Utilization of Oriented Peptide Libraries to Identify Substrate Motifs Selected by ATM*

Received for publication, February 7, 2000, and in revised form, April 26, 2000
Published, JBC Papers in Press, May 2, 2000, DOI 10.1074/jbc.M001002200

Ted O’Neill,a,b,c Alison J. Dwyer,a,b,d Yael Ziv,e Doug W. Chan,f Susan P. Lees-Miller,f Robert H. Abraham,g Jack H. Lai,h David Hill,i Yossi Shiloh,j Lewis C. Cantley,g,d and Gary A. Rathbuna,b,h

From the aCenter for Blood Research, Department of Pediatrics, Children’s Hospital, Harvard Medical School, Boston, Massachusetts 02115, the bDepartment of Human Genetics and Molecular Medicine, Sackler School of Medicine, Tel Aviv University, Ramat Aviv 69978, Israel, the cDepartment of Biological Chemistry, Department of Biological Sciences, University of Calgary, Calgary AB T2N 1 N4, Canada, the dDepartment of Pharmacology and Cancer Biology, Duke University, Durham, North Carolina 27710, the eDepartment of Medicine, Beth Israel Deaconess Hospital, Harvard Medical School, Boston, Massachusetts 02115, and the fOncogene Science, Inc., Cambridge, Massachusetts 02142

The ataxia telangiectasia mutated (ATM) gene encodes a serine/threonine protein kinase that plays a critical role in genomic surveillance and development. Here, we use a peptide library approach to define the in vitro substrate specificity of ATM kinase activity. The peptide library analysis identified an optimal sequence with a central core motif of LSQE that is preferentially phosphorylated by ATM. The contributions of the amino acids surrounding serine in the LSQE motif were assessed by utilizing specific peptide libraries or individual peptide substrates. All amino acids comprising the LSQE sequence were critical for maximum peptide substrate suitability for ATM. The DNA-dependent protein kinase (DNA-PK), a Ser/Thr kinase related to ATM and important in DNA repair, was compared with ATM in terms of peptide substrate selectivity. DNA-PK was found to be unique in its preference of neighboring amino acids to the phosphorylated serine. Peptide library analyses defined a preferred amino acid motif for ATM that permits clear distinctions between ATM and DNA-PK kinase activity. Data base searches using the library-derived ATM sequence identified previously characterized substrates of ATM, as well as novel candidate substrate targets that may function downstream in ATM-directed signaling pathways.

Cells respond to DNA damage by activating specific signaling pathways that culminate into cell cycle checkpoints, in which cell cycle progression is arrested and DNA repair is effected (1). These checkpoints are regulated, in part, by members of the phosphoinositide kinase (PIK) family of high molecular weight proteins, with homologues in yeast, Drosophila, and mammals (2, 3, 16). The mammalian PIK representatives include ATM, ataxia telangiectasia gelated (ATR), and mammalian target of rapamycin (mTOR, also known as FRAP) (4). The catalytic subunit of DNA-PK (DNA-PKcs) is a mammalian PIK family member important for DNA double strand break repair (5, 6). ATM, ATR, DNA-PKcs, and mTOR have all been shown to express a serine/threonine protein kinase activity (3–5, 7–9). DNA-PKcs and ATM have been shown to associate with DNA (6–8). For DNA-PKcs, binding to double strand DNA is essential for its catalytic activity, whereas ATM kinase activity appears mildly enhanced for certain protein substrate targets in the presence of single- and double-stranded DNA. PIK family members in lower eukaryotes with homology to mammalian counterparts in both sequence and function include Mei-41 (Drosophila) and yeast Mec1p, Rad3p, Tel1p, Tor1p, and Tor2p (9). Genetic analyses of yeast PIK-related genes have in large part established their various roles as cell cycle checkpoint genes, and defined signaling pathways that are modulated by these gene products (10).

ATM functional loss in humans results in ataxia telangiectasia (A-T) (11), an autosomal recessive, pleiotrophic disorder characterized by a progressive cerebellar degeneration, immunodeficiencies, chromosomal instability, sensitivity to ionizing radiation, sterility, defects in cell-mediated immunity, and telangiectases of the eyes, and sometimes, ears, face, and hands (12). A-T patients exhibit a high predisposition for malignancies, predominantly lymphoid in nature, and A-T-derived cell lines demonstrate cell cycle checkpoint dysfunction in G1/S and G2/M in response to DNA-damaging agents (12, 13). Targeted mutations of ATM in mice recapitulate many aspects of the human phenotype; however, a major difference is that the mutant mice exhibit a minimal or mild neurogenic phenotype (14–16).

Naturally occurring DNA-PKcs mutations have been characterized in mice and Arabian foals (17–19); in both species, these mutations result in a severe combined immunodeficiency syndrome (Scid). The Scid phenotype results from loss of efficient DNA repair in V(D)J recombination reactions (20–22). In mice, DNA-PK deficiencies result in a comparatively milder predisposition for lymphoid malignancies than is observed in ATM−/− animals. Although Scid-derived cells demonstrate sensitivity to ionizing radiation, these cells exhibit no obvious cell cycle checkpoint defects (23). The functional loss phenotypes suggest fundamental differences in the downstream substrates/effectors of DNA-PKcs and ATM.
Studies investigating localization have found ATM in both the nucleus and cytoplasm, and the extent of ATM in these locations appears cell-specific. Thus, in proliferating lymphoblastoid cells, ATM is predominantly in the nucleus and less in the cytoplasm, whereas in post-mitotic Purkinje cells, ATM resides almost exclusively in the cytoplasm, suggesting intriguing differences in ATM function and tissue-specific distinctions in ATM-directed signaling pathways (24–27). Since ATM has been shown to be in several locales, this complicates a unifying characterization of ATM-signaling pathways, except for a general dependence on protein kinase activity likely required in most cases (28). The multiple, wide ranging disorders resulting from ATM loss of function hint at the impact of ATM upon a multiplicity of biological systems. Some of these disorders, such as development of lymphoid malignancies and Purkinje cell degeneration (in human), on the surface appear only remotely connected. However, the consequences of ATM function, or loss of such even in quite different tissues, may be ultimately related, at least in part, through appropriate regulation of decisions pertaining to cell survival or cell death.

Downstream targets of ATM kinase activity that partially delineate DNA damage-activated cell cycle checkpoint signaling pathways have been recently described. The best characterized of these checkpoint pathways involve p53 and Chk2/Cds1, as internally critical components. A DNA damage-induced ATM-dependent p53 activation pathway has been shown to involve direct phosphorylation of Ser15 in p53 by ATM (29–32). ATM phosphorylation and apparent ATM-directed de-phosphorylation of certain serines in the C-terminal region of p53 has been shown to modulate multiple aspects of p53 activity including expression, DNA binding, transcriptional activation, and negative regulation (29–33). Cds1/Chk2 is modified in p53 signaling, and negative regulation (29–33). ATM phosphorylation and apparent ATM-directed dephosphorylation has been shown to be dependent on the presence of ATM (34–36). Several other target substrates for ATM have been described, for example, c-Abl, RPA, and BRCA1 (6, 7, 23, 37–39), but the participating signaling pathways are less clear.

We have employed an oriented degenerate peptide library approach to determine an optimal substrate motif for ATM protein kinase activity. This motif was used as a probe to distinguish ATM and DNA-PK catalytic activities. The peptide library-derived substrate sequence was also utilized in data base searches that identified both previously defined substrates as well as potential downstream targets of ATM.

**EXPERIMENTAL PROCEDURES**

**Cells**—G361, a human melanoma cell line that produces a high level of ATM protein, and C3ABR and L6, two EBV-transformed human B cell lines, have been described previously (29). All were grown in RPMI containing 10% heat-inactivated fetal calf serum (Atlanta Biologicals), 200 μM l-Glu and penicillin/streptomycin.

**Peptide Libraries**—The following degenerate peptide libraries were used as substrates in this study: SF, MAXXXXSFXXXXAKK; SI, MAXXXXSIXXXAKK; SK, MAXXXSISXXXXXXAKK; SX, MAXXXXRMXXXXXXAKK; SQ, MAXXXXSXXXXXXAKK; SX, MAXXXSISXXXXXXAKK; T, MAXTTTXXXAKK; 4Y +/− 1, MAXXXXSXXXXXXAKK (the + and − designations denote the presence or absence of Ser and Thr in the degenerate positions). The peptide library also contains Ser, Thr, and Tyr at degenerate positions. The other Ser-based libraries contained all amino acids but Cys, Trp, Ser, Thr, and Tyr. The Tyr-based libraries contained all amino acids but Cys, Trp, and Tyr, and the + or − signs designate the presence of Ser and Thr. The SQII library is comprised of the following: MA (G,H,R,A)−(A,V,F,M,L,Q,H,A,R,F,P,M,Y,A,V,F,M,L,Y,A,R,F,H,Q,N,Y,N,F,A)−SQ(Q,S,R,F,P,E,A,Y,F,P,S,V,Q(H,A,Q,F,L,H,A,N) AKKK. Amino acids at each position are in the parentheses.

**ATM Protein Kinase Assays, Peptide Libraries—**ATM Protein kinase assays were performed as described (31). Cell pellets per 150-cm plate containing 1.0−2.0 × 10^7 cells were harvested when cell growth was approximately 90% confluent for G361 melanoma cell lines. Cells (7.5 × 10^5−1.0 × 10^6), harvested from 100 ml C3ABR or L6, were washed twice in ice-cold phosphate-buffered saline, transferred to 1.5-ml Eppendorf microcentrifuge tubes, and microcentrifuged three times at 3000 rpm, 5 min. The supernatant was discarded, and the cell pellets were either used immediately or stored at −80 °C prior to use. Previously frozen cells were best for optimal ATM kinase activity, and pellets that had been stored for up to 6 months were usable in these assays. Cell extracts, ranging between 1.5 and 2.0 mg of protein, were generated as described above and immediately as storage of extracts usually resulted in a substantial drop on ATM kinase activity. To the Lysis, High Salt, and Base buffers (see below and Ref. 31), 1 mM each of NaF and NaVO_3 were added as phosphatase inhibitors. Generally, 20 μl of spent media, containing the monoclonal anti-ATM 10GG1-producing hybridoma directed against a peptide to positions 819–844 in ATM (29), was added to 1.5−2.0 mg of total lysate and rocked gently at 4 °C for 3 h. Sheep anti-mouse IgG complexed to magnetic beads (100 μl/sample) (Dynabeads, Dynal Corp.) was washed three times with Lysis buffer using a Magnetic Particle Concentrator (Dynal), reconstituted with 100 μl of Lysis buffer, added to the antibody/lysate mixture, and subsequently incubated, rocking gently, at 4 °C for 1.5 h.

The ATP-directed precipitation period with the anti-mouse Ig-containing magnetic beads, the supernatant was discarded, and the beads were washed three times with Lysis buffer. A High Salt buffer was added to the metal beads followed by incubation on ice for 5 min to de-activate DNA-PK (31). The High Salt buffer was removed, and the beads were washed once with Buffer (10 mM MgCl_2, 150 mM NaCl, 100 mM HEPES, pH 7.5) and once with Kinase buffer containing both 10 mM MgCl_2 and 10 mM MnCl_2. From these steps, 120 μM cold ATP, 2 mM dithiothreitol, 50 μM microcystin (Alexis Corp.), 4 μg/ml each of leupeptin, pepstatin, and 4-(2-aminoethyl)benzenesulfonyl fluoride (Sigma) and 10 μCi of [γ-^32P]ATP (NEC Life Science Products). (Initial experiments depicted in Fig. 1A only were carried out with approximately 100 μCi of [γ-^32P]ATP.) All reagents utilized for washes contained NaF and NaVO_3 except for the Kinase buffer that contained microcystin. Samples were prepared and kept on ice prior to incubation at 30 °C. Reactions were carried out in a total volume of 30 μl of kinase buffer added to the magnetic beads. Where indicated, wortmannin was added, usually at 0.5 or 1 μM final concentration, to samples that had been washed three times in Lysis buffer and once in Base buffer and were brought up to volume (usually 50 μl) in Kinase buffer. The samples were allowed to react with wortmannin by incubation at room temperature for 20 min. This room temperature preincubation stage resulted in some net loss in overall ATM activity in the absence of wortmannin. Wortmannin inhibition persisted for at least 16−18 h without significant loss of inhibiting activity. Phosphorylation assays were measured by removing 2−3 μl from kinase reactions at given time points and adding these aliquots to an equal volume of 30% acetic acid to stop the reaction. Aliquots (3.5−4.0 μl) of the halted reactions were then spotted on p81 phosphocellulose paper, allowed to dry, washed three times in 1% orthophosphoric acid (5−10 min/wash), and counted.

**DNA-PK Protein Kinase Assays—**Protein kinase activity of DNA-PK was assayed using purified DNA-PKcs, and Ku70 and Ku80 were prepared as described previously (40). Briefly, the kinase reaction was carried out in the presence of DNA-PKcs (120 ng), Ku80/Ku70 (30 ng), and a final concentration of 100 μM ATP, 10 μM of [γ-^32P]ATP, 10 mM MgCl_2, or MnCl_2, 200 ng of sonicated calf thymus DNA, 50 mM cold HEPES, pH 7.5, 1 mM dithiothreitol, 20 mM cold ATP and calculated to phosphorylate approximately 1% of the peptides in the library (41).
kinase reaction (300 \(\mu\)l) was terminated by adding an equal volume of 30% acetic acid and then desalted and purified over a 1-ml DEAE-Sephacel (Sigma) column to remove excess ATP. The void volume was usually discarded, and the samples, monitored through the presence of radioisotope, were lyophilized. Lyophilized samples were reconstituted in distilled, sterile H2O and applied to a ferric chelation nitritotriacetic acid column as described (41). Eluted samples were pooled and sequenced. Each cycle was normalized internally to all amino acids included at that position, and amino acid selectivity at a given cycle was assessed using a program developed specifically for that purpose.3 The SQ and SQII libraries were each analyzed in independent experiments at least twice.

RESULTS

Preferential Phosphorylation of a Fixed, Degenerate Peptide Library by ATM—Immunoprecipitated ATM has been shown to express a Ser/Thr protein kinase activity that phosphorylates certain protein substrates (6, 29–31). To establish an initial priority for peptide substrate targets, we compared two degenerate peptide libraries as substrates for immunoprecipitated ATM protein kinase. These libraries contained four degenerate positions on both the N- and C-terminal sides of a fixed serine or threonine and were designated the Ser and Thr libraries, respectively. Our analysis showed that the Ser degenerate library was phosphorylated more than the Thr library (Fig. 1A).

Due to the preferred Ser selection, we proceeded to screen additional second generation degenerate peptide libraries, several of which contained a second fixed residue at the +1-position (C-terminal to Ser). These analyses revealed a strong preference for the SQ library that exceeded a second tier of selected libraries severalfold (Fig. 1B). Libraries with hydrophobic amino acids at the +1-position (SI and SF) were somewhat preferred over the Ser library where the +1-position was degenerate. Arginine at position −3 (the RXXS library) was selected equivalently to the SI and SF libraries. Proline at the +1-position resulted in minimal phosphorylation (Fig. 1B).

ATM showed little or no phosphorylation of a library with a central, fixed tyrosine containing no Ser or Thr (Fig. 1B). Strikingly, the 4Y4+ library, which does contain Ser and Thr in the degenerate positions, was selected equivalently to the SF library, consistent with a preference by ATM for hydrophobic aromatic residues at the +1-position or elsewhere within the sequence. We found that utilization of Mn\(^{2+}\) versus Mg\(^{2+}\) in the kinase reaction was essential for efficient phosphorylation of the SF, SI, and 4Y4+ libraries (data not shown). The SQ library was more highly phosphorylated in the presence of Mn\(^{2+}\), although significant phosphorylation still occurred in Mg\(^{2+}\). On the basis of the results in Fig. 1B, the SQ peptide library was

3 J. H. Lai and L. C. Cantley, unpublished observations.
Amino acid selections at positions N- and C-terminal to the phosho-Ser in peptides preferentially phosphorylated by ATM

The SQ library used had the composition of MAXXXXXXXXSQXXXXAKKKK, where X stands for all residues except Cys and Trp. Approximately 1% of the library was phosphorylated by ATM. The phosphopeptides were separated from the non-phosphorylated peptides and sequenced as a batch by Edman degradation. Residues exhibiting the strongest preference values are shown. The amino acid selection values at each position were calculated as described (64); based on these calculations, 1.00 is defined as the normalized base-line value; residues with preference values at 1.00 or below are considered non-selected (see Ref. 64). The SQII library was biased toward amino acids selected from the SQ library and significantly less degenerate. Composition of the SQII library is MA(G,H,R,A)(A,V,F,M,L,Q)(H,A,R)(F,M,Y,A)(Y,F,M,L,V,A)(A,R,P,H,Q,N,L,Y,N,F,A) SQ)(Q,S,R,P,V,E,A,L,F,P,S,V,Q)(A,Q,F,L,H,N) AKKK. Amino acids at each position are in parentheses.

### Table I

| Position | X | G | S | R | P | V | E | A |
|----------|---|---|---|---|---|---|---|---|
| 2        | Gly 1.66 Phe 1.46 | His 1.06 Met 1.18 Ala 0.99 | Tyr 2.27 Phe 1.53 Gln 1.15 Tyr 2.21 Ser 1.05 | Val 1.05 Leu 1.42 Asn 1.05 | Leu 1.32 Phe 1.32 |
| 3        | Gln 1.15  |
| 4        | Met 1.16 Leu 1.27  |
| 5        | Asn 1.20  |

### ATMT SQ library selection

Position -7 Position -6 Position -5 Position -4 Position -3 Position -2 Position -1 Position +1 Position +2 Position +3 Position +4 Position +5

**Starting library**

| Starting library | MAXXXXXXXXSQXXXXAKKKK |
|------------------|------------------------|
| Gly              | Phe 1.46               |
| His              | 1.06                   |
| Met              | 1.18                   |
| Ala              | 0.99                   |
| Tyr              | 2.27                   |
| Phe              | 1.53                   |
| Gln              | 1.15                   |
| Tyr              | 2.21                   |
| Ser              | 1.05                   |
| Val              | 1.05                   |
| Leu              | 1.42                   |
| Asn              | 1.05                   |
| Leu              | 1.32                   |
| Phe              | 1.32                   |
| Ser              | 1.05                   |
| Gln              | 2.27                   |
| Glu              | 2.33                   |
| Phe              | 1.33                   |
| X                 |

### ATMT SQII library selection

Position 1 Position 2 Position 3 Position 4 Position 5

| Position | Leu 1.61 Met 1.44 Ala 1.85 Gln 1.20 Met 1.19 Gln 1.62 Leu 1.86 Ser 1.06 Glu 2.07 Glu 1.65 Phe 1.18 Pro 1.08 |
|----------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| Gly      | 1.66             | Phe 1.02         | Met 1.01         | Met 1.03         |
| Leu      | 1.44             | Phe 1.10         | Phe 1.17         | Phe 1.15         |
| Ala      | 1.85             | Gln 1.20         | Met 1.19         | Gln 1.62         |
| Gln      | 1.62             | Leu 1.86         | Ser 1.06         | Glu 2.07         |
| Met      | 1.19             | Gln 1.62         | Leu 1.86         | Glu 1.65         |
| Phe      | 1.02             | Phe 1.17         | Met 1.19         | Phe 1.18         |
| Gln      | 2.07             | Glu 1.65         | Leu 1.86         | Pro 1.08         |
| Phe      | 1.18             | Pro 1.08         |

### Predicted ATMT optimal substrate

G L/F/M A Q/F/M M/F Q/P L S Q E/Q V/Q/P F A

Sequence of ATMT phosphorylation site in p53

P S V E P P L S Q E Q E T F S

selected for further analyses.

*Phosphorylation of the SQ Library Depends on the ATM Protein and Its Catalytic Activity*—The C3ABR and L6 EBV-transformed human B cell lines represent cells with wild-type ATM and mutant (null) ATM, respectively; the latter contains a stop codon at position 35 in ATM (42). As expected, this library was highly phosphorylated in ATM immunoprecipitation assays of protein extracts derived from C3ABR, whereas immunoprecipitates from L6 demonstrated substantially less activity (Fig. 1C). Previous studies have shown that ATM is highly sensitive to wortmannin, a fungal herbicide inhibitor (9, 29, 30, 34). Fig. 1C shows that 1 μM wortmannin severely diminished phosphorylation of the SQ library. We conclude from these analyses that the kinase in the immunoprecipitates responsible for phosphorylation of the SQ library is ATM.

*Determination of an Optimal Motif for ATMT Protein Kinase Activity*—Large scale preparative peptide library analyses were undertaken to isolate the peptides in the SQ library that were selectively phosphorylated by ATM. The phosphopeptides were eluted from a ferric chelation column, as described (41), pooled and sequenced. Table I summarizes the amino acids selected for further analyses.

![Figure 2. Increased phosphorylation of successively more mature fixed SQ peptide libraries.](image)

The library-derived optimal peptide sequence at the major ATM phosphorylation site in p53. In particular, both the optimal substrate sequence and p53 share a core motif of LSQE (Table I). The library-derived optimal motif served as the reference for synthesis of individual peptides utilized for further analyses (see below).

*Evaluation of Predicted Optimal Peptide Substrates of ATMT*—We compared the optimal peptide generated from the library analysis, designated ATide, versus a peptide derived from the N-terminal sequence of p53 (p53-WT) (see Table II). These two peptides both contained the core LSQE sequence and demonstrated nearly identical $K_m$ and $V_{max}$ values (Table III).
TABLE II

Sequences of the ATide and p53 peptides

ATide peptides are derived from the optimal sequence derived from peptide library analyses (see Table I) and used for the analyses shown in Tables II and III. The p53-WT peptide contains the sequences surrounding Ser15 position in p53. For purposes of simplicity, the Ser residues at positions 9 and 20 and Thr at position 18 in the native sequence was changed to Ala (5).

| Peptide       | Sequence                        |
|---------------|---------------------------------|
| ATide         | AQMQLSQEVEFN GGKKK             |
| ATide Leu-1 → Ala | AQPASQSEVEFN GGKKK            |
| ATide Glu-2 → Ala | AQPASQAVFN GGKKK            |
| P53-WT        | AEPLSQAFA GGKKK                |
| P53 Leu-1 → Ala | AEPPASQAFA GGKKK              |
| P53 Glu-2 → Ala | AEPLSQQAFA GGKKK              |

TABLE III

$V_{max}$ and $K_m$ values of ATide and p53 peptides

ATide and p53 peptides were assayed at several concentrations ranging from 10 μM to 2 mM in immunoprecipitated ATM kinase assays to derive $K_m$ and $V_{max}$ values.

| Peptide       | $V_{max}$ (μM) | $K_m$ (μM) | $V_{max}/K_m$ (μM) |
|---------------|----------------|------------|--------------------|
| ATide         | 1240           | 23         | 54.0               |
| ATide Leu-1 → Ala | 1060           | 34         | 31.2               |
| ATide Glu-2 → Ala | 1285           | 305        | 4.2                |
| P53-WT        | 1270           | 22         | 57.7               |
| P53 Leu-1 → Ala | 1290           | 55         | 23.5               |
| P53 Glu-2 → Ala | 1195           | 90         | 13.3               |

a The estimated values for DNA-PK from phosphorylation kinetics were determined for each of the tested peptides (see text).

To determine the importance of the Leu at -1 and Glu at +2 for substrate utilization, these residues were changed to Ala in the context of the ATide and p53 sequences. In both cases, conversion of the Leu-1 to Ala caused a 1.7- and 2.4-fold decrease, respectively, in the $V_{max}/K_m$ values due to an increase in the $K_m$. Conversion of Glu+2 to Ala (Table II) had a more severe effect, resulting in an approximate 10-fold increased $K_m$ of the ATide sequence and a 4-fold increase in the $K_m$ of the p53 sequence (Table III). These results are consistent with the peptide library experiment where Glu at +2 exhibited a higher selectivity score than Leu at -1 (Table I).

Comparison of DNA-PK to ATM with Regard to Peptide Substrate Preference—The DNA-PK heterotrimeric complex has been shown to actively phosphorylate a multitude of target substrates; however, it remains unclear which, if any, of these targets have true physiological relevance. A central SQ motif has been previously determined to be an important sequence for DNA-PK phosphorylation (5, 43). We screened several degenerate peptide libraries for selection in terms of DNA-PK kinase activity. We found that the library preferences by DNA-PK were similar to that of ATM in that the SQ library was by far the most highly phosphorylated of the degenerate peptide libraries (Fig. 3 and data not shown). DNA-PK also phosphorylated the SI, SF, and 4Y+ libraries better in Mn$^{2+}$ than in Mg$^{2+}$ with the interesting exception of the SQ library, the latter of which was more highly phosphorylated in the presence of Mg$^{2+}$ (Fig. 3 and data not shown). Phosphorylation of the 4Y+ library was presumably on Ser/Thr residues since the same library lacking these residues (4Y-) was not phosphorylated (Fig. 3). The modest preference of DNA-PK for the SP library was equivalent in the presence of either cation.

Although these results indicate similarities between DNA-PK and ATM in peptide substrate preferences, there were several significant differences. ATM phosphorylated the SQII library about 30% better than the SQ library (see Fig. 2). By contrast, there was an approximately 350% increase in the phosphorylation of the SQ library versus the significantly less degenerate SQII library by DNA-PK (Fig. 3). The difference between the two kinases was also reflected in their relative preferences for the various ATide and p53 peptides. Although these peptides were readily phosphorylated by DNA-PK, the $K_m$ values in general were apparently much higher than observed with ATM and could not be readily determined; we were unable to reach a $V_{max}$ value (i.e. saturate DNA-PK catalytic activity) with the peptides used in this study. Table IV presents the relative ratio of peptide phosphorylation by DNA-PK at a single peptide concentration (100 μM). Of interest, the Leu-1 → Ala substitution in the context of either the ATide or p53 peptides resulted in an increase rather than decrease in phosphorylation rate, indicating that DNA-PK differs from ATM in amino acid selections at the -1-position. In contrast to ATM, there was a clear preference for the ATide peptide versus the p53-WT peptide by DNA-PK (Tables III and IV). Similar to ATM, the +2-position also plays a critical role in substrate preference in that the Glu+2 → Ala substitution severely diminished DNA-PK phosphorylation in both the ATide and p53 peptides. Calculated $V_{max}/K_m$ values for DNA-PK were in complete agreement with the normalized data presented in Table IV (see also Table III).

Silver stain analyses show an approximate level of 200–250 ng of immunoprecipitated ATM in G361 cells in a typical kinase reaction (data not shown). Utilization of an estimated value of 250 ng per reaction for ATM allowed us to compare the relative velocities of ATM and DNA-PK (the latter previously determined at 120 ng per reaction.4 Phosphorylation of Atide peptide (100 μM) by the two PIK kinase was assayed after 10 min at 30 °C. Under these conditions, our analysis revealed that ATM catalytic activity was about 78 pmol/min/μg and DNA-PK was 410 pmol/min/μg. Thus, under our specific assay conditions, DNA-PK appeared to demonstrate approximately

4 D. Chan and S. Lees-Miller, personal communication.
TABLE IV
Substrate Motifs Selected by ATM

| Peptide | Activity $^a$ |
|---------|---------------|
| ATide   | 100           |
| ATide Leu$^1$ → Ala | 157           |
| ATide Glu$^2$ → Ala | 18            |
| P53-WT  | 37            |
| P53 Leu$^1$ → Ala | 62            |
| P53 Glu$^2$ → Ala | 11            |

$^a$ Relative rate of DNA-PK activity is normalized to phosphorylation of the ATide peptide using 100 μM of each peptide substrate.

DISCUSSION

A major advantage of the peptide library approach is that it permits an essentially unbiased query of an enormous library of individual peptides (in the case of the SQ library, approximately $1.8 \times 10^{12}$ peptides) for derivation of an optimal substrate motif for ATM kinase activity. Within the sequence of this optimal substrate, we have identified a central core motif comprised of LSQE shared by both the derived ATM peptide substrate and amino acids 14–17 in p53, an endogenous downstream substrate. Our results indicate that the -1-, +1-, and +2-positions relative to Ser are all important for ATM and DNA-PK catalytic activities. However, these positions are weighted in a comparative sense, with the C-terminal +1- and +2-positions playing significantly more critical roles in defining substrate preference. Additionally, several positions N-terminal to Ser (e.g., Gln/Pro at -2 and Ala at -5) and the C-terminal +3-position (Gln/Val) also demonstrated relatively high preference values in the library analyses. Comparative analyses of ATM and DNA-PK peptide preferences permitted clear distinctions to be drawn between the substrate requirements of the two PIK family kinases.

Differences in ATM and DNA-PK catalytic activity were readily apparent by several criteria. The less degenerate SQII library was much more highly selected by ATM than the SQ library in contrast to the reverse for DNA-PK, in which the SQII library is only moderately selected over the second group of libraries (i.e. the 4Y4+, SF, and SI libraries) phosphorylated by the latter. Significantly, these data further support the critical role of amino acids surrounding the SQ motif in dictating or defining substrate preferences for ATM and DNA-PK. Of interest, we were unable to saturate DNA-PK catalytic activity with the peptides used in this study. By contrast, $V_{max}$ values were readily obtained for ATM. ATM exhibited little or no selectivity for the SP degenerate library in contrast to the modest preference demonstrated by DNA-PK. Conversely, the RXXS library was mildly selected by ATM but not DNA-PK.

A significant point of departure of the two PIK kinases is reflected in the strong preference by DNA-PK for the ATide sequence versus the p53-WT peptide. This preference is in contrast to the equivalent phosphorylation of the two peptides by ATM. Distinct from ATM, the Leu$^1$ → Ala interchange is preferred by DNA-PK. Indeed, we have assayed DNA-PK substrate preferences using the SQ library and find notable differences in residue preferences of DNA-PK at the -1-position and elsewhere. Similar to the situation with ATM, the Glu$^2$ → Ala substitution severely compromised both the AT and p53 peptides as substrates for DNA-PK. Thus, in the context of the peptide libraries and individual peptides assayed in this study, DNA-PK demonstrated a unique substrate preference pattern when compared with ATM. By inference, ATM and DNA-PK should preferentially phosphorylate fundamentally different downstream substrates. The null phenotypes of ATM and DNA-PK exhibit profound differences (14–16, 44–46) also arguing for a significant divergence of ATM and DNA-PK downstream signaling pathways.

We find that similar to DNA-PK (which is assayed in the context of highly purified DNA-PKcs, Ku80/Ku70, and DNA), immunoprecipitated ATM is also a robust Ser/Thr protein kinase, although less active than DNA-PK. This variance may reflect inherent differences in the two kinases or indicate that ATM is being assayed under less ideal conditions. Nevertheless, the substantial level of protein kinase activity manifested by ATM in vitro has potentially important implications regarding its in vivo functions, both as a component in cell cycle checkpoint control as well as its role in development. The activities of both ATM and DNA-PK likely require stringent control processes or physically associated factors that modulate accessibility of the kinase activity of the two PIK family members if, in fact, the observed high level kinase activities extend to in vivo function. For DNA-PK, there is an essential requirement for the presence of DNA for catalytic activity. Additionally, autophosphorylation of the DNA-PKcs subunit of the complex has been correlated with down-regulation of protein kinase activity (63). Regulation of ATM kinase activity remains unknown.

DNA-PK has been previously shown to prefer SQ sequences (5, 43). We have also found an extremely high selectivity by DNA-PK protein kinase activity for the SQ degenerate library over several other libraries with different fixed +1 amino acids. Overall, the library and individual peptide selectivity pattern of DNA-PK indicates a strong possibility for substrate target overlap between ATM and DNA-PK. In this regard, ATR is a third mammalian PIK kinase with apparent specificity for Ser$^2$ in p53 and may also participate through shared substrate preferences of ATM and DNA-PK (3, 4). PIK kinase localization and developmental expression patterns, together with accessibility of a given substrate within the microenvironment or complex containing active PIK kinases, would likely play key roles in dictating phosphorylation of a substrate and subsequent physiological consequences.

Our results show that substrate requirements for DNA-PK and ATM are modestly flexible with respect to the +1-position in that there are two tiers of substrate preference. In addition to the more highly selected SQ library, we observed a preference for libraries with hydrophobic residues at +1. These lesser preferences may well reflect a relevant class of substrate targets for both ATM and DNA-PK. Examples of alternative substrate motifs other than SQ have been described for DNA-PK (see Ref. 5). Additionally, DNA-PK has been shown to preferentially phosphorylate a Ser$^5$-Tyr site in Ku70 (47). This site selection, however, is supported by results obtained in the peptide library analyses. We have likewise observed selectivity for an aromatic hydrophobic residue at the +1-position as shown by the modest SP peptide library preference by DNA-PK and inferred by 4Y4+ library selectivity, i.e., indicating a possible selection for Tyr nearby the targeted Ser/Thr. Phosphoamino acid analysis of the DNA-PK-phosphorylated 4Y4+ library showed mainly phosphoserine and little or no evidence for phosphotyrosine. Additionally, the 4Y4- library, which con-

---

$^a$ A. Dwyer, D. Chan, T. O’Neill, S. Lees-Miller, and G. Rathbun, unpublished data.

$^b$ C. Crovello, G. Rathbun, and L. C. Cantley, unpublished data.
tains only the fixed Tyr and no Ser, Thr, or Tyr residues in the degenerate positions, was not phosphorylated by DNA-PK. The 4Y4+ and SF preferences by DNA-PK are also observed for ATM.

A recent study reported derivation of a consensus substrate sequence of (P,L,I,M)X(L,I,D,E)SQ for ATM using a glutathione S-transferase (GST) peptide fusion approach that adopted the wild-type p53 sequence as a template (48). In this analysis, a short sequence consisting of the N-terminal amino acids surrounding Ser\(^{15}\) in p53 was fused to GST, and various amino acid substitutions within the GST-fused peptides were tested in ATM kinase gel-based assays for overall substrate suitability. Similarities in results obtained in this approach with the present study include a preference for Ser versus Thr, selection for Gln at the 1\(^{1}\) position, and a preference for hydrophobic residues at the 2\(^{1}\) and 3\(^{1}\) positions. The R\(^{XX}\)S library, containing a basic residue at the 2\(^{3}\) position, was modestly selected among the degenerate libraries and may be important as a relevant “second tier” target motif. However, in more extensive analyses (see Table I), we found little or no preference by ATM for basic residues. Positively charged amino acids appeared to adversely affect phosphorylation of the GST peptides (48). The results of the peptide library analysis significantly extend the consensus substrate motif of ATM both N-ter-

---

**Table V**

**ATM phosphorylation sites**

Candidate substrate targets for ATM listed are within \(\leq 0.01\) percentile match-selected using a data base search engine developed in the Cantley lab.\(^7\) H and M indicate human and mouse homologues, respectively. The position of the serine or threonine within the protein selected as a presumed substrate target is superscript. Data base searches utilized a preference matrix produced and weighted relative to the peptide library preference values shown in Table I for the SQ and SQII libraries. The resulting percentile matches range from the highest (0.001% for p53) to the lowest (1.7% for Ser at position 357 in Cds1) percentiles, for the best suited and less preferred motifs, respectively, within putative targets selected in data base searches. H indicates human, M indicates mouse. See text for discussion.

| Known Sites                        | Sequence                                      |
|-----------------------------------|-----------------------------------------------|
| p53                               | EFQSDFPVEPTLS\(^{14}\)QETFSD                  |
| Brca-1                            | LQNRNYP\(^{15}\)TQEELIKV                      |
|                                   | AVLEQGH\(^{14}\)QPSNSYP?                      |
|                                   | KSSSYP\(^{16}\)QPNPSGLS                       |
| Possible sites on known substrates|                                               |
| ABL                               |                                               |
| PHAS-1                            |                                               |
| RAD53 (Saccharomyces cerevisiae)   |                                               |
| CDS1 (Schizosaccharomyces pombe)   |                                               |
| Cds1/CHK2 (H)                     |                                               |
| Other Possible sites              |                                               |
| MDM2                              |                                               |
| MCM3                              |                                               |
| Xenopus DNA replication licensing  |                                               |
| PMS2                              |                                               |
| DNA mismatch repair               |                                               |
| XP-A                              |                                               |
| DNA excision repair               |                                               |
| CYCLIN E2                         |                                               |
| activates Cdk2                    |                                               |
| Dynactin                          |                                               |
| binds to dynein; role in vesicle transport in neurons |                                               |
| TIA1 (T-lymphoma invasion and metastasis, inducing protein) |               |
| GDS protein specific for Rho-like GTPases-implicated in T-lymphoma development; may play a role in neuronal development | |
| PP2C                              |                                               |
| putative protein phosphatase 2C   |                                               |
| Nibrin                            |                                               |
| forms triplex with Mre11 and Rad50-mutated in Nijmegan breakage syndrome |       |
| ATM                               |                                               |
| cell cycle checkpoints, development |                                               |
| DNA-PK                            |                                               |
| repair of DNA double strand breaks|                                               |
| MLK2 (mixed lineage kinase-2)     |                                               |
| Ser/Thr kinase with SH3 and Leu zipper expressed in brain and muscle |               |
| Wee1                              |                                               |
| inhibitory Y15 phosphorylation of Cdc2 |                                               |
| Rad17 (human)                     |                                               |
| Conserved with S. pombe rad17; potential role in cell cycle damage and replication checkpoints |               |

---

\(^7\) H. Leparc and L. C. Cantley, manuscript in preparation.
nal, and in particular, C-terminal to the so-called “SQ motif.” Searches through the genomic data base using consensus sequences derived from the two separate approaches yielded several similar possible substrate targets.

Data base searches using sequences obtained from the peptide library analysis of ATM substrate preference (see Table I) revealed a number of intriguing potential downstream target molecules that might conceivably play a role in ATM-signaling pathways. Table V is a partial listing of a group of candidate molecules that contain sites that most resemble the sequences derived from the SQ and SQII library preference values (Table I). Sites of ATM phosphorylation have been identified in p53 and Brca-1, and some of these were also identified in data base searches using the derived sequence from library analysis thus providing additional validation for the library-derived sequence. Several of the Brca-1 serines apparently phosphorylated by ATM were at SQ sites (39). The tumor suppressor p53 has been previously shown to be a relevant in vitro target for ATM, and our study strongly supports p53 as an important downstream target of ATM. Conversely, the comparatively weaker p53 peptide selectivity by DNA-PK kinase activity indicates a less compelling role for p53 in signaling pathways that utilize DNA-PK (49). Although we are able to observe robust ATM and DNA-PK phosphorylation of several in vitro targets (e.g. Mdm2, Cds1, Chk1, Nbs1, and XRCC4; data not shown), the true value of any predicted substrate is determined through in vivo analyses, and these experiments are clearly the next step in this study.1

We find potential sites on several substrates previously identified as targets for ATM phosphorylation, including PHAS-1 (which contains multiple sites), Abl, and Cds1/Chk2. PHAS-1 has been used as an indicator substrate for ATM phosphorylation activity; no relevance has been established for this substrate in vivo. Cds1/Chk2 has been identified as a downstream component of ATM, and Cds1/Chk2 contains several promising sites, the best of which are shown in Table V. Conservation of certain aspects of mammalian checkpoint signaling pathways involving ATM and Cds1/Chk2 with those of yeast utilizing PIK homologues to ATM is also apparent if similar motifs are targeted as substrates in yeast. For example, Rad3 (an ATM homologue in Schizosaccharomyces pombe) is upstream of yeast Cds1 (see, for example, Ref. 50). Recent evidence suggests that in Saccharomyces cerevisiae, the MECl PIK family member and RAD9 cooperate to activate independent Rad53 and Chk1-signaling pathways that collectively form a DNA damage-activated cell cycle checkpoint (51).

ATM is cytoplasmic in several types of neuronal cells including Purkinje cells (24–27, 52). A yeast two-hybrid screen revealed a β-adaptin association with ATM, and a β-neuronal adaptin-like protein (β-NAP)-ATM complex could be demonstrated in vitro (26). β-NAP and dynactin (see Table V) may play important roles in vesicle transport in neurons. ATM has been localized to peroxisomes in primary fibroblast and established cell lines, and catalase activity has been shown to be decreased in A-T cell lines (53). β-NAP and catalase contain potential substrate motifs (at Ser511 and Ser245, respectively), for ATM kinase activity; however, these sites fall within the lower match percentiles that are 0.233% for catalase and β-NAP at 0.650%. β-Adaptin contains motifs that are predicted as substantially less suited. Nevertheless, substrate accessibility and localization may allow a less ideal sequence (as defined by the present search criteria) to be offered as a relevant ATM substrate. Together with the several studies mentioned above, data base searches also predict substrates that might play tissue-specific roles in terms of ATM-mediated regulation of essential cytoplasmic metabolism as well as vesicle transport and function.

Other possible substrate targets of ATM include proteins with roles in DNA repair and cell cycle checkpoint functions, including Mdm2 and Nbs-1 (Table V). Mdm2 functions as a negative regulator of p53 (54), and Ser175(Gln) in Mdm2 has been identified as a site for in vitro DNA-PK phosphorylation that blocked Mdm2 association with p53 (55). There are several potential sites in Mdm2 that might serve as substrate targets for ATM; the best of these, located in the C-terminal region of the molecule, are presented in Table V. Phosphorylation of Mdm2 by ATM may represent an additional level of cell cycle checkpoint control (56). Nbs-1 (Nibrin) has been identified as the gene mutated in Nijmegen breakage syndrome, a defect with repercussions similar in many ways to ATM functional loss (57–59). Nbs-1 contains several potential substrate motifs for ATM kinase activity. The Leu-Ser397-Glu-Asp sequence (see Table V) is intriguing in that phosphorylation of Ser397 occurs in response to ionizing radiation, and this site is phosphorylated in vitro by ATM (60). Mammalian Wee1 CDK tyrosine kinase has been identified as a positive upstream activator of cyclin dependent kinase Cdc2 activity; hyperphosphorylation may target Wee1 for degradation during certain stages of the cell cycle (60–62). Both ATM and DNA-PK contain potential motifs that might serve as regulatory auto- or trans-phosphorylation sites (63). These and several other molecules identified as containing reasonable substrate motifs for ATM or DNA-PK in Table V may be useful in suggesting additional signaling pathways modulated by the kinase activities of the two PIK family kinases.

Acknowledgments—We are grateful to Drs. Fred Rosen and Fred Alt for support and encouragement. Dr. Martin Lavin kindly provided the C3ABR cell line. We thank Dr. David Weaver for critical comments on the manuscript. We appreciate the secretarial assistance of Talise Dow.

REFERENCES

1. Matsuoka, S., Huang, M., and Elledge, S. J. (1998) Science 282, 1883–1897
2. Hoekstra, M. F. (1997) Curr. Opin. Genet. & Dev. 7, 170–175
3. Lakin, N. D., Henn, B. C., and Jackson, S. P. (1999) Oncogene 18, 2989–2995
4. Trottet, R. S., Brumback, K. M., Williams, J. M., Sarkaria, J. N., Chiby, W. A., Shieh, S. Y., Taya, Y., Prives, C., and Abraham, R. T. (1999) Genes Dev. 13, 152–157
5. Anderson, C. W. (1993) Trends Biochem. Sci. 18, 433–437
6. Gatley, D. P., Hittle, J. C., Chan, G. K., and Yen, T. J. (1998) Mol. Biol. Cell 9, 2361–2374
7. Suzuki, K., Kodama, S., and Watanabe, M. (1999) J. Biol. Chem. 274, 25571–25575
8. Smith, G. C., Cary, R. B., Lakin, N. D., Henn, B. C., Teo, S. H., Chen, D. J., and Jackson, S. P. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 11134–11139
9. Abraham, R. T. (1996) Curr. Opin. Immunol. 8, 412–415
10. Carr, A. M., and Hoekstra, M. F. (1995) Trends Cell Biol. 5, 32–40
11. Savitsky, K., Barshira, A., Gildad, S., Rotman, G., Zivan, Y., Vanagaite, L., Tagle, D. A., Smith, S., Uziel, T., Sfez, S., Ashkenazi, M., Pocker, I., Frydman, M., Hanirik, R., Patanjali, S. R., Simons, A., Chinis, G. A., Sarted, A., Gatti, R. A., Chessa, L., Sanal, O., Lavin, M. F., Jaspers, N. G. J., Malcolm, A., Taylor, R., and Shilooh, Y. (1996) Science 265, 1749–1753
12. Sedgwick, R. P., and Boden, E. (1993) Handsb. Clin. Neural. 18, 347–423
13. Lavin, M. F., and Shilooh, Y. (1997) Annu. Rev. Immunol. 15, 177–202
14. Barlow, C., Hirotsune, S., Paylor, R., Liyanage, M., Eckhaus, M., Collins, F., Shiloh, Y., Crawley, J. N., Ried, T., Tagle, D., and Wynshaw-Boris, A. (1996) Cell 86, 159–171
15. Xu, Y., Ashley, T., Brainard, E. E., Bronson, R. T., Meyn, M. S., and Baltimore, D. (1996) Genes Dev. 10, 2411–2422
16. Eileen, A., Wang, Y., Daugherty, C. J., Morton, C. C., Zhou, F., Campos-Torres, J., and Leder, P. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 13084–13089
17. Blum, T., Finnie, N. J., Taccioli, G. E., Smith, G. C. M., Demengeot, J., Gottlieb, T. M., Mirza, R., Varghese, A. J., Alt, F. W., Jeggo, P. A., and Jackson, S. P. (1995) Cell 80, 813–823
18. Hartley, K. O., Gell, D., Smith, G. C. M., Zhang, H., Divicha, N., Connelly, M. A., Adron, A., Leesimer, S. P., Anderson, C. W., and Jackson, S. P. (1995) Cell 82, 849–856
19. Wiler, R., Leber, R., Moore, B. B., VanDyk, L. F., Perryman, L. E., and Meek, K. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 11485–11489
20. Schuler, W., Weiler, I. J., Schuler, A., Phillips, R. A., Ruebgan, N., Mak, T. W., Kearney, J. F., Perry, R. P., and Bosma, M. J. (1986) Cell 46, 983–972
21. Maley, B. A., Blackwell, T. K., Fulp, G. M., Rathbun, A. G., Furlay, A. J., Ferrier, P., Heinke, L. B., Phillips, R. A., Yancopoulos, G. D., and Alt, F. W. (1988) Cell 54, 453–460

Substrate Motifs Selected by ATM
