c-Abl is a non-receptor tyrosine kinase that is involved in a variety of signaling pathways. Activated forms of c-Abl are associated with some forms of human leukemia. Presently, no high resolution structure of the tyrosine kinase domain of Abl is available. We have developed a structural homology model of the catalytic domain of Abl based on the crystal structure of the insulin receptor tyrosine kinase. Using this model as a guide, we selected residues near the active site predicted to play a role in peptide/protein substrate recognition. We expressed and purified 15 mutant forms of Abl with single amino acid substitutions at these positions and tested their peptide substrate specificity. We report here the identification of seven residues involved in recognition of the P-1, P+1, and P+3 positions of bound peptide substrate. Mutations in these residues cause distinct changes in substrate specificity. The results suggest features of Abl substrate recognition that may be relevant to related tyrosine kinases.

The c-abl proto-oncogene encodes a multidomain non-receptor tyrosine kinase that is expressed ubiquitously in human tissues (reviewed in Refs. 1–4). Mutant forms of c-abl are found in patients with Philadelphia chromosome-positive chronic myelogenous leukemia and acute lymphocytic leukemia (1–4). In these diseases, a chromosomal translocation event produces a chimeric oncogene consisting of 5′-sequences of bcr fused to abl. The BCR-Abl fusion protein has elevated tyrosine kinase activity relative to c-Abl, and the tyrosine kinase activity of the BCR-Abl fusion protein is necessary for disease progression. Similarly, tyrosine kinase activity is necessary for transformation of fibroblasts or hematopoietic cells by BCR-Abl (5, 6).

In addition to its tyrosine kinase catalytic domain, c-Abl has a short amino-terminal unique domain followed by SH3 and SH2 domains (1–4). This domain organization is found in many non-receptor tyrosine kinases. Abl also possesses a large carboxy-terminal region that includes a DNA-binding domain, an F-actin-binding domain, a nuclear localization signal, and a proline-rich region implicated in mediating protein-protein interactions. Studies aimed at understanding the normal physiological role of c-Abl have shown the enzyme to be involved in signal transduction, cytoskeletal rearrangement, RNA polymerase II activation, DNA repair, and cell cycle control (1–4). c-Abl has been shown to physically associate with at least seven unique proteins, including p53 and the nuclear Rb protein (4). Mice with targeted disruptions in the c-abl gene have high neonatal mortality rates and are more susceptible to infection, suggesting a role for c-abl in B-lymphocyte development (7).

At least eight in vivo substrates for Abl have been identified (4). The amino acid sequences surrounding the phosphorylation sites for two of these proteins, RNA polymerase II (8) and c-Crk (9), have been described. These sequences do not share a common primary sequence motif, suggesting that Abl may have a broad range of substrate specificity. Studies using synthetic peptides have been used to examine the substrate specificity of Abl and to define any primary sequence determinants for substrate recognition (10, 11). These studies suggest that, although Abl does not have an absolute consensus sequence for phosphorylation, the best in vitro peptide substrates for Abl contain the sequence Ile-Tyr-Ala-Xaa-Pro, where Xaa is any amino acid. These studies indicate that the P-1 (Ile) and the P+3 (Pro) positions are most important for substrate recognition.

The molecular basis of peptide/protein substrate recognition for tyrosine kinases is not well understood. Presently, a single high resolution crystal structure of a tyrosine kinase in complex with peptide substrate is available: the tyrosine kinase domain of the insulin receptor (IRK)1 complexed with a peptide substrate (12). This structure reveals interactions between enzyme and substrate that govern substrate specificity. Two adjacent hydrophobic pockets on the surface of the C-terminal lobe of IRK accommodate Met side chains C-terminal to the phosphorylated tyrosine on the peptide substrate. The crystal structure of the activated IRK-peptide complex provides a structural basis for understanding the primary signaling specificity of IRK and serves as a general model for tyrosine kinase substrate recognition.

In this paper, we have developed a molecular homology model of the kinase catalytic domain of Abl (Abl-CAT) to help identify amino acids that may be important in substrate recognition. A similar approach was used to propose a molecular model of the Bruton tyrosine kinase and to provide a structural basis for understanding mutations in this enzyme associated with the disease X-linked agammaglobulinemia (13). Our molecular homology model is based on the crystal structure of the ternary complex of IRK with peptide substrate and AMP-PNP bound (12). Using the model as a guide, we have targeted seven residues in Abl-CAT for amino acid substitutions to examine effects on substrate specificity. Site-directed mutants of Abl-CAT were engineered and tested with a series of peptide substrates to monitor changes in specificity. Kinetic analyses of these mutants with the peptide substrates show distinct changes in substrate preferences.

**Experimental Procedures**

Homology Model—Amino acids 362–625 of v-Abl were used to generate primary sequence alignments using the CLUSTALW alignment algorithm (14) and were imported into Swiss-PDB Viewer (15). The

---

* This work was supported by National Institutes of Health Grant CA58530 (to W. T. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel.: 516-444-3533; Fax: 516-444-3432; E-mail: miller@physiology.pnb.sunysb.edu.

1 The abbreviations used are: IRK, insulin receptor tyrosine kinase domain; Abl-CAT, Abl catalytic domain; AMP-PNP, 5′-adenylylimidodiphosphate; Nle, norleucine.
primary model of the Abl catalytic domain was prepared using the spatial coordinates of α-carbons from the crystal structure of the ternary form of IRK with peptide substrate and AMP-PNP bound (12). Gaps in the sequence alignment were ligated manually, minimizing large steric clashes. Energy minimization and loop insertions were carried out using SwissModel automated modeling system (16). Additional rounds of energy minimization were carried out using Sculpt for Power Macintosh (17). The stereochemical quality of the model was checked using PROCHECK Version 3.3 (18), which reports no distorted main-chain bonds, five distorted main-chain angles, and no distorted planar groups. The distorted main-chain angles were outside of the predicted range for aromatic residues and do not interfere with our interpretation of the model. Solvent-accessible surface area was calculated and visualized using WebLab Viewer (Molecular Simulations Inc.) with a 1.4 Å probe. Figs. 1A and 2 were prepared using Strata Studio Pro (Strata Inc., St. George, UT).

Mutagenesis, Expression, and Purification—Our experiments were carried out on the isolated catalytic domain of v-Abl, expressed in Escherichia coli as described previously (19). The sequence numbering used is from the gag-Abl fusion protein of the Abelson murine leukemia virus (20). Mutagenesis of the Abl catalytic domain was carried out using a QuickChange mutagenesis kit (Stratagene). Mutagenesis primers complementary to wild-type template were designed with single, double, or triple nucleotide substitutions. The DNA sequences encoding the entire catalytic domains of the mutants were confirmed by DNA sequencing on an ABI373 automated DNA sequencer. Wild-type and mutant proteins were expressed as glutathione S-transferase fusion proteins in E. coli strain NB42 and purified using glutathione-agarose (19). All proteins expressed to similar levels, were of the expected size, and purified to ≥98% homogeneity.

Peptides—Synthetic peptides were prepared by solid-phase synthesis using standard Fmoc (N-(9-fluorenlymethylcarbonyl) chemistry on an Applied Biosystems automated 431A peptide synthesizer. Peptides were purified using semi-preparative reversed-phase high performance liquid chromatography. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry was used to confirm the identity of the final products.

 Kinase Assays—Tyrosine kinase activity assays were carried out using two methods. In both cases, the intact glutathione S-transferase-Abl fusion proteins were used; we showed previously that the specific activity and substrate specificity of glutathione S-transferase-Abl are similar to those of full-length Abl and that velocity versus [enzyme] plots are linear over the concentrations used here (19). For initial comparisons of several substrates, phosphocellulose binding assays were used to measure incorporation of 32P-ATP into peptides (19, 21). These reactions were carried out in triplicate at saturating concentrations of ATP (250 μM) and Mg2+ (10 mM). The reactions were carried out in volumes of 25 μL using peptide concentrations of 50 μM or 1.5 mM and 1 μg of enzyme. Picomoles of phosphate incorporated into peptides were measured after 20 min as described (19).

A continuous spectrophotometric assay (22) was used to measure initial rates of phosphorylation and to determine kinetic constants for some peptides. Reactions were carried out in volumes of 100 μL using 5 μg of purified enzyme. Saturating concentrations of ATP (500 μM) and Mg2+ (10 mM) were used, and peptide concentrations ranged from 50 μM to 2 mM. Data were sampled every 30 s to determine rates of phosphorylation. Initial rates (<10% of the reaction) were measured in triplicate. The initial rate values were averaged, and Vmax and Km were determined by fitting to the hyperbolic velocity versus [substrate] curves using the program MacCurve Fit. For some combinations of peptides and mutants, initial experiments established that their Km values were in the millimolar range (>2 mM). In addition, we observed that high concentrations of these peptides were inhibitory. Thus, concentrations over 2 mM were not employed, and we were unable to determine Km values accurately using initial rate kinetics. For these peptides, the complete time courses for phosphorylation were measured using peptide concentrations less than Km. In these cases, we analyzed the data graphically as described (23) to determine Vmax/Km.

RESULTS

Homology Model—The overall topology of the Abl structural homology model reflects the typical bilobal structure shared by all eukaryotic protein kinases (24) with a five-stranded β-sheet and a single α-helix in the amino-terminal lobe, responsible for MgATP binding, and a highly helical carboxyl-terminal lobe (Fig. 1). Based on results for other protein kinases, the C-terminal lobe is predicted to make most of the contacts with peptide/protein substrates. The total root mean square deviation of the polypeptide backbone between the Abl model and the IRK structure is 1.8 Å. The crystal structure of the activated insulin receptor catalytic domain with peptide bound (12) served as a model for the orientation and structure a peptide may adopt in the Abl active site. Comparison of the C-terminal peptide-binding domains of IRK and Abl indicated that the greatest differences are in the regions responsible for binding the P+3 amino acid side chain. The P–1 region differs only slightly from that of IRK. The activation loop (amino acids 500–521; see Fig. 1) is extremely flexible when examined using molecular mechanics (Sculpt, Interactive Simulations, Inc.), and the structure of the loop in our model is one of many possible conformations. Additionally, the structure of IRK used for our model is multiply phosphorylated on the activation loop (12). Abl contains a single tyrosine within the activation loop, Tyr-513, which is phosphorylated in vivo and in vitro and is believed to be involved in enzyme activation (25, 26). Although we believe that our model represents the activated form of Abl, the activation loop tyrosine is modeled in its unphosphorylated state.

Design of Peptide Substrates—Previous studies of the in vitro substrate specificity of Abl have been carried out in this laboratory and by others (for review, see Ref. 27). We designed two groups of peptide substrates to examine any changes in specificity at the P–1 and P+1/P+3 positions for engineered mutant forms of Abl (Table I). (Amino acids in peptide substrates are designated by their position relative to the phosphorylated tyrosine. For example, in the sequence Ile-Tyr- Ala-Ser-Pro, Ile is at the P–1 position, Ala is at the P+1 position, and Pro is at the P+3 position.) Peptides used to examine P+1 and P+3 specificity share the sequence Ser-Arg-Gly-Asp-Tyr-Xaa3-Thr- Xaa4-Gln-Ile-Gly, where either Xaa3 or Xaa4 is varied. These peptides are based on a peptide sequence derived from a phosphorylation site of insulin receptor substrate-1 used previously in our laboratory to examine the substrate specificity of wild-type Abl at the P+1, P+2, and P+3 positions (19). In these earlier studies, we found that amino acids at the P+2 position do not strongly influence substrate recognition. Moreover, in the ternary structure of IRK, specificity in peptide binding is achieved through interactions with the P+1 and P+3 residues (12). For these reasons, we did not examine the effects of residue changes at the P+2 position in this study.

We chose to use a different series of peptides to examine P–1 specificity (Table I). This is because, in the context of the insulin receptor substrate-1 peptide, we did not observe a strong dependence on the amino acid at the P–1 position for Abl phosphorylation.2 Peptides designed to examine specificity at the P–1 position share the sequence Leu-Ile-Glu-Ala-Xaa3-Tyr-Xaa4-Ala-Arg-Gly, where Xaa3 is varied. This sequence is based on the autophosphorylation site of Src and has been previously used in our laboratory to examine P–1 specificity in wild-type Abl (11). Amino acids for the substituted position were chosen to explore the effect of size, charge, and hydrophobic character. Because the two groups of peptides are dissimilar in sequence, we did not attempt to draw conclusions about the relative importance of P–1 versus P+1/P+3 recognition for each mutant.

Mutations That Affect P+3 Peptide Recognition—Experiments with synthetic peptides and peptide libraries have demonstrated that Abl prefers proline at the P+3 residue of a substrate. In the crystal structure of IRK, Leu-1219 is part of a binding pocket that surrounds the P+3 methionine side chain

---

2 D. A. Hinds, W. T. Miller, and S. E. Shoelson, unpublished observations.
of the peptide substrate (12). In the Abl homology model, the residue homologous to Leu-1219 is Tyr-569. Tyr-569 is partially solvent-exposed in our model and could adopt the same role as Leu-1219 in IRK (Fig. 2). Three separate mutant forms of Abl-CAT were engineered with amino acid substitutions at Tyr-569: Y569L, Y569A, and Y569W. Alanine was chosen as a substitute for Tyr-569 to examine the effects of minimizing the amino acid side chain length. In c-Src, the residue corresponding to Tyr-569 is a leucine (28), and the specificity of c-Src at the P<sub>1</sub> position differs from that of Abl (10). For this reason, we chose to introduce leucine as a substitute for Tyr-569. We also mutated Tyr-569 to Trp because the Abl homology model suggests that a large side chain at this position might cause the putative P<sub>1</sub> substrate-binding pocket to be too narrow to accommodate amino acids with large side chains such as proline.

Initial screens of Abl mutants were carried out using peptide concentrations of 1.5 mM and measuring [32P]ATP incorporation after 20 min to assess any changes in substrate phosphorylation (Fig. 3). Experiments using low (50 µM) peptide concentrations showed no changes in the rank order of specificity compared with those using higher peptide concentrations (data not shown). These initial activity measurements indicated that the Y569W and Y569L mutants had an altered specificity (Fig. 3); in particular, recognition of Pro at the P+3 position was greatly reduced. These changes in specificity were characterized further by more detailed kinetic analyses (Table II). The kinetic measurements show that the Y569W mutant phosphorylated the P<sub>1</sub>Ala/P<sub>3</sub>Met peptide best, with a V<sub>max</sub>/K<sub>m</sub> value of 5.6. The P<sub>1</sub>Met/P<sub>3</sub>Pro peptide, the best for the wild type with a V<sub>max</sub>/K<sub>m</sub> value of 10.7, was phosphorylated less efficiently by this mutant, with a V<sub>max</sub>/K<sub>m</sub> value of 1.3 (Table II). The phosphorylation of other peptides by the Y569W mutant was similar to the wild type (Fig. 3 and Table II). The difference in specificity observed in the Y569W mutant therefore arises from a decrease in V<sub>max</sub>/K<sub>m</sub> for the P<sub>1</sub>Pro peptide specifically rather than an overall decrease in the phosphorylation of all peptides. Kinetic measurements also showed a decrease in the phosphorylation of the P<sub>1</sub>Met/P<sub>3</sub>Pro peptide by Y569L, with a V<sub>max</sub>/K<sub>m</sub> value of 4.4 (wild-type V<sub>max</sub>/K<sub>m</sub> 10.7) (Table II). The Y569A mutant did not display any changes in phosphorylation of the peptides tested compared with wild-type Abl (Fig. 3), and we did not carry out kinetic analysis of this mutant.

**Mutations That Affect P+1 Peptide Recognition**—Two residues in Abl-CAT, Phe-521 and Ile-523, were identified as residues that might contribute to P+1 substrate specificity (Fig. 2). Ile-523 was chosen based on the homology model. A solvent-accessible surface representation suggests that Ile-523 is partially solvent-exposed and forms the edge of a groove into which a substrate amino acid side chain may fit. Two substitutions...
were made for Ile-523: I523V and I523A. These amino acid substitutions were chosen to examine the effect of shortening the side chain on accommodating P-1 residues. Phe-521 was selected to test the possibility that residues near the activation loop contribute to specificity; based on homology to other protein kinases, the activation loop in Abl is predicted to span residues 500–521. We examined the effects of smaller (F521A) and larger (F521W) side chains at this position. Initial measurements of activity were used to assess changes in specificity and were indistinguishable from wild-type Abl. Wild-type enzyme has a preference for Leu at the P-1 position and an increased Km value for P-1Ile relative to wild-type Abl. The L444E mutant demonstrated a specificity different from that of wild-type Abl. The L444K mutant has a Km value for P-1Ile (12.0) is similar to the value of 12.3 for wild-type Abl and the Vmax/Km value of this mutant for P-1Ile, 5.4, is close to the value of 6.8 for wild-type Abl (Table IV). The Vmax/Km value of this mutant for P-1Glu, 0.08, is substantially lower than the corresponding value of 1.3 for wild-type Abl phosphorylating P-1Glu (Table IV).

Mutations That Affect P-1 Peptide Recognition—Isoleucine at the P-1 position is a strong determinant for substrate recognition in wild-type Abl (10, 11). We produced Abl mutants with amino acid substitutions at residues predicted to be near the P-1-binding region of the enzyme. Abl-CAT residue Leu-444 was chosen as a target for mutagenesis based on the role of the corresponding residue in IRK, Lys-1085. In IRK, Lys-1085 extends from the α-helix and makes a water-mediating hydrogen bond with the P-1 residue (Asp) of the substrate peptide (12). In our model, Leu-444 extends from the structurally homologous helix and is partially solvent-exposed. We chose to change Leu-444 to lysine to examine the possibility of a change in substrate specificity to that of IRK (which prefers Glu at P-1) (10). An L444E mutation was made to examine the possibility that introduction of an acidic residue at this position may favor the binding of a basic amino acid at the P-1 position of peptide substrate. The L444K and L444E mutants retained the overall preference for Ile at the P-1 position, although they differed from the wild type in phosphorylation of other substrates (Fig. 5 and Table IV). The L444K mutant has Vmax/Km values of 8.0 for the P-1His peptide and 21.0 for the P-1Phe peptide, whereas wild-type Abl has Vmax/Km values of 6.8 for P-1His and 12.3 for P-1Ile (Table IV). Thus, L444K shows enhanced recognition of P-1His relative to wild-type Abl. The L444E mutant demonstrated a specificity different from that of wild-type Abl. The L444K mutant has a Km value for P-1Ile, 12.0, is similar to the value of 12.3 for wild-type Abl and the Vmax/Km value of this mutant for P-1Ile, 5.4, is close to the value of 6.8 for wild-type Abl (Table IV). The Vmax/Km value of this mutant for P-1Glu, 0.08, is substantially lower than the corresponding value of 1.3 for wild-type Abl phosphorylating P-1Glu (Table IV).

Abl-CAT residues Gly-556 and Ser-558 were chosen because these residues are solvent-accessible in the homology model, in close proximity to Leu-444 (Fig. 2). We produced G556A, G556V, S558V, and S558N mutants. The G556A and S558A mutants were indistinguishable from the wild type in our initial screens (Fig. 5). The G556V mutant, however, has a reduced preference for Ile at the P-1 position and an increased preference for Leu at the P-1 position. For wild-type enzyme, the Vmax/Km value for P-1Ile, 6.8, is substantially lower than the corresponding value of 1.3 for wild-type Abl phosphorylating P-1Ile (Table IV).

We also chose to mutate Trp-525 of Abl based on our modeling studies. To identify residues in Abl that may make contacts with the P-1 side chain, we used the activated IRK structure to model Ile (in place of Asp) at the P-1 position of the peptide substrate. The indole of Trp-1175 of IRK packs against the side chain of one energetically favorable rotamer of the modeled Ile. In our Abl model, the residue homologous to Trp-1175, Trp-525, lies at the bottom of the putative P-1-binding pocket (Fig. 2). Trp at this position is conserved in protein tyrosine kinases (28). We substituted the bulky residues Phe and His for Trp-525 to minimize perturbations in the structure. The W525F and W525H mutants preferred the P-1His peptide overall and showed modest changes toward other substrates. For example, the enzymes showed different abilities to phosphorylate the P-1His and P-1Glu peptides (Table IV). W525H had a de-
Increased selectivity for P\textsubscript{1}\textsubscript{His} and an increased selectivity for P\textsubscript{2}\textsubscript{Glu} relative to wild-type Abl (Table IV). These results suggest that Trp-525 of Abl is in the vicinity of the P\textsubscript{2} position of bound substrate, although it appears not to play a dominant role in substrate selection.

**Fig. 2.** Three pairs of models showing amino acids targeted for mutagenesis. A and B show residues involved in P\textsubscript{−1} substrate specificity. C and D show residues involved in P\textsubscript{+1} specificity. E and F show the residue involved in P\textsubscript{+3} specificity. A, C, and E, molecular surface representations of the Abl homology model showing residues targeted for mutagenesis in red-orange. B, D, and F, tube schematic of the Abl homology model showing residues targeted for mutagenesis in ball and stick representation.

**Indirect Effects**—We examined the possibility that mutations designed to affect recognition of one position in the substrate might have effects at other positions as well. Mutants
FIG. 3. Initial comparisons of wild-type and mutant forms of Abl (Y569W, Y569L, and Y569A) with P+1/P+3 peptide variants. Wild-type Abl and the three mutant forms of Abl were tested with a panel of eight peptides. The incorporation of $[^32P]$phosphate into peptides was determined after a 20-min reaction using the phosphocellulose paper assay.

| Enzyme                  | $V_{max}$ (nmol/min/mg) | $K_m$ (mM) | $V_{max}/K_m$ | $V_{max}/K_m$ ratio |
|-------------------------|-------------------------|------------|---------------|---------------------|
| Wild type               |                         |            |               |                     |
| P+1Ala/P+3Pro          | 7.5 ± 0.9               | 0.7 ± 0.2  | 10.7          | 1.0                 |
| P+1Ala/P+3Thr          | 2.3 ± 0.5               | 1.8 ± 0.6  | 1.3           | 1.0                 |
| P+1Ala/P+3Met          | 2.5 ± 0.3               | 0.4 ± 0.1  | 6.3           | 1.0                 |
| P+1Ala/P+3Pro, Y569L   | 3.4 ± 0.4               | 0.5 ± 0.2  | 6.8           | 1.0                 |
| P+1Ala/P+3Thr, Y569L   | 4.4 ± 0.4               | 1.0 ± 0.4  | 4.4           | 0.4                 |
| P+1Ala/P+3Met, Y569L   | 2.0 ± 0.2               | 0.9 ± 0.1  | 2.2           | 1.7                 |
| P+1Ala/P+3Ala, Y569L   | 1.9 ± 0.2               | 0.8 ± 0.2  | 2.4           | 0.4                 |
| P+1Ala/P+3Thr, Y569L   | 3.6 ± 0.4               | 0.5 ± 0.1  | 7.2           | 1.1                 |
| P+1Ala/P+3Met          | 2.5 ± 0.3               | 1.9 ± 0.1  | 1.3           | 0.1                 |
| P+1Ala/P+3Thr, Y569A   | 2.8 ± 0.3               | 0.5 ± 0.1  | 5.6           | 0.9                 |

TABLE II

Kinetic measurements for wild-type Abl and two mutant forms of Abl (Y569W and Y569L)

Enzyme/peptide combinations were chosen based on initial comparisons of enzymatic activity toward P+1/P+3 peptides (Fig. 3). Kinetic constants for these enzyme/peptide combinations with significant changes from the wild type are shown. To compare phosphorylation of a particular peptide by mutant versus wild type, the $V_{max}/K_m$ value for the mutant was divided by the value for the wild type. This ratio is given in the last column.

with predicted changes in P+3 recognition were tested with five peptides containing Met, Ile, Ala, Glu, or Nle at the P+1 position (Fig. 3). Whereas the pattern of peptide phosphorylation by Y569W and Y569A closely resembled that of wild-type Abl, Y569L showed some differences from the wild type (Fig. 3). We examined these differences more closely by kinetic analysis of P+1Ala/P+3Met and P+1Ala/P+3Met (Table II). These measurements showed a decrease in the phosphorylation of P+1Ala/P+3Met, with $V_{max}/K_m = 2.4$ (wild-type $V_{max}/K_m = 6.3$). Thus, in the Y569L mutant, a change in the P+3 region has effects on P+1 recognition as well. We also screened the P+3 mutants with the following peptides varying at P+1: P+1His, P+1Glu, P+1Gln, P+1Lys, and P+1Ala. These comparisons showed no differences between wild-type Abl and any of the mutants (data not shown).

We tested for indirect effects using the P+1 and P−1 mutants as well. We carried out initial comparisons of the P+1 mutants using four peptides that vary at the P+3 position: P+1Met/P+3Met, P+1Met/P+3Thr, P+1Met/P+3Ala, and P+1Met/P+3Pro (Fig. 4). In these experiments, there was no observable difference between the substrate specificities of wild-type Abl and the P+1 mutants toward the peptides varying at P+3 (Fig. 4). Similarly, we observed no differences between the P+1 mutants and the wild type in recognition at the P−1 position (data not shown).

We screened the following four P−1 mutants of Abl for indirect effects: L444K, L444E, S558N, and S558A. We tested the following P+1/P+3 peptides: P+1Met/P+3Met, P+1Ala/P+3Met, P+1Glu/P+3Met, P+1Thr/P+3Met, P+1Ala/P+3Ala, and P+1Ala/P+3Thr. All four of the mutants displayed the same rank order of substrate preference as the wild type in this experiment (P+1Ala/P+3Met > P+1Met/P+3Met > P+1Glu/P+3Met > P+1Ala/P+3Ala > P+1Met/P+3Thr) (data not shown). We conclude from these studies on indirect effects that the sites on Abl for recognition of the P−1 and P+1/P+3 positions are distinct. On the other hand, the P+1 and P+3 sites may have some overlap, as at least one mutation (Y569L) had an effect on recognition of both positions.

DISCUSSION

Although Abl is capable of phosphorylating a wide range of peptide and protein substrates, the best peptide substrates for Abl contain the sequence Ile-Tyr-Ala-Xaa-Pro, as shown in peptide library studies (10, 11). Ile at the P−1 position and Pro at the P+3 position are the most important determinants of substrate specificity for Abl. Here, we have identified residues in the catalytic domain of Abl involved in peptide substrate binding and specificity. These residues are located primarily in the C-terminal lobe of the catalytic domain, which has been implicated previously in substrate binding for other protein kinases (12, 24, 27).

The mutant forms of Abl described here fall into three classes with respect to substrate specificity. 1) Two mutants (Y569W and Y569L) have altered substrate specificity. These mutants no longer prefer proline at the P+3 position in peptide substrates. 2) Many of the mutants (e.g., I523V, W525H, and L444E) showed no change in the major determinants for substrate recognition, but differed from the wild type in their phosphorylation of other peptide substrates. For example, mutations aimed at altering recognition of the P−1 position resulted in enzymes that still preferred Ile at P−1, but that
diverged from the wild type when screened against peptides containing other amino acids at P–1. In these cases, these residues may not be involved in direct interactions with the P–1 residue of substrate. Instead, because of their vicinity to the P–1 position, they may act indirectly, stabilizing the local structure to interact favorably with Ile at the P–1 position. 3) Some mutants (e.g. Y569A, I523A, and G556A) showed no changes in specificity when assayed against a variety of peptide substrates. These mutants were not characterized by kinetic analysis.

The Y569W mutation in Abl has the most dramatic effect on substrate specificity of the mutants we report here. Wild-type Abl phosphorylates a peptide substrate with Pro at the P–3 position best. The Y569W mutant phosphorylates P+1_Met/P+3_Met, P+3_Phe ~10 times less efficiently than does wild-type Abl (Table II). All other amino acid side chains tested at the P–3 position of the substrate were phosphorylated at the same level as in the wild type (Fig. 3). Our structural model suggests that this change in specificity arises from a steric clash between the side chain of Trp-569 in the mutant and the side chain of proline in the substrate. This is not the case with the other peptide substrates tested that have smaller amino acid side chains at the P–3 position. Proline at the P–3 position plays a role in substrate recognition in vivo in at least one case: Abl phosphorylates c-Crk at Tyr-221 within the sequence Tyr-Ala-Gln-Pro (9). Phosphorylation by Abl is believed to modulate the protein binding and transforming activity of Crk (29). Preliminary experiments indicate that, in contrast to wild-type Abl, the Y569W mutant has no activity toward Crk in vitro.3

Mutations predicted to affect substrate recognition at the P+1 position (I523V) or at the P–1 position (L444K, L444E, G556V, S558N, W525F, and W525H) do so in a more subtle manner. These mutations do not change the overall preference for Ile at P–1 or Ala at P+1; however, we observed effects on specificity when we screened these mutants against peptides containing other residues at P–1 or P+1. For example, the I523V mutant still phosphorylated the P+1_Ala/P+3_Met peptide best (of the peptides tested), but the V_{max}/K_{m} value for the P+1/Ile/P+3_Met peptide was 3.5 times higher in this mutant than in the wild type (Table III). The L444E mutant, while still preferring Ile at the P–1 position, was 2.1 times more efficient at phosphorylating the P–1_His peptide than the wild type and 19 times less efficient at phosphorylating the P–1_Glu peptide than the wild type (Table IV). There are at least two explanations for these subtle effects on substrate specificity. (i) The residues may not make direct contact with bound substrate, but might instead be involved indirectly in maintaining the three-dimensional structure of Abl to favor certain amino acids in the substrate. (ii) Additionally or alternatively, substrate specificity at P–1 or P+1 may be achieved by a combination of residues, such that single amino acid substitutions do not cause complete alterations in substrate recognition. Indirect effects could explain why, for example, the L444K mutant phosphorylates the P–1_His peptide better than the wild type and shows a decrease in the phosphorylation of P–1_Glu when compared with the wild type (Table IV).

There are residues in the three-dimensional structures of Src family kinases that appear to correspond to residues identified in our study (30, 31). The substrate specificity of Src differs from that of Abl at the P–3 position (10). Src prefers a phenylalanine at the P+3 position in a peptide substrate, whereas Abl prefers proline. The residue homologous to tyrosine 569, a

---

3 J. H. Till and W. T. Miller, unpublished observations.
FIG. 5. Initial comparisons of wild-type and mutant forms of Abl (G556V, G556A, L444K, L444E, S558N, S558A, W525H, and W525F) with P–1 peptide variants. Each of the eight mutant forms of Abl shown here was tested with a panel of five peptides: P–1Ala, P–1Glu, P–1Ile, P–1Gln, and P–1His. The incorporation of [32P]phosphate into peptides was determined after a 20-min reaction using the phosphocellulose paper assay.

TABLE IV

Kinetic measurements for wild-type Abl and six mutant forms of Abl (L444K, L444E, S558N, G556V, W525H, and W525F)

Enzyme/peptide combinations were chosen based on initial comparisons of enzymatic activity toward all P–1 peptides (Fig. 5). Kinetic constants for those enzyme/peptide combinations with significant changes from the wild type are shown. To compare phosphorylation of a particular peptide by mutant versus wild type, the $V_{\text{max}}/K_{\text{m}}$ value for the mutant was divided by the value for the wild type. This ratio is given in the last column.

| Enzyme | $V_{\text{max}}$ | $K_{\text{m}}$ | $V_{\text{max}}/K_{\text{m}}$ | $V_{\text{max}}/K_{\text{m}}$ ratio |
|--------|-----------------|----------------|-----------------------------|---------------------------------|
| Wild type |                  |                |                            |                                 |
| P–1Ile | 4.9 ± 0.5       | 0.4 ± 0.05     | 12.3                        | 1.0                             |
| P–1Glu | 1.8 ± 0.3       | 1.4 ± 0.1      | 1.3                         | 1.0                             |
| P–1Ile | 0.6 ± 0.1       | 0.8 ± 0.1      | 0.8                         | 1.0                             |
| P–1Glu | 4.1 ± 0.4       | 0.6 ± 0.1      | 6.8                         | 1.0                             |
| L444K  | P–1Ile 18.9 ± 3.3| 0.9 ± 0.3      | 21.0                        | 1.7                             |
|        | P–1Glu ND                             |                | 0.5                         | 0.4                             |
|        | P–1Ile 10.4 ± 1.9| 1.3 ± 0.3      | 8.0                         | 1.2                             |
| L444E  | P–1Ile 12.2 ± 1.4| 0.7 ± 0.1      | 18.9                        | 1.5                             |
|        | P–1Glu ND                             |                | 0.07                        | 0.05                            |
|        | P–1Ile 5.8 ± 0.7| 0.4 ± 0.1      | 14.5                        | 2.1                             |
| S558N  | P–1Ile 3.6 ± 0.4| 0.3 ± 0.04     | 12.0                        | 1.0                             |
|        | P–1Glu ND                             |                | 0.08                        | 0.06                            |
|        | P–1Ile 1.3 ± 0.1| 0.2 ± 0.03     | 5.4                         | 0.8                             |
| G556V  | P–1Ile 2.5 ± 0.3| 0.6 ± 0.1      | 4.2                         | 0.3                             |
|        | P–1Glu 1.3 ± 0.1| 0.7 ± 0.1      | 1.9                         | 2.3                             |
| W525H  | P–1Ile 5.1 ± 0.6| 0.4 ± 0.04     | 12.8                        | 1.0                             |
|        | P–1Glu 2.3 ± 0.3| 0.7 ± 0.1      | 3.3                         | 2.5                             |
|        | P–1Ile 3.0 ± 0.3| 1.5 ± 0.2      | 2.0                         | 0.3                             |
| W525F  | P–1Ile 4.5 ± 0.5| 0.6 ± 0.1      | 7.5                         | 0.6                             |
|        | P–1Glu 2.0 ± 0.4| 1.3 ± 0.2      | 1.5                         | 1.2                             |
|        | P–1Ile 3.4 ± 0.4| 0.5 ± 0.1      | 6.8                         | 1.0                             |

* ND, not determined (see "Experimental Procedures").
residue involved in P+3 specificity in Abl (Fig. 3), is a leucine (Leu-472) in Src. This sequence difference may account for the differences seen in substrate specificity; an L472Y mutant of Src might phosphorylate P+3-containing peptides more efficiently. A tryptophan substitution at this position could prevent large side chain amino acids from binding in this region, as we observed for Abl.

The substrate specificities of tyrosine kinase catalytic domains are important in maintaining the fidelity of cellular signal transduction pathways. This is best illustrated in the case of the RET receptor. A naturally occurring mutation in the signal transduction pathways. This is best illustrated in the case of the RET receptor. A naturally occurring mutation in the catalytic domain of this receptor changes a methionine residue to threonine residue in a region homologous to the region of Abl kinase domain of this receptor changes a methionine residue to a threonine residue in a region homologous to the region of Abl shown here to be involved in substrate recognition of the P+1 residue (32, 33). This change affects the substrate specificity of the enzyme at the P+1 position, changing the preference from methionine at that position in the substrate to alanine (10). This mutant form of the RET receptor is implicated in multiple endocrine neoplasia type 2A (32, 33).

Mutations throughout the Abl protein have been reported previously (4). Many of these mutations affect the regulation of the enzyme in vivo. One such mutation, which is sufficient to activate c-Abl enzymatic activity in vivo, is found in the catalytic domain. This mutation changes a tyrosine to phenylalanine within the ATP-binding fold of the enzyme (34). Mutations that affect substrate recognition by the catalytic domain, however, have not been reported previously. Our studies on Abl have highlighted seven residues as playing important roles in peptide substrate recognition. We also show that a single amino acid change of Tyr-569 to tryptophan can affect the substrate specificity dramatically. The results raise the possibility of altering tyrosine kinase substrate specificity in vivo by protein engineering.

Acknowledgment—We thank Professor Stevan Hubbard (New York University Medical School) for helpful comments on the manuscript.

REFERENCES

1. Wang, J. Y. J. (1993) Curr. Opin. Genet. Dev. 3, 35–43
2. Laneuville, P. (1995) Semin. Immunol. 7, 255–266
3. Pendergast, A. M. (1996) Curr. Opin. Cell Biol. 8, 174–181
4. Raitano, A. B., Whang, Y. E., and Sawyers, C. L. (1997) Biochem. Biophys. Acta 1335, P201–P216
5. Lugo, T. G., Pendergast, A. M., Muller, A. J., and Witte, O. N. (1990) Science 247, 1079–1082
6. McLaughlin, J., Chianese, E., and Witte, O. N. (1989) Mol. Cell. Biol. 9, 1866–1874
7. Tybulewicz, V. L., Crawford, C. E., Jackson, P. K., Bronson, R. T., and Mulligan, R. C. (1991) Cell 65, 1153–1163
8. Boettcher, R., Dahnus, M. E., and Wang, J. Y. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 11167–11171
9. Feller, S. M., Knudsen, B., and Hanafusa, H. (1994) EMBO J. 13, 2341–2351
10. Zhou, S., Carraway, R. L., III, Eck, M. J., Harrison, S. C., Feldman, R. A., Mohammadi, M., Schlessinger, J., Hubbard, S. R., Smith, D. P., Engel, C., Lorenzo, M. J., Ponder, B. A. J., Mayer, B. J., and Cantley, L. C. (1995) Nature 373, 536–539
11. Till, J. H., Annan, R. S., Carr, S. A., and Miller, W. T. (1994) J. Biol. Chem. 269, 7423–7428
12. Hubbard, S. R. (1997) EMBO J. 16, 5572–5581
13. Vihinen, M., Vetrie, D., Maniar, H. S., Ochs, H. D., Zhu, Q., Vorechovsky, I., Webster, A. D., Notarangelo, L. D., Nilsson, L., and Sowadski, J. M. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 12803–12807
14. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Nucleic Acids Res. 22, 4673–4680
15. Guex, N., and Peitsch, M. C. (1997) Electrophoresis 18, 2714–2723
16. Peitsch, M. C. (1996) Biochem. Soc. Trans. 24, 274–279
17. Surles, M. C., Richardson, J. S., Richardson, D. C., and Brooks, F. P., Jr. (1994) Protein Sci. 3, 198–210
18. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallogr. 26, 283–291
19. Garcia, P., Shoelson, S. E., George, S. T., Hinds, D. A., Goldberg, A. R., and Miller, W. T. (1993) J. Biol. Chem. 268, 25146–25151
20. Reddy, E. P., Smith, M. J., and Steivens, A. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 3623–3627
21. Casmelie, J. E. (1991) Methods Enzymol. 200, 115–120
22. Boerner, R. J., Barker, S. C., and Knight, W. B. (1995) Biochemistry 34, 16419–16423
23. Thomas, N. E., Bramson, H. N., Nairn, A. C., Greengard, P., Fry, D. C., Mildvan, A. S., and Kaiser, E. T. (1987) Biochemistry 26, 4471–4474
24. Taylor, S. S., and Radzio-Andzelm, E. (1994) Structure 2, 345–355
25. Golub, T. R., Goga, A., Barker, G. F., Affar, D. R., McLaughlin, J., Bohlander, S. K., Rowley, J. W., Witte, O. N., and Gilliland, D. G. (1996) Mol. Cell. Biol. 16, 4107–4116
26. More, M., Wallenhorst, J., Goga, A., Witte, O. N., and Superti-Furga, G. (1997) Leukemia (Basingstoke) 11, Suppl. 3, 313–315
27. Pinna, L. A., and Ruzzene, M. (1996) Biochim. Biophys. Acta 1314, 191–225
28. Hanks, S. K., Quinn, A. M., and Hunter, T. (1988) Science 241, 42–52
29. Rosen, M. K., Yamazaki, T., Gish, G. D., Kay, C. M., Pawson, T., and Kay, L. E. (1995) Nature 374, 477–479
30. Sicheri, F., Murelli, I., and Kurtyjan, J. (1997) Nature 385, 602–609
31. Xu, W., Harrison, S. C., and Eck, M. J. (1997) Nature 385, 595–602
32. Carlson, K. M., Dou, S., Chi, D., Scavarda, N., Toshima, K., Jackson, P. K., Bronson, R. T., and Mulligan, R. C. (1991) Cell 65, 1153–1163
33. Hofstra, R. M., Landsvater, R. M., Ceccherini, I., Stulp, R. P., Stelwagen, T., Luo, Y., Pasini, B., Hoppener, J. W., van Amstel, H. K., Rome, G., Lips, C. J. M., and Buys, C. H. C. M. (1994) Nature 376, 375–376
34. Allen, P. B., and Wiedemann, L. M. (1996) J. Biol. Chem. 271, 19585–19591