GENERATION AND EVALUATION OF A FUNCTIONAL SINGLE-CHAIN VARIABLE FRAGMENT (scFv) INTRABODY AGAINST HIV-1 CAPSID PROTEIN (P24)

MOHAMMAD TASYRIQ BIN CHE OMAR

UNIVERSITI SAINS MALAYSIA

2017

GENERATION AND EVALUATION OF A FUNCTIONAL SINGLE-CHAIN VARIABLE FRAGMENT (scFv) INTRABODY AGAINST HIV-1 CAPSID PROTEIN (P24)

by

MOHAMMAD TASYRIQ BIN CHE OMAR

Thesis submitted in the fulfillment of the requirements

for the degree of

Doctor of Philosophy

January 2017

ACKNOWLEDGEMENT

In the name of Allah, the most Gracious and Merciful. I would like to convey my gratitude to my superb supervisor Dr. Syed Atif Ali for his guidance, concern, understanding, and his support throughout the development of this fantastic project. The thesis would not have been possible without his help in the technical and analysis aspect of the project. Thanks to Dr. Muhamad Yusri Bin Musa and Prof. Dr. Narazah Mohd Yusoff for encouraging me throughout the study. My amazing peers; Siti Aisyah Binti Mualif, Ronald Syeang Sin Teow, Chew Yik Wei, Muhammad Alif Che Nordin, Nurdianah Harif Fadzilah, Warren Lee and industrial students, you all have amazed me in different ways.

Mothers; Siti Eshah Binti Che Mat and Fuziyah Binti Abdullah and fathers; Che Omar Bin Ibrahim and Maksir Bin Said, I am so thankful for everything that you have taught me and how you have always surrounded me with so much love and happiness. My deep appreciation goes to my brothers and sisters in supporting me indirectly. Finally, my lovely wife Nur Syuhanis Binti Maksir and two beautiful daughters Fatimah & Zainab, thank you very much for offering your help during the three and half years of Ph.D. study. Mommy and kids, you have helped daddy more than you will ever know. Thank you for making daddy so happy, I love you sweethearts.

For my financial support during this study, I would like to thank School of Distance Education, Universiti Sains Malaysia for offering me with the Academic Staff Training Scheme (ASTS) and also to Ministry of Higher Education (MOHE) for covering the university fees. This Ph.D. study was supported by FRGS grant (203/CIPPT/6711206) and ERGS grant (203/CIPPT/6730104) from Ministry of Higher Education (MOHE) of Malaysia.

TABLE OF CONTENT

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

2D	two-dimensional
A_{260}	absorbance at 260nm
A_{450}	absorbance at 450nm
A_{600}	absorbance at 600nm
AIDS	Acquired immunodeficiency syndrome
Amp	Ampicillin
APS	Ammonium persulfate
ART	Anti-retroviral therapy
ATCC	American Type Culture Collection
Bis	N,N-bis (2-hydroxyethyl)-2-aminoethanesulfonic acid
BFR	Bacterioferritin
bp	Base pair
BSA	bovine serum albumin
CA	capsid p24
CDR	complementary-determining regions
CFU	Colony-forming unit
CHO	Chinese Hamster Ovary
cm	Centimeter
CTD	C-terminal domain
CV	column volume
Da	Dalton
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethylsulfoxide
DNA	Deoxy ribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleoside triphosphate
DsbC	Disulfide bond C isomerase
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbent assay
Env	Envelope
FBS	Fetal Bovine Serum
Fc	Constant fragment

FDA	Food and Drug Adiministration
FPLC	Fast Protein Liquid Chromatography
Fv	variable fragment
GFP	Green-fluorescence protein
gp	glycoprotein
H_2O	water
HAMA	Human anti-mouse antibody
HCl	Hydrochloric acid
HEK293T	Human embroyonic kidney 293 cells T
His	Histidine
HIV	Human immunodeficiency virus
HRP	Horse Reddish Peroxidase
h	Hour
IG	immunoglobulin
IgG	Imunoglobulin gamma
IL-2	Interleukin-2
IMAC	Immobilised metal affinity chromatography
IMGT	The International Immunogenetics Information System®
IMGT/V-QUEST	IMGT (V-QUEry and STandardization)
IN	Integrase
IPTG	Isopropyl-β-D-thiogalactopyranoside
kb	kilo base
kDa	kiloDalton
LB	Luria Bertani media
LTR	Long tandem repeat
MA	Matrix p17
mAb	monoclonal antibdoy
MBP	maltose binding protein
MES	2-(N-morpholine)-ethamesulfonic acid
ml	Mililiter
mm	Milimeter
mM	Milimolar
MW	Molecular weight
Nc	nitrocellulose

NC	nucleocapsid p7
Nef	Negative factor
NIH	National Institute of Health
nm	nanometer
NP-40	Nonidet P-40 (detergent)
NTD	N-terminal domain
NMR	Nuclear magnetic resonance spectroscopy
NusA	N-utilising substance A protein
OD	Optical density
ORF	Open reading frame
PCR	Polymerase chain reaction
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
pH	Potential hydrogen
PHA	Phytohemagglutinin
PHA-M	Phytohemagglutinin M form
Pol	Polymerase
PR	Protease
RE	Restriction enzyme
Rev	Regulatory of expression of viral protein
RNA	Ribonucleic acid
rt	Room temperature (22°C-25°C)
RT	Reverse transcriptase
RT-PCR	Reverse transcriptase- polymerase chain reaction
Sec.	Section
scFv	single chain variable fragment
sdAb	single domain antibody
SDS	Sodium dedecyl sulfate
SDS-PAGE	Sodium dedecyl sulfate polyarylamide gel electrophoresis
T cells	Thymus cells (originated from thymus)
TAE	Tris-Acetate-EDTA
Tat	Transactivator of transcription
Taq	Thermus aquaticus

TEMED	N,N,N,N’-tetramethylethylenediamine
Tm	melting temperature
Tris	Tris(hydroxymethyl)aminomethane
U.S.A	United state of America
UV	Ultraviolet
V	volts
v/v	volume/volume
V_H	variable regions of heavy chain
Vif	Virus infectivity factor
V_L	variable regions of light chain
Vpr	Viral protein r
Vpu	Viral protein u
VSV-G	Vesicular Stomatitis Virus-G protein
w/v	weight/volume
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
µm	Micrometer
µM	Micromolar
µL	Microliter
°C	Degree Celsius

PENGHASILAN DAN PENILAIAN SINGLE-CHAIN VARIABLE FRAGMENT (scFv) INTRABODI YANG BERFUNGSI TERHADAP PROTEIN KAPSID HIV-1 (P24)

ABSTRAK

HIV/AIDS adalah wabak yang menyerang di seluruh dunia. Walaupun kemajuan dalam rawatan HIV telah mengurangkan perkembangan penyakit ini, namun penyakit lain yang dikaitkan dengan AIDS dan kadar kematian masih kekal tinggi. Ini disebabkan oleh kelemahan dalam pelaksanaan rawatan dan kemunculan baka HIV yang mempunyai daya tahan terhadap ubat yang sedia ada. Oleh itu, terdapat keperluan yang berterusan untuk meneroka pendekatan baru bagi membangunkan kaedah rawatan alternatif terhadap HIV/AIDS. Kapsid protein HIV, p24 membentuk teras konikal yang meliputi genom virus dan memainkan peranan yang penting dalam peringkat awal dan akhir kitaran hidup virus. Sebagai sasaran terapeutik yang berpotensi, protein p24 telah menarik perhatian yang agak besar sejak beberapa tahun kebelakangan ini. Penemuan penting yang memperlihatkan faktor-faktor sekatan perumah seperti Trim5α dan Mx2 yang menyasarkan teras konikal p24 memperlihatkan potensi p24 sebagai sasaran pencegahan terhadap virus HIV. Kajian sedia ada bertujuan untuk menghasilkan antibodi single chain variable fragment (scFv) yang khusus terhadap protein p24 HIV-1 dan menilai aktivitinya sebagai intrabodi. Urutan pengekodan rantaian ringan (V_L) dan rantaian berat (V_H) pada bahagian boleh ubah yang diklon dari cDNA sel hybridoma (klon 183-H12-5C) menghasilkan antibodi terhadap p24 HIV. Urutan V_H dan V_L telah dihubungkan dengan satu turutan pengikat peptide yang fleksibel dan diekpres sebagai scFv. scFv anti-p24 yang terhasil berinteraksi secara khusus dengan p24 dari virus dan p24

rekombinan dan menghalang penggabungan p24 di dalam ekperimen tabung uji. Apabila diekspresikan di dalam sel, scFv anti-p24 secara ketara menghalang replikasi virus di dalam sel MAGI, sel Jurkat T, dan PBMC. Kesimpulannya, scFv anti-p24 antibodi yang dihasilkan dalam kajian ini boleh digunapakai sebagai medium yang berharga bagi membangunkan kaedah bio-terapeutik berasaskan antibodi yang berupaya menentang HIV/AIDS.

GENERATION AND EVALUATION OF A FUNCTIONAL SINGLE-CHAIN VARIABLE FRAGMENT (scFv) INTRABODY AGAINST HIV-1 CAPSID PROTEIN (P24)

ABSTRACT

HIV/AIDS is now present in every country in the world. Although advances in the treatment of HIV have reduced the progress of the disease, AIDS related illnesses and deaths remain high due to poor compliance with treatment and emergence of drug -resistant HIV strains. Thus there is a persistent need to explore new approaches for developing alternative treatment modalities against HIV/AIDS. The HIV capsid protein p24 forms the conical core that encapsulates the viral genome inside the virus and plays crucial roles in both early and late stages of viral life cycle. As a potential therapeutic target, the p24 protein has attracted considerable attention in recent years. Key findings that host restriction factors such as Trim5α and Mx2 target incoming conical core accentuate the potential of p24 as a feasible anti-HIV target. The aim of the present study was to generate a single chain variable fragment (scFv) antibody specific to HIV-1 p24 protein and evaluate its anti-HIV activity as intrabody. Sequences encoding variable regions of heavy (V_H) and light (V_L) chains were cloned from cDNA isolated from a hybridoma cell line producing HIV p24 mAb (clone 183-H12-5C). The V_H and V_L sequences were linked by a flexible peptide linker sequence and expressed as scFv. The resulting anti-p24 scFvs interacted specifically with both recombinant and virus -derived p24 and inhibited p24 polymerisation in an in vitro assay. Intracellularly expressed anti-p24 scFv markedly inhibited viral replication in CD4/CCR5 -expressing MAGI, Jurkat T cells, and PBMCs. In conclusion, anti-p24 scFv antibody generated in this work may serve

as a valuable tool for the development of antibody -based biotherapeutics against HIV/AIDS.

CHAPTER 1: INTRODUCTION

1.1 AIDS Epidemic

Acquired immunodeficiency syndrome (AIDS) is one of the biggest threats to human health around the world since it's discovery in 1981 (Centers for Disease Control (CDC) 1981). In 1983, the human immunodeficiency virus (HIV) had been identified as the leading cause of AIDS (Barré-Sinoussi et al. 1983; Gallo et al. 1983). Two viruses being detected liable for the development of AIDS in humans are HIV-1 and HIV-2 which represent multiple independent zoonotic transmission of simian immunodeficiency virus (SIV). HIV-1 was passed into a human through chimpanzee virus (SIVcpz), while HIV-2 was transferred from Sooty mangabeys (Cohen et al. 2008; Sharp & Hahn 2011). In the deficiency of precautionary vaccines, the viruses remain a human pathogen that threatens people's health for the coming decades. Ongoing efforts in designing and developing anti-HIV-1 drugs and vaccines have proven to be insufficient for eradicating or averting HIV/AIDS (Ensoli et al. 2014).

The HIV pandemic documents that approximately 36.7 million people infected by HIV-1 globally (UNAIDS 2015). An estimated 1-2 million infections caused by HIV-2 in Africa and India regions which most of the patient progresses more slowly (Ingole et al. 2013; de Silva & Weiss 2010; Campbell-Yesufu & Gandhi 2011). Developing countries show high cases of AIDS -related illnesses and deaths yearly. As evident, Malaysia has shown a stabilised HIV/AIDS cases over last five years. Since 2002 average of 9 new cases of HIV/AIDS was reported each day (Shitan & Mondal 2011; HIV/STI Section MOH Malaysia 2011; HIV/STI Section Ministry of Health Malaysia 2015)

Once infections take place (Figure 1.1), human can harbour the virus for a decade or more without showing clear clinical appearances (Buchbinder et al. 1994; Learmont et al. 1992). As presented in Figure 1.2, in the lack of treatment, ultimately, the virus replication reduces the circulating $CD4^+$ T cells below 200 cells per cm^3. In contrast, an average of 800 cells per cm^3 is detected for healthy people thus suggesting the depletion of severe immune cells for people that progressed HIV-1/AIDS (Shete et al. 2010; Rodger et al. 2011). This observation promotes to an extensive series of AIDS –related illness (Ledergerber et al. 1999; Bakshi 2004). Indeed, HIV-1 infected $CD4^+$T can produce up to $10x10^9$ virions per day in an infected person and having a half-life of 2.2 days (Perelson et al. 1996).

Implementation of antiretroviral therapy (ART) prolongs life likelihood of individuals who newly diagnosed with AIDS. ART is still the best available treatment for managing HIV/AIDS. A significant effect of drug combinations in the treatment had improved the prognosis of HIV-1 patients which lead to less transmission of the virus to the environment (Bartlett et al. 2001; Bartlett et al. 2006; Egger et al. 2002). Table 1.1 showed FDA approval drugs and the combinations using in ART. It is successful to constrain the pandemic and lowers the mortality. However, drug-resistant HIV-1 strains still can cause the increment of the death rate (Emamzadeh-Fard et al. 2013). Thus, the exploration of new approaches is needed in developing efficient cures to combat HIV/AIDS.

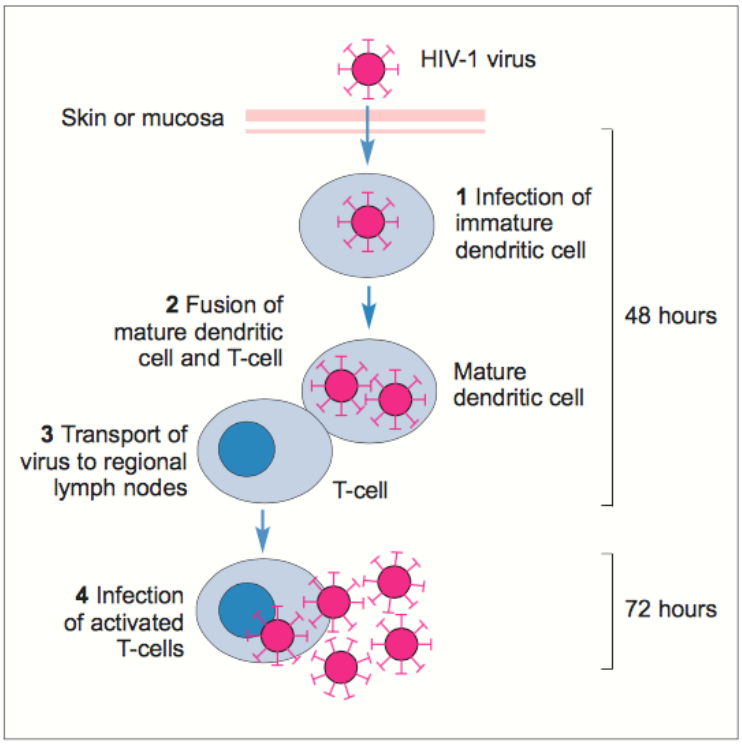

Figure 1.1 HIV transmission and early immunological events. Obtained from (Rodger et al. 2011)

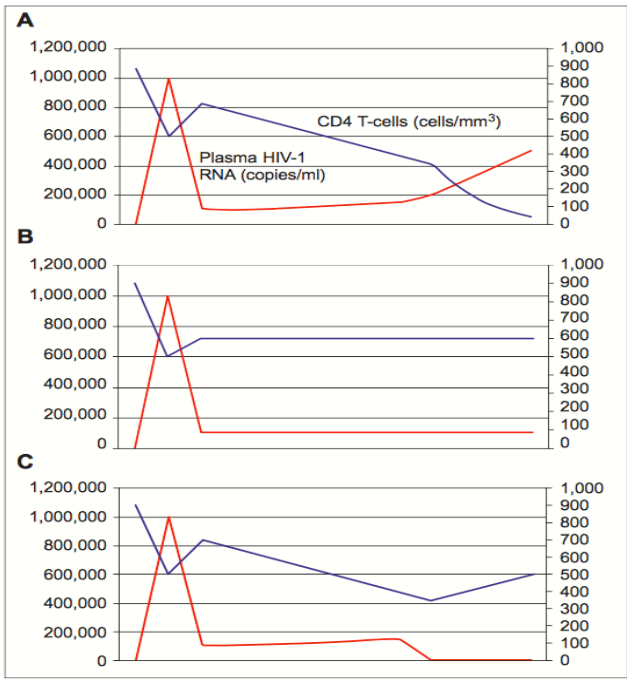

Figure 1.2 Natural history of HIV infection in: (A) untreated progressors; (B) untreated nonprogressors; (C) treated progressors. Obtained from (Rodger et al. 2011)

Table 1.1 List of FDA-approved anti-HIV drugs. Drugs used in anti-retroviral therapy targeting different HIV-1 proteins and their combination partners.

Drug	Generic name	Brand Name	FDA Approval Date
Non-Nucleoside Reverse Transcriptase Inhibitors (NRTIs)			
NRTIs block reverse transcriptase, an enzyme HIV needs to make copies of itself	abacavir	Ziagen	December 17, 1998
	didanosine	Videx	October 9, 1991
	emtricitabine	Emtriva	July 2, 2003
	lamivudine	Epivir	November 17,1995
	stavudine	Zerit	June 24, 1994
	Tenofovir disoproxil fumarate	Viread	October 26, 2001
	zidovudine	Retrovir	March 19, 1987
Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs)			
NNRTIs bind to and later alter reverse transcriptase, an enzyme HIV needs to make copies of itself	efavirenz	Sustiva	September 17, 1998
	etravirine	Intelence	January 18, 2008
	nevirapine	Viramune	June 21, 1996
	rilpivirine	Edurant	May 20, 2011
Protease Inhibitors (PIs)			
PIs block HIV protease, an enzyme HIV needs to make copies of itself	atazanavir	Reyataz	June 20, 2003
	darunavir	Prezista	June 23, 2006
	fosamprenavir	Lexiva	October 20, 2003
	indinavir	Crixivan	March 13, 1996
	nelfinavir	Viracept	March 14, 1997
	ritonavir	Norvir	March 1, 1996
	saquinavir	Invirase	December 6, 1995
	tipranavir	Aptivus	June 22, 2005

Table 1.1. Continued

Drug	Generic name	Brand Name	FDA Approval Date
Integrase Inhibitors (PIs)			
Integrase inhibitors block HIV integrase, an enzyme HIV needs to make copies of itself.	dolutegravir	Tivicay	August 13, 2013
	elvitegravir	Vitekta	September 24, 2014
	raltegravir	Isentress	October 12, 2007
Fusion inhibitors			
Fusion inhibitors block HIV from entering the CD4 cells of the immune system	enfuvirtide	Fuzeon	March 13, 2003
Entry inhibitors			
Entry inhibitors block proteins on the CD4 cells that HIV needs to enter the cells.	maraviroc	Selzentry	August 6, 2007
Pharmacokinetic Enhancers			
Pharmacokinetic enhancers are used in HIV treatment to increase the effectiveness of an HIV medicine included in an HIV regimen	cobicistat	Tybost	September 24, 2014
Combination HIV Medicines			
Combination HIV medicines contain two or more HIV medicines from one or more drug classes	abacavir a lamivudine	Epzicom	August 2, 2004
	abacavir, dolugtegravir, and lamivudine	Triumeq	August 22, 2014
	abacavir, lamivudine, and zidovudine	Trizivir	November 14, 2000

Table 1.1. Continued

Drug	Generic name	Brand Name	FDA Approval Date
Combination HIV Medicines			
Combination HIV medicines contain two or more HIV medicines from one or more drug classes	efavirenz, emtricitabine, and tenofovir disoproxil fumarate	Atripla	July 12, 2006
	elvitegravir, cobicistat, emtricitabine, and tenofovir alafenamide fumarate	Genvoya	November 5, 2015
	elvitegravir, cobicistat, emtricitabine, and tenofovir disoproxil fumarate	Stribild	August 27, 2012
	emtricitabine, rilpivirine and tenofovir alafenamide	Odefsey	March 1, 2016
	emtricitabine, rilpivirine and tenofovir disoproxil fumarate	Complera	August 10, 2011
	emtricitabine, and tenofovir alafenamide	Descovy	April 4, 2016
	emtricitabine, and tenofovir disproxil fumarate	Truvada	August 2, 2004
	lamivudine and zidvudine	Combivir	September 27, 1997
	lopinavir and ritonavir	Kaletra	September 15, 2000
	atazanavir and cobicistat	Evotaz	January 29, 2015
	darunavir and cobicistat	Prezcobix	January 29, 2015

Obtained and modified from https://aidsinfo.nih.gov/education-materials/fact-sheets/21/58/fda-approved-hiv-medicines

1.2 HIV-1 Structure and Genome Organisation

HIV type 1 (HIV-1) is categorised under the human lentivirus genus in *retroviridae* family (Cullen 1991; Rodger et al. 2011). The size of HIV-1 virus ranged between 119 to 207 nm with average diameter of 145 nm (Briggs et al. 2004; Briggs et al. 2003). The ~9.8 kb HIV-1 provirus comprise of the genes that encode at least nine proteins which can be divided into three classes: 1) The structural protein; 2) The regulatory protein; 3) The accessory protein (Frankel & Young 1998; Bruggeman et al. 1994). The ~9.2 kb primary transcript (HIV-1 mRNAs unspliced) is a homodimer of linear, positive-sense, and single-stranded RNA which generated from the expression of 5' LTR to 3' LTR (Moore & Hu 2009; Bohne et al. 2005). It is necessary for the production of mature infectious HIV-1 viruses (Karn & Stoltzfus 2012). Besides, the distribution of Gag and Gag-Pol protein at a ratio of 20:1 in virus-producing cells is also critical for viral infectivity (Shehu-Xhilaga & Crowe 2001; Sundquist & Krausslich 2012).

The Gag, Pol, and Env proteins are the major proteins that are common to all retroviruses (Katz & Skalka 1994; Wills & Craven 1991). In order to form the core structure of HIV-1 virion, the HIV-1 Gag polyprotein undergoes the proteolytic processing by the viral protease (PR) that generates matrix (MA; p17), capsid (CA; p24), nucleocapsid (NC; p7) and p6 (Ganser-Pornillos et al. 2012). The formation of HIV-1 viral envelope is driven by the cleavage of HIV-1 Env precursors (gp160) by the cellular protease furin in the Golgi that produces surface (SU; gp120) and transmembrane (TM; gp41) (Freed 2013). The enzymes that necessarily involve in HIV-1 replication are generated from the proteolytic process of HIV-1 Gag-Pol polyprotein by the viral PR. They are protease (PR), reverse transcriptase (RT), and integrase (IN) (Konnyu et al. 2013). HIV-1 also encodes two regulatory proteins

such Tat and Rev together with four accessory proteins like Vif, Vpr, Nef and Vpu which facilitate in the enhancing of viral infectivity, regulatory functions, and proper viral assembly (Frankel & Young 1998). Figure 1.3 shows the structure and genome organisation of HIV-1.

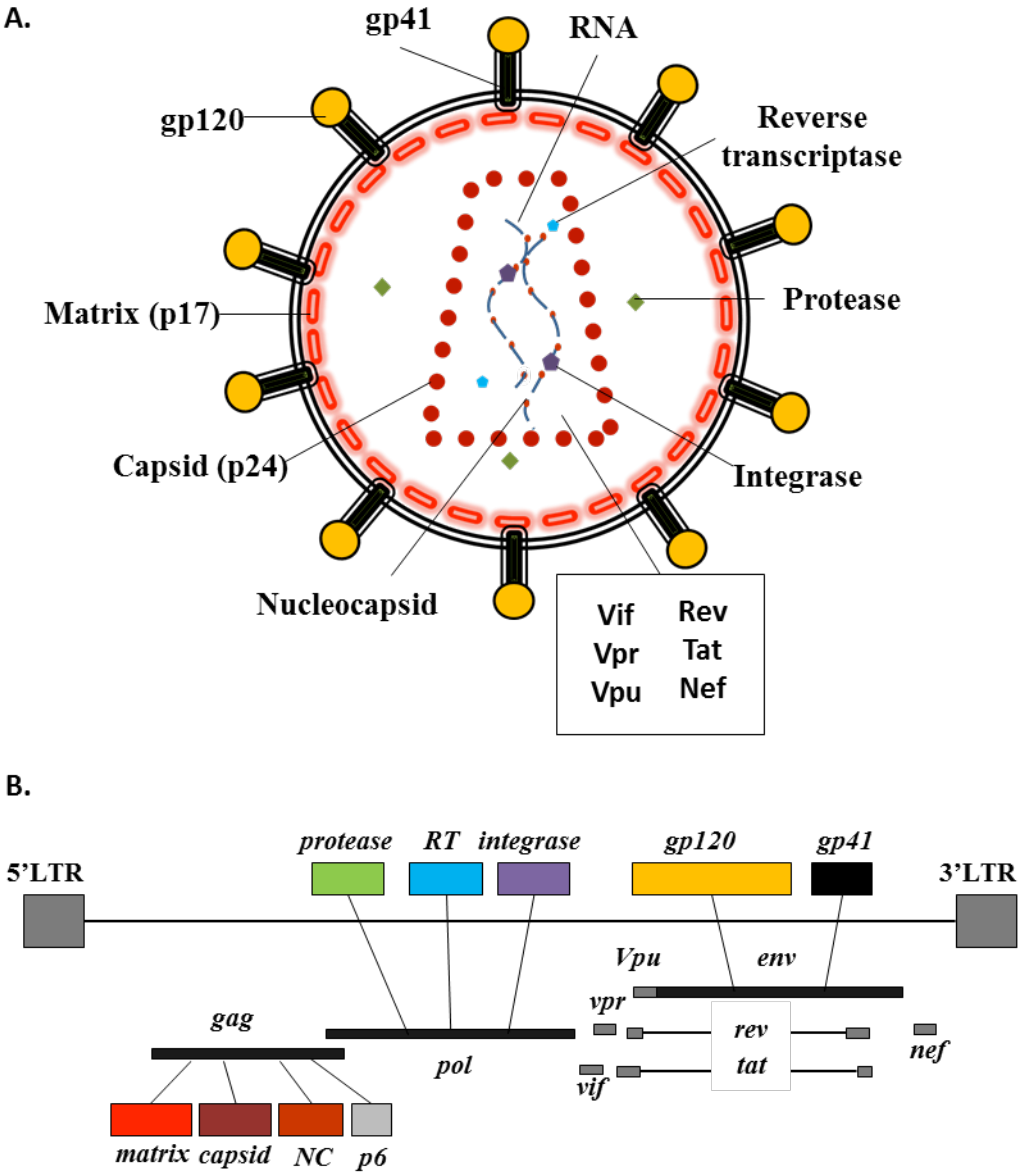

Figure 1.3 HIV-1 structure and genome. A. Cross section of the HIV-1 virion. B. HIV-1 genome is encoding the nine viral proteins and six accessory proteins. Modified from (Frankel & Young 1998)

1.3 HIV-1 Virus Replication Cycle

The HIV replication begins with a series of events that can be categorised in two phases which are known as early and late. The early phase starts when the mature viruses recognise the target cells until the genomic DNA integrates into the chromosome of the host cell. Meanwhile, the late phase starts when integrated proviral genome is regulated until the virus reaches the maturation stage (Figure 1.4).

1.3.1 Early Phase

HIV-1 attachment (step 1) is mediated via the binding of the surface envelope glycoprotein gp120 to the primary host receptor CD4 (Mondor et al. 1998; Wilen et al. 2012) followed by the binding to the chemokine receptors either CCR5 or CXCR4. R5- and X4-tropic HIV-1 isolates utilise CCR5 or CXCR4, respectively, as a co-receptor (Collman et al. 2000; Xiang et al. 2010). Dual-tropic HIV-1 can use either CXCR4 or CCR5 as co-receptor (Yi et al. 1999; Xiang et al. 2011; Yi et al. 2005). Once HIV-1 binds CD4 and co-receptor on the cell surface, its triggering conformational changes in the transmembrane Env glycoprotein gp41 that promote fusion of the viral and cellular membranes (step 2) (Pancera et al. 2010). Fusion process releases the viral capsid core and leads the entry into the cytoplasm of the target cell where HIV-1 undergoes a series of complex viral core uncoating steps. Several findings proposed that the capsid core remains intact post-entry rather than fully disassembled, to an undetermined extent, until the core docks with the nuclear pore to release the pre-integration complex (step 3) (Wilen et al. 2012; Blumenthal et al. 2012; Ambrose & Aiken 2014; Nathalie Arhel 2010). The newly formed pre-integration complex, composed of viral RNA, MA, Vpr, and IN, is transported across the nuclear envelope (step 4) (Popov et al. 1998; Piller et al. 2003; Tsurutani et al. 2007). Viral RNA is then reverse transcribed into double stranded DNA by the viral

enzyme RT (step 5) (Ben-Artzi et al. 1992; Hu & Hughes 2012). Once completely reverse transcribed, the proviral DNA is integrated into the host cellular genome using HIV-1 IN (step 6) (Butler et al. 2001; Craigie & Bushman 2012).

1.3.2 Late Phase

Viral RNA expression is regulated from the HIV-1 long terminal repeat (LTR) promoter and viral and cellular proteins (step 7) (Karn & Stoltzfus 2012). Viral proteins and RNA undergo assembly (step 8) into immature viral particles that bud from the cell surface (step 9) (Sundquist & Krausslich 2012; Freed 2015). The 55-kDa Gag polyprotein is the protein that mediates the assembly and budding of the immature virion. Once released, proteolytic processing via the viral enzyme PR promotes virus maturation (step 10), which is necessary to create an infectious virion (Konvalinka et al. 2015). Since capsid and capsid core involve in both phases of HIV-1 replication, targeting the p24 protein is a viable option for developing alternative therapies against HIV-1/AIDS.

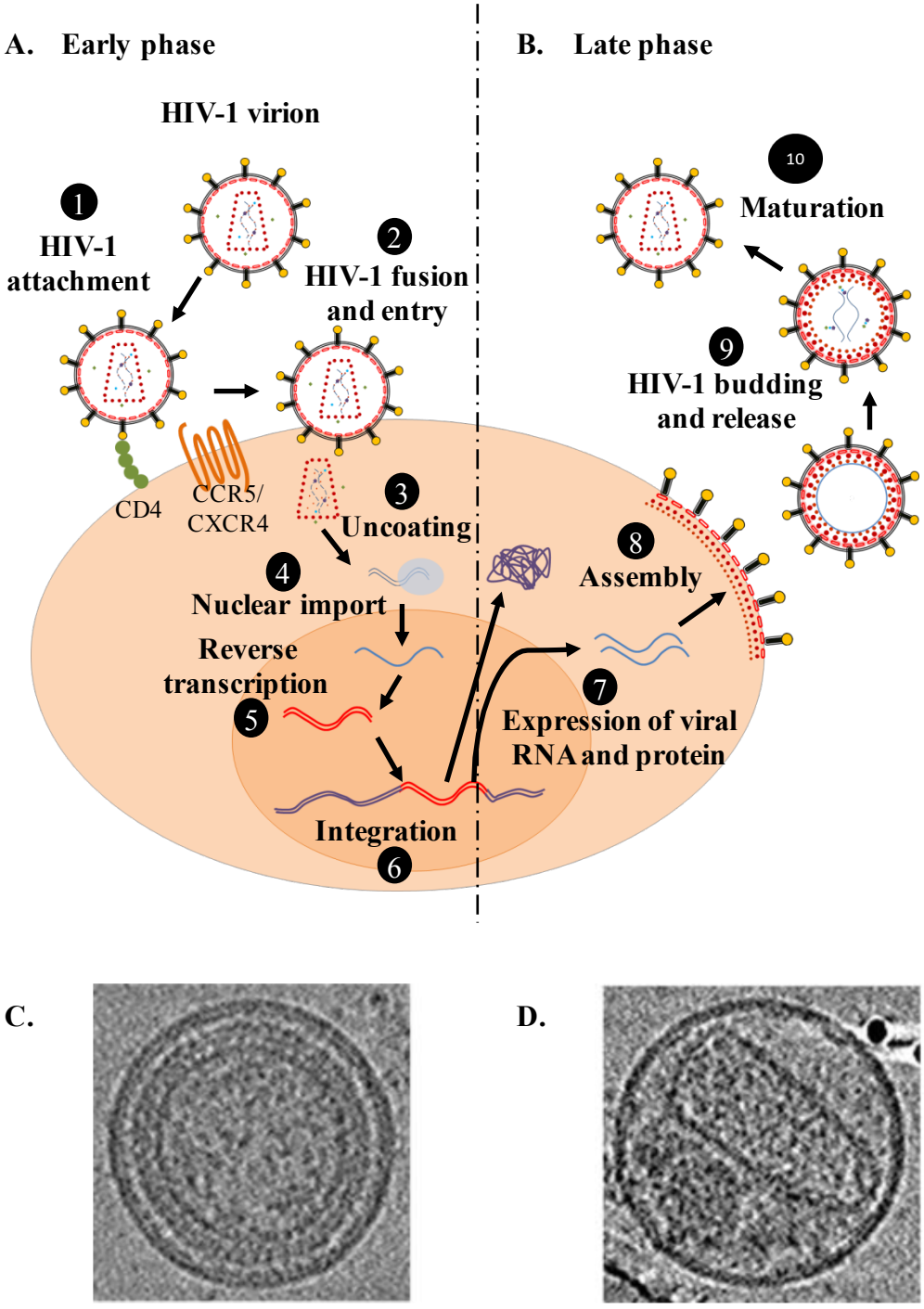

Figure 1.4 HIV-1 life cycle. A. Early phase. B. Late phase. C. Central section from a cryo-EM tomographic reconstruction of an immature HIV-1 virion D. Central section from a tomographic reconstruction of a mature HIV-1 virion. Image C and D obtained from (Sundquist & Krausslich 2012)

1.4 HIV-1 Capsid protein

1.4.1 HIV-1 Capsid p24 Structure

HIV-1 CA p24 comprises of a 231 amino-acid protein encoded by HIV-1 *gag* gene. N-terminal domain (NTD) and C-terminal domain (CTD) are two different domains of HIV-1 CA p24 that connected by a flexible 5-amino acid peptide linker (146-150) (Jiang et al. 2013; Arvidson et al. 2003). The NTD (aa 1-145) contains seven helices (numbered 1–7) and one amino-terminal hairpin that is thought to be involved in the virus infectivity (Morikawa et al. 2000). The C-terminal domain (aa 146–231) comprises of four short helices (numbered 8–11) and a single-turn 310-helix that might contribute to the assembly process during the late stage of virus replication (Ternois et al. 2005).

The NTD is apparently forming the hexamers while the CTD is forming dimeric linkers that connect neighboring hexamers. The presence of NTD-CTD interactions identified from biochemical and genetic experiments suggest that the interaction is essential for hexameric CA p24 formation (Ganser-Pornillos et al. 2007)(Figure 1.5B). Cyclophilin A (CypA) binds CA p24 on residues 85–93 between helices 4 and 5 that is exposed at the surface of the hexameric CA (Gamble et al. 1996; Luban 1996) (Figure 1.5A). The p24 protein of mature virus forms the fullerene conical structure that encapsulates the genomic RNA-nucleocapsid complex (Erdemci-Tandogan et al. 2016). The core with an average size of 100–120 nm in length and 50–60 nm in wide possessed of approximately 250 CA hexameric rings which form lattice structure (Briggs et al. 2003)(Figure 1.5C). The hexameric lattice comprises of about 1,000–1,500 assembled CA p24 proteins with a 10 nm spacing (Briggs et al. 2004)(Figure 1.5D). Both canonical CA p24 structural domains have shown to be involved in HIV replication and infectivity (Fassati 2012).

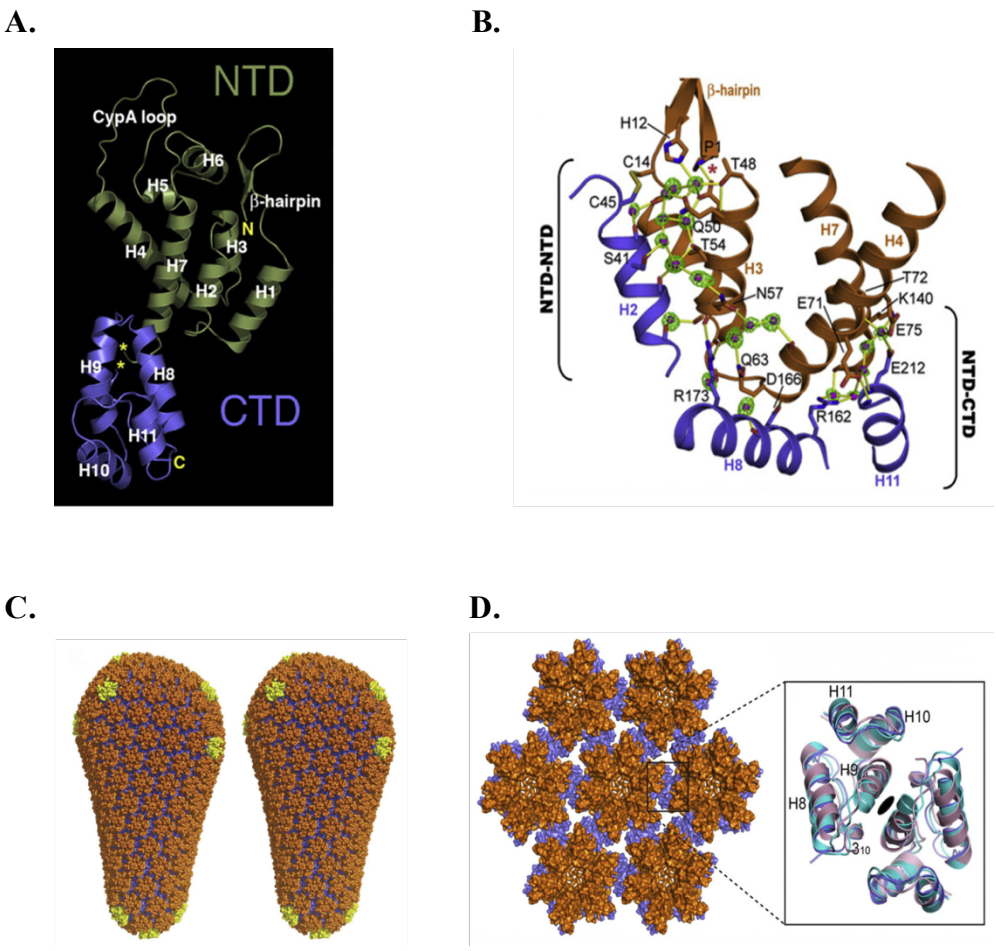

Figure 1.5 Structure of the CA protein, HIV-1 core, and the CA hexameric lattice. (A) Side view of the pseudoatomic monomer model of full-length CA, with the secondary structural elements, labeled, including the cypA-binding loop. Yellow asterisks indicate the carboxy terminus of the NTD and the amino terminus of the CTD (approximately 4 ˚A apart). (B) Polar and water-mediated intermolecular CA contacts within a hexamer. (C) Stereo view of a backbone-only fullerene cone model composed of 1056 CA subunits. (D) Top view of one sheet in the templated CA crystals, which recapitulates the hexameric lattice of authentic capsids at its planar limit. The NTDs are colored orange, and the CTDs are blue. Interactions between neighboring hexamers are mediated only by the CTD. Right-hand box, a top view of the CTD–CTD interface that connects adjacent hexamers. The black oval represents the twofold symmetry axis. Obtained from (Fassati 2012)

1.5 HIV-1 Capsid p24 Inhibitor

1.5.1 Innate immune response factors

1.5.1(a) TRIM5α

The species-specific retroviral restriction factor, the coiled-coil domain TRIM5, and the related protein TRIMCyp, which are a part of the innate immune response factors have been revealed to react with incoming core CA p24 and evade infection of HIV-1 (Matthew et al. 2004; Sayah et al. 2004; Neagu et al. 2009). The binding expedites the uncoating process, hypothetically as a result of TRIM5 assembling on top of the hexameric CA lattice (Stremlau et al. 2006; Owen Pornillos, Barbie K. Ganser-Pornillos 2011). However, the human TRIM5 demonstrated less efficient in inhibition of HIV-1 compared to rhesus monkey TRIM5. Removal of arginine 332 in the human TRIM5 domain and replaced with proline which is residue find in rhesus monkey TRIM5 showed a potent inhibition of both HIV-1 and SIV replication (Li et al. 2006; Javanbakht et al. 2006)

1.5.1(b) Mx2

Human myxovirus resistance 2 (Mx2 or MxB) is a group of the IFN-inducible GTPase superfamily, comprises of the proteins that functioning in cellular processes like cytokinesis, vesicular transport, and blocking of the intracellular pathogens (Liu et al. 2013; Goujon et al. 2013). The protein shows potent activity against primate immunodeficiency virus HIV-1 and does not affect the non-primate viruses such as EIAV, MLV, and FIV. Downregulation of the protein by RNA interference decreased the anti-HIV-1 potency of IFNα (Kane et al. 2013). Mx2 reduces the 2-LTR circles levels and proviruses by ~90%. These findings supporting that Mx2 prevent the HIV-1 nuclear import and DNA integration. Mx2 also increases the amount of pelletable CA, suggesting its mechanism in stabilising the mature capsid.

1.5.1(c) CCDC8

Coiled-coil domain containing protein 8 (CCDC8) is a human membrane-associated protein that showed to inhibit the HIV-1 particle production. The protein mainly acts by binding towards Gag matrix region on the plasma membrane and was detected in virion of HIV-1. The interaction of Gag and CCDC8 led to Gag polyubiquitination and endocytosis degradation (Wei et al. 2015).

1.5.2 Capsid-based assembly inhibitor

Currently, small molecule inhibitors representing either natural compounds or synthetic peptides are being tested for their antiviral activity against HIV-1 CA p24. The most important compound is CAP-1 that had been identified by Tang and group in 2003 (Tang et al. 2003). Study on CAP-1 by different groups resulted in understanding the mechanism of viral inhibition by specifically targeting HIV-1 CA p24 (Kelly et al. 2007). Almost a decade of its finding, the binding pocket of the amino-terminal domain of CA (NTD CA)/CAP-1 has been used in high-throughput screening of the potential anti-HIV compounds. Although the pocket model successfully identified the potential compounds, most of them were inactive in *in vivo* assessment. Lately, the trend in the identification of the potential compound has carried out by a combination of both approaches. This strategy led to identifying the compound such acylhydrazone, benzodiazepine, benzimidazole, PF-3450074, I-XW-053, H22, and pyrrole pyrazolones that act against HIV CA p24.

Meanwhile, the most known synthetic peptide is CAI that had been published by Sticht et al. (2005) three years after CAP-1. The 12-mer α-helical peptide was successfully identified by phage display technology that binds the carboxyl-terminal domain of CA (CTD CA). The binding led to the assembly inhibition of immature- and mature-like capsid particles *in vitro* (Sticht et al. 2005). A few years later,

several studies had shown the improvements of the peptide. Hydrocarbon stapling technique that stabilises the α-helical structure of CAI resulted in the conversion of the peptide into a cell-penetrating peptide (CPP). This modified peptide called as NYAD was found to increase the cell permeability which ultimately led to the inhibition of HIV replication. Series of studies by Zhang et al. (2008) concluded that the peptide bind to CTD of CA at post-entry of HIV replication (Zhang et al. 2008; Zhang et al. 2011; Zhang et al. 2013). In the meantime, a less known peptide that was identified almost at the same time with CAP-1 is *CAC1*. The 20-mer synthetic peptide showed binding against the carboxyl-terminal domain of CA (CTD CA) *in vitro* setting (Garzón et al. 2004). Same as CAI, the *CAC1* was unable to penetrate the cell membrane and needed the cell penetrating peptide to cross the cells. *CAC1M* was shown to inhibit CA polymerisation *in vitro* and the combination of *CAC1M* with established peptides such *CAC1*, and H8 resulted in 90% inhibition of HIV particle production (Bocanegra et al. 2011).

Like most of the small inhibitors and peptide inhibitors are targeting the NTD of CA and CTD of CA respectively, recent finding showed the inhibitors like quinolinone, phenyl, Ebselen, and 2-Arylquinazolines also could bind to CTD of CA (Curreli et al. 2011; Thenin-Houssier et al. 2016; Machara et al. 2016).

Figure 1.6 shows the capsid inhibitors against HIV-1 CA p24 and next sections are a brief introduction of each of the inhibitors.

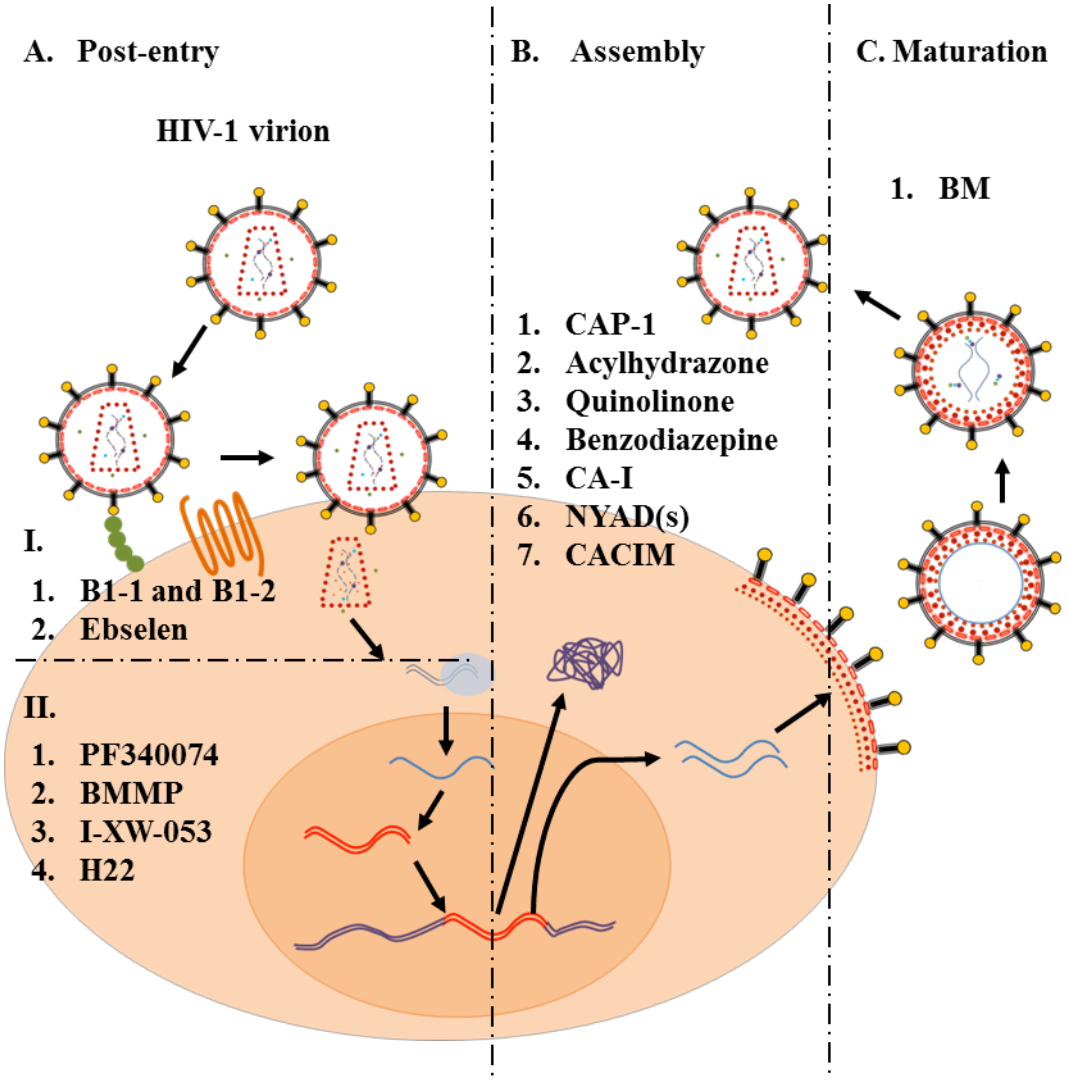

Figure 1.6 HIV-1 capsid p24 inhibitors. A. Capsid-based post-entry inhibitors. B. Capsid-based assembly inhibitors. C. Capsid-based maturation inhibitors.

1.5.2(a) Small molecule

1.5.2(a)(i) CAP-1

As mentioned before, CA p24 is involved in both phases in HIV replication. At the early stage of HIV replication, the HIV-1 Gag polyprotein will assemble and start to form a budding at the membrane of cells. Disruption or destabilising of this process is believed to block the virus production or lead to the production of non-infectious viruses. More than a decade researchers found that there were inhibitors that responsible for this process and some of them classified these inhibitors as capsid-based assembly inhibitor. The discovery of the earliest inhibitors in the series was CAP-1. Summers and the group had found this small molecule by computationally screening based on docking technique (AUTODOCK4.0). The molecule binds to the amino-terminal domain (NTD) of the immature capsid between helices 1 and 2, in residue 59-63 during the early step in HIV replication verified by NMR titration assay (Kelly et al. 2007).

The compound was found to decrease the rate of the p24 assembly during *in vitro* CA p24 assembly and inhibited viral infectivity up to 95% and 98% in latently infected U1 cells and MAGI cells, respectively. Interestingly, the high or low concentration of CAP-1 had no effect on HIV particle production and Gag expression. Morphology analysis of the virions produced from CAP-1-treated cells by an electron microscope showed the presence of greater size heterogeneity virions with most of them do not have a cone-shaped core. These observations suggest that the molecule interacted with HIV CA p24 of immature HIV at the late phase HIV replication that led to Gag assembly defect and subsequently production of less infectious viruses (Tang et al. 2003). Although the binding is well documented, the low affinities, K_d ~800 μM could not consider as a potent drug candidate. Despite

that, the discovery of the binding pocket between CAP-1 and HIV-1 CA p24 has facilitated in identification of other potent compounds as mentioned below

1.5.2(a)(ii) Acylhydrazone (14i and 14l)

The binding pocket between CAP-1 and HIV-1 CA p24 was used by Tien and group to evaluated the binding of the molecules in series of acylhydrazone derivative. The molecules bind to NTD of HIV-1 CA p24 as well as CAP-1 but with two additional grooves of the protein. High-throughput *in vitro* CA p24 assembly showed two molecules (14i and 14l) significantly decreased the rate of p24 assembly in a dose-dependent manner. These two molecules contain an L-phenylalanine side chain that demonstrated to inhibit 50% of SIV virus replication in CEM cells with low effective dose (Tian et al. 2009). Further modification of the side chain by replacing with L-histidine exhibited significant anti-SIV activity, but the EC_{50} were still not as good as previously reported molecules (Jin et al. 2010).

1.5.2(a)(iii) Quinolinone and phenyl (6 and 50)

As mentioned above, CAP-1 and acylhydrazone derivative bind to the hydrophobic cavity of NTD of the HIV-1 CA p24. Small molecule inhibitors identified using the binding pocket between CA-1 and HIV-1 CA p24 could bind to the hydrophobic cavity of carboxyl-terminal domain (CTD) of the HIV-1 CA p24 at residue 165-215. Two potent compounds known as quinolinone and phenyl (6 and 50) inhibited the formation of the mature-like particle but not for the immature-like particle suggesting that the compound target CA instead of CANC. Unlike CAP-1, which do not affect the virus particle production, both quinolinone derivative (6) and phenyl derivative (50) decreased the HIV-1 particle release of different HIV-1 subtype B viruses and RT- or protease-resistant strains from compound-treated -MT-2 and -PBMC cells. Infectivity assay using normalised p24-virus infected MT-2 cells showed about 40%