Université de Paris

École doctorale: Médicament- Toxicologie-Chimie-Imagerie (MTCI)

Laboratoire : Cibles Thérapeutiques et Conception de Médicaments (CiTCoM), CNRS UMR8038, Faculté de Pharmacie, Université de Paris

Structures and interactions of primatelentiviral accessory proteins Vpr, Vpx and p6 in the presence of DPC micelles

Xiao WANG

Thèse de doctorat de Biologie Structurale (ED 563)

Dirigée par Serge Bouaziz

Co-dirigée par Sylvie Nonin-Lecomte

Présentée et soutenue publiquement le 21 Septembre 2021

Devant un jury composé de :

Dr Serena Bernacchi	DR, Université de Strasbourg	Rapportrice
Dr Olivier Mauffret	DR, Université Paris-Saclay	Rapporteur
Dr Florence Margottin-Goguet	DR, Institut Cochin	Examinatrice
Dr Emeric Miclet	MCU, Sorbonne Université	Examinateur
Dr Sylvie Nonin-Lecomte	CR, Université de Paris	Co-directrice de thèse
Dr Serge Bouaziz	DR, Université de Paris	Directeur de thèse

Acknowledgements

This research was performed in the research unit Cibles Thérapeutiques et Conception de Médicaments (CiTCoM) at the Faculté de Pharmacie in Université de Paris. Firstly, I would like to thank the committee members for generously offering their time, support, guidance and judging my work. I would also like to thank Prof. Nicolas Leulliot, the director of the CiTCoM, for accepting and allowing me in the laboratory. I would also like to thank the people in the lab who helped me during these four years since 2017.

I would like to give a big thank to my Ph.D. supervisor Serge Bouaziz, the group leader of our team, Structural and functional studies of new therapeutic targets by NMR. Thank you very much for your support, understanding, advice and experience for my work. Thank you for your kindness and patience in learning French language and in English pronunciation. Thank you for helping me on nuclear magnetic resonance in biology (bioNMR) and for the manipulation of NMR experiments in the chemistry of natural substances institute (ICSN), Gif sur Yvette. Thank you for the installation of some software, the papers modification and submission. Thank you for your translation in every meeting. You are very friendly, funny, and patient, while I am a straight mind man with acute temper. Sometimes, when we have some debates and contrary views, you can be tolerant and explain them again and again.

I would also like to thank my co-supervisor Sylvie Nonin-Lecomte for her advice on the expression and purification of protein. Thank you for the manipulation of NMR experiments and the assignment of proteins. Thanks for offering your time to the modification of papers and my Ph.D. thesis. I also want to thank the members of my group for supporting and helping me during the four years. Thanks to Hesna Kara, Pascale Coric, Édith Gincel, Philippe Laugâa, and Valéry Larue for your advice on experiments, the utilization of programs ARIA, XPLOR-NIH, and NMR spectroscopy. You gave me a nice and relaxed research atmosphere.

I also want to thank all the people that have helped and supported me. Thanks to Cyril Garnier, Lila Delbos, Luc Ponchon, Gilles Phan, Cathy Bonamour Du Tartre, Magali Blaud, and Jean-Francois Gaucher in the lab for the reagent order, the mass spectrometry, and the manipulation of different equipment during my thesis. Thanks to Andrea Cimarelli, Centre International de Recherche en Infectiologie (CIRI), Université Lyon 1, for offering the plasmid pET-11d SIVmac Vpx. I also want to thank Nelly Morellet for collecting NMR data at the ICSN, Gif sur Yvette, France. Thanks to Rodrigue Marquant, CNRS Labo des Biomolécules - UMR 7203, for the syntheses of p6 peptides.

Finally, I would like to thank my families and friends for their love and support. Especially since the Covid-19 break out, they have been worrying about me. I would especially like to thank my beloved wife, Dr. Xiaowei Chen for supporting me for everything. She helped me analysis the experiment data, tried to make beautiful figures, modified my manuscripts and Ph.D. thesis. She is such a good wife always cheering me up and always optimistically spreading positive energy to me. Thank you for always accompanying with me and supporting me. I am happy and honored to spend the rest of my life with you.

Contents

Abbreviations

aa: Amino acid

AIDS: Acquired immune deficiency syndrome

ALIX: ALG-2-interacting protein X

ANT: Adenine nucleotide translocator

ARDs: Antiretroviral drugs

ART: Antiretroviral treatment

ARIA: Ambiguous restraints for iterative assignment

ASP: Antisense protein

AP-1: Activator protein 1

APOBEC3: Apolipoprotein B mRNA editing enzyme catalytic polypeptide-like 3

CA: Capsid

CCR5/CXCR4: CC-chemokine receptors 5/CXC-chemokine receptor 4

CCPNMR: Collaborative computing project for NMR

CD4: Cluster of differentiation 4

CDK9: Cyclin dependent kinase 9

CNS: Crystallography & NMR system

cryo-EM: Cryo-electron microscopy

CSP: Chemical shift perturbation

CT: Cytoplasmic N-terminal tail

CTD: C-terminal domain

CypA: Cyclophilin A

DCs: Dendritic cells

DPC: Dodecylphosphocholine

ECD: Extracellular domain

ESCRT: Endosomal sorting complex required for transport

GPI: Glycosyl phosphatidylinositol

HBR: Highly basic region

HDAC: Histone deacetylase

hCG1: Human nucleoporin CG1

HIF-1α: Hypoxia inducible factor 1 alpha

His6: Hexa-histidine

HIV: Human immunodeficiency virus

HPLC: High performance liquid chromatography

HSQC: Heteronuclear single quantum coherence

HUSH: Human silencing hub

IBB: Importin-β binding domain

IN: Integrase

Initiator tRNA: eIF2-GTP/Met-tRNAi

IFN-α: Interferon-α

IRES: Internal ribosome entry site

LTR: Long terminal repeat

MA: Matrix

MHR: Major homology region

MPP8: Matrix metalloproteinase-8

NC: Nucleocapsid

NF-κB: Nuclear factor kappa-light-chain-enhancer of activated B cells

NLS: Nuclear localization signal

NMR: Nuclear magnetic resonance

NMT: N-myristyltransferase

NPCs: Nuclear pore complexes

NTD: N-terminal domain

PCNA: Proliferating cell nuclear antigen

PEV: Position-effect variegation

PI4,5P2: Phosphatidylinositol 4,5-bisphosphate

PICs: Pre-integration complexes

Pol II: polymerase II

Pr55Gag: Gag precursor

Pr160GagPol: GagPol precursor

pPLK1: Phosphorylated PLK1

RPA: Replication protein A

SAMHD1: Sterile alpha motif and HD-domain–containing protein 1

SDS: Sodium dodecyl sulphate

SETDB1: SET domain bifurcated 1

SIV: Simian immunodeficiency virus

SIVgor: SIV in wild gorillas

Spl: Spironolactone

TAK1: Transforming growth factor-β-activated kinase 1

TAR: Transactivation response

TASOR: Transgene activation supreressor

TFA: Trifluoroacetic acid

TFE: Trifluoroethanol

TMD: Transmembrane domain

TRIM5α: Tripartite motif containing 5 α

TSG101: Tumor susceptibility gene 101

UNG2: Uracil-DNA glycosylase-2

Short Abstract (English)

The primate lentiviral accessory proteins Vpr and Vpx are specifically incorporated into virions by direct interaction with p6 of the precursor Pr55Gag ($p6^{gag}$) in the early stages of viral replication. The identified functions of $p6^{gag}$ usually occur close to the plasma membrane. However, the structure of p6 under membrane mimicking conditions has never been characterized and the mechanistic of the recognition of Vpr and Vpx by $p6^{gag}$ remains elusive either. Current studies focus on treatments targeting viral enzymes to reduce the viral load in cells at different stages of the replication cycle. Our objective is to identify new areas to block the viral replication. The structural determination of p6 and their interactions with Vpr/Vpx will bring a new view on how occurs the incorporation of these proteins in newly formed virions, which helps to identify compounds capable of disrupting the formation of these complexes. In this study, we characterized, by NMR, the 3D structures of HIV-1 and SIVmac p6 in the presence of DPC micelles and provided structural details on the mode of recognition of Vpr and Vpx by p6. Compared to previous p6 structures obtained in aqueous trifluoroethanol solution, our results disclose the existence of an additional short helix (residues 24-27) in HIV-1 p6. Using chemical shift perturbations experiments, we show that both Vpr and Vpx interact with p6. HIV-1 Vpr hot region I, containing I74, G75, C76, R77, I81, G82, I84 and Q85, and hot region II, including V31, I37, S41, L42, Y50, D52, I63 and Q66, interact with the two motifs of HIV-1 p6, ^{15}FXFG^{18} and ^{41}LXXLF^{45}. Similarly, hot regions I, including V29, E43, and W49, and hot region II, containing R7, L25, M62, Q76, and F80 in SIVmac Vpx mainly target SIVmac p6 residues K36 and R38, as well as the ^{17}DXAXXLL^{23} motif. A weak cross-interaction was also identified between SIVmac Vpx and HIV-1 p6. Our study thus provides atomic-level insights of the incorporation of Vpr and Vpx into virions mediated by p6.

Keywords:

Vpr, Vpx, p6, incorporation, structure, interaction, DPC

Short Abstract (French)

Les protéines accessoires lentivirales de primate Vpr et Vpx sont spécifiquement incorporées dans les virions par interaction directe avec p6, domaine C-Terminal du précurseur Pr55Gag ($p6^{gag}$). Les fonctions identifiées de $p6^{gag}$ s'effectuent généralement à proximité de la membrane plasmique. Cependant, la structure de p6 dans des conditions mimant l'environnement membranaire n'a jamais été caractérisée et le mécanisme de reconnaissance de Vpr et Vpx par $p6^{gag}$ reste incompris egalement. Dans cette étude, nous avons caractérisé, par RMN, les structures 3D de HIV-1 et SIVmac p6 en présence de micelles de DPC et fourni des détails structuraux sur le mode de reconnaissance de Vpr et Vpx par p6. Par rapport aux structures de p6 obtenues précédemment en solution aqueuse de trifluoroéthanol, nos résultats révèlent l'existence d'une hélice courte supplémentaire (résidus 24-27) dans HIV-1 p6. En utilisant des expériences de perturbations de déplacements chimiques, nous avons montré que Vpr et Vpx interagissent avec p6. HIV-1 Vpr deux possède deux regions: (1) contenant I74, G75, C76, R77, I81, G82, I84 et Q85, et (2) comportant les résidus V31, I37, S41, L42, Y50, D52, I63 et Q66. Ces deux régions interagissent avec les deux motifs de HIV-1 p6, ^{15}FXFG^{18} et ^{41}LXXLF^{45}. De même, la région I comportant V29, E43 et W49, et la région II, contenant R7, L25, M62, Q76 et F80 dans SIVmac Vpx ciblent principalement les résidus de SIVmac p6 K36 et R38, ainsi que le motif ^{17}DXAXXLL^{23}. Une faible interaction croisée a également été identifiée entre SIVmac Vpx et HIV-1 p6. Notre étude fournit ainsi un aperçu au niveau atomique de l'interaction permettant l'incorporation de Vpr et Vpx dans les virions par l'intermédiaire de p6.

Mots clés:

Vpr, Vpx, p6, incorporation, structure, interaction, DPC

Résumé en Français

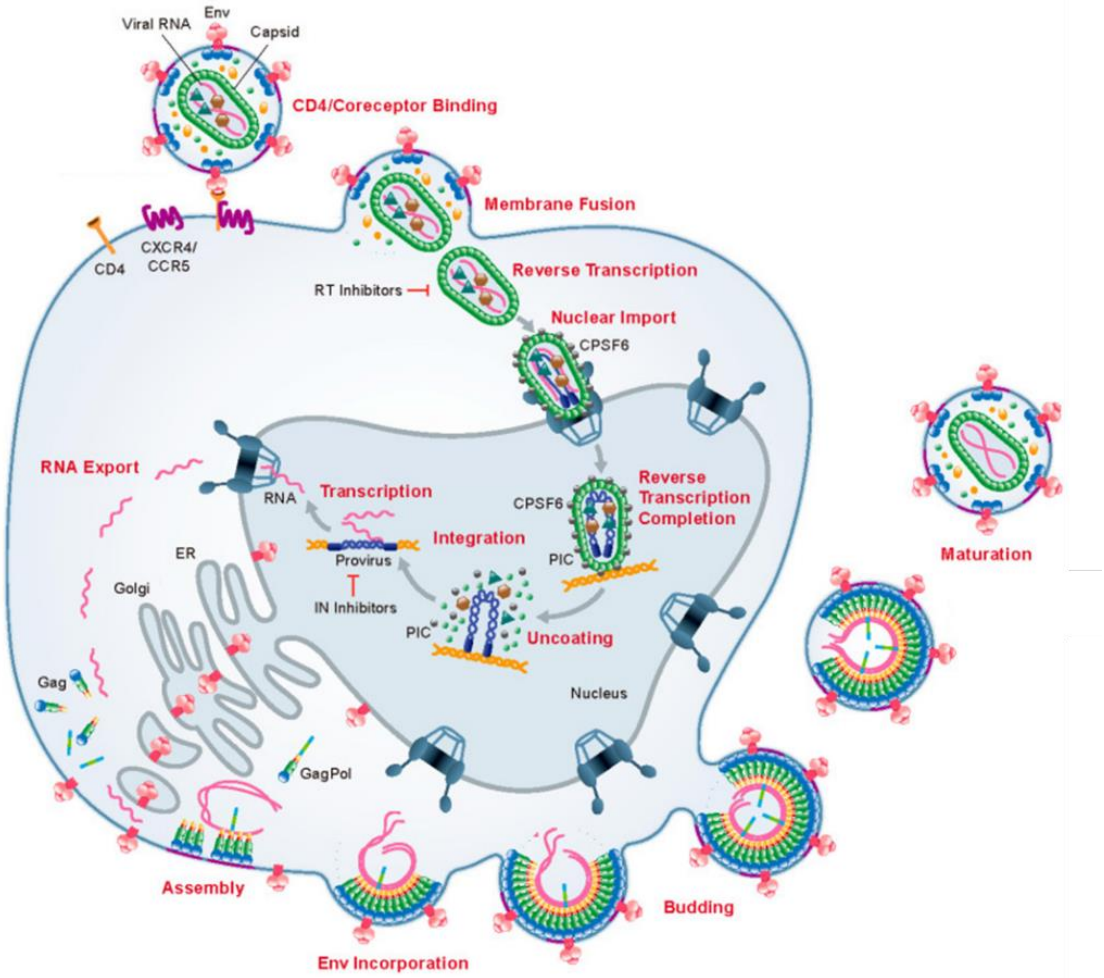
I. Introduction

Cycle réplicatif du VIH

Les lentivirus sont une sous-famille de rétrovirus, dont l'ARN est rétro-transcrit puis intégré à l'ADN du génome de l'hôte. La famille des lentivirus des primates comprend les virus de l'immunodéficience simienne (VIS) et les virus de l'immunodéficience humaine (VIH). Les VIH sont classés en deux grands types, le VIH-1 et le VIH-2. Le syndrome d'immunodéficience acquise (SIDA) est causé par les virus de l'immunodéficience humaine (VIH). Actuellement, le traitement antirétroviral (ART) n'élimine pas l'infection par le VIH en raison de l'infection persistante latente et de la non-reconnaissance par le système immunitaire. En outre, les médicaments à base d'anticorps entraînent une lourde charge financière et leur application à grande échelle n'est pas disponible dans les pays en développement. Par conséquent, pour éradiquer complètement le VIH, des études sont encore nécessaires pour surmonter la résistance aux médicaments, identifier de nouvelles cibles et développer de nouveaux médicaments, anticorps ou vaccins.

Le cycle de réplication du VIH-1 comprend principalement deux phases : la phase précoce comportant l'attachement viral, la fusion membranaire, la transcription inverse, l'entrée nucléaire et l'intégration. La phase tardive comporte la transcription, la traduction, l'assemblage, le bourgeonnement et la maturation (Turner et Summers, 1999). Dans la phase précoce, les particules virales s'attachent d'abord à la surface et pénètrent dans les cellules hôtes exprimant la glycoprotéine cluster de différenciation 4 (CD4) à leur surface. Pour envahir les cellules hôtes, les pointes de la protéine virale Env formées par l'association non covalente de gp41 et gp120 subissent d'abord des réarrangements structurels. Ensuite, les peptides de fusion gp41 hydrophobes sont insérés dans la membrane de la cellule hôte pour faciliter la fusion de l'enveloppe virale et de la membrane plasmique et permettre à la nucléocapside du virus de pénétrer dans la cellule. L'ARN simple brin est utilisé comme matrice par la transcriptase inverse pour synthétiser un ADNc simple brin, qui à son tour sert de matrice pour la synthèse d'ADN double brin par l'ADN

polymérase dans le cytoplasme. L'ADN viral utilisé comme matrice est transcrit en ARN la polymérase II (Pol II) de l'hôte dans le noyau. Ensuite, certaines molécules d'ARN sont modifiées en 5’ et 3’ et pour devenir le génome viral et les autres sont épissées en ARNm viraux, qui sont traduits en polyprotéines précurseurs. L'assemblage viral commence par la synthèse de la polyprotéine Gag dans le cytosol. La polyprotéine Gag et l'ARN dimère s'assemblent dans la particule virale naissante et immature avec le précurseur GagPol au niveau de la membrane plasmique. Au cours du processus de bourgeonnement, la protéase virale clive les polyprotéines Gag et GagPol pour déclencher la maturation du virus, ce qui entraînera des changements majeurs dans la morphologie du virion, la génération du noyau de capside conique et la production des MA, CA, NC entièrement transformées ainsi que les protéines p6, PR, RT et IN. Ces protéines transformées se réarrangent pour générer le virion mature et infectieux.



Cycle de réplication du VIH-1 (Kleinpeter et Freed, 2020).

Structure et fonction de Vpr / Vpx / p6

Les protéines accessoires lentivirales de primates Vpr et Vpx sont toutes deux incorporées dans des virions par interaction directe avec p6, le domaine C-terminal du précurseur Pr55Gag (Cohen et al., 1990 ; Kappes et al., 1993 ; Paxton et al., 1993). Cette incorporation est essentielle pour la réplication virale dans les cellules au repos, la translocation du complexe de pré-intégration (PIC) vers le noyau et pour la neutralisation des facteurs antiviraux cellulaires tels que l'uracile DNA glycosylase 2 (UNG2), Domaine SAM et protéine contenant le domaine HD 1 (SAMHD1) et le complexe Humain Silencing Hub (HUSH). De plus, Vpr et Vpx sont nécessaires pour une réplication efficace du virus dans les cellules ne se divisant pas telles que les macrophages. Vpr est impliqué dans le transport des PIC vers le noyau. Il transactive la longue répétition terminale (LTR) virale et induit un arrêt du cycle cellulaire (Fukumori et al., 1998 ; Somasundaran et al., 2002 ; Yasuda et al., 2001 ; Zhao et al., 2011). Partageant des similitudes de séquences d'acides aminés avec Vpr, Vpx est également impliquée dans la translocation du PIC dans le noyau. Au cours des dernières décennies, une série d'études sur les protéines liées aux lentivirus ont révélé l'importance de Vpr et Vpx dans le cycle viral et dans l'antagonisme immunitaire inné.

Le domaine p6 du précurseur Pr55Gag est impliqué dans la phase tardive du cycle de réplication virale et permet l'incorporation de Vpr et Vpx dans les virions. La liaison de p6 à TSG101 et ALIX permet le recrutement du complexe ESCRT nécessaire à l'assemblage et au bourgeonnement de la particule virale (Friedrich et al., 2016 ; Garrus et al., 2001 ; Strack et al., 2003). L'analyse mutationnelle a mis en évidence que le motif ^{41}LXXLF^{45} dans la région C-terminale de HIV-1 p6 est requis pour l'incorporation de Vpr (Kondo et Göttlinger, 1996 ; Lu et al., 1995). Cependant, une étude contradictoire basée sur des analyses de mutation et de délétion montre que le domaine qui englobe le motif ^{15}FXFG^{18} à l'extrémité N-terminale, était suffisant pour l'incorporation de Vpr (Zhu et al., 2004), suggérant ainsi que le motif ^{41}LXXLF^{45} est superflu. Ces résultats sont contredits par une récente étude de mutagenèse (Ala scaning) qui montre l'importance des motifs ^{15}FXFG^{18} et ^{41}LXXLF^{45} pour le recrutement de Vpr du VIH-1 et que des mutations dans les résidus

^{34}ELY^{36} altèrent le recrutement de Vpr (Wanaguru et Bishop, 2021). En revanche, un motif ^{17}DXAXXLL^{23} spécifique de SIVmac p6 a été identifié comme étant requis pour l'incorporation de Vpx (Accola et al., 1999). En conclusion, aucun résidu consensus et/ou motif essentiel pour les interactions Vpr-p6 et Vpx-p6 n'a été identifié à ce jour.

Des analyses mutagènes sur Vpr ont montré que des résidus de 84 à 94 près de l'extrémité C-terminale sont nécessaires pour son incorporation dans les virions (Paxton et al., 1993). Par la suite, des résultats similaires ont montré que les délétions des résidus 73-96 ou 78-96 de Vpr entraînaient une altération de l'incorporation de Vpr dans les virions (Yao et al., 1995). Cependant, quatre mutants Vpr, L23F, E25K, A30F et I63F présentaient également une réduction de l'incorporation de Vpr. Comme les acides aminés L23, E25 et A30 sont tous inclus dans un domaine formant une hélice alpha amphipathique (résidus 16-34), cette hélice a été considérée comme importante pour l'incorporation de Vpr. La structure 3D de Vpr, résolue par la suite, a montré que les résidus L23, E25 et A30 sont distribués dans la première hélice de HIV-1 Vpr (Morellet et al., 2003 ; Schwefel et al., 2014 ; Wu et al., 2016), tandis que I63 est dans la troisième hélice. En revanche, les résidus 73-96 sont répartis dans la troisième hélice et dans l'extrémité C-terminale. Il semble que la première hélice, la troisième hélice et l'extrémité C-terminale de Vpr soient impliquées dans l'interaction avec p6 pour son incorporation. Ainsi, une étude détaillée de la structure du complexe est nécessaire. De même, les résidus 73-89 de HIV-2 Vpx se sont avérés importants pour l'incorporation de virions (Pancio et Ratner, 1998). Fait intéressant, une étude ultérieure utilisant les exxpériences d’alanine scaning a montré que la suppression des résidus 73-89 n'abolit pas l'incorporation du virion, mais que sa substitution réduisait significativement l'incorporation dans le virus (Jin et al., 2001). En général, il n'y a pas de données structurelles directes sur l'incorporation de Vpr et Vpx et une détermination détaillée et une investigation précise au niveau atomique sont nécessaires.

II. Objectifs de la these

Les protéines accessoires lentivirales Vpr et Vpx sont spécifiquement incorporées dans les virions par interaction directe avec p6^{gag} pour participer aux premiers stades de la

réplication virale. L'incorporation des protéines accessoires de primate Vpr/Vpx est essentielle pour la réplication virale dans les cellules au repos, la translocation du complexe de pré-intégration vers le noyau et pour la neutralisation des facteurs de restriction cellulaires antiviraux. Nos objectifs sont d'étudier la structure et les interactions de Vpr/Vpx avec leur partenaire p6 en présence de micelles DPC mimant l'environnement hydrophobe de la membrane. De plus, nous avons développé un protocole très simple destiné à remplacer les micelles DPC protonées par des micelles DPC deutériées de manière rentable. En effet, le DPC non deutéré utilisé pour solubiliser Vpr et Vpx empêche l'enregistrement de spectres RMN de bonne qualité. Nos objectifs sont divisés en quatre parties : (1) optimiser le protocole pour l'élimination des micelles DPC entièrement protonées tout en maintenant la structure native de la protéine, (2) élucider les structures de HIV-1 Vpr, SIVmac Vpx et de leurs partenaires p6 en présence de micelles DPC, (3) caractériser au niveau atomique les interactions VIH-1 Vpr-p6 et SIVmac Vpx-p6, (4) tester les interactions croisées entre SIVmac Vpx et VIH-1 p6.

III. Méthodes

Expression et purification de HIV-1 Vpr et SIVmac Vpx.

Pour l'expression et la purification de HIV-1 Vpr et SIVmac Vpx, onze colonies simples de pET-11d-SIVmac Vpx et quatre colonies simples de pET-11d-HIV-1 Vpr LB ont été précultivées individuellement dans 5 mL de Milieu M9a à 37°C pendant la nuit. Un volume approprié de préculture a été dilué à une DO_{600} de 0.3 dans 10 mL de milieu M9a modifié, puis inoculé à 37 °C pendant 60 min jusqu'à ce que la valeur DO_{600} de la nouvelle culture atteigne 0.6. La culture a été récoltée par centrifugation et les culots ont été remis en suspension avec 10 mL de milieu M9b modifié. L'IPTG a été ajouté à une concentration finale de 0.5 mM et la culture a été incubée à 18 °C pendant 20 heures. Les niveaux d'expression ont été évalués par SDS-PAGE. Les meilleures cellules avec un niveau d'expression plus élevé ont été stockées dans une solution de glycérol à -80 °C et utilisées pour l'expression à grande échelle.

Pour l'expression des protéines marquées $^{15}N/^{13}C$, 100 µL de stock de glycérol d'*E. coli* Rosetta 2 (DE3)-pET11d SIVmac Vpx et pET-11d-HIV-1 Vpr ont été inoculés dans des

erlenmeyers en verre de 200 mL avec 50 mL de milieu M9a modifié et 100 µg/mL d'ampicilline. Les précultures ont été cultivées une nuit à 37 °C. Lorsque la DO_{600} a atteint une valeur de 0.3, les précultures ont été chacune transférées dans un erlenmeyer en verre de 2.5 L contenant 1.0 L de milieu M9a modifié. Lorsque les valeurs de DO_{600} des cultures ont atteint 0.6, les cellules ont été centrifugées à 6000 g pendant 30 min à 18 °C et les culots remis en suspension dans 1.0 L du deuxième milieu M9b modifié contenant seulement 1 g de ^{13}C-glucose et 0.5 g de $^{15}NH_4Cl$. L'expression a été induite en ajoutant l'IPTG à une concentration finale de 0.5 mM et les cellules cultivées pendant 20 h à 18 °C puis récoltées. Les protéines non marquées ont été produites en utilisant les mêmes méthodes en présence de 1 g de glucose non marqué et 0.5 g de NH_4Cl. Après purification, SIVmac Vpx et HIV-1 Vpr ont été obtenus en grande quantité.

Synthèse chimique de HIV-1 p6, SIVmac p6 et du peptide SIVmac p6$^{16-28}$.

Les peptides SIVmac p6 et HIV-1 p6 ont été synthétisés chimiquement par Rodrigue Marquant, notre collaborateur du LBM, CNRS, UMR 7203, Sorbonne Université. Les SIVmac p6 et HIV-1 p6 non marqués ont été synthétisés comme décrit précédemment. En bref, les HIV-1 p6 et SIVmac p6 complets ont été produits individuellement par synthèse automatisée en phase solide en utilisant la stratégie Fmoc. La purification a été effectuée par chromatographie liquide à haute performance en phase inverse (HPLC) dans un mélange H_2O/0.05 % d'acide trifluoroacétique (TFA) et d'acétonitrile en utilisant des procédures déjà rapportées pour produire d'autres protéines rétrovirales. Un gradient linéaire de 0% à 100% d'acétonitrile a été réalisé pour optimiser la purification et les peptides ont finalement été analysés par spectroscopie de masse. Pour la synthèse de HIV-1 p6 marqué, six acides aminés marqués (95% ^{15}N, 15% ^{13}C) ont été introduits dans les deux motifs ^{15}FXFG^{18} et ^{41}LXXLF^{45} de HIV-1 p6 aux positions F15, F17, S40, L41, L45 et F46. De même, sept acides aminés marqués (95 % ^{15}N, 15 % ^{13}C) ont été introduits dans le motif ^{17}DXAXXLL^{23} de SIVmac p6 aux positions D17, P18, A19, V20, D21, L22 et L23. En outre, trois acides aminés marqués (95% ^{15}N, 15% ^{13}C) supplémentaires ont été introduits dans SIVmac p6 au niveau de la région C-terminale de la première hélice aux positions L55, L58 et F59.

Etudes des structures protéiques et de leurs interactions par RMN.

Cela fait plus de 50 ans que la RMN en solution est utilisée pour l'étude des macromolécules biologiques. La première structure en solution de l'inhibiteur de la protéinase IIA du plasma séminal de taureau a été résolue par RMN en solution en 1985, ce qui a ouvert la voie à l'analyse RMN des macromolécules biologiques. La RMN, ainsi que la cristallographie aux rayons X et la cryomicroscopie électronique (cryo-EM), sont les principales techniques de recherche pour la biologie structurale des complexes protéiques. Par ailleurs, la RMN a la capacité unique de fournir des informations sur la structure et la dynamique des protéines et permet la détermination structurale en solution, aux conditions physiologiques mais elle est limitée par le poids moléculaire des protéines. La détermination de la structure des macromolécules par RMN est généralement divisée en quatre étapes principales comprenant (1) la préparation de l'échantillon, (2) les mesures RMN, (3) les attributions des résonances et (4) la détermination de la structure à partir des données RMN.

Pour les mesures par RMN, le spectre RMN ^{1}H 1D est encombré en raison d'un grand nombre de signaux de protons. Ainsi, les spectres bidimensionnels et tridimensionnels sont réalisés avec des protéines marquées au ^{15}N et/ou au ^{13}C pour surmonter les superpositions de déplacements chimiques. Jusqu'à présent, la méthode la plus simple d'attribution de protéines implique les expériences d'attribution de squelette et de chaîne latérale. Les expériences d'attribution de squelette sont les séquences CBCANNH, CBCA(CO)NNH, HNCA, HN(CO)CA, HNCO et HN(CA)CO, tandis que les attributions de chaînes latérales sont obtenues par les expériences CC(CO)NNH, HBHA(CO)ONH et H(CCCO)NNH. L'attribution du squelette des protéines marquées $^{15}N/^{13}C$ est principalement basée sur les expériences CBCANNH et CBCA(CO)NNH. L'expérience CBCANNH permet de corréler chaque groupement amide d'un résidu avec les déplacements chimiques de ses propres Cα et Cβ et ceux du résidu précédent. De plus, l'expérience CBCA(CO)NNH ne fait que corréler le groupement amide d'un résidu avec les déplacements chimiques Cα et Cβ du précédent. Les expériences HNCA et HN(CO)CA donnent accès aux déplacements chimiques des CO. Enfin, la spectroscopie à effet Overhauser nucléaire (NOESY) permet la détermination de contraintes de distance pour les calculs de structure.

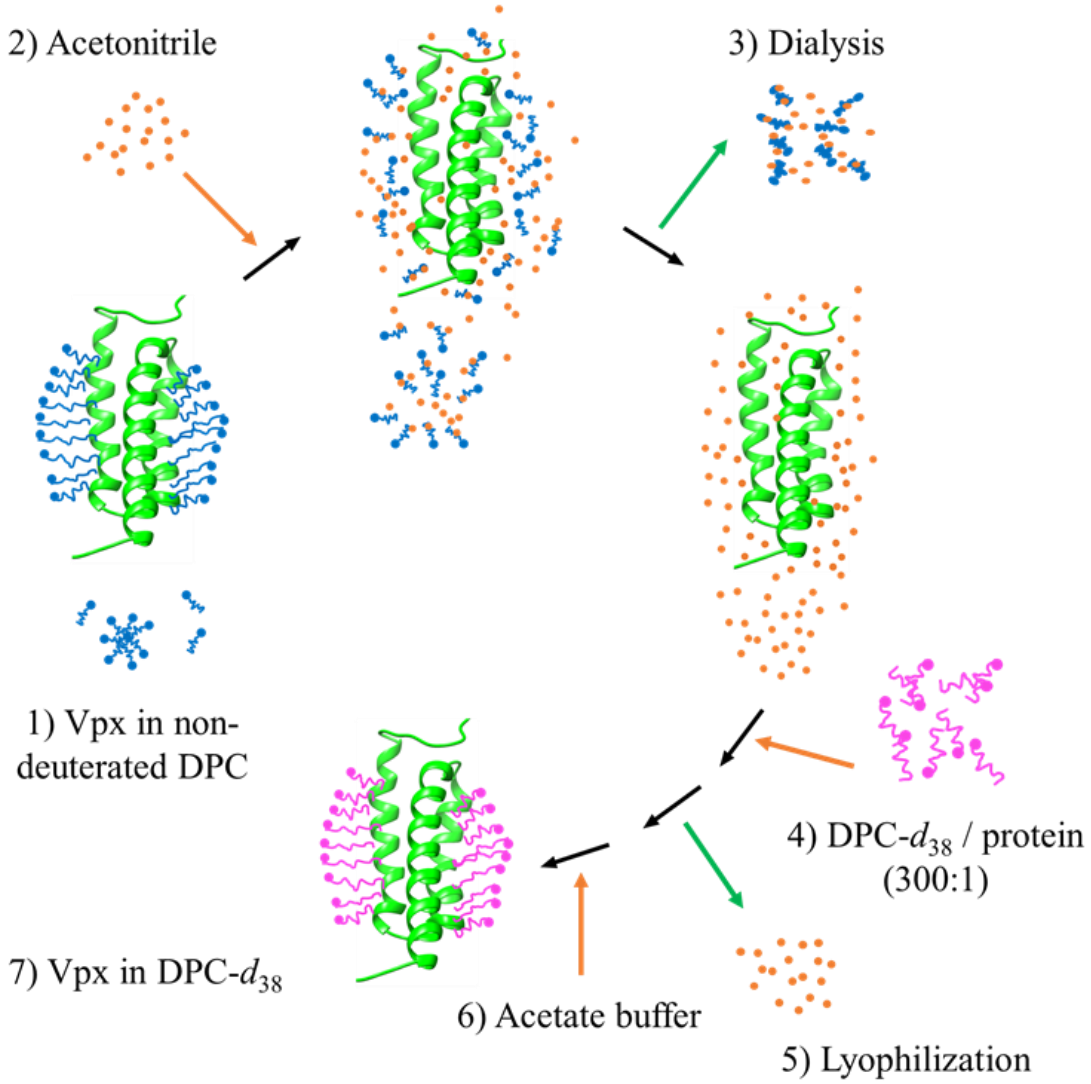
IV. Résultats

Expression et purification de HIV-1 Vpr et SIVmac Vpx.

Les protéines VIH-1 Vpr et SIVmac Vpx ont été exprimées dans des cellules *E. coli* Rosetta 2 (DE3) et la production induite par 0.5 mM d'IPTG à 18 °C pendant 20 heures. Les culots cellulaires ont été remis en suspension dans 20 mM de phosphate de sodium pH 7,4, 200 mM de NaCl, 5 mM de BME et 6 mM de DPC et rompus par sonication comme décrit précédemment. Le surnageant a été collecté à 15000 g à 4 °C pendant 30 min et la protéine dans le surnageant a été purifiée par colonne Ni^{2+}-NTA. Le contenu des fractions a été analysé par SDS-PAGE.

Remplacement du DPC non deutéré par du DPC deutéré pour l'étude RMN.

Cette partie décrit un nouveau protocole simple pour le remplacement indirect du DPC non deutéré par du DPC deutéré (DPC-d_{38}). Les signaux de protons DPC entrainent l'enregistrement de spectres RMN de mauvaise qualité. Ainsi, nous avons développé et organisé un protocole qui a été publié dans le journal Protein Science. Ce protocole assure l'élimination totale des lipides tout en conservant les propriétés structurales de la protéine hydrophobe sur la base de notre hypothèse selon laquelle les molécules de détergent recouvrant la protéine peuvent être facilement et totalement remplacées par des molécules de solvant organique pendant la dialyse selon le schéma ci-dessous.



Protocole de remplacement du DPC non deutéré par le DPC-d_{38} deutéré.

Analyse RMN des structures de Vpr et Vpx en présence de micelles de DPC.

Vpr marquée $^{15}N/^{13}C$ a été concentré dans 10 mM d'acétate de sodium, pH 4, 50 mM de NaCl, 5 mM de β-mercaptoéthanol, 90 mM de DPC-d_{38} à une concentration finale de 0.3 mM. Nous avons réalisé toutes les expériences d'attribution du squelette et des chaînes latérales sur un spectromètre 950 MHz équipé d'une cryosonde à 323 K au CNRS, à Gif sur Yvette. Topspin a été utilisé pour traiter les expériences et CcpNMR pour l'analyse des spectres et l'attribution de Vpr marquée $^{15}N/^{13}C$.

SIVmac Vpx marquée $^{15}N/^{13}C$ a été concentrée dans 10 mM d'acétate de sodium, pH 4, 50 mM de NaCl, 5 mM de β-mercaptoéthanol, 280 mM de DPC à la concentration finale de 0.7 mM. Les expériences d'attribution du squelette et des chaînes latérales ont été enregistrées sur un spectromètre AVANCE 950 MHz équipé d'une cryosonde. L'acquisition et le traitement des données ont été obtenus avec Topspin et CcpNMR a été utilisé pour l'analyse et l'attribution des spectres. Tous les résidus ont été attribués à l'exception des prolines et les quatre résidus M1, E54, W56 et H82.

La structure de solution de SIVmac Vpx à 323 K a été déterminée et calculée à l'aide de XPLOR-NIH sur la base des contraintes NOE et des contraintes d'angle dièdre dérivées de TALOS. Un ensemble des dix structures de plus faible énergie de SIVmac Vpx a été sélectionné pour l'analyse structurale. Ces structures montrent un RMSD par paires sur les résidus 23-90 de 0.55 Å pour les atomes du squelette et de 1.12 Å pour les atomes lourds, respectivement. Cinq contraintes longue distance ont été obtenues, mais peu de NOE sont disponibles. En comparaison avec la structure cristalline de SIVmac Vpx, la structure obtenue en solution a montré un repliement du squelette similaire.

RMN des protéines p6 du VIH-1 et du SIVmac en présence de micelles de DPC.

Notre groupe a découvert précédemment que HIV-1 p6 présente une affinité élevée pour les bicouches membranaires, ce qui améliore considérablement son interaction avec Vpr. Les fonctions identifiées de HIV-1 p6 se produisent généralement dans des conditions hydrophobes à proximité de la membrane cytoplasmique. Ici, les micelles DPC-d_{38} ont été utilisées pour l'étude de HIV-1 p6 dissous dans 10 mM d'acétate de sodium, pH 4, 50 mM de NaCl, 5 mM de β-mercaptoéthanol, 100 mM de DPC-d_{38} à une concentration finale de 50 μM. La dispersion spectrale des résonances NH est meilleure en présence de micelles DPC-d_{38} qu'en leur absence, suggérant que HIV-1 p6 est bien structuré en présence de DPC-d_{38}.

L'attribution du squelette et des chaînes latérales de p6 a été complétée par 2D NOESY et 2D TOCSY et la structure en solution de p6 a été déterminée à l'aide du programme ARIA sous contraintes de distance RMN et des angles dièdres dérivés de DANGLE dans CcpNmr. Les superpositions des dix structures de plus faible énergie sur chacune des hélices, montrent que l'orientation de chaque hélice par rapport aux deux autres est très variable,

même si chaque hélice présente une bonne convergence entre les modèles calculés, avec un RMSD respectif de 1.40 Å, 0.48 Å et 0.73 Å sur les atomes du squelette.

De même, la structure du SIVmac p6 complet a été étudiée par RMN dans 90 % de tampon acétate (10 mM d'acétate de sodium, pH 4, 50 mM de NaCl, 5 mM de β-mercaptoéthanol, 100 mM de DPC-d_{38}) et 10 % de D_2O à la concentration finale de 1 mM dans un tube RMN de 3 mm. Les attributions ont été effectuées sur la base des pics croisés dans les spectres bidimensionnels ^{1}H NOESY avec de courtes distances entre HN, Hα et Hβ de l'acide aminé (i+1) et HN Hα, et Hβ de l'acide aminé (i) et les spectres TOCSY. La structure en solution de SIVmac p6 a été déterminée à l'aide du programme ARIA sous contraintes de distances RMN expérimentales et d'angles dièdres dérivés de DANGLE dans CcpNmr. La structure de plus faible énergie présente deux longues hélices aux positions 18-39 et 45-59. L'orientation relative variable des deux hélices est essentiellement due à l'absence de NOE longue distance. Les dix structures de plus faible énergie ont été superposées soit à sur la première hélice (RMSD de 1.88 Å) soit sur la seconde (RMSD de 1.39 Å).

Analyse RMN de l'interaction entre Vpr/p6 du VIH-1 et Vpx/p6 du SIVmac.

Pour cartographier les acides aminés impliqués dans l'interaction entre Vpr et p6 du VIH-1, Vpr marqué ^{15}N/^{13}C a été fixé à 300 µM avec 90 mM de DPC-d_{38}. Les perturbations de déplacements chimiques (CSP) ont été mesurés sur des spectres ^{1}H-^{15}N HMQC lors de l'ajout de VIH-1 p6 non marqué jusqu'à un rapport Vpr/p6 de 1:4. Les résidus dont les CSP sont supérieurs à la moyenne + 0.5 SD sont concentrés sur deux régions principales I et II. La région I est formée par les résidus V31, I37, S41, L42, Y50, D52, I63 tandis que la région II contient I74, G75, C76, R77, I81, G82, I84 et Q85. Les résidus E17, I37, Y50, D52, I63, Q66 et S96 sont significativement perturbés par la présence de HIV-1 p6. Les changements de déplacement chimique des résidus E17 et S96 résultent probablement de l'effet indirect de la flexibilité des extrémités N- et C-terminales.

Nous avons effectué l'étude de cartographie inverse dans le même tampon en utilisant le VIH-1 p6 marqué au ^{15}N et la protéine VIH-1 Vpr non marquée. Les analyses de délétion et de mutation impliquant les résidus 37-52 dans HIV-1 p6 ont révélé que le motif ^{41}LXXLF^{45} est requis pour l'incorporation de HIV-1 Vpr. Cependant, les résidus 1-23 comprenant le motif ^{15}FXFG^{18} ont été identifiés comme étant suffisants pour

l'incorporation de HIV-1 Vpr, sans la présence du motif ^{41}LXXLF^{45}. Six résidus conservés marqués au ^{15}N (F15, F17, S40, L41, L44 et F45) ont été introduits pendant la synthèse chimique de p6 total pour déterminer et confirmer les sites de liaison de Vpr sur p6 en présence de micelles de DPC. Nos résultats ont révélé que les six résidus marqués ont subi des changements de déplacements chimiques lors de la liaison de HIV-1 Vpr. De plus, nous avons remarqué que L44 et F45 Hα, F15, F17, S40, et L44 Hβ, L44 Hγ et L41 Hδ présentent également des perturbations des déplacements chimiques. De même, les expériences RMN de CSP de SIVmac Vpx marqué au ^{15}N, ^{13}C ont ensuite été réalisées avec SIVmac p6 non marqué par ^{1}H-^{15}N SOFAST-HMQC.

L'attribution séquentielle a été complétée réalisée en combinant les deux expériences NOESY tridimensionnelles de la forme libre du SIVmac Vpx marqué ^{15}N / ^{13}C et du complexe avec le SIVmac p6 complet. Les déplacements chimiques des protons amides ont été mesurés sur des spectres HMQC ^{1}H-^{15}N. Deux régions principales avec un CSP supérieur à la moyenne + 0.5 SD caractérisent les régions I et II d'interaction basées respectivement sur les structures RMN et cristallines. La région I est formée par les résidus V29, E43 et W49, tandis que la région II est formée par R7, L25, M62, Q76 et F80. Les résidus D3 et A112 des extrémités N- et C-terminales ont une variation de déplacement chimique résultant probablement d'effets indirects dus à leur flexibilité.

L'étude de la liaison inverse de SIVmac p6 a été réalisée en utilisant SIVmac p6 sélectivement marqué. Le SIVmac p6 sélectivement marqué subit des CSP plus grands conformément à ce qui a été observé avec les expériences inverses. Fait intéressant, les résonances des arginines R36 et R38 situées près de la fin de la première hélice présentent des variations plus appréciables. De plus, le proton amide de la chaîne latérale de R38 subi une importante variation de déplacement chimique.

V. Discussion

La cytotoxicité et l'hydrophobicité de HIV-1 Vpr et SIVmac Vpx entravent la production et la purification de ces protéines. Des études antérieures ont utilisé deux constructions, NusA-Vpr et His_6-MBP-intein1-Vpr-intein2-Cyt b5-His6 pour l'expression de Vpr (Hänel et al., 2014 ; Wu et al., 2016). Au début, nous avons essayé de produire du VIH-1 Vpr

marqué au ^{15}N, ^{13}C en utilisant ces constructions, mais nous n'avons pas réussi en raison du milieu de culture différent. Cependant, l'utilisation de bactéries *E. coli* Rosetta 2 (DE3) et de micelles DPC a permis la production et l'extraction de ces protéines. Les cellules Rosetta 2 (DE3) peuvent supporter la cytotoxicité des protéines et les micelles DPC peuvent être utilisées comme détergent pour surmonter les problèmes d'hydrophobicité lors des étapes d'extraction et de purification. Dans ce travail, nous avons modifié la composition du milieu M9+ (Cai et al., 2016) utilisé pour l'induction et l'expression de la protéine deutérée. Le milieu riche en préculture a été remplacé par notre milieu modifié M9a. De plus, un deuxième milieu M9b modifié a été utilisé pour l'induction et l'expression de protéines enrichies isotopiquement. Enfin, nous avons produit HIV-1 Vpr et SIVmac Vpx marquées ^{13}C et ^{15}N par des cellules Rosetta 2 (DE3) et les avons purifiées en présence de micelles DPC. Cependant, le grand nombre de protons issus des micelles DPC empêche l'enregistrement de spectres RMN avec un rapport signal/bruit acceptable, compromettant ainsi les analyses ultérieures telles que la détermination de la structure et les études d'interaction avec des partenaires protéiques ou ligands pour l'identification de nouveaux médicaments et cibles. Dans notre protocole, l'acétonitrile permet un échange efficace de détergents lipidiques non deutérés par des détergents deutérés, tout en conservant la structure 3D de la protéine. Il est alors possible d'enregistrer des spectres RMN à haute résolution, exempts de signaux protons interférents provenant de détergents, nécessaires aux études RMN de protéines isolées ou en interaction avec leurs partenaires.

De plus, la dispersion spectrale des résonances des protéines VIH-1 Vpr et SIVmac Vpx dans le DPC n'est pas optimale. Pour obtenir des spectres de bonne qualité et bien éclatés, un nouveau plasmide dans lequel le gène Vpr est fusionné à celui de la protéine de liaison au maltose (MBP) est conçu pour augmenter la solubilité de Vpr. Ainsi, la protéine recombinante est soluble en l'absence de détergents. De plus, les micelles DPC pourraient ne pas être le meilleur détergent pour Vpr ou Vpx. À l'avenir, nous effectuerons un criblage de détergents tels que la 1,2-dihexanoyl-sn-glycéro-3-phosphocholine (DHPC) et la 1,2-dioléoyl-sn-glycéro-3-phosphocholine (DOPC) pour obtenir le meilleur détergent pour Vpx et Vpr et des spectres RMN bien dispersés. En outre, plusieurs mélanges de détergents seront également testés pour obtenir des spectres RMN bien dispersés et de meilleure

qualité, ce qui sera utile et important pour obtenir des spectres ^{13}C- et ^{15}N-NOE de haute qualité pour en extraire les contraintes de distance nécessaires au calcul de la structure.

La structure de SIVmac p6 n'a jamais été rapportée, ni dans sa forme mature, ni au sein de Gag. La structure du VIH-1 p6 n'a été étudiée que dans des conditions de solvant organique (Fossen et al., 2005). Dans cette étude, nous avons caractérisé par RMN les structures 3D de HIV-1 et SIVmac p6 en présence de micelles DPC et fourni des détails structurels de la reconnaissance de Vpr et Vpx par p6. Par rapport aux structures précédentes du VIH-1 p6 obtenues en solution aqueuse de trifluoroéthanol, nos résultats révèlent l'existence d'une hélice courte supplémentaire (résidus 24-27) en présence de micelles DPC-d_{38} suggérant que le VIH-1 p6 adopte une conformation différente dans des conditions différentes. Pour l'interaction entre Vpr et p6 du VIH-1, nos résultats affichent deux régions importantes : les résidus V31, I37, S41, L42, Y50, D52, I63 et Q66 (région I), et I74, G75, C76, R77, I81, G82, I84 et Q85 (région II). La région I correspond à l'hélice II, tandis que la région II correspond à l'hélice III. Ceci est cohérent avec des études antérieures qui ont identifié des résidus importants 73-96 dans HIV-1 Vpr (Paxton et al., 1993). Cependant, d'autres études ont montré que les mutants Vpr du VIH-1 : L23F, E25K, A30F et I63F altèrent son incorporation dans le virion par p6 (Yao et al., 1995). Les trois mutants (L23F, E25K et A30F) situés dans l'hélice I de HIV-1 Vpr réduisent également l'incorporation de la protéine ce qui suggère que cette région participe à l'incorporation dans le virion. Les résidus L23 et E25 sont loin des deux régions identifiées par RMN. La structure RMN en solution de HIV-1 Vpr indique que les résidus L23 et E25 sont impliqués dans les contraintes moyennes et que les résidus A30 Hβ* ont des contraintes à longue distance avec L39 Hδ*, L64 Hδ* et L68 Hδ* par des NOE basés sur des liaisons hydrogène, ce qui semble important pour le maintien de la structure Vpr du VIH-1 (Morellet et al., 2003). Ainsi, les mutants Vpr du VIH-1 L23F, E25K et A30F peuvent perdre les contraintes de liaison hydrogène et sa structure perturbée et conduire à altérer l'interaction avec p6. Aucune région de SIVmac Vpx n'a encore été identifiée comme étant responsable de l'interaction avec SIVmac p6. Cependant, HIV-2 Vpx partage presque la même séquence avec SIVmac Vpx (79.1 % d'identité et 86.4 % de similarité). Les résidus 73-89 dans HIV-2 Vpx ont été identifiés comme étant importants pour l'incorporation de Vpx (Pancio et Ratner, 1998) mais non essentiels, ce qui suggère que d'autres résidus peuvent être

impliqués. Dans nos études, deux régions I (V29, E43 et W49) et II (R7, L25, M62, Q76 et F80) dans Vpx, ont été identifiées pour interagir avec p6. Les résidus Q76 et F80 sont contenus dans le domaine 73-89 de HIV-2 Vpx.

En résumé, en utilisant des expériences de perturbations du déplacements chimiques, nous avons montré que Vpr et Vpx interagissent avec p6. Les régions I (V31, I37, S41, L42, Y50, D52, I63 et Q66) et II (I74, G75, C76, R77, I81, G82, I84 et Q85) de HIV-1 Vpr interagissent avec deux motifs de VIH-1 p6, ^{15}FXFG^{18} et ^{41}LXXLF^{45}. De même, les régions I (V29, E43 et W49) et II (R7, L25, M62, Q76 et F80) de SIVmac Vpx ciblent principalement les résidus K36, R38 et le motif ^{17}DXAXXLL^{23} de SIVmac p6. À l'avenir, nous exprimerons un VIH-1 p6 et un SIVmac p6 entièrement marqués ^{15}N/^{13}C par le système *E. coli*, ce qui sera utile pour une analyse plus approfondie des propriétés structurales de p6 et de ses interactions avec Vpr ou Vpx.

List of Figures and Tables

capsid then penetrates the channel at step III. Subsequently, the intact HIV-1 capsid translocates into the nucleus at step IV. After the departure from the NPC central channel, CPSF6 interacts with the capsid, and then the capsid uncoats and releases the genome at step V. At steps V and VI, the viral genome is released and integrated into the host genome.

Figure 11. New HIV-1 replication cycle.

Figure 12. Cartoon representation of 15 HIV proteins and their functional domains. The structure of ASP protein is not known, and it is not there.

Figure 13. The networks of HIV protein-protein interaction. Green arrows, well-known interactions cited more than 300 times or reported by more than 100 citations at least 3 publications; gray arrows, little-known interactions reported with fewer than 100 citations at a single paper; blue arrows, lesser-known interactions reported by fewer citation.

Figure 14. Summary of HIV protein-protein associations. Fifteen HIV-1 proteins are displayed at the left and top, and protein functions are highlighted in different colors. The interaction between Vpr and p6 was shown in red squire.

Figure 15. Models of HIV-1 Gag.

Figure 16. Molecular model of HIV-1 Gag assembly.

Figure 17. The HIV-1 capsid. Each part is shown in different colors. Orange, β-hairpin; light green, helix 1; turquoise, helix 2; dark blue, helix 3; purple, helix 4; gray, cyclophilin A-binding loop; pink, helix 5; rose, helix 6; red, helix 7; dark red, helix 8; salmon, helix 9; blue, helix 10; and green, helix 11.

Figure 18. Structures of HIV-1 Vpr. (A) Solution structure of the chemically synthesized HIV-1 Vpr determined by NMR (PDB:1M8L). (B) Crystal structure of HIV -1Vpr extracted from the DDB1-DCAF1-Vpr-UNG2 complex (PDB:5JK7). (C) Superposition of solution NMR (A) and the crystal (B) structures of HIV-1 Vpr.

Figure 19. Different models of HIV-1 Vpr. The parallel dimer model (A), the antiparallel dimer model (B), Hexamer model top view (C), and hexamer model side view (D).

Figure 20. Structures of SIVmac Vpx and HIV-1 Vpr. (A) Crystal structure of SIVmac Vpx (4CC9). Vpx is made up of a three-helical bundle and stabilized by a zinc finger motif in blue. Information about the structure at residues 91-99 is missing. (B) Superposition of structures of SIVmac Vpx and HIV-1 Vpr.

Figure 21. Important functional regions or sites of HIV-1 p6 adapted from [141]. The highly conserved residue S40 in red is important for viral replication. ERK-2 phosphorylation sites, ubiquitin and SUMO-1 attachment sites, and binding domains for TSG101, Vpr, and Alix are shonwn in grey.

Figure 22. Solution structure of HIV-1 p6 (PDB:2C55). Blue, helix I with residues 14-21; red, helix II with residues 33-47.

Figure 23. Comparison of HIV-1 immature and mature virions.

Figure 24. The pathway of ESCRT recruitment.

Figure 25. The pathway of TRIM5α in retrovirus infection.

Figure 26. Structure of APOBEC3 (A3G) (PDB 4N9F).

Figure 27. Polyubiquitination and subsequent degradation of APOBEC3 mediated by Vif.

Figure 28. Schematic representation of Tetherin with the transmembrane domain (TMD), a helical extracellular domain (ECD) and glycophosphatidylinositol (GPI). Tetherin is characterized by a short cytoplasmic N-terminus, then followed by an α-helical single-pass TM domain, an extended coiled-coil extracellular domain, and a C-terminal GPI anchor in the plasma membrane. For disulfide-bond formation, two N-glycosylation sites and three cysteine residues are noted.

Figure 29. Degradation of tetherin by Vpu.

Figure 30. Degradation of SAMHD1 by Vpx.

Figure 31. Pictures of Kaempferia pulchra (A) and Picrasama javanica (B).

Figure 32. High-throughput screening of Vpr inhibitors. A, a workflow of the screening process. B, inhibitor 3 shown in arrow. C, growth curves of hit compound 1, compound 3, and DMSO. D, western blotting assays of Vpr expression. Lane D, DMSO; lane 1, compound 1; lane 3, compound 3 (87 μM), lane L, leucine; lane Dp, 4,6-diamidino-2-phenyl-indole (DAPI). E, structure of the hit compound 3.

Figure 33. Backbone and sidechain assignment experiments measured by NMR.

Figure 34. Backbone assignment based CBCANNH and CBCA(CO)NNH

Figure 35. Backbone assignment based HNCA and HN(CO)CA experiments.

Figure 36. Structure calculation flow chart.

Figure 37. Chemical shift perturbation (CSP) (A) and NMR titration experiments (B).

Figure 38. The dissociation constant (K_d).

Figure 39. Level of expression of HIV-1 Vpr by SDS-PAGE. The expression of HIV-1 Vpr (~12-14 kDa) was highlighted in red squares.

Figure 40. Purification of the HIV-1 Vpr protein by Ni^{2+} binding affinity. The target protein is highlighted in red squares. A band at position ~ 24 kDa may be Vpr dimer.

Figure 41. SIVmac Vpx expression levels.

Figure 42. Purification of the SIVmac Vpx protein by Ni^{2+} binding affinity (A and B) and anion exchange chromatography (C and D). The protein is eluted at peaks in red arrows. The molecular weitht of eluted SIVmac Vpr is ~14 kDa in red square.

Figure 43. A flowchart of nondeuterated DPC indirectly replaced by deuterated DPC-d_{38}. Binding of DPC detergent favors and maintains SIVmac Vpx conformation and prevents aggregation. The interaction of SIVmac Vpx with DPC molecules was disrupted by acetonitrile and DPC was then present in free form. The original conformation of the protein was conserved under the hydrophobic environment. Prior to lyophilization, the deuterated DPC-d_{38} was supplied to maintain and protect SIVmac Vpx protein. Finally, the lyophilized sample was dissolved in acetate buffer for NMR studies.

Figure 44. Efficiency of DPC removal followed by NMR. All spectra were recorded in acetate buffer A at 323 K and pH 4.0. (A) 1D proton spectrum of DPC (30 mM). The seven proton resonances (labelled I to VII) are from DPC. (B) 0.3 mM SIVmac Vpx proton spectrum with 100 equivalents protonated DPC. (C) 0.3 mM SIVmac Vpx with 100 equivalents DPC-d_{38} after lyophilization and resuspension according to the protocol described here. All DPC proton signals were suppressed allowing proper observation on the protein spectrum.

Figure 45. ^{1}H-^{15}N HMQC spectrum of 0.3 mM ^{13}C, ^{15}N-labeled SIVmac Vpx in 30% acetonitrile recorded at 323 K and 600 MHz.

Figure 46. ^{1}H-^{15}N SOFAST-HMQC spectra at different SIVmac Vpx / DPC-d_{38} molar ratios. Spectra were recorded in 3 mm NMR tube at 323 K on Bruker Avance 600 MHz spectrometer equipped with cryogenic triple resonance probe for molar ratios of 1:100 (A), 1:200 (B), 1:300 (C), 1: 400 (D), and 1:500 (E). (F) Superimposition of HSQC spectra of SIVmac Vpx in the presence of DPC-d_{38} at different ratios.

Figure 47. Appearance of SIVmac Vpx lyophilizate in the absence and in the presence of DPC-d_{38}. (A) SIVmac Vpx lyophilized without detergents. SIVmac Vpx in the 30% acetonitrile aqueous was directly lyophilized. The protein aggregated and the powder is attached on the wall of the eppendorf. (B) The lyophilized SIVmac Vpx with DPC-d_{38}. Prior to lyophilization, the deuterated DPC-d_{38} was added. The lyophilized SIVmac Vpx powder in the presence of DPC-d_{38} was porous and fibrous.

Figure 48. 2D ^{1}H-^{15}N SOFAST-HMQC spectra of 300 μM HIV-1 Vpr in 10 mM sodium acetate, pH 4, 50 mM NaCl, and 5 mM β-mercaptoethanol at 323 K.

Figure 49. Assignment of HIV-1 Vpr. The sequence of HIV-1 Vpr is at the top. The residues in red were not assigned.

Figure 50. Secondary structure prediction of HIV-1 Vpr. The secondary chemical shift of Hα, Cα, Cβ, and Co were shown by DANGLE predicted secondary structure.

Figure 51. 2D ^{1}H-^{15}N SOFAST-HMQC spectra of 700 μM SIVmac Vpx in 10 mM sodium acetate, pH 4, 50 mM NaCl, and 5 mM β-mercaptoethanol at 323 K.

Figure 52. Assignment of SIVmac Vpx. (A) The sequence of SIVmac Vpx is at the top. The residues in red were not assigned. (B) The assignment of SIVmac Vpx by 2D SOFAST-HMQC spectra in the presence of DPC micelles. The square is enlarged in (C).

Figure 53. Prediction of secondary structures of SIVmac Vpx protein based on TALOS-N analysis.

Figure 54. Structures of SIVmac Vpx. (A) Superposition of the tenlowest energy structures. (B) Superposition of the crystal (light blue, PDB 4CC9) and solution NMR (grey) structures of SIVmac Vpx.

Figure 55. Spectra of HIV-1 p6 in the absence and presence of DPC micelles. (A) The sequence of the full-length HIV-1 p6. (B) 1D spectra of HIV-1 p6 in the absence (Red) and presence of DPC micelles (Blue). (C) 2D ^{1}H-^{15}N SOFAST-HMQC spectraof HIV-1 p6 in the absence (Red) and presence of DPC micelles (Blue). These experiments were recorded on a 600 MHz spectrometer equipped with cryoprobe at 323 K.

Figure 56. Determination of the selectively labeled HIV-1 p6 with the addition of DPC-d_{38}. (A) Overlap of ^{1}H-^{15}N SOFAST-HMQC spectra of HIV-1 p6 in complex with DPC-d_{38} at molar ratio of 1:100 (blue), 1:200 (magenta), 1:300 (cyan), 1:400 (green), and 1:450 (yellow), respectively. (B) Variations of chemical shift perturbations (CSP) of three residues (L41, L44, and F45).

Figure 57. HIV-1 p6 spectra in acetate buffer with 100 mM DPC-d_{38}. (A) Superposition of 2D-NOESY (blue) and 2D-TOCSY (red) spectra recorded at 600 MHz and 323 K. The

sequential NOEs connectivities [dαN (i, i + 1)] and [dN (i, i + 1)] are outlined. (B) Distance restraints chart.

Figure 58. Superposition of ten lowest energy structures of HIV-1 p6 determined using ARIA, showing the superimposition on each alpha-helix respectively on residues 11-19 (A), 24-27 (B) and 35-45 (C) in the presence of DPC-d_{38} micelles.

Figure 59. Solution structure of HIV-1 p6 in the presence of DPC-d_{38} micelles. Red, helix I with residues 11-19; blue, helix II with residues 24-27; cyan, helix III with residues 35-45.

Figure 60. Evolution of the 2D ^{1}H-^{15}N SOFAST-HMQC natural abundance spectra of the peptide SIVmac p6$^{16-28}$ at different time after dissolution.

Figure 61. Spectra of the peptide SIVmac p6$^{16-28}$ in the absence and presence of DPC micelles. Sequence of SIVmac p6$^{16-28}$ peptide is at the top. Natural abundance 2D spectra ofSIVmac p6$^{16-28}$ peptide in the absence (Red) and presence (Blue) of DPC micelles. These experiments were recorded on a 600 MHz spectrometer equipped with cryoprobe at 323 K.

Figure 62. SIVmac p6$^{16-28}$ spectra in acetate buffer with 100 mM DPC-d_{38}. (A) Superposition of 2D-NOESY (blue) and 2D-TOCSY (red) spectra recorded at 600 MHz and 323 K. The sequential NOEs connectivities [dαN (i, i + 1)] and [dN (i, i + 1)] are outlined. (B) Distance restraints chart.

Figure 63. Structural properties of the peptide SIVmac p6$^{16-28}$. (A) Superposition of the ten SIVmac p6$^{16-28}$ NMR structures of lowest energy (left). The side chains of the helix were shown (right). (B) Helical wheel representation of the helix 18-26 in the fragment SIVmac p6$^{16-28}$ by Pepwheel. The hydrophobic and hydrophilic residues were respectively boxed in red and blue. The color coding reflects the hydrophobic properties: orange, strong; light orange, medium; light blue, weak; blue, hydrophilic. (C) Hydrophobic surface of the fragment SIVmac p6$^{16-28}$ using same color coding as in (B).

Figure 64. Characterization of the SIVmac p6 structure in the presence of DPC-d_{38} micelles. Superposition of natural abundance 2D ^{1}H-^{15}N HMQC spectra of SIVmac p6 recorded at 600 MHz and 323 K in the absence (red) and the presence (blue) of DPC-d_{38} micelles. The sequence of the full-length SIVmac p6 at the top.

Figure 65. SIVmac p6 spectra in acetate buffer with 100 mM DPC-d_{38}. (A) Superposition of 2D-NOESY (blue) and 2D-TOCSY (red) spectra recorded on a 600 MHz and at 323 K. The sequential NOEs connectivities [dαN (i, i + 1)] and [dN (i, i + 1)] are outlined. (B) Distance restraints chart.

Figure 66. Superposition of the ten lowest energy structures on residues 18-39 (left, red) and 45-59 (right, green).

Figure 67. Ribbon representation for the lower energy conformer of SIVmac p6.

Figure 68. Superimposition of the NMR structures of SIVmac p6 and of SIVmac p6$^{16-28}$ peptide.

Figure 69. Mapping the interaction sites between HIV-1 Vpr and HIV-1 p6 in the presence of DPC micelles. (A) Overlay of ^{1}H-^{15}N HMQC spectra of ^{15}N, ^{13}C labeled HIV-1 Vpr in free form (green) and perturbed by the addition of unlabeled HIV-1 p6 (red). The ^{15}N, ^{13}C labeled HIV-1 Vpr sample was prepared at final concentration of 300 μM in the presence of 90 mM DPC-d_{38} (300 eq.) in acetate buffer. The molar ratio of HIV-1 Vpr to HIV-1 p6 is 1:4. (B) Amide proton chemical shift perturbation of ^{15}N, ^{13}C labeled HIV-1 Vpr in complex with HIV-1 p6 at molar ratio of 1:4. The average value is shown in solid line and 1 / 0.5 standard deviation (SD) is shown in dash. (C) Mapping of the surface of the HIV-1 Vpr (PDB code 1M8L) according to chemical shift perturbations (CSPs) observed on the spectra. CSPs > average + 1 standard deviation (SD) in red; CSPs > average + 0.5 SD and < average + 1 SD in orange; CSPs > average and < average + 0.5 SD in yellow.

Figure 70. Mapping of the interaction between p6 and Vpr of HIV-1 by NMR studied by variation of chemical shift of labeled amino acids in p6 after addition of Vpr. HIV-1 p6 and Vpr are in a 1: 4 molar ratio. (A) Superposition of the ^{1}H-^{15}N SOFAST-HMQC spectra

of HIV-1 p6 comprising six labeled residues in free form (blue) and in the presence of HIV-1 Vpr (red). (B) Superposition of the plans extracted from the 3D NOESY-HSQC spectra at the ^{15}N frequency of each labeled residue of HIV-1 p6 in free form (blue) and in the presence of HIV-1 Vpr (red).

Figure 71. Interaction between SIVmac Vpx and the peptide SIVmac $p6^{16-28}$ by NMR. Overlap of ^{1}H-^{15}N SOFAST-HMQC spectra of the labeled ^{15}N and ^{13}C SIVmac Vpx/unlabeled $p6^{16-28}$ at molar ratio of 1:3 in its free form (blue) and in complex with the peptide SIVmac $p6^{16-28}$ (magenta).

Figure 72. Mapping of SIVmac $p6^{16-28}$-Vpx interactions by NMR. Overlap of natural abundance ^{1}H-^{15}N SOFAST-HMQC spectra of SIVmac $p6^{16-28}$ in free form (blue) and in complex with the SIVmac Vpx (magenta).

Figure 73. Interaction between SIVmac Vpx and the full-length SIVmac p6 by NMR. ^{1}H-^{15}N SOFAST-HMQC spectra of the labeled ^{15}N and ^{13}C SIVmac Vpx in its free form (blue) and in the presence of p6 at molar ratio of 1:3 (red). Blue is in free form of the labeled ^{15}N and ^{13}C SIVmac Vpx and red is in complex with the full-length SIVmac p6.

Figure 74. Mapping the interaction sites between SIVmac Vpx and SIVmac p6 in the presence of DPC micelles. (A) Chemical shifts perturbation (CSP) of labeled ^{15}N and ^{13}C SIVmac Vpx/unlabeled p6 at molar ratio of 1:3. Amide proton chemical shift perturbations analysis of ^{15}N, ^{13}C labeled SIVmac Vpx residues interacting with the full-length SIVmac p6. The average value is shown in solid line and 1 / 0.5 standard deviation (SD) is shown in dash line. (B) Surface of the hot spots displayed based the crystal structure of SIVmac Vpx (PDB code 4CC9). (C) Surface of the hot spots displayed based the solution NMR structure of SIVmac Vpx. CSPs > average + 1 standard deviation (SD) (red); CSPs > average + 0.5 SD and < average + 1 SD (orange); CSPs > average and < average + 0.5 SD (yellow).

Figure 75. Mapping of SIVmac p6-Vpx interactions by NMR. (A) Overlap of ^{1}H-^{15}N SOFAST-HMQC spectra of labeled SIVmac p6 in free form (blue), in complex with the

SIVmac Vpx at molar ratio of 1:0.05 (green), and in complex with the SIVmac Vpx at molar ratio of 1:0.1 (pink). The side chain amide proton in R38 is highlighted *R38.

Figure 76. Sequence alignment. (A) Sequence alignment of primate Vpr and Vpx proteins. (B) Sequence alignment of different p6 proteins. The motifs and strains used in this study were highlighted in red and black boxes, respectively.

Figure 77. Mapping of SIVmac Vpx-HIV-1 p6 interactions by NMR. Overlap of ^{1}H-^{15}N SOFAST-HMQC spectra of labeled SIVmac Vpx in free form (blue), in complex with the HIV-1 p6 at molar ratio of 1:3 (light green), and in complex with the HIV-1 p6 at molar ratio of 1:5 (red).

Figure 78. Interaction of labeled HIV-1 p6 with unlabeled SIVmac Vpx. (A) Overlap of ^{1}H-^{15}N SOFAST-HMQC spectra of HIV-1 p6 in free form (blue)and in complex with the unlabeled SIVmac Vpx at molar ratio of 1:0.3 (cyan) of 1:0.8 (light green), and of 1:1.2 (red). (B) Variations of chemical shift perturbations (CSPs) of three residues involved in the interaction with SIVmac Vpx. The value of their K_d was computed.

Figure 79. Solution NMR study of DO1148 binding to HIV-1 Vpr. (A) ^{1}H-^{15}N SOFAST-HMQC spectra of HIV-1 Vpr in free form (blue) and in complex with DO1148 at molar ratio of 1:10 (red) recorded at 323 K on a 600 MHz spectrometer equipped with cryoprobe. (B) Chemical shifts perturbation (CSP) at HIV-Vpr/DO1148 at molar ratio of 1:10. The average value is shown in solid line, and the standard deviations (SD) of 0.5 and 1 are shown in dashes.

Table 1. Multiple roles of Vpr in the HIV-1 replication cycle.

Table 2. The composition of the modified M9a and secondly modified M9b medium.

Table 3. Summary of 10 structures of Vpx in the presence of DPC micelles.

Table 4. Summary of long distance restraints of SIVmac Vpx.

Table 5. Restraints and structural statistics for the 10 final structures of the full-length HIV-1 p6 in the presence of DPC-d_{38} micelles.

Table 6. Structural statistics for the 10 final structures of SIVmac $p6^{16-28}$ in the presence of DPC-d_{38} micelles.

Table 7. Structural statistics for the 10 final structures of SIVmac p6 in the presence of DPC-d_{38} micelles.

Introduction

1. The situation of AIDS in the world and importance of further studies on HIV

Acquired immune deficiency syndrome (AIDS) is caused by human immunodeficiency viruses (HIVs). Since the identification of the first case of AIDS in 1980 in the United States, the number of people who died of HIV/AIDS has been up to 33 million. Furthermore, 38.0 million people were living with HIV/AIDS at the end of 2019 according to the report by the World Health Organization (https://www.who.int/data/gho/data/themes/hiv-aids).

HIV is transmitted by unprotected sex and by direct contact with bodily fluids including blood, breast milk, semen and vaginal fluids. As for now, HIV infected people cannot be cured except for two cases: the “Berlin patient” with acute myeloid leukemia and the “London patient” with cancer who were successfully cured by the transplant of mutant CCR5 stem cells, which encourages the infected people and bring the dawn of hope [1,2]. Fortunately, HIV infection is fully preventable. Condoms can prevent sexually HIV transmission and the transmission from mother to child in the process of pregnancy, delivery and breast feeding can be effectively blocked by antiretroviral treatment (ART), which can reduce the viral load and enable people to live longer. Four categories of antiretroviral drugs (ARDs) including nucleoside/non-nucleoside reverse transcriptase inhibitors, protease inhibitors, integrase inhibitors and entry inhibitors have been developed and used for HIV therapy. In addition, some antibodies are being studied for HIV treatment. However, drug-resistant HIV caused by error-prone HIV replication occurs within a few weeks or several years after starting ART [3–5]. Furthermore, ART does not eradicate HIV infection due to the persistent infection of the latently and unrecognition by the immune system. Besides, antibody drugs cause a heavy financial burden, and their widespread application is unavailable in developing countries. Therefore, to completely eradicate HIV, studies on HIV are still necessary for overcoming drugs resistance, identifying new targets, and developing new drugs, antibodies or vaccines.

2. Primate lentiviral genome and replication cycle

2.1 Primate lentiviral genome

Lentiviruses are a subfamily of retroviruses, whose RNA is reverse-transcribed and then integrated into DNA of the host genome. The primate lentivirus family includes simian immunodeficiency viruses (SIVs) and human immunodeficiency viruses (HIVs). Sequential analysis reveals that the primate lentiviral genomes are divided into 8 or 9 genes, which encode for 15 different proteins (**Figure 1**) [6].

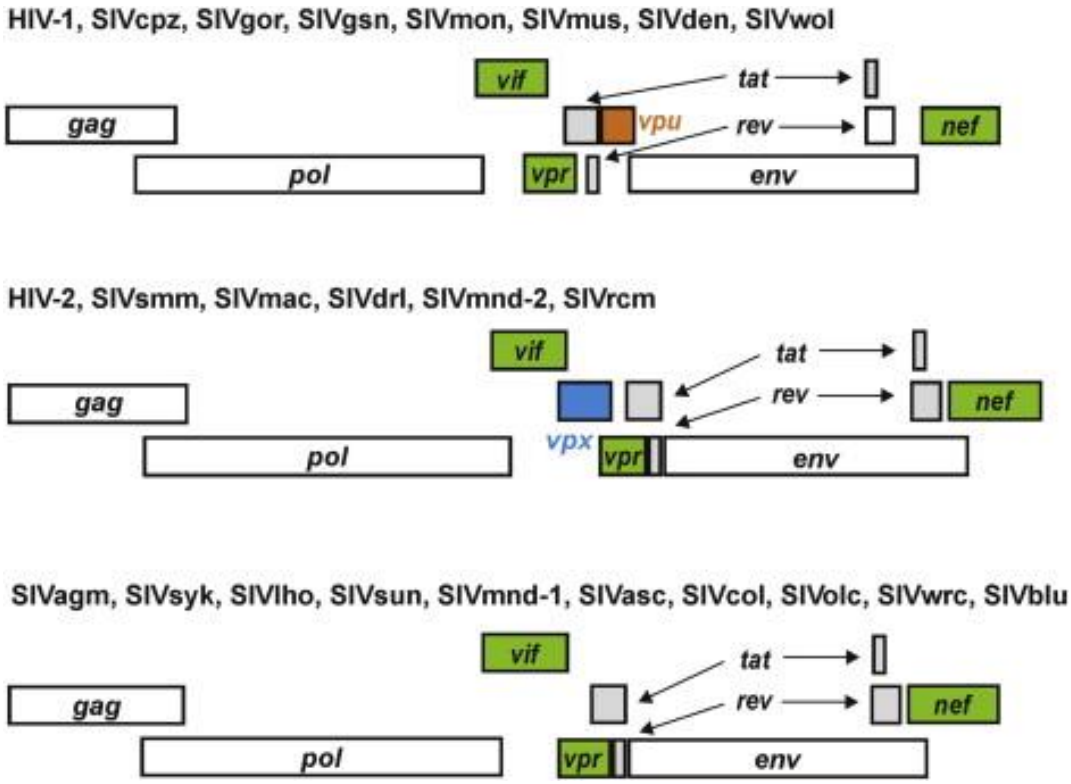


Figure 1. *Schematic representations of different primate lentiviral genomes (extracted from [7]).*

HIVs are classified into two broad types, HIV-1 and HIV-2. Currently, HIV-1 spreads the global pandemic, while the infection by HIV-2 occurs mainly in West Africa [7]. Both HIV types result from cross-species transmissions of different simian immunodeficiency viruses.

The transmissions from different SIVs of chimpanzees and gorillas to humans occurred in the 1900s and produced three HIV-1 groups, M, N, and O (**Figure 2**) [8,9]. Subsequently, in 2009, a new group P was reported to share greater similarities of sequence with a SIV of wild gorillas (SIVgor). By contrast, HIV-2 was closely related to a simian virus that is found in West African primate including the captive macaques and the sooty mangabey [10].

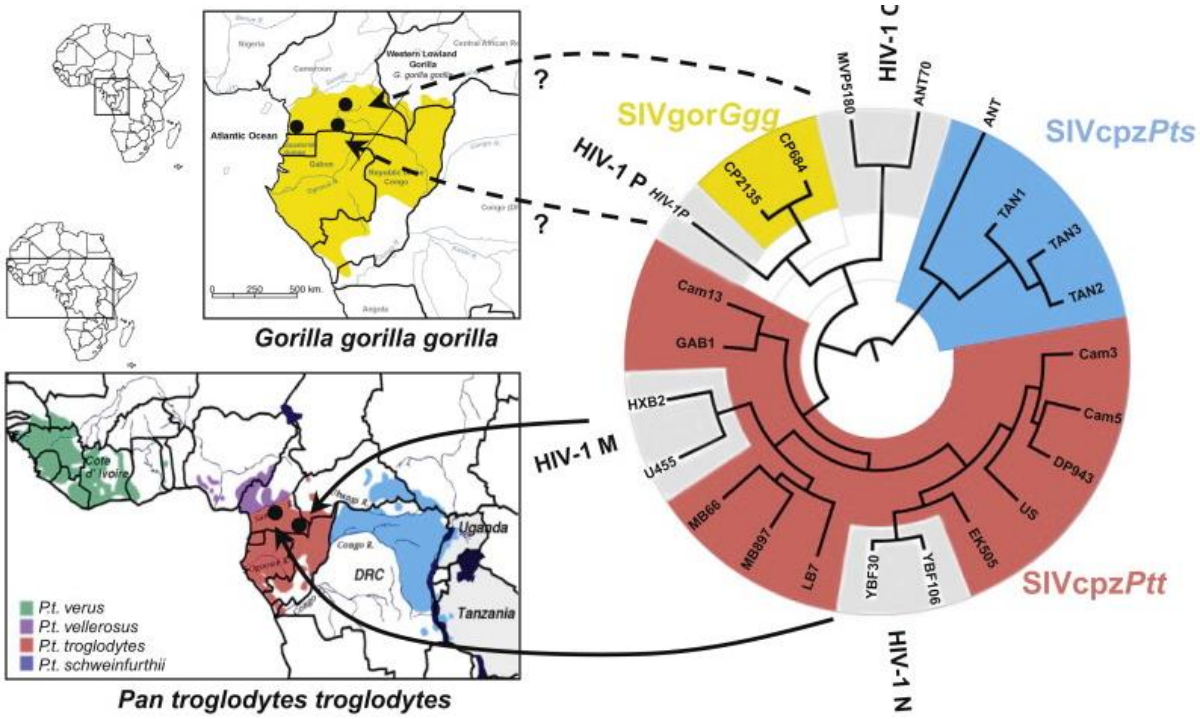


Figure 2. *Origin of HIV-1 groups from the analyses of SIV and HIV-1 genomes (extracted from [8]). gorGgg, cpzPts and cpzPtt are species subspecies. SIVgorGgg, SIVgor Gorilla gorilla gorilla; SIVcpzPts, SIVcpz P. t. schweinfurthii; SIVcpzPtt, SIVcpz Pan troglodytes troglodytes.*

Given that HIV-1 spreads the global pandemic, we focused on HIV-1 genome with nine open reading frames that encode for 15 viral proteins (**Figure 1**). Subsequently, antisense protein (ASP) encoded by the negative strand of HIV-1 was identified. Three major genes: *gag*, *pol*, and *env*, individually encode the precursor of the polyprotein Gag (Pr55Gag), the precursor of GagPol (Pr160GagPol) and the precursor of the envelop (gp160). The Gag precursor is cleaved by the HIV protease (PR) into six proteins (matrix or MA, capsid or

CA, Spacer peptide 1 or SP1, nucleocapsid or NC, Spacer peptide 2 or SP2 and p6). The GagPol precursor is cleaved into three viral enzymes: (protease (PR), reverse transcriptase (RT), and integrase (IN). The envelop protein precursor (gp160) is cleaved into mature surface glycoprotein (gp120) and the transmembrane glycoprotein (gp41) by the cellular proteases. Regulatory proteins (Tat and Rev) and accessory proteins (Vif, Vpu, Vpr and Nef) are encoded by the remaining genes (**Figure 3**).

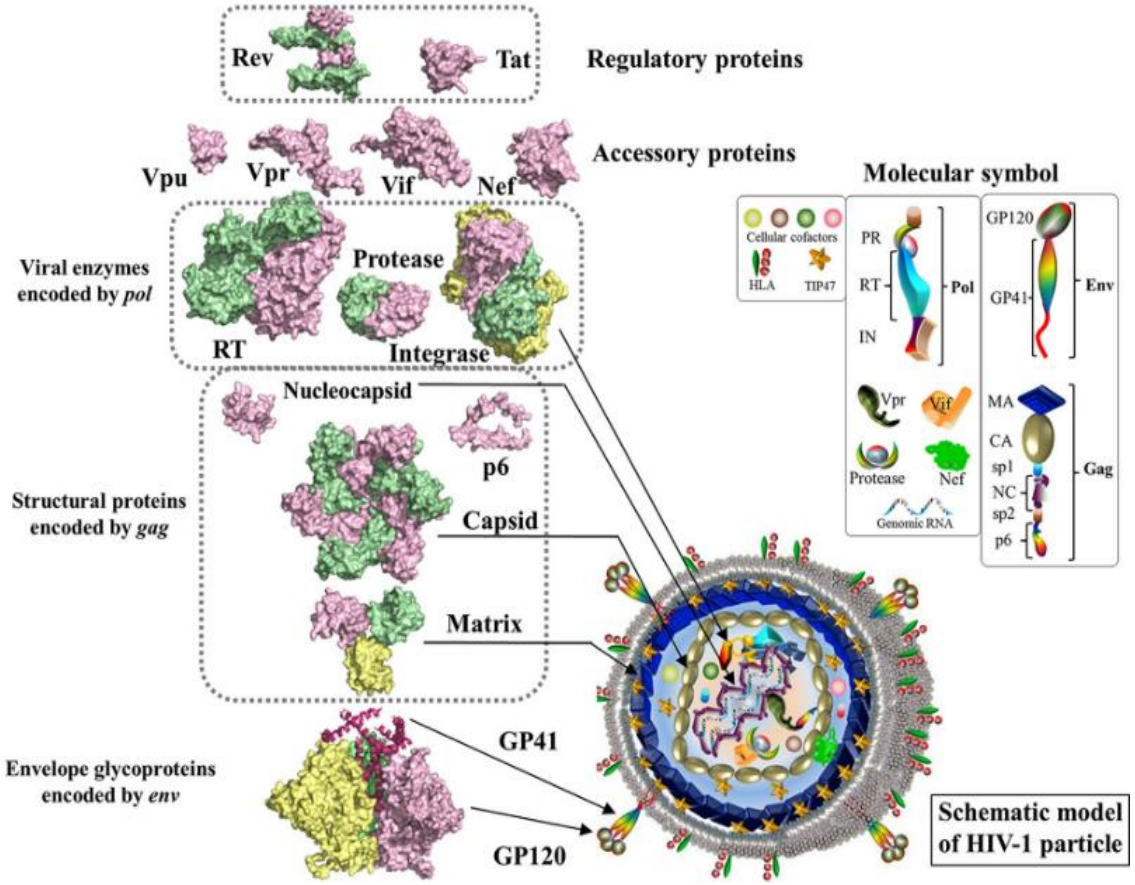


Figure 3. *Surface representation and schematic view of the HIV-1 particle [11].*

2.2 Primate lentiviral replication cycle

Albeit the primate lentiviruses are divergent, the structure and biological functions of HIVs and SIVs are both well conserved. Here, we present the primate lentiviral replication cycle using the well-characterized HIV-1 as a prototype.

Previous studies showed two phases of HIV-1 replication cycle [12]: the early phase (viral attachment and membrane fusion, reverse transcription, nuclear entry, and integration) and the late phase (transcription, translation and assembly, budding and maturation) (**Figure 4**).

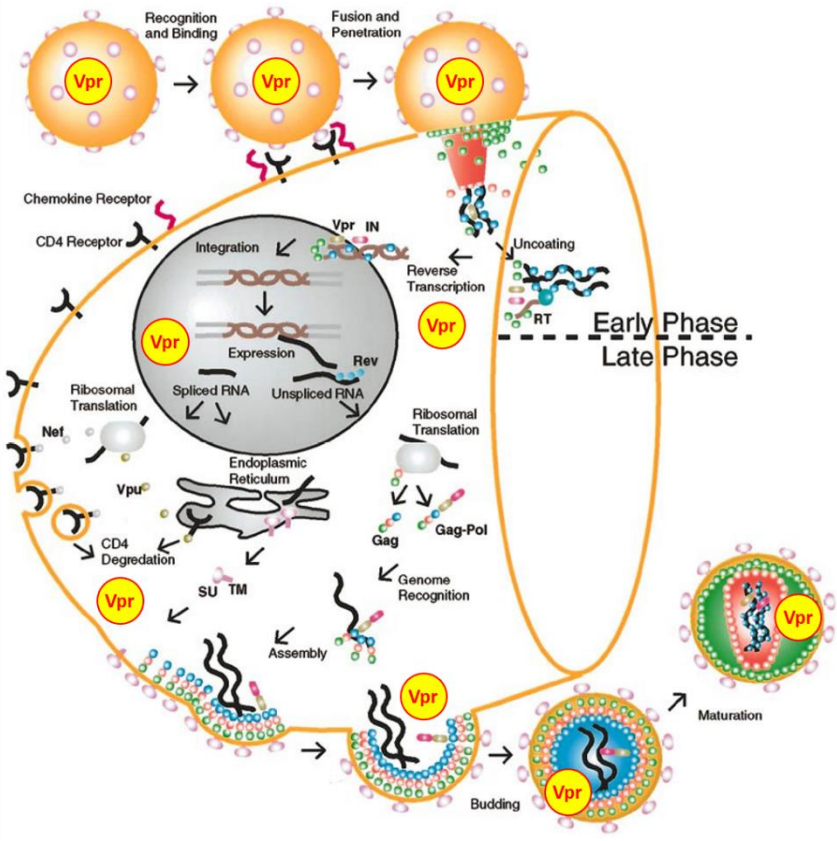


Figure 4. *HIV-1 replication cycle [12].*

In the early phase, the viral particles firstly attach at the surface and penetrate the host cells expressing the cluster of differentiation 4 (CD4) glycoprotein on their surface. The interaction between the two glycoproteins gp120 and gp41 of Env in the outer membrane

of HIV makes it possible to maintain the metastable form of the viral spike without ligand. In order to invade host cells, HIV Env spikes formed by the non-covalent association of gp41 and gp120 firstly undergo structural rearrangements [13–23]. Then, gp120 first binds to the main receptor CD4 and then binds to the co-receptors CCR5/CXCR4 after a conformational change. Subsequently, the conformation keeps changing to exposes gp41 [24]. gp41 mainly contains the ectodomain, the transmembrane domain, and the cytoplasmic domain. The ectodomain is composed of the hidden fusion peptide region, the helical N-terminal heptad repeat (NHR) and C-terminal heptad repeat (CHR). Once exposed, the hidden hydrophobic gp41 fusion peptides are inserted in the host cell membrane to facilitate the fusion of the viral envelope and the host cell membrane and allow the virus nucleocapsid to penetrate into the cell [25] (**Figure 5**). Due to the unknown process of the reverse transcription and integration of reverse transcription complexes (RTC) and preintegration complexes (PIC) [26,27], it is though that the viral nucleocapsid penetrates the cell and that the viral nucleic acid, the reverse transcriptase, integrase, protease, etc., required for virus replication are released into the cytoplasm of the host cells after fusion. Then, the single-stranded RNA was used as a template for the reverse transcriptase to synthesize a single-stranded cDNA, which in turn serves as a template for double-stranded DNA synthesis by the DNA polymerase in the cytoplasm. The integrase cleaves specific phosphodiester bonds near the viral double-stranded DNA ends during the 3'-processing reaction to form a pre-integration complex (PIC) that includes the integrase, MA, Vpr and and the viral double-stranded DNA [28]. Following the translocation of the PIC to the nucleus induced by Vpr [29,30], the integrase catalyzes the 3'-end processing in the presence of a divalent metal ion and at the same time makes a staggered cut of the DNA of the chromosome to open a gap for the insertion of the viral double-stranded DNA. Finally, the double-stranded DNA is inserted into the host genome by the integrase [31]. This integrated double-stranded DNA is called the proviral DNA or provirus.

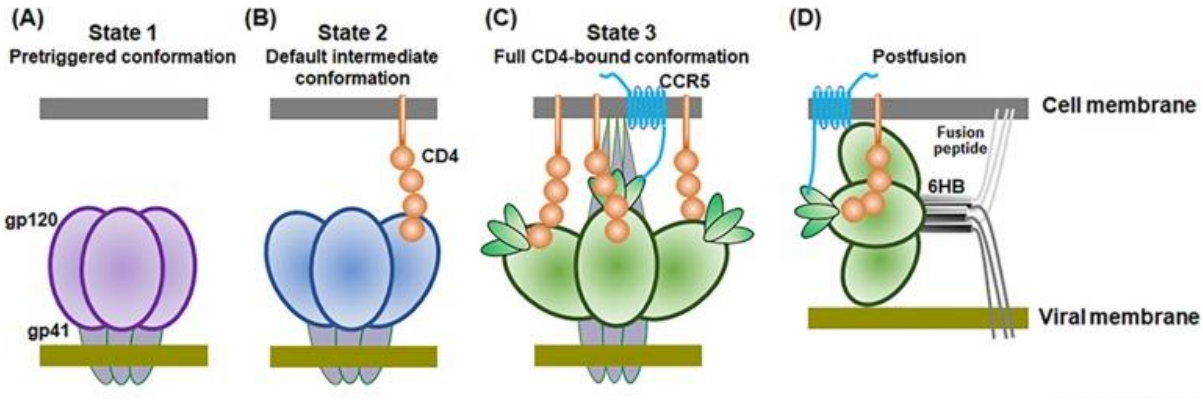


Figure 5. *Cartoons of HIV-1 attachment and fusion including pretriggered conformation (A), default intermediate conformation (B), full CD4-bound conformation (C), and postfusion (D)* [32].

In the late phase, proviral transcription is active and initiated by the viral promoter in the U3 region of the 5' LTR. The viral DNA used as a template is transcribed into RNA by the host polymerase II (Pol II) in the nucleus. Initially, the transcriptional output is very low due to less efficient elongation of viral transcripts. However, the transcription is regulated by regulatory protein Tat, which is absolutely required for the activation of the transcriptional elongation [33]. Tat can dramatically improve the efficiency of the elongation by Pol II through interactions with the transactivation response (TAR) RNA element (**Figure 6**). Then, some RNA molecules are capped and tailed to become the viral genome and the some other are spliced into viral mRNAs, which are translated into precursor polyproteins.

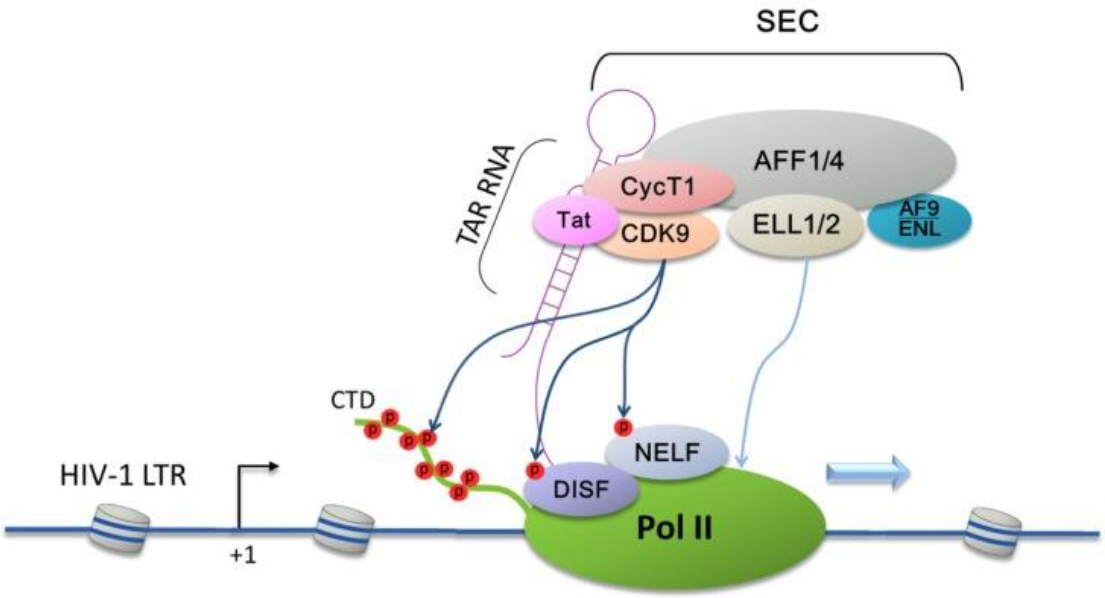


Figure 6. *Tat activates transcriptional elongation to promote viral transactivation* [34].

The translation initiation occurs for most eukaryotic mRNAs by a scanning mechanism, of which the 40S ribosomal subunit is recruited to the cap structure and scans in the 5' to 3' direction until encountering the initiation codon. Once the initiation codon has been identified, the 80S ribosome is formed and the translation elongation begins. 80S ribosome then mediates the addition of amino acids and elongates the polypeptide chain (**Figure 7**). In addition, internal ribosome entry site (IRES) also initiates HIV-1 translation by directly recruiting the 40S ribosome. The HIV-1 5' UTR contains TAR, Poly(A) loop, PBS domain, the SD, DIS and Ψ. These elements are very important for different stages of the HIV-1 replication cycle. Prior to translation initiation, the 5' UTR of HIV-1 mRNAs is capped with a 7-methylguanosine (m^7G). The cap-dependent translation initiation relies on the recognition of the 5' UTR cap structure by the eIF4F complex to promote the recruitment of 43S PIC.

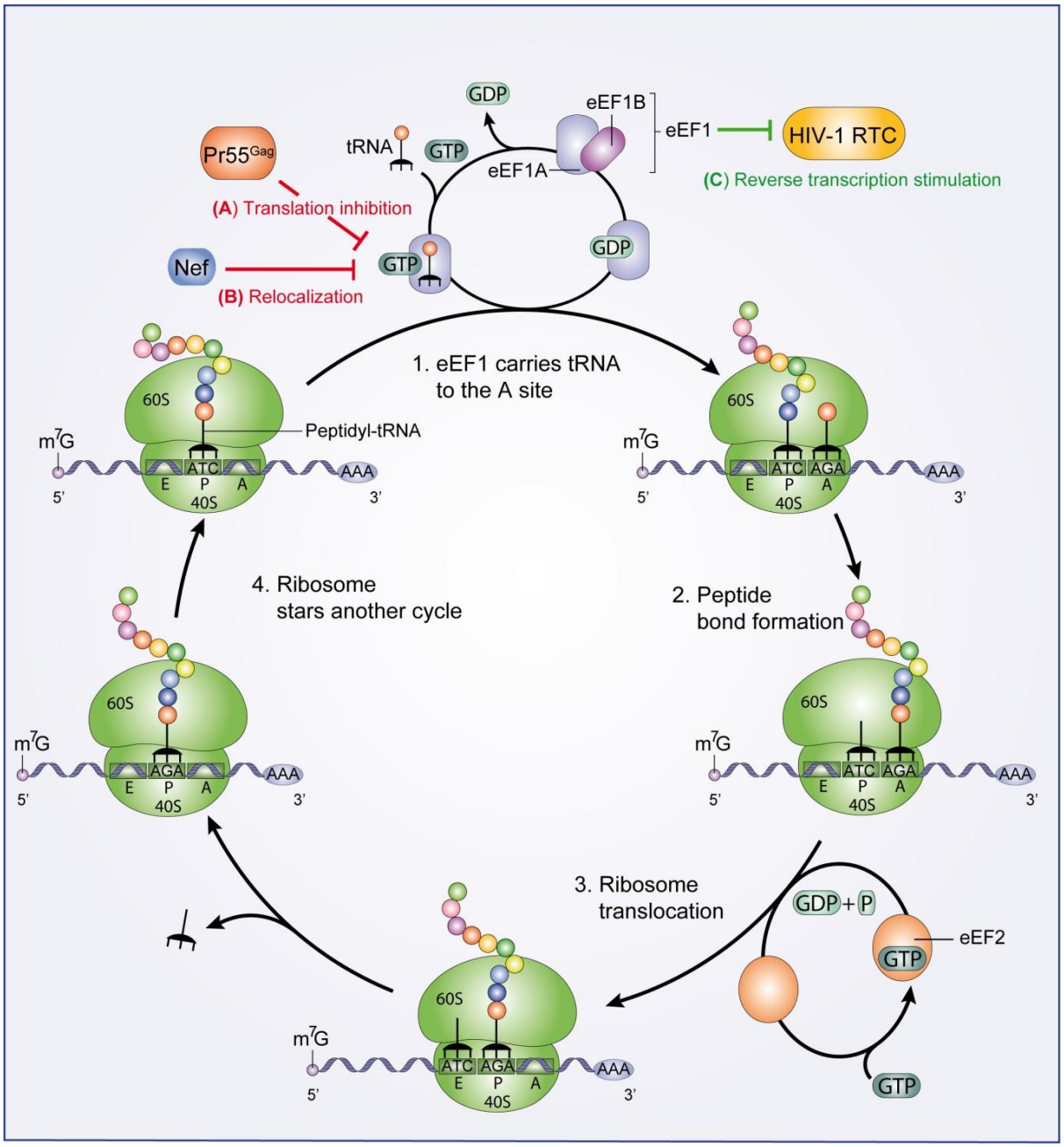


Figure 7. *Eukaryotic translation elongation* [35]. *The aminoacylated tRNA was transported into the A site of the 80S ribosome by eEF1A at step 1. The 80S ribosome mediates peptide bond formation at step 2. The 80S ribosome translocation is mediated by eEF2 at step 3. Then, the ribosome continues next cycle of peptide elongation at step 4.*

The viral assembly begins with the synthesis of the Gag polyprotein in the cytosol [36]. The NC protein of the gag contains two zinc fingers that bind the RNA genome. The NC-RNA interaction can package two copies of the genome. Thus, the Gag polyprotein and the

dimeric RNA assemble into the nascent, immature virus particle together with the GagPol precursor at the plasma membrane (**Figure 8**). Gag can directly interact with the inositol headgroup of phosphatidylinositol 4,5-bisphosphate, also known simply as PI-4,5-P_2, which is a minor phospholipid component of cell membranes. PI-4,5-P2 further triggers the exposure of N-terminal myristate of Gag and then facilitates its insertion into the plasma membrane. Once the assembly of the immature Gag lattice into a hexagonal structure at the plasma membrane is finished, the nascent particle undergoes a membrane fission event mediated by the cellular endosomal sorting complex required for transport (ESCRT) (**Figure 8**). ESCRT drives different cellular processes such as endosomal sorting, vesicular trafficking, maintenance of plasma membrane integrity, membrane fission and enveloped virus budding. The PTAP motif from the C-terminal p6 domain of Gag directly interacts with tumour susceptibility gene 101 (TSG101), which is a subunit of ESCRT-I. ESCRT-III can assemble into circular arrays and help to constrict the membrane [37,38]. In the process of budding, the viral protease cleaves the Gag and GagPol polyproteins to trigger virus maturation, which will further drive major changes in virion morphology, the generation of the conical capsid core, and production of the fully processed MA, CA, NC, p6, PR, RT and IN proteins. Furthermore, these processed proteins rearrange to generate the mature and infectious virion [36,39].

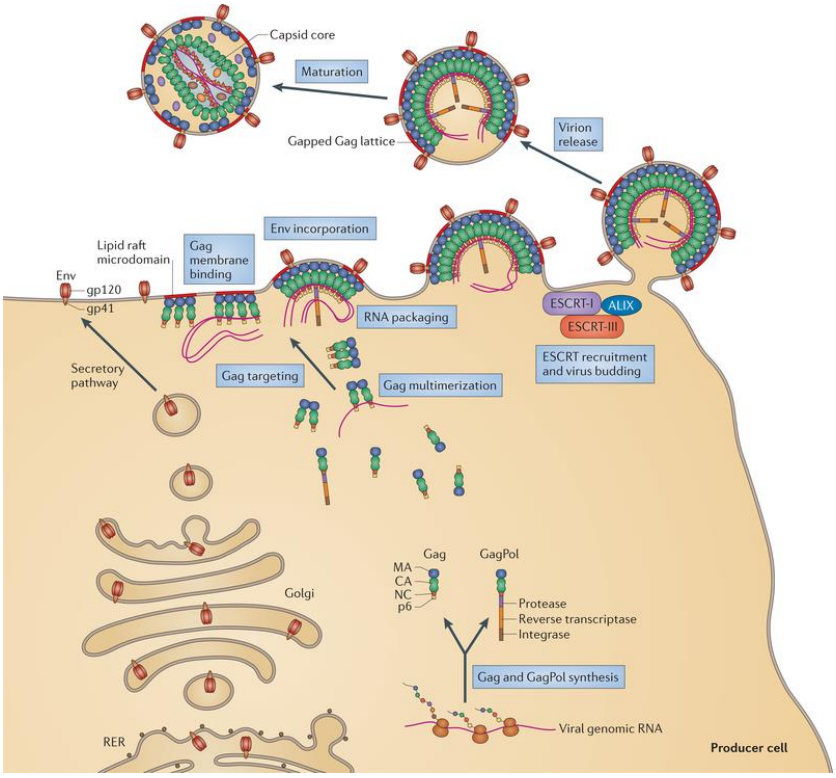


Figure 8. *Schematic of the HIV-1 assembly, budding, and maturation extracted from* [36].

Recent results from 3D correlative fluorescence light and electron microscopy, and cryo-electron tomography reported that microtubules could mediate the uncoating process of the HIV-1 capsid (**Figure 9**) and that cone-shaped HIV-1 capsids can be directly transported into the nucleus (**Figure 10**) [40–43]. Thus, a new model for HIV-1 replication cycle was proposed (**Figure 11**) [44]. The capsid penetrates into the cell membrane and enters into cytoplasm [25,40,44]. After fusion, the single-stranded RNA serves as a template for DNA synthesis by the reverse transcriptase as the capsid is transported into the nuclear envelope. Previous model showed the reverse transcription happens in the cytoplasm after uncoating (**Figure 4**). By contrast, here, the new model shows the reverse transcription starts in the cytoplasm and ends in the nucleus before uncoating.

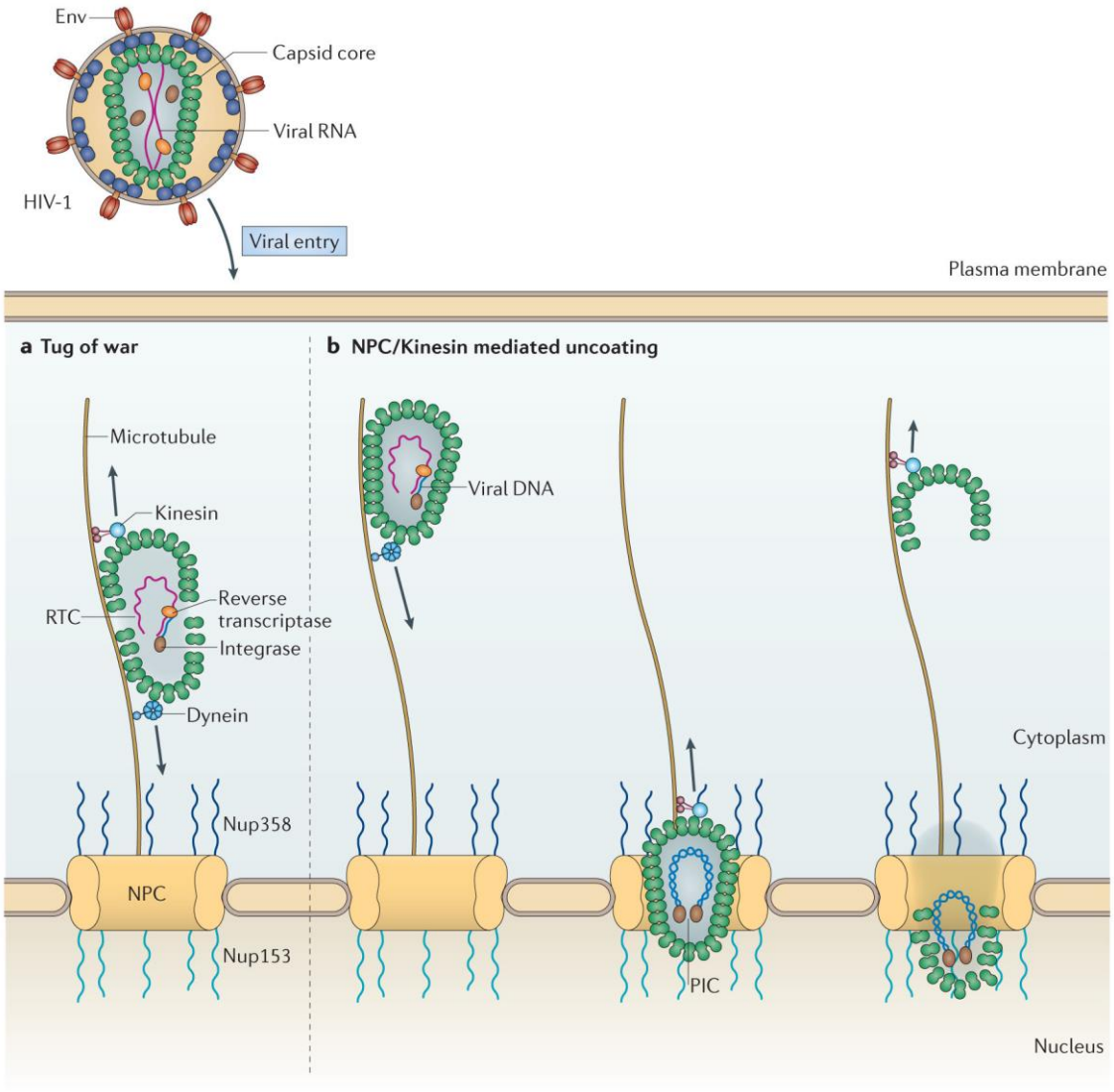


Figure 9. *Uncoating of the HIV-1 capsid mediated by the microtubules* [41]. *(A) Uncoating occurs by the tug-of-war mediated by dynein and kinesin motor proteins. (B) Dynein mediates and facilitates the interaction between the capsid core and components of the nuclear pore complex (NPC). Env, viral envelope glycoproteins; PIC, pre-integration complex; RTC, reverse transcription complex.*

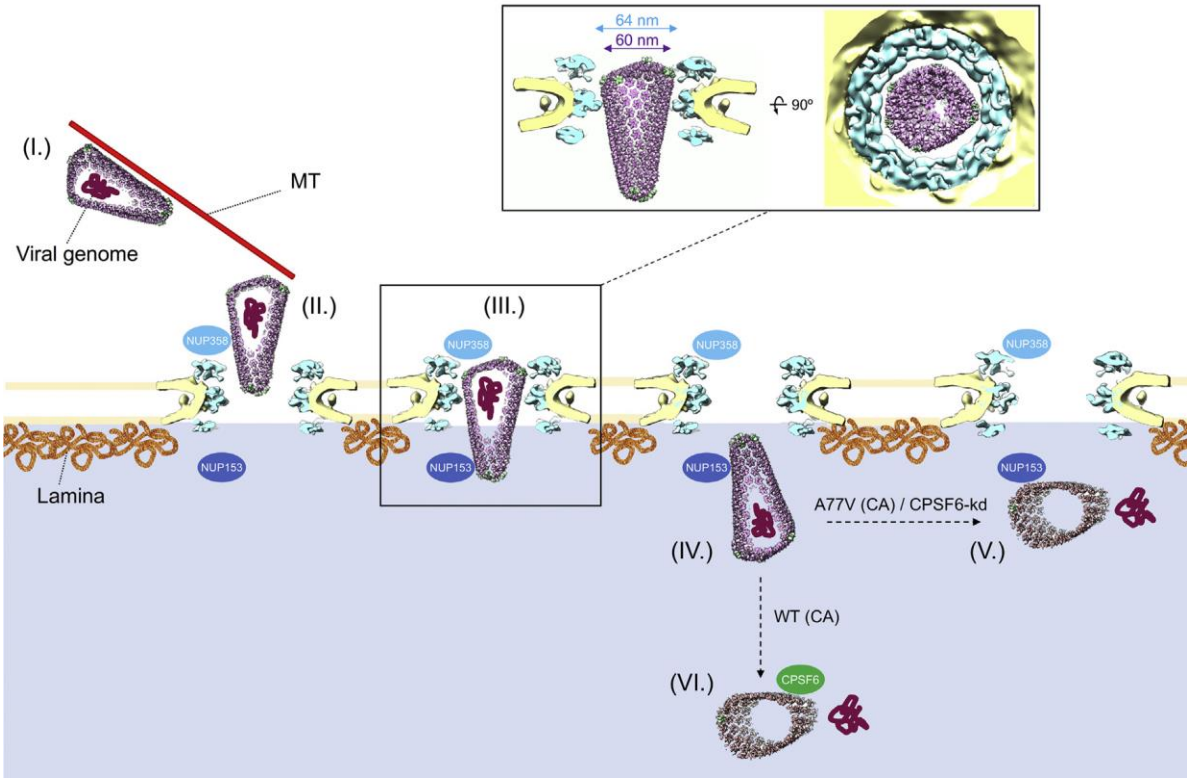


Figure 10. *The model of HIV-1 nuclear import in T cells* [40]. *The intact cone-shaped HIV-1 capsid including the viral genome is transported towards the nuclear periphery thanks to the microtubules (MT) at step I. NUP358 binds to NPC and docks to the NPC at step II. The capsid then penetrates the channel at step III. Subsequently, the intact HIV-1 capsid translocates into the nucleus at step IV. After the departure from the NPC central channel, CPSF6 interacts with the capsid, and then the capsid uncoats and releases the genome at step V. At steps V and VI, the viral genome is released and integrated into the host genome.*

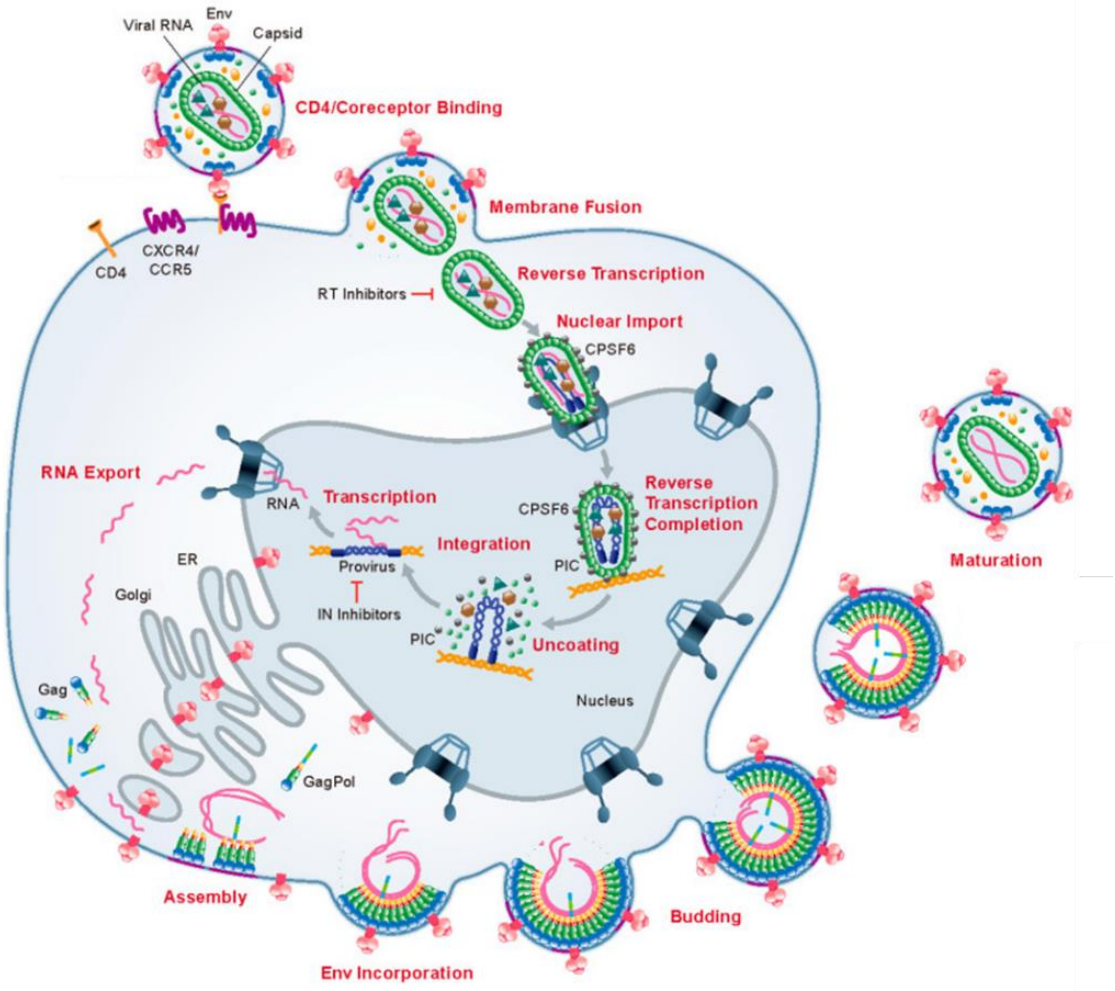
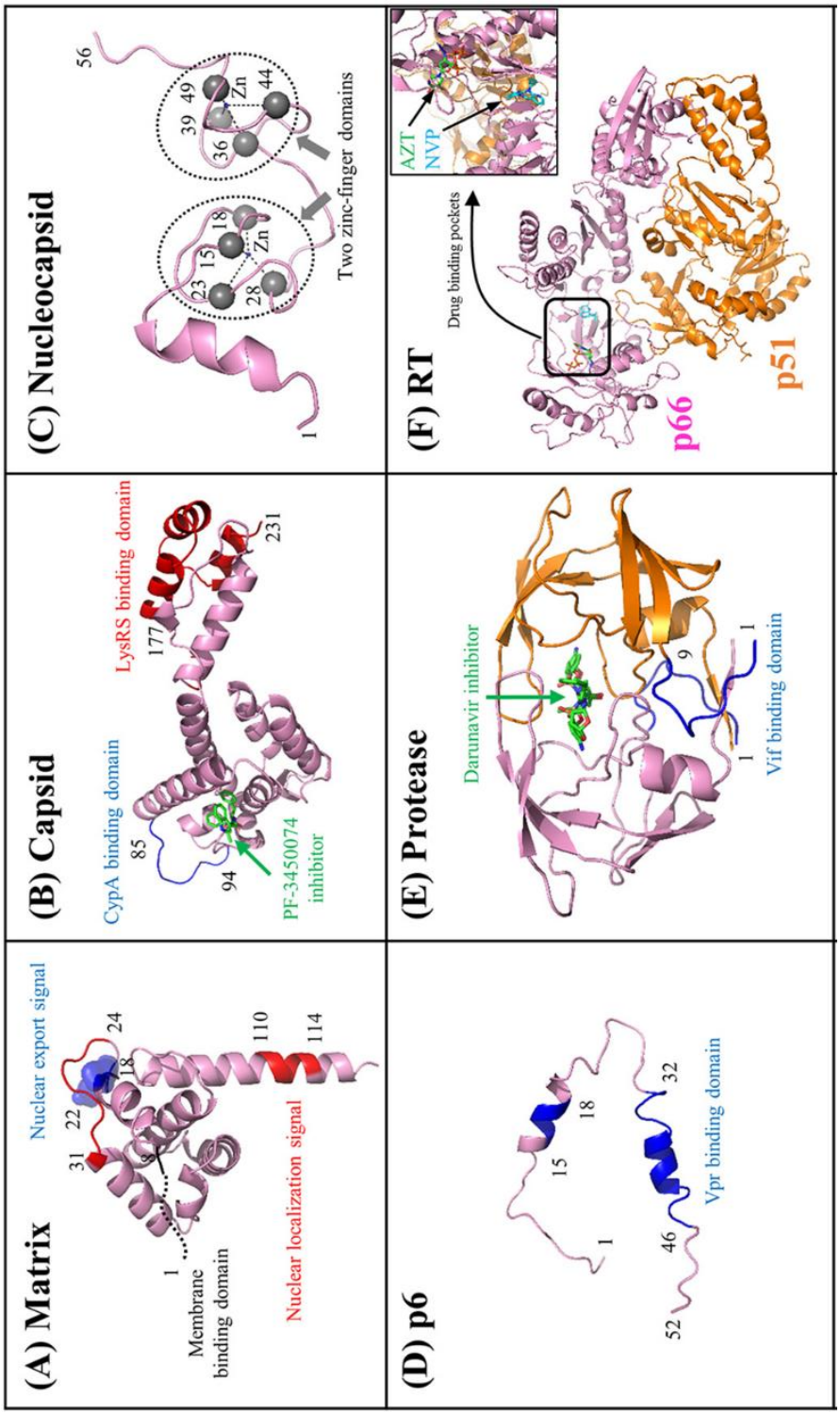


Figure 11. *New HIV-1 replication cycle from* [44].

3. HIV / SIV proteins and their interactions

HIVs and SIVs almost share common structures and biological function. We thus chose to focus on the HIV proteins. The HIV genome encodes 16 viral proteins, which are structural proteins (matrix, capsid, nucleocapsid and p6), three viral enzymes (integrase, protease and reverse transcriptase), two envelope proteins (gp120 and gp41), two regulatory proteins (Tat and Rev) and five accessory proteins (Vif, Vpu/Vpx, Vpr and Nef). Interestingly, Vpu is only carried by HIV-1, while Vpx can be found in HIV-2 and in a limited range of SIVs including SIVmac/SIVsmm and SIVrcm/mnd2. Here, we list the structures of the 16 proteins and their interaction domains (**Figure 12**). Furthermore, the networks of the

interactions between pairs of HIV proteins are shown in **Figure 13** and the summary of the HIV-1 pairwise protein associations are shown in **Figure 14**.



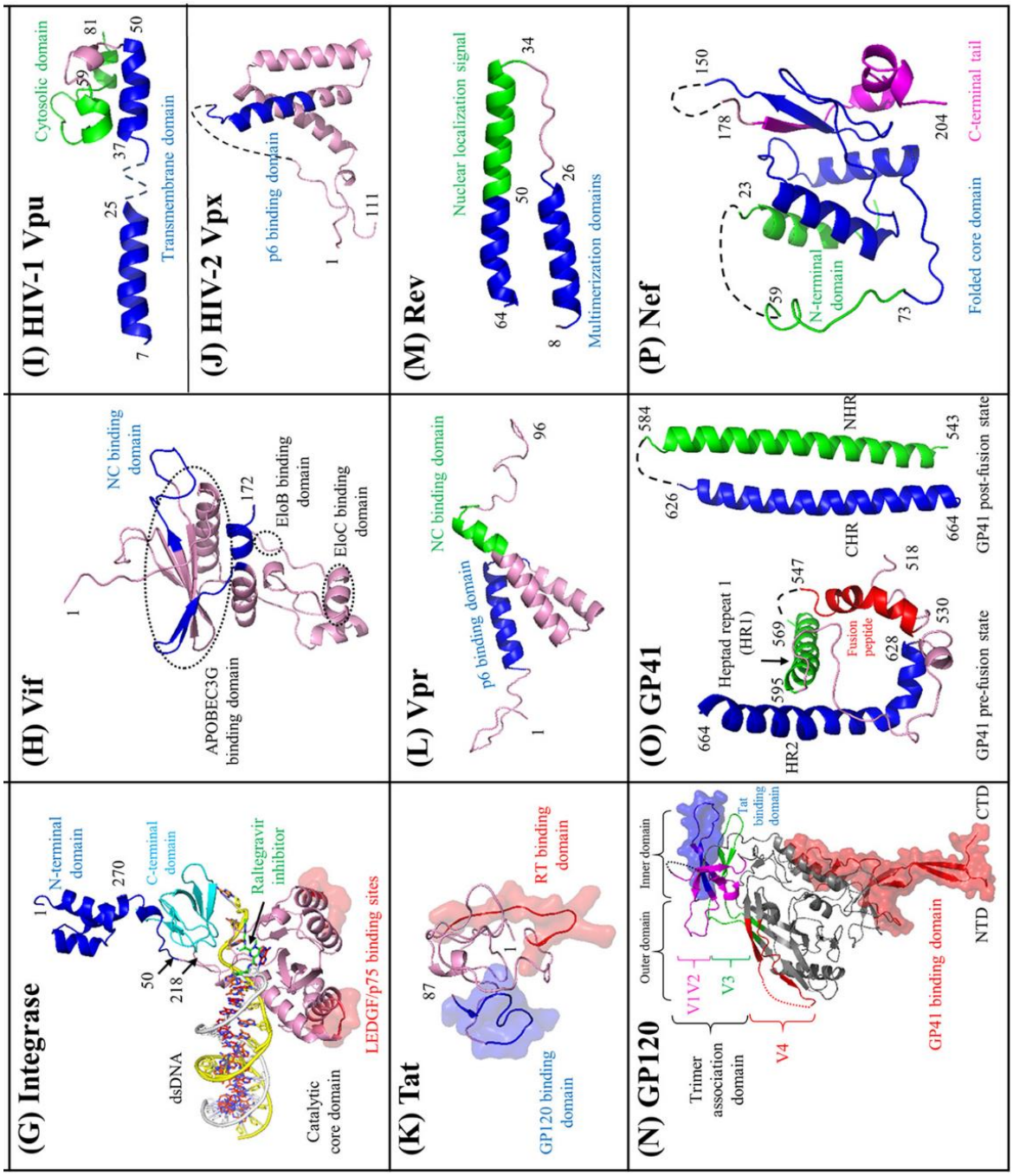


Figure 12. *Cartoon representation of 15 HIV proteins and their functional domains from* [11]. *The structure of ASP protein is not known, and it is not there.*

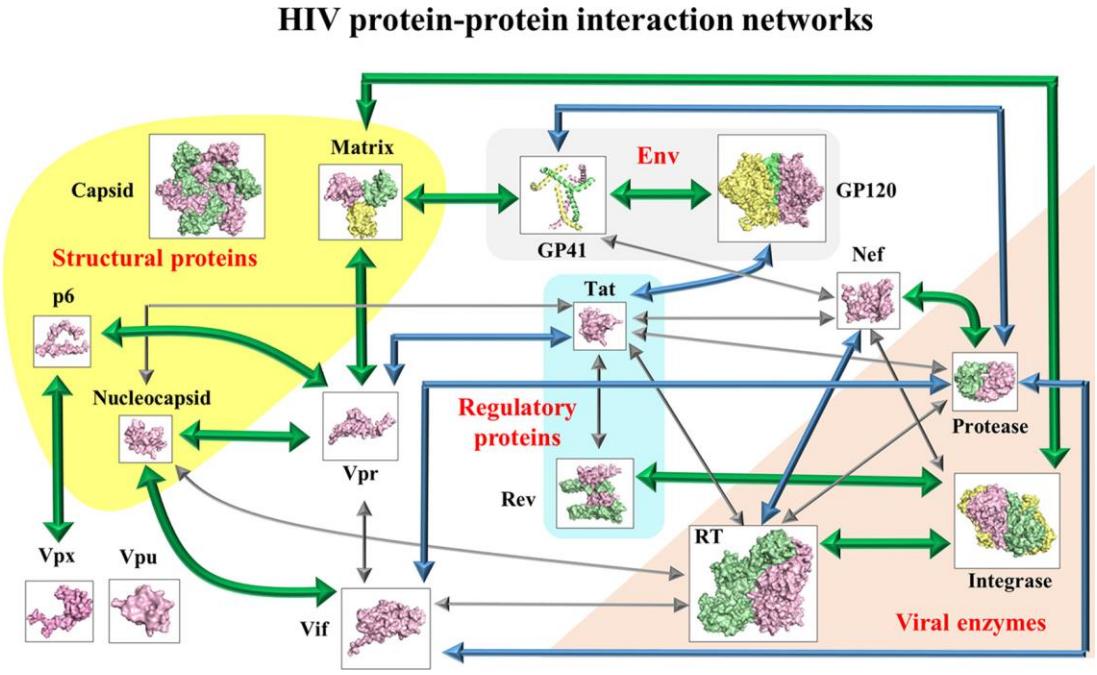


Figure 13. *The networks of HIV protein-protein interactions, extracted from [11]. Green arrows, well-known interactions cited more than 300 times or reported by more than 100 citations at least 3 publications; gray arrows, little-known interactions reported with fewer than 100 citations at a single paper; blue arrows, lesser-known interactions reported by fewer citations.*

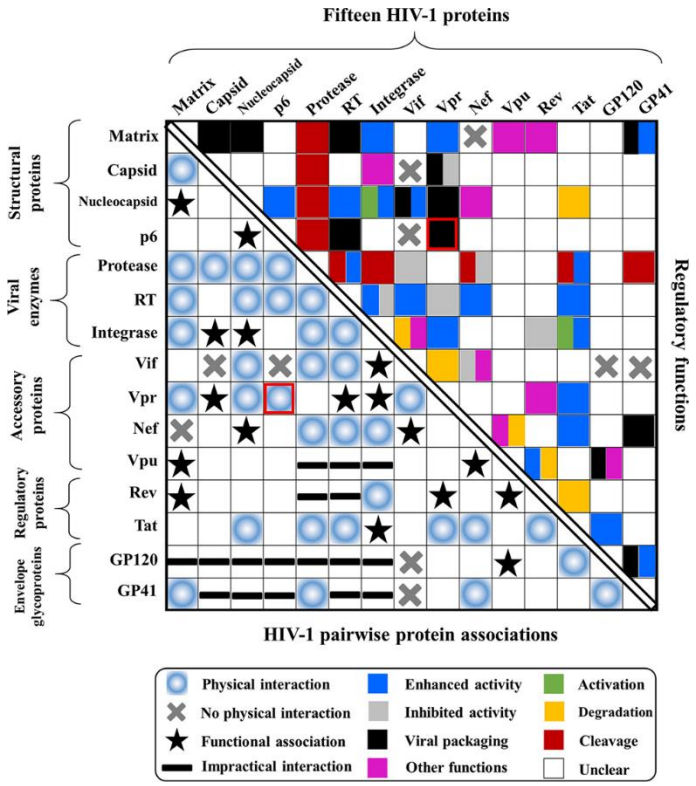


Figure 14. *Summary of HIV protein-protein associations, extracted from [11]. Fifteen HIV-1 proteins are displayed at the left and top, and protein functions are highlighted in different colors. The interaction between Vpr and p6 was shown in red square.*

4. HIV-1 Gag precursor

The Gag polyprotein is mainly responsible for the viral assembly and budding. It contains four proteins, the matrix (MA), the capsid (CA), the nucleocapsid (NC) proteins and one small protein p6, and two spacer peptides SP1 and SP2 [36,39,45] (**Figure 15**).

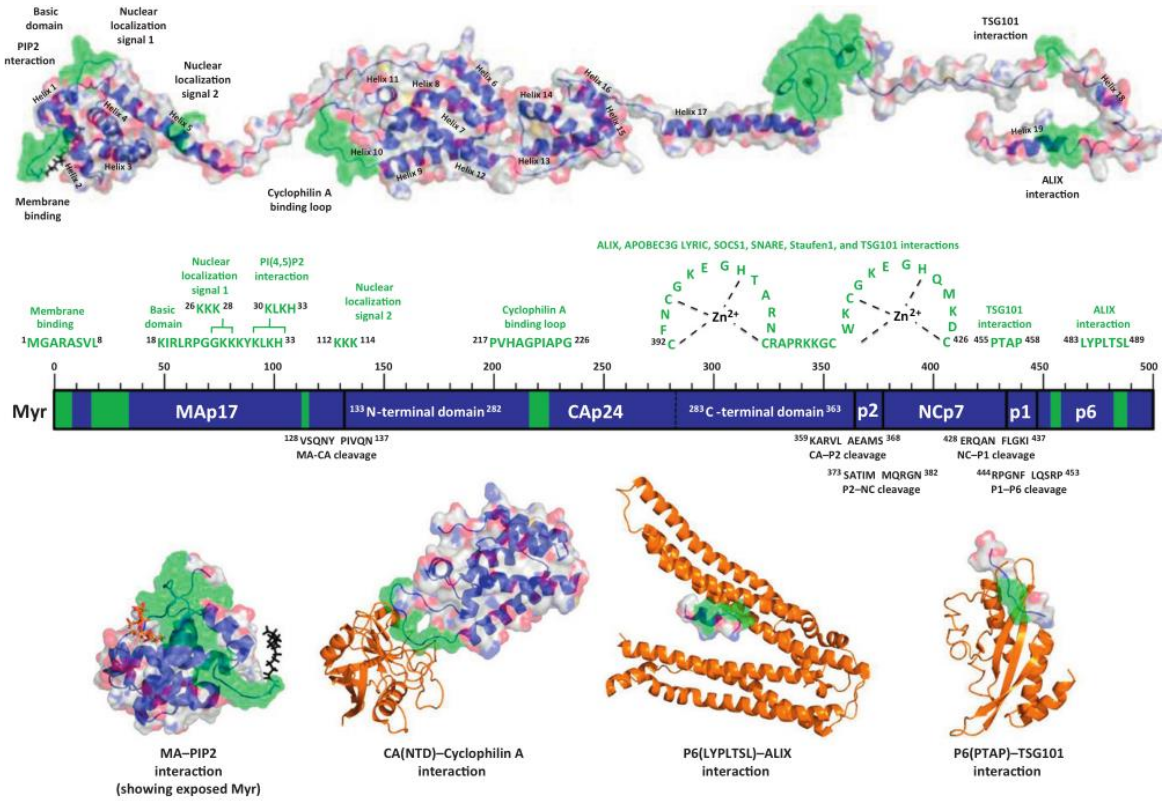


Figure 15. *Models of HIV-1 Gag from* [46].

The matrix (MA): The MA protein with 132 amino acids, comprises five alpha helices and a three-strand mixed beta sheet [47]. N-terminal myristylation (Myr) plays an important role in plasma membrane targeting and viral assembly. In the presence of N-myristyltransferase (NMT), myristic acid is covalently linked to Gly-2 of the matrix protein by covalent attachment within the consensus sequence MGXXX(S/T)XX. The MA domain of Gag is myristoylated and together with the first four helical domains of MA, myristic acid including a 14-carbon saturated fatty acid, forms a hydrophobic pocket to bind to the plasma membrane and fit into the lipid bilayer as anchor [48]. In addition, the highly basic region (HBR) at the N-terminal of the MA was also found to interact by electrostatic interactions with phosphatidylinositol 4,5-bisphosphate [PI (4,5) P_2] of the plasma membrane. Generally, the viral assembly on the plasma membrane is mainly mediated by the interaction of the myristoylated matrix domain of Gag with PI(4,5)P_2 of the plasma membrane [46,49–53] (**Figure 16**).

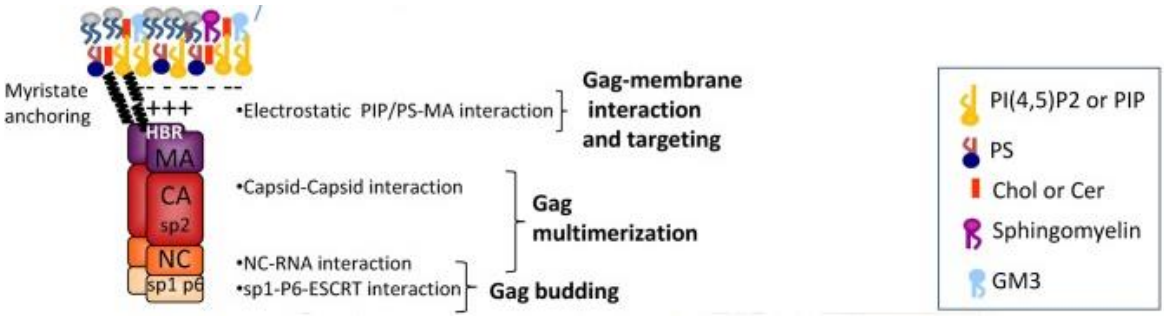


Figure 16. *Molecular model of HIV-1 Gag assembly from* [54].

The capsid (CA): The HIV genome is protected by a CA shell that is constructed from ∼1000−1500 copies of the CA protein [55–57]. As shown in **Figure 17**, CA protein is divided into two domains, the N-terminal domain CA-NTD in green and the C-terminal domain CA-CTD in blue, linked by a flexible linker [58]. Structurally, the CA-NTD domain displays seven α-helices and a cyclophilin A (CypA)-binding loop connecting helices 4 and 5 [59,60] (**Figure 17**). CA-NTD domains interact with each other, forming an hexagonal in the immature capsid [61]. CypA, a prolyl peptidyl isomerase, can be packaged into HIV particles by binding to Gag. The CA-CTD domain contains a short 3_{10} helix followed by an extended strand of four alpha helices [62,63]. A major homology region (MHR), ^{284}DIRQGPKEPFRDYVDRFYKTL^{304}, was found in CA-CTD. MHR is highly conserved in the gag gene of all retroviruses. Mutations of the MHR can inhibit viral maturation [64–66].

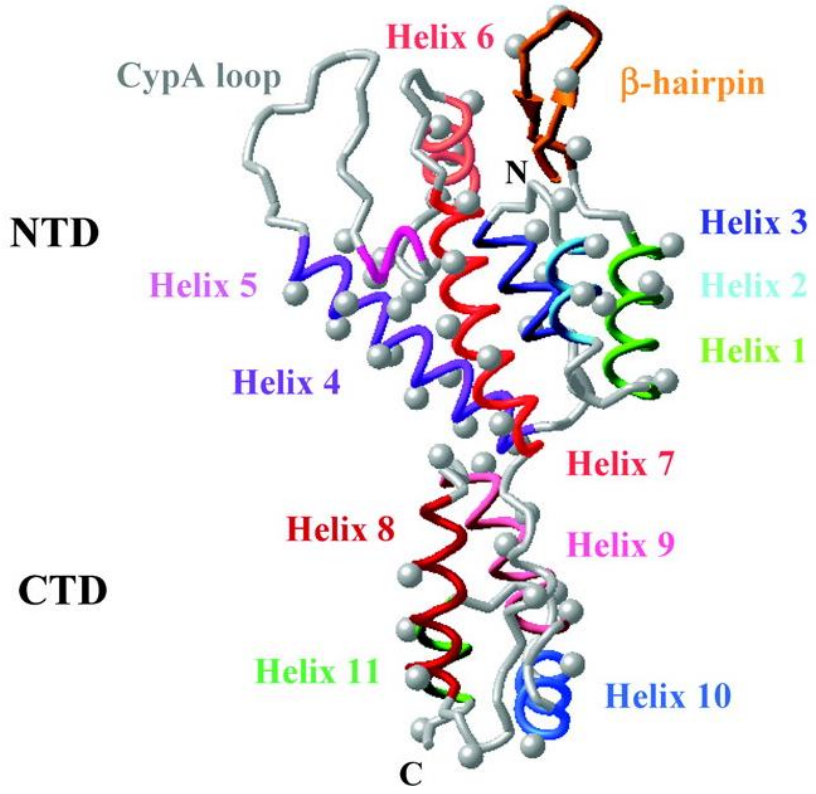


Figure 17. *The HIV-1 capsid* [67]. *Each part is shown in different colors. Orange, β-hairpin; light green, helix 1; turquoise, helix 2; dark blue, helix 3; purple, helix 4; gray, cyclophilin A-binding loop; pink, helix 5; rose, helix 6; red, helix 7; dark red, helix 8; salmon, helix 9; blue, helix 10; and green, helix 11.*

The SP1 Spacer peptide: The SP1 spacer peptide, also called P2, is composed of 14 amino acids. The HIV CA assembly is regulated by SP1. Mutations in both CA–SP1 or SP1–NC junctions influence the Gag–Gag interactions. Structurally, SP1, together with CA C-terminal and NC N-terminal domains forms an alpha helix from residue L343 of CA to Val390 of NC in distilled water/TFE-d_2 (70:30; v/v) as seen by solution nuclear magnetic resonance (NMR) [68,69]. The cleavage of CA–SP1 occurs in the late phase of viral processing, which is important for CA maturation. Thus, the CA-SP1 junction was regarded as an attractive drug target. Bevirimat and its derivative EP39 studied by our laboratory, both leads to CA–SP1 cleavage inhibition [70,71] and are promising drugs to work on for the development of a new class of antiretroviral (ART) compounds.

The nucleocapsid (NC): NC is a fifty-five amino acid protein with two zinc fingers [72]. It plays important roles in the recognition and interaction of viral RNA. NC is a chaperon protein can facilitate viral RNA dimerization and stabilize the dimeric RNA genome [73,74]. NC is involved in reverse transcription of the viral genome by promoting the annealing of the (+)/(-) primer-binding site (PBS) of the viral genomic RNA during the second strand transfer reaction [75]. During HIV-1 assembly, the NC domain (of Gag) is required for selection and packaging of HIV-1 genomic RNA via binding to the domain comprising the stems loops SL1 to SL4, in 5 ' of the genomic RNA [76].

The SP2 spacer peptide: The SP2 spacer peptide, also called P1, is composed of 16 amino acids. Two conserved proline residues P439 and P445 are important for the incorporation of Gag and Pol into the virion [77,78].

The small protein p6: As shown in **Figure 16**, p6, located at the C terminus of Gag, is composed of 52 amino acids. It can be translated into two forms, the p6 in Gag and the p6* in Gag-Pol. The solution structure of isolated p6 was determined by solution NMR and displays two helices, spaning residues ^{432}SFRSG^{466} and ^{481}KELYPLTSLRSL^{492} [79].

5. Structure and function of Vpr / Vpx / p6

Primate lentiviral accessory proteins Vpr and Vpx are both incorporated into virions by direct interaction with p6, the C-terminal domain of the precursor Pr55Gag [80–82]. This incorporation is essential for the viral replication in resting cells, the translocation of the pre-integration complexe (PIC) to the nucleus and for the neutralization of the cellular antiviral factors such as the uracil DNA glycosylase 2 (UNG2), the sterile alpha motif, the HD-domain–containing protein 1 (SAMHD1) and the human silencing hub (HUSH). Furthermore, Vpr and Vpx are particularly required for efficient virus replication in nondividing cells such as macrophages. Vpr is involved in the transport of PICs to the nucleus. It transactivates the viral long terminal repeat (LTR) and induces cell cycle arrest [83–88]. Sharing amino acid sequence similarities with Vpr, Vpx is also involved in PIC translocation into the nucleus. In the last few decades, a series of studies on lentivirus-

related proteins have revealed the importance of Vpr and Vpx in the viral life cycle and in innate immune antagonism [85].

5.1 Structure and function of Vpr

Vpr is distributed to all primate lentiviruses such as simian immunodeficiency viruses (SIVs), human immunodeficiency virus HIV-1 and HIV-2. The structure of HIV-1 Vpr with three α-helices was well defined by solution nuclear magnetic resonance (NMR) and by X-ray crystallography (**Figure 18**) [89,90].

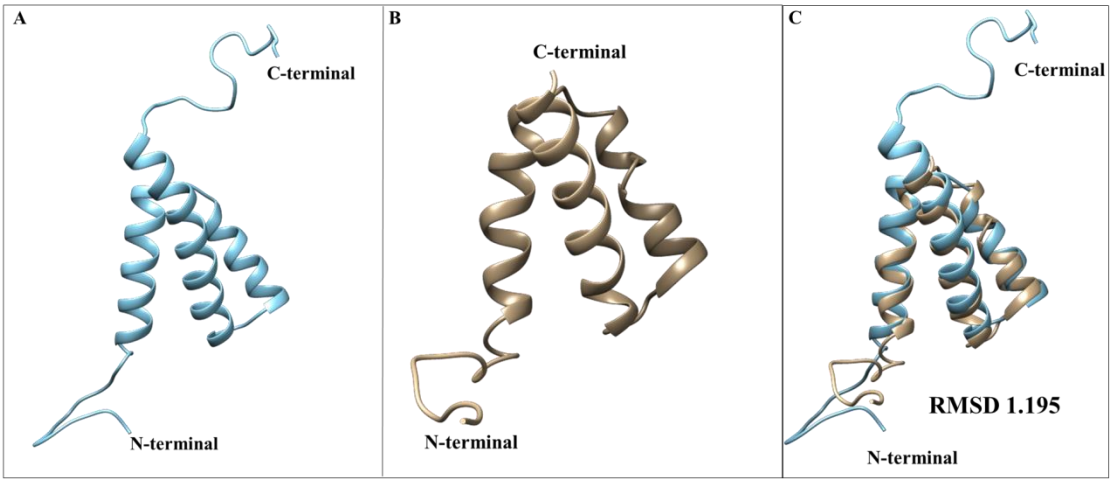


Figure 18. *Structures of HIV-1 Vpr. (A) Solution structure of the chemically synthesized HIV-1 Vpr determined by NMR (PDB:1M8L). (B) Crystal structure of HIV -1Vpr extracted from the DDB1-DCAF1-Vpr-UNG2 complex (PDB:5JK7). (C) Superposition of solution NMR (A) and crystal (B) structures of HIV-1 Vpr.*

However, dimers and higher order oligomers of Vpr have also been reported. Vpr oligomers can be packaged into virus particles and have the ability to provide nuclear translocation, while non-oligomeric Vpr proteins retain cell cycle arrest and apoptosis [91]. Vpr molecules mainly exist as dimers or trimers, co-distributed in the nucleus and cytoplasm, but mainly in the nucleus as a nucleocytoplasmic shuttling protein [92]. To gain a better understanding of the oligomerization process, docking models of parallel or

antiparallel dimers and hexamers were built using Cluspro and Rosetta software (**Figure 19**) [93–96].

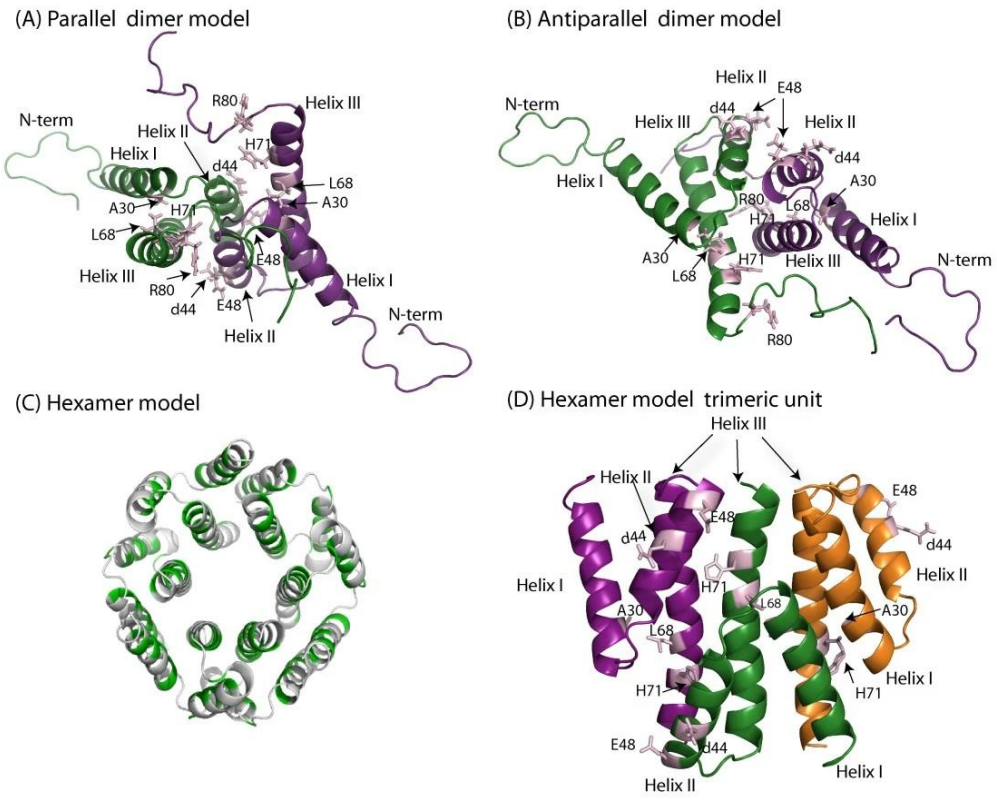


Figure 19. *Different models of HIV-1 Vpr. The parallel dimer model (A), the antiparallel dimer model (B), Hexamer model top view (C), and hexamer model side view (D).*

HIV-1 Vpr is produced late in the viral replication cycle and incorporated into the virion through a direct interaction with p6, the carboxy-terminal domain of the precursor Pr55Gag, also known as $p6^{gag}$ [97,98]. Although Vpr is expressed at a late stage of the virus replication cycle, it is present during the early steps of infection and plays different roles in viral replication cycle since the virion is released during uncoating (**Table 1**).

Table 1. Multiple roles of Vpr in the HIV-1 replication cycle

HIV-1 life stage	Vpr function	Reference(s)
Binding or attachment	-	-
Fusion and releasing	-	-
Reverse transcription	Contributes to the fidelity of viral reverse transcription by interacting with Uracil-DNA glycosylase-2 (UNG2)	[99–103]
	Promotes the assembly of deacetylated Lys-tRNA into virus particles by inhibiting the acetylation of Lys-tRNA	[104–108]
Viral integration	Promotes the nuclear transport and localization of HIV-1 PIC in the non-dividing cells	[29,30,88,109–116]
	Enhances the expression of unintegrated viral DNA expression	
	Promotes the binding of integrase to dsDNA by inducing DNA double-strand breaks	
Viral transcription	Activates HIV-1 LTR-directed transcription through the interaction with a TATA box of LTR	[117–122]
	Promotes viral gene transcription by directly trans-activing ubiquitous cellular transcription factor Sp1	
	Facilitates the transcriptional activity of LTR by interacting with Tat	
	Indirectly regulates LTR activity by relieving LTR activity inhibition from UNG2	
	Vpr activates the oxidative stress pathway and induce the expression of HIF-1α to regulate transcription	
	Induces the phosphorylation of transforming growth factor-β-activated kinase 1 (TAK1) and activates NF-κB and AP-1 to stimulate HIV-1 LTR promoter	
	Reactivates the viral promoter by inducing the degradation of histone deacetylase 1 (HDAC1) and HDAC3	
Viral translation	Regulates viral translation by the interaction of Vpr with human nucleoporin CG1 (hCG1)	[123]
Assembly	-	-
Budding	Incorporated into virions by binding to NCp7 and p6	[98,124–127]

Vpr contributes to the fidelity of viral reverse transcription. The reverse transcription process is extremely error-prone and results in the frequent introduction of point mutations. A wrong nucleotide is incorporated every 2000-5000 bp. The reverse transcriptase of HIV-1 does not have a proofreading function, which gives retroviruses a very high mutation rate

[128]. Although the high mutation rate is a greater chance for HIV-1 to evolve and escape the immune system, mutation rates are almost disastrous and increase the risk of lethal non-functional genomes. Vpr can balance this condition by interacting with Uracil-DNA glycosylase-2 (UNG2) in the cell. UNG2 is a DNA repair protein of the base excision repair system that specifically removes the RNA base uracil from single or double-stranded DNA [99,100]. Vpr interacts with UNG2 and then UNG2 is packaged into virus particles to remove uracil and maintain the fidelity in the process of the viral reverse transcription [101]. Usually, cytidine deamination or dUTP is mistakenly incorporated into DNA. If this error cannot be removed and repaired, C->T and G->A mutations will occur during the next round of replication. Mutant viruses in Vpr W54R can abolish the interaction with UNG2. In addition, HIV-1 RT does not possess 3' - and 5' -exonucleolytic proofreading activity and is shown to be error-prone in cell free systems explaining that C->T and G->A base pair transition mutations are increased 4-fold in the process of HIV-1 reverse transcription [102,103]. In addition, the reverse transcription initiation of HIV-1 needs deacetylated Lys-tRNA as a primer. Lys-tRNA synthetase interacts with Vpr. Furthermore, in the presence of Vpr, the aminoacylation of Lys-tRNA mediated by the Lys-tRNA synthetase is inhibited, suggesting that the interaction of Vpr with this synthetase may influence the initiation of the HIV-1 reverse transcription [104–108].

Vpr promotes the nuclear transport and localization of HIV-1 PIC in the non-dividing cells. There are several different nuclear transport pathways. Although the mechanism of nuclear transport pathways is complicated, most pathways include a binding step to the transporter, transportation to the nuclear envelope, translocation into the nucleus, and release from the cargo. The classic nuclear protein import pathway involves importin-α and importin-β as adaptor and carrier respectively. The adapter importin-α binds to the nuclear localization signal (NLS) sequence motifs on the cargo proteins and interacts with importin-β through its N-terminal importin-β binding domain (IBB) to form importin-β-importin-α-cargo complex [109,110]. Subsequently, the importin-β mediates the translocation of the cargo proteins into the nucleus through nuclear pore complexes (NPCs) that span the outer and inner nuclear membranes. In the nucleus, the interaction of importin-β with Ran GTP, a protein involved in transport in and out of the cell nucleus

during interphase, dissociates the import complex and then releases the cargo proteins [111]. Some researchers found that Vpr binds to the adapter importin-α [29,30,113] in the PIC, [88,112]. Thus, the import complex including importin-β, importin-α and the PIC is hypothesized to be translocated through NPCs and finally be released. However, the results of nuclear transport assay *in vitro* showed that the nuclear import of Vpr was not affected by the depletion of importin-β upon using siRNAs technology, while depletion of importin-α significantly decreased the efficiency of the nuclear import of Vpr by the same nuclear transport assay [114]. Thus, some researchers hypothesize that Vpr could interact with importin-α and transports the PIC into nucleus in an unknown way. In addition, Vpr was shown to interact specifically with phenylalanine-glycine (FG)-repeat domains of the nucleoporin protein, the constituent building blocks of the NPCs, suggesting that this interaction is critical for importin-mediated nuclear import [115]. Furthermore, Vpr disrupts the nuclear envelope suggesting Vpr may carry the HIV-1 PIC to the nucleus as a nucleon-cytoplasmic shuttling protein [116]. In conclusion, the reason why HIV-1 can replicate in non-dividing cells is partly attributed to the ability of Vpr to promote the translocation of the PIC into the nucleus. However, the specific nuclear transport mechanism of Vpr is still needed to be further studied.

Vpr enhances the expression of unintegrated viral DNA, induces DNA double-strand breaks and promotes the binding of integrase to dsDNA. Unintegrated viral DNA, also called retroviral DNA, can express viral RNA and proteins under some circumstances and plays a significant role in HIV-1 pathogenesis. In the presence of Vpr, integration-defective HIV-1 mutants were found to express unintegrated viral DNA at almost same levels as that of wild-type HIV-1, while the expression of unintegrated viral DNA decreases 10- to 20-fold in the absence of Vpr based on in luciferase activity [129]. Besides, Vpr can stimulate the expression of Nef protein from unintegrated viral DNA [130].

Although Vpr does not exhibit endonuclease activity *in vitro*, Vpr promotes double strand breaks in the presence of isolated nucleic acids [131]. Furthermore, a truncated Vpr mutant without DNA-binding activity cannot induce double-strand break formation in isolated nucleic acids [131], suggesting that the break of double-strand depends on Vpr DNA-binding activity.

Vpr regulates viral transcription. Vpr can regulate viral transcription in a variety of ways. Firstly, once the viral DNA is integrated into host chromosomes, Vpr specifically activates HIV-1 LTR-directed transcription through the interaction with a TATA box of LTR [117]. Secondly, Vpr also directly trans-activates ubiquitous cellular transcription factor Sp1 to promote viral gene transcription [117]. Besides, Tat is involved into the use of cyclin T1 and cyclin -dependent kinase 9 (CDK9) to increase the level of LTR transcription. Vpr can facilitate the transcriptional activity of LTR by the interaction with Tat [118]. Thirdly, UNG2 was shown to inhibit LTR activity under the stimulation mediated by Tat trans-activator or TNF-alpha. Vpr counteracts UNG2 to indirectly upregulate the LTR activity [119]. Fourthly, the hypoxia inducible factor 1 alpha (HIF-1α) activates the HIV-1 promoter through the GC-rich binding domain in the LTR. Vpr activates the oxidative stress pathway and induces the expression of HIF-1α to upregulate transcription [120]. Fifthly, Vpr can activate NF-κB and AP-1 to stimulate HIV-1 LTR promoter based on the phosphorylation of transforming growth factor-β-activated kinase 1 (TAK1) induced by Vpr [121]. Sixthly, degradation of the histone deacetylases 1 and 3 (HDAC1 and HDAC3) was regarded as markers of transcription activation. Vpr reactivates the viral promoter by inducing the degradation of HDAC1 and HDAC3 in J-Lat cells latently infected with HIV-1 [122]. Currently, the process of transcription is well described and involves large molecular complexes based on nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB), spironolactone (Spl) and activator protein 1 (AP-1) [132]. Vpr may act as a cofactor within these complexes to directly or indirectly regulate the activity of HIV-1 LTR.

Vpr regulates viral translation. Vpr is primarily localized into the nucleus, but a significant fraction is also shown to concentrate at the nuclear envelope [113]. Furthermore, the interaction of Vpr with human nucleoporin CG1 (hCG1) in the nuclear pore complex was revealed by the yeast two-hybrid system suggesting that Vpr may interact with hCG1 to facilitate mRNA nuclear export [123].

Vpr is incorporated into virions during budding. The incorporation of Vpr in the virion could be mediated by NCp7 and p6 located in the C-terminal domain of the Pr55Gag polyprotein precursor. The two zinc fingers in NCp7 were required for the interaction with

Vpr [124]. Two motifs in p6 were shown to be required for the binding to Vpr [98,125–127,133].

5.2 Structure and function of Vpx

Conversely to Vpr, Vpx is found only to HIV-2 and a limited range of SIVs including SIVmac/SIVsmm and SIVrcm/mnd2. The structure of SIVmac Vpx is made up of a three-helical bundle and stabilized by a zinc finger motif (**Figure 20A**). Interestingly, the structure of SIVmac Vpx displays a structure similar as the structure of HIV-1 Vpr with three well defined α-helices [134] (**Figure 20B**). Vpx also shares similar functions to Vpr such as the translocation of the pre-integration complex to the nucleus and the neutralization of cellular antiviral factors such as SAMHD1 and Trim5α. However, there are differences between the two proteins. Vpx can induce the degradation of SAMHD1 with dNTP triphosphatase as a cellular enzyme, but not HIV-1 Vpr [135,136]. Besides, Vpx cannot mediate apoptosis.

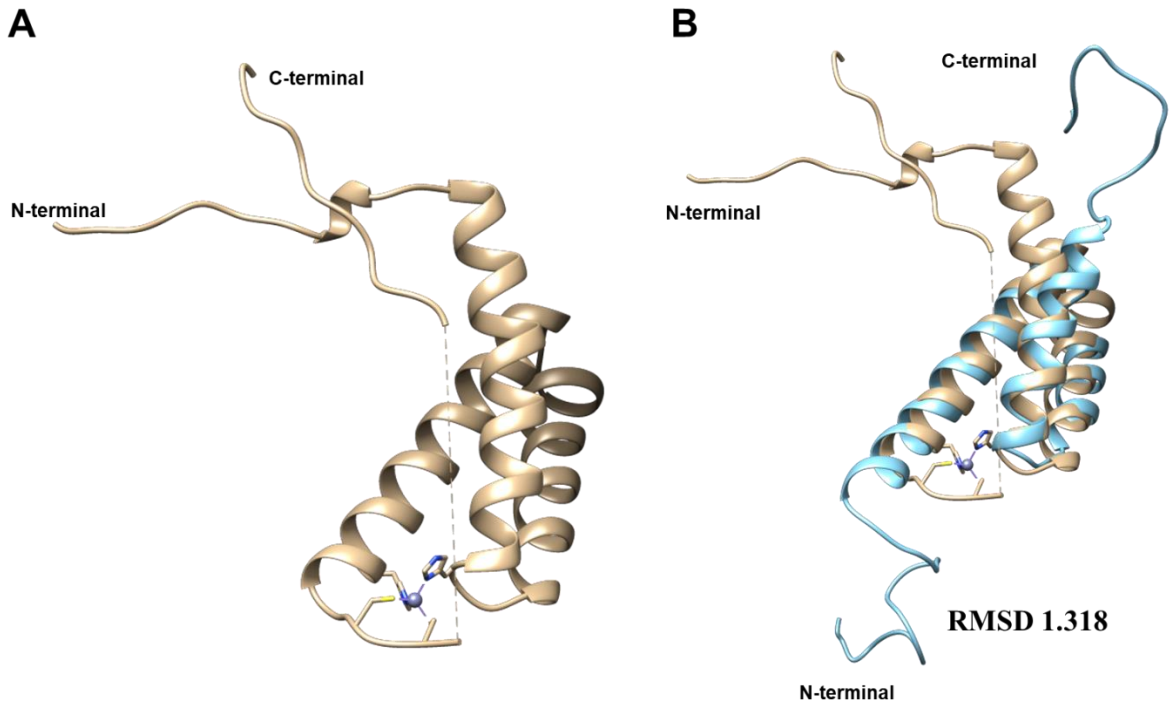


Figure 20. *Structures of SIVmac Vpx and HIV-1 Vpr. (A) Crystal structure of SIVmac Vpx (PDB 4CC9). Vpx is made up of a three-helical bundle and stabilized by a zinc finger motif*

in blue. Information about the structure at residues 91-99 is missing. (B) Superposition of structures of SIVmac Vpx (PDB 4CC9) and HIV-1 Vpr (PDB 1M8L).

5.3 Structure and function of p6

The domain p6 of the Pr55Gag precursor is involved in the late phase of the lentiviral replication cycle and mediates the incorporation of Vpr and Vpx into virion particles. The motif PTAP of p6 is mainly responsible for the recruitment of the cellular factor Tsg101 to facilitate the viral assembly. The second motif YPLASL in p6 is involved in the binding of another cellular budding factor, ALG-2-interacting protein X (ALIX), which allows the recruitment of the endosomal sorting complex required for transport (ESCRT) necessary for assembly and budding of the virus particle [137–139]. **Figure 21** summarizes all the important functional interacting regions of HIV-1 p6 [140].

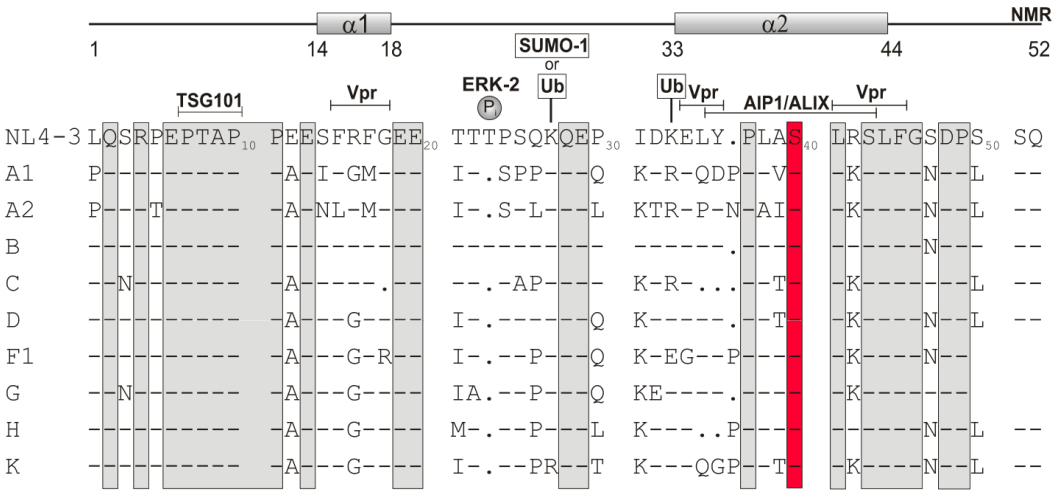


Figure 21. *Important functional regions or sites of HIV-1 p6 adapted from* [141]. *The highly conserved residue S40 in red is important for the viral replication. ERK-2 phosphorylation sites, ubiquitin and SUMO-1 attachment sites, and binding domains for TSG101, Vpr, and Alix are shown in grey.*

The structure of HIV-1 p6 in distilled water/TFE-d_2 (1:1 v/v) was determined by solution NMR and shows that p6 adopts a flexible helical structure with two short helices ranging from amino acids 14 to 21 and 33 to 47 (**Figure 22**) [142].

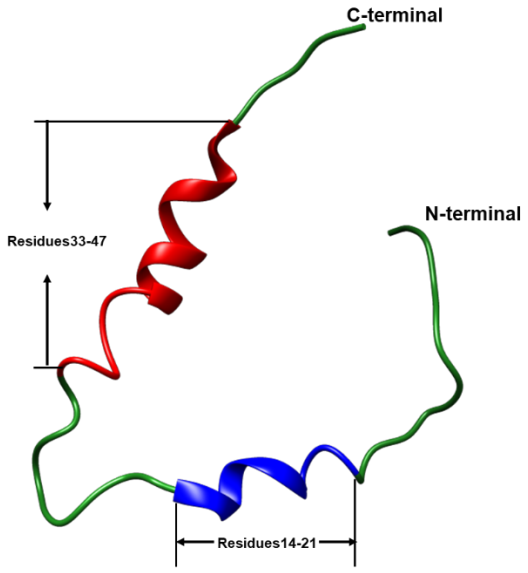


Figure 22. *Solution structure of HIV-1 p6 (PDB:2C55). Blue, helix I with residues 14-21; red, helix II with residues 33-47.*

6. Incorporation of Vpr / Vpx mediated by p6

P6 of the Pr55Gag precursor is involved in the late phase of the lentiviral replication cycle. The binding of p6 to TSG101 and ALIX allows the recruitment of the ESCRT complex necessary for assembly and budding of the viral particle [137–139]. The incorporation of Vpr and Vpx into virion particle is mediated by p6. Mutational analysis evidenced that the motif ^{41}LXXLF^{45} at the C-terminal region of HIV-1 p6 is required for Vpr incorporation [125,126]. However, a contradictory study based on mutational and deletion analyses shows that residues 1-23, which encompass a motif ^{15}FXFG^{18} at the N-terminus, were sufficient to drive HIV-1 Vpr incorporation [143], thus suggesting that the ^{41}LXXLF^{45} motif is dispensable. These results are contradicted by a recent alanine-scanning mutagenesis study that shows the importance of ^{15}FXFG^{18} and ^{41}LXXLF^{45} motifs for HIV-1 Vpr recruitment and that mutations in residues ^{34}ELY^{36} impair HIV-1 Vpr recruitment [144]. By contrast, a ^{17}DXAXXLL^{23} motif specific to SIVmac p6 was identified to be required for the incorporation of SIVmac Vpx [145]. As a conclusion, no consensus

residues and/or motifs essential for the Vpr-p6 and Vpx-p6 interactions have been identified to date.

Mutagenic analyses on HIV-1 Vpr showed that residues from 84 to 94 near the carboxy terminus are required for its incorporation into HIV-1 virions [82]. Subsequently, similar results showed that deletions of residues 73-96 or 78-96 of Vpr resulted in the impairment of virion incorporation of Vpr [146]. However, four Vpr mutants, L23F, E25K, A30F and I63F also exhibited reduction in Vpr incorporation. As amino acids L23, E25 and A30 are all included within a domain forming an amphipathic alpha helix (residues 16-34), this helix was considered important for virion incorporation of Vpr [146]. The subsequently resolved 3D structure of HIV-1 Vpr showed that residues L23, E25, and A30 are distributed into the first helix of HIV-1 Vpr [89,90,147], while I63 is located in the third helix. By contrast, residues 73-96 are distributed into the third helix and in the carboxy terminus. It seems that the first helix, third helix and the carboxy terminus are involved in the interaction with p6 for HIV-1 Vpr incorporation. Thus, a detailed investigation of the structure is necessary. Similarly, residues 73-89 of HIV-2 Vpx were shown to be important for virion incorporation [148]. Interestingly, a subsequent study using alanine scanning showed that the deletion of residues 73–89 does not abolish virion incorporation, but that its substitution significantly reduced incorporation into virus [149]. In general, there is no direct structural data on the incorporation of Vpr and Vpx and a detailed determination and a precise atomic-scale level investigation are necessary.

7. Assembly, budding and maturation

The HIV-1 Pr55Gag polyprotein mainly drives the assembly of HIV-1 particles at the plasma membrane. The Pr55Gag polyprotein is divided into four major domains: MA, CA, NC, p6 and two spacer peptides SP1 and SP2 [36,39,45]. Currently, the mechanism of assembly of Pr55Gag molecules is not fully understood. It is generally accepted that a myristic acid located at the N-terminal Glycine of MA anchors Pr55Gag precursor to the plasma membrane [49]. Following the binding to the plasma membrane, the multimerization of Pr55Gag forms a hexameric lattice, which promotes viral particle

assembly and leads to the budding and the release of nascent HIV virions, also called immature virions (**Figure 23**).

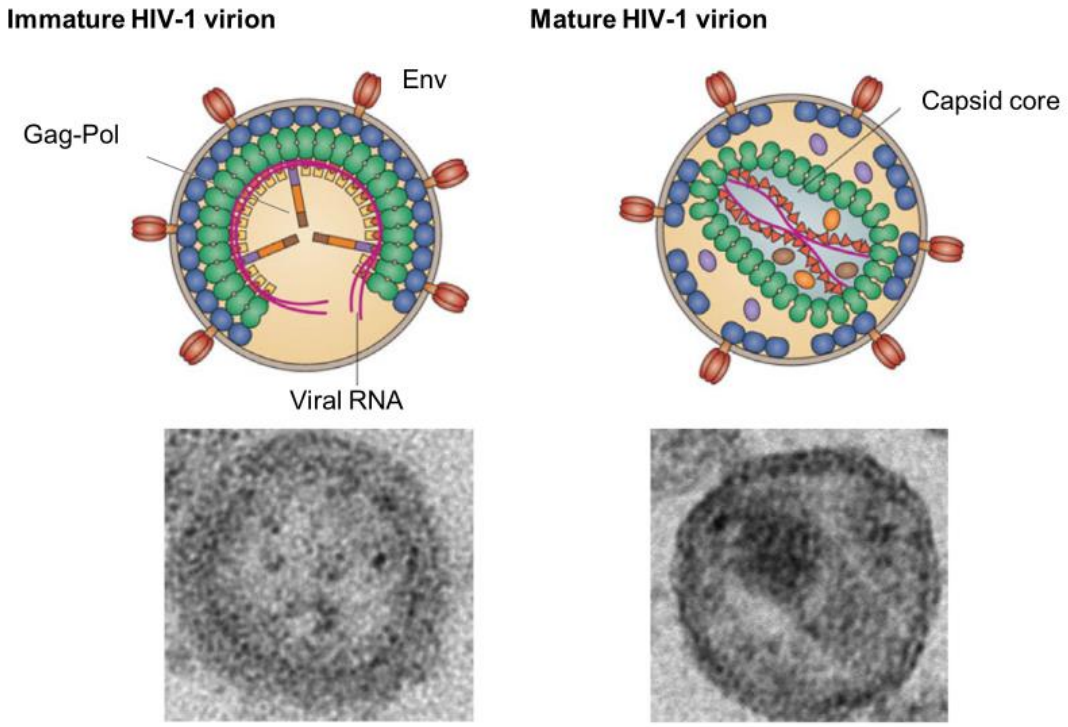


Figure 23. *Comparison of HIV-1 immature virion and mature virion from* [150].

The interactions of the p6 domain of Pr55Gag with the ESCRT directly drives the viral particle release [151,152] (**Figure 24**).

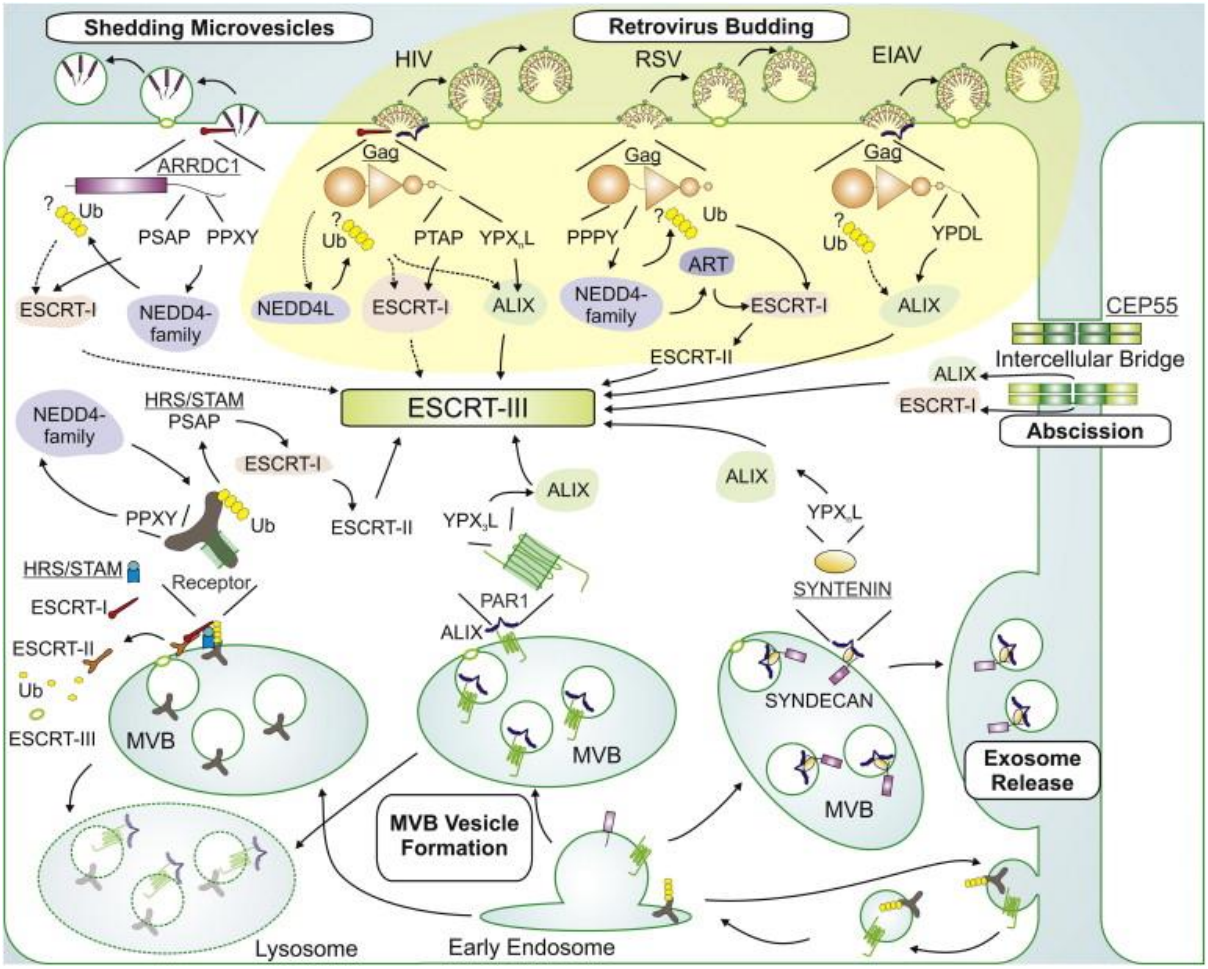


Figure 24. *The pathway of ESCRT recruitment* [151].

About 2,400±700 copies of Gag precursors were measured in the immature virus particle by scanning transmission electron microscopy and tomography [153]. The ratio of Gag to Gag-Pol was quantified to approximately 20:1 by dot blot assays [56,154]. After the release of the immature viral particles, PR protease is released from Gag-Pol by autocleavage [155]. Subsequently, Gag precursor is cleaved into MA, CA, NC, p6, SP1 and SP2 by PR, while Gag-Pol is cleaved into MA, CA, SP1, NC, SP2, p6*, PR, RT and IN [155]. This step called maturation is necessary to obtain new infectious viruses.

8. HIV Restriction factors

The antiviral response mediated by intracellular antiviral factors is the first line of defense against retroviral infection. Cellular restriction factors can play an inhibitory role in retrovirus infection. Many intracellular antiviral molecules were discovered in the last decade of studies on HIV. These restriction factors usually have four basic characteristics including (1) they significantly reduce the viral infection, (2) reverse and antagonistic of that of HIV, (3) rapid evolution due to the antagonistic mechanism, and (4) the relevance to innate immune system [156]. Here, we review several important restriction factors including the tripartite motif containing 5 α protein (TRIM5α), the apolipoprotein B mRNA editing enzyme catalytic polypeptide-like 3 (APOBEC3), tetherin, SAMHD1, human silencing hub (HUSH), Uracil DNA Glycosylase 2 (UNG2) and lysosomal-associated transmembrane protein 5 (LAPTM5). The studies of the interactions between restriction factors and HIV help us better understand the defense mechanisms of the host against an HIV infection, but also how HIV antagonizes these host restriction factors. All these interactions can be viewed as interesting targets for new HIV treatments.

8.1 TRIM5α

The TRIM family is composed by 70 members, including the HIV restriction factor TRIM5α. All members share similar domains, including the N-terminal RING finger that has a E3 ubiquitin ligase activity, one or two B-box domains, and a central coiled-coil domain [157]. In addition, the PRYSPRY sequence at the C-terminus was discovered in over 500 different TRIM proteins. It is responsible for the interaction with their target proteins such as HIV-1 capsid and antibody IgG [158,159]. Rhesus macaque TRIM5α was identified to be against HIV-1 in primate cells [160,161]. Interestingly, rhesus macaque TRIM5α does not display any antiviral activity against SIV or humans TRIM5α does not display any antiviral activity against HIV-1, while humans TRIM5α can effectively inhibit SIV suggesting that long-term co-evolution of the virus allowed adaptation to its host cell defenses [162,163]. Besides, the replacement of the residue R332 by a proline in the rhesus *macaque* TRIM5α has a similar level of anti-HIV activity [164].

TRIM5α can bind to HIV CA protein and inhibit lentivirus replication by blocking post-entry infection of HIV [165,166]. The antiviral activities of TRIM5 against viral infections are dependent on its multi-functional domains. The N-terminal RING finger and the C-terminal domain containing the PRYSPRY motif are closely related to inhibitory activity. Currently, it is agreed that the PRYSPRY domain interacts with the HIV-1 capsid lattice, which greatly increases the activity of UBC13/UEVA-dependent E3 in cooperation with the E3 ubiquitin ligase activity of the N-terminal RING finger [160]. Then, TRIM5 catalyzes the synthesis of K63-linked ubiquitin, which will activate the TAK1 (MAP3K7) kinase complex and stimulate AP-1 and activate the innate immune response [166] (**Figure 25**).

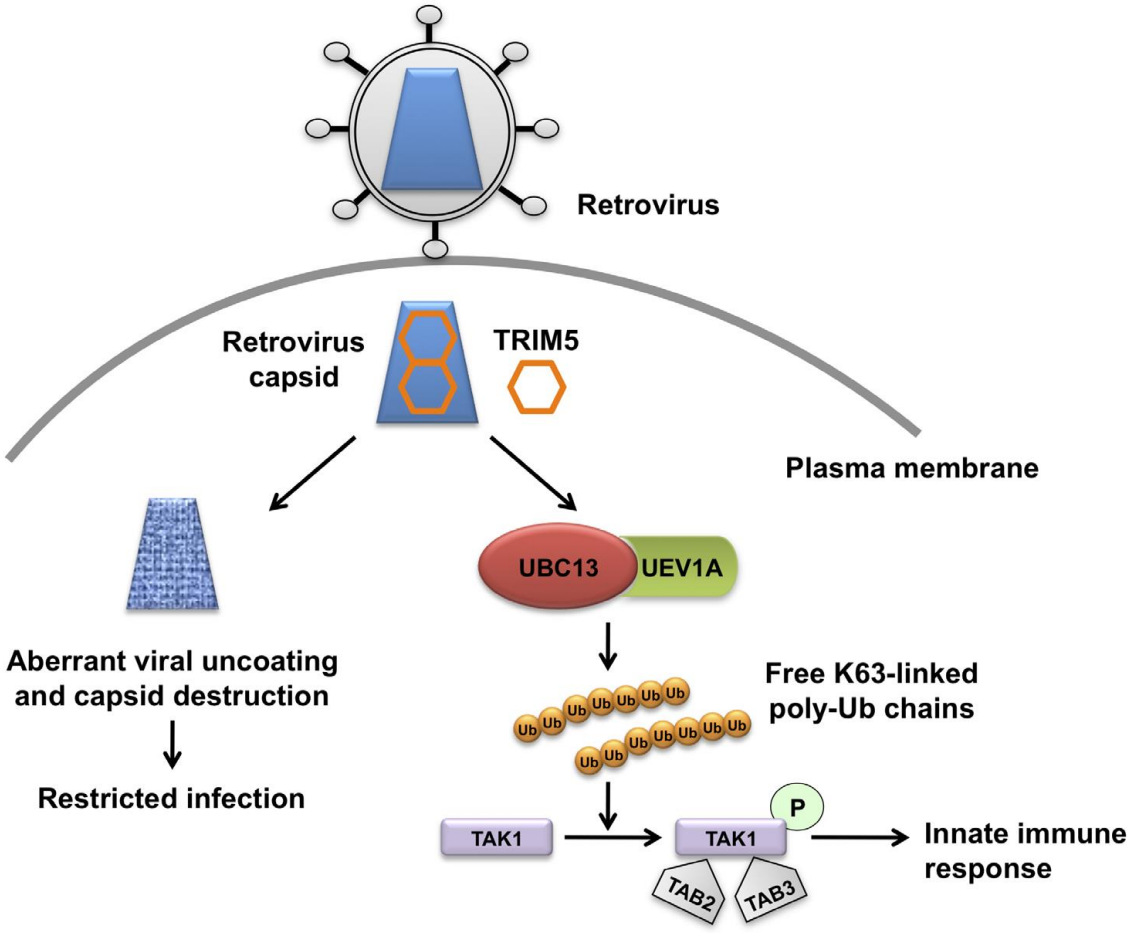


Figure 25. *The pathway of TRIM5α in retrovirus infection* [167].

8.2 APOBEC3

Apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3 (APOBEC3 or A3) belongs to the APOBEC family of cytidine deaminase enzymes with seven members (A, B, C, D, F, G and H). Structurally, APOBEC3 proteins share different sequences for the recognition of DNA and have a conserved Zn^{2+} catalytic motif HXE and PCXXC (**Figure 26A**). The N-terminal domain (NTD) of APOBEC3 is responsible for binding DNA/RNA or HIV NC, while the C-terminal domain (CTD) has a deaminase activity [168]. APOBEC3G is the first discovered restriction factor with anti-HIV activity. APOBEC3F and APOBEC3G have the strongest antiviral activity, while APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3D, and APOBEC3H also have antiviral activities [169–173]. APOBEC3G and APOBEC3F mediate HIV restriction by deaminase activity whichindividually converts G to A and blocks viral reverse transcription [174–176]. However, their sites for deamination are different in the process of viral reverse transcription. APOBEC3G can recognize deoxythymidine at position 0, while APOBEC3F recognizes deoxycytosine at position -1 and -2 [177,178]. Besides, APOBEC3G has a higher mutagenic potential to antagonize HIV infection [179–181].

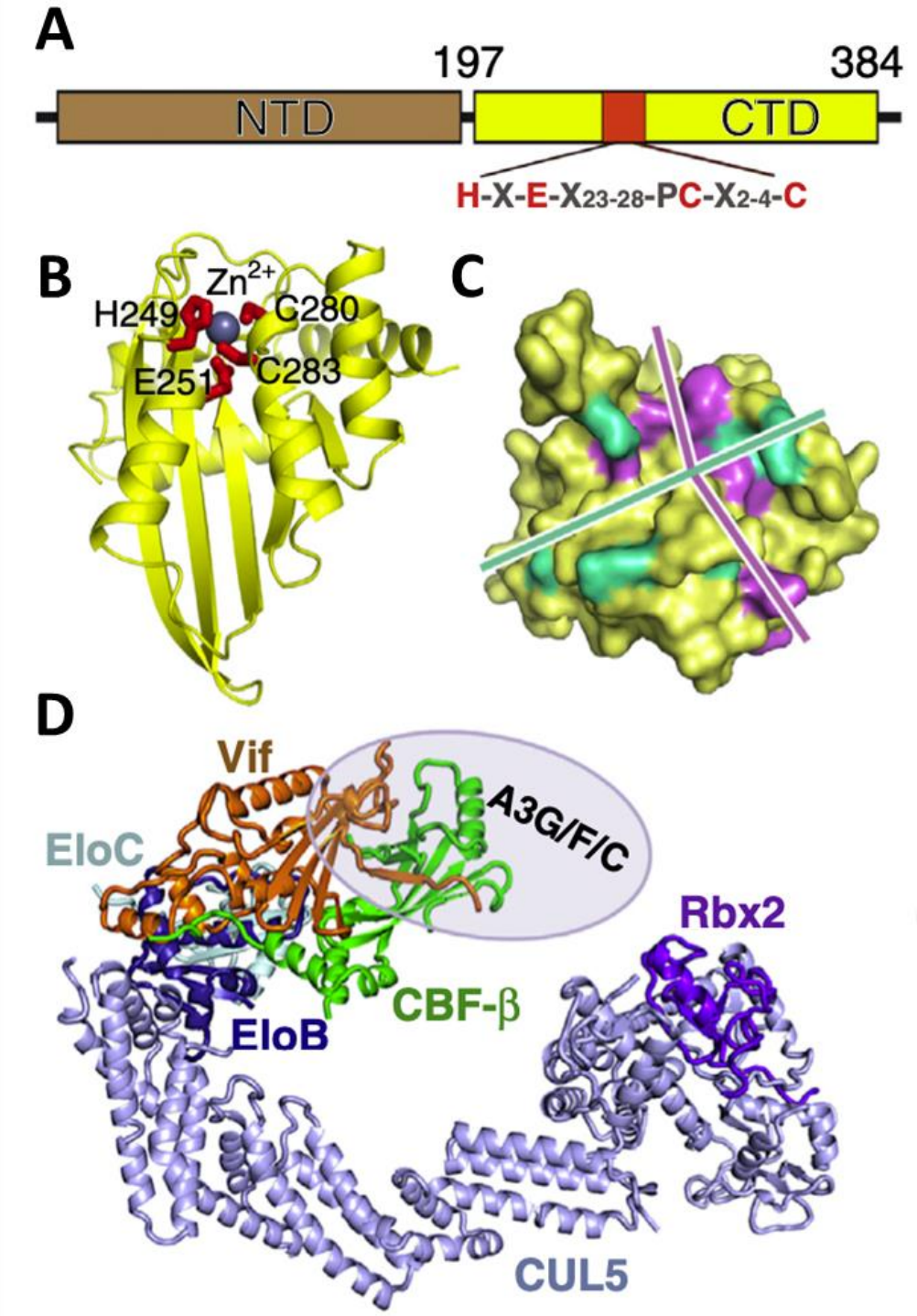


Figure 26. *Structure of APOBEC3 (A3G) (PDB 4N9F)* [182].

The viral Vif protein can specifically bind to APOBEC3F or APOBEC3G and mediates the polyubiquitination and subsequent degradation of APOBEC3F or APOBEC3G by the

26S proteasome [183,184]. HIV-1 Vif, in complex with CBFβ, EloBC and Cul5-NTD, hijacks the E3 ligase to degrade APOBEC3F/G/C (**Figure 26D** and **Figure 27**). However, the mechanism of APOBEC3 recruitment is unclear.

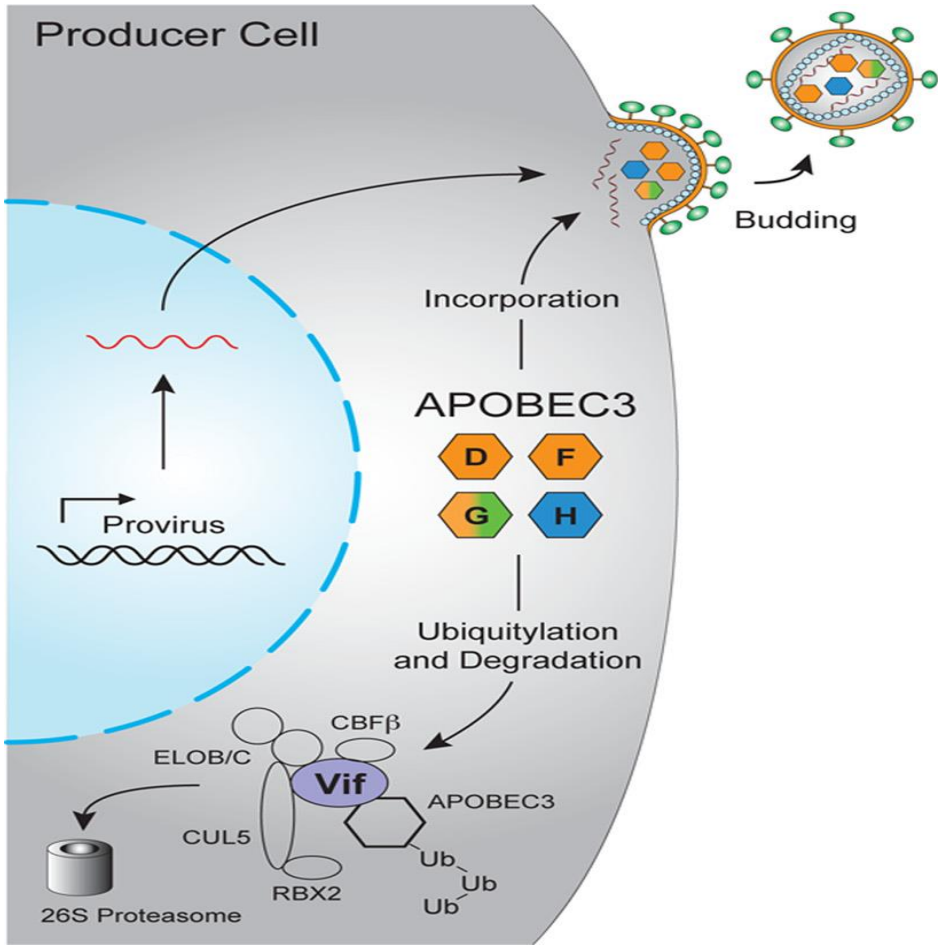


Figure 27. *Polyubiquitination and subsequent degradation of APOBEC3 mediated by Vif* [156].

8.3 Tetherin

Tetherin, also named BST-2, CD317 or HM1.24, can be induced by interferon-α (IFN-α) following infection with HIV. Tetherin is a type II transmembrane glycoprotein of 180 amino acids with a molecular weight ranging from 28 to 36 kDa [185]. The N-terminal contains a short, cytoplasmic N-terminal tail (CT) followed by a transmembrane domain,

a helical extracellular domain (EC) flanked by extended parallel coiled-coil domains, and a C-terminal anchoring glycosyl phosphatidylinositol (GPI) (**Figure 28**).

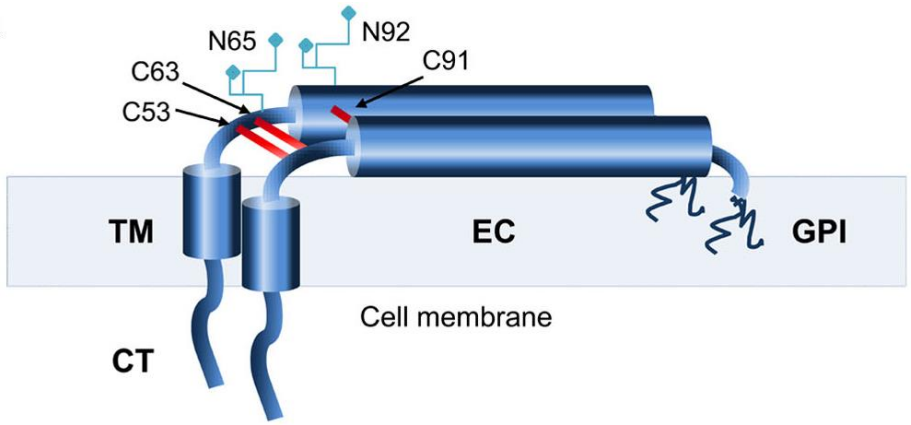


Figure 28. *Schematic representation of tetherin with the transmembrane domain (TMD), the helical extracellular domain (ECD) and the C-terminal glycophosphatidylinositol (GPI), from* [186]. *Tetherin is characterized by a short cytoplasmic N-terminus, then followed by an α-helical single-pass TM domain, an extended coiled-coil extracellular domain, and a C-terminal GPI anchor in the plasma membrane. For disulfide-bond formation, two N-glycosylation sites and three cysteine residues are noted.*

Tetherin can retain HIV viral particles on cell surface and block the release of viral particles in the late stage [187,188]. However, HIV Vpu protein can antagonize tetherin. It is agreed that Vpu antagonizes tetherin by reducing the quantity of tetherin on the cell surface [188]. Vpu recruits β-TrCP linker molecules to form ubiquitin ligase to degrade tetherin by lysosomal degradation (**Figure 29**).

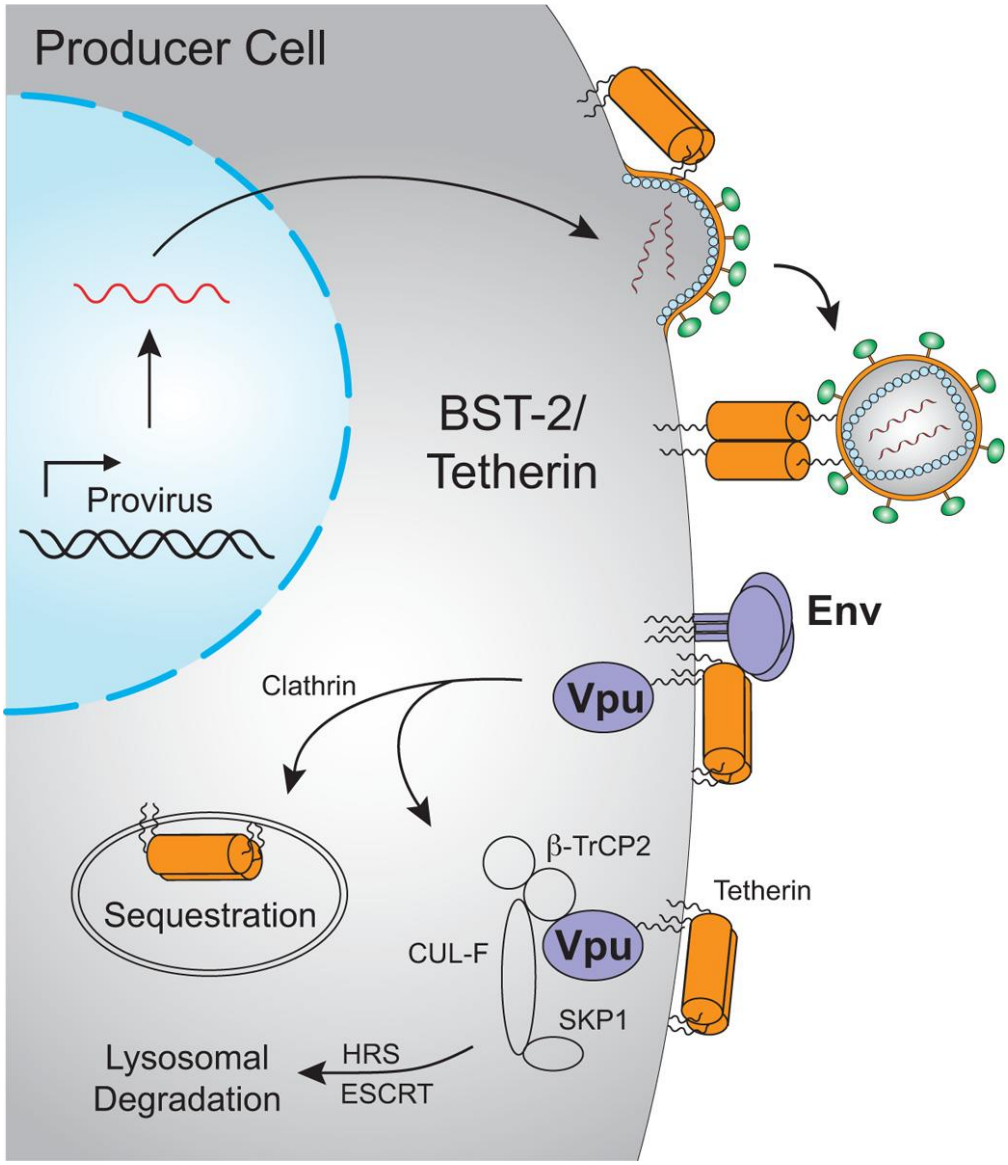


Figure 29. *Degradation of tetherin by Vpu* [156].

8.4 SAMHD1

SAMHD1 protein is a nuclear protein with deoxyribonucleoside triphosphate triphosphohydrolase (dNTPase) activity. It is mainly composed of two structural domains, the SAM domain and the HD domain. The SAM domain contains a nuclear localization signal [189,190]. The HD domain has dGTP triphosphatase [191–195]. The antiviral actions of SAMHDl were discovered by the study of HIV-1 and SIV lentiviral vectors. Dendritic cells (DCs) are regarded as important regulatory cells of the immune system and

good targets for gene therapy. However, the transduction efficiency of DCs is very low with (50 % - 90 %) [196]. Interestingly, when DCs are pretreated with SIV virus-like particles containing SIVmac Vpx, the transduction efficiency of the HIV-1/SIVmac lentiviral vectors can be greatly improved, which indicates that a unknown host restriction factor is antagonized by Vpx [197]. Subsequently, SAMHD1 was identified as a Vpx-interacting protein by mass spectrometry [135,136]. SAMHD1 is degraded by the interaction of Vpx with CUL4A-DDB1-DCAF1 (**Figure 30**). CUL4A-DDB1-DCAF1 is an E3 ubiquitin ligase complex that mediates protein ubiquitination and degradation by a proteasome-dependent pathway. Vpx in complex CUL4A-DDB1-DCAF1-Vpx can recruit SAMHD1 to promote the polyubiquitination and degradation of SAMHD1 (79, 80, 86).

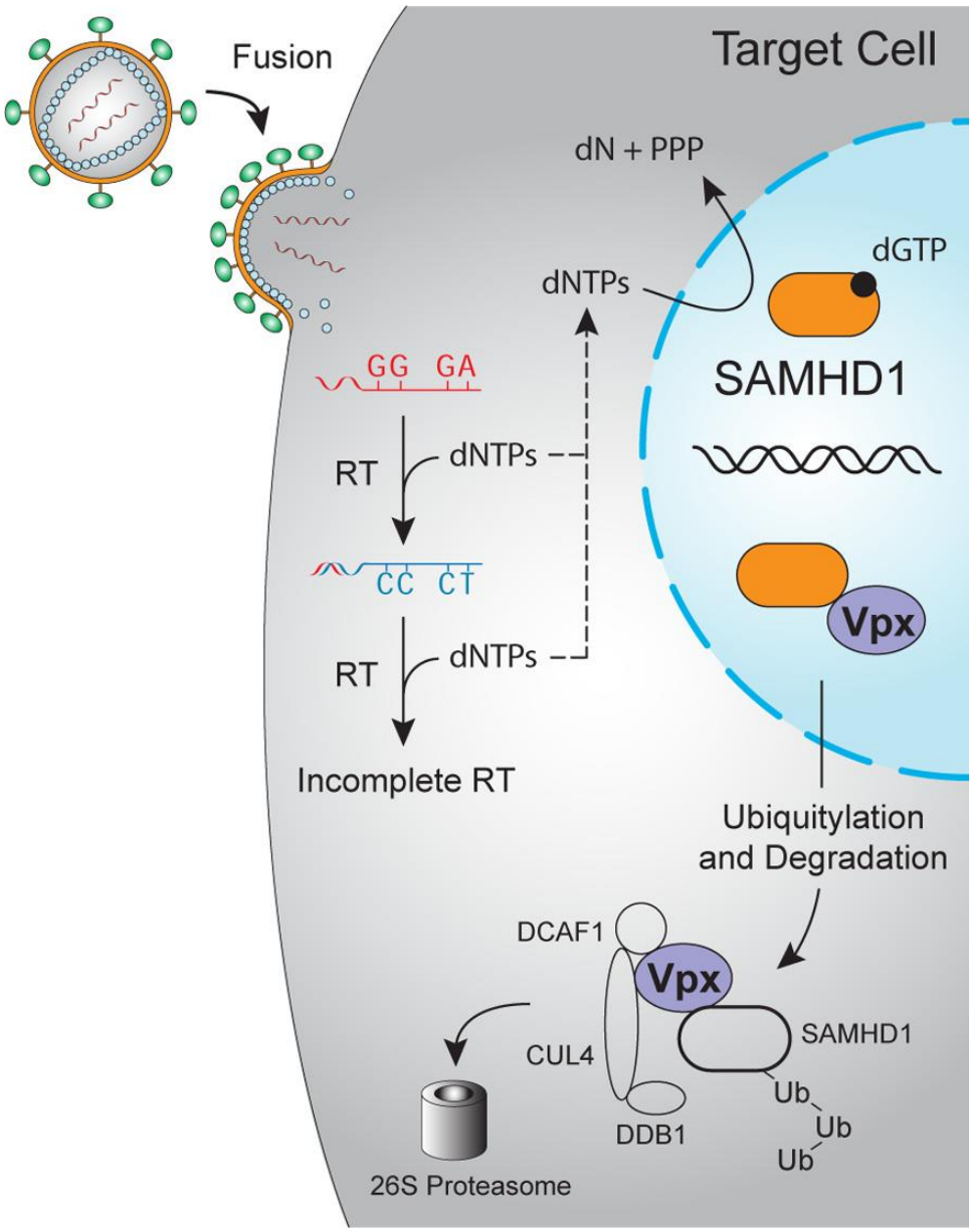


Figure 30. *Degradation of SAMHD1 by Vpx* [156].

8.5 HUSH

HUSH is a complex including the transgene activation suppressor (TASOR), the matrix metalloproteinase-8 (MPP8) and Periphilinis. It is a newly discovered restriction factor [198,199]. The HUSH complex is involved in position-effect variegation (PEV) and mediates the epigenetic silencing by recruiting the H3K9me3 methyltransferase SET

domain bifurcated 1 (SETDB1) in the nucleus [200]. Currently, many restriction factors acting at early and late stages of the viral replication cycle are identified in the cytoplasm but not in the nucleus. Thus, some researchers proposed a hypothesis that some restriction factors against HIV can act inside the nucleus. HUSH complex was identified by a screening of these restriction factors [198,199]. The antagonism of HUSH was mediated by the hijacking of DCAF1 by Vpx, using the same antagonism mechanism based on CUL4A-DDB1-DCAF1 as for SAMHD1. Here this mechanism leads to the degradation of TASOR within the HUSH complex. Interestingly, not all Vpr/Vpx can antagonize HUSH complex. Vpx from SIVrcm-GAB1 and Vpr from SIVagm.gri-677 do not degrade the human HUSH complex [199].

8.6 UNG2

Nuclear UNG2, composed of 313 amino acids, belongs to the classical large and highly conserved family of uracil-DNA glycosylases that exists in vertebrates, yeast, most bacteria and even in some viruses [201–203]. The main function of UNG2 is to remove the misincorporated uracils [204]. UNG2 is packaged into new HIV-1 virions by binding to HIV-1 Vpr or IN proteins [101,103,205,206]. The overexpression of UNG2 was found to inhibit HIV-1 replication by decreasing the HIV-1 LTR transcription efficiency [119]. Interestingly, UNG2 is essential for HIV cDNA stability and has positive effects on HIV replication by repairing uracilated viral cDNA [207]. Similarly, HIV-1 Vpr can degrade UNG2 by using the same mechanism involving CUL4A-DDB1-DCAF1 as for SAMHD1 and HUSH [208].

8.7 SERINC5

The serine incorporator 5 (SERINC5) cellular protein is incorporated into the viral particle and disrupts its ability to fuse with the host cell. The HIV-1 Env envelope glycoprotein is sensitive to restriction mediated by SERINC5, but the mechanism by which SERINC5 interferes with HIV-1 fusion is not understood. This effect is antagonized by the viral protein Nef. Tedbury et al. have shown that the incorporation of SERINC5 into the virion in the absence of Nef inhibits the formation of fusion pores between the virus and the cell

[209]. SERINC5 promotes the inactivation of Env glycoproteins and increases the exposure of conserved domains of gp41. Although the interaction between SERINC5 and Env has not been demonstrated, the authors conclude that SERINC5 blocks the fusion of HIV-1 at a stage before pore formation by inactivating the Env glycoproteins or by modifying their conformation. The increased sensitivity of HIV-1 to anti-gp41 antibodies and peptides suggests that SERINC5 also delays the folding of unfused Env trimers. Understanding the molecular mechanism by which SERINC5 interferes with HIV-1 fusion could help develop new strategies against HIV / AIDS.

8.8 LAPTM5

Lysosomal-associated transmembrane protein 5 (LAPTM5), a lysosomal transmembrane protein, can induce protein degradation by lysosomes. Lysosomes are ubiquitous organelles and degrade varieties of different substances such as nucleic acids, glycosaminoglycans, and proteins [210,211]. HIV-1 Vpr increases infection in macrophages but not in CD4+ T cells. A newly published results show that LAPTM5 is highly expressed in macrophages and that in the absence of LAPTM5, Vpr does not increase HIV-1 infection. On the other hand, in the absence of Vpr, silencing of LAPTM5 reproduces the same effect as Vpr on HIV-1 infection. Expression of LAPTM5 in primary CD4+ T cells increases HIV-1 infection, as in macrophages. Zhao et al. therefore proposed a molecular mechanism by which Vpr overcomes restriction by LAPTM5 in macrophages. LAPTM5 inhibits virion infectivity by transporting HIV-1 envelope glycoproteins to lysosomes for degradation and Vpr counteracts this restrictive effect of LAPTM5 by causing its degradation via DCAF1 [212]. Thus, the accessory protein HIV-1 Vpr, can promote infection by overcoming the inhibitory effects of restriction factors in the host cell. This molecular mechanism could provide a potential strategy for the discovery of new anti-HIV / AIDS therapies.

9. Cell cycle arrest and apoptosis

Besides, Vpr has two significant functions on cell cycle G2 arrest and apoptosis. The cell cycle includes interphase (G1, S, and G2 phases), mitotic phase (mitosis and cytokinesis), and G0 phase. A very significant function of Vpr is the arrest of the cell cycle at the G2 step [213–217]. This provides an advantageous environment for viral transcription. The specific mechanism of Vpr-induced cell cycle arrest is involved in a SLX4 complex and a specific cullin ubiquitin E3 ligase known as Cul4A-DDB1-DCAF1 E3 ubiquitin ligase complex [218–221]. Cul4A-DDB1-DCAF1 E3 ubiquitin ligase complex can be recruited by Vpx to degrade SAMHD1 or HUSH complex [198,222–224]. Vpr recruits phosphorylated PLK1 (pPLK1) and DCAF1 to the SLX4 complex, leading to the phosphorylation of EME1 and the activation of the SLX4 complex. The activated SLX4 complex results in cell cycle G2 arrest by cleavage of the replication forks. The activated SLX4 complex also leads to the process of HIV-1 DNA into escaping from innate immune sensing. In addition, Vpr usurps the CUL4A-DDB1-DCAF1 E3 ubiquitin ligase complex leading to the ubiquitination of MUS81. Subsequently, the decrease of MUS81 results in the inability of cells to resolve ultra fine bridges (UFBs). UFBs are unresolved DNA interlinks that persist into anaphase generate DNA structures. Thus, the decrease of MUS81 may contribute to G2/M arrest [225]. However, the effects of the MUS81-UFBs pathway on G2/M arrest is unclear. In conclusion, Vpr creates a favorable environment which contributes to viral transcription enhancement by induction of a cell cycle G2 arrest.

The induction of apoptosis is the second significant function of HIV-1 Vpr. Apoptosis is triggered via two fundamentally distinct pathways, namely the extrinsic and intrinsic pathways [226]. Currently, the mechanism of Vpr-mediated apoptosis remains unresolved, and several proposed mechanisms are under investigation. It is widely accepted that Vpr is involved into the binding to adenine nucleotide translocator (ANT), a mitochondrial reverse transport carrier. The interaction of Vpr with ANT disrupts the mitochondrial membrane and releases membrane gap proteins such as cytochrome c and apoptosis-inducing factor finally resulting in forming apoptotic bodies and inducing the activation of effectors and cell programmed death [227]. However, because of its high hydrophobicity, Vpr can directly interact with mitochondria membrane to disrupt it and induce apoptosis.

Besides, Vpr can disrupt the mitochondria stability to induce apoptosis by inducing the mitochondrial anti-interference protein HAX-1 out of its normal position. In conclusion, current results only showed correlation between Vpr and apoptosis pathway. Vpr-mediated apoptosis is still to be clarified.

10. HIV and Vpr inhibitors

Given that Vpr can increase HIV replication, Vpr inhibitors are supposed to be invaluable target against AIDS. Thus, some Vpr inhibitors have been found. A few kinds of natural plants from Myanmar were identified to display anti-Vpr activities. For example, two Myanmar medicinal plants used in folk medicines, *Kaempferia pulchra* rhizomes and *Picrasama javanica* bark, were identified to be active against Vpr [228,229] (**Figure 31**). Two compounds, isopimarane diterpenoids and picrasane quassinoids, were identified to be effective for inhibiting Vpr activity involved in apoptosis. Crude extracts of *swertia chirata* plant, bis-iridoid and iridoid glycosides from *Picrorhiza kurroa* collected in Myanmar were also found to inhibit the expression of Vpr in Hela cells [194,195].

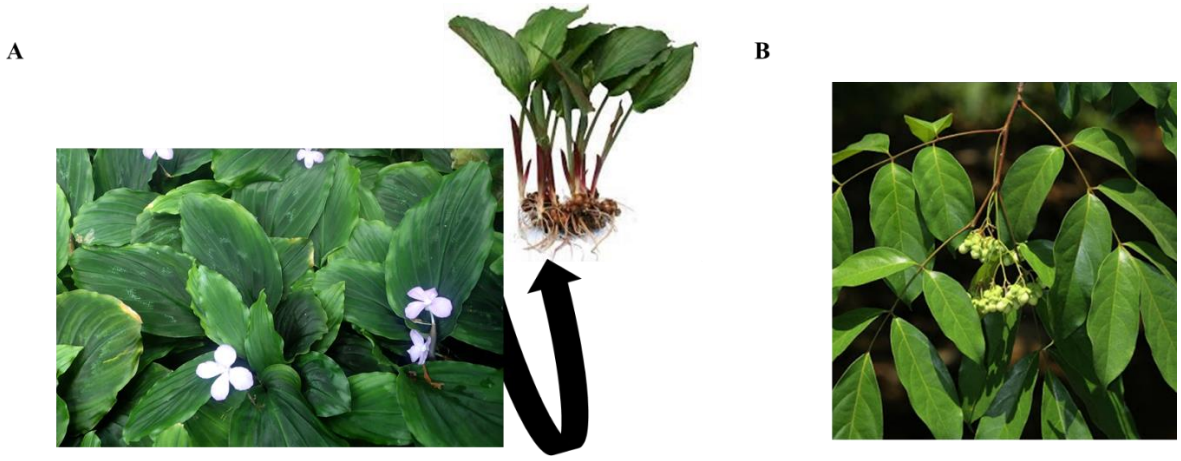


Figure 31. *Pictures of Kaempferia pulchra (A) and Picrasama javanica (B).*

Apart from natural medicinal plants, flavonoids, widely spreading in the plant kingdom and commonly consumed in the form of fruit, vegetables and drinks such as wine and tea, were identified to show anti-HIV activity [232]. Quercetin, a flavonoid, can inhibit Vpr-induced cell cycle arrest [233]. Interestingly, fumagillin, isolated from fungal metabolites, also inhibits cell cycle arrest activity induced by Vpr [234]. Damnacanthal, a component of noni, is identified to be a Vpr inhibitor by inhibiting Vpr-associated cell death with no effect on cell cycle arrest [235]. Besides, a high-throughput screening (HTS) system was developed to screen of Vpr inhibitors in a library of compounds in the RIKEN NPDepo (**Figure 32**) [236].

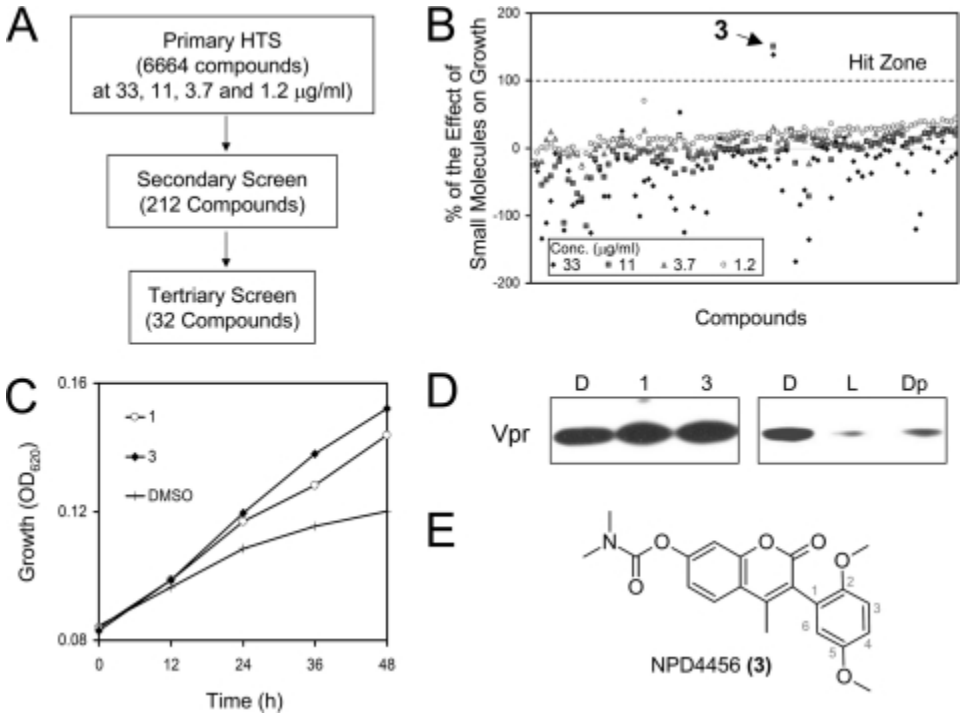


Figure 32. *High-throughput screening of Vpr inhibitors* [236]*. (A) a workflow of the screening process. (B) inhibitor 3 shown in arrow. (C) growth curves of hit compound 1, compound 3, and DMSO. (D) western blotting assays of Vpr expression. Lane D, DMSO; lane 1, compound 1; lane 3, compound 3 (87 μm), lane L, leucine; lane Dp, 4,6-diamidino-2-phenyl-indole (DAPI). (E) structure of the hit compound 3.*

Objectives of the Thesis

The lentiviral accessory proteins Vpr and Vpx are specifically incorporated into virions by direct interaction with $p6^{gag}$ to participate to the early stages of the viral replication. The incorporation of primate lentiviral accessory proteins Vpr/Vpx is essential for the viral replication in resting cells, the translocation of the pre-integration complex to the nucleus and for the neutralization of the cellular antiviral factors. Our objectives are to study the structure and the interactions of Vpr / Vpx with their partner protein p6 in the presence of DPC micelles that mimic the hydrophobic environment of the membrane. In addition, we have developed a very simple protocol intended to replace the fully protonated DPC micelles used to solubilize Vpr and Vpx, and that prevent recording nice NMR spectra, by perdeuterated DPC micelles in a cost-effective manner. Our objectives are divided fourparts: (1) optimize protocol for the elimination of fully protonated DPC micelles while maintaining native conditions for the protein, (2) elucidate structures of HIV-1 Vpr, SIVmac Vpx and their p6 partners in the presence of DPC micelles, (3) characterize at the atomic level the interactions HIV-1 Vpr-p6 and SIVmac Vpx-p6, (4) test the cross interactions between SIVmac Vpx and HIV-1 p6.

1. Indirect replacement of non-deuterated DPC by deuterated DPC in native conditions.

Lipidic detergents are alternatively employed to successfully solubilize Vpr and Vpx. Unfortunately, the large number of protons contained in these detergents complicates our NMR spectra analysis. The effectiveness of detergent removal challenges the downstream determination of protein-protein or protein-ligand interaction for the identification of new drugs and their targets since some detergents such as dodecylphosphocholine (DPC) and sodium dodecyl sulphate (SDS) can bind to detergent-soluble proteins (DSPs) so tightly that they cannot be removed. Therefore, we developed a new efficient and economical protocol to remove lipidic detergents while maintaing the native structure of the protein.

2. Structures of HIV-1 Vpr, of SIVmac Vpx and of their p6 partners in the presence of DPC micelles.

The structures of HIV-1 Vpr and SIVmac Vpx are similar and well characterized with three defined α-helices [89,90,134]. However, Vpr and Vpx have cytostatic and cytotoxic effects in the yeast cells [237], and HIV-1 Vpr can induce apoptosis in huan cells [131,238]. In addition, HIV-1 Vpr can inhibit prokaryotic cell growth [239]. The cytostatic and cytotosic effects directly lead to the difficulties in protein production. Thus, at first step, we set up a protocol allowing production and purification of milligrams of Vpr and Vpx, which is a prerequisite to the NMR studies.

The previous structure of HIV-1 Vpr was solved by NMR in aqueous buffer containing 30% acetonitrile, at low pH [89]. The acetonitrile is an organic solvent, very different from the physiological conditions. The identified functions of HIV-1 p6 usually occur close to the plasma membrane [133,240,241]. Previous studies conducted in the group demonstrated the existence of an interaction in vitro between p6 and HIV-1 Vpr in membrane conditions [133,240]. Here, DPC micelles are employed to mimic a membrane environment. The structures of HIV-1 Vpr, SIVmac Vpx, and of their p6 protein partners will be investigated in the presence of DPC micelles.

3. Characterization at the atomic level of HIV-1 Vpr-p6 and SIVmac Vpx-p6 interactions in the presence of DPC micelles

The incorporation of Vpr and Vpx into virion particle is mediated by $p6^{gag}$. However, the mechanistic of incorporation of Vpr/Vpx by p6 has been controversial. For HIV-1 Vpr, deletion and mutational analyses suggest that residues 84-94 near the carboxyl terminus are important for the incorporation of Vpr into HIV-1 virions (8,11). However, mutations in the N-terminal region (L23F, E25K, and A30F) also reduce the protein incorporation [146]. These mutations located in the first alpha-helix of HIV-1 Vpr suggest that the

regions of the first amphipathic helix are involved in the incorporation into the virion. In addition, the mutant I63F at third helix was also identified to impair the incorporation of HIV-1 Vpr into virions. No consensus has been thus established yet for the residues of Vpr involved in its encapsidation. For SIVmac Vpx, no region was identified as being responsible for the interaction with SIVmac p6. On the contrary, for HIV-2 Vpx, residues 73-89 were identified as important for Vpx incorporation mediated by p6 [148], but not essential [149], which indicates that some important residues are still unidentified. Thus, a detailed investigation of the molecular determinants for the incorporation of Vpr and Vpx by p6 is necessary.

The $p6^{gag}$ protein is involved in the late phase of lentiviral replication cycle. Previous studies showed that HIV-1 p6 interacts with the cytoplasmic membrane and that the interaction with its partners occurs under the cytoplasmic membrane [133,240,241]. Mutational analysis shows that the motif ^{41}LXXLF^{45} at the C-terminal region of HIV-1 p6 is required for Vpr incorporation [125,126]. However, a contradictory studies of mutational and deletion analyses show that residues 1-23 including a motif ^{15}FXFG^{18} at N-terminal region of HIV-1 p6, not the motif ^{41}LXXLF^{45}, was sufficient for HIV-1 Vpr incorporation [143]. Recent alanine-scanning mutagenesis studies show the importance of both motifs for HIV-1 Vpr recruitment [144]. In addition, the mutant in residues ^{34}ELY^{36} impaired HIV-1 Vpr recruitment [144]. In comparison with HIV-1 p6, a unique motif ^{17}DXAXXLL^{23} in SIVmac p6 was identified to be required for the incorporation of SIVmac Vpx [145].

Our objective is to characterize the interaction between HIV-1 Vpr / SIVmac Vpx and p6 at the atomic level. These data could help us to identify new targets to block the interaction between Vpr / Vpx and p6 and disrupt viral replication. These studies would be complementary to those currently being carried out on treatments targeting viral enzymes and would make it possible to reduce the viral load in cells at different stages of the replication cycle.

4. Cross interactions between SIVmac Vpx and HIV-1 p6

Sequence analyses revealed that SIVmac p6 and HIV-1 p6 both have a conserved motif LXXLF. SIVmac Vpx and HIV-1 Vpr have a similar structure with three well defined α-helices. Thus, we hypothesize that the motif LXXLF of HIV-1 p6 could interact with SIVmac Vpx, and we will investigate this possibility.

Materials and methods

The “Materials and Methods” includes all the expression, synthesis and purification protocols for the various proteins studied in this thesis. Besides, the NMR techniques are also introduced.

1. Expression and purification of HIV-1 Vpr and SIVmac Vpx.

The plasmid pET-11d-SIVmac Vpx was kindly provided by Andrea Cimarelli, Centre International de Recherche en Infectiologie (CIRI), Université Lyon 1. Briefly, the fragment for expression in *E.coli* of SIVmac Vpx with the His_6 at the N terminal was amplified by PCR. Subsequently, the PCR products were cloned into the pET-11d with ampicillin resistance via the restriction enzymes *Hin*dIII-*Bam*HI. Finally, the plasmid was confirmed by sequencing. Using the same vector and enzyme sites, HIV-1 Vpr with a N-terminal hexa-histidine (His_6) affinity tag was ordered from GenScript.

1.1 Plasmid transformation

Take one eppendorf with 100 µL of chemically competent Rosetta 2 (DE3) *E. coli* cells and allow the eppendorf to thaw on wet ice. Then, add 1 µL (about 50 ng) of expression plasmid pET11d-HIV-1 Vpr and pET11d-SIVmac Vpx directly into the competent cells, respectively. Mix them gently and incubate the mixture on wet ice for 30 min. Transfer and keep the mixture at 42 °C in a water bath for 1 min. Move and keep the mixture onto wet ice for 3 min. Add 200 µL of LB medium into the eppendorf and mix by gentle tapping. Transfer and incubate on a shaker incubator at 180 r.p.m. and at 37 °C for 45 min. Pipette 100 µL of the transformed cells suspension and spread them on an LB agar plate (100 µg/mL ampicillin) using a sterile spreader. Incubate the plate at 37 °C for 16 h.

1.2 Cell culture

Eleven single colonies of pET-11d-SIVmac Vpx and four single colonies of pET-11d-HIV-1 Vpr from the LB agar plate were individually precultured each into 5 mL of modified M9a medium (see **Table 2** for composition) [242] at 37 ℃ overnight. An appropriate

volume of preculture was diluted to an OD_{600} of 0.3 into a fresh 10 mL of modified M9a medium, and then inoculated at 37 °C for 60 min until the OD_{600} value of the new culture reached 0.6. The culture was harvested by centrifugation and the pellets were resuspended with 10 mL of secondly modified M9b medium (**Table 2**). IPTG was added to a final concentration of 0.5 mM and the culture was incubated at 18 °C for 20 hours. The expression levels were assessed by SDS-PAGE. The best cells with higher expression level were stored as glycerol stocks at -80 °C and used for the large-scale expression.

For $^{15}N/^{13}C$ labeled protein expression, 100 µL of glycerol stock of *E. coli* Rosetta 2 (DE3)-pET11d SIVmac Vpx and pET-11d-HIV-1 Vpr were each inoculated into a 200 mL baffled glass erlenmeyer flask with 50 mL modified M9a medium and 100 µg/mL ampicillin. The precultures were grown overnight at 37 °C. When the OD_{600} reached a value of 0.3, the precultures were each transferred into a 2.5 L baffled glass erlenmeyer flask filled with 1.0 L of modified M9a medium. When the OD_{600} values of the cultures reached 0.6, the cells were harvested at 6000 g for 30 min at 18 °C and re-suspended in 1.0 L of the secondly modified M9b medium containing only 1 g of ^{13}C-glucose and 0.5 g of $^{15}NH_4Cl$. The expression was induced by adding IPTG to a final concentration of 0.5 mM and the cells were cultured for 20 h at 18 °C. The cells were finally harvested.

The unlabeled proteins were produced using the same methods. The differences are that 1 g of ^{13}C-glucose and 0.5 g of ^{15}N-NH_4Cl in secondly modified M9b media were individually replaced by unlabeled 1 g of glucose and 0.5 g of NH_4Cl.

Table 2. Compositions of the modified M9a and secondly modified M9b medium.

Modified M9a media		Secondly modified M9b media	
K_2HPO_4	19.0 g	K_2HPO_4	19.0 g
KH_2PO_4	5.0 g	KH_2PO_4	5.0 g
Na_2HPO_4	9.0 g	Na_2HPO_4	9.0 g
K_2SO_4	2.4 g	K_2SO_4	2.4 g
Glucose	10 g	^{13}C-Glucose	1 g
NH_4Cl	5 g	$^{15}NH_4Cl$	0.5 g
1M $CaCl_2$	200 μL	1M $CaCl_2$	200 μL
1M $MgSO_4$	1 mL	1M $MgSO_4$	1 mL
Vitamins	0.5 x (final concentration)	Vitamins	0.5 x (final concentration)
Trace 100x	0.5 x (final concentration)	Trace 100x	0.5 x (final concentration)

1.3 HIV-1 Vpr and SIVmac Vpx purification.

2. Chemical synthesis of HIV-1 p6, SIVmac p6 and the peptide SIVmac p6$^{16-28}$.

2.1 The peptide SIVmac p6$^{16-28}$

The chemically synthesized peptide SIVmac p6$^{16-28}$ with 98% purity was purchased from ProteoGenix and directly dissolved for NMR experiments.

2.2 SIVmac p6 and HIV-1 p6

3. Nuclear magnetic resonance (NMR) used for protein structures and protein-protein interactions studies.

It has been more than 50 years since solution NMR was used for the study of the biological macromolecules [244]. The first solution structure of the proteinase inhibitor IIA from bull seminal plasma was solved by solution NMR in 1985 [245], which paved the way for NMR analysis of biological macromolecules. NMR, together with X-ray crystallography and cryo-electron microscopy (cryo-EM), is main research techniques for the structural biology of protein complexes. Besides, NMR has the unique ability to provide information on the structure and dynamics of proteins and allows the structural determination in solution, at physiological conditions and temperature, while it is limited by the molecular weight of proteins. The determination of the protein structure by NMR is usually divided into four main steps including (1) sample preparation, (2) NMR measurements, (3) resonance assignments, and (4) structure determination from NMR data [246].

^{15}N and/or ^{13}C labeled proteins are usually essential for molecular weights above 12 kDa. To obtain ^{15}N and/or ^{13}C labeled proteins, chemically synthesized labeled peptides/proteins and various expression systems like bacteria, yeast, mammals and cell-free expression *in vitro* are employed. For large proteins with high molecular weights above 30 kDa, triple ^{15}N, ^{13}C, ^{2}H-labelling protein is required, and the proteins are expressed in minimal medium supplemented with D_2O instead of H_2O. By deuterating the protein, most signals from ^{1}H atoms (protons) can be removed, and the relaxation properties are improved again. Subsequently, the obtained protein is purified by chromatographic methods including for example affinity binding, size exclusion and/or ion exchange. Finally, the protein is dissolved in water with some salts or all kinds of buffer to ensure a near physiological condition. A slightly acidic pH from 3 to 5 is recommended [246,247].

For measurements by NMR, the one-dimensional ^{1}H NMR spectrum is crowed due to a large number of protons signals. Thus, two-dimensional and three-dimensional spectra are

best achieved with ^{15}N and/or ^{13}C labeled proteins to overcome chemical shift overlaps. So far, the simplest method of protein assignment involves the backbone and sidechain assignment experiments. The backbone assignment experiments are mainly composed of CBCANNH, CBCA(CO)NNH, HNCA, HN(CO)CA, HNCO and HN(CA)CO spectra, while sidechain assignments are derived from CC(CO)NNH, HBHA(CO)ONH and H(CCCO)NNH spectra (**Figure 33**). In addition, the nuclear Overhauser effect spectroscopy (NOESY) allows the determination of distance restrains for structure calculations.

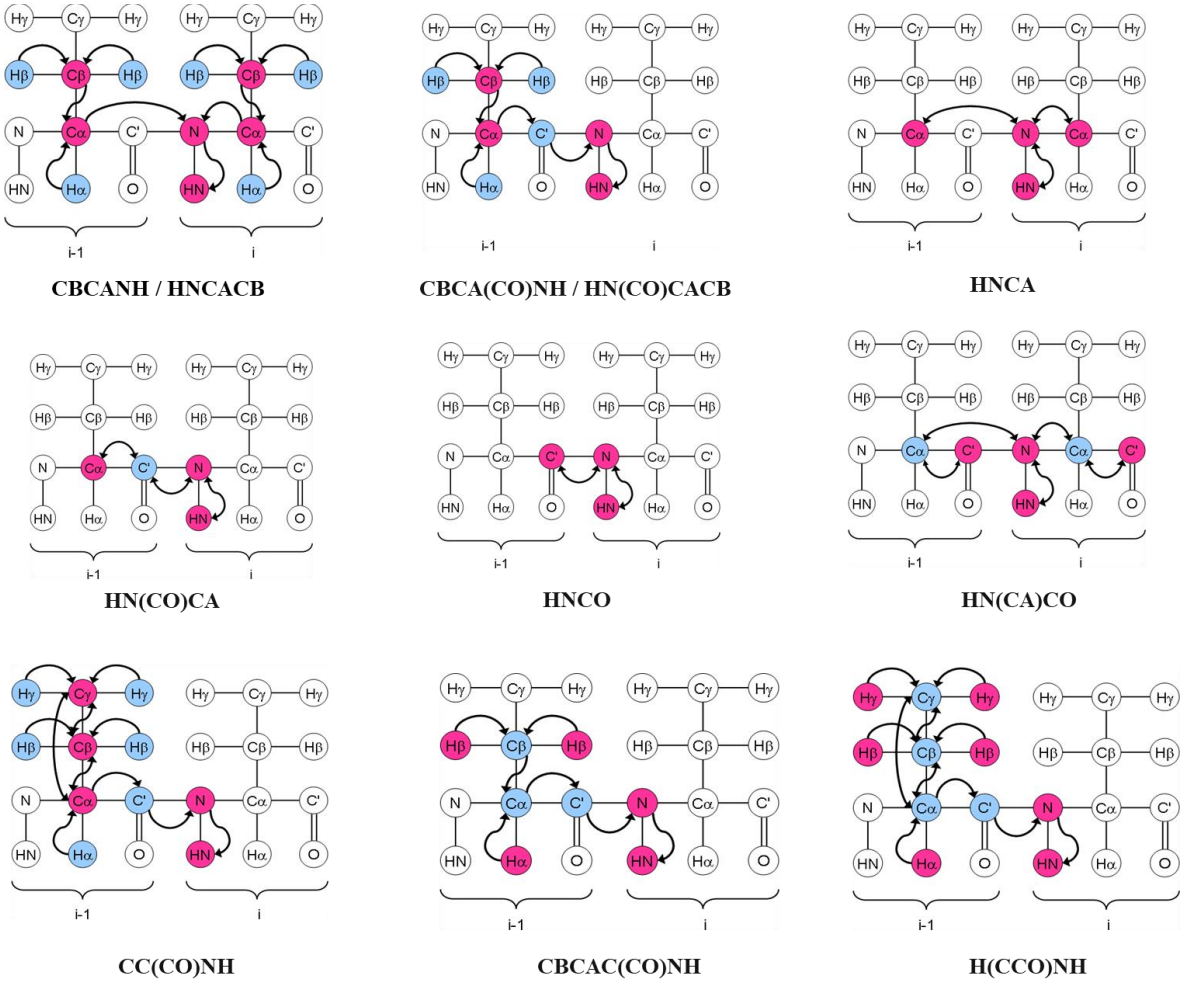


Figure 33. *Backbone and sidechain assignment experiments measured by NMR (from the website https://www.protein-nmr.org.uk/.)*

3.1 Triple Resonance Backbone Assignment

The backbone assignment of labeled ^{15}N / ^{13}C proteins is mainly based on the CBCANNH and CBCA(CO)NNH experiments. As shown in **Figure 34**, the CBCANNH experiment makes it possible to correlate each amide group of a residue with the chemical shifts of its own Cα and Cβ and those of the preceding residue. In addition, the CBCA(CO)NNH experiment only correlates the amide group of a residue with the Cα and Cβ chemical shifts of the preceding one. Similarly, the HNCA and HN(CO)CA experiments give access to the CO chemical shifts (**Figure 35**).

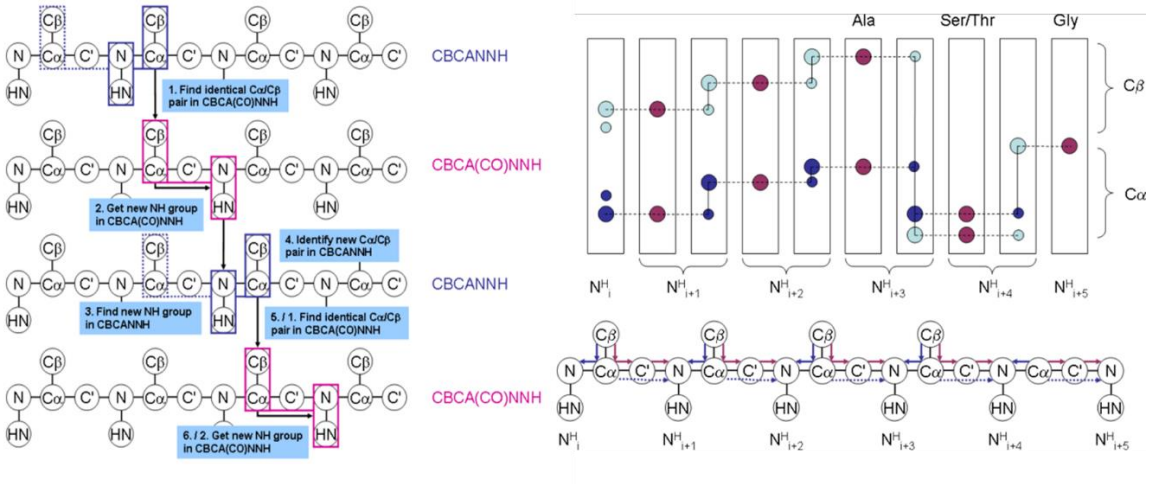


***Figure 34**. Backbone assignment based CBCANNH and CBCA(CO)NNH (from [https://www.protein-nmr.org.uk/solution-nmr/assignment-theory/triple-resonance-backbone-assignment/)](https://www.protein-nmr.org.uk/solution-nmr/assignment-theory/triple-resonance-backbone-assignment/).*

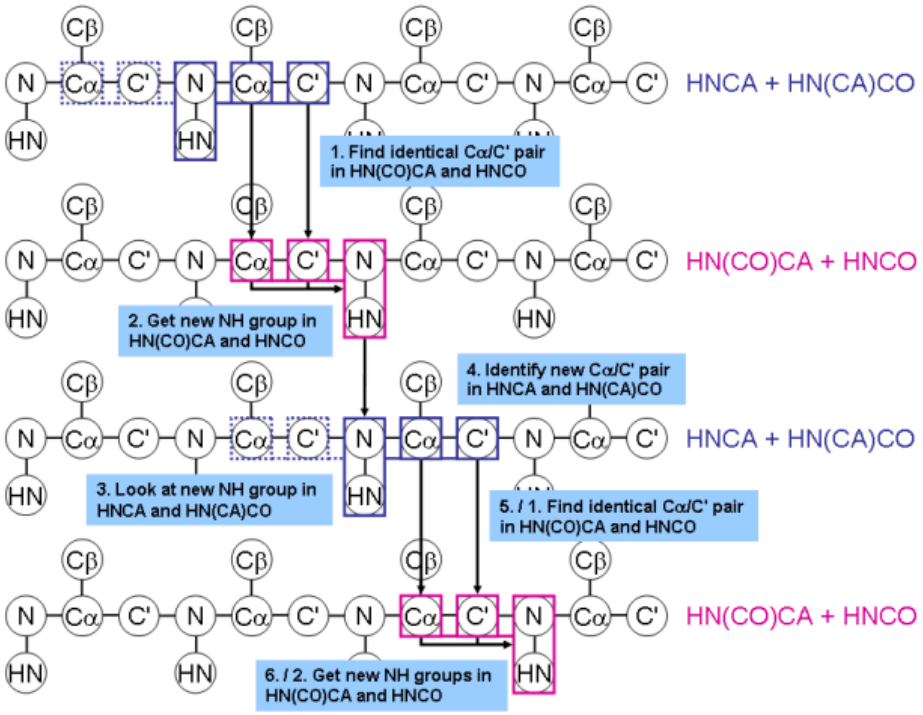


Figure 35. *Backbone assignment based HNCA and HN(CO)CA experiments from https://www.protein-nmr.org.uk/solution-nmr/assignment-theory/triple-resonance-backbone-assignment/.*

3D structure of protein is obtained mainly by through-space NMR experiments involving nuclear overhauser effect (NOE). Thus, NOE experiments can be used to obtain restraints. The intensity of NOE peaks is used to compute distance restraints which are usually divided into 3 distance bins, the strong (less than 3 Å), the medium (less than 4 Å apart) and the weak (less than 6 Å) [248]. In addition, NOE experiments can also be used for the sequencial assignment for small molecular weight proteins.

3.2 Protein Structure Calculation

Software packages are required for the analysis of NMR data and structure calculation. Here, the program, ambiguous restraints for iterative assignment (ARIA), is used for automated NOE assignment and structure calculation [249,250]. Together with the collaborative computing project for NMR (CcpNMR) analysis [251], ARIA uses the

program Crystallography & NMR System (CNS) to execute the structure calculation using NOEs data, chemical shifts, and dihedral angles [252] (**Figure 36**).

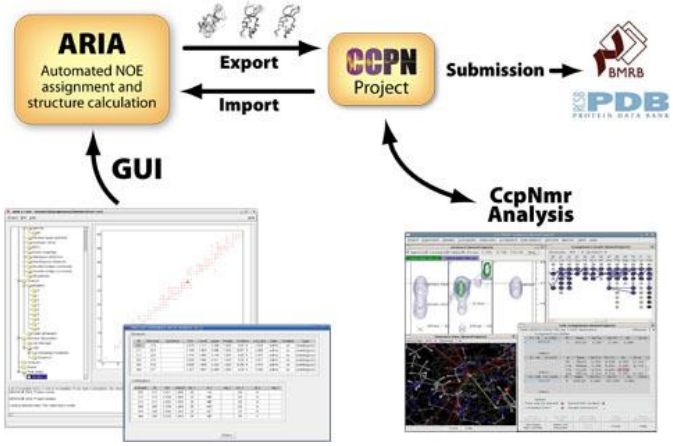


Figure 36. *Structure calculation flow chart from http://aria.pasteur.fr/*

3.3 Protein-Protein/Ligand Interaction by Chemical Shift Perturbation (CSP)

So far, chemical shift perturbation (CSP) is widely used for the determination of protein-protein or protein-ligand interactions [253–255]. The chemical shifts from nuclei are considerably sensitive to residue electronic environment once perturbed by binding or interaction events. The analysis based on the variations of chemical shifts can provide rich information about the process of the binding event induced by protein-protein or protein-ligand interactions. Usually, a typical CSP experiment relies on the analysis of 2D-heteronuclear single quantum coherence (HSQC) spectra recorded in the absence and in the presence of the partner at different concentrations and their comparison (**Figures 37A**).

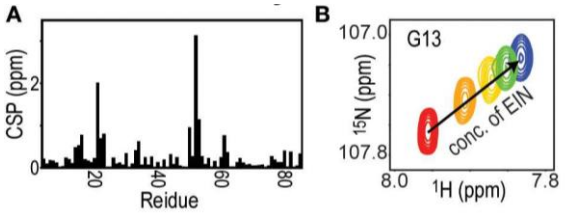


Figure 37. *Chemical shift perturbation (CSP) (A) and NMR titration experiments (B)* [255].

NMR titration experiments allow the quantification of protein-protein or protein-ligand interactions ranging from sub-µM to nM (**Figure 37B**) [253,254]. The dissociation constant (K_d) can be obtained by following the CSP versus the ligand concentration (**Figure 38**).

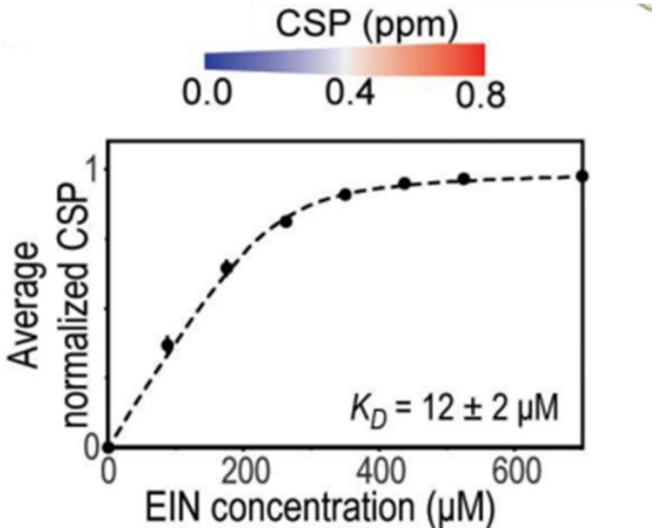


Figure 38. *The dissociation constant (K_d)* [255]

All NMR CSP experiments were performed at 323 K on Bruker 950 MHz and 600 MHz spectrometers both equipped with cryogenic triple resonance probes. To investigate the characterize of selectively labeled HIV-1 p6 in presence of DPC-d_{38}, the concentrations of selectively ^{15}N-labeled HIV-1 p6 was set to 50 µM, and DPC-d_{38} was added at molar ratios ranging from 0:1 to 500:1. ^{1}H-^{15}N HMQC spectra was recorded at different DPC-d_{38}-to-HIV-1 p6 molar ratios.

The concentrations of the ^{15}N, ^{13}C labeled HIV-1 Vpr and ^{15}N, ^{13}C labeled SIVmac Vpx NMR samples were respectively set to 0.3 mM with 90 mM DPC (*ie* 300 eq) and 0.35 mM with 105 mM DPC (*ie* 300 eq). To detect the interaction of HIV-1 Vpr with p6, CSP experiments of ^{15}N, ^{13}C labeled HIV-1 Vpr were performed by adding the unlabeled HIV-

1 p6 at 1:4 molar ratio. The reverse study was performed on selectively ^{15}N-labeled HIV-1 p6 at 50 μM with 300 equivalents DPC-d_{38} by adding the unlabeled HIV-1 Vpr at 1:4 molar ratio. Similarly, the CSP experiments for ^{15}N, ^{13}C labeled SIVmac Vpx with unlabeled either SIVmac p6$^{16-28}$ or SIVmac p6 was performed at molar ratio 1:3.
For the reverse study, the concentrations of unlabeled SIVmac p6$^{16-28}$ and selectively ^{15}N-labeled SIVmac p6 were both set to 1mM with 100mM DPC-d_{38}. The final concentration proportion of either unlabeled SIVmac p6$^{16-28}$ or selectively labeled SIVmac p6 to unlabeled SIVmac p6 was 1:0.1.

For the cross-interaction between SIVmac Vpx and HIV-1 p6, ^{15}N, ^{13}C labeled SIVmac Vpx was titrated with unlabeled HIV-1 p6 at molar ratio 1:3 and 1:5. NMR titration was performed on selectively HIV-1 p6 with unlabeled SIVmac Vpx at molar ratio 1:0, 1:0.1, 1:0.3, 1:0.5, 1:0.8, 1:1 and 1:1.2.

The chemical shift perturbations (CSPs) were computed from the equation: $\Delta\delta=[(\Delta\delta H)^2+(\Delta\delta N/6.5)^2]^{1/2}$ (41), ΔδH and ΔδN standing respectively for the chemical shift variations in the ^{1}H and ^{15}N dimensions. To map the interacting sites, the residues with more than the average of CSPs were analysed on the surface of HIV-1 Vpr and SIVmac Vpx, and residues with CSPs over than the average + 0.5 standard deviation (SD) are regarded as hot spots. The equilibrium dissociation constants K_d were derived from one-site binding model and non-linear fitting with the equation: Y¼Bmax*X/(K_d+X), using Graphpad Prism software. X is the ligand concentration, Y is the chemical shift perturbation, and Bmax is the maximum specific binding.

Results

General discussion

References

[1] G. Huetter, D. Nowak, M. Mossner, S. Ganepola, A. Muessig, K. Allers, T. Schneider, J. Hofmann, C. Kuecherer, O. Blau, I. Blau, W.K. Hofmann, E. Thiel, Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation., N. Engl. J. Med. (2009). https://doi.org/10.1056/NEJMoa0802905.

[2] R.K. Gupta, D. Peppa, A.L. Hill, C. Gálvez, M. Salgado, M. Pace, L.E. McCoy, S.A. Griffith, J. Thornhill, A. Alrubayyi, L.E.P. Huyveneers, E. Nastouli, P. Grant, S.G. Edwards, A.J. Innes, J. Frater, M. Nijhuis, A.M.J. Wensing, J. Martinez-Picado, E. Olavarria, Evidence for HIV-1 cure after CCR5Δ32/Δ32 allogeneic haemopoietic stem-cell transplantation 30 months post analytical treatment interruption: a case report, Lancet HIV. 7 (2020) e340–e347. https://doi.org/10.1016/S2352-3018(20)30069-2.

[3] A.M. Vandamme, K. Van Laethem, E. De Clercq, Managing resistance to anti-HIV drugs: an important consideration for effective disease management, Drugs. 57 (1999) 337–361. https://doi.org/10.2165/00003495-199957030-00006.

[4] L. Lu, P. Tong, X. Yu, C. Pan, P. Zou, Y.-H. Chen, S. Jiang, HIV-1 variants with a single-point mutation in the gp41 pocket region exhibiting different susceptibility to HIV fusion inhibitors with pocket- or membrane-binding domain, Biochim. Biophys. Acta. 1818 (2012) 2950–2957. https://doi.org/10.1016/j.bbamem.2012.07.020.

[5] B. Zheng, K. Wang, L. Lu, F. Yu, M. Cheng, S. Jiang, K. Liu, L. Cai, Hydrophobic mutations in buried polar residues enhance HIV-1 gp41 N-terminal heptad repeat-C-terminal heptad repeat interactions and C-peptides' anti-HIV activity, AIDS Lond. Engl. 28 (2014) 1251–1260. https://doi.org/10.1097/QAD.0000000000000255.

[6] D. Sauter, F. Kirchhoff, Chapter 4 - Properties of Human and Simian Immunodeficiency Viruses, in: A.A. Ansari, G. Silvestri (Eds.), Nat. Hosts SIV, Elsevier, Amsterdam, 2014: pp. 69–84. https://doi.org/10.1016/B978-0-12-404734-1.00004-8.

[7] O.T. Campbell-Yesufu, R.T. Gandhi, Update on Human Immunodeficiency Virus (HIV)-2 Infection, Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am. 52 (2011) 780–787. https://doi.org/10.1093/cid/ciq248.

[8] L. Etienne, E. Delaporte, M. Peeters, 26 - Origin and Emergence of HIV/AIDS, in: M. Tibayrenc (Ed.), Genet. Evol. Infect. Dis., Elsevier, London, 2011: pp. 689–710. https://doi.org/10.1016/B978-0-12-384890-1.00026-1.

[9] C.J. Miller, M. Marthas, Simian Immunodeficiency Virus: Animal Models of Disease, in: B.W.J. Mahy, M.H.V. Van Regenmortel (Eds.), Encycl. Virol. Third Ed., Academic Press, Oxford, 2008: pp. 594–603. https://doi.org/10.1016/B978-012374410-4.00679-8.

[10] P.M. Sharp, B.H. Hahn, Origins of HIV and the AIDS Pandemic, Cold Spring Harb. Perspect. Med. 1 (2011). https://doi.org/10.1101/cshperspect.a006841.

[11] G. Li, E. De Clercq, HIV Genome-Wide Protein Associations: a Review of 30 Years of Research, Microbiol. Mol. Biol. Rev. MMBR. 80 (2016) 679–731. https://doi.org/10.1128/MMBR.00065-15.

[12] B.G. Turner, M.F. Summers, Structural biology of HIV, J. Mol. Biol. 285 (1999) 1–32. https://doi.org/10.1006/jmbi.1998.2354.

[13] M. Pancera, S. Majeed, Y.-E.A. Ban, L. Chen, C. Huang, L. Kong, Y.D. Kwon, J. Stuckey, T. Zhou, J.E. Robinson, W.R. Schief, J. Sodroski, R. Wyatt, P.D. Kwong, Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility, Proc. Natl. Acad. Sci. 107 (2010) 1166–1171. https://doi.org/10.1073/pnas.0911004107.

[14] A. Bartesaghi, A. Merk, M.J. Borgnia, J.L.S. Milne, S. Subramaniam, Prefusion structure of trimeric HIV-1 envelope glycoprotein determined by cryo-electron microscopy, Nat. Struct. Mol. Biol. 20 (2013) 1352–1357. https://doi.org/10.1038/nsmb.2711.

[15] M. Pancera, T. Zhou, A. Druz, I.S. Georgiev, C. Soto, J. Gorman, J. Huang, P. Acharya, G.-Y. Chuang, G. Ofek, G.B.E. Stewart-Jones, J. Stuckey, R.T. Bailer, M.G. Joyce, M.K. Louder, N. Tumba, Y. Yang, B. Zhang, M.S. Cohen, B.F. Haynes, J.R. Mascola, L. Morris, J.B. Munro, S.C. Blanchard, W. Mothes, M. Connors, P.D. Kwong, Structure and immune recognition of trimeric pre-fusion HIV-1 Env, Nature. 514 (2014) 455–461. https://doi.org/10.1038/nature13808.

[16] A. Finzi, S.-H. Xiang, B. Pacheco, L. Wang, J. Haight, A. Kassa, B. Danek, M. Pancera, P.D. Kwong, J. Sodroski, Topological Layers in the HIV-1 gp120 Inner

Domain Regulate gp41 Interaction and CD4-Triggered Conformational Transitions, Mol. Cell. 37 (2010) 656–667. https://doi.org/10.1016/j.molcel.2010.02.012.

[17] J.-P. Julien, A. Cupo, D. Sok, R.L. Stanfield, D. Lyumkis, M.C. Deller, P.-J. Klasse, D.R. Burton, R.W. Sanders, J.P. Moore, A.B. Ward, I.A. Wilson, Crystal structure of a soluble cleaved HIV-1 envelope trimer, Science. 342 (2013) 1477–1483. https://doi.org/10.1126/science.1245625.

[18] D. Lyumkis, J.-P. Julien, N. de Val, A. Cupo, C.S. Potter, P.-J. Klasse, D.R. Burton, R.W. Sanders, J.P. Moore, B. Carragher, I.A. Wilson, A.B. Ward, Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer, Science. 342 (2013) 1484–1490. https://doi.org/10.1126/science.1245627.

[19] H.E. Drummer, M.K. Hill, A.L. Maerz, S. Wood, P.A. Ramsland, J. Mak, P. Poumbourios, Allosteric modulation of the HIV-1 gp120-gp41 association site by adjacent gp120 variable region 1 (V1) N-glycans linked to neutralization sensitivity, PLoS Pathog. 9 (2013) e1003218. https://doi.org/10.1371/journal.ppat.1003218.

[20] Y.D. Kwon, M. Pancera, P. Acharya, I.S. Georgiev, E.T. Crooks, J. Gorman, M.G. Joyce, M. Guttman, X. Ma, S. Narpala, C. Soto, D.S. Terry, Y. Yang, T. Zhou, G. Ahlsen, R.T. Bailer, M. Chambers, G.-Y. Chuang, N.A. Doria-Rose, A. Druz, M.A. Hallen, A. Harned, T. Kirys, M.K. Louder, S. O'Dell, G. Ofek, K. Osawa, M. Prabhakaran, M. Sastry, G.B.E. Stewart-Jones, J. Stuckey, P.V. Thomas, T. Tittley, C. Williams, B. Zhang, H. Zhao, Z. Zhou, B.R. Donald, L.K. Lee, S. Zolla-Pazner, U. Baxa, A. Schön, E. Freire, L. Shapiro, K.K. Lee, J. Arthos, J.B. Munro, S.C. Blanchard, W. Mothes, J.M. Binley, A.B. McDermott, J.R. Mascola, P.D. Kwong, Crystal structure, conformational fixation and entry-related interactions of mature ligand-free HIV-1 Env, Nat. Struct. Mol. Biol. 22 (2015) 522–531. https://doi.org/10.1038/nsmb.3051.

[21] N. Alsahafi, O. Debbeche, J. Sodroski, A. Finzi, Effects of the I559P gp41 Change on the Conformation and Function of the Human Immunodeficiency Virus (HIV-1) Membrane Envelope Glycoprotein Trimer, PLOS ONE. 10 (2015) e0122111. https://doi.org/10.1371/journal.pone.0122111.

[22] Y. Mao, L. Wang, C. Gu, A. Herschhorn, A. Désormeaux, A. Finzi, S.-H. Xiang, J.G. Sodroski, Molecular architecture of the uncleaved HIV-1 envelope glycoprotein

trimer, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 12438–12443. https://doi.org/10.1073/pnas.1307382110.

[23] Y. Mao, L. Wang, C. Gu, A. Herschhorn, S.-H. Xiang, H. Haim, X. Yang, J. Sodroski, Subunit organization of the membrane-bound HIV-1 envelope glycoprotein trimer, Nat. Struct. Mol. Biol. 19 (2012) 893–899. https://doi.org/10.1038/nsmb.2351.

[24] E.A. Berger, P.M. Murphy, J.M. Farber, Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease, Annu. Rev. Immunol. 17 (1999) 657–700. https://doi.org/10.1146/annurev.immunol.17.1.657.

[25] J.A. Esté, A. Telenti, HIV entry inhibitors, Lancet Lond. Engl. 370 (2007) 81–88. https://doi.org/10.1016/S0140-6736(07)61052-6.

[26] A.N. Engelman, P.K. Singh, Cellular and molecular mechanisms of HIV-1 integration targeting, Cell. Mol. Life Sci. 75 (2018) 2491–2507. https://doi.org/10.1007/s00018-018-2772-5.

[27] W.-S. Hu, S.H. Hughes, HIV-1 Reverse Transcription, Cold Spring Harb. Perspect. Med. 2 (2012). https://doi.org/10.1101/cshperspect.a006882.

[28] V. Scoca, F. Di Nunzio, The HIV-1 Capsid: From Structural Component to Key Factor for Host Nuclear Invasion, Viruses. 13 (2021) 273. https://doi.org/10.3390/v13020273.

[29] S. Popov, M. Rexach, L. Ratner, G. Blobel, M. Bukrinsky, Viral Protein R Regulates Docking of the HIV-1 Preintegration Complex to the Nuclear Pore Complex*, J. Biol. Chem. 273 (1998) 13347–13352. https://doi.org/10.1074/jbc.273.21.13347.

[30] S. Popov, M. Rexach, G. Zybarth, N. Reiling, M.-A. Lee, L. Ratner, C.M. Lane, M.S. Moore, G. Blobel, M. Bukrinsky, Viral protein R regulates nuclear import of the HIV-1 pre-integration complex, EMBO J. 17 (1998) 909–917. https://doi.org/10.1093/emboj/17.4.909.

[31] E.M. Anderson, F. Maldarelli, The role of integration and clonal expansion in HIV infection: live long and prosper, Retrovirology. 15 (2018) 71. https://doi.org/10.1186/s12977-018-0448-8.

[32] Q. Wang, A. Finzi, J. Sodroski, The Conformational States of the HIV-1 Envelope Glycoproteins, Trends Microbiol. 28 (2020) 655–667. https://doi.org/10.1016/j.tim.2020.03.007.

[33] K.A. Jones, Taking a new TAK on tat transactivation, Genes Dev. 11 (1997) 2593–2599. https://doi.org/10.1101/gad.11.20.2593.

[34] R. Liu, J. Wu, R. Shao, Y. Xue, Mechanism and factors that control HIV-1 transcription and latency activation, J. Zhejiang Univ. Sci. B. 15 (2014) 455–465. https://doi.org/10.1631/jzus.B1400059.

[35] S. Guerrero, J. Batisse, C. Libre, S. Bernacchi, R. Marquet, J.-C. Paillart, HIV-1 Replication and the Cellular Eukaryotic Translation Apparatus, Viruses. 7 (2015) 199–218. https://doi.org/10.3390/v7010199.

[36] E.O. Freed, HIV-1 assembly, release and maturation, Nat. Rev. Microbiol. 13 (2015) 484–496. https://doi.org/10.1038/nrmicro3490.

[37] P.I. Hanson, R. Roth, Y. Lin, J.E. Heuser, Plasma membrane deformation by circular arrays of ESCRT-III protein filaments, J. Cell Biol. 180 (2008) 389–402. https://doi.org/10.1083/jcb.200707031.

[38] Q.-T. Shen, A.L. Schuh, Y. Zheng, K. Quinney, L. Wang, M. Hanna, J.C. Mitchell, M.S. Otegui, P. Ahlquist, Q. Cui, A. Audhya, Structural analysis and modeling reveals new mechanisms governing ESCRT-III spiral filament assembly, J. Cell Biol. 206 (2014) 763–777. https://doi.org/10.1083/jcb.201403108.

[39] W.I. Sundquist, H.-G. Kräusslich, HIV-1 Assembly, Budding, and Maturation, Cold Spring Harb. Perspect. Med. 2 (2012). https://doi.org/10.1101/cshperspect.a006924.

[40] V. Zila, E. Margiotta, B. Turoňová, T.G. Müller, C.E. Zimmerli, S. Mattei, M. Allegretti, K. Börner, J. Rada, B. Müller, M. Lusic, H.-G. Kräusslich, M. Beck, Cone-shaped HIV-1 capsids are transported through intact nuclear pores, Cell. 184 (2021) 1032-1046.e18. https://doi.org/10.1016/j.cell.2021.01.025.

[41] C. Em, H. Tj, HIV-1 capsid: the multifaceted key player in HIV-1 infection, Nat. Rev. Microbiol. 13 (2015). https://doi.org/10.1038/nrmicro3503.

[42] P. Pawlica, L. Berthoux, Cytoplasmic dynein promotes HIV-1 uncoating, Viruses. 6 (2014) 4195–4211. https://doi.org/10.3390/v6114195.

[43] Z. Lukic, A. Dharan, T. Fricke, F. Diaz-Griffero, E.M. Campbell, HIV-1 uncoating is facilitated by dynein and kinesin 1, J. Virol. 88 (2014) 13613–13625. https://doi.org/10.1128/JVI.02219-14.

[44] A.B. Kleinpeter, E.O. Freed, HIV-1 Maturation: Lessons Learned from Inhibitors, Viruses. 12 (2020). https://doi.org/10.3390/v12090940.

[45] J.R. Lingappa, J.C. Reed, M. Tanaka, K. Chutiraka, B.A. Robinson, How HIV-1 Gag assembles in cells: Putting together pieces of the puzzle, Virus Res. 193 (2014) 89–107. https://doi.org/10.1016/j.virusres.2014.07.001.

[46] N.M. Bell, A.M.L. Lever, HIV Gag polyprotein: processing and early viral particle assembly, Trends Microbiol. 21 (2013) 136–144. https://doi.org/10.1016/j.tim.2012.11.006.

[47] C.P. Hill, D. Worthylake, D.P. Bancroft, A.M. Christensen, W.I. Sundquist, Crystal structures of the trimeric human immunodeficiency virus type 1 matrix protein: implications for membrane association and assembly, Proc. Natl. Acad. Sci. 93 (1996) 3099–3104. https://doi.org/10.1073/pnas.93.7.3099.

[48] J.S. Saad, J. Miller, J. Tai, A. Kim, R.H. Ghanam, M.F. Summers, Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly, Proc. Natl. Acad. Sci. 103 (2006) 11364–11369. https://doi.org/10.1073/pnas.0602818103.

[49] M. Barros, F. Heinrich, S.A.K. Datta, A. Rein, I. Karageorgos, H. Nanda, M. Lösche, Membrane Binding of HIV-1 Matrix Protein: Dependence on Bilayer Composition and Protein Lipidation, J. Virol. 90 (2016) 4544–4555. https://doi.org/10.1128/JVI.02820-15.

[50] V. Chukkapalli, S.J. Oh, A. Ono, Opposing mechanisms involving RNA and lipids regulate HIV-1 Gag membrane binding through the highly basic region of the matrix domain, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 1600–1605. https://doi.org/10.1073/pnas.0908661107.

[51] B. Ganser-Pornillos, M. Yeager, W.I. Sundquist, The Structural Biology of HIV Assembly, Curr. Opin. Struct. Biol. 18 (2008) 203. https://doi.org/10.1016/j.sbi.2008.02.001.

[52] J.R. Lingappa, J.C. Reed, M. Tanaka, K. Chutiraka, B.A. Robinson, How HIV-1 Gag assembles in cells: Putting together pieces of the puzzle, Virus Res. 193 (2014) 89–107. https://doi.org/10.1016/j.virusres.2014.07.001.

[53] P.Y. Mercredi, N. Bucca, B. Loeliger, C.R. Gaines, M. Mehta, P. Bhargava, P.R. Tedbury, L. Charlier, N. Floquet, D. Muriaux, C. Favard, C.R. Sanders, E.O. Freed, J. Marchant, M.F. Summers, Structural and Molecular Determinants of Membrane Binding by the HIV-1 Matrix Protein, J. Mol. Biol. 428 (2016) 1637–1655. https://doi.org/10.1016/j.jmb.2016.03.005.

[54] C. Mariani, M. Desdouits, C. Favard, P. Benaroch, D.M. Muriaux, Role of Gag and lipids during HIV-1 assembly in CD4+ T cells and macrophages, Front. Microbiol. 5 (2014). https://doi.org/10.3389/fmicb.2014.00312.

[55] B.K. Ganser, S. Li, V.Y. Klishko, J.T. Finch, W.I. Sundquist, Assembly and analysis of conical models for the HIV-1 core, Science. 283 (1999) 80–83. https://doi.org/10.1126/science.283.5398.80.

[56] J.A.G. Briggs, M.N. Simon, I. Gross, H.-G. Kräusslich, S.D. Fuller, V.M. Vogt, M.C. Johnson, The stoichiometry of Gag protein in HIV-1, Nat. Struct. Mol. Biol. 11 (2004) 672–675. https://doi.org/10.1038/nsmb785.

[57] G. Zhao, J.R. Perilla, E.L. Yufenyuy, X. Meng, B. Chen, J. Ning, J. Ahn, A.M. Gronenborn, K. Schulten, C. Aiken, P. Zhang, Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics, Nature. 497 (2013) 643–646. https://doi.org/10.1038/nature12162.

[58] O. Pornillos, B.K. Ganser-Pornillos, B.N. Kelly, Y. Hua, F.G. Whitby, C.D. Stout, W.I. Sundquist, C.P. Hill, M. Yeager, X-ray Structures of the Hexameric Building Block of the HIV Capsid, Cell. 137 (2009) 1282–1292. https://doi.org/10.1016/j.cell.2009.04.063.

[59] T.R. Gamble, F.F. Vajdos, S. Yoo, D.K. Worthylake, M. Houseweart, W.I. Sundquist, C.P. Hill, Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid, Cell. 87 (1996) 1285–1294. https://doi.org/10.1016/s0092-8674(00)81823-1.

[60] R.K. Gitti, B.M. Lee, J. Walker, M.F. Summers, S. Yoo, W.I. Sundquist, Structure of the amino-terminal core domain of the HIV-1 capsid protein, Science. 273 (1996) 231–235. https://doi.org/10.1126/science.273.5272.231.

[61] T.A.M. Bharat, N.E. Davey, P. Ulbrich, J.D. Riches, A. de Marco, M. Rumlova, C. Sachse, T. Ruml, J.A.G. Briggs, Structure of the immature retroviral capsid at 8 Å resolution by cryo-electron microscopy, Nature. 487 (2012) 385–389. https://doi.org/10.1038/nature11169.

[62] D.K. Worthylake, H. Wang, S. Yoo, W.I. Sundquist, C.P. Hill, Structures of the HIV-1 capsid protein dimerization domain at 2.6 Å resolution, Acta Crystallogr. D Biol. Crystallogr. 55 (1999) 85–92. https://doi.org/10.1107/S0907444998007689.

[63] T.R. Gamble, S. Yoo, F.F. Vajdos, U.K. von Schwedler, D.K. Worthylake, H. Wang, J.P. McCutcheon, W.I. Sundquist, C.P. Hill, Structure of the Carboxyl-Terminal Dimerization Domain of the HIV-1 Capsid Protein, Science. 278 (1997) 849–853. https://doi.org/10.1126/science.278.5339.849.

[64] P.M. Lokhandwala, T.-L.N. Nguyen, J.B. Bowzard, R.C. Craven, Cooperative role of the MHR and the CA dimerization helix in the maturation of the functional retrovirus capsid, Virology. 376 (2008) 191–198. https://doi.org/10.1016/j.virol.2008.03.001.

[65] J.G. Purdy, J.M. Flanagan, I.J. Ropson, K.E. Rennoll-Bankert, R.C. Craven, Critical role of conserved hydrophobic residues within the major homology region in mature retroviral capsid assembly, J. Virol. 82 (2008) 5951–5961. https://doi.org/10.1128/JVI.00214-08.

[66] R. Doménech, R. Bocanegra, A. Velázquez-Campoy, J.L. Neira, The isolated major homology region of the HIV capsid protein is mainly unfolded in solution and binds to the intact protein, Biochim. Biophys. Acta BBA - Proteins Proteomics. 1814 (2011) 1269–1278. https://doi.org/10.1016/j.bbapap.2011.06.011.

[67] U.K. von Schwedler, K.M. Stray, J.E. Garrus, W.I. Sundquist, Functional Surfaces of the Human Immunodeficiency Virus Type 1 Capsid Protein, J. Virol. 77 (2003) 5439–5450. https://doi.org/10.1128/JVI.77.9.5439-5450.2003.

[68] N. Morellet, S. Druillennec, C. Lenoir, S. Bouaziz, B.P. Roques, Helical structure determined by NMR of the HIV-1 (345–392)Gag sequence, surrounding p2:

Implications for particle assembly and RNA packaging, Protein Sci. Publ. Protein Soc. 14 (2005) 375–386. https://doi.org/10.1110/ps.041087605.

[69] S.A.K. Datta, L.G. Temeselew, R.M. Crist, F. Soheilian, A. Kamata, J. Mirro, D. Harvin, K. Nagashima, R.E. Cachau, A. Rein, On the Role of the SP1 Domain in HIV-1 Particle Assembly: a Molecular Switch?▿, J. Virol. 85 (2011) 4111–4121. https://doi.org/10.1128/JVI.00006-11.

[70] C.S. Adamson, M. Sakalian, K. Salzwedel, E.O. Freed, Polymorphisms in Gag spacer peptide 1 confer varying levels of resistance to the HIV- 1maturation inhibitor bevirimat, Retrovirology. 7 (2010) 36. https://doi.org/10.1186/1742-4690-7-36.

[71] X. Chen, P. Coric, V. Larue, S. Turcaud, X. Wang, S. Nonin-Lecomte, S. Bouaziz, The HIV-1 maturation inhibitor, EP39, interferes with the dynamic helix-coil equilibrium of the CA-SP1 junction of Gag, Eur. J. Med. Chem. 204 (2020) 112634. https://doi.org/10.1016/j.ejmech.2020.112634.

[72] J. Godet, C. Boudier, N. Humbert, R. Ivanyi-Nagy, J.-L. Darlix, Y. Mély, Comparative nucleic acid chaperone properties of the nucleocapsid protein NCp7 and Tat protein of HIV-1, Virus Res. 169 (2012) 349–360. https://doi.org/10.1016/j.virusres.2012.06.021.

[73] J.-L. Darlix, C. Gabus, M.-T. Nugeyre, F. Clavel, F. Barré-Sinoussi, Cis elements and Trans-acting factors involved in the RNA dimerization of the human immunodeficiency virus HIV-1, J. Mol. Biol. 216 (1990) 689–699. https://doi.org/10.1016/0022-2836(90)90392-Y.

[74] Y.-X. Feng, T.D. Copeland, L.E. Henderson, R.J. Gorelick, W.J. Bosche, J.G. Levin, A. Rein, HIV-1 Nucleocapsid Protein Induces ``Maturation'' of Dimeric Retroviral RNA in vitro, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 7577–7581.

[75] J.G. Levin, M. Mitra, A. Mascarenhas, K. Musier-Forsyth, Role of HIV-1 nucleocapsid protein in HIV-1 reverse transcription, RNA Biol. 7 (2010) 754–774. https://doi.org/10.4161/rna.7.6.14115.

[76] S.E. El Meshri, D. Dujardin, J. Godet, L. Richert, C. Boudier, J.L. Darlix, P. Didier, Y. Mély, H. de Rocquigny, Role of the nucleocapsid domain in HIV-1 Gag oligomerization and trafficking to the plasma membrane: a fluorescence lifetime

imaging microscopy investigation, J. Mol. Biol. 427 (2015) 1480–1494. https://doi.org/10.1016/j.jmb.2015.01.015.

[77] Y.X. Feng, T.D. Copeland, L.E. Henderson, R.J. Gorelick, W.J. Bosche, J.G. Levin, A. Rein, HIV-1 nucleocapsid protein induces “maturation” of dimeric retroviral RNA in vitro., Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 7577–7581.

[78] M. Hill, A. Bellamy-McIntyre, L. Vella, S. Campbell, J. Marshall, G. Tachedjian, J. Mak, Alteration of the Proline at Position 7 of the HIV-1 Spacer Peptide p1 Suppresses Viral Infectivity in a Strain Dependent Manner, Curr. HIV Res. 5 (2007) 69–78. https://doi.org/10.2174/157016207779316323.

[79] B. Xue, M.J. Mizianty, L. Kurgan, V.N. Uversky, Protein intrinsic disorder as a flexible armor and a weapon of HIV-1, Cell. Mol. Life Sci. CMLS. 69 (2012) 1211–1259. https://doi.org/10.1007/s00018-011-0859-3.

[80] E.A. Cohen, E.F. Terwilliger, Y. Jalinoos, J. Proulx, J.G. Sodroski, W.A. Haseltine, Identification of HIV-1 vpr product and function, J. Acquir. Immune Defic. Syndr. 3 (1990) 11–18.

[81] J.C. Kappes, J.S. Parkin, J.A. Conway, J. Kim, C.G. Brouillette, G.M. Shaw, B.H. Hahn, Intracellular transport and virion incorporation of vpx requires interaction with other virus type-specific components, Virology. 193 (1993) 222–233. https://doi.org/10.1006/viro.1993.1118.

[82] W. Paxton, R.I. Connor, N.R. Landau, Incorporation of Vpr into human immunodeficiency virus type 1 virions: requirement for the p6 region of gag and mutational analysis, J. Virol. 67 (1993) 7229–7237. https://doi.org/10.1128/JVI.67.12.7229-7237.1993.

[83] S.A. Stewart, B. Poon, J.B.M. Jowett, Y. Xie, I.S.Y. Chen, Lentiviral delivery of HIV-1 Vpr protein induces apoptosis in transformed cells, Proc. Natl. Acad. Sci. 96 (1999) 12039–12043. https://doi.org/10.1073/pnas.96.21.12039.

[84] J. Yasuda, T. Miyao, M. Kamata, Y. Aida, Y. Iwakura, T Cell Apoptosis Causes Peripheral T Cell Depletion in Mice Transgenic for the HIV-1 vpr Gene, Virology. 285 (2001) 181–192. https://doi.org/10.1006/viro.2001.0964.

[85] T. Fukumori, H. Akari, S. Iida, S. Hata, S. Kagawa, Y. Aida, A.H. Koyama, A. Adachi, The HIV-1 Vpr displays strong anti-apoptotic activity, FEBS Lett. 432 (1998) 17–20. https://doi.org/10.1016/S0014-5793(98)00824-2.

[86] M. Somasundaran, M. Sharkey, B. Brichacek, K. Luzuriaga, M. Emerman, J.L. Sullivan, M. Stevenson, Evidence for a cytopathogenicity determinant in HIV-1 Vpr, Proc. Natl. Acad. Sci. 99 (2002) 9503–9508. https://doi.org/10.1073/pnas.142313699.

[87] E. Le Rouzic, S. Benichou, The Vpr protein from HIV-1: distinct roles along the viral life cycle, Retrovirology. 2 (2005) 11. https://doi.org/10.1186/1742-4690-2-11.

[88] R.Y. Zhao, G. Li, M.I. Bukrinsky, Vpr-Host Interactions During HIV-1 Viral Life Cycle, J. Neuroimmune Pharmacol. 6 (2011) 216–229. https://doi.org/10.1007/s11481-011-9261-z.

[89] N. Morellet, S. Bouaziz, P. Petitjean, B.P. Roques, NMR structure of the HIV-1 regulatory protein VPR, J. Mol. Biol. 327 (2003) 215–227. https://doi.org/10.1016/s0022-2836(03)00060-3.

[90] Y. Wu, X. Zhou, C.O. Barnes, M. DeLucia, A.E. Cohen, A.M. Gronenborn, J. Ahn, G. Calero, The DDB1–DCAF1–Vpr–UNG2 crystal structure reveals how HIV-1 Vpr steers human UNG2 toward destruction, Nat. Struct. Mol. Biol. 23 (2016) 933–940. https://doi.org/10.1038/nsmb.3284.

[91] L. Selig, J.-C. Pages, V. Tanchou, S. Prévéral, C. Berlioz-Torrent, L.X. Liu, L. Erdtmann, J.-L. Darlix, R. Benarous, S. Benichou, Interaction with the p6 Domain of the Gag Precursor Mediates Incorporation into Virions of Vpr and Vpx Proteins from Primate Lentiviruses, J. Virol. 73 (1999) 592–600. https://doi.org/10.1128/JVI.73.1.592-600.1999.

[92] M.P. Sherman, C.M.C. de Noronha, L.A. Eckstein, J. Hataye, P. Mundt, S.A.F. Williams, J.A. Neidleman, M.A. Goldsmith, W.C. Greene, Nuclear export of Vpr is required for efficient replication of human immunodeficiency virus type 1 in tissue macrophages, J. Virol. 77 (2003) 7582–7589. https://doi.org/10.1128/jvi.77.13.7582-7589.2003.

[93] N.J. Venkatachari, L.A. Walker, O. Tastan, T. Le, T.M. Dempsey, Y. Li, N. Yanamala, A. Srinivasan, J. Klein-Seetharaman, R.C. Montelaro, V. Ayyavoo,

Human immunodeficiency virus type 1 Vpr: oligomerization is an essential feature for its incorporation into virus particles, Virol. J. 7 (2010) 119. https://doi.org/10.1186/1743-422X-7-119.

[94] W. Schüler, K. Wecker, H. de Rocquigny, Y. Baudat, J. Sire, B.P. Roques, NMR structure of the (52–96) C-terminal domain of the HIV-1 regulatory protein vpr: molecular insights into its biological functions11Edited by P. E. Wright, J. Mol. Biol. 285 (1999) 2105–2117. https://doi.org/10.1006/jmbi.1998.2381.

[95] E. Coeytaux, D. Coulaud, E. Le Cam, O. Danos, A. Kichler, The Cationic Amphipathic α-Helix of HIV-1 Viral Protein R (Vpr) Binds to Nucleic Acids, Permeabilizes Membranes, and Efficiently Transfects Cells*, J. Biol. Chem. 278 (2003) 18110–18116. https://doi.org/10.1074/jbc.M300248200.

[96] J.V. Fritz, P. Didier, J.-P. Clamme, E. Schaub, D. Muriaux, C. Cabanne, N. Morellet, S. Bouaziz, J.-L. Darlix, Y. Mély, H. de Rocquigny, Direct Vpr-Vpr Interaction in Cells monitored by two Photon Fluorescence Correlation Spectroscopy and Fluorescence Lifetime Imaging, Retrovirology. 5 (2008) 87. https://doi.org/10.1186/1742-4690-5-87.

[97] J.C. Paillart, H.G. Göttlinger, Opposing effects of human immunodeficiency virus type 1 matrix mutations support a myristyl switch model of gag membrane targeting, J. Virol. 73 (1999) 2604–2612. https://doi.org/10.1128/JVI.73.4.2604-2612.1999.

[98] Y. Jenkins, O. Pornillos, R.L. Rich, D.G. Myszka, W.I. Sundquist, M.H. Malim, Biochemical Analyses of the Interactions between Human Immunodeficiency Virus Type 1 Vpr and p6Gag, J. Virol. 75 (2001) 10537–10542. https://doi.org/10.1128/JVI.75.21.10537-10542.2001.

[99] E.C. Friedberg, A history of the DNA repair and mutagenesis field: The discovery of base excision repair, DNA Repair. 37 (2016) A35–A39. https://doi.org/10.1016/j.dnarep.2015.12.003.

[100] T. Lindahl, My Journey to DNA Repair, Genomics Proteomics Bioinformatics. 11 (2013) 2–7. https://doi.org/10.1016/j.gpb.2012.12.001.

[101] M. Bouhamdan, S. Benichou, F. Rey, J.M. Navarro, I. Agostini, B. Spire, J. Camonis, G. Slupphaug, R. Vigne, R. Benarous, J. Sire, Human immunodeficiency virus type

1 Vpr protein binds to the uracil DNA glycosylase DNA repair enzyme., J. Virol. 70 (1996) 697–704. https://doi.org/10.1128/JVI.70.2.697-704.1996.

[102] T. Yoshikawa, K. Matsuo, K. Matsuo, Y. Suzuki, A. Nomoto, S.-I. Tamura, T. Kurata, T. Sata, Total viral genome copies and virus-Ig complexes after infection with influenza virus in the nasal secretions of immunized mice, J. Gen. Virol. 85 (2004) 2339–2346. https://doi.org/10.1099/vir.0.79892-0.

[103] L.M. Mansky, S. Preveral, L. Selig, R. Benarous, S. Benichou, The Interaction of Vpr with Uracil DNA Glycosylase Modulates the Human Immunodeficiency Virus Type 1 In Vivo Mutation Rate, J. Virol. 74 (2000) 7039–7047. https://doi.org/10.1128/JVI.74.15.7039-7047.2000.

[104] L.A. Stark, R.T. Hay, Human Immunodeficiency Virus Type 1 (HIV-1) Viral Protein R (Vpr) Interacts with Lys-tRNA Synthetase: Implications for Priming of HIV-1 Reverse Transcription, J. Virol. 72 (1998) 3037–3044. https://doi.org/10.1128/JVI.72.4.3037-3044.1998.

[105] E.J. Arts, J.T. Miller, B. Ehresmann, S.F.J. Le Grice, Mutating a Region of HIV-1 Reverse Transcriptase Implicated in tRNALys-3 Binding and the Consequences for (−)-Strand DNA Synthesis*, J. Biol. Chem. 273 (1998) 14523–14532. https://doi.org/10.1074/jbc.273.23.14523.

[106] S. Cen, M. Niu, L. Kleiman, The connection domain in reverse transcriptase facilitates the in vivo annealing of tRNALys3 to HIV-1 genomic RNA, Retrovirology. 1 (2004) 33. https://doi.org/10.1186/1742-4690-1-33.

[107] A. Khorchid, H. Javanbakht, S. Wise, R. Halwani, M.A. Parniak, M.A. Wainberg, L. Kleiman, Sequences within Pr160gag-pol affecting the selective packaging of primer tRNALys3 into HIV-111Edited by J. Karn, J. Mol. Biol. 299 (2000) 17–26. https://doi.org/10.1006/jmbi.2000.3709.

[108] E. Dufour, J. Reinbolt, M. Castroviejo, B. Ehresmann, S. Litvak, L. Tarrago-Litvak, M.-L. Andreola, Cross-linking localization of a HIV-1 reverse transcriptase peptide involved in the binding of primer tRNALys311Edited by M. Yaniv, J. Mol. Biol. 285 (1999) 1339–1346. https://doi.org/10.1006/jmbi.1998.2430.

[109] M. Marfori, A. Mynott, J.J. Ellis, A.M. Mehdi, N.F.W. Saunders, P.M. Curmi, J.K. Forwood, M. Bodén, B. Kobe, Molecular basis for specificity of nuclear import and

prediction of nuclear localization, Biochim. Biophys. Acta BBA - Mol. Cell Res. 1813 (2011) 1562–1577. https://doi.org/10.1016/j.bbamcr.2010.10.013.

[110] M. Marfori, T.G. Lonhienne, J.K. Forwood, B. Kobe, Structural Basis of High-Affinity Nuclear Localization Signal Interactions with Importin-α, Traffic. 13 (2012) 532–548. https://doi.org/10.1111/j.1600-0854.2012.01329.x.

[111] M. Stewart, Molecular mechanism of the nuclear protein import cycle, Nat. Rev. Mol. Cell Biol. 8 (2007) 195–208. https://doi.org/10.1038/nrm2114.

[112] S. Lyonnais, R.J. Gorelick, F. Heniche-Boukhalfa, S. Bouaziz, V. Parissi, J.-F. Mouscadet, T. Restle, J.M. Gatell, E. Le Cam, G. Mirambeau, A protein ballet around the viral genome orchestrated by HIV-1 reverse transcriptase leads to an architectural switch: From nucleocapsid-condensed RNA to Vpr-bridged DNA, Virus Res. 171 (2013) 287–303. https://doi.org/10.1016/j.virusres.2012.09.008.

[113] M.A. Vodicka, D.M. Koepp, P.A. Silver, M. Emerman, HIV-1 Vpr interacts with the nuclear transport pathway to promote macrophage infection, Genes Dev. 12 (1998) 175–185. https://doi.org/10.1101/gad.12.2.175.

[114] Y. Nitahara-Kasahara, M. Kamata, T. Yamamoto, X. Zhang, Y. Miyamoto, K. Muneta, S. Iijima, Y. Yoneda, Y. Tsunetsugu-Yokota, Y. Aida, Novel nuclear import of Vpr promoted by importin alpha is crucial for human immunodeficiency virus type 1 replication in macrophages, J. Virol. 81 (2007) 5284–5293. https://doi.org/10.1128/JVI.01928-06.

[115] R.A. Fouchier, B.E. Meyer, J.H. Simon, U. Fischer, A.V. Albright, F. González-Scarano, M.H. Malim, Interaction of the human immunodeficiency virus type 1 Vpr protein with the nuclear pore complex, J. Virol. 72 (1998) 6004–6013. https://doi.org/10.1128/JVI.72.7.6004-6013.1998.

[116] C.M.C. de Noronha, Dynamic Disruptions in Nuclear Envelope Architecture and Integrity Induced by HIV-1 Vpr, Science. 294 (2001) 1105–1108. https://doi.org/10.1126/science.1063957.

[117] L. Wang, S. Mukherjee, F. Jia, O. Narayan, L.-J. Zhao, Interaction of Virion Protein Vpr of Human Immunodeficiency Virus Type 1 with Cellular Transcription Factor Sp1 and trans-Activation of Viral Long Terminal Repeat (∗), J. Biol. Chem. 270 (1995) 25564–25569. https://doi.org/10.1074/jbc.270.43.25564.

[118] B.E. Sawaya, K. Khalili, J. Gordon, R. Taube, S. Amini, Cooperative Interaction between HIV-1 Regulatory Proteins Tat and Vpr Modulates Transcription of the Viral Genome*, J. Biol. Chem. 275 (2000) 35209–35214. https://doi.org/10.1074/jbc.M005197200.

[119] D. Fenard, L. Houzet, E. Bernard, A. Tupin, S. Brun, M. Mougel, C. Devaux, N. Chazal, L. Briant, Uracil DNA Glycosylase 2 negatively regulates HIV-1 LTR transcription, Nucleic Acids Res. 37 (2009) 6008–6018. https://doi.org/10.1093/nar/gkp673.

[120] S.L. Deshmane, S. Amini, S. Sen, K. Khalili, B.E. Sawaya, Regulation of the HIV-1 promoter by HIF-1α and Vpr proteins, Virol. J. 8 (2011) 477. https://doi.org/10.1186/1743-422X-8-477.

[121] R. Liu, Y. Lin, R. Jia, Y. Geng, C. Liang, J. Tan, W. Qiao, HIV-1 Vpr stimulates NF-κB and AP-1 signaling by activating TAK1, Retrovirology. 11 (2014) 45. https://doi.org/10.1186/1742-4690-11-45.

[122] B. Romani, R. Kamali Jamil, M. Hamidi-Fard, P. Rahimi, S.B. Momen, M.R. Aghasadeghi, E. Allahbakhshi, HIV-1 Vpr reactivates latent HIV-1 provirus by inducing depletion of class I HDACs on chromatin, Sci. Rep. 6 (2016) 31924. https://doi.org/10.1038/srep31924.

[123] E. Le Rouzic, A. Mousnier, C. Rustum, F. Stutz, E. Hallberg, C. Dargemont, S. Benichou, Docking of HIV-1 Vpr to the Nuclear Envelope Is Mediated by the Interaction with the Nucleoporin hCG1*, J. Biol. Chem. 277 (2002) 45091–45098. https://doi.org/10.1074/jbc.M207439200.

[124] H. de Rocquigny, P. Petitjean, V. Tanchou, D. Decimo, L. Drouot, T. Delaunay, J.-L. Darlix, B.P. Roques, The Zinc Fingers of HIV Nucleocapsid Protein NCp7 Direct Interactions with the Viral Regulatory Protein Vpr*, J. Biol. Chem. 272 (1997) 30753–30759. https://doi.org/10.1074/jbc.272.49.30753.

[125] Y.L. Lu, R.P. Bennett, J.W. Wills, R. Gorelick, L. Ratner, A leucine triplet repeat sequence (LXX)4 in p6gag is important for Vpr incorporation into human immunodeficiency virus type 1 particles, J. Virol. 69 (1995) 6873–6879. https://doi.org/10.1128/JVI.69.11.6873-6879.1995.

[126] E. Kondo, H.G. Göttlinger, A conserved LXXLF sequence is the major determinant in p6gag required for the incorporation of human immunodeficiency virus type 1 Vpr., J. Virol. 70 (1996) 159–164. https://doi.org/10.1128/JVI.70.1.159-164.1996.

[127] F. Bachand, X.-J. Yao, M. Hrimech, N. Rougeau, É.A. Cohen, Incorporation of Vpr into Human Immunodeficiency Virus Type 1 Requires a Direct Interaction with the p6 Domain of the p55 Gag Precursor*, J. Biol. Chem. 274 (1999) 9083–9091. https://doi.org/10.1074/jbc.274.13.9083.

[128] Y. Li, Z. Zhang, J.K. Wakefield, S.M. Kang, C.D. Morrow, Nucleotide substitutions within U5 are critical for efficient reverse transcription of human immunodeficiency virus type 1 with a primer binding site complementary to tRNA(His), J. Virol. 71 (1997) 6315–6322. https://doi.org/10.1128/JVI.71.9.6315-6322.1997.

[129] B. Poon, I.S.Y. Chen, Human immunodeficiency virus type 1 (HIV-1) Vpr enhances expression from unintegrated HIV-1 DNA, J. Virol. 77 (2003) 3962–3972. https://doi.org/10.1128/jvi.77.7.3962-3972.2003.

[130] B. Poon, M.A. Chang, I.S.Y. Chen, Vpr is required for efficient Nef expression from unintegrated human immunodeficiency virus type 1 DNA, J. Virol. 81 (2007) 10515–10523. https://doi.org/10.1128/JVI.00947-07.

[131] H. Tachiwana, M. Shimura, C. Nakai-Murakami, K. Tokunaga, Y. Takizawa, T. Sata, H. Kurumizaka, Y. Ishizaka, HIV-1 Vpr Induces DNA Double-Strand Breaks, Cancer Res. 66 (2006) 627–631. https://doi.org/10.1158/0008-5472.CAN-05-3144.

[132] J.M. Elinoff, L.-Y. Chen, E.J. Dougherty, K.S. Awad, S. Wang, A. Biancotto, A.H. Siddiqui, N.A. Weir, R. Cai, J. Sun, I.R. Preston, M.A. Solomon, R.L. Danner, Spironolactone-induced degradation of the TFIIH core complex XPB subunit suppresses NF-κB and AP-1 signalling, Cardiovasc. Res. 114 (2018) 65–76. https://doi.org/10.1093/cvr/cvx198.

[133] G.F. Salgado, R. Marquant, A. Vogel, I.D. Alves, S.E. Feller, N. Morellet, S. Bouaziz, Structural Studies of HIV-1 Gag p6ct and Its Interaction with Vpr Determined by Solution Nuclear Magnetic Resonance †, ‡, Biochemistry. 48 (2009) 2355–2367. https://doi.org/10.1021/bi801794v.

[134] D. Schwefel, V.C. Boucherit, E. Christodoulou, P.A. Walker, J.P. Stoye, K.N. Bishop, I.A. Taylor, Molecular Determinants for Recognition of Divergent

SAMHD1 Proteins by the Lentiviral Accessory Protein Vpx, Cell Host Microbe. 17 (2015) 489–499. https://doi.org/10.1016/j.chom.2015.03.004.

[135] N. Laguette, B. Sobhian, N. Casartelli, M. Ringeard, C. Chable-Bessia, E. Ségéral, A. Yatim, S. Emiliani, O. Schwartz, M. Benkirane, SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx, Nature. 474 (2011) 654–657. https://doi.org/10.1038/nature10117.

[136] K. Hrecka, C. Hao, M. Gierszewska, S.K. Swanson, M. Kesik-Brodacka, S. Srivastava, L. Florens, M.P. Washburn, J. Skowronski, Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein, Nature. 474 (2011) 658–661. https://doi.org/10.1038/nature10195.

[137] B. Strack, A. Calistri, S. Craig, E. Popova, H.G. Göttlinger, AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding, Cell. 114 (2003) 689–699. https://doi.org/10.1016/s0092-8674(03)00653-6.

[138] J.E. Garrus, U.K. von Schwedler, O.W. Pornillos, S.G. Morham, K.H. Zavitz, H.E. Wang, D.A. Wettstein, K.M. Stray, M. Côté, R.L. Rich, D.G. Myszka, W.I. Sundquist, Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding, Cell. 107 (2001) 55–65. https://doi.org/10.1016/s0092-8674(01)00506-2.

[139] M. Friedrich, C. Setz, F. Hahn, A. Matthaei, K. Fraedrich, P. Rauch, P. Henklein, M. Traxdorf, T. Fossen, U. Schubert, Glutamic Acid Residues in HIV-1 p6 Regulate Virus Budding and Membrane Association of Gag, Viruses. 8 (2016) 117. https://doi.org/10.3390/v8040117.

[140] J. Votteler, L. Neumann, S. Hahn, F. Hahn, P. Rauch, K. Schmidt, N. Studtrucker, S.M.Ø. Solbak, T. Fossen, P. Henklein, D.E. Ott, G. Holland, N. Bannert, U. Schubert, Highly conserved serine residue 40 in HIV-1 p6 regulates capsid processing and virus core assembly, Retrovirology. 8 (2011) 11. https://doi.org/10.1186/1742-4690-8-11.

[141] Highly conserved serine residue 40 in HIV-1 p6 regulates capsid processing and virus core assembly | SpringerLink, (n.d.). https://link.springer.com/article/10.1186/1742-4690-8-11 (accessed April 6, 2021).

[142] T. Fossen, V. Wray, K. Bruns, J. Rachmat, P. Henklein, U. Tessmer, A. Maczurek, P. Klinger, U. Schubert, Solution Structure of the Human Immunodeficiency Virus

Type 1 p6 Protein*, J. Biol. Chem. 280 (2005) 42515–42527. https://doi.org/10.1074/jbc.M507375200.

[143] H. Zhu, H. Jian, L.-J. Zhao, Identification of the 15FRFG domain in HIV-1 Gag p6 essential for Vpr packaging into the virion, Retrovirology. 1 (2004) 26. https://doi.org/10.1186/1742-4690-1-26.

[144] M. Wanaguru, K.N. Bishop, HIV-1 Gag recruits oligomeric Vpr via two binding sites in p6, but both mature p6 and Vpr are rapidly lost upon target cell entry, J. Virol. (2021) JVI0055421. https://doi.org/10.1128/JVI.00554-21.

[145] M.A. Accola, A.A. Bukovsky, M.S. Jones, H.G. Göttlinger, A conserved dileucine-containing motif in p6(gag) governs the particle association of Vpx and Vpr of simian immunodeficiency viruses SIV(mac) and SIV(agm), J. Virol. 73 (1999) 9992–9999. https://doi.org/10.1128/JVI.73.12.9992-9999.1999.

[146] X.J. Yao, R.A. Subbramanian, N. Rougeau, F. Boisvert, D. Bergeron, E.A. Cohen, Mutagenic analysis of human immunodeficiency virus type 1 Vpr: role of a predicted N-terminal alpha-helical structure in Vpr nuclear localization and virion incorporation, J. Virol. 69 (1995) 7032–7044. https://doi.org/10.1128/JVI.69.11.7032-7044.1995.

[147] D. Schwefel, H.C.T. Groom, V.C. Boucherit, E. Christodoulou, P.A. Walker, J.P. Stoye, K.N. Bishop, I.A. Taylor, Structural basis of lentiviral subversion of a cellular protein degradation pathway, Nature. 505 (2014) 234–238. https://doi.org/10.1038/nature12815.

[148] H.A. Pancio, L. Ratner, Human immunodeficiency virus type 2 Vpx-Gag interaction, J. Virol. 72 (1998) 5271–5275. https://doi.org/10.1128/JVI.72.6.5271-5275.1998.

[149] L. Jin, Y. Zhou, L. Ratner, HIV type 2 Vpx interaction with Gag and incorporation into virus-like particles, AIDS Res. Hum. Retroviruses. 17 (2001) 105–111. https://doi.org/10.1089/08892220150217193.

[150] M. Novikova, Y. Zhang, E.O. Freed, K. Peng, Multiple Roles of HIV-1 Capsid during the Virus Replication Cycle, Virol. Sin. 34 (2019) 119–134. https://doi.org/10.1007/s12250-019-00095-3.

[151] J. Votteler, W.I. Sundquist, Virus budding and the ESCRT pathway, Cell Host Microbe. 14 (2013) 232–241. https://doi.org/10.1016/j.chom.2013.08.012.

[152] J. Lippincott-Schwartz, E. o. Freed, S. b. van Engelenburg, A Consensus View of ESCRT-Mediated Human Immunodeficiency Virus Type 1 Abscission, Annu. Rev. Virol. 4 (2017) 309–325. https://doi.org/10.1146/annurev-virology-101416-041840.

[153] L.-A. Carlson, J.A.G. Briggs, B. Glass, J.D. Riches, M.N. Simon, M.C. Johnson, B. Müller, K. Grünewald, H.-G. Kräusslich, Three-dimensional analysis of budding sites and released virus suggests a revised model for HIV-1 morphogenesis, Cell Host Microbe. 4 (2008) 592–599. https://doi.org/10.1016/j.chom.2008.10.013.

[154] B. Houck-Loomis, M.A. Durney, C. Salguero, N. Shankar, J.M. Nagle, S.P. Goff, V.M. D'Souza, An equilibrium-dependent retroviral mRNA switch regulates translational recoding, Nature. 480 (2011) 561–564. https://doi.org/10.1038/nature10657.

[155] C. Tang, J.M. Louis, A. Aniana, J.-Y. Suh, G.M. Clore, Visualizing transient events in amino-terminal autoprocessing of HIV-1 protease, Nature. 455 (2008) 693–696. https://doi.org/10.1038/nature07342.

[156] R.S. Harris, J.F. Hultquist, D.T. Evans, The restriction factors of human immunodeficiency virus, J. Biol. Chem. 287 (2012) 40875–40883. https://doi.org/10.1074/jbc.R112.416925.

[157] S. Nisole, J.P. Stoye, A. Saïb, TRIM family proteins: retroviral restriction and antiviral defence, Nat. Rev. Microbiol. 3 (2005) 799–808. https://doi.org/10.1038/nrmicro1248.

[158] L.C. James, A.H. Keeble, Z. Khan, D.A. Rhodes, J. Trowsdale, Structural basis for PRYSPRY-mediated tripartite motif (TRIM) protein function, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 6200–6205. https://doi.org/10.1073/pnas.0609174104.

[159] J.-S. Woo, J.-H. Imm, C.-K. Min, K.-J. Kim, S.-S. Cha, B.-H. Oh, Structural and functional insights into the B30.2/SPRY domain, EMBO J. 25 (2006) 1353–1363. https://doi.org/10.1038/sj.emboj.7600994.

[160] M. Stremlau, C.M. Owens, M.J. Perron, M. Kiessling, P. Autissier, J. Sodroski, The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys, Nature. 427 (2004) 848–853. https://doi.org/10.1038/nature02343.

[161] D.M. Sayah, E. Sokolskaja, L. Berthoux, J. Luban, Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV-1, Nature. 430 (2004) 569–573. https://doi.org/10.1038/nature02777.

[162] Y. Li, X. Li, M. Stremlau, M. Lee, J. Sodroski, Removal of arginine 332 allows human TRIM5alpha to bind human immunodeficiency virus capsids and to restrict infection, J. Virol. 80 (2006) 6738–6744. https://doi.org/10.1128/JVI.00270-06.

[163] M. Stremlau, M. Perron, S. Welikala, J. Sodroski, Species-specific variation in the B30.2(SPRY) domain of TRIM5alpha determines the potency of human immunodeficiency virus restriction, J. Virol. 79 (2005) 3139–3145. https://doi.org/10.1128/JVI.79.5.3139-3145.2005.

[164] M.W. Yap, S. Nisole, J.P. Stoye, A single amino acid change in the SPRY domain of human Trim5alpha leads to HIV-1 restriction, Curr. Biol. CB. 15 (2005) 73–78. https://doi.org/10.1016/j.cub.2004.12.042.

[165] M.G. Grütter, J. Luban, TRIM5 structure, HIV-1 capsid recognition, and innate immune signaling, Curr. Opin. Virol. 2 (2012) 142–150. https://doi.org/10.1016/j.coviro.2012.02.003.

[166] T. Pertel, S. Hausmann, D. Morger, S. Züger, J. Guerra, J. Lascano, C. Reinhard, F.A. Santoni, P.D. Uchil, L. Chatel, A. Bisiaux, M.L. Albert, C. Strambio-De-Castillia, W. Mothes, M. Pizzato, M.G. Grütter, J. Luban, TRIM5 is an innate immune sensor for the retrovirus capsid lattice, Nature. 472 (2011) 361–365. https://doi.org/10.1038/nature09976.

[167] S. de Silva, L. Wu, TRIM5 acts as more than a retroviral restriction factor, Viruses. 3 (2011) 1204–1209. https://doi.org/10.3390/v3071204.

[168] H. Aydin, M.W. Taylor, J.E. Lee, Structure-guided analysis of the human APOBEC3-HIV restrictome, Struct. Lond. Engl. 1993. 22 (2014) 668–684. https://doi.org/10.1016/j.str.2014.02.011.

[169] C. Chaipan, J.L. Smith, W.-S. Hu, V.K. Pathak, APOBEC3G Restricts HIV-1 to a Greater Extent than APOBEC3F and APOBEC3DE in Human Primary CD4+ T Cells and Macrophages, J. Virol. 87 (2013) 444–453. https://doi.org/10.1128/JVI.00676-12.

[170] R.M. Kohli, R.W. Maul, A.F. Guminski, R.L. McClure, K.S. Gajula, H. Saribasak, M.A. McMahon, R.F. Siliciano, P.J. Gearhart, J.T. Stivers, Local sequence targeting in the AID/APOBEC family differentially impacts retroviral restriction and antibody diversification, J. Biol. Chem. 285 (2010) 40956–40964. https://doi.org/10.1074/jbc.M110.177402.

[171] K. Bourara, T.J. Liegler, R.M. Grant, Target Cell APOBEC3C Can Induce Limited G-to-A Mutation in HIV-1, PLOS Pathog. 3 (2007) e153. https://doi.org/10.1371/journal.ppat.0030153.

[172] A. Harari, M. Ooms, L.C.F. Mulder, V. Simon, Polymorphisms and splice variants influence the antiretroviral activity of human APOBEC3H, J. Virol. 83 (2009) 295–303. https://doi.org/10.1128/JVI.01665-08.

[173] Y. Dang, L.M. Siew, X. Wang, Y. Han, R. Lampen, Y.-H. Zheng, Human cytidine deaminase APOBEC3H restricts HIV-1 replication, J. Biol. Chem. 283 (2008) 11606–11614. https://doi.org/10.1074/jbc.M707586200.

[174] J.P. Vartanian, A. Meyerhans, B. Asjö, S. Wain-Hobson, Selection, recombination, and G----A hypermutation of human immunodeficiency virus type 1 genomes, J. Virol. 65 (1991) 1779–1788. https://doi.org/10.1128/JVI.65.4.1779-1788.1991.

[175] M. Janini, M. Rogers, D.R. Birx, F.E. McCutchan, Human immunodeficiency virus type 1 DNA sequences genetically damaged by hypermutation are often abundant in patient peripheral blood mononuclear cells and may be generated during near-simultaneous infection and activation of CD4(+) T cells, J. Virol. 75 (2001) 7973–7986. https://doi.org/10.1128/jvi.75.17.7973-7986.2001.

[176] M. Imahashi, M. Nakashima, Y. Iwatani, Antiviral Mechanism and Biochemical Basis of the Human APOBEC3 Family, Front. Microbiol. 3 (2012). https://doi.org/10.3389/fmicb.2012.00250.

[177] K. Luo, B. Liu, Z. Xiao, Y. Yu, X. Yu, R. Gorelick, X.-F. Yu, Amino-terminal region of the human immunodeficiency virus type 1 nucleocapsid is required for human APOBEC3G packaging, J. Virol. 78 (2004) 11841–11852. https://doi.org/10.1128/JVI.78.21.11841-11852.2004.

[178] D. Ebrahimi, F. Anwar, M.P. Davenport, APOBEC3G and APOBEC3F rarely co-mutate the same HIV genome, Retrovirology. 9 (2012) 113. https://doi.org/10.1186/1742-4690-9-113.

[179] X. Wang, P.T. Dolan, Y. Dang, Y.-H. Zheng, Biochemical differentiation of APOBEC3F and APOBEC3G proteins associated with HIV-1 life cycle, J. Biol. Chem. 282 (2007) 1585–1594. https://doi.org/10.1074/jbc.M610150200.

[180] S. Gallois-Montbrun, R.K. Holmes, C.M. Swanson, M. Fernández-Ocaña, H.L. Byers, M.A. Ward, M.H. Malim, Comparison of cellular ribonucleoprotein complexes associated with the APOBEC3F and APOBEC3G antiviral proteins, J. Virol. 82 (2008) 5636–5642. https://doi.org/10.1128/JVI.00287-08.

[181] R.K. Holmes, F.A. Koning, K.N. Bishop, M.H. Malim, APOBEC3F can inhibit the accumulation of HIV-1 reverse transcription products in the absence of hypermutation. Comparisons with APOBEC3G, J. Biol. Chem. 282 (2007) 2587–2595. https://doi.org/10.1074/jbc.M607298200.

[182] X. Jia, Q. Zhao, Y. Xiong, HIV suppression by host restriction factors and viral immune evasion, Curr. Opin. Struct. Biol. 31 (2015) 106–114. https://doi.org/10.1016/j.sbi.2015.04.004.

[183] B.P. Doehle, A. Schäfer, B.R. Cullen, Human APOBEC3B is a potent inhibitor of HIV-1 infectivity and is resistant to HIV-1 Vif, Virology. 339 (2005) 281–288. https://doi.org/10.1016/j.virol.2005.06.005.

[184] J.F. Hultquist, J.A. Lengyel, E.W. Refsland, R.S. LaRue, L. Lackey, W.L. Brown, R.S. Harris, Human and rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-deficient HIV-1, J. Virol. 85 (2011) 11220–11234. https://doi.org/10.1128/JVI.05238-11.

[185] S. Kupzig, V. Korolchuk, R. Rollason, A. Sugden, A. Wilde, G. Banting, Bst-2/HM1.24 Is a Raft-Associated Apical Membrane Protein with an Unusual Topology, Traffic. 4 (2003) 694–709. https://doi.org/10.1034/j.1600-0854.2003.00129.x.

[186] J.F. Arias, Y. Iwabu, K. Tokunaga, Structural Basis for the Antiviral Activity of BST-2/Tetherin and Its Viral Antagonism, Front. Microbiol. 2 (2011). https://doi.org/10.3389/fmicb.2011.00250.

[187] S.J.D. Neil, T. Zang, P.D. Bieniasz, Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu, Nature. 451 (2008) 425–430. https://doi.org/10.1038/nature06553.

[188] N. Van Damme, D. Goff, C. Katsura, R.L. Jorgenson, R. Mitchell, M.C. Johnson, E.B. Stephens, J. Guatelli, The interferon-induced protein BST-2 restricts HIV-1 release and is downregulated from the cell surface by the viral Vpu protein, Cell Host Microbe. 3 (2008) 245–252. https://doi.org/10.1016/j.chom.2008.03.001.

[189] C.A. Kim, M. Gingery, R.M. Pilpa, J.U. Bowie, The SAM domain of polyhomeotic forms a helical polymer, Nat. Struct. Biol. 9 (2002) 453–457. https://doi.org/10.1038/nsb802.

[190] F. Qiao, J.U. Bowie, The many faces of SAM, Sci. STKE Signal Transduct. Knowl. Environ. 2005 (2005) re7. https://doi.org/10.1126/stke.2862005re7.

[191] L. Aravind, E.V. Koonin, The HD domain defines a new superfamily of metal-dependent phosphohydrolases, Trends Biochem. Sci. 23 (1998) 469–472. https://doi.org/10.1016/s0968-0004(98)01293-6.

[192] N. Beloglazova, P. Petit, R. Flick, G. Brown, A. Savchenko, A.F. Yakunin, Structure and activity of the Cas3 HD nuclease MJ0384, an effector enzyme of the CRISPR interference, EMBO J. 30 (2011) 4616–4627. https://doi.org/10.1038/emboj.2011.377.

[193] I.A. Oussenko, R. Sanchez, D.H. Bechhofer, Bacillus subtilis YhaM, a member of a new family of 3'-to-5' exonucleases in gram-positive bacteria, J. Bacteriol. 184 (2002) 6250–6259. https://doi.org/10.1128/jb.184.22.6250-6259.2002.

[194] M.D. Zimmerman, M. Proudfoot, A. Yakunin, W. Minor, Structural insight into the mechanism of substrate specificity and catalytic activity of an HD-domain phosphohydrolase: the 5'-deoxyribonucleotidase YfbR from Escherichia coli, J. Mol. Biol. 378 (2008) 215–226. https://doi.org/10.1016/j.jmb.2008.02.036.

[195] S.R. Kornberg, I.R. Lehman, M.J. Bessman, E.S. Simms, A. Kornberg, Enzymatic cleavage of deoxyguanosine triphosphate to deoxyguanosine and tripolyphosphate, J. Biol. Chem. 233 (1958) 159–162.

[196] P.H. Tan, S.C. Beutelspacher, S.-A. Xue, Y.-H. Wang, P. Mitchell, J.C. McAlister, D.F.P. Larkin, M.O. McClure, H.J. Stauss, M.A. Ritter, G. Lombardi, A.J.T. George,

Modulation of human dendritic-cell function following transduction with viral vectors: implications for gene therapy, Blood. 105 (2005) 3824–3832. https://doi.org/10.1182/blood-2004-10-3880.

[197] C. Goujon, L. Jarrosson-Wuillème, J. Bernaud, D. Rigal, J.-L. Darlix, A. Cimarelli, With a little help from a friend: increasing HIV transduction of monocyte-derived dendritic cells with virion-like particles of SIV(MAC), Gene Ther. 13 (2006) 991–994. https://doi.org/10.1038/sj.gt.3302753.

[198] G. Chougui, S. Munir-Matloob, R. Matkovic, M.M. Martin, M. Morel, H. Lahouassa, M. Leduc, B.C. Ramirez, L. Etienne, F. Margottin-Goguet, HIV-2/SIV viral protein X counteracts HUSH repressor complex, Nat. Microbiol. 3 (2018) 891–897. https://doi.org/10.1038/s41564-018-0179-6.

[199] L. Yurkovetskiy, M.H. Guney, K. Kim, S.L. Goh, S. McCauley, A. Dauphin, W.E. Diehl, J. Luban, Primate immunodeficiency virus proteins Vpx and Vpr counteract transcriptional repression of proviruses by the HUSH complex, Nat. Microbiol. 3 (2018) 1354–1361. https://doi.org/10.1038/s41564-018-0256-x.

[200] R.T. Timms, I.A. Tchasovnikarova, P.J. Lehner, Position-effect variegation revisited: HUSHing up heterochromatin in human cells, BioEssays News Rev. Mol. Cell. Dev. Biol. 38 (2016) 333–343. https://doi.org/10.1002/bies.201500184.

[201] M.V. Andersen, Purification and characterization of an unusual DNA glycosylase in diatoms, (n.d.) 80.

[202] H.E. Krokan, R. Standal, G. Slupphaug, DNA glycosylases in the base excision repair of DNA., Biochem. J. 325 (1997) 1–16.

[203] H.E. Krokan, F. Drabløs, G. Slupphaug, Uracil in DNA – occurrence, consequences and repair, Oncogene. 21 (2002) 8935–8948. https://doi.org/10.1038/sj.onc.1205996.

[204] H. Nilsen, I. Rosewell, P. Robins, C.F. Skjelbred, S. Andersen, G. Slupphaug, G. Daly, H.E. Krokan, T. Lindahl, D.E. Barnes, Uracil-DNA glycosylase (UNG)-deficient mice reveal a primary role of the enzyme during DNA replication, Mol. Cell. 5 (2000) 1059–1065. https://doi.org/10.1016/s1097-2765(00)80271-3.

[205] K.E. Willetts, F. Rey, I. Agostini, J.M. Navarro, Y. Baudat, R. Vigne, J. Sire, DNA repair enzyme uracil DNA glycosylase is specifically incorporated into human

immunodeficiency virus type 1 viral particles through a Vpr-independent mechanism, J. Virol. 73 (1999) 1682–1688. https://doi.org/10.1128/JVI.73.2.1682-1688.1999.

[206] H. Kara, N. Chazal, S. Bouaziz, Is Uracil-DNA Glycosylase UNG2 a New Cellular Weapon Against HIV-1?, Curr. HIV Res. 17 (2019) 148–160. https://doi.org/10.2174/1570162X17666190821154331.

[207] S. Priet, N. Gros, J.-M. Navarro, J. Boretto, B. Canard, G. Quérat, J. Sire, HIV-1-associated uracil DNA glycosylase activity controls dUTP misincorporation in viral DNA and is essential to the HIV-1 life cycle, Mol. Cell. 17 (2005) 479–490. https://doi.org/10.1016/j.molcel.2005.01.016.

[208] J. Ahn, T. Vu, Z. Novince, J. Guerrero-Santoro, V. Rapic-Otrin, A.M. Gronenborn, HIV-1 Vpr loads uracil DNA glycosylase-2 onto DCAF1, a substrate recognition subunit of a cullin 4A-ring E3 ubiquitin ligase for proteasome-dependent degradation, J. Biol. Chem. 285 (2010) 37333–37341. https://doi.org/10.1074/jbc.M110.133181.

[209] P.R. Tedbury, S.G. Sarafianos, Exposing HIV's weaknesses, J. Biol. Chem. 292 (2017) 6027–6028. https://doi.org/10.1074/jbc.H117.777714.

[210] W.W.-Y. Yim, N. Mizushima, Lysosome biology in autophagy, Cell Discov. 6 (2020) 1–12. https://doi.org/10.1038/s41421-020-0141-7.

[211] P. Saftig, J. Klumperman, Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function, Nat. Rev. Mol. Cell Biol. 10 (2009) 623–635. https://doi.org/10.1038/nrm2745.

[212] L. Zhao, S. Wang, M. Xu, Y. He, X. Zhang, Y. Xiong, H. Sun, H. Ding, W. Geng, H. Shang, G. Liang, Vpr counteracts the restriction of LAPTM5 to promote HIV-1 infection in macrophages, Nat. Commun. 12 (2021) 3691. https://doi.org/10.1038/s41467-021-24087-8.

[213] J. He, S. Choe, R. Walker, P. Di Marzio, D.O. Morgan, N.R. Landau, Human immunodeficiency virus type 1 viral protein R (Vpr) arrests cells in the G2 phase of the cell cycle by inhibiting p34cdc2 activity., J. Virol. 69 (1995) 6705–6711. https://doi.org/10.1128/JVI.69.11.6705-6711.1995.

[214] J.B. Jowett, V. Planelles, B. Poon, N.P. Shah, M.L. Chen, I.S. Chen, The human immunodeficiency virus type 1 vpr gene arrests infected T cells in the G2 + M phase of the cell cycle, J. Virol. 69 (1995) 6304–6313.

[215] V. Planelles, F. Bachelerie, J.B. Jowett, A. Haislip, Y. Xie, P. Banooni, T. Masuda, I.S. Chen, Fate of the human immunodeficiency virus type 1 provirus in infected cells: a role for vpr., J. Virol. 69 (1995) 5883–5889. https://doi.org/10.1128/JVI.69.9.5883-5889.1995.

[216] F. Re, D. Braaten, E.K. Franke, J. Luban, Human immunodeficiency virus type 1 Vpr arrests the cell cycle in G2 by inhibiting the activation of p34cdc2-cyclin B., J. Virol. 69 (1995) 6859–6864. https://doi.org/10.1128/JVI.69.11.6859-6864.1995.

[217] M.E. Rogel, L.I. Wu, M. Emerman, The human immunodeficiency virus type 1 vpr gene prevents cell proliferation during chronic infection., J. Virol. 69 (1995) 882–888. https://doi.org/10.1128/JVI.69.2.882-888.1995.

[218] S. Angers, T. Li, X. Yi, M.J. MacCoss, R.T. Moon, N. Zheng, Molecular architecture and assembly of the DDB1–CUL4A ubiquitin ligase machinery, Nature. 443 (2006) 590–593. https://doi.org/10.1038/nature05175.

[219] Y.J. He, C.M. McCall, J. Hu, Y. Zeng, Y. Xiong, DDB1 functions as a linker to recruit receptor WD40 proteins to CUL4-ROC1 ubiquitin ligases, Genes Dev. 20 (2006) 2949–2954. https://doi.org/10.1101/gad.1483206.

[220] L.A. Higa, H. Zhang, Stealing the spotlight: CUL4-DDB1 ubiquitin ligase docks WD40-repeat proteins to destroy, Cell Div. 2 (2007) 5. https://doi.org/10.1186/1747-1028-2-5.

[221] J. Jin, E.E. Arias, J. Chen, J.W. Harper, J.C. Walter, A Family of Diverse Cul4-Ddb1-Interacting Proteins Includes Cdt2, which Is Required for S Phase Destruction of the Replication Factor Cdt1, Mol. Cell. 23 (2006) 709–721. https://doi.org/10.1016/j.molcel.2006.08.010.

[222] A. Brandariz-Nuñez, J. Valle-Casuso, T.E. White, N. Laguette, M. Benkirane, J. Brojatsch, F. Diaz-Griffero, Role of SAMHD1 nuclear localization in restriction of HIV-1 and SIVmac, Retrovirology. 9 (2012) 49. https://doi.org/10.1186/1742-4690-9-49.

[223] H. Hofmann, E.C. Logue, N. Bloch, W. Daddacha, S.B. Polsky, M.L. Schultz, B. Kim, N.R. Landau, The Vpx Lentiviral Accessory Protein Targets SAMHD1 for Degradation in the Nucleus, J. Virol. 86 (2012) 12552–12560. https://doi.org/10.1128/JVI.01657-12.

[224] L. Yurkovetskiy, M.H. Guney, K. Kim, S.L. Goh, S. McCauley, A. Dauphin, W. Diehl, J. Luban, Primate immunodeficiency virus Vpx and Vpr counteract transcriptional repression of proviruses by the HUSH complex, Molecular Biology, 2018. https://doi.org/10.1101/293001.

[225] N. Laguette, C. Brégnard, P. Hue, J. Basbous, A. Yatim, M. Larroque, F. Kirchhoff, A. Constantinou, B. Sobhian, M. Benkirane, Premature Activation of the SLX4 Complex by Vpr Promotes G2/M Arrest and Escape from Innate Immune Sensing, Cell. 156 (2014) 134–145. https://doi.org/10.1016/j.cell.2013.12.011.

[226] G. Kroemer, L. Galluzzi, C. Brenner, Mitochondrial Membrane Permeabilization in Cell Death, Physiol. Rev. 87 (2007) 99–163. https://doi.org/10.1152/physrev.00013.2006.

[227] T. Roumier, H.L. Vieira, M. Castedo, K.F. Ferri, P. Boya, K. Andreau, S. Druillennec, N. Joza, J.M. Penninger, B. Roques, G. Kroemer, The C-terminal moiety of HIV-1 Vpr induces cell death via a caspase-independent mitochondrial pathway, Cell Death Differ. 9 (2002) 1212–1219. https://doi.org/10.1038/sj.cdd.4401089.

[228] N.N. Win, H. Ngwe, I. Abe, H. Morita, Naturally occurring Vpr inhibitors from medicinal plants of Myanmar, J. Nat. Med. 71 (2017) 579–589. https://doi.org/10.1007/s11418-017-1104-7.

[229] N.N. Win, T. Ito, T. Matsui, S. Aimaiti, T. Kodama, H. Ngwe, Y. Okamoto, M. Tanaka, Y. Asakawa, I. Abe, H. Morita, Isopimarane diterpenoids from Kaempferia pulchra rhizomes collected in Myanmar and their Vpr inhibitory activity, Bioorg. Med. Chem. Lett. 26 (2016) 1789–1793. https://doi.org/10.1016/j.bmcl.2016.02.036.

[230] S.-Y. Woo, N.N. Win, W.M. Noe Oo, H. Ngwe, T. Ito, I. Abe, H. Morita, Viral protein R inhibitors from Swertia chirata of Myanmar, J. Biosci. Bioeng. 128 (2019) 445–449. https://doi.org/10.1016/j.jbiosc.2019.04.006.

[231] N.N. Win, T. Kodama, K.Z.W. Lae, Y.Y. Win, H. Ngwe, I. Abe, H. Morita, Bis-iridoid and iridoid glycosides: Viral protein R inhibitors from Picrorhiza kurroa collected in Myanmar, Fitoterapia. 134 (2019) 101–107. https://doi.org/10.1016/j.fitote.2019.02.016.

[232] B.H. Havsteen, The biochemistry and medical significance of the flavonoids, Pharmacol. Ther. 96 (2002) 67–202. https://doi.org/10.1016/s0163-7258(02)00298-x.

[233] M. Shimura, Y. Zhou, Y. Asada, T. Yoshikawa, K. Hatake, F. Takaku, Y. Ishizaka, Inhibition of Vpr-Induced Cell Cycle Abnormality by Quercetin: A Novel Strategy for Searching Compounds Targeting Vpr, Biochem. Biophys. Res. Commun. 261 (1999) 308–316. https://doi.org/10.1006/bbrc.1999.0994.

[234] N. Watanabe, Y. Nishihara, T. Yamaguchi, A. Koito, H. Miyoshi, H. Kakeya, H. Osada, Fumagillin suppresses HIV-1 infection of macrophages through the inhibition of Vpr activity, FEBS Lett. 580 (2006) 2598–2602. https://doi.org/10.1016/j.febslet.2006.04.007.

[235] M. Kamata, R.P. Wu, D.S. An, J.P. Saxe, R. Damoiseaux, M.E. Phelps, J. Huang, I.S.Y. Chen, Cell-based chemical genetic screen identifies damnacanthal as an inhibitor of HIV-1 Vpr induced cell death, Biochem. Biophys. Res. Commun. 348 (2006) 1101–1106. https://doi.org/10.1016/j.bbrc.2006.07.158.

[236] E.B.B. Ong, N. Watanabe, A. Saito, Y. Futamura, K.H. Abd El Galil, A. Koito, N. Najimudin, H. Osada, Vipirinin, a Coumarin-based HIV-1 Vpr Inhibitor, Interacts with a Hydrophobic Region of VPR*, J. Biol. Chem. 286 (2011) 14049–14056. https://doi.org/10.1074/jbc.M110.185397.

[237] C. Zhang, C. Rasmussen, L.J. Chang, Cell cycle inhibitory effects of HIV and SIV Vpr and Vpx in the yeast Schizosaccharomyces pombe, Virology. 230 (1997) 103–112. https://doi.org/10.1006/viro.1997.8459.

[238] L.J. Chang, C.H. Chen, V. Urlacher, T.Z. Lee, Differential apoptosis effects of primate lentiviral Vpr and Vpx in mammalian cells, J. Biomed. Sci. 7 (2000) 322–333. https://doi.org/10.1007/BF02253252.

[239] M. Bodéus, F. Margottin, H. Durand, E. Rouer, R. Benarous, Inhibition of prokaryotic cell growth by HIV1 Vpr, Res. Virol. 148 (1997) 207–213. https://doi.org/10.1016/s0923-2516(97)83990-8.

[240] G.F. Salgado, A. Vogel, R. Marquant, S.E. Feller, S. Bouaziz, I.D. Alves, The Role of Membranes in the Organization of HIV-1 Gag p6 and Vpr: p6 Shows High Affinity for Membrane Bilayers Which Substantially Increases the Interaction between p6 and Vpr., J. Med. Chem. 52 (2009) 7157–7162. https://doi.org/10.1021/jm901106t.

[241] S.M.Ø. Solbak, T.R. Reksten, F. Hahn, V. Wray, P. Henklein, P. Henklein, Ø. Halskau, U. Schubert, T. Fossen, HIV-1 p6 — a structured to flexible multifunctional membrane-interacting protein, Biochim. Biophys. Acta BBA - Biomembr. 1828 (2013) 816–823. https://doi.org/10.1016/j.bbamem.2012.11.010.

[242] M. Cai, Y. Huang, R. Yang, R. Craigie, G.M. Clore, A simple and robust protocol for high-yield expression of perdeuterated proteins in Escherichia coli grown in shaker flasks, J. Biomol. NMR. 66 (2016) 85–91. https://doi.org/10.1007/s10858-016-0052-y.

[243] H. de Rocquigny, D. Ficheux, C. Gabus, M.-C. Fournié-Zaluski, J.-L. Darlix, B.P. Roques, First large scale chemical synthesis of the 72 amino acid HIV-1 nucleocapsid protein NCp7 in an active form, Biochem. Biophys. Res. Commun. 180 (1991) 1010–1018. https://doi.org/10.1016/S0006-291X(05)81166-0.

[244] M. Saunders, A. Wishnia, J.G. Kirkwood, THE NUCLEAR MAGNETIC RESONANCE SPECTRUM OF RIBONUCLEASE1, (2002). https://doi.org/10.1021/ja01569a083.

[245] M.P. Williamson, T.F. Havel, K. Wüthrich, Solution conformation of proteinase inhibitor IIA from bull seminal plasma by 1H nuclear magnetic resonance and distance geometry, J. Mol. Biol. 182 (1985) 295–315. https://doi.org/10.1016/0022-2836(85)90347-x.

[246] K. Wüthrich, Protein structure determination in solution by NMR spectroscopy, J. Biol. Chem. 265 (1990) 22059–22062.

[247] K. Wüthrich, NMR with Proteins and Nucleic Acids, Europhys. News. 17 (1986) 11–13. https://doi.org/10.1051/epn/19861701011.

[248] M.J. Howard, Protein NMR spectroscopy, Curr. Biol. 8 (1998) R331–R333. https://doi.org/10.1016/S0960-9822(98)70214-3.

[249] M. Nilges, Calculation of protein structures with ambiguous distance restraints. Automated assignment of ambiguous NOE crosspeaks and disulphide connectivities, J. Mol. Biol. 245 (1995) 645–660. https://doi.org/10.1006/jmbi.1994.0053.

[250] M. Nilges, M.J. Macias, S.I. O'Donoghue, H. Oschkinat, Automated NOESY interpretation with ambiguous distance restraints: the refined NMR solution structure of the pleckstrin homology domain from beta-spectrin, J. Mol. Biol. 269 (1997) 408–422. https://doi.org/10.1006/jmbi.1997.1044.

[251] R.H. Fogh, W. Boucher, W.F. Vranken, A. Pajon, T.J. Stevens, T.N. Bhat, J. Westbrook, J.M.C. Ionides, E.D. Laue, A framework for scientific data modeling and automated software development, Bioinforma. Oxf. Engl. 21 (2005) 1678–1684. https://doi.org/10.1093/bioinformatics/bti234.

[252] A.T. Brünger, P.D. Adams, G.M. Clore, W.L. DeLano, P. Gros, R.W. Grosse-Kunstleve, J.S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, G.L. Warren, Crystallography & NMR system: A new software suite for macromolecular structure determination, Acta Crystallogr. D Biol. Crystallogr. 54 (1998) 905–921. https://doi.org/10.1107/s0907444998003254.

[253] M.P. Williamson, Using chemical shift perturbation to characterise ligand binding, Prog. Nucl. Magn. Reson. Spectrosc. 73 (2013) 1–16. https://doi.org/10.1016/j.pnmrs.2013.02.001.

[254] A. Furukawa, T. Konuma, S. Yanaka, K. Sugase, Quantitative analysis of protein-ligand interactions by NMR, Prog. Nucl. Magn. Reson. Spectrosc. 96 (2016) 47–57. https://doi.org/10.1016/j.pnmrs.2016.02.002.

[255] J.A. Purslow, B. Khatiwada, M.J. Bayro, V. Venditti, NMR Methods for Structural Characterization of Protein-Protein Complexes, Front. Mol. Biosci. 7 (2020) 9. https://doi.org/10.3389/fmolb.2020.00009.

[256] S.-H. Lin, G. Guidotti, Purification of membrane proteins, Methods Enzymol. 463 (2009) 619–629. https://doi.org/10.1016/S0076-6879(09)63035-4.

[257] B.S. Antharavally, K.A. Mallia, M.M. Rosenblatt, A.M. Salunkhe, J.C. Rogers, P. Haney, N. Haghdoost, Efficient removal of detergents from proteins and peptides in

a spin column format, Anal. Biochem. 416 (2011) 39–44. https://doi.org/10.1016/j.ab.2011.05.013.

[258] J.-L. Rigaud, D. Levy, G. Mosser, O. Lambert, Detergent removal by non-polar polystyrene beads, Eur. Biophys. J. 27 (1998) 305–319. https://doi.org/10.1007/s002490050138.

[259] K. Ohlendieck, Removal of detergent from protein fractions, Methods Mol. Biol. Clifton NJ. 59 (1996) 305–312. https://doi.org/10.1385/0-89603-336-8:305.

[260] N. Morellet, S. Bouaziz, P. Petitjean, B.P. Roques, NMR Structure of the HIV-1 Regulatory Protein VPR, J. Mol. Biol. 327 (2003) 215–227. https://doi.org/10.1016/S0022-2836(03)00060-3.

[261] Y. Wu, X. Zhou, C.O. Barnes, M. DeLucia, A.E. Cohen, A.M. Gronenborn, J. Ahn, G. Calero, The DDB1–DCAF1–Vpr–UNG2 crystal structure reveals how HIV-1 Vpr steers human UNG2 toward destruction, Nat. Struct. Mol. Biol. 23 (2016) 933–940. https://doi.org/10.1038/nsmb.3284.

[262] R. Montserret, N. Saint, C. Vanbelle, A.G. Salvay, J.-P. Simorre, C. Ebel, N. Sapay, J.-G. Renisio, A. Böckmann, E. Steinmann, T. Pietschmann, J. Dubuisson, C. Chipot, F. Penin, NMR Structure and Ion Channel Activity of the p7 Protein from Hepatitis C Virus, J. Biol. Chem. 285 (2010) 31446–31461. https://doi.org/10.1074/jbc.M110.122895.

[263] B.P. Oestringer, J.H. Bolivar, J.K. Claridge, L. Almanea, C. Chipot, F. Dehez, N. Holzmann, J.R. Schnell, N. Zitzmann, Hepatitis C virus sequence divergence preserves p7 viroporin structural and dynamic features, Sci. Rep. 9 (2019) 8383. https://doi.org/10.1038/s41598-019-44413-x.

[264] R. Ranjith Kumar, P. Hanumantha Rao, M. Arumugam, Lipid Extraction Methods from Microalgae: A Comprehensive Review, Front. Energy Res. 2 (2015). https://doi.org/10.3389/fenrg.2014.00061.

[265] M.H.H. Roby, Synthesis and Characterization of Phenolic Lipids, IntechOpen, 2017. https://doi.org/10.5772/66891.

[266] W.K. Downey, M.K. Keogh, R.F. Murphy, Lipid separation on sephadex LH-20, Biochem. J. 110 (1968) 13P-14P. https://doi.org/10.1042/bj1100013pb.

[267] W. Rieping, M. Habeck, B. Bardiaux, A. Bernard, T.E. Malliavin, M. Nilges, ARIA2: automated NOE assignment and data integration in NMR structure calculation, Bioinforma. Oxf. Engl. 23 (2007) 381–382. https://doi.org/10.1093/bioinformatics/btl589.

[268] N. Sunseri, M. O'Brien, N. Bhardwaj, N.R. Landau, Human immunodeficiency virus type 1 modified to package Simian immunodeficiency virus Vpx efficiently infects macrophages and dendritic cells, J. Virol. 85 (2011) 6263–6274. https://doi.org/10.1128/JVI.00346-11.

[269] K. Hänel, L. Möckel, M. Brummel, K. Peiris, R. Hartmann, A.J. Dingley, D. Willbold, A. Loidl-Stahlhofen, Expression and purification of soluble HIV-2 viral protein R (Vpr) using a sandwich-fusion protein strategy, Protein Expr. Purif. 95 (2014) 156–161. https://doi.org/10.1016/j.pep.2013.12.007.

[270] Q. Fu, A. Piai, W. Chen, K. Xia, J.J. Chou, Structure determination protocol for transmembrane domain oligomers, Nat. Protoc. 14 (2019) 2483–2520. https://doi.org/10.1038/s41596-019-0188-9.

[271] L. Columbus, J. Lipfert, K. Jambunathan, D.A. Fox, A.Y.L. Sim, S. Doniach, S.A. Lesley, Mixing and Matching Detergents for Membrane Protein NMR Structure Determination, J. Am. Chem. Soc. 131 (2009) 7320–7326. https://doi.org/10.1021/ja808776j.

[272] Q. Wang, S. Su, J. Xue, F. Yu, J. Pu, W. Bi, S. Xia, Y. Meng, C. Wang, W. Yang, W. Xu, Y. Zhu, Q. Zheng, C. Qin, S. Jiang, L. Lu, An amphipathic peptide targeting the gp41 cytoplasmic tail kills HIV-1 virions and infected cells, Sci. Transl. Med. 12 (2020). https://doi.org/10.1126/scitranslmed.aaz2254.

[273] C. Beauvineau, Conception et synthese de nouvelles sondes ciblees pour l'imagerie moleculaire, phdthesis, Université Pierre et Marie Curie - Paris VI, 2011. https://pastel.archives-ouvertes.fr/pastel-00644998 (accessed October 3, 2021).

Appendices

1. Assigned chemical shifts of HIV-1 Vpr

The assigned chemical shifts of HIV-1 Vpr in a buffer 10mM sodium acetate, pH 4, 50mM NaCl, 5mM β-mercaptoethanol and 120 mM DPC at final concentration of 0.3 mM.

2. Assigned chemical shifts of HIV-1 p6

The assigned chemical shifts of HIV-1 p6 in a buffer 10mM sodium acetate, pH 4, 50mM NaCl, 5mM β-mercaptoethanol and 100 mM DPC-d_{38} at final concentration of 1 mM.

3. Assigned chemical shifts of SIVmac Vpx

The assigned chemical shifts of SIVmac Vpx in a buffer 10mM sodium acetate, pH 4, 50mM NaCl, 5mM β-mercaptoethanol and 280 mM DPC-d38 at final concentration of 0.7 mM.

4. Assigned chemical shifts of SIVmac p6$^{16-28}$

The assigned chemical shifts of SIVmac p616-28 in a buffer 10mM sodium acetate, pH 4, 50mM NaCl, 5mM β-mercaptoethanol and 100 mM DPC-d38 at final concentration of 1 mM.

5. Assigned chemical shifts of SIVmac p6

The assigned chemical shifts of SIVmac p6 in a buffer 10mM sodium acetate, pH 4, 50mM NaCl, 5mM β-mercaptoethanol and 100 mM DPC-d_{38} at final concentration of 1 mM.

6. Titration of HIV-1 p6 by DPC micelles
7. Chemical shift perturbation of labeled HIV-1 Vpr by unlabeled HIV-1 p6
8. Chemical shift perturbation of labeled SIVmac Vpx by unlabeled SIVmac p6
9. Chemical shift perturbation of labeled HIV-1 p6 by unlabeled SIVmac Vpx
10. Chemical shift perturbation of labeled HIV-1 Vpr by DO1148
11. Paper published

Received: 28 June 2021 | Revised: 17 August 2021 | Accepted: 17 August 2021

DOI: 10.1002/pro.4174

METHODS AND APPLICATIONS

THE PROTEIN SOCIETY WILEY

Acetonitrile allows indirect replacement of nondeuterated lipid detergents by deuterated lipid detergents for the nuclear magnetic resonance study of detergent-soluble proteins

Xiao Wang | **Xiaowei Chen** | **Sylvie Nonin-Lecomte** | **Serge Bouaziz**

CiTCoM, CNRS, UMR 8038, Université de Paris, Paris, France

Correspondence
Serge Bouaziz, CiTCoM, CNRS, UMR 8038, Université de Paris, 4 Avenue de l'Observatoire, Paris 75270, France.
Email: serge.bouaziz@parisdescartes.fr

Funding information
Centre National de la Recherche Scientifique; China Scholarship Council, Grant/Award Number: 201709110124; Université Paris Descartes

Abstract

Detergent-soluble proteins (DSPs) are commonly dissolved in lipid buffers for NMR experiments, but the huge lipid proton signal prevents recording of high-quality spectra. The use of costly deuterated lipids is thus required to replace nondeuterated ones. With conventional methods, detergents like dodecylphosphocholine (DPC) cannot be fully exchanged due to their high binding affinity to hydrophobic proteins. We propose an original and simple protocol which combines the use of acetonitrile, dialysis and lyophilization to disrupt the binding of lipids to the protein and allow their indirect replacement by their deuterated equivalents, while maintaining the native structure of the protein. Moreover, by this protocol, the detergent-to-protein molar ratio can be controlled as it challenges the protein structure. This protocol was applied to solubilize the Vpx protein that was followed upon addition of DPC-d_{38} by ^{1}H-^{15}N SOFAST-HMQC spectra and the best detergent-to-DSPs molar ratio was obtained for structural studies.

KEYWORDS

3D structure, acetonitrile, detergent-soluble proteins, DPC-to-protein molar ratio, NMR

1 | INTRODUCTION

The hydrophobicity of some proteins challenges their purification in liquid phase as they are not soluble in aqueous buffers. Denaturation agents such as urea and guanidine hydrochloride can be used for solubilization,[1–5] but at the expense of the protein native structure. A renaturation step is therefore required, but the process of renaturation remains poorly controlled and the ultimately obtained structure may be different from the native conformation. Lipid detergents such as DPC and SDS are alternatively employed to successfully solubilize these proteins called DSPs of which the human amylin and the SARS-CoV-2 spike proteins.[6–10] Unfortunately, the large number of protons of these detergents prevents the recording of NMR spectra with acceptable signal-to-noise ratio, thus compromising subsequent analyses like structure determination and interaction studies with protein or ligand partners for the identification of new drugs and targets. Detergents, often lipids, should be replaced by their deuterated equivalents. The price of deuterated detergents is high and prevents their use from the first to

Abbreviations: CD, circular dichroism; DPC, dodecylphosphocholine; DPC-d_{38}, deuterated DPC; DSP, detergent-soluble protein; HMQC, heteronuclear multiple quantum coherence; NMR, nuclear magnetic resonance; SDS, sodium dodecyl sulphate; SIVmac, macaque simian immunodeficiency virus; TFA, trifluoroacetic acid; TFE, trifluoroethanol; TMD, transmembrane domain.

the last steps of purification. Various detergent removal methods were described such as gel filtration, dialysis, precipitation, and ion-exchange chromatography,[11–14] but the effectiveness of their removal is still uncertain since some detergents such as DPC and SDS can bind to DSPs so tightly that they cannot be removed. A solution might be to perform the full extraction and purification procedures in deuterated detergents, but the high cost of such a protocol is also a deterrent. Therefore, a new efficient and economical solution is required for the removal and exchange of lipid detergents.

Literature shows that some DSPs were well solubilized and folded in organic solvents such as acetonitrile or TFE, albeit not under near physiological conditions. For example, the synthetic HIV-1 Vpr adopts a three α-helical structure in the presence of acetonitrile,[15] which is almost identical to the structure determined later by X-ray crystallography.[16] Similarly, the structure of HCV p7 is mainly α-helical, whether in detergents or in organic solvent/water mixtures as shown by CD analyses.[17] Subsequently, the proteins HIV-1 Vpr and HCV p7 were reported to be well structured in detergent micelles.[18,19] Thus, these proteins are well folded in organic solvents and detergents micelles. Lipids are soluble and can be purified in organic solvents.[20–22]

Structure, multimerization, and aggregation of DSPs strongly depend on the detergent-to-DSP molar ratio. For example, the oligomerization of HCV p7 protein was reported to be highly sensitive to the detergent-to-protein molar ratio.[23] It is important for structural and biophysical studies to retain the native structure of the protein and thus to determine the optimal lipid-to-DSP ratio to ensure native and monodisperse structuration of the protein. A slight change in the detergent-to-protein ratio can induce chemical shift perturbations in NMR spectra,[24] which hampers further analyses and identification of protein–protein or protein–ligand interactions by the method of chemical shifts perturbations induced upon partner binding.

In this study, we present a new protocol which ensures total removal of the lipids while retaining the structural properties of the DSP based on our assumption that the detergent molecules coating the protein can be easily and totally replaced by organic solvent molecules during dialysis. In a second step, the deuterated detergent can be added to the buffer, and a near membrane condition can be obtained by eliminating the organic solvent by lyophilization. Thus, the aim of this work was to replace the nondeuterated DPC by deuterated DPC (DPC-d_{38}). Then to determine the best detergent-to-protein ratio, a series of high-resolution ^{1}H-^{15}N SOFAST-HMQC[25] spectra were recorded in different detergent-to-DSP molar ratios. Finally, a 300:1 molar ratio was defined and obtained when the molar ratio of 400:1 only induced very slight change compared with the equivalent at 300:1. In conclusion, we performed an indirect replacement of detergents by their deuterated equivalents and determined best detergent-to-DSPs molar ratio.

2 | RESULTS

2.1 | Acetonitrile disrupts the binding of detergent molecules

As a prototype, we used the highly hydrophobic protein Vpx, a DSP from SIVmac. Vpx extraction after bacteria cells lysis and purification were both conducted in protonated DPC buffer. To eliminate DPC, acetonitrile was used to efficiently disrupt the binding of DPC to Vpx (Figure 1). Proton NMR spectra confirm that DPC can be removed completely during dialysis steps against acetonitrile (Figure 2). Organic solvents are thus effective to disrupt the binding of the coating detergent molecules.

2.2 | Acetonitrile retains the structural properties of the DSP

Maintaining the native structure of the protein is very important for the structural studies. The elimination of detergents leads to the precipitation or unfold conformation of DSPs. Some DSPs are well-folded in both detergent and organic solvent. Here, to confirm the structural character of Vpx in the context of acetonitrile, a ^{1}H-^{15}N SOFAST-HMQC spectra was recorded (Figure 3). Our NMR spectra show that acetonitrile buffer retains the structural conformation of Vpx and prevents its aggregation by offering adequate hydrophobic conditions.

2.3 | DPC-d_{38} prevents aggregation during lyophilization

After the elimination of DPC micelles, we directly lyophilized the sample and found that the lyophilized sample without DPC-d_{38} was aggregated and somehow attached to the wall of the Eppendorf (Figure 4a). Then, we dissolved the lyophilized sample with DPC-d_{38} in acetate buffer, but only a small part of them was solubilized. We propose that DSPs without DPC may aggregate during the freeze-drying process. Thus, prior to lyophilization, the sample was supplied with 100 equivalents of DPC-d_{38} to avoid protein aggregation upon acetonitrile lyophilization. We found that the lyophilized sample in the presence of DPC-d_{38} was porous and fibrous (Figure 4b)

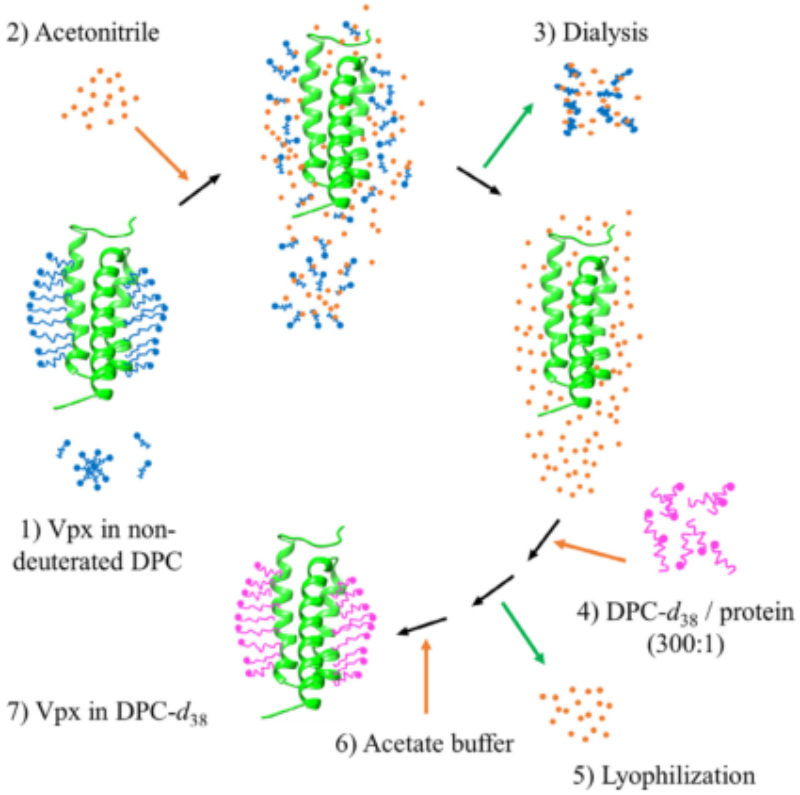


FIGURE 1 A flowchart of nondeuterated DPC indirectly replaced by deuterated DPC-d_{38}. Binding of DPC detergent favors and maintains Vpx conformation and prevents aggregation. The interaction of Vpx with DPC molecules was disrupted by acetonitrile and DPC was then present in free form. The original conformation of the protein was conserved under the hydrophobic environment. Prior to lyophilization, the deuterated DPC-d_{38} was supplied to maintain and protect Vpx protein. Finally, the lyophilized sample was dissolved in acetate buffer for NMR studies

indicating that DPC-d_{38} can act as a freeze-drying protective agent.

2.4 | The best detergent-to-DSP molar ratio is determined by NMR

DSPs usually have different spectrum at different detergent-to-DSP molar ratios suggesting that the structure of DSPs evolves at different detergent-to-DSP molar ratios. Thus, structure, multimerization, and aggregation of DSPs strongly depend on the detergent-to-DSP molar ratio. To obtain the best detergent-to-DSP molar ratio, the precise DPC-d_{38}-to-Vpx ratio for the structural and interaction studies is determined by NMR, using only a small amount of labeled protein (~60 nmol per tube). Different detergent-to-Vpx molar ratios, starting with 100:1, were assayed by ^{1}H-^{15}N SOFAST-HMQC spectra until no significant chemical shift perturbations of the protein resonances occurred (Figure 5a–e). We found only very slight changes between molar ratios of 300:1 and 400:1 (Figure 5f). Therefore, a 300:1 DPC-d_{38}-to-Vpx ratio was considered as suitable for structural analysis by NMR. In addition, the spectrum at a 300:1 DPC-d_{38}-to-Vpx molar ratio is identical to the spectrum at the same DPC-to-Vpx ratio (Figure S1). Two residues A96 and A112 located at the C-terminus of Vpx have slight change in chemical shift probably resulting from indirect effects of the flexibility of the C-terminus.

3 | DISCUSSION

The purification of DSPs is usually performed using denaturing agents such as urea or guanidinium salts, in which DSPs may have different conformations before and after renaturation. Maintaining the native structure of the protein is very important for structural studies and the use of detergents can avoid the denaturation or aggregation of DSPs during extraction and purification. Unfortunately, the large number of protons contained in these detergents complicates NMR spectra analysis and compromises the downstream structure determination of the protein and its interaction with ligands for the identification of new drugs. Detergent removal efficiency is still a challenge since various methods cannot remove

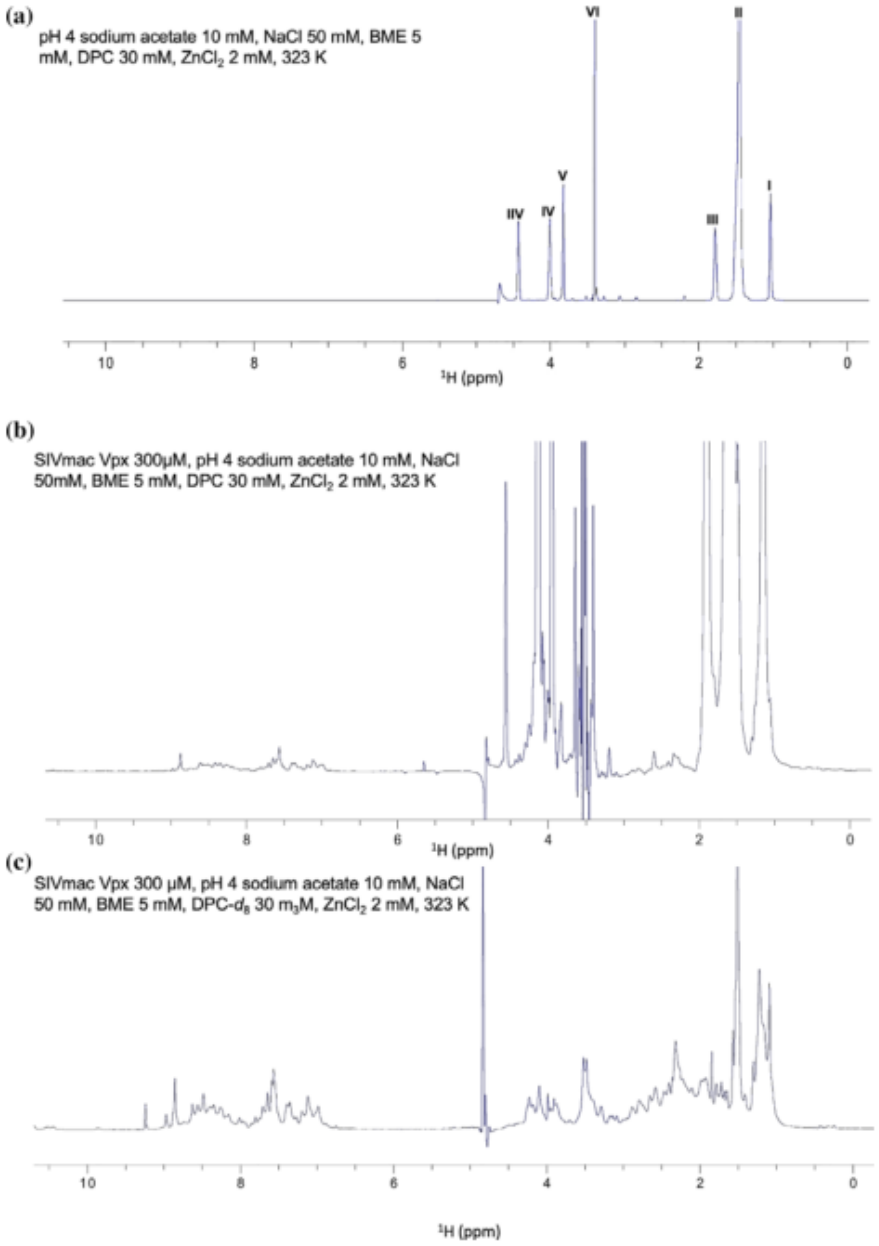


FIGURE 2 Efficiency of DPC removal followed by NMR. All spectra were recorded in acetate buffer A at 323 K and pH 4.0. (a) 1D proton spectrum of DPC (30 mM). The seven proton resonances (labelled I–VII) are from DPC. (b) 0.3 mM Vpx proton spectrum with 100 equivalents protonated DPC. (c) 0.3 mM Vpx with 100 equivalents DPC-d_{38} after lyophilization and resuspension according to the protocol described here. All DPC proton signals were suppressed allowing proper observation on the protein spectrum

efficiently detergent bound to the protein. The extremely high price of deuterated detergents must also be considered.

Organic solvents are widely used to study the protein properties in biochemical physics, biotechnology, and biomedicine. The advantages of the organic solvents are

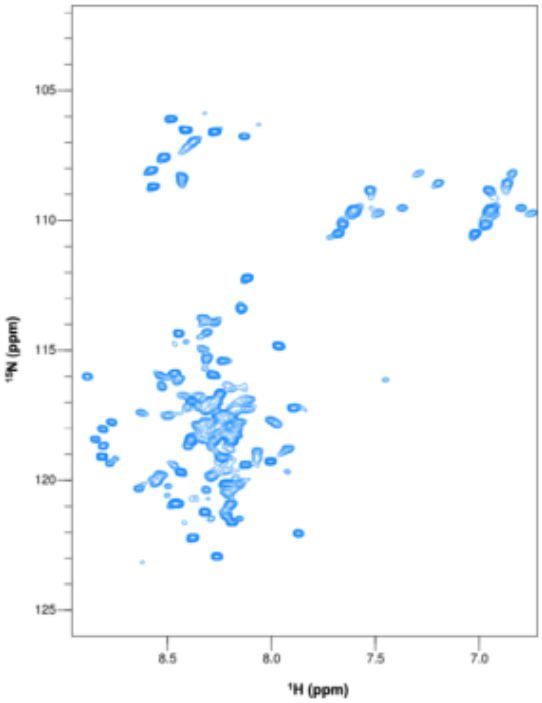


FIGURE 3 ^{1}H-^{15}N SOFAST-HMQC spectrum of 0.3 mM ^{13}C, ^{15}N-labeled Vpx in 30% acetonitrile recorded at 323 K and 600 MHz

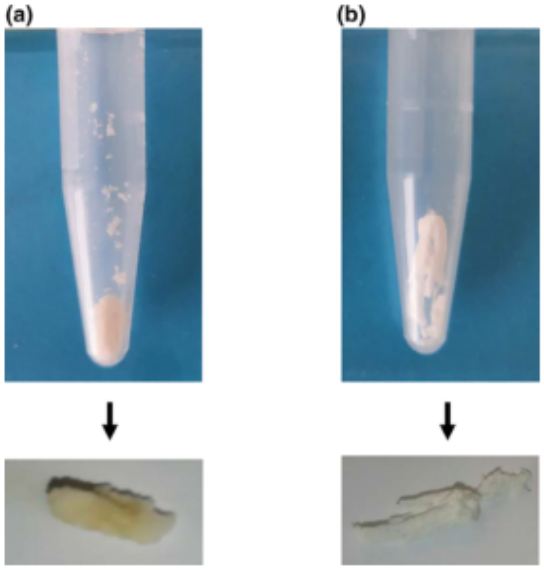


FIGURE 4 DPC as a freeze-drying protective agent. (a) Vpx solubilized in 30% acetonitrile and DPC-free buffer. After lyophilization, the protein self-aggregates, and powder is attached on the wall of the Eppendorf. (b) Deuterated DPC-d_{38} added in a 100:1 M ratio to Vpx in acetonitrile, prior to lyophilization. The resulting powder is loose and porous

that they overcome unwanted water-dependent side reactions and provide a hydrophobic environment in which hydrophobic force is involved in maintain 3D structure of proteins.[26–28] Among several organic solvents, acetonitrile and TFE are most widely utilized due to their low viscosity and high hydrophobicity.[29] In our protocol, acetonitrile allows efficient exchange of nondeuterated lipid detergents by deuterated ones, while retaining the 3D structure of the protein. It is possible, then, to record high-resolution NMR spectra, free of interfering proton signals from detergent, required for NMR studies of proteins isolated or in interaction with their partners. Importantly, Vpx in 30% acetonitrile solution still maintains its 3D structure during the substitution between DPC and DPC-d_{38}. Similar with the structure of Vpx in acetonitrile, our previous study showed that HIV-1 Vpr is organized around three well defined α-helices in 30% acetonitrile solution.[30] Other proteins in acetonitrile aqueous such as subtilisin Carlsburg and human insulin are well characterized in their 3D structure.[31–33] In addition, solution NMR structure of protein-G B1 domain displays native-like β-hairpin tertiary structure in 30% TFE solution.[34] TFE and acetonitrile are polar solvents that can easily repel water from the protein surface. Consequently, side chains of polar residues in proteins become more rigid, and favor interactions with other residues by intramolecular hydrogen bonds, which explains why certain organic solvents maintain 3D structures and secondary structural elements. However, the 3D structure of proteins could be perturbed in organic solvents including TFA and acetonitrile. Acetonitrile and TFE have strong electronegativity so that they can form strong hydrogen bonds with various hydrogen donors.[35] Thus, acetonitrile and TFE perturb 3D structure of proteins by disrupting intramolecular hydrogen bonds. In addition, Gekko et al. reported that the stability of the 3D structure of protein depends on the organic solvent concentration using hen-egg lysozyme as a prototype.[36] Their results suggest that the perturbation of 3D structure does not occur at low concentrations of acetonitrile (below 40% acetonitrile). However, the hydrophobic bonding capacity for nonpolar sidechains is weakened with the addition of acetonitrile, consequently, leading to destabilization of the 3D structure of proteins.[36] In our study, 30% acetonitrile, a low concentration used, is favorable for maintaining native state of Vpx. Thus, a preferential concentration of acetonitrile or other organic solvents such as TFE and methanol needs to be investigated to ensure the degree of solubilization and folding of DSPs when applying our protocol.

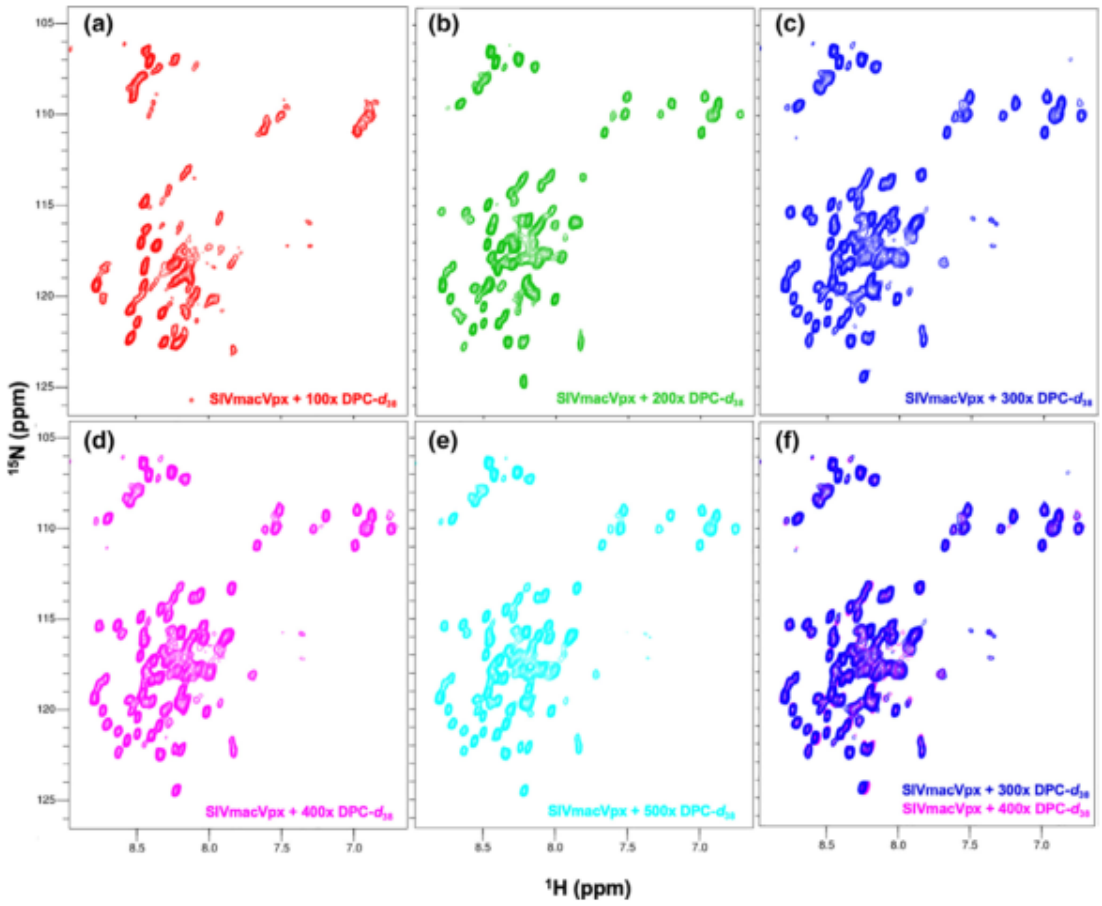


FIGURE 5 Series of ^{1}H-^{15}N SOFAST-HMQC spectra of 0.3 mM ^{13}C, ^{15}N -labeled Vpx recorded in acetate buffer A at 323 K in the presence of DPC-d_{38} for DPC-d_{38}-to-DSP molar ratios of 100:1 (a), 200:1 (b), 300:1 (B), 400:1 (d), and 500:1 (e). Vpx structure was stabilized as DPC-d_{38} quantity increased. (f) Superposition of spectra recorded for 300:1 and 400:1 ratio. A molar ratio of 400:1 only induces slight changes compared to 300:1

On the other hand, this protocol deserves to be tested even if the 3D structure of the protein is lost in the presence of organic solvent. Indeed, in the case where the protein would be able to fold after elimination of the organic solvent by lyophilization, this would make it possible to widen the applicability to other classes of proteins solubilized by a detergent.

Nondeuterated DPC was used to solubilize Vpx and was removed by dialysis against acetonitrile. Then, DPC-d_{38} was supplied, the sample lyophilized, and acetate buffer added for NMR studies. However, all DSPs are not always well-folded in the presence of DPC micelles, although DPC is the most widely used detergent in solution NMR, and it is possible to adapt our protocol to other detergents. Similar with DPC, SDS is a lipid detergent as well. DPC and SDS, not only create a hydrophobic environment through their nonpolar tails preventing protein self-aggregation, but they also provide a suitable polyelectrolytic environment through polar heads allowing protein–protein interactions.[28] Furthermore, SDS has the same long hydrophobic carbon chain $CH_3(CH_2)_{11}$ as DPC suggesting that SDS binds to DSPs in the same way as DPC. Thus, our protocol could be applied to the substitution of SDS. In comparison with the monolayer of DPC or SDS, some other lipid detergents such as 1,2-diheptanoyl-sn-glycero-3-phosphocholine and 1,2-dioleoyl-sn-glycero-3-phosphocholine have bilayer properties. Thus, whether acetonitrile or other organic solvents can be used for the substitution has to be checked.

Vpx is a small size hydrophobic protein, which remains folded under a variety of experimental conditions in this study. Most membrane proteins have one or several TMDs,[37–39] and their TMDs are usually composed by small size α-helical domains. Furthermore, many TMDs are well organized with α-helices in detergents. For example, α-helical structures of TMDs from lysosome-associated membrane protein type 2a, influenza A virus M2, phospholamban, and human nicastrin are well organized in DPC micelles.[40–43] Currently, DPC has been used to obtain 40% of the membrane protein structures determined by solution NMR, making it the most frequently used detergent.[45] Interestingly, organic solvents, including TFA and acetonitrile, can promote or maintain secondary structural elements and mostly work for helical proteins suggesting that many transmembrane domains possibly fold well in certain organic solvents and in detergents. For example, the TMDs of the cannabinoid receptor subtype 2, a member of the well-characterized G-protein coupled receptor superfamily, folds well in acetonitrile and in DPC micelles.[44] If membrane proteins fold well in organic solvents, our protocol could be widely applied for DSPs, membrane proteins, or even protein with more complex topology.

In this study, acetonitrile was used to break down the hydrophobic forces between protein and detergent and was then removed by dialysis. Subsequently, a few tests were made on samples adding very small amount of deuterated detergent prior to acetonitrile removal by lyophilization. Our protocol allows maintaining the original state of the native protein and compared to TMDs[1] protocols, it avoids the process of protein denaturation and renaturation. Our protocol is thus expected to have a broad impact on the research community studying hydrophobic proteins in solution.

4 | MATERIALS AND METHODS

4.1 | Elimination of detergents

The ^{13}C,^{15}N-labeled Vpx was produced in *Escherichia coli*, extracted, and purified by liquid chromatography in acetate buffer A (10 mM sodium acetate, 50 mM NaCl, and 5 mM β-mercaptoethanol buffer at pH 4.0) supplemented with 6 mM protonated DPC. The purified protein was concentrated on a 3 kDa Amicon Ultra-15 and dialyzed at room temperature against 30% acetonitrile/70% H_2O at pH 2.5 to remove salts and DPC. The quantity of protein was determined from the OD_{280} absorbance value using the acetonitrile dialysis buffer as blank control. All NMR spectra were recorded in 3 mm tubes at 323 K on a Bruker Avance III 600 MHz spectrometer equipped with triple resonance cryoprobe and Z-gradients. The efficiency of DPC removal was determined by proton 1D NMR spectra.

4.2 | Acetonitrile removal by lyophilization

Prior to lyophilization, DPC-d_{38} was added to the sample in Eppendorf that was frozen in liquid nitrogen. The frozen sample was lyophilized in a freeze dryer (Cryotec, Lyophilisateur Crios model) and acetonitrile was gradually removed. After lyophilization, acetate buffer was used to re-dissolve the freeze-dry sample. As a control, the sample without DPC-d_{38} was directly lyophilized.

4.3 | Quantification of the best ratio between detergents and DSPs

To determine the best detergent-to-DSP ratio by NMR, a small portion of labeled Vpx protein was aliquoted into 5 Eppendorf (i.e., ~60 nmol per sample) and DPC-d_{38} was individually added at molar ratios ranging from 100:1 to 500:1 prior to lyophilization. After lyophilization, the five samples were resuspended in 200 µl of acetate buffer A. ^1H-^{15}N SOFAST-HMQC[25] spectra were recorded and compared. The best ratio (300:1) is obtained when the addition of DPC-d_{38} no longer induces chemical shifts perturbations on the SOFAST-HMQC spectrum.

ACKNOWLEDGMENTS

We gratefully thank Dr Andrea Cimarelli, Centre International de Recherche en Infectiologie (CIRI), Université de Lyon 1, for kindly providing the plasmid pET11d-SIVmac Vpx. We are grateful to Dr Valéry Larue for his help in setting up the NMR experiments. Xiao Wang was granted by China Scholarship Council (n° 201709110124). This work was supported by the French Centre National de la Recherche Scientifique (CNRS) and the University of Paris.

AUTHOR CONTRIBUTIONS

Xiao WANG and Serge Bouaziz: Conceptualization (equal); data curation (equal); formal analysis (equal); project administration (equal); writing – original draft (equal). **Xiaowei Chen:** Data curation (equal); formal analysis (equal); methodology. **Sylvie Nonin-Lecomte and Serge Bouaziz:** Investigation (equal); methodology (equal); supervision (equal); validation; writing – review and editing (equal).

ORCID

Xiao Wang https://orcid.org/0000-0001-5811-8712
Xiaowei Chen https://orcid.org/0000-0001-8894-9429
Sylvie Nonin-Lecomte https://orcid.org/0000-0001-5155-8088
Serge Bouaziz https://orcid.org/0000-0001-7792-8307

REFERENCES

1. Fu Q, Piai A, Chen W, Xia K, Chou JJ. Structure determination protocol for transmembrane domain oligomers. Nat Protoc. 2019;14:2483–2520.
2. Rossky PJ. Protein denaturation by urea: Slash and bond. Proc Natl Acad Sci U S A. 2008;105:16825–16826.
3. Maxwell KL, Bona D, Liu C, Arrowsmith CH, Edwards AM. Refolding out of guanidine hydrochloride is an effective approach for high-throughput structural studies of small proteins. Protein Sci. 2003;12:2073–2080.
4. Hagihara Y, Aimoto S, Fink AL, Goto Y. Guanidine hydrochloride-induced folding of proteins. J Mol Biol. 1993; 231:180–184.
5. Huang SK, Shin K, Sarker M, Rainey JK. Apela exhibits isoform- and headgroup-dependent modulation of micelle binding, peptide conformation and dynamics. Biochim Biophys Acta Biomembr. 2017;1859:767–778.
6. Rabilloud T, Luche S, Santoni V, Chevallet M. Detergents and chaotropes for protein solubilization before two-dimensional electrophoresis. Methods Mol Biol. 2007;355:111–119.
7. Matar-Merheb R, Rhimi M, Leydier A, et al. Structuring detergents for extracting and stabilizing functional membrane proteins. PLOS ONE. 2011;6:e18036.
8. Linke D. Detergents: an overview. Methods Enzymol. 2009; 463:603–617.
9. Zhang J, Cai Y, Xiao T, et al. Structural impact on SARS-CoV-2 spike protein by D614G substitution. Science. 2021;372:525–530.
10. Patil SM, Xu S, Sheftic SR, Alexandrescu AT. Dynamic alpha-helix structure of micelle-bound human amylin. J Biol Chem. 2009;284:11982–11991.
11. Lin S-H, Guidotti G. Purification of membrane proteins. Methods Enzymol. 2009;463:619–629.
12. Antharavally BS, Mallia KA, Rosenblatt MM, et al. Efficient removal of detergents from proteins and peptides in a spin column format. Anal Biochem. 2011;416:39–44.
13. Rigaud J-L, Levy D, Mosser G, Lambert O. Detergent removal by non-polar polystyrene beads. Eur Biophys J. 1998;27: 305–319.
14. Ohlendieck K. Removal of detergent from protein fractions. Methods Mol Biol. 1996;59:305–312.
15. Morellet N, Bouaziz S, Petitjean P, Roques BP. NMR structure of the HIV-1 regulatory protein VPR. J Mol Biol. 2003;327: 215–227.
16. Wu Y, Zhou X, Barnes CO, et al. The DDB1–DCAF1–Vpr–UNG2 crystal structure reveals how HIV-1 Vpr steers human UNG2 toward destruction. Nat Struct Mol Biol. 2016;23: 933–940.
17. Montserret R, Saint N, Vanbelle C, et al. NMR structure and ion channel activity of the p7 protein from hepatitis C virus. J Biol Chem. 2010;285:31446–31461.
18. Salgado GF, Marquant R, Vogel A, et al. Structural studies of HIV-1 Gag p6ct and its interaction with Vpr determined by solution nuclear magnetic resonance. Biochemistry. 2009;48: 2355–2367.
19. Oestringer BP, Bolivar JH, Claridge JK, et al. Hepatitis C virus sequence divergence preserves p7 viroporin structural and dynamic features. Sci Rep. 2019;9:8383.
20. Ranjith Kumar R, Hanumantha Rao P, Arumugam M. Lipid extraction methods from microalgae: A comprehensive review. Front Energy Res. 2015;2:1–9.
21. Roby MHH. Synthesis and Characterization of Phenolic Lipids. In: Soto-Hernandez M, Ed. (2017) Phenolic Compounds - Natural Sources, Importance and Applications. London, UK: IntechOpen; 2017. pp 89–116.
22. Downey WK, Keogh MK, Murphy RF. Lipid separation on sephadex LH-20. Biochem J. 1968;110:13P–14P.
23. Chen W, OuYang B, Chou JJ. Critical effect of the detergent: Protein ratio on the formation of the hepatitis C virus p7 channel. Biochemistry. 2019;58:3834–3837.
24. Williamson MP. Using chemical shift perturbation to characterise ligand binding. Prog Nucl Magn Reson Spectrosc. 2013; 73:1–16.
25. Favier A, Brutscher B. NMRlib: User-friendly pulse sequence tools for Bruker NMR spectrometers. J Biomol NMR. 2019;73: 199–211.
26. Fetzner S, Lingens F. Bacterial dehalogenases: Biochemistry, genetics, and biotechnological applications. Microbiol Rev. 1994;58:641–685.
27. Torres S, Castro GR. Non-aqueous biocatalysis in homogeneous solvent systems. Food Technol Biotechnol. 2004;42: 271–277.
28. Maurel P, Douzou P, Waldmann J, Yonetani T. Enzyme behaviour and molecular environment. The effects of ionic strength, detergents, linear polyanions and phospholipids on the pH profile of soluble cytochrome oxidase. Biochim Biophys Acta. 1978;525:314–324.
29. Rubinstein M. Preparative high-performance liquid partition chromatography of proteins. Anal Biochem. 1979;98:1–7.
30. Schwefel D, Groom HCT, Boucherit VC, et al. Structural basis of lentiviral subversion of a cellular protein degradation pathway. Nature. 2014;505:234–238.
31. Fitzpatrick PA, Ringe D, Klibanov AM. X-ray crystal structure of cross-linked subtilisin Carlsberg in water vs. acetonitrile. Biochem Biophys Res Commun. 1994;198:675–681.
32. Fedorov AA, Joseph-McCarthy D, Fedorov E, Sirakova D, Graf I, Almo SC. Ionic interactions in crystalline bovine pancreatic ribonuclease a. Biochemistry. 1996;35:15962–15979.
33. Bocian W, Sitkowski J, Bednarek E, Tarnowska A, Kawecki R, Kozerski L. Structure of human insulin monomer in water/acetonitrile solution. J Biomol NMR. 2008;40:55–64.
34. Blanco FJ, Jiménez MA, Pineda A, Rico M, Santoro J, Nieto JL. NMR solution structure of the isolated N-terminal fragment of protein-G B1 domain. Evidence of trifluoroethanol induced native-like beta-hairpin formation. Biochemistry. 1994;33: 6004–6014.
35. Cringus D, Yeremenko S, Pshenichnikov MS, Wiersma DA. Hydrogen bonding and vibrational energy relaxation in water–acetonitrile mixtures. J Phys Chem B. 2004;108:10376–10387.

36. Gekko K, Ohmae E, Kameyama K, Takagi T. Acetonitrile-protein interactions: Amino acid solubility and preferential solvation. Biochim Biophys Acta Protein Struct Mol Enzymol. 1998;1387:195–205.
37. Sharpe HJ, Stevens TJ, Munro S. A comprehensive comparison of transmembrane domains reveals organelle-specific properties. Cell. 2010;142:158–169.
38. White SH, Wimley WC. Membrane protein folding and stability: Physical principles. Annu Rev Biophys Biomol Struct. 1999; 28:319–365.
39. Killian JA, von Heijne G. How proteins adapt to a membrane–water interface. Trends Biochem Sci. 2000;25:429–434.
40. Rout AK, Strub M-P, Piszczek G, Tjandra N. Structure of transmembrane domain of lysosome-associated membrane protein type 2a (LAMP-2A) reveals key features for substrate specificity in chaperone-mediated autophagy. J Biol Chem. 2014;289:35111–35123.
41. Oxenoid K, Chou JJ. The structure of phospholamban pentamer reveals a channel-like architecture in membranes. Proc Natl Acad Sci U S A. 2005;102:10870–10875.
42. Ma C, Polishchuk AL, Ohigashi Y, et al. Identification of the functional core of the influenza A virus A/M2 proton-selective ion channel. Proc Natl Acad Sci U S A. 2009;106:12283–12288.
43. Li Y, Liew LSY, Li Q, Kang C. Structure of the transmembrane domain of human nicastrin-A component of γ-secretase. Sci Rep. 2016;6:19522.
44. Zheng H, Zhao J, Sheng W, Xie X-Q. A transmembrane helix-bundle from G-protein coupled receptor CB2: Biosynthesis, purification, and NMR characterization. Biopolymers. 2006;83: 46–61.
45. Chipot C, Dehez F, Schnell JR, et al. Perturbations of native membrane protein structure in alkyl phosphocholine detergents: A critical assessment of NMR and biophysical studies. Chem Rev. 2018;118:3559–3607.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Wang X, Chen X, Nonin-Lecomte S, Bouaziz S. Acetonitrile allows indirect replacement of nondeuterated lipid detergents by deuterated lipid detergents for the nuclear magnetic resonance study of detergent-soluble proteins. Protein Science. 2021;1–9. https://doi.org/10.1002/pro.4174

FIGURE LEGENDS

Figure 1. A flowchart of nondeuterated DPC indirectly replaced by deuterated DPC-d_{38}. Binding of DPC detergent favors and maintains SIVmac Vpx conformation and prevents aggregation. The interaction of SIVmac Vpx with DPC molecules was disrupted by acetonitrile and DPC was then present in free form. The original conformation of the protein was conserved under the hydrophobic environment. Prior to lyophilization, the deuterated DPC-d_{38} was supplied to maintain and protect SIVmac Vpx protein. Finally, the lyophilized sample was dissolved in acetate buffer for NMR studies.

Figure 2. Efficiency of DPC removal followed by NMR. All spectra were recorded in acetate buffer A at 323 K and pH 4.0. (A) 1D proton spectrum of DPC (30 mM). The seven proton resonances (labelled I to VII) are from DPC. (B) 0.3 mM SIVmac Vpx proton spectrum with 100 equivalents protonated DPC. (C) 0.3 mM SIVmac Vpx with 100 equivalents DPC-d_{38} after lyophilization and resuspension according to the protocol described here. All DPC proton signals were suppressed allowing proper observation on the protein spectrum.

Figure 3. ^{1}H-^{15}N HMQC spectrum of 0.3 mM ^{13}C, ^{15}N-labeled SIVmac Vpx in 30% acetonitrile recorded at 323 K and 600 MHz.

Figure 4. DPC as a freeze-drying protective agent. (A) SIVmac Vpx solubilized in 30% acetonitrile and DPC-free buffer. After lyophilization, the protein auto-aggregates, and the powder is attached on the wall of the eppendorf. (B) Deuterated DPC-d_{38} added in a 100: 1 molar ratio to SIVmac Vpx in acetonitrile, prior to lyophilization. The resulting powder is loose and porous.

Figure 5. Series of ^{1}H-^{15}N HMQC spectra of 0.3 mM ^{13}C, ^{15}N -labeled SIVmac Vpx recorded in acetate buffer A at 323 K in the presence of DPC-d_{38} for DPC-d_{38} -to-DSP molar ratios of 100:1 (A), 200:1 (B), 300:1 (B), 400:1 (D), and 500:1 (E). SIVmac Vpx structure was stabilized as DPC-d_{38} quantity increased. (F) Superposition of spectra recorded for 300:1 and 400:1 ratios. A molar ratio of 400:1 only induces slight changes compared to 300:1.