated with buffer A. The column was washed with 150 ml of 100mM NaCl in buffer A and then the kinase activities were eluted by increasing the concentration of NaCl to 200mM. Step 3: Ammonium sulphate was added to the pooled 100-200mM NaCl fractions from the DEAE column and brought to 55% of the saturated concentration in ice. The precipitate was collected by centrifugation at 30,000 g for 30 min at 4°C, resuspended in 5 ml of buffer A and then applied to a S300 Sephacryl HR column (Pharmacia, 1.5x100cm) pre-equilibrated with buffer A. At this gel filtration step, histones H1 kinase activities were separated into two sharp peaks: S1 and S2, (eluting at 100 and 150 ml, respectively). Step 4: The S1 pool (30 ml) was applied to a 5ml Econo-Pac Heparin Cartridge (Bio Rad) equilibrated with buffer A which was then washed with 50 ml of 100mM NaCl in buffer A. The *cdc2* kinase activity was eluted with a solution containing 200mM NaCl in buffer B (25 mM Tris-HCl, pH 7.5, 1 mM DTT, 0.2 mM PMSF, 0.5 mM EGTA, 5 mM β-glycerophosphate, 0.2 mg/L antipain, 0.1 mg/L leupeptin, 0.02 % Nonidet P40, 20 % glycerol). Step 5: The pooled (10 ml) eluate was diluted 5 times with buffer B

and immediately subjected to a 2ml Q-Sepharose column (Pharmacia). After washing the column with 150mM NaCl in buffer B, the kinase activity was eluted by 250 mM NaCl in buffer B. Step 6: Active fractions (10 ml) were pooled, diluted 5 times with buffer B and loaded onto a 2ml S-Sepharose column (Pharmacia). After washing the column with 50 mM NaCl in Buffer B, the kinase activity was eluted with 200 mM NaCl in Buffer B. Step 7: The active pool of S-Sepharose (10 ml) was diluted 5 times and applied to a 1.5 ml casein-Sepharose column (Sigma). After washing the column with 50 ml of 150 mM NaCl in buffer B, kinase activity was eluted with 0.3 M NaCl in buffer B. Step 8: p13^{suc1} protein, a gift from Drs. D. Dudits and F. Fellöldi, was purified by FPLC before use and conjugated to CNBr-activated Sepharose 4B (Pharmacia). 3.5 mg of protein/ml gel was used in the reaction, which was carried out as described by the manufacturer. The eluate of the previous step (5 ml) was directly applied to the 1.5 ml *suc1* affinity column pre-equilibrated with 300 mM NaCl in buffer B. The column was successively washed with 50 ml of 500 mM NaCl in buffer B and then with 50 ml of buffer B. The kinase activity was eluted from the column with 10 ml of *suc1* protein solution (2mg/ml in buffer B). Step 9: In order to eliminate high concentrations of *suc1* from the kinase preparation, the eluent of the previous step was subjected to a final heparin chromatography (see Step 4) and *cdc2* kinase was recovered from the column by eluting with 200mM NaCl in buffer B.

SDS gel electrophoresis of the purified material showed bands corresponding to p^{34cdc2}, cyclin B and *suc1* (Figure S1). Western blotting analysis gave positive results with anti-P34 *cdc2* (ICN, Biomedical Inc., Irvine, CA, Cat. No. 69-900) and with anti human cyclin-B monoclonal antibodies (Upstate Biotechnology Incorporated, Lake Placid, New York, Cat. No. 14-103). No hybridisation was obtained with anti-human cyclin-A monoclonals (Upstate Biotechnology Incorporated, Cat. No. 05-115), showing that enzyme was cyclin-B/p^{34cdc2} complex (data not shown). The enzyme preparation was tested with a number of substrates/inhibitors of other kinases and the results (Table S2) did not show the presence of contaminating kinase activities.

Table S1. Purification of cyclin-b/p^{34cdc2} from HeLa cells

No	Step	Total protein ¹ (mg)	Total activity (cpm x 10 ⁻⁶)	Recovery (%)	Specific activity (cpm x 10 ⁻⁶)	Purification (fold)
1	Supernatant ²	812	55.2	100	0.066	1
2	DEAE-Sepharose	43	74.25	134.5	1.73	26
3	S300-Sephacryl	18	37.5	67.9	2.08	31
4	Heparin-I	2.5	29.8	54	11.92	180
5	Q-Sepharose	1	20	36.2	20	303
6	S-Sepharose	0.6	15	27.2	25	378
7	Casein-Sepharose	0.2	9	16.3	35	530
8	p13 ^{suc1} -Sepharose ³	-	12.7	-	-	-
9	Heparin-II	0.006	0.9	1.6	150	2272

¹Protein concentration was determined by Micro BCA kit (Pierce, cat. no. 23235, USA) except in the last purification step (Heparin II) where the quantity was estimated from silver stained SDS gels (41).

²The apparent activity of the crude extract is lower than what can be expected on the basis of the later fractions. This is probably due to inhibitors present in the crude extract, as suggested by Clisek and Cordan (21).

³The amount of p^{34cdc2} could not be determined at this step due to the high concentration of *suc1* present.

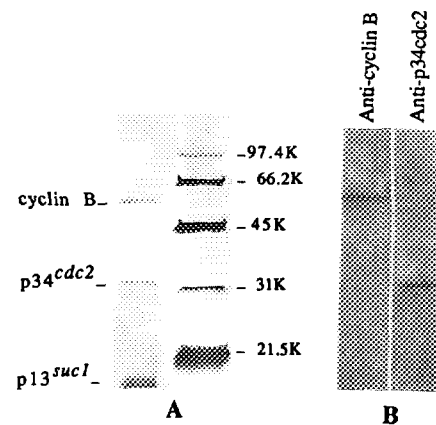


Figure S1. Characterisation of the cyclin-B/*cdc2* kinase preparation. A: Silver staining gel of purified *cdc2*/cyclin B complex after Heparin II. The bands corresponding to cyclin B/p^{34cdc2} and p13^{suc1} (eluted in the front) as well as the molecular mass markers are indicated. B: Immunoblotting of purified *cdc2*/cyclin B complex after Heparin II. Monoclonal antibodies against human p^{34cdc2} and human cyclin-B1 were from ICN (Cat. No. 69-900) and UBI (Cat. No. 14-103) respectively, and detected with 35S-protein A (Amersham, Cat. SJ 444) after autoradiography.

Table S2. K_m and IC_{50} of cyclin-B/cdc2 for different compounds

Compound	K_m	IC_{50}
Histone H1	1.2 μ M	
p-1 ¹	3.5 μ M	
Casein	No phosphorylation	
Kemptide ²	No phosphorylation	
ZnCl ₂ ³		1.3 mM
Heparin ⁴		No effect to 1 μ M
PKI peptide ⁵		No effect to 1 μ g/reaction
p-nitrophenyl phosphate ⁶		No effect to 80 mM
Inositol hexasulfate ⁷		No effect to 800 μ M

¹Sequence shown in Table 1.

²Kemptide (LRRASLG, Sigma, cat # K 1127) is the substrate of cAMP-dependent kinase (53).

³ZnCl₂ is an inhibitor of cdc2 kinase with an IC_{50} of 1 mM (33)

⁴Heparin is an inhibitor of S6 kinase (54) and of casein kinase II (55)

⁵PKI peptide (TTYADFIASGRTGRRHAIHD, Sigma, cat # P 3294) is and inhibitor of cAMP-dependent kinase (56)

⁶p-nitrophenyl phosphate is the inhibitor of S6 kinase (54)

⁷Inositol hexasulphate is an inhibitor of casein kinase II (55)