Identification of an inhibitor-binding site to HIV-1 integrase with affinity acetylation and mass spectrometry

Nick Shkriabai*, Sachindra S. Patil[†], Sonja Hess[‡], Scott R. Budihas[§], Robert Craigie[‡], Terrence R. Burke, Jr.[†], Stuart F. J. Le Grice[§], and Mamuka Kvaratskhelia*[¶]

*Ohio State University Health Sciences Center, College of Pharmacy, Center for Retrovirus Research and Comprehensive Cancer Center, Columbus, OH 43210; [†]Laboratory of Medicinal Chemistry and [§]HIV Drug Resistance Program, National Cancer Institute, Frederick, MD 21702; and [‡]National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892

Edited by John M. Coffin, Tufts University School of Medicine, Boston, MA, and approved March 31, 2004 (received for review February 6, 2004)

We report a methodology that combines affinity acetylation with MS analysis for accurate mapping of an inhibitor-binding site to a target protein. For this purpose, we used a known HIV-1 integrase inhibitor containing aryl di-*O*-acetyl groups (Acetylated-Inhibitor). In addition, we designed a control compound (Acetylated-Control) that also contained an aryl di-*O*-acetyl group but did not inhibit HIV-1 integrase. Examination of the reactivity of these compounds with a model peptide library, which collectively contained all 20 natural amino acids, revealed that aryl di-*O*-acetyl compounds effectively acetylate Cys, Lys, and Tyr residues. Acetylated-Inhibitor and Acetylated-Control exhibited comparable chemical reactivity with respect to these small peptides. However, these two compounds differed markedly in their interactions with HIV-1 integrase. In particular, Acetylated-Inhibitor specifically acetylated K173 at its inhibitory concentration (3 μM) whereas this site remained unrecognized by Acetylated-Control. Our data enabled creation of a detailed model for the integrase:Acetylated-Inhibitor complex, which indicated that the inhibitor selectively binds at an architecturally critical region of the protein. The methodology reported herein has a generic application for systems involving a variety of ligand–protein interactions.

HIV-1 encodes three enzymes that are essential for virus replication: reverse transcriptase, protease, and integrase (IN). To date, only the first two proteins have been exploited as therapeutic targets. The continuing emergence of new HIV-1 variants resistant to current treatments, together with cytotoxicity problems, makes the search for new anti-HIV-1 drugs imperative. HIV-1 IN is an attractive antiviral drug target that has no known human counterparts. Furthermore, the use of a common active site for 3′ end processing and DNA strand transfer constrain the range of mutations that can contribute to drug resistance without compromising catalytic activity.

HIV-1 IN catalyzes integration of the viral DNA, made by reverse transcription, into the host chromosome in a two-step reaction (reviewed in ref. 1). In the first step, called 3′ processing, two nucleotides are removed at each 3′ end of the viral DNA. In the next step, called DNA strand transfer, concerted transesterification reactions integrate the viral DNA ends into the host genome. *In vivo*, the enzyme acts in the context of a large nucleoprotein complex termed the "preintegration complex" (PIC). A number of viral proteins and host factors assemble within the PIC to orchestrate the integration process (2–10).

HIV-1 IN is composed of three distinct structural and functional domains: the N-terminal domain (residues 1–50), which contains the HHCC zinc-binding motif; the core domain (residues 51–212), which contains the catalytic site; and the C-terminal domain (residues 213–270), which is thought to provide a platform for DNA binding. Crystallographic or NMR structural data are available for each of the individual domains (11–15). In addition, two-domain crystal structures [either the core and C-terminal domains (16), or N-terminal and the core domains (17) of HIV-1 IN] have been recently determined. However, tremendous efforts to obtain a structure of the full-length protein have been impeded by poor protein solubility.

Systematic screening for potential inhibitors has been conducted predominantly by using purified IN-based assays. These studies have revealed several classes of compounds with anti-HIV-1 IN activity including diketo analogs, pyranodipyrimidines, nucleotide analogs, hydroxylated aromatic compounds, DNA-interacting agents, peptides, and antibodies (18–30). Whereas a number of promising drug candidates targeting HIV-1 IN are emerging, structural details concerning inhibitor:protein interactions lag substantially behind.

Crystallographic data on the interaction of HIV-1 IN with small molecules are very limited. Goldgur *et al*. (31) showed that a diketo group-containing inhibitor [1-(5-chloroindol-3-yl)-3-hydroxy-3-(2*H*-tetrazol-5-yl)-propenone] binds to the active site. However, this structure did not include the cognate DNA substrate, which is believed to contribute to the inhibition through formation of a ternary inhibitor:protein:DNA complex. In separate work, a small molecule (3,4-dihydroxyphenyltriphenylarsonium bromide) inhibitor was shown to bind at the HIV-1 IN dimer interface (32). These structural analyses were performed by using the core domain of IN rather than the full-length protein (31, 32). Therefore, there is a need to devise new approaches that are complementary to crystallography and NMR to enable rapid and accurate mapping of inhibitor:protein contacts under biologically relevant conditions.

Here, we report a methodology that combines affinity acetylation and MS analysis to map interactions of small molecule inhibitors with target proteins. In particular, we have found that aryl di-*O*-acetyl groups are capable of effectively acetylating Cys, Lys, and Tyr residues. Using a known HIV-1 IN inhibitor containing aryl di-*O*-acetyl groups, we have demonstrated a specific ability of the inhibitor to acetylate an architecturally critical region of the protein. Our methodology enabled analysis of inhibitor:protein interactions under biologically relevant conditions using reaction mixtures containing full-length IN, divalent metals, and DNA substrates. In addition to providing structural data at a single amino acid resolution, other benefits of this MS approach include the speed with which the analysis can be performed as well as a requirement for only very limited quantities of protein and inhibitor. These features lay a foundation for implementing high-throughput schemes.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: IN, integrase; MALDI-TOF, matrix-assisted laser desorption ionization/time-of-flight; PSD, post source decay; MS/MS, tandem MS.

[¶]To whom correspondence should be addressed at: Ohio State University, 500 West 12th Avenue, 238 L. M. Parks Hall, Columbus, OH 43210. E-mail: kvaratskhelia.1@osu.edu.

www.pnas.org/cgi/doi/10.1073/pnas.0400873101

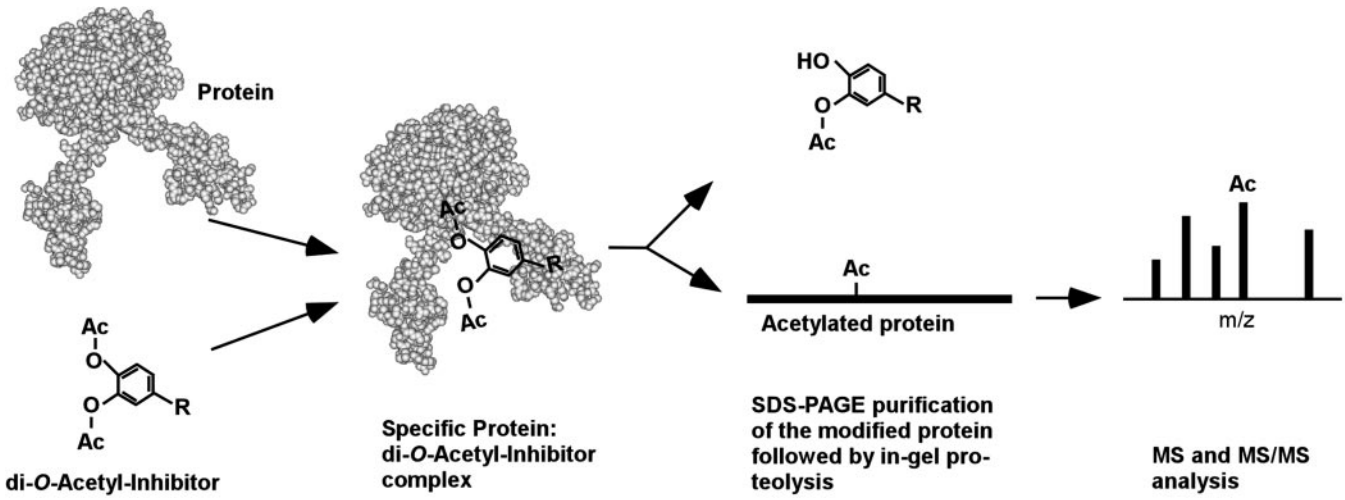


Fig. 1. Schematic representation of methodology. An aryl di-*O*-acetyl-containing inhibitor is incubated with target protein at $\approx IC_{50}$ concentration. Formation of a specific complex results in covalent modification of the protein at the inhibitor-binding site. Subsequent SDS/PAGE purification ensures unfolding of the protein, which is essential for effective proteolysis. Finally, specifically acetylated peptide fragments and individual residues are identified from MS and MS/MS analysis, respectively.

Materials and Methods

Preparation of Recombinant HIV-1 Integrase and Aryl Di-*O*-acetyl Compounds. Full-length integrase (F185K/C280S) was expressed in *Escherichia coli* and purified as described [see Jenkins *et al*. for detailed procedure (33)]. Acetylated-Inhibitor and Acetylated-Control were prepared according to the previously described protocol (34). Measurements of the *in vitro* IC_{50} values of the compounds were carried out as described (34).

Acetylation of Short Peptides and HIV-1 IN. The acetylation reactions for short peptides were performed in 50 mM Hepes (pH 8.0) buffer containing 5 μM peptide and acetyl compounds. Acetylation of HIV-1 IN was carried out in the same buffer that also included 50 mM NaCl, 5 mM $MgCl_2$, 1 μM DNA, 5 μM full-length protein, and acetyl compounds. The reactions were incubated at 37°C for 30 min and quenched by adding Lys in its free amino acid form (5 mM final concentration). The peptides were mixed with α-cyano-4-hydroxy-cinnamic acid and analyzed by matrix-assisted laser desorption ionization/time-of-flight (MALDI-TOF). HIV-1 IN was processed as follows. The protein was unfolded by adding 40 mM DTT and incubating at 70°C for 20 min. To modify Cys residues, 100 mM iodoacetamide was added to the mix, and the reaction was incubated at 25°C for 45 min. The process was quenched by addition of 100 mM DTT. The modified protein samples were subjected to SDS/PAGE. The integrase band was excised and digested with 1 μg of trypsin or Lys C (Roche) in 50 mM ammonium-bicarbonate buffer (pH 8.0) at 25°C for 16 h. Proteolytic peptide fragments were desiccated in a SpeedVac (Thermo Savant, Holbrook, NY) and then resuspended in 15 μl of aqueous 0.1% trifluoroacetic acid.

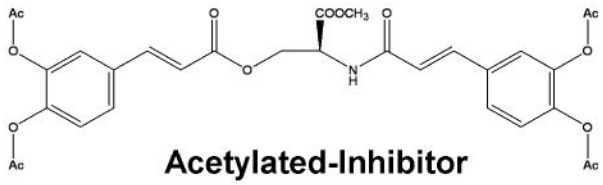
MS Analyses. Acetylation of short peptides was monitored with a Kratos MALDI-TOF instrument equipped with a curved field reflectron feature (Kratos Analytical Instruments) enabling generation of post source decay (PSD) data, i.e., amino acid sequence information. MALDI-TOF experiments were performed by using α-cyano-4-hydroxy-cinnamic acid as a matrix. MS and tandem MS (MS/MS) analyses of the proteolytic digests of HIV-1 IN were performed with a Waters Q-TOF-II instrument equipped with an electrospray source and a Waters cap-LC. The peptides were separated with a Waters Symmetrie300 precolumn and a MicroTech Scientific (Vista, CA) ZC-10-C18SBWX-150 column by using two sequential linear gradients of 5–40% and 40–90% of acetonitrile for 35 and 25 min, respectively.

Modeling of IN:Acetylated-Inhibtor Complex. The model was generated and minimized by using the Builder module of INSIGHTII (v. 2000.1, Accelrys, San Diego). Polar hydrogens and electrostatic potentials were added to the IN core structure (16) by using INSIGHTII. Docking was performed with AUTODOCK 3 (35). The grid was centered on the side chain of Tyr 99. One hundred dockings were calculated by using the Lamarckian genetic algorithm (35). All other parameters remained at their default values.

Results

Our experimental strategy for accurate mapping of inhibitor: protein contacts using affinity acetylation and MS is depicted in Fig. 1. Burke and coworkers have previously identified a new promising class of potent HIV-1 IN inhibitors, acetylated bis-caffeoyl derivatives of glycidic acid and amino acids such as serine and β-aminoalanine (34). These compounds emerged after significant structural changes in chicoric acid, such as substituting the dihydroxy-containing catechol groups with di-*O*-acetyl counterparts and replacing the two carboxylic groups with methyl ester. The goal of the previous study (34) was to develop new inhibitors with reduced toxicity and increased bioavailability. In the present work, we exploited Acetylated-Inhibitors as affinity modifiers that, together with MS, were envisaged to enable accurate mapping of the inhibitor-binding site on the protein. In particular, we postulated that acetyl groups might serve as tags that would be transferred from the inhibitor to the protein upon formation of a specific IN:inhibitor complex (Fig. 1). The IN residues specifically acetylated could subsequently be identified by MS (Fig. 1). We chose to test this scheme using the Acetylated-Inhibitor depicted in Fig. 2. This compound had been shown to be one of the most potent HIV-1 IN inhibitors ($IC_{50} \approx 3$ μM) among a large number of acetylated inhibitors tested (34). At the same time, we also used Acetylated-Control depicted in Fig. 2 to confirm that Acetylated-Inhibitor functioned as an affinity rather than as a nonspecific acetylator because the control did not exhibit any anti-HIV-1 IN activity at 300 μM (the highest concentration tested). Therefore, comparative analysis of Acetylated-Inhibitor and Acetylated-Control

BIOCHEMISTRY



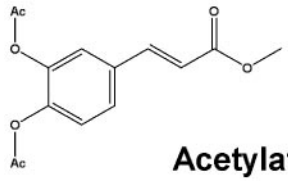


Fig. 2. Structures of Acetylated-Inhibitor and Acetylated-Control. Acetylated-Inhibitor impairs both HIV-1 integrase activities (3′ end processing and strand transfer) with very similar IC_{50} values of ≈ 3 μM. In contrast, no detectable inactivation of HIV-1 integrase activity was observed with 300 μM Acetylated-Control.

was intended to distinguish between specific and nonspecific acetylation.

Because acetylation of proteins with di-*O*-acetyl derivative has not been described previously, our initial efforts focused on revealing the amino acid residues that could be modified by these compounds. Therefore, we used commercially available peptides as a model system to rapidly identify acetylation sites. Five peptides were selected that collectively contained all 20 natural amino acids. The sequences of the peptides were as follows: peptide 1, HCKFWW; peptide 2, RGDSPASSKP; peptide 3, DRVYVHPF; peptide 4, HDMNKVLDL; peptide 5, CQDSETRTFY.

Formation of acetyl-derivatized peptides was monitored by MS. Representative data in Fig. 3*A* show that upon acetylation peptide 1 was converted into three distinct species. The molecular masses of the new peptide peaks peptide 1 plus 42 Da, peptide 1 plus 84 Da, and peptide 1 plus 126 Da corresponded precisely to the addition of one, two, and three acetyl groups to the peptide, respectively (Fig. 3*A*). Our next efforts focused on dissection of the modification sites within these newly formed peptides.

We exploited the curved field reflectron capability of our MALDI-TOF instrument to identify the acetylation sites in the peptides (see Fig. 3*B*) by analyzing various modified species as well as the original unmodified peptides. The representative PSD data of the triply acetylated peptide is depicted in Fig. 3. These results clearly indicate that acetyl groups are localized on Cys, Lys, and the peptide N terminus. PSD analysis of the peptide 1 plus 42 Da species indicated that Cys was a primary target of acetylation (data not shown). The fragmentation pattern of the peptide 1 plus 84 Da peak revealed acetylation of both Cys and Lys residues (data not shown), indicating that Lys is also very efficiently modified by di-*O*-acetyl compounds. Finally, we observed a triply acetylated peptide due to the peptide N terminus gaining an additional acetyl group (Fig. 3*B*). The data in Fig. 3*A* clearly show that the latter reaction was substantially less efficient than modification of Cys or Lys residues.

Identical detailed analyses were also performed with the other four peptides. Interestingly, these studies not only confirmed productive acetylation of Lys and Cys and residual modification of the N terminus but also revealed that Tyr residues were efficiently acetylated with our compounds. For example, peptide 3 gained two acetyl groups, with PSD analysis unequivocally identifying modifications of Tyr and the peptide N terminus. Peptide 5 gained three acetyl groups on Cys, Tyr, and the N

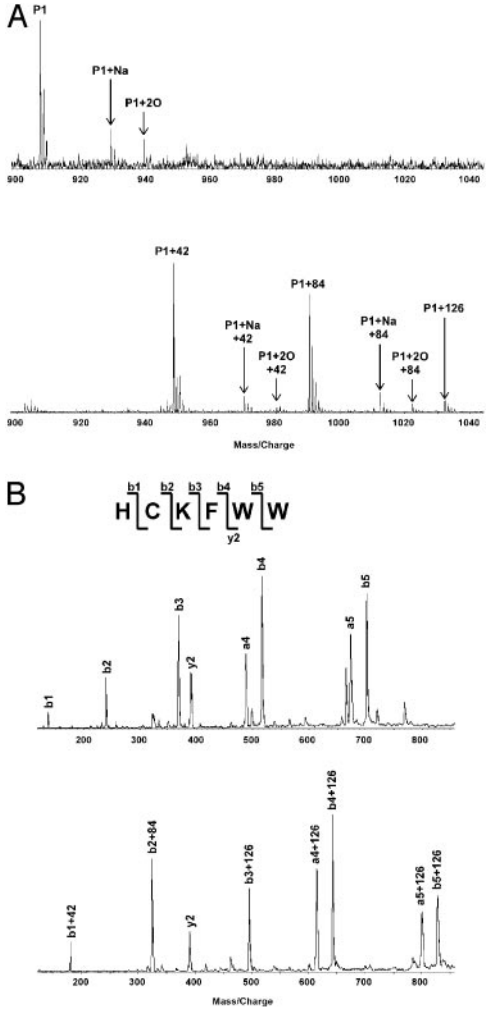


Fig. 3. MALDI-TOF data on acetylation of peptide 1 (P1). (*A*) MS data of the original peptide before (*Upper*) and after (*Lower*) treatment with Acetylated-Inhibitor. Incubation of P1 with Acetylated-Inhibitor resulted in the formation of new peptide peaks: P1 plus 42, P1 plus 84, and P1 plus 126, corresponding to the addition of one, two, and three acetyl groups, respectively. The original peptide preparation contained small amounts of sodium adduct (P1 plus Na) and an oxidized species (P1 plus 2O), which were also susceptible to acetylation (see *Lower*). (*B*) PSD data of the unmodified peptide: P1 (*Upper*) and the acetylated peak P1 plus 126 (*Lower*). The ''b'' and ''y'' ions peaks are derived from internal fragmentation of peptide bonds and provide amino acid sequence information read from the peptide N terminus to C terminus and from the C terminus to N terminus, respectively. The PSD data confirm the known amino acid sequence of this commercially available peptide. After acetylation (*Lower*) b1, b2, and b3 ions were shifted by 42, 84, and 126 Da, respectively. These data indicate acetylation of the N terminus, Cys and Lys. The 126-Da gap persisted for b4, b5, and the parent ion, indicating that there are only three acetylation sites in the peptide. The fact that the y2 ion remained unchanged confirms that two C-terminal Trp residues were not acetylated.

terminus. Thus, this screening approach using a model peptide library and MS analysis enabled us to identify the three amino acid residues Cys, Lys and Tyr as target sites in proteins that can effectively be modified by aryl di-*O*-acetyl group-containing compounds.

Our next goal was to compare the relative chemical reactivity of Acetylated-Inhibitor and Acetylated-Control with respect to short peptides. The representative data depicted in Fig. 4 demonstrate that these two compounds exhibited very similar reactivity with respect to peptide 4, which contained a single internal acetylation site (the Lys residue). In particular, high

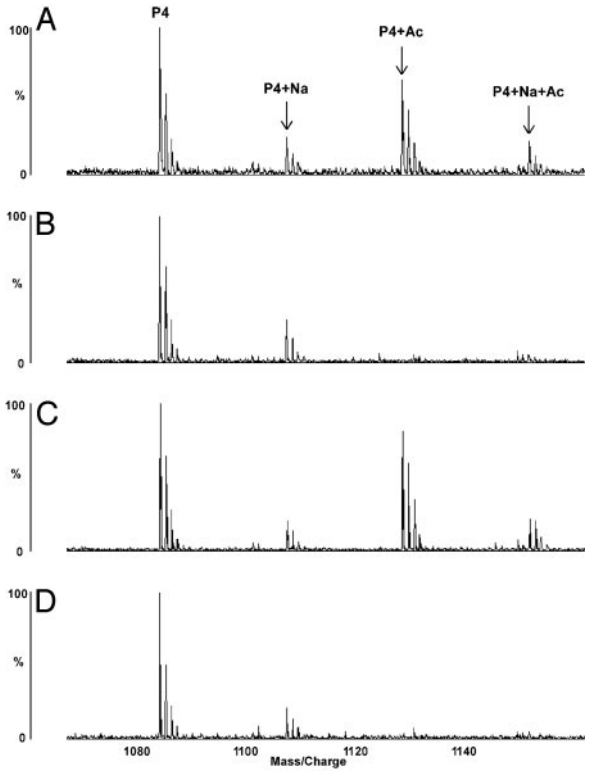


Fig. 4. MALDI-TOF data demonstrating similar chemical reactivities of Acetylated-Inhibitor and Acetylated-Control with respect to model peptide 4 (P4). (*A*) P4 plus 500 μM Acetylated-Inhibitor. (*B*) P4 plus 3 μm Acetylated-Inhibitor. (*C*) P4 plus 500 μM Acetylated-Control. (*D*) P4 plus 3 μm Acetylated-Control. PSD analysis revealed acetylation of Lys in the P4 plus Ac peak (data not shown). P4 plus Na and P4 plus Na plus Ac are sodium adducts of the original and acetylated peptides, respectively.

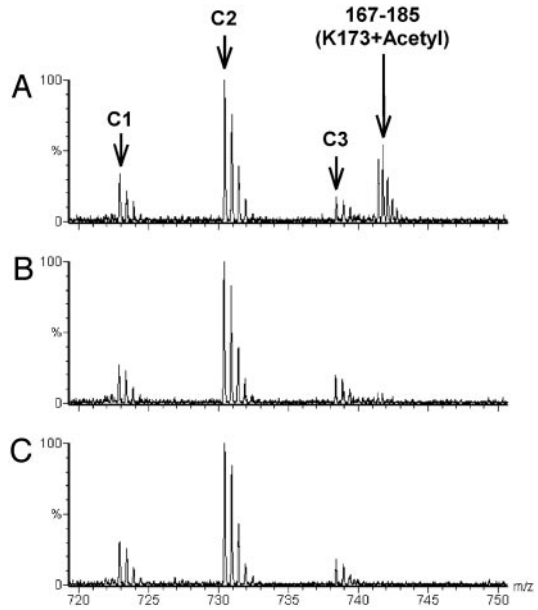


Fig. 5. MS data showing specific acetylation of HIV-1 IN with Acetylated-Inhibitor. (A*)* IN plus 3 μm Acetylated-Inhibitor. (*B*) IN plus 3 μm Acetylated-Control. (*C*) Free IN. MS spectra illustrate that treatment of IN with 3 μM Acetylated-Inhibitor resulted in a single new peak: 167–185 plus Acetyl (*A*). The molecular mass of this triply charged ion peak [741.03^{3+} = (741.03 − 1.008) × 3 = 2,220.13 Da] corresponds to that of the 167–185 peptide (molecular mass = 2,178.12 Da) plus one acetyl group (molecular mass = 42.03 Da). Importantly, this acetylated peptide peak was not formed in the presence of Acetylated-Control (*B*). C1, C2, and C3 represent unmodified tryptic peptides of HIV-1 IN.

concentrations of the compounds (≈500 μM) were necessary for efficient acetylation of the peptide whereas, at low μM concentrations of Acetylated-Inhibitor or Acetylated-Control, no acetylated peptide peaks were detected. Obviously, the short peptide fragments did not possess a tertiary structure that could be specifically recognized by Acetylated-Inhibitor whereas acetylation at high concentration was due to nonspecific interactions. Importantly, Acetylated-Inhibitor and Acetylated-Control exhibited comparable chemical reactivity with short peptides (Fig. 4), indicating that the latter compound could be used as a valid control when studying the interaction of Acetylated-Inhibitor with full-length HIV-1 IN.

We next analyzed interactions of the acetyl compounds with HIV-1 IN. The enzyme contains 5 Cys, 27 Lys, and 8 Tyr that represent "reporter" amino acids for inhibitor binding. Lys residues are the most abundant in HIV-1 IN, and their distribution throughout the protein provides an excellent environment for detailed analysis of modified sites. Mapping tryptic or Lys C peptides detected by MS revealed ≈98% coverage of the entire amino acid sequence of the protein. Thus, we were able to probe a large fraction of possible binding sites of Acetylated-Inhibitor to the full-length HIV-1 IN. Interestingly, treatment of HIV-1 IN with 3 μM (≈IC_{50} concentration) of Acetylated-Inhibitor resulted in formation of a single new peak in the MS spectra when compared with digestions of free IN (Fig. 5). The molecular mass of this peak (2,220.13 Da) corresponded to the 167–185 peptide fragment (2,178.12 Da) plus one acetyl group (42.03 Da). Importantly, this peak was not formed in the presence of Acetylated-Control (compare Fig. 5 *A* and *B*). Particularly striking is the contrast between Acetylated-Inhibitor and Acetylated-Control with respect to their interaction with HIV-1 IN protein (Fig. 5) whereas these compounds behaved identically with the short peptides (Fig. 4). These data indicate that the observed modification of IN was due to specific binding of Acetylated-Inhibitor to the protein whereas this site was unrecognized by Acetylated-Control.

Our next goal was to identify the modified residue within the IN peptide fragment 167–185 (DQAEHLKTAVQMAVFIHNK). This peptide contains two potential acetylation sites: K173 and K185. The former residue remained resistant to trypsin hydrolysis whereas K185 was readily digested by the protease. These data suggested acetylation of K173 because modified Lys residues are known to be resistant to trypsin cleavage. Unequivocal proof for K173 modification emerged from MS/MS analysis of the acetylated peak (Fig. 6).

In our experiments, we tested effects of divalent metal ions and DNA substrates on the acetylation of HIV-1 IN. The presence or omission of these components in the reaction mix had no detectable effect on specific modification of K173. The MS results enabled us to create a detailed model of the Acetylated-Inhibitor:IN complex (Fig. 7). Our model reveals that the Acetylated-Inhibitor specifically binds at an architecturally critical region of HIV-1 IN, which is located at the protein dimer interface (Fig. 7). Of four *O*-acetyl groups present on Acetylated-Inhibitor, only one is positioned to modify a "reporter" residue, K173. Thus, our model and experimental observations are consistent.

Discussion

We report a methodology for accurate mapping of inhibitor: protein interactions that combines affinity acetylation and MS analysis. Using a library of model peptides, we found that aryl di-*O*-acetyl compounds can productively acetylate Cys, Lys, and

BIOCHEMISTRY

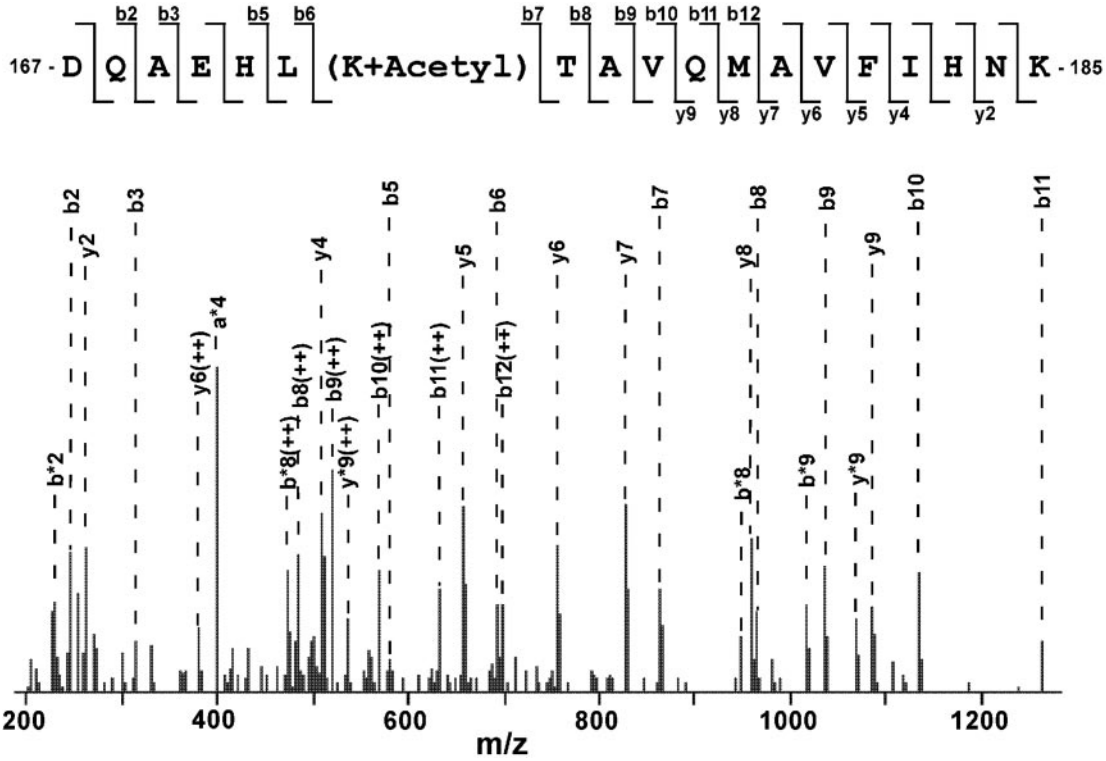


Fig. 6. MS/MS data on the acetylated IN peptide (167–185 plus Acetyl). Internal fragmentation of the peptide provided its amino acid sequence information. Sequential fragmentation of peptide bonds from the peptide N and C terminus resulted in ''b'' and ''y'' ions, respectively (*Upper*). *, These peaks were generated due to the loss of the NH_3 group from fragments. ++, Doubly charged ions (all other fragments possessed a single charge). The mass increment between ''b6'' and ''b7'' ions corresponds to an acetylated lysine (K173 plus Acetyl). Moreover, the masses of ions preceding and following ''b'' or ''y'' ions assign perfectly to the IN sequence only when K173 is considered to be acetylated.

Tyr residues. We exploited this chemical reactivity to study interaction of Acetylated-Inhibitor with HIV-1 IN. Importantly, our analyses were performed under biologically relevant conditions with reaction mixtures containing full-length IN enzyme, divalent metal ions, and DNA substrate.

Our data indicate that the inhibitor specifically binds at the protein dimer interface, modifying K173. Whereas this particular residue serves as a reporter amino acid for inhibitor binding, the model depicted in Fig. 7 reveals interactions of Acetylated-Inhibitor with a number of architecturally important residues (K173, T174, and M178). Sire and coworkers have recently demonstrated that viruses with IN mutated within this region (172/173, 174, 178, and 180/181) are impaired for HIV-1 replication (36). It has also been reported that the peptide reproducing the IN amino acid sequence segment, residues 167–187, strongly inhibits enzyme activity by interfering with protein oligomerization (37). It is now intriguing to reveal how a small molecule such as Acetylated-Inhibitor may affect the IN architecture.

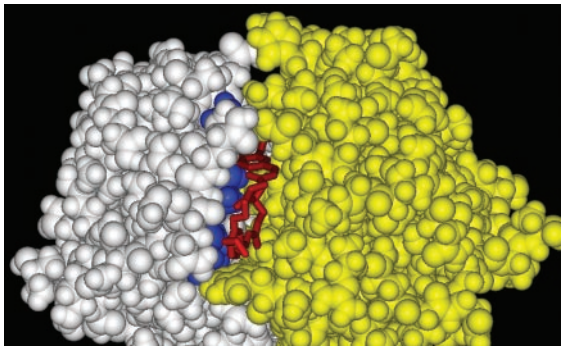


Fig. 7. A model of IN:Acetylated-Inhibitor complex. The spacefilling model shows that the inhibitor (colored red) binds at the protein dimer interface contacting residues K173, T174, and M178 (all colored in blue). Separate protein monomers are colored in yellow and white.

Alterations introduced in the IN structure upon Acetylated-Inhibitor binding may be detrimental for a number of protein: protein interactions that are essential for viral replication. For example, Sire and coworkers (36) have shown that the region of HIV-1 IN encompassing residues 170–181 is important for packaging of the host uracil DNA glycosylase (UNG2) enzyme into virions. In fact, the integration process *in vivo* is orchestrated by a number of viral and cellular proteins through their interaction with IN. Therefore, identification of inhibitors selectively disrupting multiprotein complexes essential for the virus replication is of significant interest.

Our finding that Acetylated-Inhibitor binds with high affinity at an architecturally critical region of HIV-1 IN *in vitro* highlights the potential for further development of this compound as a selective IN inhibitor. Indeed, a very limited number of unique structural pockets present in this enzyme have complicated development of commercially available drugs targeting HIV-1 IN. Therefore, a specific inhibitor-binding pocket, depicted in this work, represents an interesting structural target.

Comparison of Acetylated-Inhibitor with 3,4-dihydroxyphenyltriphenylarsonium bromide, an HIV-1 IN inhibitor that has been shown to also bind at the IN dimer interface, is interesting (32). Our data indicate that Acetylated-Inhibitor binds deep into the dimer interface pocket owing to its ''slim'' structure (see Figs. 2 and 7). 3,4-Dihydroxyphenyltriphenylarsonium bromide binds to the same region of the protein, but the bulky tetra phenyl

structure limits interaction of this compound to the outer surface of the dimer interface (32). The IC_{50} values for Acetylated-Inhibitor and 3,4-dihydroxyphenyltriphenylarsonium bromide are $\approx$3 μM and $\approx$150 μM, respectively (32). Taken together, these findings enable us to begin to observe a correlation between structural interactions and inhibitory potency, which is important for better understanding of HIV-1 IN as a therapeutic target.

Comparison of Acetylated-Inhibitor with chicoric acid is also noteworthy. Robinson and coworkers (38, 39) reported that HIV-1 resistance to chicoric acid results from a single mutation of integrase amino acid G140. Pluymers *et al*. (40) have shown that the primary target of chicoric acid is viral envelope glycoprotein gp 120, indicating a dual function for this compound. Because our affinity acetylation approach is limited to modification of Cys, Lys, and Tyr residues, we were unable to detect whether Acetylated-Inhibitor could also bind to G140 of HIV-1 IN, leaving open the possibility of multiple binding sites. However, it is important to point out that Acetylated-Inhibitor and chicoric acid exhibit key structural differences. For example, Acetylated-Inhibitor contains one methyl ester group originating from the central linker region (Fig. 2) whereas chicoric acid possesses two carboxylic acid moieties on its linker segment. Previous studies have shown that structural variations in this region could markedly alter functionality of IN inhibitors (34). Therefore, the mechanism of inhibition of IN with Acetylated-Inhibitor could significantly differ from that of chicoric acid.

Our MS assays can now be applied to examine interactions of a large number of different di-*O*-acetyl inhibitors with HIV-1 IN to dissect the structural features essential for specific recognition of the architecturally critical IN protein dimer interface. These experiments may also reveal new interesting pockets in HIV-1 IN that could be targeted by small molecule inhibitors. Equally, the methodology described herein could be exploited to probe the interaction sites of acetylated compounds with respect to other retroviral proteins such as gp120. These studies could lead to identification of new drug-binding sites and elucidation of a detailed mechanism of inhibition, and thereby facilitate developing improved antiviral therapies.

In summary, we have demonstrated that combining affinity acetylation with MS analysis is a powerful tool for accurate mapping of inhibitor:protein interactions. In so doing, we have developed an effective model system where commercially available short peptides are used to rapidly dissect the chemistry of new covalent modifiers. This technology has a generic application for studying a wide variety of inhibitor:protein interactions. The ability to perform high-resolution structural analysis of inhibitor:protein complexes under biologically relevant conditions using only very limited sample quantities makes such an MS approach very attractive and opens a doorway for implementing high-throughput schemes.

This study has been supported by American Cancer Society Grant IRG-98-278-03 (to M.K.). Funding was also provided in part by the National Cancer Institute's intramural Center for Cancer Research, which supports the HIV Drug Resistance Program.

1. Brown, P. O. (1997) in *Retroviruses*, eds. Coffin, J. M., Hughes, S. H. & Varmus, H. E. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 161–204.
2. Bukrinsky, M. I., Sharova, N., McDonald, T. L., Pushkarskaya, T., Tarpley, W. G. & Stevenson, M. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 6125–6129.
3. Buckman, J. S., Bosche, W. J. & Gorelick, R. J. (2003) *J. Virol.* **77,** 1469–1480.
4. Carteau, S., Batson, S. C., Poljak, L., Mouscadet, J. F., de Rocquigny, H., Darlix, J. L., Roques, B. P., Kas, E. & Auclair, C. (1997) *J. Virol.* **71,** 6225–6229.
5. Carteau, S., Gorelick, R. J. & Bushman, F. D. (1999) *J. Virol.* **73,** 6670–6679.
6. Chen, H. & Engelman, A. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 15270–15274.
7. Farnet, C. M. & Bushman, F. D. (1997) *Cell* **88,** 483–492.
8. Kalpana, G. V., Marmon, S., Wang, W., Crabtree, G. R. & Goff, S. P. (1994) *Science* **266,** 2002–2006.
9. Lee, M. S. & Craigie, R. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 9823–9827.
10. Lee, M. S. & Craigie, R. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 1528–1533.
11. Cai, M., Zheng, R., Caffrey, M., Craigie, R., Clore, G. M. & Gronenborn, A. M. (1997) *Nat. Struct. Biol.* **4,** 567–577.
12. Dyda, F., Hickman, A. B., Jenkins, T. M., Engelman, A., Craigie, R. & Davies, D. R. (1994) *Science* **266,** 1981–1986.
13. Goldgur, Y., Dyda, F., Hickman, A. B., Jenkins, T. M., Craigie, R. & Davies, D. R. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 9150–9154.
14. Eijkelenboom, A. P., Lutzke, R. A., Boelens, R., Plasterk, R. H., Kaptein, R. & Hard, K. (1995) *Nat. Struct. Biol.* **2,** 807–810.
15. Lodi, P. J., Ernst, J. A., Kuszewski, J., Hickman, A. B., Engelman, A., Craigie, R., Clore, G. M. & Gronenborn, A. M. (1995) *Biochemistry* **34,** 9826–9833.
16. Chen, J. C., Krucinski, J., Miercke, L. J., Finer-Moore, J. S., Tang, A. H., Leavitt, A. D. & Stroud, R. M. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 8233–8238.
17. Wang, J. Y., Ling, H., Yang, W. & Craigie, R. (2001) *EMBO J.* **20,** 7333–7343.
18. Nicklaus, M. C., Neamati, N., Hong, H., Mazumder, A., Sunder, S., Chen, J., Milne, G. W. & Pommier, Y. (1997) *J. Med. Chem.* **40,** 920–929.
19. Neamati, N., Marchand, C. & Pommier, Y. (2000) *Adv. Pharmacol.* **49,** 147–165.
20. Zhang, X., Neamati, N., Lee, Y. K., Orr, A., Brown, R. D., Whitaker, N., Pommier, Y. & Burke, T. R., Jr. (2001) *Bioorg. Med. Chem.* **9,** 1649–1657.
21. Zhao, L., O'Reilly, M. K., Shultz, M. D. & Chmielewski, J. (2003) *Bioorg. Med. Chem. Lett.* **13,** 1175–1177.
22. Pommier, Y., Marchand, C. & Neamati, N. (2000) *Antiviral Res.* **47,** 139–148.
23. De Clercq, E. (2001) *Curr. Med. Chem.* **8,** 1543–1572.
24. Lipford, J. R., Worland, S. T. & Farnet, C. M. (1994) *J. Acquired Immune Defic. Syndr.* **7,** 1215–1223.
25. Mazumder, A., Neamati, N., Sommadossi, J. P., Gosselin, G., Schinazi, R. F., Imbach, J. L. & Pommier, Y. (1996) *Mol. Pharmacol.* **49,** 621–628.
26. Yi, J., Cheng, H., Andrake, M. D., Dunbrack, R. L., Jr., Roder, H. & Skalka, A. M. (2002) *J. Biol. Chem.* **277,** 12164–12174.
27. Pannecouque, C., Pluymers, W., Van Maele, B., Tetz, V., Cherepanov, P., De Clercq, E., Witvrouw, M. & Debyser, Z. (2002) *Curr. Biol.* **12,** 1169–1177.
28. Hazuda, D. J., Felock, P., Witmer, M., Wolfe, A., Stillmock, K., Grobler, J. A., Espeseth, A., Gabryelski, L., Schleif, W., Blau, C. & Miller, M. D. (2000) *Science* **287,** 646–650.
29. Espeseth, A. S., Felock, P., Wolfe, A., Witmer, M., Grobler, J., Anthony, N., Egbertson, M., Melamed, J. Y., Young, S., Hamill, T., *et al*. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 11244–11249.
30. Grobler, J. A., Stillmock, K., Hu, B., Witmer, M., Felock, P., Espeseth, A. S., Wolfe, A., Egbertson, M., Bourgeois, M., Melamed, J., Wai, J. S., *et al*. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 6661–6666.
31. Goldgur, Y., Craigie, R., Cohen, G. H., Fujiwara, T., Yoshinaga, T., Fujishita, T., Sugimoto, H., Endo, T., Murai, H. & Davies, D. R. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 13040–13043.
32. Molteni, V., Greenwald, J., Rhodes, D., Hwang, Y., Kwiatkowski, W., Bushman, F. D., Siegel, J. S. & Choe, S. (2001) *Acta Crystallogr. D Biol. Crystallogr.* **57,** 536–544.
33. Jenkins, T. M., Engelman, A., Ghirlando, R. & Craigie, R. (1996) *J. Biol. Chem.* **271,** 7712–7718.
34. Lin, Z., Neamati, N., Zhao, H., Kiryu, Y., Turpin, J. A., Aberham, C., Strebel, K., Kohn, K., Witvrouw, M., Pannecouque, C., *et al*. (1999) *J. Med. Chem.* **42,** 1401–1414.
35. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K. & Olson, A. J. (1998) *J. Comput. Chem.* **19,** 1639–1662.
36. Priet, S., Navarro, J. M., Querat, G. & Sire, J. (2003) *J. Biol. Chem.* **278,** 20724–20730.
37. Maroun, R. G., Gayet, S., Benleulmi, M. S., Porumb, H., Zargarian, L., Merad, H., Leh, H., Mouscadet, J. F., Troalen, F. & Fermandjian, S. (2001) *Biochemistry* **40,** 13840–13848.
38. King, P. J. & Robinson, W. E., Jr. (1998) *J. Virol.* **72,** 8420–8424.
39. King, P. J., Lee, D. J., Reinke, R. A., Victoria, J. G., Beale, K. & Robinson, W. E., Jr. (2003) *Virology* **306,** 147–161.
40. Pluymers, W., Neamati, N., Pannecouque, C., Fikkert, V., Marchand, C., Burke, T. R., Jr., Pommier, Y., Schols, D., De Clercq, E., Debyser, Z. & Witvrouw, M. (2000) *Mol. Pharmacol.* **58,** 641–648.