Functional conservation of HIV-1 gag: implications for rational drug design

LI, G.[1], **VERHEYEN, J.**[2], **RHEE, S.-Y.**[1,3], **VOET, A.**, **EUSÉBIO, M.**[5], **VANDAMME, A.-M.**[1,5], **THEYS, K.**[1]

1.Laboratory for Clinical and Epidemiological Virology, Rega Institute for Medical Research, Department of Microbiology and Immunology, KU Leuven, Belgium; **2.**Institute of Virology, University Hospital, University Duisburg-Essen, Essen, Germany **3.**Division of Infectious Diseases, Department of Medicine, Stanford University, Stanford, California, USA; **4.**Zhang IRU, RIKEN Institute Laboratories, Hirosawa 2-1, Wako-shi, Saitama, Japan; **5.**Centro de Malária e Outras Doenças Tropicais, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Portugal

14th EUROPEAN AIDS CONFERENCE
OCTOBER 16–19, 2013
BRUSSELS, BELGIUM

Objectives

HIV-1 replication can be successfully blocked by targeting gag gene products, offering a promising strategy for new drug classes that complement current HIV-1 treatment options. However, advanced clinical phase studies of investigational gag inhibitors have shown that naturally occurring polymorphisms in drug binding sites can severely compromise the antiviral activity. Therefore, a comprehensive understanding of gag natural diversity is needed.

Introduction

A curative therapy or preventive vaccine for HIV-1 infected patients remains elusive to date and standard treatment is confronted with the emergence of antiviral resistance to existing drug classes, urging for inhibitors with new mechanisms of action[1]. The gag polyprotein, essential for HIV-1 morphogenesis, comprises four major domains - matrix, capsid, nucleocapsid, p6 and two small spacer peptides p1 and p2[2]. Recently, HIV-1 inhibitors that target different stages of virion morphogenesis have demonstrated promising antiviral activity, mainly by inhibiting capsid assembly, disrupting nucleocapsid binding with viral RNA/DNA or blocking proteolytic processing of polyproteins during maturation[2, 3, 4, 5]. Studies that investigated the implications of extensive HIV-1 diversity for gag-directed drug development are lacking to date. In this large-scale analysis, we examined the distribution of naturally occurring sequence variability in full-length gag sequences of major HIV-1 subtypes. Moreover, we evaluated the impact of HIV-1 subtypes on the conservation of gag drug binding positions and multisite binding pockets published to date.

Methodology

We retrieved 12543 gag sequences of 8 major HIV-1 subtypes spanning all 1500 base pairs from the HIV Los Alamos database. Hypermutated sequences were detected using the Los Alamos hypermut tool. HIV-1 subtype was determined by the Rega and COMET subtyping tools. Information on 50 known gag candidate inhibitors and 136 binding sites was retrieved from literature. To quantify the degree of functional conservation, a conservation index was calculated for each position by averaging pairwise dissimilarity scores between all AAs using BLOSUM62 matrix.

References

[1] Engekman,A., Cherepanov, P. (2012). *Nat Rev Microbiol*, 10, 279-90.
[2] Waheed, A.A., Freed E.O. (2012). *AIDS Res Hum Retroviruses*, 28, 54-75.
[3] Bocanegra, R. *et al.* (2012). *Virus Res*, 169, 388-410.
[4] Dau, B., Holodniy, M. (2009). *Drugs*, 69, 31-50.
[5] Salzwedel, K. *et al.* (2007). *AIDS Rev*, 9, 162-72.
[6] Gaschen, B. *et al.* (2002). *Science*, 296, 2354-60.
[7] Rhee, S.Y. *et al.* (2008). *Retrovirology*, 5, 74.
[8] Zhang, Z. *et al.* (2011). *Bioinformatics*, 27, 2083-8.
[9] Blair, W.S. *et al.* (2010). *PLoS Pathog*, 6, e1001220.
[10] Zetner, I. *et al.* (2013). *Bioorg Med Chem Lett*, 23, 1132-5.

Acknowledgments

Collaborative HIV and Anti-HIV Drug Resistance Network (CHAIN, grant Health-F3-2009-223131, European Community's Seventh Framework Programme FP7/2007-2013)
BEST HOPE: funded through HIVERA, grant 249697)
Fonds voor Wetenschappelijk Onderzoek - Flanders (FWO) grant G.0611.09N
FWO K8.012.12N
China Scholarship Council
JSPS funding

Main Results

Complete conservation across all subtypes was detected in 147 out of 500 positions (29%), with the highest level of conservation observed for capsid protein. Almost half (41%) of the 136 known drug binding sites were overall conserved, but all inhibitors were confronted with natural occurring polymorphisms in their binding sites, of which some clearly correlate with the HIV-1 subtype. Integration of sequence and structural information revealed one drug binding pocket with minimal genetic variability, which is situated at the N-terminal domain of the capsid protein.

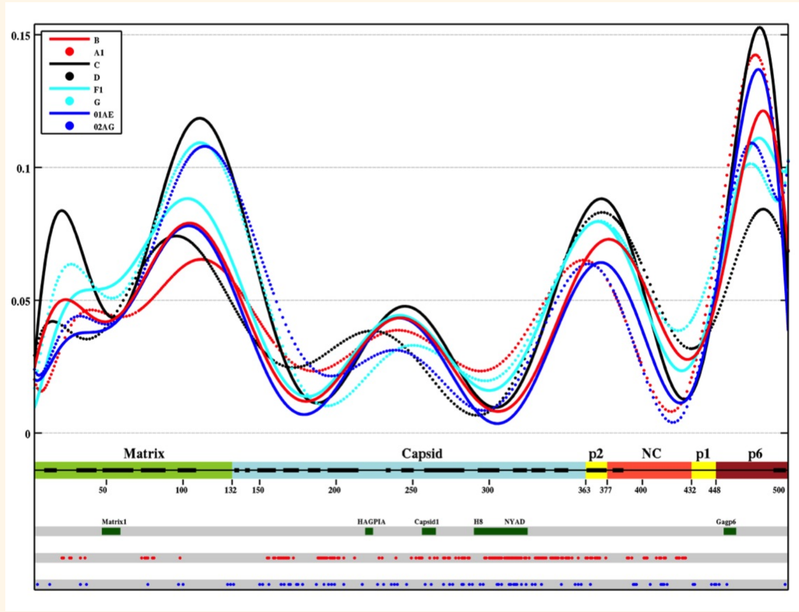


Figure 1. Density plots of CI values are shown for 8 HIV-1 subtypes. Within each protein region, the secondary structure is shown: thick lines for helices and thin lines for coiled-coil structures. Positions conserved in all subtypes are shown in blue (layer [1]), the known drug binding sites are shown in red (layer [2]) and the regions where HIV-1 peptide inhibitors have been derived are indicated in green (layer [3]).

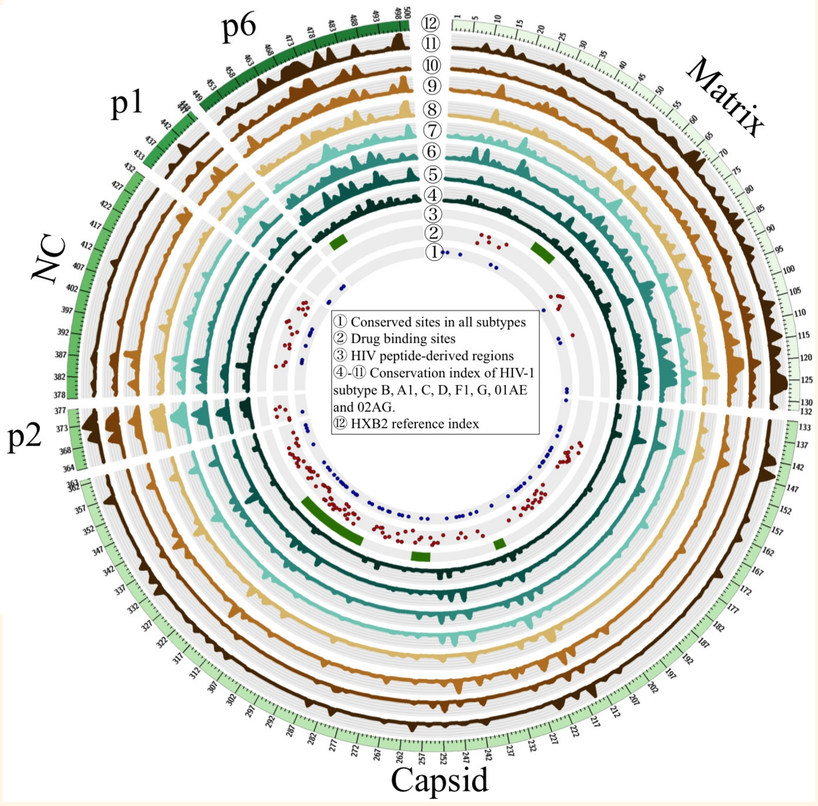


Figure 2. The distribution of CI values in 500 gag positions and annotation with drug binding sites. Visualization software: Circos v0.64 (http://circos.ca/).

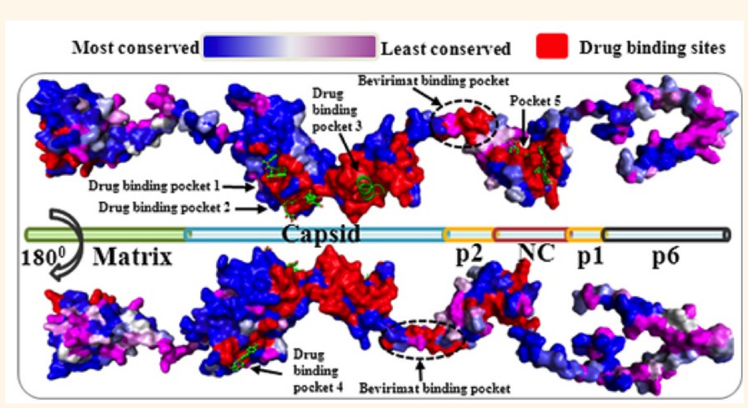


Figure 3. Mapping of drug bidding sites and drug binding pockets to HIV-1 gag protein structures. Known drug bibding sites are colored red. the front and back views of gag structures are shown. The surface spectral colors indicate the most conserved (blue CI = 0) to the least conserved positions (pink CI $\geq$ 0.1). Hypothesized bevirimat binding sites are also annotated.

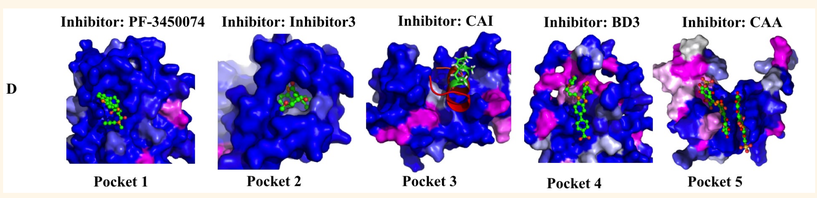
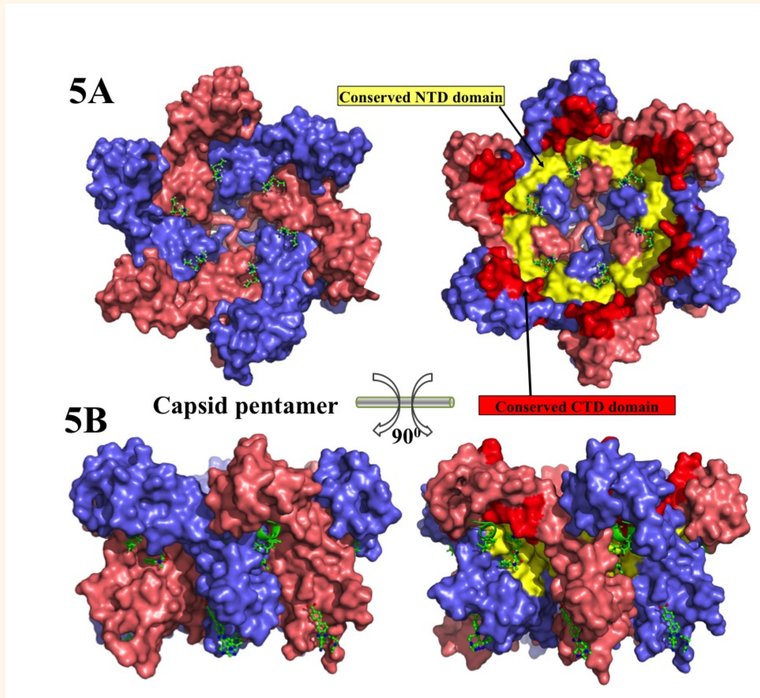


Figure 4. The surface representation of five drug binding pockets in HIV-1 subtype B. PDB data of gag proteins: matrix, 1HIW; capsid, 3NTE; p2, 1U57; nucleocapsid, 2M3Z; p6, 2C55. PDB data of capsid inhibitors: 2BUO, 2L6E, 2XDE, 4E91, 4E92, 2JPR and 4INB, each of which was superimposed to 3H4E using PyMOL V1.5 (http://www.pymol.org/). PDBs of 5 binding pockets: pocket 1, 2XDE; pocket 2, 4INB; pocket 3, 2BUO; pocket 4; 4E91 and pocket 5, 2M3Z.



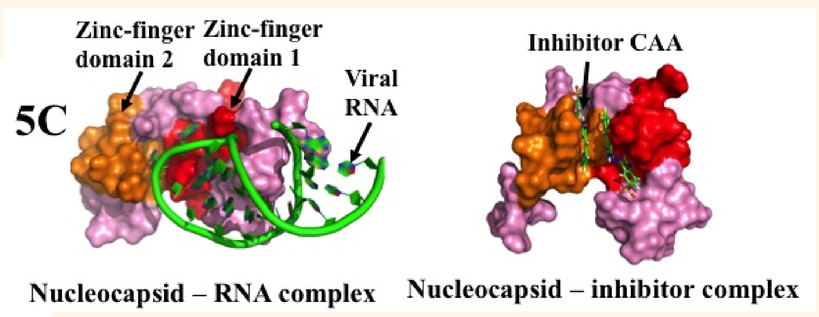


Figure 5. Conserved regions in capsid and nucleocapsid. The capsid hexamer structure is shown in top and side views in Fig (A) and (B) respectively (PDB: 3H4E). Conserved NTD-NTD interaction domains are colored yellow. Conserved NTD-CTD interaction domains are colored red. Fig (C) shows the conserved zinc-finger domains in nucleocapsid. The nucleocapsid - RNA and nucleocapsid - inhibitor complexes are shown on the left and right side, respectively (PDB: 1A1T, 2M3Z). The first zinc-finger domain (nucleocapsid position: 14-29, corresponding to gag HXB2 position: 389-404) and the second zinc-finger domain (nucleocapsid position: 35-50, corresponding to 410-425) are colored red and orange, respectively.

Conclusions

This first large-scale analysis of full-length HIV-1 gag provided a detailed mapping of natural diversity across major subtypes, and highlighted the conserved drug binding pockets in capsid. Our results contribute to the optimization of gag inhibitors in rational drug design, given that drug binding sites should ideally be conserved across all HIV-1 subtypes.