
DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU


HARVARD LIBRARY
Office for Scholarly Communication

# Integration Requires a Specific Interaction of the Donor DNA Terminal 5#-Cytosine with Glutamine 148 of the HIV-1 Integrase Flexible Loop

## Citation

Johnson, Allison A., Webster Santos, Godwin C. G. Pais, Christophe Marchand, Ronak Amin, Terrence R. Burke, Gregory Verdine, and Yves Pommier. 2006. "Integration Requires a Specific Interaction of the Donor DNA Terminal 5#-Cytosine with Glutamine 148 of the HIV-1 Integrase Flexible Loop." Journal of Biological Chemistry281 (1): 461–67. https://doi.org/10.1074/jbc.M511348200.

## Permanent link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:41511259

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

# Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 1, pp. 461–467, January 6, 2006
Printed in the U.S.A.

# Integration Requires a Specific Interaction of the Donor DNA Terminal 5′-Cytosine with Glutamine 148 of the HIV-1 Integrase Flexible Loop*

Received for publication, October 19, 2005 Published, JBC Papers in Press, October 27, 2005, DOI 10.1074/jbc.M511348200

**Allison A. Johnson[‡], Webster Santos[§], Godwin C. G. Pais[¶], Christophe Marchand[‡], Ronak Amin[‡], Terrence R. Burke, Jr.[¶], Gregory Verdine[§], and Yves Pommier[‡1]**

*From the [‡]Laboratory of Molecular Pharmacology and [¶]Laboratory of Medicinal Chemistry, Center for Cancer Research, NCI, National Institutes of Health, Bethesda, Maryland 20892 and the [§]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138*

**Integration is essential for retroviral replication and gene therapy using retroviral vectors. Human immunodeficiency virus, type 1 (HIV-1), integrase specifically recognizes the terminal sequences of each long terminal repeat (LTR) and cleaves the 3′-end terminal dinucleotide 5′-GT. The exposed 3′-hydroxyl is then positioned for nucleophilic attack and subsequent strand transfer into another DNA duplex (target or chromosomal DNA). We report that both the terminal cytosine at the protruding 5′-end of the long terminal repeats (5′-C) and the integrase residue Gln-148 are critical for strand transfer. Proximity of the 5′-C and Gln-148 was demonstrated by disulfide cross-linking. Cross-linking is inhibited by the inhibitor 5CITEP 1-(5-chloroindol-3-yl)-3-hydroxy-3-(2*H*-tetrazol-5-yl)-propenone. We propose that strand transfer requires a conformational change of the integrase-viral (donor) DNA complex with formation of an H-bond between the N-3 of the 5′-C and the amine group of Gln-148. These findings have implications for the molecular mechanisms coupling 3′-processing and strand transfer as well as for the molecular pharmacology of integrase inhibitors.**

Integration of the HIV-1[2] DNA into a host chromosome is required to complete viral infection. Integration is also required for retroviral gene transfer. HIV-1 integrase (integrase) catalyzes integration in two steps (1–3). First, an endonucleolytic cleavage reaction releases a terminal 5′-GT dinucleotide from the end of each viral long terminal repeat (LTR) in a reaction called 3′-processing (3′-P). Subsequently, the viral and host DNAs are joined by insertion of each exposed viral DNA 3′-hydroxyl end into a host chromosome (strand transfer, ST). The integrase-DNA complex during ST consists of several integrase monomers, the two ends of the viral DNA, and the target DNA. Although the arrangement of the integrase dimers has been elucidated by x-ray diffraction of the structures of various integrase domains (4–10), determination of the structure of an integrase-DNA co-crystal has remained elusive.

HIV-1 integrase binds to the tips ("att" sites) of each viral LTR. The roles of several bases at the extremities of the viral DNA have been examined. The release of the 5′-GT dinucleotide during 3′-P leaves a complementary 5′-AC overhang (see Fig. 1*A* for the sequence of the terminal 21 bp of the HIV-1 U5 LTR). The 5′-AC overhang is required for ST (11) and has been proposed to stabilize the viral DNA end in the proper position for ST (11, 12). Yet the roles of these individual nucleotides, as well as their functional association with an integrase amino acid during ST have not been established.

A flexible loop consisting of amino acids 140–149 resides over the integrase active site. The conformational flexibility of this loop is suggested to be important for the catalytic step following DNA binding (4). Amino acids Gln-148 and Tyr-143 have been identified through photocross-linking as possible contacts for the 5′-adenine of the 5′-AC overhang (13). Additionally, a Q148L integrase mutation reduces the stability of the integrase-viral DNA complex as well as ST activity (14). Together, these results suggest that integrase Gln-148 resides in the region of the 5′-AC overhang and plays a role in the ST reaction.

The aim of the present study was to elucidate the molecular interactions between integrase and the viral LTR DNA ends. We find the cytosine in the 5′-AC overhang (hereafter referred to as the "5′-C", Fig. 1*A*) and integrase Gln-148 to be critical for ST especially in the presence of magnesium, the probable biological metal cofactor. We demonstrate the proximity of the 5′-C and Gln-148 through disulfide cross-linking, and propose that ST requires the formation of a hydrogen bond between the cytosine N-3 and the glutamine amine group. These results directly relate to the mechanism and structure of the ST reaction as well as to the mechanism of action of integrase inhibitors that block 3′-P and ST.

## EXPERIMENTAL PROCEDURES

*Oligonucleotide Synthesis*—Oligonucleotides, except those used for disulfide cross-linking, were commercially synthesized by (Integrated DNA Technologies, Coralville, IA). O-4-Triazolyl-dU-CE phosphoramidite was purchased from Glen Research (Sterling, VA). Oligonucleotides containing convertible nucleosides were synthesized on an Applied Biosystems 392 DNA synthesizer. The two carbon tether (cystamine) was added post-synthetically using the convertible nucleoside approach (15), reduced with excess dithiothreitol, and purified using a NAP-5 column. Incubation of the DNA with a 10-fold molar excess of 5,5′-dithiobis(2-nitrobenzoic acid) (Sigma) using pH 8.5 phosphate buffer and subsequent ethanol precipitation provided the activated DNA (16). The activated DNA was annealed with the complementary strand before cross-linking experiments were performed.

*Mutagenesis*—Integrase mutants were created using the Stratagene QuickChange site-directed mutagenesis kit (La Jolla, CA), according to the manufacturer's recommendations. Primers containing mutations were as follows: for C56S, 5′-CAAGTAGACAGTAGCCCAGGA-3′;

* This research was supported in part by the Intramural Research Program of NCI, Center for Cancer Research, National Institutes of Health. 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.

[1] To whom correspondence should be addressed: Laboratory of Molecular Pharmacology, Bldg. 37, Rm. 5068, National Institutes of Health, Bethesda, MD 20892. Tel.: 301-496-5944; Fax: 301-402-0752; E-mail: pommier@nih.gov.

[2] The abbreviations used are: HIV-1, human immunodeficiency virus, type 1; Mops, 4-morpholinepropanesulfonic acid; 3′-P, 3′-processing; 5′-C, referring to the cytosine closest to the 5′-end of the HIV-1 U5 LTR; LTR, long terminal repeat; ST, strand transfer; SSS, triple serine; 5CITEP, 1-(5-chloroindol-3-yl-3-hydroxy-3-(2*H*-tetrazol-5-yl)-propenone.

Downloaded from http://www.jbc.org/ by guest on October 8, 2019

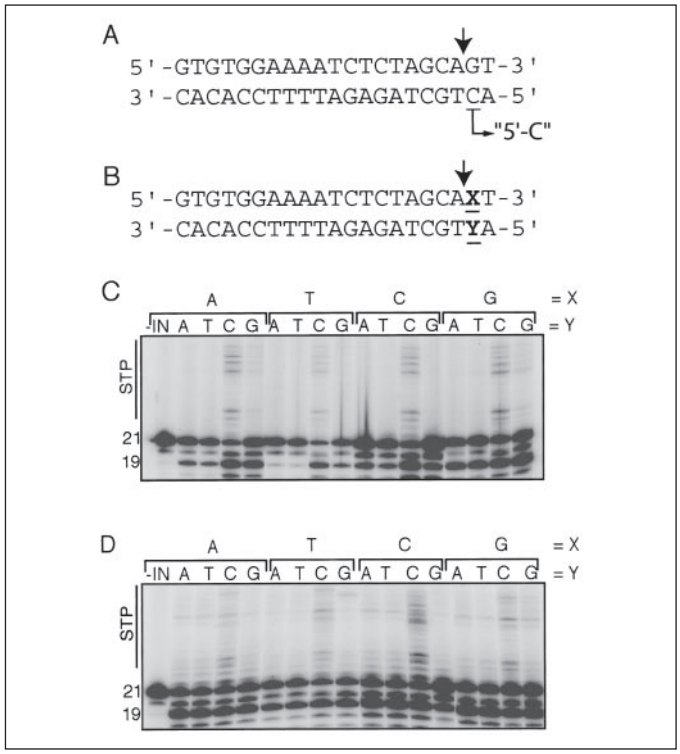


FIGURE 1. **Importance of the 5′-end cytosine for integrase ST activity.** *A,* sequence of the terminal 21 bp of the HIV-1 U5 LTR. *B,* positions of base changes (*X* and *Y*, in *boldface and underlined*). The *vertical arrows* indicate the 3′-processing site. *C,* gel showing the effect of the 16 possible base pair combinations on integrase activity in the presence of magnesium. Each lane contains a different DNA substrate as indicated by *X* and *Y* (see *B*). For simplicity, the only DNA shown before reaction with integrase is the *X*/*Y* = A/A substitution (*lane 1*). The other 15 DNAs appeared similar prior to addition of integrase. *D,* same as *C* except that manganese was used instead of magnesium as metal cofactor.

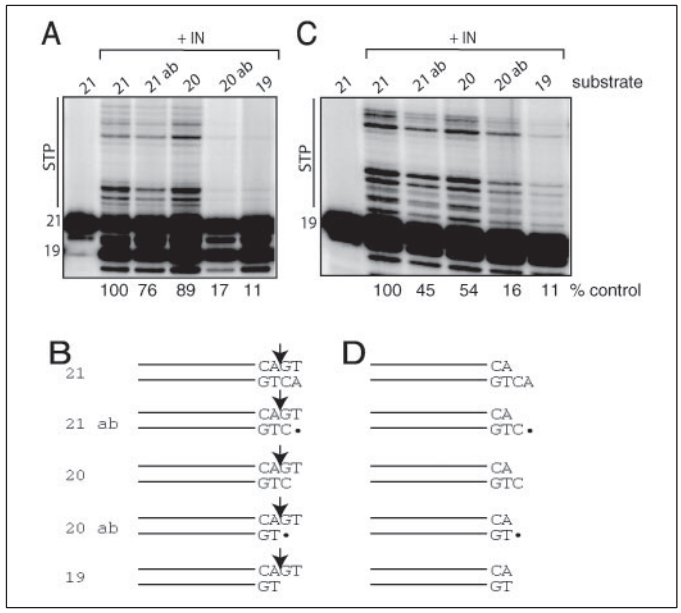


FIGURE 2. **Requirement of 5′-C for strand transfer.** Integrase (*IN*) reactions are shown in *A* and *C* (full-length and precleaved upper strand, respectively). DNA substrates are described in *B* and *D* (full-length and precleaved upper strand, respectively). An abasic site is designated by a *dot*. *21 ab* is the full-length lower strand with an abasic adenosine as the 5′-nucleotide. *20* is the lower strand without the entire 5′-terminal adenosine. *20 ab* is the lower strand without the 5′-terminal adenosine and the neighboring cytosine base. *19* is the lower strand without the two 5′-terminal nucleotides. Note that removal of the 5′-C, as in substrates *20 ab* and *19*, selectively blocked ST without changing 3′-P. Strand transfer efficiency for each substrate is indicated *below* each lane in *A* and *C*.

for C65S, 5′-GGCAGCTAGATTCTACACATTTAG-3′; for C280S, 5′-GGCAGGTGATGATAGTGTGGCAAGTAG-3′; for Q148N, 5′-CCCCAAAGTAACGGGGTAATAG-3′; for Q148C, 5′-CCCCAAAGTTGCGGGGTAATAG-3′. A complementary primer was used for each mutant. The presence of desired mutations and the integrity of the remainder of the integrase sequence were verified by DNA sequencing.

*Integrase Purification*—Recombinant wild-type or mutant integrase was purified from *Escherichia coli* as described (17) with the addition of 10% glycerol to all buffers. Reducing agents were omitted from the elution and dialysis steps of proteins used in disulfide cross-linking reactions.

*Integrase Reactions*—Integrase reactions contained 400 nM integrase, 20 nM DNA, 7.5 mM $MgCl_2$ or $MnCl_2$, 7.5 mM NaCl, 14 mM 2-mercaptoethanol, and 20 mM Mops, pH 7.2. Magnesium was used as the divalent metal, unless otherwise noted. Reactions were at 37 °C for 1 h and were quenched by addition of an equal volume of gel loading dye (formamide containing 0.25% bromphenol blue and xylene cyanol). Products were separated on 20% polyacrylamide denaturing sequencing gels. Dried gels were visualized using a 445 SI PhosphorImager (Amersham Biosciences). Densitometry analyses were performed using ImageQuant software from Amersham Biosciences.

*Disulfide Cross-linking*—Integrase (10 $\mu$M) was incubated with DNA duplexes (10 $\mu$M) containing tethered thiols in the presence of 20 mM Tris, pH 7.4, 7.5 mM $MgCl_2$, and 10% glycerol for 20 min (unless otherwise noted) at 37 °C. Reactions were capped by the addition of 20 mM methyl methanethiosulfonate. Nonreducing gel loading buffer (100 mM Tris-Cl, pH 6.8, 4% SDS, 0.2% bromphenol blue, 20% glycerol) was added, and samples were heated at 95 °C prior to loading onto 16% polyacrylamide gels (Invitrogen). Gels were stained with Microwave Blue according to the manufacturer's recommendations (Protiga, Gaithersburg, MD).

Alternatively, 500 nM integrase was incubated with 5CITEP in the presence of 20 mM Tris, pH 7.4, 7.5 mM $MgCl_2$, and 10% glycerol for 20 min. DNA (20 nM) containing a 5′-$^{32}$P label on one strand and a thiol-modified cytosine on the other strand was added, and reactions were capped with methyl methanethiosulfonate at 0.016, 0.33, 0.66, 1, 2, 5, 7.5, 10, 15, and 20 min or at 1 min (see Fig. 6). Following capping, nonreducing gel loading buffer was added, and samples were directly loaded on 16% polyacrylamide gels.

*Schiff Base Cross-linking*—Inhibition of DNA binding by 5CITEP was examined by using 5′-$^{32}$P-labeled oligonucleotide containing an abasic site (*X*) substitution for adenine of the conserved 5′-CA, 5′-GTGTGGAAAATCTCTAGC*X*GT annealed to the thiol-modified DNA. Q148C/SSS integrase (500 nM) was incubated with 111 $\mu$M 5CITEP for 15 min at room temperature in the presence of 50 mM Mops, pH 7.2, 143 mM 2-mercaptoethanol, and 7.5 mM $MgCl_2$. Schiff base DNA (20 nM) was added, and the reaction was incubated at 37 °C for 10 min. 100 mM $NaHB_4$ was added to stabilize the Schiff base cross-link, followed by an equal volume of SDS-polyacrylamide gel loading buffer. Samples were heated at 95 °C for 3 min and electrophoresed by SDS-PAGE. For a more detailed description of the assay see Ref. 18.

## RESULTS

*The 5′-C Is Required for ST*—Integrase 3′-P and ST activities can be measured by using a duplex DNA derived from the terminal 21 bp of the HIV-1 U5 LTR (Fig. 1*A*). The base pair on the 3′-side of the integrase 3′-P site, which normally is guanine/cytosine, was replaced by each of the other 15 possible base pair combinations at the *X*/*Y* position (Fig. 1*B*) to examine the importance of the base pair sequence for the 3′-P and ST reactions. By using a full-length substrate, the 3′-P reaction must

Downloaded from http://www.jbc.org/ by guest on October 8, 2019

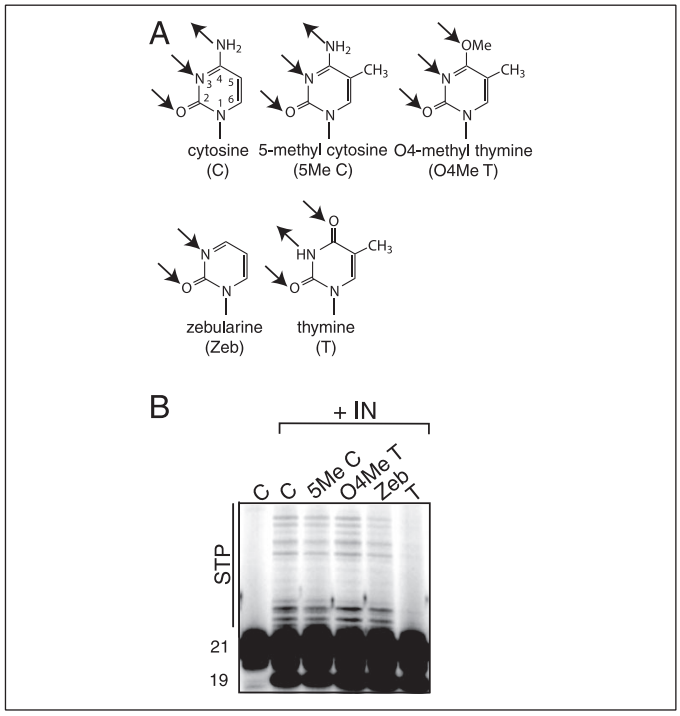


FIGURE 3. **Requirement for ST of cytosine N-3 as hydrogen bond acceptor.** The modified bases were introduced at position *Y* in Fig. 1*B* (*X* = G). *A,* structures of modified bases used. *Inward* and *outward arrows* indicate acceptor and donor hydrogen bonds, respectively. *B,* representative gel shows the effects of base modifications on integrase (*IN*) ST and 3′-P.

occur prior to ST. Only two combinations fully inhibited 3′-P (T/A and T/T) and therefore also inhibited subsequent ST in magnesium-containing reactions (Fig. 1*C*). The presence of an upper strand guanine (*X* = G) or a lower strand cytosine (*Y* = C), both corresponding to the natural sequence, led to the most efficient 3′-P in the presence of magnesium. Notably, cytosine as the lower strand base (*Y* = C) was strongly preferred for ST, regardless of the identity of the upper strand base. When reactions were performed with manganese, the DNA base requirements for 3′-P and ST were less stringent (Fig. 1*D*). The identity of the base pair had no effect on 3′-P, and all of the base combinations yielded some level of ST. Yet, even with manganese, the 5′-C on the lower strand was preferred for ST. Removal of the terminal 5′-nucleotide (dA; see Fig. 1*A*) had no effect on ST (Fig. 2). By contrast, when the 5′-C was removed, integrase failed to produce ST products but retained a normal level of 3′-P (Fig. 2). These results demonstrate the selective requirement of the 5′-C for ST.

To examine which portion of the 5′-C confers efficient ST, exocyclic substitutions were introduced sequentially to convert the base from cytosine (preferred) to thymine (nonpermissive) (Fig. 3*A*). Addition of a 5-methyl group had no effect on ST (Fig. 3*B*). Conversion of the 4-amine to an O-4-methyl group resulted in a slight enhancement of the ST reaction. The O-4-methyl group can be a weak hydrogen bond acceptor compared with the hydrogen bond donating amine group of cytosine. ST remained effective without any substitution at position 4 (zebularine), indicating this position does not play a role in ST. Most critically, ST was lost when the N-3 group was changed to NH, as in thymine, which reverses the direction of hydrogen bonding at this position. Also, the ST in the presence of the O-4-methyl group indicates the thymine carbonyl is not inhibitory to ST (Fig. 3). These results suggest the N-3 group, as a hydrogen bond acceptor, makes a critical contact with integrase or the target DNA for ST.

**TABLE 1**

**Alignment of flexible loop residues for selected retroviruses**

The base refers to the identity of the base corresponding to the 5′-C in the HIV LTR (see Fig. 1*A*). *Boldface letters* indicate amino acid residues differing from the HIV sequence. The underline indicates the conserved glutamine (Gln-148 in HIV-1). SIV indicates simian immunodeficiency virus; FIV indicates feline immunodeficiency virus; ASV indicates avian sarcoma virus; and RSV indicates Rous sarcoma virus.

| Species | Loop residues | Base |
|---|---|---|
| HIV-1 | $^{140}$GIPYNPQSQGVVE$^{152}$ | C |
| SIV | $^{140}$GIPYNPQSQGVVE$^{152}$ | C |
| FIV | $^{142}$GIP**G**NPQSQ**A**LVE$^{154}$ | C |
| ASV | $^{145}$GIP**G**N**S**Q**G**Q**A**MVE$^{157}$ | A |
| RSV | $^{145}$GIP**G**N**S**Q**G**Q**A**MVE$^{157}$ | A |

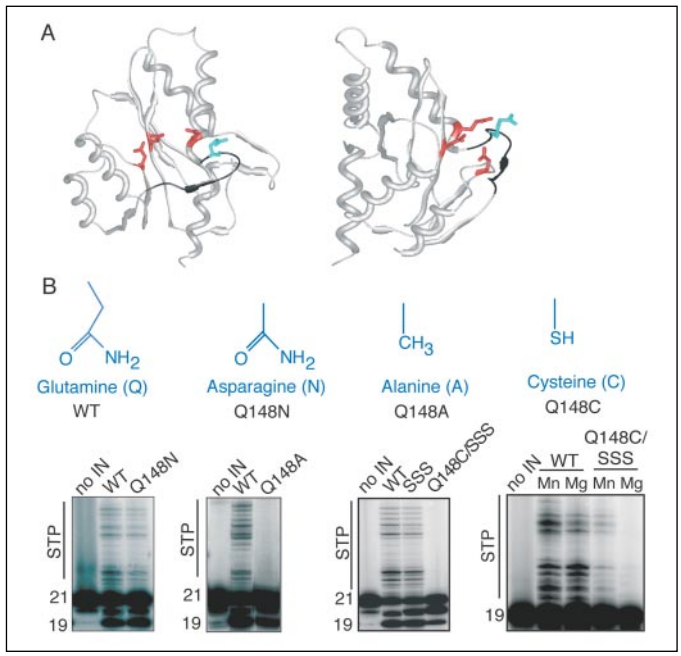


FIGURE 4. **Reduced ST activity of Gln-148 integrase mutants.** *A,* illustration of the location of Gln-148 within the integrase core domain taken from Ref. 21. Front and side views are shown on the *left* and *right*, respectively. The catalytic amino acids Asp-64, Asp-116, and Glu-152 are *red*; Gln-148 is *blue*; and the flexible loop is *black*. *B,* diagram of the amino acid side chains and representative gels showing reactions of wild-type (*WT*), Q148N, Q148A, a triple serine mutant SSS, and Q148C/SSS integrase (*IN*) with full-length DNA, and wild-type and Q148C/SSS integrase with precleaved DNA. *SSS* is the mutant C56S/C65S/C280S. *Q148C/SSS* contains the mutations C56S/C65S/C280S/Q148C.

*Glutamine 148 of Integrase Promotes ST*—We next looked for a potential integrase hydrogen bond donor residue. An alignment of the retroviral flexible loop residues through the adjacent catalytic glutamate residue and the DNA base corresponding to the 5′-C are shown in Table 1. Of the potential hydrogen bond donating residues, only Asn-144 and Gln-148 are conserved and required for viral replication (19, 20). We chose those residues as potential hydrogen bond donors, as well as Tyr-143 because of published results indicating this amino acid cross-linked the 5′-A of the LTR tip (13). Additionally, residues Thr-122 and Lys-127 were selected from modeling studies of integrase and viral DNA.[3] Mutating Thr-122 (to Ala or Cys) or Tyr-143 (to Ser, Phe, or Ile) had no effect on ST. Mutating Asn-144 (to Lys, Ala, or Cys) had a minimal effect on ST. Mutation of Lys-127 (to Ala or Cys) blocked 3′-P and ST (data not shown). Therefore, positions 122, 127, 144, and 143 were not investigated further.

Fig. 4*A* shows the close proximity of Gln-148 (*blue*) to the three acidic, catalytic amino acids (*red*) (from the crystal structure 1BI4

[3] R. Karki and M. Nicklaus, personal communication.

Downloaded from http://www.jbc.org/ by guest on October 8, 2019

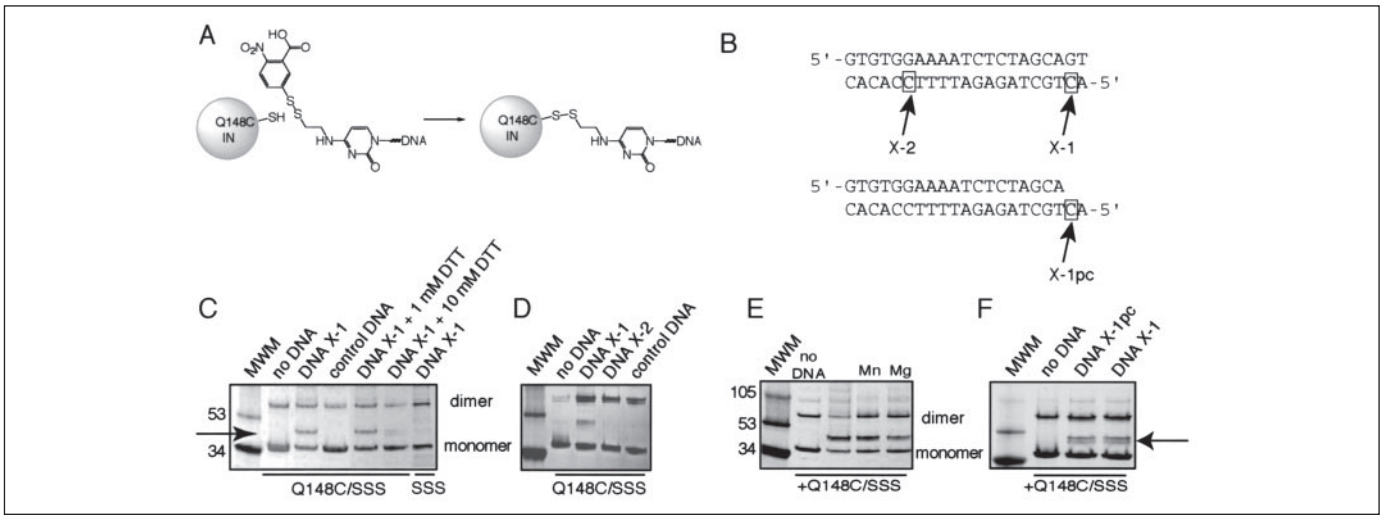


FIGURE 5. **Disulfide cross-linking between the Q148C integrase mutant and the 5′-C.** *A,* schematic diagram of the cross-linking reaction. *B,* duplex DNA sequence showing the location of cross-linking groups. *DNA X-1pc* is a precleaved substrate. *C–F,* Coomassie-stained SDS-polyacrylamide gels showing the disulfide cross-linking products. Integrase (*IN*) migrates through the gel as a mixture of monomer and dimer under nonreducing conditions. The integrase mutants are labeled *below* the gel, and the *arrow* indicates the position of the cross-linked integrase/DNA product. The size of molecular weight markers (*MWM*) in kDa is indicated on the *left side* of the gels. *C,* cross-linking requires thiol-modified DNA, Q148C mutation, and nonreducing conditions. *D,* Gln-148 cross-linking to DNA X-1 compared with DNA X-2. *E,* divalent metal is not required for cross-linking. *F,* cross-linking with precleaved DNA. *DTT*, dithiothreitol.

encompassing the flexible loop (21)). We created the mutants Q148N and Q148A to examine the efficiency of ST when the amino acid 148 side chain is shortened by one methylene group or changed to a methyl group (Fig. 4*B*). Q148N completed 3′-P at a level similar to wild-type integrase but had a 40% reduction in ST activity (Fig. 4*B* and data not shown). Q148A had a slight decrease in 3′-P and lacked ST. Other Gln-148 mutants (Glu and Lys) were also deficient for ST (data not shown). The 5′-C requirement was tested with the Q148N mutant, and the results were similar to wild-type integrase (see Fig. 1) for both 3′-P and ST, except for the globally lower ST efficiency (data not shown).

*Direct Cross-linking of Integrase Gln-148 to the 5′-C*—Because changes to either the 5′-C and Gln-148 diminished ST, and both the 5′-C and Gln-148 are plausible partners for hydrogen bonding, we tested their potential direct interaction using disulfide cross-linking. Disulfide cross-linking between proteins and DNA has been accomplished for several protein-DNA complexes, including HIV-1 reverse transcriptase and DNA repair proteins (for review see Ref. 22). Additionally, the integrase mutant E246C has been cross-linked previously with viral DNA substrate at the seventh base (adenine) from the 5′-end of the lower DNA strand by using a similar method (23).

Three surface cysteine residues of integrase were mutated to serine to eliminate background cross-linking. Therefore, in cross-linking experiments the control protein had the mutations C56S/C65S/C280S ("SSS"). Sequential mutations indicated Cys-56 was responsible for background cross-linking, in contrast to a previous study where Cys-65 and Cys-280 provided background cross-linking to the 5′-adenine of the lower DNA strand (23). The mutation Q148C was added to the triple serine mutant integrase to place a reactive sulfur at residue 148 ("Q148C/SSS").

The activity of these mutants is shown in Fig. 4*B*. The triple serine (SSS) retains wild-type activity for both 3′-P and ST. The Q148C/SSS mutant exhibits no ST and altered 3′-P specificity. 3′-P was reduced at the canonical cleavage site (CA ↓ GT-3′ generating the 19-mer) and enhanced at the adjacent site CAG ↓ T-3′, resulting in the formation of a 20-mer product and a similar level of total 3′-P. Altered 3′-P has also been reported for a Q148L mutant of HIV-2 integrase with enhanced generation of circular dinucleotide products (24). Thus, it is possible that Gln-148 helps position the 3′-P reactants, allowing dinucleotide release, and the role of Gln-148 in ST becomes important only once the 5′-C is unpaired. Activity of the Q148C/SSS mutant was tested with precleaved DNA in the presence of manganese. These conditions permitted a partial "rescue" (19%) of ST. Presumably cysteine can interact with manganese, and not magnesium, because of the ability of manganese to bind cysteine (25).

The cross-linking components and reaction are shown schematically in Fig. 5*A*. DNA was synthesized with an alkanethiol tether at the 5′-C position. The tether was activated by 5,5′-dithiobis(2-nitrobenzoic acid) to yield the substrate shown before reaction with integrase in Fig. 5*A*.

The DNA substrates used to examine position-specific cross-linking are shown in Fig. 5*B*. Cross-linking of the Q148C/SSS integrase to the thiol-modified 5′-C-containing DNA (DNA X-1) was observed as the band migrating between the monomer and dimer species of integrase (highlighted by *horizontal arrow* in Fig. 5*C*). Cross-linking efficiency was between 30 and 50% relative to the monomer concentration. The cross-link band was not observed when unmodified DNA was used (Fig. 5, *C* and *D*). As expected, dithiothreitol reduced the level of protein-DNA cross-link, and no cross-link was observed with the SSS integrase lacking Q148C (Fig. 5*C*). Thus, the core domain surface residue Cys-130 does not cross-link with the thiol-containing DNA. Cross-linking was not observed when the reactive thiol group was placed distal from the 5′-end of the duplex (DNA X-2), which rules out nonspecific cross-linking (Fig. 5*D*). Finally, cross-linking was observed in the presence of manganese and magnesium, precleaved or full-length DNA (Fig. 5, *E* and *F*) in agreement with the previous observation of DNA binding in the absence and presence of metal (26, 27). The recently reported flexibility of the viral DNA end may allow cross-linking of nonprocessed substrate (28). Together, these experiments indicate that the integrase residue 148 is in the vicinity of the cytosine of the 5′-overhang and that catalysis is not required for DNA cross-linking.

*Cross-linking Interference by Integrase Inhibitors*—Diketo acid derivatives have emerged recently as specific ST inhibitors with antiviral activity (29, 30) and are reviewed in Refs. 3 and 31. Interaction between 5CITEP and Gln-148 was observed crystallographically (25). Inhibition of cross-linking by the integrase inhibitor 5CITEP was examined at concentrations of integrase and DNA used in standard catalytic exper-

Downloaded from http://www.jbc.org/ by guest on October 8, 2019

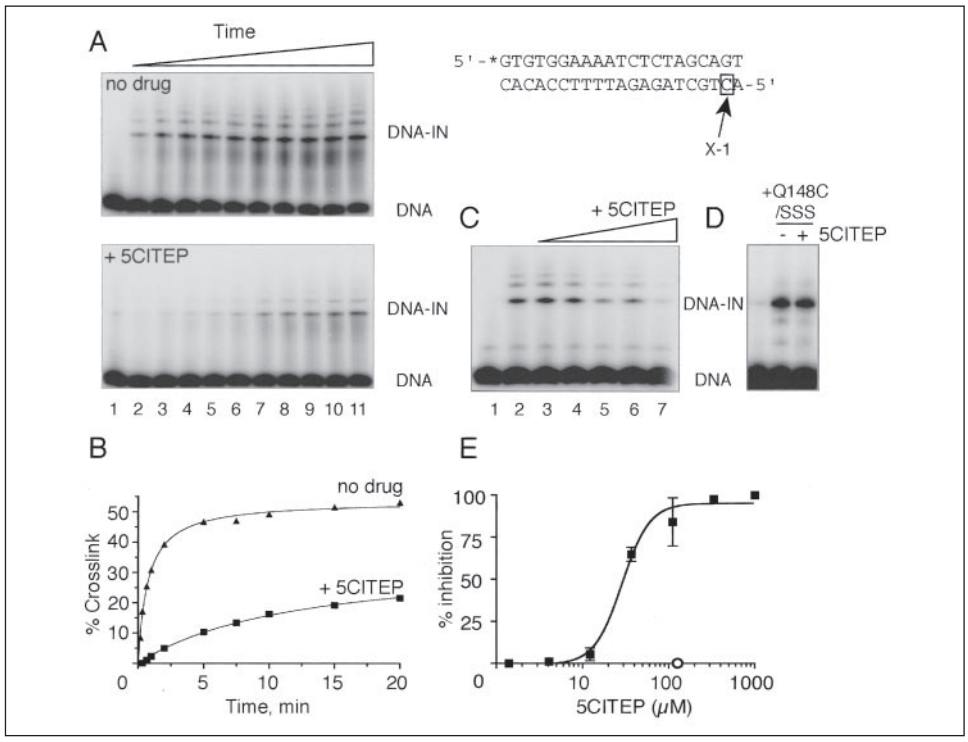


FIGURE 6. **Inhibition of integrase-DNA cross-linking by 5CITEP.** *A,* drug inhibition of integrase/DNA cross-linking was analyzed using 5′-$^{32}$P-labeled DNA. Cross-linking products were visualized by autoradiography following SDS-PAGE. *Left panels,* time course of cross-link formation in the absence and presence of 100 $\mu$M 5CITEP. Time points were 0.16, 0.33, 0.66, 1, 2, 5, 7.5, 10, 15, and 20 min (*lanes 2–11*). *Lane 1* has no integrase (*IN*). *B,* graph of time course of cross-link formation in the absence and presence of 100 $\mu$M 5CITEP. *C,* inhibition of disulfide cross-linking by 5CITEP. *Lane 1,* DNA control; *lanes 2–7* contain Q148C/SSS integrase plus 0–333 $\mu$M 5CITEP. *D,* 5CITEP (111 $\mu$M) failed to block DNA X-1 binding as measured by Schiff base assay. *E,* quantitation of *C.* Inhibition of disulfide cross-linking is shown as *closed squares.* Lack of inhibition of integrase-DNA binding at 111 $\mu$M 5CITEP is shown as the *open circle.*

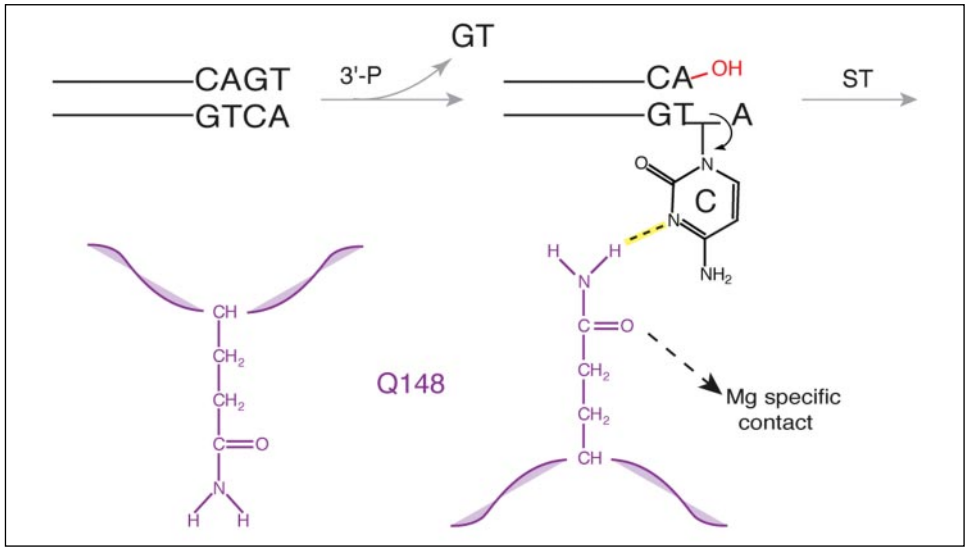


FIGURE 7. **Model of integrase Gln-148–5′-C interaction.** Integrase amino acid residue Gln-148, as part of the flexible loop, is diagrammed in *violet*. A proposed hydrogen bond between the Gln-148 amine and 5′-C N-3 groups is highlighted in *yellow*. Movement of the flexible loop and/or rotation of the 5′-C out of the DNA helix facilitate the interaction. Specificity in the presence of magnesium may be achieved through interaction with the glutamine carbonyl group.

iments in the presence of magnesium. The DNA was radioactively labeled to permit visualization of integrase-DNA complexes. A time course experiment indicated cross-linking was complete in less than 2 min (Fig. 6, *A* and *B*). Addition of 5CITEP decreased the rate of cross-link formation. A 1-min cross-linking time was chosen to examine the inhibition of cross-linking at different 5CITEP concentrations (Fig. 6*C*). The $IC_{50}$ of cross-link inhibition was 29 $\mu$M. At 111 $\mu$M, 5CITEP had no effect on overall binding of integrase to DNA, as measured by Schiff base assay (Fig. 6, *D* and *E*), suggesting inhibition of disulfide cross-linking reflects the 5CITEP-binding site.

## DISCUSSION

The present study demonstrates that efficient ST requires both the presence of a 5′-C at the end of the donor (viral) DNA and Gln-148 in the flexible loop of HIV-1 integrase. The 5′-C is conserved in the LTR sequences of HIV-1, HIV-2, simian immunodeficiency virus, and feline immunodeficiency virus (Table 1). In contrast, avian sarcoma virus and Rous sarcoma virus LTRs contain an adenine at that position, where the nitrogen at the first ring position could contribute to a similar interaction in an appropriately structured active site. Gln-148 is conserved among retroviral integrases, including HIV-1, simian immunodeficiency virus, avian sarcoma virus, Rous sarcoma virus, and feline immunodeficiency virus (32). Gln-148 resides in a flexible loop disordered in most integrase crystal structures or crystallized in a conformation that is unsuitable for catalysis (4–9). It has been suggested that the loop becomes ordered upon DNA binding and acts to stabilize the 5′-end of the viral DNA (5, 13).

Mutations that reduce the flexibility of the loop containing Gln-148 impair catalysis without affecting DNA binding (4). A conformational change following 3′-P has been observed (23). This flexibility might

Downloaded from http://www.jbc.org/ by guest on October 8, 2019

allow Gln-148 and the 5′-C to interact following the 3′-P reaction, allowing efficient ST. This structural change could involve movement of the flexible loop, rotation of the 5′-C base out of the DNA duplex, or both, as illustrated in Fig. 7. In our model, release of the 5′-GT dinucleotide provides the rotational freedom to the lower strand 5′-C, leaving the area near the catalytic 3′-hydroxyl available for subsequent ST. The 5′-AC overhang is then anchored away from the catalytic 3′-hydroxyl by a hydrogen bond between the Gln-148 side chain amine and the cytosine N-3 group. Note that integrase is active as a multimer, and our model does not confine the active site performing catalysis and the Gln-148 facilitating ST to the same subunit.

A precedent for a similar glutamine-cytosine N-3 hydrogen-bonding interaction exists for human deoxycytidine kinase (33). Deoxycytidine kinase efficiently phosphorylates deoxycytidine but also deoxyguanosine and deoxyadenosine by interaction with Gln-97. The bound base dictates rotation of the glutamine side chain into an appropriate hydrogen-bonding position. Discrimination of deoxycytidine by deoxycytidine kinase is achieved through H-bonding of the cytosine N-3 with the amino groups to the Gln-97 (33).

The use of magnesium in the integrase active site is an important component of our model. Divalent metals influence the effectiveness of integrase inhibitors (34, 35), integrase sequence specificity during DNA binding (13, 36), 3′-P (Fig. 1) and ST (37), 3′-P nucleophile selection (24, 38), and the sensitivity of drug-resistant mutant integrases to inhibitors (34). Integrase undergoes a divalent metal-induced protein conformational change (39), which may enhance specific recognition of the viral DNA end (26). The strict coordination requirements of magnesium (40) and the base specificity we observed in the presence of magnesium *in vitro* may point to the importance of the cytosine base requirement during ST *in vivo*.

This specificity may be reduced in our *in vitro* studies done in the presence of manganese because of altered metal coordination (40). Magnesium, a hard metal, prefers coordination to hard atoms such as oxygen (reviewed in Ref. 41). Coordination of magnesium may dictate that the Gln-148 carbonyl group supplies a direct- or water-mediated contact to magnesium (*Mg specific contact* in Fig. 7). This contact restricts rotation of the Gln-148 side chain and pulls electrons from the glutamine amino group, facilitating H-bond formation (Fig. 7). As a result, the side chain amine is directed to specifically contact and stabilize the 5′-C during ST. In contrast, the Gln-148 carbonyl group may not be needed in the same capacity in the presence of manganese, a soft metal. An altered metal coordination could allow the conformation of both of the glutamine side chain functional groups to accommodate any of the four DNA bases; hence, a loss of stringency for the base that stabilizes the viral DNA during ST in the presence of manganese (may be biologically irrelevant).

Because of the flexibility of the loop containing Gln-148 and the single-stranded nature of the 5′-C following 3′-P, other loop residues may contact the 5′-C. The loop contains residues $^{140}$GIPYNPQSQG$^{149}$, with the underlined residues capable of acting as H-bond donors. The functional relevance of Tyr-143 and Gln-146 can be excluded because mutants have no effect on viral replication (19), and Ser-147 is not conserved among retroviruses. We tested N144C in the disulfide cross-linking assay. Only a small amount of cross-linking was observed, and furthermore, aggregation of the protein, presumably from protein self-cross-linking precluded accurate assessment of integrase-DNA cross-link specificity (data not shown).

Gln-148 may form part of the binding pocket for diketo acid integrase inhibitors, which are selective against the integrase ST reaction (for review see Ref. 3). A co-crystal between integrase and 5CITEP revealed a hydrogen bond between Gln-148 and the nitrogen of the indole ring of 5CITEP (42), and docking studies suggest the participation of Gln-148 in the drug-binding pocket for several other integrase inhibitors (35, 43, 44). A drug-resistant mutation Q148K developed after exposure to S-1360, a diketo acid integrase inhibitor in clinical trial, and this mutant exhibited poor viral replication lending support to the important role of Gln-148 (20). Diketo acids have been proposed to bind the integrase-DNA complex and block ST by interacting with the 5′-AC overhang (45). We observed that cross-link between the LTR 5′-C and residue 148 of integrase was inhibited by 5CITEP at an $IC_{50}$ of 29 $\mu$M (Fig. 6, *C* and *D*). 5CITEP inhibits 3′-P and ST at $IC_{50}$ values of 35 and 0.65 $\mu$M, respectively (45). Inhibition of Q148C/5′-C cross-linking by 5CITEP suggests the binding site for this inhibitor overlaps the region of the Gln-148/5′-C interaction. This is supported by the lack of inhibition of Schiff base formation between integrase and an abasic site substitution for the conserved adenine at the 3′-P site (Fig. 6, *D* and *E*). The cross-linking approach could be used to scan the molecular contacts between integrase and its DNA substrates, as well as inhibitor-binding sites. Stabilization of integrase-DNA complexes may also enable co-crystal structure determination.

## REFERENCES


1. Engelman, A., Mizuuchi, K., and Craigie, R. (1991) *Cell* **67,** 1211–1221
2. Katz, R. A., and Skalka, A. M. (1994) *Annu. Rev. Biochem.* **63,** 133–173
3. Pommier, Y., Johnson, A. A., and Marchand, C. (2005) *Nat. Rev. Drug Discov.* **4,** 236–248
4. Greenwald, J., Le, V., Butler, S. L., Bushman, F. D., and Choe, S. (1999) *Biochemistry* **38,** 8892–8898
5. Dyda, F., Hickman, A. B., Jenkins, T. M., Engelman, A., Craigie, R., and Davies, D. R. (1994) *Science* **266,** 1981–1986
6. Goldgur, Y., Dyda, F., Hickman, A. B., Jenkins, T. M., Craigie, R., and Davies, D. R. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 9150–9154
7. Wang, J. Y., Ling, H., Yang, W., and Craigie, R. (2001) *EMBO J.* **20,** 7333–7343
8. Bujacz, G., Alexandratos, J., Qing, Z. L., Clement-Mella, C., and Wlodawer, A. (1996) *FEBS Lett.* **398,** 175–178
9. Chen, J. C., Krucinski, J., Miercke, L. J., Finer-Moore, J. S., Tang, A. H., Leavitt, A. D., and Stroud, R. M. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 8233–8238
10. Chiu, T. K., and Davies, D. R. (2004) *Curr. Top. Med. Chem.* **4,** 965–977
11. Bushman, F. D., and Craigie, R. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 3458–3462
12. Ellison, V., and Brown, P. O. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 7316–7320
13. Esposito, D., and Craigie, R. (1998) *EMBO J.* **17,** 5832–5843
14. Gerton, J. L., Ohgi, S., Olsen, M., DeRisi, J., and Brown, P. O. (1998) *J. Virol.* **72,** 5046–5055
15. Ferentz, A. E., and Verdine, G. L. (1991) *J. Am. Chem. Soc.* **113,** 4000–4002
16. Wolfe, S. A., Ferentz, A. E., Grantcharova, V., Churchill, M. E., and Verdine, G. L. (1995) *Chem. Biol.* **2,** 213–221
17. Leh, H., Brodin, P., Bischerour, J., Deprez, E., Tauc, P., Brochon, J. C., LeCam, E., Coulaud, D., Auclair, C., and Mouscadet, J. F. (2000) *Biochemistry* **39,** 9285–9294
18. Mazumder, A., Neamati, N., Pilon, A., Sunder, S., and Pommier, Y. (1996) *J. Biol. Chem.* **271,** 27330–27338
19. Lu, R., Limon, A., Ghory, H. Z., and Engelman, A. (2005) *J. Virol.* **79,** 2493–2505
20. Neamati, N. (2002) *Expert Opin. Ther. Patents* **12,** 709–724
21. Maignan, S., Guilloteau, J. P., Zhou-Liu, Q., Clement-Mella, C., and Mikol, V. (1998) *J. Mol. Biol.* **282,** 359–368
22. Verdine, G. L., and Norman, D. P. (2003) *Annu. Rev. Biochem.* **72,** 337–366
23. Gao, K., Butler, S. L., and Bushman, F. (2001) *EMBO J.* **20,** 3565–3576
24. van Gent, D. C., Groeneger, A. M. O., and Plasterk, R. H. A. (1993) *Nucleic Acids Res.* **21,** 3373–3377
25. Gao, K., Wong, S., and Bushman, F. (2004) *J. Virol.* **78,** 6715–6722
26. Yi, J., Asante-Appiah, E., and Skalka, A. M. (1999) *Biochemistry* **38,** 8458–8468
27. Neamati, N., Hong, H., Sunder, S., Milne, G. W., and Pommier, Y. (1997) *Mol. Pharmacol.* **52,** 1041–1055
28. Renisio, J. G., Cosquer, S., Cherrak, I., El Antri, S., Mauffret, O., and Fermandjian, S. (2005) *Nucleic Acids Res.* **33,** 1970–1981
29. Hazuda, D. J., Young, S. D., Guare, J. P., Anthony, N. J., Gomez, R. P., Wai, J. S., Vacca, J. P., Handt, L., Motzel, S. L., Klein, H. J., Dornadula, G., Danovich, R. M., Witmer, M. V., Wilson, K. A., Tussey, L., Schleif, W. A., Gabryelski, L. S., Jin, L., Miller, M. D., Casimiro, D. R., Emini, E. A., and Shiver, J. W. (2004) *Science* **305,** 528–532
30. Hazuda, D. J., Felock, P., Witmer, M., Wolfe, A., Stillmock, K., Grobler, J. A., Espesath, A., Gabryelski, L., Schlelf, W., Blau, C., and Miller, M. D. (2000) *Science* **287,** 646–650


Downloaded from http://www.jbc.org/ by guest on October 8, 2019

31. Johnson, A. A., Marchand, C., and Pommier, Y. (2004) *Curr. Top. Med. Chem.* **4,** 1059–1077
32. Engelman, A., and Craigie, R. (1992) *J. Virol.* **66,** 6361–6369
33. Sabini, E., Ort, S., Monnerjahn, C., Konrad, M., and Lavie, A. (2003) *Nat. Struct. Biol.* **10,** 513–519
34. 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., Young, S., Vacca, J., and Hazuda, D. J. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 6661–6666
35. Marchand, C., Johnson, A. A., Karki, R. G., Pais, G. C. G., Zhang, X., Cowansage, K., Patel, T. A., Nicklaus, M. C., Burke, T. R., Jr., and Pommier, Y. (2003) *Mol. Pharmacol.* **64,** 600–609
36. Katzman, M., Katz, R. A., Skalka, A. M., and Leis, J. (1989) *J. Virol.* **63,** 5319–5327
37. Johnson, A. A., Sayer, J. M., Yagi, H., Kalena, G. P., Amin, R., Jerina, D. M., and Pommier, Y. (2004) *J. Biol. Chem.* **279,** 7947–7955
38. Katzman, M., and Sudol, M. (1996) *J. Virol.* **70,** 2598–2604
39. Asante-Appiah, E., and Skalka, A. M. (1997) *J. Biol. Chem.* **272,** 16196–16205
40. Martin, R. B. (1990) in *Metal Ions in Biological Systems* (Sigel, H., Sigel, A., ed) pp. 1–13, Marcel Dekker, Inc., New York
41. Verma, S., and Eckstein, F. (1998) *Annu. Rev. Biochem.* **67,** 99–134
42. Goldgur, Y., Craigie, R., Cohen, G. H., Fujiwara, T., Yoshinaga, T., Fujishita, T., Sugimoto, H., Endo, T., Murai, H., and Davies, D. R. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 13040–13043
43. Pais, G. C. G., Zhang, X., Marchand, C., Neamati, N., Cowansage, K., Svarovskaia, E. S., Pathak, V. K., Tang, Y., Nicklaus, M. C., Pommier, Y., and Burke, T. R., Jr. (2002) *J. Med. Chem.* **45,** 3184–3194
44. Sotriffer, C. A., Ni, H., and McCammon, J. A. (2000) *J. Med. Chem.* **43,** 4109–4117
45. Marchand, C., Zhang, X., Pais, G. C., Cowansage, K., Neamati, N., Burke, T. R., Jr., and Pommier, Y. (2002) *J. Biol. Chem.* **277,** 12596–12603


Downloaded from http://www.jbc.org/ by guest on October 8, 2019

# Integration Requires a Specific Interaction of the Donor DNA Terminal 5′ -Cytosine with Glutamine 148 of the HIV-1 Integrase Flexible Loop

Allison A. Johnson, Webster Santos, Godwin C. G. Pais, Christophe Marchand, Ronak Amin, Terrence R. Burke, Jr., Gregory Verdine and Yves Pommier

*J. Biol. Chem. 2006, 281:461-467.*
*doi: 10.1074/jbc.M511348200 originally published online October 27, 2005*

---

Access the most updated version of this article at doi: 10.1074/jbc.M511348200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 44 references, 24 of which can be accessed free at
http://www.jbc.org/content/281/1/461.full.html#ref-list-1

Downloaded from http://www.jbc.org/ by guest on October 8, 2019