

## Modelación molecular y variación estructural de las integrasas de dos retrovirus humanos: HTLV-I y VIH-1

Molecular modeling and structural variation of two human retrovirus integrases: HTLV-I and HIV-1

Felipe García Vallejo<sup>1</sup>, Yesid Cuesta Astroz<sup>2</sup>, Martha C. Domínguez<sup>3</sup>, Adalberto Sánchez<sup>4</sup>, Mercedes Salcedo Cifuentes<sup>5</sup>, Diane María Díaz<sup>6</sup>, Milton Quintana<sup>7</sup>

### Resumen

**Objetivo:** Analizar las características moleculares y de variación de secuencias de las integrasas del HTLV-I y del VIH-1 y sus variantes poblacionales.

**Metodología:** Análisis de secuencias y estructuras obtenidas de diferentes bases de datos; para ello se utilizaron programas computacionales de modelación de estructuras proteicas e identificación de sustituciones polimórficas en secuencias de aminoácidos de integrasas del HTLV-I y VIH-1 previamente reportadas.

**Materiales y métodos:** Tanto la integrasa del HTLV-I como la del VIH-1 son proteínas compuestas por 288 residuos de aminoácidos. Se encontró un parecido de estructuras terciarias entre los dominios catalíticos de las IN de VIH-1, ASV y RSV con la del HTLV-I. A partir de 103 secuencias completas de la integrasa del VIH-1 se registraron, en 46 codones, un total de 53 sustituciones que se localizaron en diferentes posiciones de la proteína nativa; las más frecuentes fueron: N27G (32,1%), A265V (30,1%), L101I (31,1%) y T123A (27,0%). Ninguna de las sustituciones más frecuentemente encontradas generó un cambio en el plegamiento nativo de la correspondiente región.

**Conclusión:** La estructura tridimensional del dominio central catalítico de la integrasa condicionaría su actividad y su relación con moléculas potencialmente inhibitoras. Las sustituciones observadas fueron neutrales sin alterar la estructura nativa. Los resultados obtenidos confirman que la integrasa es un nuevo y promisorio blanco para el desarrollo de terapias antirretrovirales más efectivas en el siglo XXI.

**Palabras clave:** Retrovirus, integrasa, modelación molecular, estructura proteica, polimorfismos proteicos, resistencia antiviral.

<sup>1</sup> Ph. D. Director Científico. [labiomol@gmail.com](mailto:labiomol@gmail.com)

<sup>2</sup> Biólogo. Investigador. [yesid.cuesta@gmail.com](mailto:yesid.cuesta@gmail.com)

<sup>3</sup> Magister en Microbiología. Investigadora.

<sup>4</sup> Ph. D. Profesor auxiliar e investigador. [asanchez6911@yahoo.com](mailto:asanchez6911@yahoo.com)

<sup>5</sup> Magister en Epidemiología. [mercysal2003@yahoo.com](mailto:mercysal2003@yahoo.com)

<sup>6</sup> Bióloga. Investigadora del Laboratorio de Biología Molecular y Patogénesis del Departamento de Ciencias Fisiológicas. Universidad del Valle. Estudiante de Maestría en Ciencias Biomédicas. [dianemave2001@yahoo.com](mailto:dianemave2001@yahoo.com)

<sup>7</sup> M. S. Profesor. Facultad de Medicina. Universidad del Sinú. [Monteria.quintanaso@yahoo.com](mailto:Monteria.quintanaso@yahoo.com)

<sup>1-6</sup> Miembros del Laboratorio de Biología Molecular y Patogénesis. Departamento de Ciencias Fisiológicas. Universidad del Valle

**Correspondencia:** Edificio 116 Luis Maria Borrero, oficina 508, Universidad del Valle, sede San Fernando, calle 4B N.º 36-00, A.A. 25360, Cali (Colombia), teléfono: (572)-5185601, fax:(572)-5185657.

Fecha de recepción: 4 de febrero de 2009  
Fecha de aceptación: 20 de marzo de 2009

  
Vol. 25, n.º 1, 2009  
ISSN 0120-5552

## Abstract

**Objective:** *To analyze the molecular characteristics and amino acid sequence variations of HTLV-I and of HIV-1 integrases and their population variants.*

**Materials and methods:** *Data mining and analysis of integrase sequences and protein structure data bases by using appropriate software for modelling and search for polymorphic substitutions in HTLV-I and HIV-1 integrase amino acid sequences previously reported.*

**Results:** *HTLV-I and HIV-1 integrases are proteins of 288 amino acid residues. Structural modeling of tertiary folding of HTLV-I integrase catalytic central domain's, showed closed structural characteristic with those of HIV-1, ASV and RSV. From 103 full amino acid sequences of HIV-1 integrase, 53 substitutions located in 46 different codons were recorded. The more frequent correspond to N27G (32,1%), L101I (31,1%), A265V (30,1%) and T123A (27,0%). None of these frequent substitutions introduced changes in the folding of HIV-1 native integrase.*

**Conclusion:** *The tridimensional structure of central catalytic domain would influence the integrase activity and its relationship with potentially inhibitory molecules. Those observed amino acid substitutions were neutral and do not alter the native protein structure. Our data confirm those previously published, and enable us to propose that IN is a new and promissory target for develop more effective antiviral therapies in the XXI century.*

**Keywords:** *Retrovirus, Integrase, molecular modeling, protein structure, protein polymorphisms, antiviral resistance.*

## INTRODUCCIÓN

Las infecciones por el virus linfotrópico humano tipo I (HTLV-I) y el virus de la inmunodeficiencia humana 1 (VIH-1) se han convertido en problemas serios de salud pública a nivel mundial. La infección por el HTLV-I registra una cifra mayor a 20 millones de personas portadoras; de estas una proporción variable desarrolla bien sea una oncopatía, la leucemia de células T en el adulto (LCTA), o bien una mielopatía asociada a la infección por este virus, denominada paraparesia espástica tropical (PET/MAH); epidemiológicamente la infección por HTLV-I se asocia con una distribución en zonas endémicas alrededor del mundo (1).

De acuerdo con el reporte anual en 2007 del Programa de las Naciones Unidas para el SIDA/VIH, se estima que entonces había, aproximadamente, 33 millones de personas infectadas por el VIH-1, causante del síndrome de inmunodeficiencia adquirida (sida). Durante este mismo año, se reportaron 2,7 millones de nuevos casos en comparación con el 2006 y ocasionó 2 millones de muertes (2).

Los tratamientos empleados actualmente para frenar la infección por el VIH-1 y controlar el sida, han mostrado un éxito limitado. Recientemente, se pensó que la terapia de alta eficiencia (*Highly Active Antiretroviral Therapy*-HAART-), que consiste en combinaciones de dos inhibidores de la transcriptasa reversa (TR) y/o uno de la proteasa (Pr) (dos enzimas codificadas por el genoma del virus) y que son claves para su replicación, que sería una terapia efectiva; sin embargo, posteriormente se demostró que presentaba fallas. Al respecto se han reportado más de 200 mutaciones, que están asociadas con la resistencia al tratamiento

por seis diferentes clases de drogas antirretrovirales, actualmente en uso (3). Si bien el problema en el VIH-1 es preocupante, en el caso del tratamiento de la infección HTLV-I es aun más grave, puesto que los esquemas terapéuticos no han dado los resultados esperados en el control de la infección (4).

Dentro de los nuevos enfoques farmacológicos para controlar la infección por el VIH-1 se consideran dos alternativas novedosas: el uso de moléculas inhibidoras de la fusión del virus con la célula y la inhibición de la integración del ADN viral dentro del genoma celular (5). En este último enfoque, y desde hace varios años, se han desarrollado estudios que han permitido definir como una potencial terapia antirretroviral a moléculas inhibidoras de la actividad de la integrasa (IN), una enzima viral que participa, de manera primordial, en la integración del ADN vírico dentro del genoma de la célula infectada (6).

Una de las ventajas de la estrategia de inhibición de la integración del cADN viral se basa en impedir, de manera definitiva, la infección del virus en la célula. Esta es radicalmente diferente de aquellas que involucran TR y Pr, puesto que la integración del cADN viral es necesaria para que se expresen estas dos enzimas y se inhiban por la acción de fármacos. Así pues, los diseños de terapias que inhiban la integración del cADN viral se perfilan como los más promisorios, pues evitarían el continuo cambio de esquemas terapéuticos que deben hacerse por la selección de cuasiespecies resistentes a drogas. Además, como la IN no tiene análogo celular, sus efectos son mínimos en la expresión y regulación de genes celulares (7).

Durante más de 10 años de investigación sobre inhibidores de la integrasa retroviral se han desarrollado y ensayado una amplia gama de compuestos químicamente diferentes entre los que se incluyen los cetoácidos y derivados, las Dihydroxypyrimidina-4-carboxamidas, N-metilpirimidonas y las pirimidonas bicíclicas, las pirrolquinolonas, entre otras; todos estos compuestos se encuentran en fases de experimentación. Otras moléculas, como la GS-9137 (Elvitegravir), el GSK-364735, el BMS-707035 y el MK-2048, son drogas que actualmente están en fases preclínicas y ensayos clínicos, y prometen ser una segunda generación de drogas inhibitoras de la integrasa del VIH. Sin embargo, en la actualidad solo el Raltegravir, (MK-0518), cuyo nombre comercial es Isentress™, es el único fármaco inhibidor de integrasa aprobado por la FDA como droga antiviral para uso combinado (8). Para obtener información acerca de la estructura y variación de secuencias de las IN del VIH-1 y del HTLV-I, se presentarán los principales resultados obtenidos de un análisis *in silico* de las secuencias y estructuras de las integrasas que se encuentran en las principales bases de datos; también se describirán los polimorfismos naturales encontrados en 103 secuencias completas de la IN del VIH-1. Estos resultados refuerzan otros que consideran a la integrasa como un blanco muy promisorio en el tratamiento antirretroviral del siglo XXI (9, 10).

## MATERIALES Y MÉTODOS ESTRUCTURAS DE LAS INTEGRASAS

Para la obtención de estructuras tridimensionales de la IN del VIH-1 se realizó una búsqueda en el Protein Data Bank (<http://www.rcsb.org/pdb/>). Para la modelación

y obtención de las interacciones de las IN se escogieron tres estructuras cristalizadas y analizadas previamente mediante difracción de rayos X cuyos códigos de acceso en el PDB son: 1BIS, que corresponde al dominio central catalítico de la integrasa del VIH-1 obtenido a una resolución de 1,95 Å (11); 1C0M, la estructura cristalina de dos dominios de la integrasa del virus del sarcoma de Rous (RSV) a una resolución de 2,54 Å (12) y 1BI4, la estructura cristalina del dominio catalítico de la integrasa del VIH-1 obtenida a una resolución de 2,5 Å (13). Las estructuras anteriormente descritas se visualizaron utilizando el programa PDB Protein Workshop 1.5 (<http://www.rcsb.org/pdb/>), y aplicando los comandos correspondientes para ajustar las distintas variables de estructura.

Para predecir la estructura terciaria del dominio central de las integrasas, se buscaron en el PDB estructuras conocidas que compartieran similitud con la del HTLV-I. La integrasa del virus del sarcoma aviar (ASV), cuyo código PDB es 1ASW, mostró alta homología en secuencias y en estructura con la del HTLV-1 (14). Esta se empleó para realizar algunas predicciones en la unión de la secuencia al ADN sustrato. Todos estos procesos se realizaron utilizando el programa DeepView 4.0 (15), de acceso público en <http://spdbv.niehs.nih.gov/>.

## OBTENCIÓN DE ESTRUCTURAS MOLECULARES MEDIANTE MODELACIÓN POR HOMOLOGÍA

Los mejores alineamientos obtenidos se enviaron al Swiss-Model disponible para su modelación (<http://www.expasy.org/swissmod/SWISS-MODEL.html>) (15) (16).

En la primera parte del proceso, el programa ProModII (17) generó un modelo básico de las proteínas del VIH-1, HTLV-1, ASV y RSV. Estos se optimizaron por minimización de energía, empleando el programa Gromos 96 (18); mediante este procedimiento se obtuvieron estructuras en gradiente conjugado para 300 ciclos. Con la ayuda del programa PyMol se modeló la interacción entre la integrasa del VIH-1 y el ADN blanco de integración. Puesto que no se encontró alta homología con los dominios carboxilo terminal y amino terminal empleando el procedimiento anterior, las secuencias correspondientes se sometieron a un análisis de estructura secundaria, utilizando el PHDsec Method (<http://www.predictprotein.org/>) para obtener la más probable estructura secundaria correspondiente (19) (20).

#### ANÁLISIS DE SECUENCIAS DE AMINOÁCIDOS DE LA IN DEL VIH-1

Del GenBank (<http://www.ncbi.nih.gov/>), mediante la aplicación del programa BlastP, se seleccionaron y se alinearon 103 secuencias de aminoácidos completas de la integrasa del VIH-1 provenientes de individuos no tratados con ningún inhibidor de la actividad de la integrasa. Estas se realinearon con el programa ClustalW 2.0 (21), utilizando como secuencia consenso aquella cuyo código GenBank es NP705928; a partir de estos alineamientos se determinó la presencia de sustituciones de aminoácidos y sus frecuencias correspondientes. Los sitios polimórficos se caracterizaron por tener dos sustituciones o más con relación a la secuencia consenso.

Las secuencias de las integrasas de VIH-1 incluidas en este estudio corresponden a los códigos GenBank que se referencian a continuación:

AAO61875,	AAC83546,	
ABR68155,	ABC88346,	AAC83549,
AAC83550,	AAC83549,	AAC83477,
AAC83479,	ABR68191,	AAL01914,
AAL01905,	AAL01911,	AAG35639,
ABR68177,	AAL01910,	AAL01972,
ABR68197,	AAC83547,	AAN05469,
AAL01954,	ABR68162,	AAL01973,
AAL01944,	AAL01949,	AAL01906,
AAL01906,	AAL01946,	AAL01924,
AAL01913,	AAL01935,	AAL01948,
AAL01908,	AAL01930,	AAL01912,
AAL01923,	AAL01938,	AAL01945,
AAP03070,	AAL01932,	AAL01917,
AAG35642,	AAL01970,	AAL01915,
AAL01947,	AAL01943,	AAL01952,
AAL01951,	AAL01971,	AAC83491,
AAC83490,	AAC83495,	AAC83476,
AAC83465,	AAC83465,	AAC83465,
AAC83465,	AAC83466,	AAC83496,
AAC83551,	AAC83488,	AAC83484,
AAC83511,	AAC83514,	AAC83508,
ACG71187,	ACG71175,	ACG71122,
ACG71121,	ACG71124,	ACG71162,
ACG71220,	ACG71176,	ACG71224,
ACG71225,	ACG71208,	ACG71236,
ACG71237,	ACG71129,	ACG71211,
ACG71128,	ACG71210,	ACG71119,
ACG71147,	ACG71223,	ACG71170,
ACG71115,	ACG71172,	ACG71156,
ACG71198,	ACG71146,	ACG71213,
ACG71227,	ACG71231,	ACG71132,
ACG71171,	ACG71131,	ACG71169,
ACG71154,	ACG71178,	ACG71133,
ACG71186,	ACG71139.	

Se usó la herramienta de genotipado viral (*Viral genotyping tool*) del Centro Nacional para Biotecnología de la Información de los Estados Unidos (*National Center for Biotech-*

nology Information-NCBI-) (<http://www.ncbi.nlm.nih.gov/projects/genotyping/formpage.cg>), para determinar el correspondiente subtipo viral.

## RESULTADOS Y DISCUSIÓN

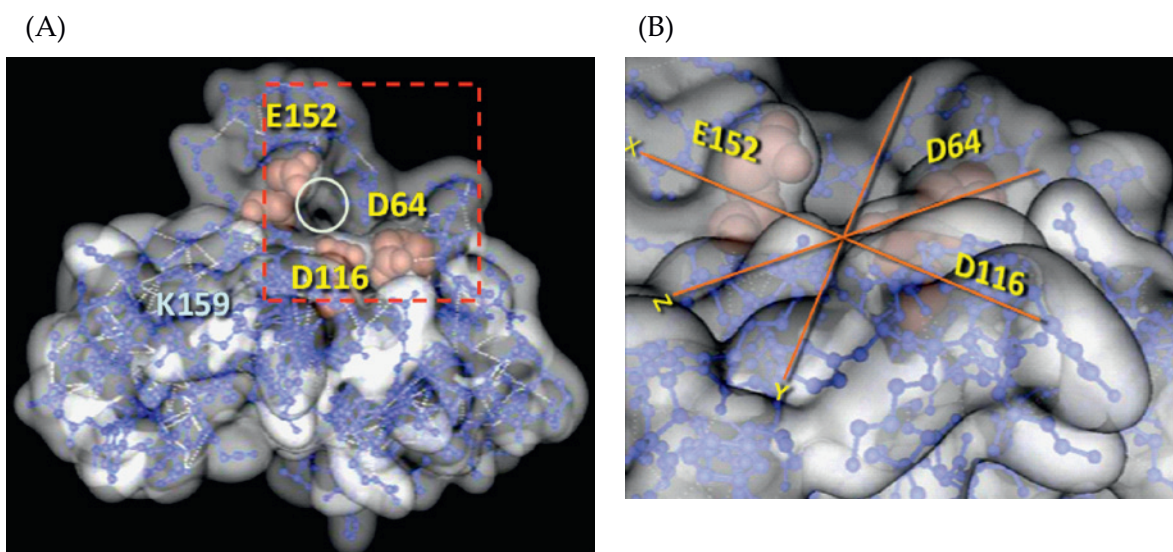
### LA ESTRUCTURA DE LAS INTEGRASAS RETROVIRALES

Los análisis de búsqueda de secuencias proteicas revelaron que las integrasas del HTLV-I y del VIH-1 son proteínas con 288 residuos de aminoácidos cuyo  $PM_r$  calculado es de 32 kDa, que son codificadas por el extremo 3' del gen Pol. Como resultado de la acción de la proteasa sobre la poliproteína precursora Pol, se producen las proteínas p51<sup>TR</sup>, p15 y p32<sup>IN</sup> (22). Inicialmente la IN es un monómero que se pliega para generar tres dominios diferentes cuyas funciones han sido identificadas mediante análisis mutacionales y estructurales. Su plegamiento terciario es muy similar a un gran número de proteínas de unión al ADN, que contienen el motivo hélice-vuelta-hélice (HTH).

La disposición espacial de los tres aminoácidos catalíticos D64, D116 y G152 localizados en el dominio central catalítica (51 a 212 aa), obtenida mediante simulación computacional para la IN del HTLV-I, le confiere al sitio activo una tridimensionalidad que es compatible con la actividad de nucleotidil transferasa (figura 1A). Estudios previos muestran que las integrasas de los retrovirus junto con las transposasas, la RNAsa H y las RNAsa argonautas asociadas a Dicer (involucrada en el paso de silenciamiento de genes) son enzimas que pertenecen al grupo de las polinucleotidil transferasas/esterasas (EC 2.7.7.49). De otra parte, se ha documentado que mutaciones sitiodirigi-

das sobre los residuos conservados del centro catalítico suprimen completamente la capacidad de la enzima para llevar a cabo la integración (23, 24).

La predicción del plegamiento de la secuencia de los 288 residuos de aminoácidos de IN del HTLV-I efectuada en este trabajo mostró que en el extremo N-terminal existen tres hélices  $\alpha$  ( $\alpha_1$ , 7-13 aa.;  $\alpha_2$ , 16-23 aa.;  $\alpha_3$ , 29-42 aa.). Esta predicción concuerda con la estructura encontrada previamente por cristalografía de rayos X para la IN del VIH-1 en la cual a este extremo se le ha llamado HHCC por los residuos conservados, pues su estructura es semejante a los dominios de dedos de zinc (25). Un estudio reciente demostró que mutantes deficientes en el primer residuo del extremo amino tienen actividades reducidas en el proceso de integración (26). En el dominio central catalítico se determinaron 5 hojas  $\beta$  ( $\beta_1$ , 64-69 aa.;  $\beta_2$ , 76-86 aa.;  $\beta_3$ , 91-97 aa.;  $\beta_4$ , 120-123 aa y  $\beta_5$ , 209-211 aa) y 4 hélices  $\alpha$  ( $\alpha_4$ , 102-114 aa.;  $\alpha_5$ , 131-147 aa.;  $\alpha_6$  158-176 aa.;  $\alpha_7$ , 183-198 aa.). El dominio del centro catalítico (CC), que se extiende del aminoácido 51 al 212, incluyó la tríada catalítica formada por los residuos D64, D116 y E152 (DD-35-E), conocida como el motivo DDE. Aunque no fue posible encontrar una estructura tridimensional altamente significativa para dominio carboxilo, se asignaron 2 hélices  $\alpha$  ( $\alpha_8$ , 248-254 aa.;  $\alpha_9$ , 273-279 aa.) y 3 hojas  $\beta$  ( $\beta_6$ , 231-236 aa.;  $\beta_7$ , 258-262 aa.;  $\beta_8$ , 269-271 aa.). Se determinó que existe una diferencia con la estructura tipo SH3 encontrada en VIH-1 y ASV en las que solamente hay hojas  $\beta$ . Se ha observado que en el extremo carboxilo de la IN existe una baja conservación de aminoácidos; esta podría explicar el porqué de la discrepancia según la predicción hecha por el método comparativo empleado en este estudio.



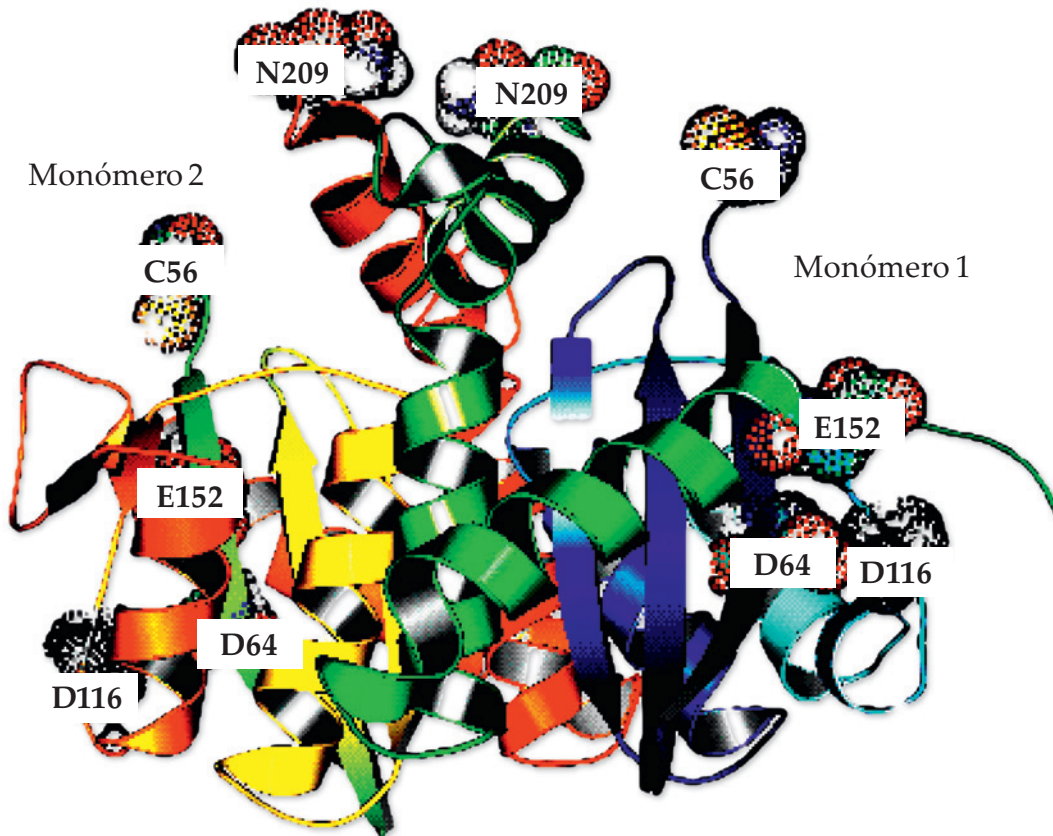
**Figura 1.** Resolución de la estructura tridimensional del dominio central catalítico de la IN del HTLV-I mediante modelación molecular. (A). Estructura tridimensional del dominio central catalítico en la que se muestran los residuos de D64, D116 y E152 y el bolsillo de interacción con el ADN sustrato; además, se detalla la localización del residuo K159 como estabilizador de la estructura tridimensional de sitio activo. (B). Vista detallada de la disposición tridimensional de los residuos D64, D116 y E152 dentro del bolsillo molecular en donde ocurre la catálisis. Las estructuras se obtuvieron mediante modelado por homología, tomando como molde el dominio catalítico (código PDB: 1BI4) de la estructura de la IN de HIV. A partir del alineamiento de las secuencias de HIV PDB: 1BI4 con la de aminoácidos de la región catalítica de la integrasa de HTLV-I (ATK-I) (NA J02029), se obtuvo el modelo del dominio central de la integrasa de HTLV-I; posteriormente, esta se minimizó utilizando el programa Spdb Viewer; por último, empleando el programa Gromos 96, se obtuvo la localización de los residuos catalíticos DDE y se calcularon las distancias, en Å, entre ellos. El círculo en blanco en la figura A muestra el dominio de unión al ADN sustrato.

**Fuente:** Programas empleados: programa Spdb Viewer y Gromos 96.

Un análisis más detallado del dominio central catalítico de la IN del HTLV-I mostró que los aminoácidos D64, D116, E152 y K159 se localizan cerca de una región de 13 aminoácidos, que está rodeada por los residuos 140 a154. En la enzima nativa, la distancia teórica calculada entre los tres residuos catalíticos fue 4,04Å para D64 -D116, 7,40Å en D64-E152 y 10,20Å entre E152 y D116. En el modelo construido, los residuos E152 y D116 se encuentran orientados hacia el interior del bolsillo del sitio activo y en el mismo plano, mientras que D64 está localizado en el plano interno de la cavidad catalítica (figura 1B).

Estudios previos de modelación de estructuras (27, 28) sugieren que la integrasa fun-

cional del VIH-1 es una proteína multimérica; aunque el número exacto de subunidades en la enzima funcional *in vivo* no se ha establecido. Los resultados previos del análisis de estructuras cristalinas obtenidas por difracción de rayos X, que se complementaron con simulaciones *in silico* y métodos de dinámica molecular realizados en este trabajo, sustentan la propuesta de un modelo dimérico (figura 2). De acuerdo con este modelo, los contactos entre subunidades están relacionados entre sí por un eje de la díada que contiene una superficie muy amplia de exclusión de solvente (1300 Å<sup>2</sup>/monómero). Los dos monómeros conforman, cada uno de ellos, cinco regiones de estructuras de hojas β plegadas, que se orientan de forma antiparalela (27).



**Figura 2.** Modelo computacional de la estructura del dímero de integrasa del VIH-1 a partir de la simulación computacional obtenida de la resolución de la estructura de difracción de rayos X a 1,60 Å del dímero de la integrasa del VIH-1 (IBIS) (11). En la figura se muestran los residuos catalíticos (D64, D116 y E152). Además, se localizan las regiones de interacción entre los monómeros con las dos glutaminas de los extremos carboxilo de cada monómero del dominio central (K209) y las dos cisteínas del extremo amino (C56).

**Fuente:** El modelo computacional se obtuvo mediante modelación por homología empleando el Swiss Model y posteriormente utilizando el programa DeepView 4.0 de acceso público en <http://spdbv.niehs.nih.gov/>.

Análisis sobre las propiedades moleculares de esta proteína muestran que las interacciones intermoleculares entre monómeros son predominantemente hidrofóbicas y ocurren entre los residuos W243, A248 y V250 de un monómero con los L242, V250, I257 and V259 del otro. El dímero se estabiliza mediante enlaces salinos y enlaces de hidrógeno que involucran tres dominios de estructura de hojas  $\beta$ -plegadas y las hélices  $\alpha$  1, 3, 5, 6. El principal enlace de hidrógeno que liga las dos subunidades es el que se establece entre el carboxilo del E246 de un monómero con el grupo carboxamida del Q252 del

otro (28). La interacción entre los monómeros de la región central genera una interfase de exclusión de solvente de, aproximadamente, 800 Å<sup>2</sup> por cada molécula.

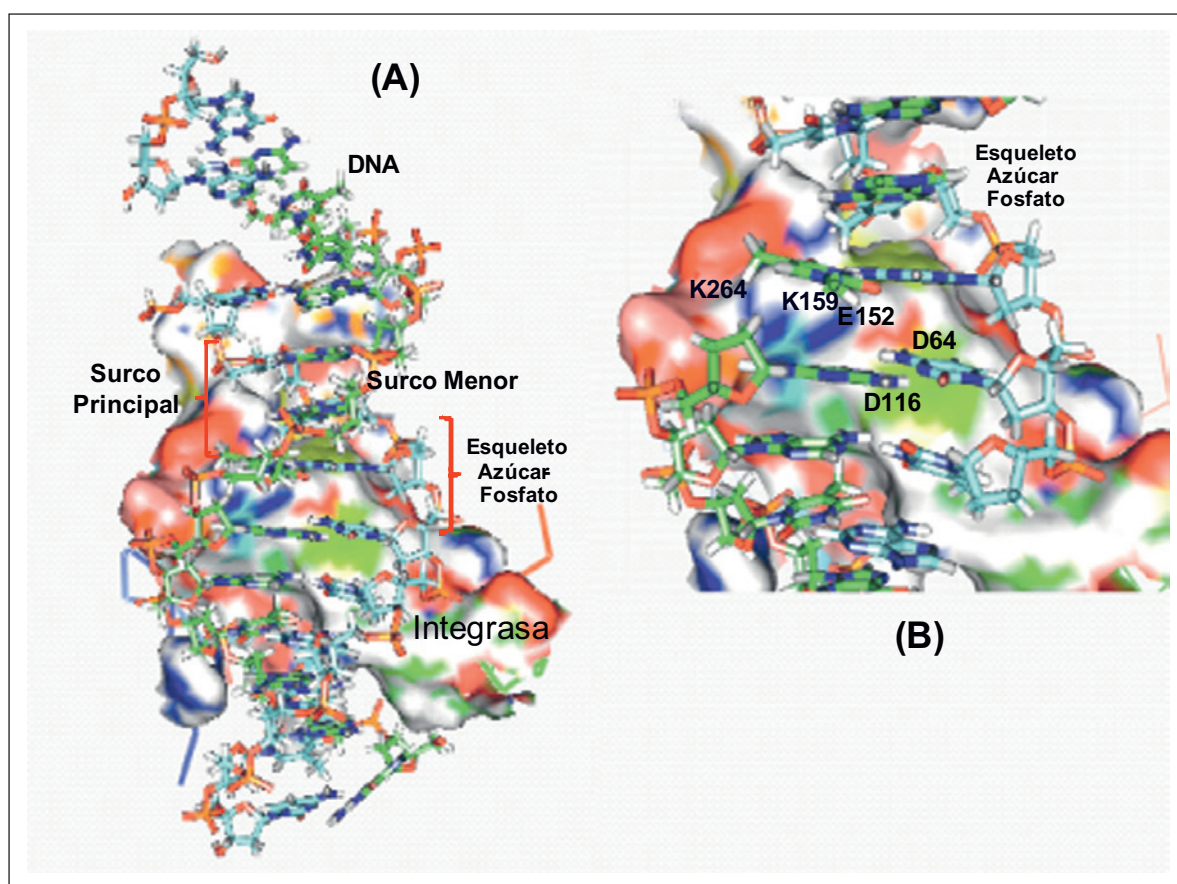
En este estudio se determinó que en el dímero, el dominio de unión al ADN está compuesto por una serie de *loops* que conectan las hojas  $\beta$  plegadas 1 y 2 del extremo amino en cada monómero. Este bolsillo tiene las dimensiones apropiadas para encajar ADN de la forma B (24 x 23 x 12 Å) (figura 3a). A partir de un modelo obtenido por difracción



de rayos X para la IN del VIH-1, cuyo código PDB es 1BIS (11), y empleando la modelación tridimensional con el programa PyMol se determinó que las cadenas laterales del residuo K264 de cada monómero interactúan con el esqueleto azúcar-fosfato del ADN blanco (figura 3b). Resultados previos obtenidos por mutagénesis sitiodirigida en esta región mostraron que el residuo K264 es clave para la interacción entre la proteína y el ADN blanco (28). Otra interacción importante es la que

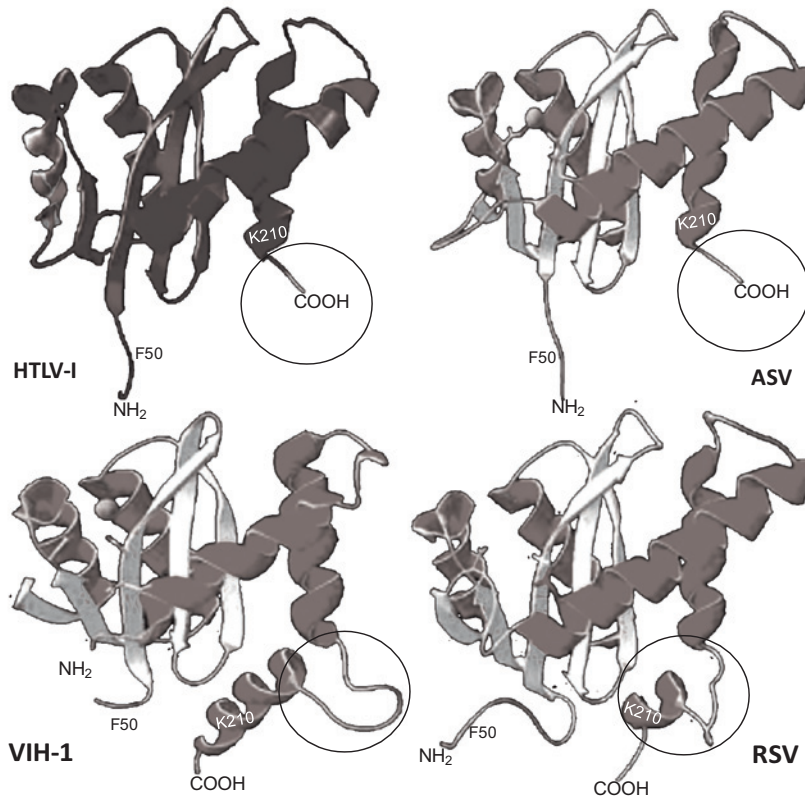
ocurre entre R231 de cada monómero y las bases nitrogenadas de la curvatura principal del ADN sustrato (29).

Del análisis comparativo efectuado con los modelos obtenidos del dominio central catalítico de la IN del HTLV-I, los del virus del sarcoma aviar (ASV), del VIH-1 y el del virus del sarcoma de Rous (RSV), se observaron arreglos tridimensionales muy semejantes (figura 4). Sin embargo, tanto en la IN



**Figura 3.** Modelación molecular de la interacción entre el dímero de la integrasa del VIH-1 con el ADN blanco de integración. Se muestra la superficie de contacto de la enzima con la doble hélice del ADN. (A). Vista panorámica de la interacción en la que se precisan los dominios del sitio activo, compuesto, entre otros, por los residuos catalíticos D64, D116 y E152; además de la región de interacción con el ADN con el K264. (B). Aproximación ampliada que muestra la interacción entre el K264 y una porción del esqueleto azúcar fosfato del ADN blanco. En el interior de la integrasa se localizan los aminoácidos catalíticos, interactuando con la porción interna de una vuelta del ADN blanco. Las figuras se obtuvieron a partir de la estructura de la integrasa del VIH-1 código PDB 1BIS, que fue obtenida por difracción de rayos X. Para ello se utilizó el programa PyMol. (D) ácido aspártico; (E) ácido glutámico; (K) lisina.

**Fuente:** Se utilizó el programa ProModII que generó un modelo básico al cual se le minimizó la energía empleando el programa Gromos 96; posteriormente con la ayuda del programa PyMol, se modeló la interacción entre la integrasa del VIH-1 y el ADN blanco de integración.



**Figura 4.** Comparación de estructuras del dominio central catalítico (50-212 aa) obtenidas por modelación molecular de las integrasas de cuatro diferentes retrovirus. Las poses de cada uno de los modelos son similares y muestran la orientación espacial de los distintos plegamientos de estructuras alfa y beta. (HTLV-I) virus linfotrópico humano tipo I. (ASV) virus del sarcoma aviar. (RSV) virus del sarcoma de Rous. (VIH-1) virus de la inmunodeficiencia humana 1. Los círculos dentro de cada una de las estructuras muestran el *loop*, (184 a 190 aa) extra en las integrasas del VIH-1 y RSV, pero ausente en las de HTLV-I y ASV; además de la alfa hélice, que se extiende de los aminoácidos 200-210 en las integrasas del VIH-1 y RSV.

**Fuente:** Usando el Swiss-Model disponible para su modelación y a partir de las secuencias cuyos códigos GenBank fueron VIH-1(1BIS\_B), HTLV-1 (EF591890), ASV (1VSK-A) y RSV (1COM\_B) se seleccionaron los mejores estructuras, éstas se les minimizó la energía utilizando el programa Gromos96 del cual se obtuvieron estructuras en gradiente conjugado para 300 ciclos

del HTLV-I como en la del ASV falta una de las hélices  $\alpha$ , característica de las IN de VIH-1 y de VSR; además, en la cuarta hélice  $\alpha$  del dominio central de la del VIH-1 existe una región de 14 residuos que en las de ASV y HTLV-I es de 21 aminoácidos.

Todos los retrovirus exhiben preferencias distintas hacia los blancos genómicos de integración; así, las características en el plegamiento terciario de regiones del dominio central catalítico pueden afectar las frecuencias de integración. Con base en los resul-

tados obtenidos sobre la simulación tridimensional de las integrasas de ASV, RSV, HTLV-I y VIH-1 es tentador especular, en el caso de los lentivirus que sería la interacción entre la integrasa y proteínas cromosómicas conocidas la que determinaría la distribución de los sitios. Previamente se ha definido que en el dominio central catalítico de la integrasa del VIH-1, las hélices  $\alpha 3$ ,  $\alpha 4$  y la  $\alpha 5$  conforman el principal sitio de reconocimiento para el factor del hospedero específico de lentivirus -LEDGF/p75- (*Lentivirus-specific host factor*) (30). Se determinó

que la organización de sitios análogos en las integrasas de ASV, HTLV-I y de Spumavirus es un poco diferente (31); estas diferencias incluyen el empaquetamiento relativo de las hélices, la orientación y tamaño del *loop* conector (21 residuos extras en la IN del HTLV-I y ASV) podrían ser factores estructurales que promoverían una interacción diferencial con proteínas celulares que conforman el complejo de preintegración. En general, los factores de unión putativos del hospedero en los Spumavirus son menos hidrofóbicos que aquellos de VIH-1 y ASV. Sin embargo, aún no ha sido posible establecer si estos dos retrovirus utilizan factores del hospedero de unión a integrasa que determinen la frecuencia de integraciones (32). De estas evidencias se puede concluir que las pequeñas variaciones en el plegamiento local de las hélice alfa del dominio central catalítico pueden intervenir, en un proceso multifactorial, en la manera como cada retrovirus se integra dentro del genoma de la célula hospedera.

### SUSTITUCIONES DE CODONES EN LA INTEGRASA DEL VIH-1

Con el fin de determinar si existen sustituciones de codones que puedan alterar la estructura terciaria de la integrasa, se determinaron los cambios de aminoácidos en diferentes posiciones en 103 secuencias completas de la IN del VIH-1 de individuos infectados consignadas en el GenBank. Se encontró que el 96,1% (99/103) de las secuencias incluidas se identificaron como subtipo B; solamente aquellas cuyos códigos de acceso eran AAL01910 (B/D), AAL01915 (B/A), AAL01917 (B/F2) y AAN05469 (B/G) correspondieron a formas recombinantes.

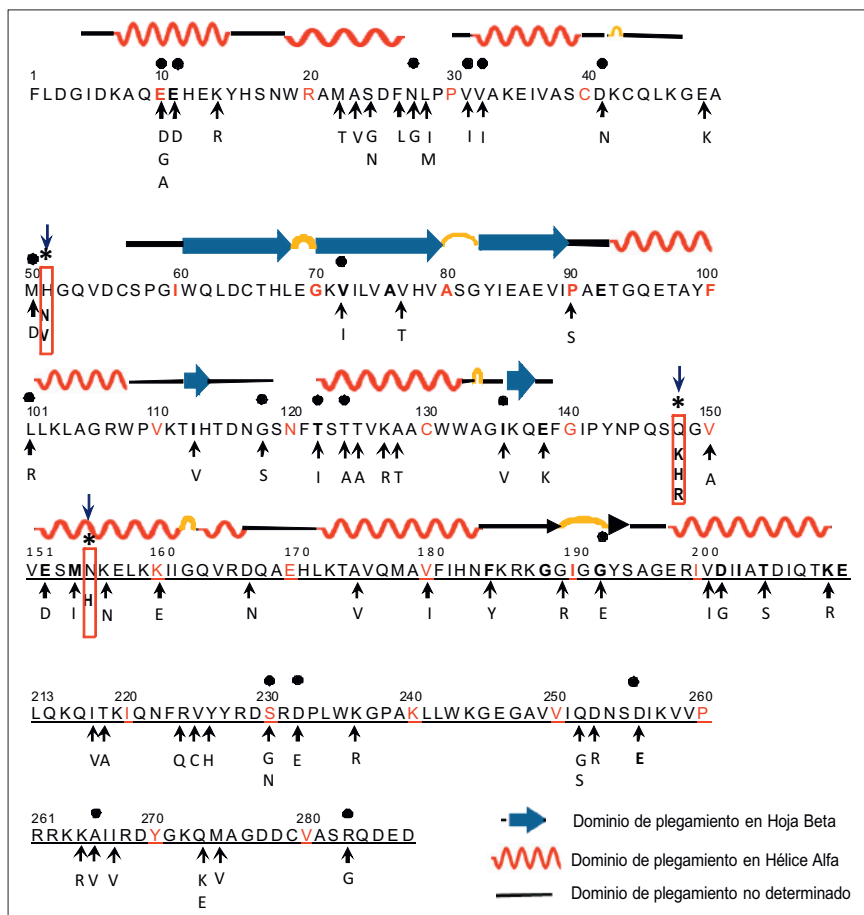
En un total de 61 codones (21,18%) se registraron 71 sustituciones, que se localizaron en diferentes posiciones en cada uno de los tres dominios de la integrasa del VIH-1 (figura 5). En el dominio amino terminal se identificaron 4 sustituciones en las secuen-

**Tabla 1.**

Distribución de las sustituciones de aminoácidos más frecuentes en la región del genoma de VIH codificada por la secuencia completa de la integrasa de 103 pacientes VIH-1.

Codón número	Consenso	Sustitución	Porcentaje (%)
10	E	D	18,4
11	E	D	12,6
27	N	G	32,1
31	V	I	9,7
32	V	I	13,6
50	M	I	14,6
72	V	I	18,4
101	L	I	31,1
123	T	A	27,0
124	T	A	16,5
201	V	I	18,4
206	T	S	11,6
211	K	R	6,8
230	S	G	14,6
230	S	N	9,7
232	D	E	13,6
256	D	E	16,5
265	A	V	30,1
284	R	G	8,7

Fuente: GenBank.



**Figura 5.** Localización de las sustituciones de aminoácidos más frecuentemente observadas en diferentes codones en los tres dominios de la IN del VIH-1. (A) alanina; (M) metionina; (V) valina; (R) arginina; (T) treonina; (Q) glutamina; (N) asparagina; (G) glicina; (S) serina; (P) prolina; (F) fenilalanina; (L) leucina; (D) ácido aspártico; (H) histidina; (E) ácido glutámico; (K) lisina; (Y) tirosina. Los círculos en negro muestran las sustituciones más frecuentes encontradas en 103 secuencias completas de la integrasa del VIH-1 de pacientes no tratados con drogas inhibitoras de la integrasa (INI). Las sustituciones marcadas con asteriscos corresponden a aquellas previamente reportadas por otros autores que están asociadas con resistencia al tratamiento con Raltegravir (34, 38). Las líneas en amarillo indican porciones de la estructura terciaria con *loops* no alfa no beta.  
**Fuente:** El cálculo y modelación de la estructura secundaria se llevó a cabo usando el programa Secondary Structure Prediction (SSP) el cual está integrado en Prime, utilizando el módulo Structure Prediction

cias analizadas, cuya frecuencia fue mayor al 5%, donde el cambio N27G (32,1%) es el más representativo (tabla 1). En el dominio central catalítico se registraron 7 cambios de aminoácidos con porcentajes mayores al 5%; en esta región, la sustitución más frecuente fue T123A identificada en el 27,0% (tabla 1). Finalmente, en el dominio carboxilo terminal se registraron 5 sustituciones que ocurrieron en más del 5% de las integrasas

en las que A265V representó el 30,2%. Sin embargo, además de estos cambios, fue posible localizar sustituciones menores en los codones 60, 74, 84, 111, 124, 125, 126, 154, 155 y 211. Los residuos E10, S24, H51, Q148 y Q273 fueron sitios polimórficos.

Un análisis de los efectos sobre la conformación tridimensional de la integrasa nativa del VIH-1, que puedan ocurrir por efecto de

las sustituciones antes descritas, mostró que solamente la sustitución T→A registrada en los codones 123 (27,0%) y 124 (16,5%) podría generar un cambio que favorecería una tendencia a un plegamiento de tipo alfa hélice en la región que se extiende del amino ácido 120 al 133. Cuando se compararon estos cambios con la estructura secundaria obtenida de la integrasa nativa, fue evidente que en esta región existe una predominancia de plegamiento  $\alpha$  (figura 4). Por lo anterior, se puede concluir que las sustituciones registradas en este estudio no introducirían ninguna alteración dramática del plegamiento en la región del dominio central catalítico que rodea al bolsillo del sitio activo de la enzima (D64, E152 y D116). Ninguna de las sustituciones observadas involucró los codones catalíticos ni al residuo K159.

Trabajos previos, en donde se empleó una metodología similar, pero que incluían secuencias completas y parciales provenientes de individuos tratados y no tratados (*naïves*) con diferentes drogas antirretrovirales (33, 34), mostraron una alta coincidencia con aquellas mutaciones naturales más frecuentemente encontradas en este estudio. Estos expandieron la gama de polimorfismos a 75 sustituciones localizadas en 58 codones distintos de la integrasa del VIH-1. Las sustituciones localizadas en los codones I72 y T124 fueron más abundantes en los subtipos A, D, C y F y en recombinantes A/E; tanto la sustitución I72V como la T124A representaron más del 70% de los aislados de los subtipos anteriormente descritos (33). Entre las sustituciones más frecuentes previamente estudiadas se incluyen N51V, Q148H/R/K y N155H, las cuales se asocian con niveles elevados de resistencia a inhibidores de la IN, como el Raltegravir, Elvitegravir y algunos derivados cetóxicos (31-34).

En el análisis de las secuencias de integrasa incluidas en el estudio, se determinó que 227 de los 288 residuos de aminoácidos estaban conservados (78,8%), lo que representa una baja variabilidad de secuencias. De este resultado, es posible sugerir que aquellos residuos asociados con la estabilidad del plegamiento de los distintos dominios, así como la multimerización, la unión al ADN, el sitio catalítico y la unión con el cofactor celular LEDGF/p75, están altamente conservados. De otra parte, los polimorfismos naturales que involucran las sustituciones L74M, T97A, S119G/R, V151I, K156N, E157Q, G163K/R, V165I, I203M, T206S, S230N, que representan frecuencias variables, estarían haciendo parte de la estructura poblacional de las cuasiespecies y reflejarían la variación natural en la región correspondiente a la integrasa; este análisis está en concordancia con los datos reportados por Myers y Pillay, 2008 (35).

El 99% de los aislamientos incluidos en estos análisis pertenecieron al subtipo B, por lo que es posible postular que las distintas sustituciones de codones registradas estarían asociadas con este subtipo en particular. En un trabajo reciente se determinó que todas las mutaciones emergentes marcadoras para la falla de los pacientes tratados con Raltegravir o Elvitegravir, así como aquellas de emergencia secundaria, no se encontraron en pacientes *naïve*, independientemente de que fueran del subtipo B o CRF02\_AG (36). Además, aquellas variaciones naturales registradas por (34), estuvieron asociadas con la prevalencia del subtipo o forma recombinante en cada población (37).

Un estudio previo acerca del efecto del cambio potencial estructural local que puedan inducir estas sustituciones asociadas con re-

sistencia a Raltegravir o Elvitegravir, mostró que ninguna de ellas, *per se*, induciría cambios drásticos en el plegamiento de dominios localizados a su alrededor. Además, en los análisis comparativos efectuados en el presente trabajo, no se encontraron en las 103 secuencias completas de IN del VIH-1, en ninguna de las sustituciones N51V, Q148H/R/K y N155H marcadoras primarias de resistencia a Raltegravir (38) o Elvitegravir (39); así pues, estas serían sustituciones seleccionadas durante los procesos de replicación de las cuasiespecies, bajo la presión selectiva del antiviral, que no producirían alteraciones importantes en el plegamiento de los dominios localizados a su alrededor.

Estos resultados, así como otros anteriormente descritos (34-41), permiten postular que la integrasa, en ausencia de la presión de inhibidores, es una enzima viral que requiere una conservación estructural y funcional de casi los dos tercios de sus aminoácidos para ejercer su función. Así, aquellas mutaciones primarias asociadas con la resistencia a inhibidores de la integrasa que son importantes, estarían ausentes o serían altamente infrecuentes en aquellos pacientes no tratados. La caracterización de residuos altamente conservados e involucrados en la interacción tanto con el ADN como con proteínas celulares ayudaría a entender y diseñar de manera más racional nuevos inhibidores contra la infección por el VIH-1 con mecanismos de acción y perfiles de resistencia más favorables.

En la actualidad, la amplia circulación de cepas VIH-1 multirresistentes a las drogas antivirales comúnmente utilizadas en el tratamiento de pacientes VIH/sida plantea la necesidad de implementar estrategias nuevas que puedan, por lo menos, disminuir

los efectos poblacionales de la resistencia. En consecuencia, la importancia de este trabajo radica en que provee un análisis de la estructura tridimensional fina de dos integrasas de retrovirus humanos de importancia clínica, las cuales se correlacionaron con aquellos polimorfismos de la región del gen *pol* que codifica por la integrasa del VIH-1. Los datos obtenidos servirán de base para que en un futuro inmediato se puedan establecer parámetros de estructura terciaria y su relación con la estabilidad genética de la integrasa; su valoración, permitirá determinar mecanismos alternativos de acción de drogas y diseñar sitios de interacción deseados, utilizando para ello mutagénesis sitio dirigida. Así pues, ante el actual panorama farmacológico y epidemiológico de la infección por retrovirus humanos, la integrasa tal cual se ha mostrado en este trabajo surge como un blanco, muy promisorio, de acción antirretroviral, que permitirá la puesta en escena de nuevos fármacos, más efectivos, para enfrentar la pandemia del sida/VIH en este inicio del siglo XXI.

### **Declaración de conflictos de interés**

Los investigadores declaran no existir ningún tipo de conflicto de intereses con respecto a los resultados de esta investigación.

### **Fuentes de financiación**

Este trabajo se realizó como parte del proyecto de investigación cofinanciado por Colciencias (código 1106-04-13095) y la Universidad del Valle, y del proyecto de investigación de la convocatoria interna 2007 de la Vicerrectoría de Investigaciones de la Universidad del Valle (contrato 1576).

## REFERENCIAS

1. Edlich RF, Hill LG, Williams FM. Global epidemic of human T-cell lymphotropic virus type-I (HTLV-I): An update. *J Long Term Eff Med Implants* 2003;13:127-40.
2. United Nations AIDS Program on HIV/AIDS. Report of the Global AIDS. UNAIDS/UN 2008.
3. Shafer RW, Schapiro JM. HIV-1 drug resistance mutations: an updated framework for the second decade of HAART. *AIDS Rev* 2008;10:67-84.
4. García-Vallejo F, Domínguez MC. La genotipificación y fenotipificación de la resistencia del virus de la inmunodeficiencia humana a los fármacos antirretrovirales. *Colombia Médica* 2003; 34:143-54.
5. Hughes A, Barber T, Nelson M. New treatment options for HIV salvage patients: an overview of second generation PIs, NNRTIs, integrase inhibitors and CCR5 antagonists. *J Infect* 2008;57:1-10.
6. Pace P, Rowley M. Integrase inhibitors for the treatment of HIV infection. *Curr Opin Drug Discov Devel* 2008;11:471-9
7. Cherepanov P, Pluymers W, Claeys A, Proost P, De Clercq E, Debyser Z. High-level expression of active HIV-1 integrase from a synthetic gene in human cells. *FASEB J* 2000;14:1389-99.
8. Serrao E, Odde S, Ramkumar K, Neamati N. Raltegravir, Elvitegravir, and Metoogravir: the birth of "me-too" HIV-1 integrase inhibitors. *Retrovirology* 2009;6:25
9. LaFemina RL, Schneider CL, Robbins HL, Callahan PL, LeGrow K, Roth E, et al. Requirement of active human immunodeficiency virus type 1 integrase enzyme for productive infection of human T-lymphoid cells. *J Virol* 1992;66:7414-9.
10. Hanson K, Hicks C. New antiretroviral drugs. *Curr HIV/AIDS Rep* 2006; 3:93-101.
11. Goldgur Y, Dyda F, Hickman AB, Jenkins TM, Craigie R, Davies DR. Three new structures of the core domain of HIV-1 integrase: an active site that binds magnesium. *Proc Natl Acad Sci USA* 1998;95:9150-4.
12. Yang ZN, Mueser TC, Bushman FD, Hyde CC. Crystal structure of an active two-domain derivative of Rous sarcoma virus integrase. *J Mol Biol* 2000; 296:535-48.
13. Maignan S, Guilloteau JP, Zhou-Liu Q, Clement-Mella C, Mikol V. Crystal structures of the catalytic domain of HIV-1 integrase free and complexed with its metal cofactor: high level of similarity of the active site with other viral integrases. *J Mol Biol* 1998 ;282:359-68.
14. Bujacz G, Jaskolski M, Alexandratos J, Wlodawer A, Merkel G, Katz RA, et al. High-resolution structure of the catalytic domain of avian sarcoma virus integrase. *J Mol Biol* 1995;253:333-46.
15. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* 1997;18: 2714-23.
16. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res* 2003;31:81-5.
17. Peitsch M.C. Protein modelling by E-Mail. *BioTechnology* 1995;13:658-60.
18. Van Gunsteren WF, Billeter SR, Eising A, Hünenberger PH, Krüger P, Mark A E, et al. *Biomolecular Simulations: The GROMOS96 Manual and User Guide*. Zürich: VdF Hochschulverlag ETHZ, 1996.
19. Rost B, Sander C. Conservation and prediction of solvent accessibility in protein families. *Proteins* 1994;20:216-26.
20. Rost B, Sander C. Secondary structure prediction of all-helical proteins in two states. *Protein Eng* 1993;6:831-6.
21. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics* 2007;23:2947-8.
22. Zeinalipour-Loizidou E, Nicolaou C, Nicolaidis A, Kostrikis LG. HIV-1 integrase: from biology to chemotherapeutics. *Curr HIV Res* 2007;5:365-88.

23. Savarino A. In-Silico docking of HIV-1 integrase inhibitors reveals a novel drug type acting on an enzyme/DNA reaction intermediate. *Retrovirology* 2007;4:79-91.
24. Tsurutani N, Kubo M, Maeda Y, Ohashi T, Yamamoto N, Kannagi M, et al Identification of critical amino acid residues in human immunodeficiency virus type 1 IN required for efficient proviral DNA formation at steps prior to integration in dividing and nondividing cells. *J Virol* 2000;74:4795-806.
25. Eijkelenboom AP, van den Ent FM, Wechslerberger R, Plasterk RH, Kaptein R, Boelens R. Refined solution structure of the dimeric N-terminal HHCC domain of HIV-2 integrase. *J Biomol NMR* 2000;18:119-28.
26. Lloyd AG, Ng YS, Muesing MA, Simon V, Mulder LC. Characterization of HIV-1 integrase N-terminal mutant viruses. *Virology* 2007;360:129-35.
27. Wang JY, Ling H, Yang W, Craigie R. Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. *EMBO J* 2001;20:7333-43.
28. Dyda F, Hickman AB, Jenkins TM, Engelman A, Craigie F, Davies DR. Crystal Structure of the Catalytic Domain of HIV-1 Integrase: Similarity to Other Polynucleotidyl Transferases. *Science* 1994;266:1981-6.
29. Lodi PJ, Ernst JA, Kuszewski J, Hickman AB, Engelman A, Craigie R, et al. Solution Structure of the DNA Binding Domain of HIV-1 Integrase. *Biochemistry* 1995;34:9826-33.
30. Cherepanov P, Ambrosio AL, Rahman S, Ellenberger T, Engelman A. Structural basis for the recognition between HIV-1 integrase and transcriptional coactivator p. 75. *Proc Natl Acad. Sci. USA* 2005;102:17308-13.
31. Valkov E, Gupta SS, Hare S, Helander A, Roversi P, McClure M, Cherepanov P. Functional and structural characterization of the integrase from the prototype foamy virus. *Nucleic Acid Res* 2009;37:243-55.
32. Delelis O, Carayon K, Guiot E, Leh H, Tauc P, Brochon JC, et al. Insight into the integrase-DNA recognition mechanism. A specific DNA-binding mode revealed by an enzymatically labeled integrase. *J. Biol. Chem* 2008;283:27838-27849.
33. Van Baelen K, Van Eygen V, Rondelez E, Stuyver LJ. Clade-specific HIV-1 integrase polymorphisms do not reduce raltegravir and elvitegravir phenotypic susceptibility. *AIDS* 2008;22:1877-80.
34. Rhee SY, Liu TF, Kiuchi M, Zioni R, Gifford RF, Holmes SP, et al. Natural variation of HIV-1 group M integrase: implications for a new class of antiretroviral inhibitors. *Retrovirology* 2008;5:74-84.
35. Myers RE, Pillay D. Analysis of natural sequence variation and covariation in human immunodeficiency virus type 1 integrase. *J Virol* 2008;82:9228-35.
36. Maïga AI, Malet I, Soulie C, Derache A, Koita V, Amellal B, et al. Genetic barriers for integrase inhibitor drug resistance in HIV type-1 B and CRF02\_AG subtypes. *Antivir Ther* 2009;14:123-9.
37. Ceccherini-Silberstein F, Malet I, D'Arrigo R, Antinori A, Marcelin AG, Perno CF. Characterization and structural analysis of HIV-1 integrase conservation. *AIDS Rev* 2009;11:17-29.
38. Malet I, Delelis O, Valantin MA, Montes B, Soulie C, Wirlden M, et al. Mutations associated with failure of raltegravir treatment affect integrase sensitivity to the inhibitor in vitro. *Antimicrob Agents Chemother* 2008;52:1351-8.
39. Goethals O, Clayton R, Van Ginderen M, Vereycken I, Wagemans E, Geluykens P, et al. Resistance mutations in human immunodeficiency virus type 1 integrase selected with elvitegravir confer reduced susceptibility to a wide range of integrase inhibitors. *J Virol* 2008;82:10366-74.
40. Pommier Y, Johnson AA, Marchard C. Integrase inhibitors to treat HIV/AIDS. *Nat Rev Drug Discov* 2005;4:236-48.
41. Jegede O, Babu J, Di Santo R, McColl DJ, Weber J, Quiñones-Mateu M. HIV type 1 integrase inhibitors: from basic research to clinical implications. *AIDS Rev* 2008;10:172-89.