SCREENING FOR SMALL MOLECULE INHIBITORS OF HIV-1 VPR: DISCOVERY OF VIPIRININ AND ITS USE AS A BIOPROBE

EUGENE ONG BOON BENG

UNIVERSITI SAINS MALAYSIA

2011

SCREENING FOR SMALL MOLECULE INHIBITORS OF HIV-1 VPR: DISCOVERY OF VIPIRININ AND ITS USE AS A BIOPROBE

by

EUGENE ONG BOON BENG

Thesis submitted in fulfillment of the requirements
for the degree of
Doctor of Philosophy

May 2011

ACKNOWLEDGMENTS

I wish to thank everyone that has made this thesis possible.

*Sensei*s,

Professor Nazalan Najimudin - for the opportunity, guidance, and passion for science.

Professor Hiroyuki Osada - for the great insights into the world of chemical biology.

Professor Nobumoto Watanabe - for the kind mentoring, patience and guidance.

Assoc Professor Mohd Razip Samian - for the sharing of ideas and good advice.

Assoc Professor Sudesh Kumar - for the opportunity, discussions and advice.

I will always be indebted to all of you for the wisdom that you have shared with me.

Friends and colleagues,

Kawan-kawan Lab 414 - for their friendship and help.

Antibiotic Lab Members - for their friendship, assistance and technical help.

Universiti Sains Malaysia - for use of facilities and candidature matters.

RIKEN - for the scholarship and opportunity to carry out this work as an Asia Program Associate.

Cindy and Oliver, and family - for their love, encouragement and support.

Last but not least,

this thesis is dedicated to my parents, Mr and Mrs Ong Soon Aun.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

α1	alpha-helix 1
α2	alpha-helix 2
α3	alpha-helix 3
aa	amino acids
AIDS	Acquired Immunodeficiency Syndrome
DCAF1	DDB1- and Cul4A-associated factor 1
DDB1	damaged-DNA specific binding protein 1
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
HAART	Highly Active Antiretroviral Therapy
HIV-1	Human Immunodeficiency Virus Type 1
HTS	High-Throughput Screening
LTR	Long terminal repeat
ml	milliliter
µl	microliter
µM	micromolar
NMR	Nuclear magnetic resonance
NPD	Natural Products Derivative
NPDepo	Natural Products Depository
OD	Optical Density
PCR	Polymerase Chain Reaction
RNA	Ribonucleic acid
SAR	Structure-Activity Relationship
SDS	Sodium Dodecyl Sulfate
UV	Ultraviolet
Vpr	Viral protein R
VprBP	Vpr Binding Protein

SARINGAN MOLEKUL KECIL PERENCAT VPR HIV-1: PENEMUAN VIPIRININ DAN PENGGUNAANNYA SEBAGAI BIOPROB

ABSTRAK

Protein virus R (Vpr) virus imunodefisiensi manusia jenis 1 (HIV-1) adalah protein aksesori yang telah diketahui mempunyai pelbagai peranan dalam serangan HIV-1. Sini, satu sistem penyaringan daya pemprosesan tinggi (HTS) telah dimajukan untuk menyaring perpustakaan kimia dalam RIKEN Natural Products Depository (NPDepo). Melalui HTS, satu sebatian berasaskan kumarin 3-fenil, iaitu NPD4456, telah dikenalpasti sebagai inhibitor (perencat) aktiviti hentian kitar sel oleh Vpr. Unit bioaktif terkecil perencat tersebut dapat ditentukan melalui kajian hubungan struktur-aktiviti dan sebatian terbitan yang lebih manjur telah dihasilkan. Perencat-perencat Vpr tersebut digunakan sebagai bioprob untuk menyelidik satu panel mutan titik tunggal alanina Vpr, dan kawasan hidrofobik sekitar asid amino E25 dan Q65 pada protein Vpr adalah berkemungkinan besar terlibat dalam pengikatan perencat. Model dok meramalkan bahawa perencat tersebut mampu mengikat kepada poket hidrofobik di kawasan yang sama. Dengan menjalankan analisis tarik-bawah menggunakan manik-manik berpautkan perencat, Vpr didapati boleh mengikat kepada perencat secara terus, dan pengikatan protein ini adalah bercorak kompetitif. Perencat tersebut juga didapati boleh menghalang jangkitan virus bergantung-Vpr terhadap sel makrofaj manusia. Keseluruhannya, keputusan yang diperolehi telah menggariskan pendekatan yang boleh digunapakai untuk menyelidik kawasan sasaran molekul pada protein Vpr dengan menggunakan perencat molekul kecil.

SCREENING FOR SMALL MOLECULE INHIBITORS OF HIV-1 VPR: DISCOVERY OF VIPIRININ AND ITS USE AS A BIOPROBE

ABSTRACT

The human immunodeficiency virus 1 (HIV-1) viral protein R (Vpr) is an accessory protein that has been known to have multiple roles in HIV-1 pathogenesis. Here, a high-throughput screening (HTS) system was developed to screen chemical libraries in RIKEN Natural Products Depository (NPDepo). Through HTS, NPD4456, a 3-phenyl coumarin based compound, was identified to be an inhibitor of the cell cycle arrest activity of Vpr in yeast. The minimal pharmacophore of the inhibitor was determined through a structure-activity relationship study and more potent derivatives were produced. The Vpr inhibitors were used as bioprobes to explore a panel of single point alanine mutants of Vpr and the hydrophobic region about residues E25 and Q65 of the Vpr protein was found to be potentially involved in the binding of the inhibitor. Docking models predicted that the inhibitors were able to bind to a hydrophobic pocket in the same region. Direct binding of Vpr to the inhibitor was detected in a pull-down assay with compound-immobilized beads and the protein was observed to bind in a competitive manner. The inhibitors were also found be able to inhibit Vpr-dependent viral infection of human macrophages. Taken together, these findings delineated a convenient approach to explore a probable drug targeting site on Vpr using small molecule inhibitors.

CHAPTER ONE

Introduction

1.1 General Introduction

The advent of chemical biology in recent years was due in part by the emergence of high-throughput screening (HTS) technologies that became available to fundamental research laboratories. Due to the high cost of equipments, manpower and facilities needed to perform HTS, it was previously only carried out by big pharmaceutical companies in their screen for novel drugs, but however, it is now also routinely used in chemical biology studies.

Chemical biology is a relatively new field that arose from the interface of chemistry and biology. It is thought that chemical biology emerged over decades as a complex hybridization of bioorganic chemistry, biochemistry, cell biology and pharmacology (Bucci et al., 2010). In the past decade, the chemical biology approach to study biological questions through the use of chemical probes, for example biologically active natural products as tools for examining protein function and studying cell biology, had been increasingly adopted.

The reductionist approach normally taken for the understanding of biological processes is based on ways to perturb a certain biological process and observe the effects of the perturbation. Established molecular biology methods use genetic knockouts, RNA interference, and site-directed mutagenesis to understand the roles of genes and gene products by perturbing protein expression at the genetic or transcriptional level. Although these methods have increased our knowledge in molecular, cellular, and developmental biology, there are still many unanswered questions. Through the use of biologically active compounds in the chemical biology approach, we now have additional molecular toolboxes to address these questions.

Small molecules have a significant advantage over classical genetic techniques in that they can be used for specific and temporal observations. For example, a small molecule that targets a specific protein can be used to 'knock out' a gene or inhibit its protein only at specific time points during the cell cycle or during an organism's developmental process. Thereby, these small molecules can be used in a temporal manner to induce or inhibit a specific biological activity, thus providing a method to investigate specific events within a specified temporal window. In this way, chemical biology provides a means to answer biological questions that may not be possible to achieve with standard genetic methods.

In contrast to genetic approaches that had been widely used, small molecule inhibitors are also useful because of the following attributes – small molecules usually acts within minutes or even seconds upon administration, allowing quick inhibition or activation of protein function, they are often reversible, thus small molecules allow a high degree of time-based control (Tolliday et al., 2006). These precise perturbations in specific periods are especially useful in studying dynamic biological processes. Furthermore, the amount of small molecule inhibitors used can be precisely controlled to vary the levels of inhibition. Additionally, small molecules, especially natural products, are usually applicable in multiple biological systems, allowing us to dissect biological mechanisms in different organisms, different cell types, and *in vitro* systems (Osada, 2000).

Natural products are produced by various microorganisms and plants. It is thought that these naturally produced small molecules with potent biological activities are the result of an evolved survival mechanism to counter external stress or threats, as observed in the textbook example of penicillin. These small molecules have various biological activities and have been used by humans historically as herbal remedies or

drugs (Remington, 1995). Initially, these natural products were used as a source or lead for a new drug candidate by pharmaceutical companies but recently researchers have found that natural products are useful for perturbing cellular processes (Osada, 2009).

Furthermore, natural products are useful to study biological systems by virtue of their being produced naturally in a living system, making them mostly cell-permeable and they usually already have specific biological targets. Combined with synthetic chemistry to produce derivatives or analogs, structure-activity relationships of natural products provide further insights into their mechanism of action and are good starting points for the development of new synthetic biological probes. Many biologically active natural products are isolated each year but not all are effective as biological tools. Only those that act as ligands for enzymes or target proteins are suitable for potential use as chemical biology probes. Such probes are also called bioprobes (Osada, 2000).

Bioprobes have been successful for characterizing and dissecting many of the cellular and signaling pathways known today such as the mammalian target of rapamycin (Hardwick et al., 1999) and in solving biological mysteries such as the target of the drug thalidomide (Ito et al., 2010) that caused severe birth defects of newborns in the 1960s (Melchert and List, 2007). However, new small molecules are needed to function as probes to further dissect the known pathways, to discover novel pathways and interactions, to validate potential drug targets, and to function as drugs to treat unmet medical needs, including more targeted therapies (Osada, 2009).

This thesis adopts the chemical biology approach to study the HIV-1 Vpr (Viral protein R), an accessory protein which has been reported to play many roles in HIV-1 pathogenesis. The approach that was taken, starting from a HTS for Vpr inhibitors, structure-activity relationship study and inhibitor-protein interaction assays for analysis is outlined in Figure 1.1.

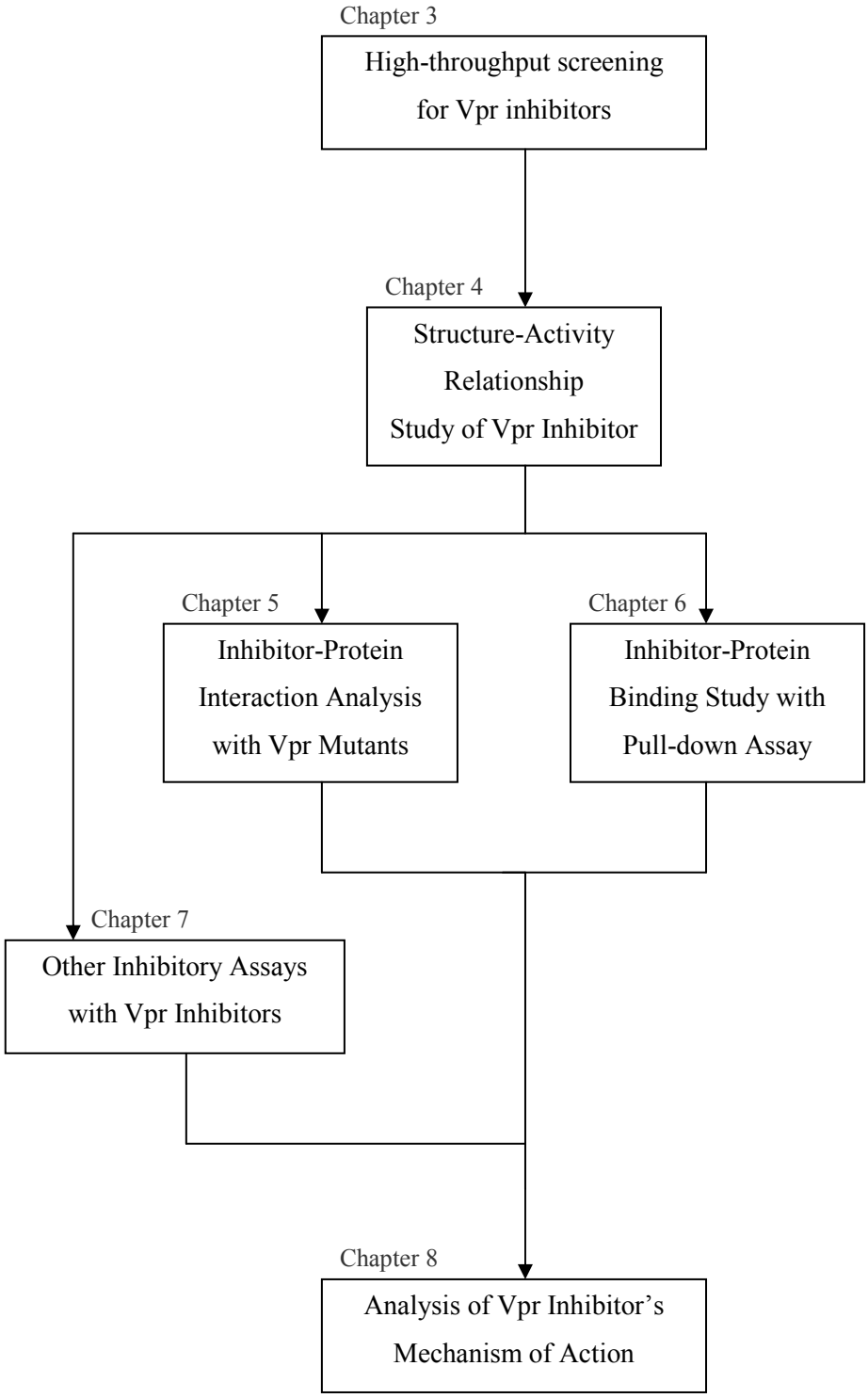


Figure 1.1 Outline and flow of thesis.

Experimental approach taken to screen for Vpr inhibitors and study of its mechanism of action.

1.2 Hypothesis and Aims of Thesis

The HIV-1 genome encodes for structural proteins, regulatory proteins and accessory proteins, yet there are still no approved anti-HIV drugs that target its accessory proteins. Therefore, increasing the targets of HIV therapy by additionally targeting accessory proteins may expand treatment options, resulting in improved therapy and better management of the disease.

Previously, a Vpr inhibitor was isolated by the Antibiotics Laboratory (where the research for this thesis was carried out) by screening fungal metabolites using a paper disc method (Watanabe et al., 2006). This thesis expands upon the previous work by developing a pragmatic screening system to interrogate a library of natural compounds for Vpr inhibitor(s) and to use the inhibitor as a bioprobe to study Vpr as well as the inhibitor's mechanism of action.

Specific aims were:

1. To establish a HTS system for Vpr inhibitors by converting the previous yeast-based paper disc method into a 96-well plate liquid assay using growth recovery as an index (Chapter 3).
2. To study the screening hit inhibitor's structure-activity relationship by designing and synthesizing novel derivatives (Chapter 4).
3. To study inhibitor-protein interaction by using Vpr mutants in an inhibitory assay and by pull-down assays with compound-immobilized beads (Chapters 5 and 6).
4. To analyze the inhibitor's effect on other Vpr-based assays (Chapter 7).

CHAPTER TWO

Literature Review

2.1 AIDS and Antiviral Therapy Overview

The human immunodeficiency virus (HIV) is the causative agent of the acquired immunodeficiency syndrome (AIDS) in humans (Levy et al., 1985). Since the first report of AIDS in the United States in 1981, AIDS had grown to become a worldwide pandemic. Although, the percentage of people living with HIV has stabilized since 2000, the overall number of people living with HIV continues to increase steadily as new infections occur. Despite the decline in annual infections, 2.7 million were infected and 2 million died of HIV-associated diseases in 2007, bringing the total to an estimated 33 million HIV+ individuals in that year (UNAIDS, 2010).

HIV establishes a persistent infection in humans that eventually results in defective cellular immunity. Although, a number of effective drugs have been approved by the US Food and Drug Administration (FDA) to reduce viral replication, their antiviral effect is only temporary due to HIV's ability to quickly develop resistance to these antiviral agents (Johnson et al., 2008, Shafer and Schapiro, 2008). To overcome this problem, the drugs are combined and used in a highly active antiretroviral therapy (HAART), which was proven to be effective at both reducing virus load and suppressing the emergence of resistance in patients. In countries where HAART is easily accessible for HIV positive (HIV+) persons, the number of HIV related deaths has dropped (Shafer and Schapiro, 2008).

Still, the success of HAART is limited as ultimately viral resistance appears in quickly in some individuals receiving treatment (Bennett et al., 2008). Viral resistance in HIV during the course of infection combined with a high viral mutation rate is one of the reasons there is still no cure for AIDS. Incomplete viral suppression is believed to

be the main cause for drug resistant variants to emerge. Even when viral plasma levels have been reduced to low levels (<50 copies/ml) by HAART, HIV replication continues in the background (Kinter et al., 2003, Zhu et al., 2002).

Although HAART is able to prolong a HIV+ person's lifespan and improve their quality of life, there are also negative implications such as toxicity and long-term side effects. It is believed that 50 % of those receiving HAART for more than a year will suffer from mitochondrial dysfunction, fat maldistribution, hyperlipidemia, insulin resistance, hypersensitivity, and, hematologic and bone changes because of the drugs (Sutinen, 2009).

Considering all these factors, until there is a cure for AIDS, we need to constantly search for new antiviral agents to not only repress viral resistance but also to suppress viral replication even further while reducing the side effects caused by the drugs.

2.2 Human Immunodeficiency Virus type 1 (HIV-1)

The origins of HIV was traced to zoonotic (cross species) transmissions of Simian Immunodeficiency Virus (SIV) to humans from at least two different kinds of non-human primates, HIV-1, from chimpanzees, and HIV-2, from sooty mangabeys (Hahn et al., 2000, Gao et al., 1999). HIV-1 and HIV-2 are distinguished on the basis of their genome organizations and phylogenetic relationships with other primate lentiviruses (Hahn et al., 2000). HIV-1 accounts for most HIV infections throughout the world, whereas HIV-2 is confined to West Africa; HIV-1 is relatively more pathogenic than HIV-2 (Reeves and Doms, 2002).

HIV-1 is a retrovirus, containing two copies of RNA as genetic material. Upon virus infection of cells, viral genome generated through reverse transcription gets inserted into host DNA to form provirus. The HIV genome encodes essential structural

viral proteins, the *gag*, *pol* and *env* genes, and contains six supplementary open reading frames termed *tat*, *rev*, *nef*, *vif*, *vpr*, *vpx* and *vpu* (Greene, 1991). The *tat* and *rev* genes encode transactivating regulatory proteins necessary for virus replication, and the other genes – *nef*, *vif*, *vpr*, *vpx* and *vpu* encode small accessory proteins (Emerman and Malim, 1998, Gelderblom et al., 1989).

Although these accessory proteins have been found to be dispensable for virus growth in *in vitro* systems, these proteins are essential for viral replication and pathogenesis *in vivo* (Connor et al., 1995, Planelles et al., 1995). The two *vpr*- and *vpx*-related genes are found only in members of the HIV-2, SIVsm, SIVmac group, whereas primate lentiviruses from other lineages (HIV-1, SIVcpz, SIVagm, SIVmnd and SIVsyk) contain a single *vpr* gene (Figure 2.1) (Emerman and Malim, 1998, Tristem et al., 1998). The organization of the HIV genome and the summary of its gene functions are shown in Table 2.1 and Figure 2.2.

The HIV infection and replication cycle can be categorized into six main steps. These steps are – fusion of the viral envelope and the cell membrane (1), uncoating of the viral nucleocapsid and reverse transcription of the viral RNA to proviral DNA (2), integration of the proviral DNA into the cellular genome (3), transcription of the proviral DNA to RNA (4), budding, virion assembly and release (5), and the final step of maturation (6), when proteases cleave polyproteins into individual functional HIV proteins and enzymes that assemble to form a mature HIV virion (Greene, 1991).

The individual steps in the viral replication cycle are currently targeted by various anti-HIV agents; although considerable progress has been made in developing successful anti- HIV-1 drugs, HIV-1 has developed resistance to all 21 of the currently approved drugs to treat it (Boyer et al., 2006). Therefore, we must seek new targets in

the HIV-1 infection pathways for pharmaceutical intervention to allow more comprehensive management of the disease.

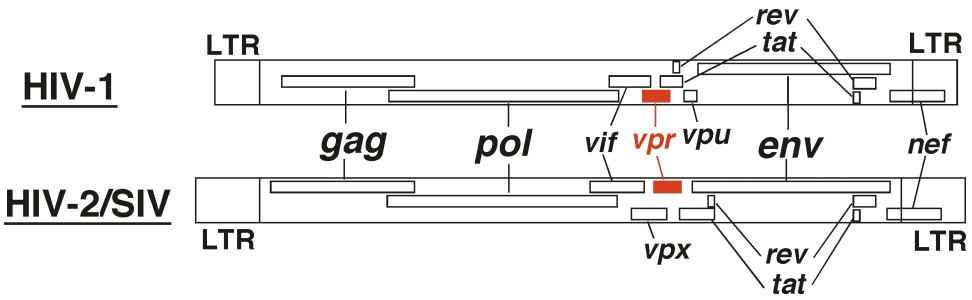
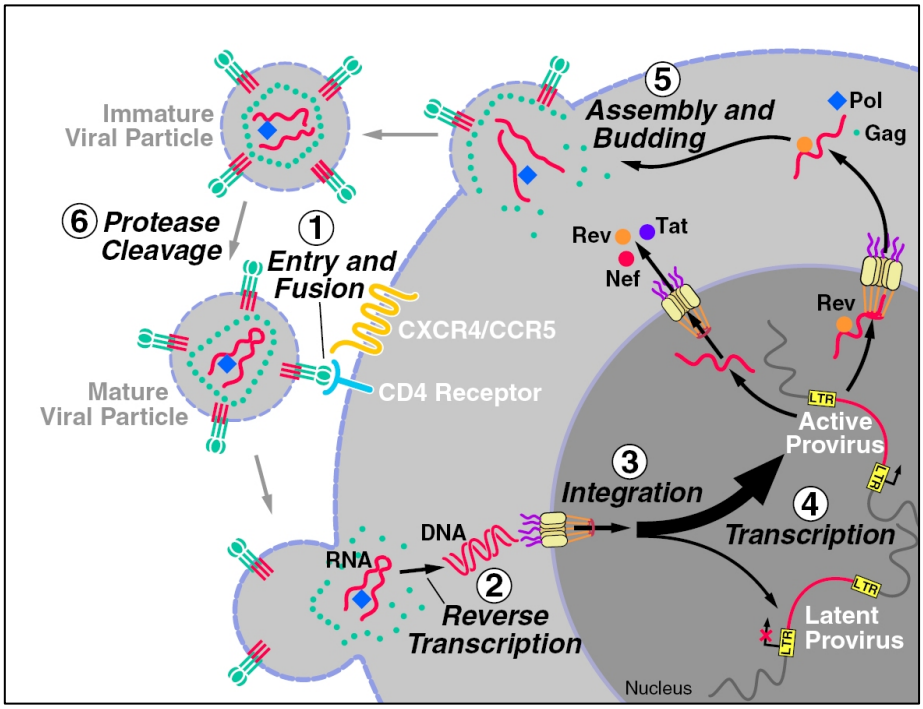


Figure 2.1 The *vpr* gene is highly conserved in both HIV-1 and HIV-2 and SIV. HIV-2 and SIV contains an additional *vpr*-related gene called *vpx* (Tristem et al., 1998). Illustration by Nobumoto Watanabe of Antibiotics Laboratory.

Table 2.1 Genes of the HIV genome (Emerman and Malim, 1998).

Structural Genes	
gag	Encodes virus core (capsid, CA; matrix, MA ; nucleocapsid, NC). CA binds cyclophilins, MA contains a nuclear import signal (NLS) and NC binds RNA.
pol	Encodes viral enzymes (reverse transcriptase, RT; RNase H; protease, PR; integrase, IN). RT and PR are major targets of antivirals. RT is responsible for viral mutation and IN contains an NLS.
env	Encodes virus envelope surface (gp120/SU) and transmembrane (gp41/TM) glycoproteins. gp120/SU binds to CD4 receptor and gp41/TM fuses with cellular membrane.
Transactivation Genes – Genes essential for replication *in vitro*	
tat	Trans-activator of viral transcription that binds to tat-responsive RNA element (TAR). Tat upregulates viral transcription.
rev	Regulator of viral expression that binds to rev-responsive RNA element (RRE). Rev regulates viral RNA transport and splicing.
Accessory Genes – Genes not essential for replication *in vitro*	
nef	Negative factor (a misnomer). Nef is essential for viral disease induction *in vivo*. It downregulates CD4 (cluster of differentiation) and MHC-I (major histocompatibility complex I). Nef also binds to cellular kinases.
vif	Virion infectivity factor. Vif facilitates virion maturation.
vpr	Viral protein regulatory. Vpr causes cell cycle arrest and contains an NLS.
vpu	Viral protein unknown. Vpu downregulates CD4, MHC-I promotes virus release. Not found in HIV-2.
vpx	Viral protein X. Vpx allows infection of macrophages and viral dissemination. Only found in HIV-2.

A



B

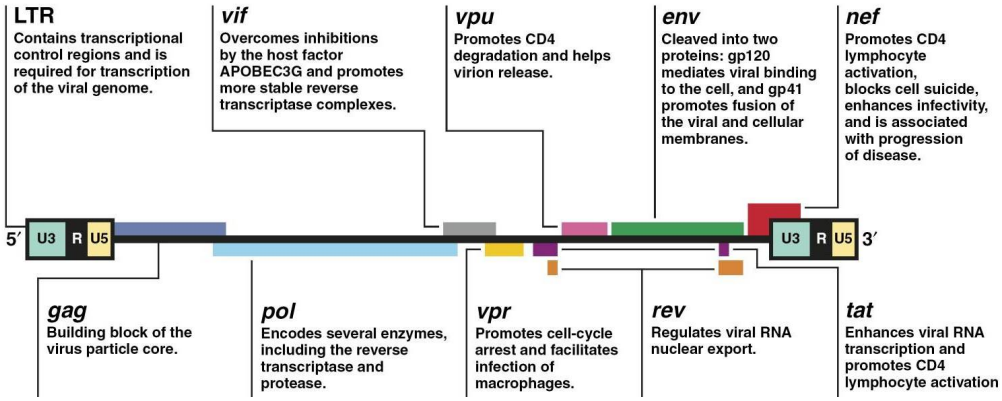


Figure 2.2 The life cycle of HIV and the function of its genes.

A, Steps in the HIV replication cycle. B, Layout of the nine genes along the ~9 kilobase genome of the HIV provirus and a brief description of its functions (Greene, 1991). Illustrations are from the Gladstone Institute of Virology and Immunology (http://www.gladstone.ucsf.edu/wp/2009/10/antiviraltherapies/, accessed on 15 January 2010).

2.3 HIV-1 Vpr

Vpr was originally termed 'viral protein, regulatory' because, when its gene was mutagenized, the mutant virus replicated with a slower kinetic (Hattori et al., 1990). HIV-1 Vpr is a small, 96-amino acid (14 kDa) protein that appears separately at two different times of the virus life cycle – it is packaged into the virus particle via a direct interaction with the Gag precursor, and therefore is present in the cytoplasm of newly infected cells and Vpr is also expressed de novo by the provirus, from a late mRNA (reviewed in Le Rouzic and Benichou, 2005).

The HIV-1 Vpr is an accessory protein that is implicated in viral pathogenesis. The Vpr gene is highly conserved in all HIV and SIV members of the lentivirus family (Hirsch et al., 1989, Huet et al., 1990, Tristem et al., 1998), denoting the importance of this gene in viral assault. In spite of its small size and simple structure, the Vpr protein plays numerous roles in the virus life cycle and has many effects on infected host cells (Figure 2.3). Studies have shown that Vpr is a multifunctional protein with many functions ascribed to it – induction of cell cycle arrest in the G2 phase, transactivation, nuclear import of preintegration complexes in macrophages and other non-dividing cells, induction of apoptosis, and enhancement of the fidelity of reverse transcription (reviewed in(Andersen and Planelles, 2005, Le Rouzic and Benichou, 2005). Vpr had also been detected in the sera and cerebrospinal fluids of AIDS patients (Xiao et al., 2008), and a recent study showed that extracellular Vpr can reactivate viral production from latently infected cells (Hoshino et al., 2010).

The Vpr proteins can form functional dimers and oligomers (Bourbigot et al., 2005, Fritz et al., 2008), possibly explaining its many biological abilities and the many protein-protein interactions between Vpr and various host cell factors (Kino et al., 2002, Le Rouzic et al., 2008, Zhao and Elder, 2000).

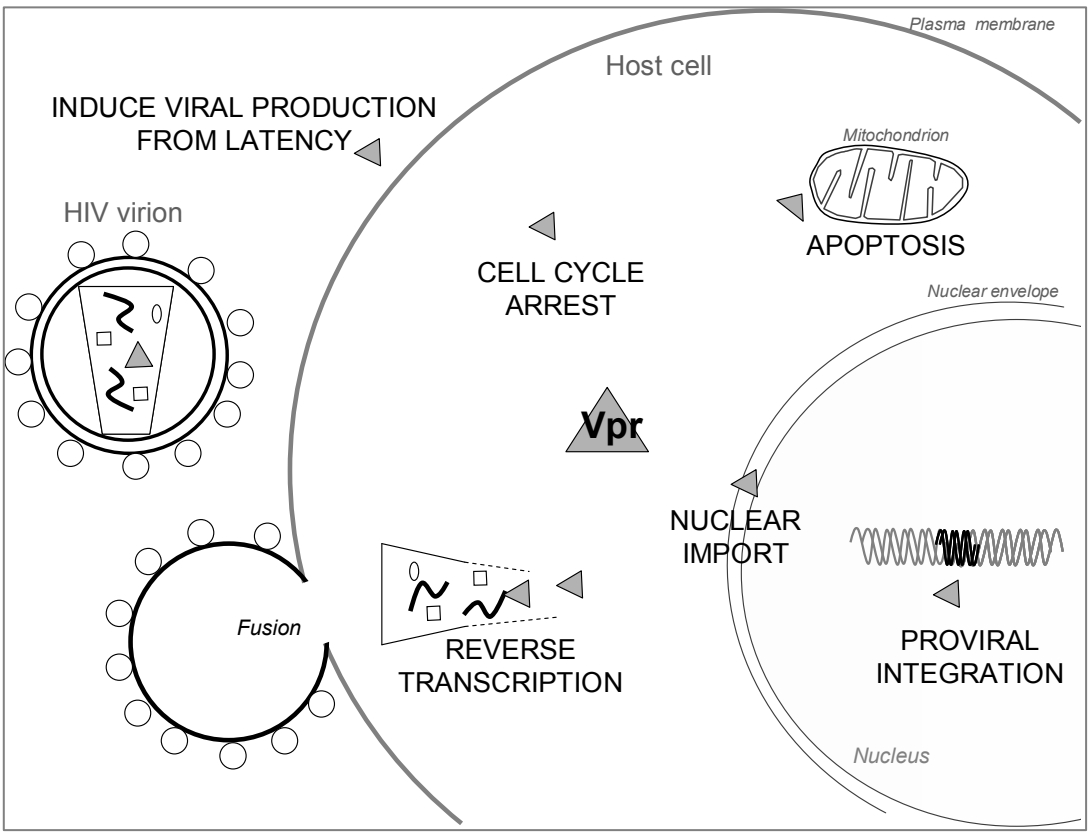


Figure 2.3 The various roles of Vpr in HIV pathogenesis.

Vpr (indicated by grey triangles) are one of the viral proteins that are packaged in the HIV virion and had been found to affect the host cells' cellular pathways in ways which are independent of one another such as G2 cell cycle arrest, apoptosis, nuclear import, retrotranscription, the integration of the provirus in the nucleus and transactivation of the HIV long terminal repeats (adapted from Le Rouzic and Benichou, 2005).

2.4 Vpr and HIV Pathogenesis

As previously mentioned, Vpr is packaged into virions and is present during the early steps of infection of target cells; it is also expressed at a late stage of the virus life cycle and is recruited into the core of the infectious virus particle budding at the cell surface (Le Rouzic and Benichou, 2005). In the core of the virion particle Vpr could interact with nucleic acids and/or matrix protein. Based on the Vpr–Gag incorporation ratio of 1:7, it is estimated that there are 275 molecules of Vpr per virion (Muller et al., 2000).

Upon cell entry, the viral core is released into the cytoplasm of the target cell and initiates the reverse transcription of the viral RNA (Figure 2.3). A large nucleoprotein complex called the reverse transcription complex (RTC) is formed and is made up of the two copies of viral RNA together with the viral proteins, RT, IN, NCp7, Vpr and a few molecules of the matrix protein (Farnet and Haseltine, 1991, Fassati and Goff, 2001, Miller et al., 1997, Nermut and Fassati, 2003). Other recent studies confirmed that Vpr co-localizes with viral nucleic acids and IN within purified HIV-1 RTCs (Fassati et al., 2003, McDonald et al., 2003, Nermut and Fassati, 2003).

After virus entry into the cell, the viral capsid is rapidly uncoated and the reverse transcription of the genomic HIV-1 RNA leading to the full length double strand DNA is completed. This viral DNA associates with viral and host cell proteins into the so-called pre-integration complex (PIC). Vpr has been reported to assist the transport of the viral DNA into the nucleus of nondividing cells (Connor et al., 1995, Gallay et al., 1996, Heinzinger et al., 1994), by promoting interactions with the cellular machinery regulating the nucleo-cytoplasmic shuttling (Fouchier et al., 1998, Le Rouzic et al., 2002, Popov et al., 1998).

Vpr shows karyophilic (attracted to the nucleus) properties and localizes in the nucleus, but a significant fraction is bound to the nuclear envelope (NE) as was visualized by nuclear rim staining in fluorescence microscopy experiments (Kamata and Aida, 2000, Vodicka et al., 1998, Waldhuber et al., 2003). The role of Vpr at the NE is still not clear but it had been proposed that virion Vpr would dock to the nuclear pore complex (NPC) before nuclear translocation occurs (de Noronha et al., 2001, Dupont et al., 1999, Haffar et al., 2000, Reil et al., 1998).

The most characteristic function of Vpr is to block the cell cycle at the G2 phase (reviewed in Section 2.6), followed by apoptosis in both HIV-1-infected cells and Vpr-expressing cells (Jowett et al., 1995). The activity of Vpr-mediated arrest of the cell cycle is conserved in HIV-2 and SIV, suggesting an important role for Vpr in the life cycle of such viruses (Planelles et al., 1996). Indeed, arrest in the G2 phase of the cell cycle may create an intracellular environment that enhances the transcription of the HIV-1 promoter, the long terminal repeat (LTR) (Goh et al., 1998). In addition, the uncoupling of the centrosome duplication cycle from the cell division cycle due to long G2 arrest induces abnormal M-phase spindle formation and subsequent apoptosis (Watanabe et al., 2000).

Vpr has been reported to induce apoptosis following G2 arrest in numerous cell types including primary T-cells, T-cell lines, and various other human transformed cell lines (Ayyavoo et al., 1997, Poon et al., 1998, Poon et al., 1997, Shostak et al., 1999, Stewart et al., 1997). However, the mechanism by which Vpr induces apoptosis is not completely understood because HIV induces apoptosis in multiple pathways, with Vpr being one of them. Eventually, apoptosis leads to the depletion of CD4+T cells in the body. Vpr-induced apoptosis was reported to be caused by permeabilization of the mitochondrial membrane and subsequent dissipation of the mitochondrial

transmembrane potential (Roshal et al., 2003). Other reports suggest that this effect is generated by either a direct action of Vpr on mitochondria (Zennou et al., 2000), or by the physical interaction of Vpr with the adenine nucleotide translocator (ANT) (Zhao et al., 1994b), which triggers caspase-mediated apoptosis (Zhou and Elledge, 2000).

2.5 Structure of Vpr

Visual analysis of the structure of Vpr protein derived from the reported NMR structure of Vpr (Protein Data Bank ID 1M8L) (Morellet et al., 2003) using PyMOL molecular visualizations (DeLano, 2002) software reveals that Vpr consists of a flexible N-terminal region, followed by three α-helices connected by loops and ends with flexible C-terminus (Figure 2.4).

The three helices are arranged in a specific manner and are classified as an orthogonal bundle in the CATH Protein Structure Classification system (http://www.cathdb.info/, accessed on 20 September 2009). The flexible N-terminal region is made up of 16 residues, followed by the first alpha helix (α1) comprising 17 residues (aa 17-33). α1 is connected to the second helix (α2) by a short loop of four residues (aa 34-37). α2 which is made up of 13 residues (aa 38-50) is followed by a second inter-helical loop of six residues that connects to the third helix (α3). The α3, comprising 23 residues (aa 55-77), is followed by the flexible C-terminal domain (aa 78-96). Helices α1, α2 and α3 are amphipathic, whereby one side of the helix contains hydrophobic residues and another side contains hydrophilic ones.

2.5.1 Amphipathic Property of the α-Helices

The amphipathic properties of Vpr's three α-helices can be analyzed by categorizing the amino acid side chains into two categories – hydrophobic and

hydrophilic. This will determine their preference for aqueous or nonpolar environments. The properties of the side chains in the respective helices are summarized in Table 2.2 and shown in Figure 2.4B.

Table 2.2 Hydrophobic and hydrophilic properties of the amino acid side chains of Vpr in helices α1, α2 and α3.

Helices	Properties of side chains	Amino acids
α1	Hydrophobic	W18, T19, L20, L22, L23 L26, A30 and V31
	Hydrophilic	D17, E21, E24, E25, K27, N28, E29 and R32
α2	Hydrophobic	W38, L39, L42, and I46
	Hydrophilic	S41, Q44 and E48
α3	Hydrophobic	V57, L60, L63, L64, L67, L68 and I74
	Hydrophilic	E58, R62, E65, E66, R73, R77 and R80

The hydrophobic side chains of the residues on the three helices form the hydrophobic core of the protein, while the hydrophilic side chains are exposed externally (Figure 2.4). Further examination also reveals that α2 and α3 have a few externally exposed hydrophobic residues thus creating a hydrophobic facet on the surface of Vpr. It is thought that this hydrophobic facet between α2 and α3 is the site for protein-protein interaction that leads to dimerization of Vpr by producing a more thermodynamically stable condition through the hydrophobic interaction of two monomers (Pandey et al., 2009). NMR studies on the C-terminal region of Vpr (aa 52-96) have showed that it has a tendency to dimerize in solution (Schuler et al., 1999).

2.5.2 Structure-Function Relationship

Extensive efforts have been channeled towards the understanding of the molecular and structural biology of the HIV-1 Vpr; however, the lack of a crystal structure for this

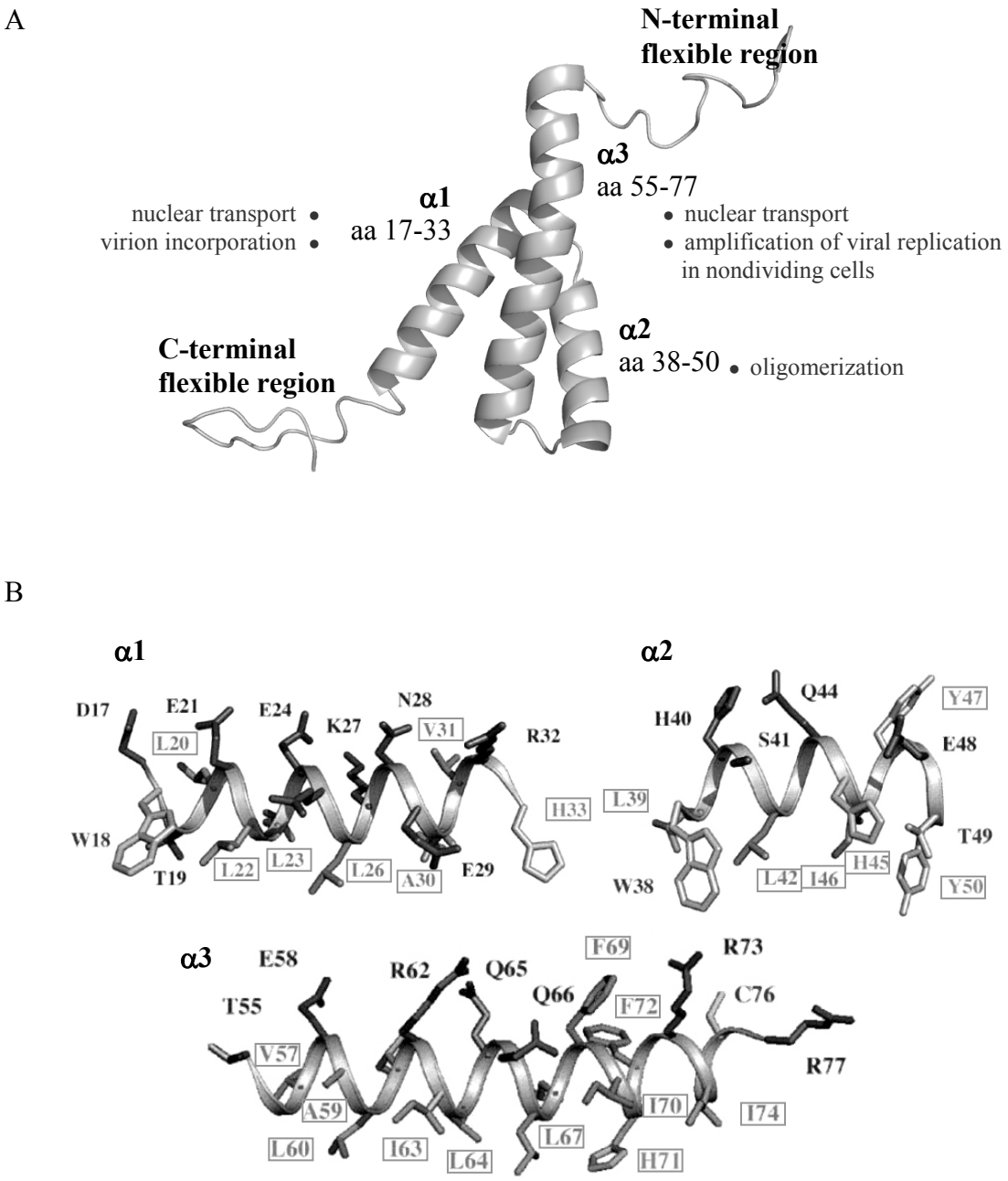


Figure 2.4 The structure of Vpr and the hydrophobicity of its three helices.

A, Ribbon diagram showing three α-helices flanked by two flexible regions. B. The three amphipathic α-helices of Vpr WT with their side-chains shown. Amino acids are labeled according to their hydrophobicity from boxed grey (hydrophobic) to black (hydrophilic) (Morellet et al., 2009, Morellet et al., 2003).

protein limits our attempts to further understand it. The three-dimensional structure of a protein is important to understand the structural basis for its functions; in drug discovery protein crystallography can tell us how small molecules bind to their target protein (Canduri and de Azevedo, 2008). In order to know the roles of the three α-helices in Vpr function, molecular biology experiments (deletion, insertion and substitution mutagenesis) were carried out before the three-dimensional structure was known. From those studies, the structure-function relationship of Vpr was deduced. Mutations that disrupted the tertiary and the secondary structure of the protein were found to abolish numerous Vpr activities such oligomerization, virion incorporation, cell cycle arrest and localization of Vpr in the nucleus (Morellet et al., 2009, Pandey et al., 2009).

Based on investigations on the the flexible domains, the inter-helical loops, and the three α-helices of Vpr, α1 appears to be responsible for nuclear transport and virion incorporation; α2 seems to play a role in the oligomerization of Vpr; and α3 in appears to have roles in nuclear transport and amplification of viral replication in nondividing cells (Kichler et al., 2000, Mahalingam et al., 1997, Schuler et al., 1999, Yao et al., 1995). A leucine-rich region, in the last 14 residues of α3 has a potential for leucine zipper formation necessary for nuclear localization, transactivation and dimerization of Vpr (Gaynor and Chen, 2001, Jacotot et al., 2000, Wang et al., 1996). The C-terminal domain of Vpr contains many basic residues and was shown to be involved in G2 arrest, upregulation of viral expression and important for the stability of the protein (Sawaya et al., 2000, Forget et al., 1998, Thotala et al., 2004).

Overall, three of the most important functions of Vpr, cell cycle arrest, LTR activation and apoptosis are inter-related because they are believed to serve optimal viral production (Figure 2.5). Vpr is thought to induce those effects by manipulating the ubiquitin proteasome system (UPS) to lead to cell cycle arrest.

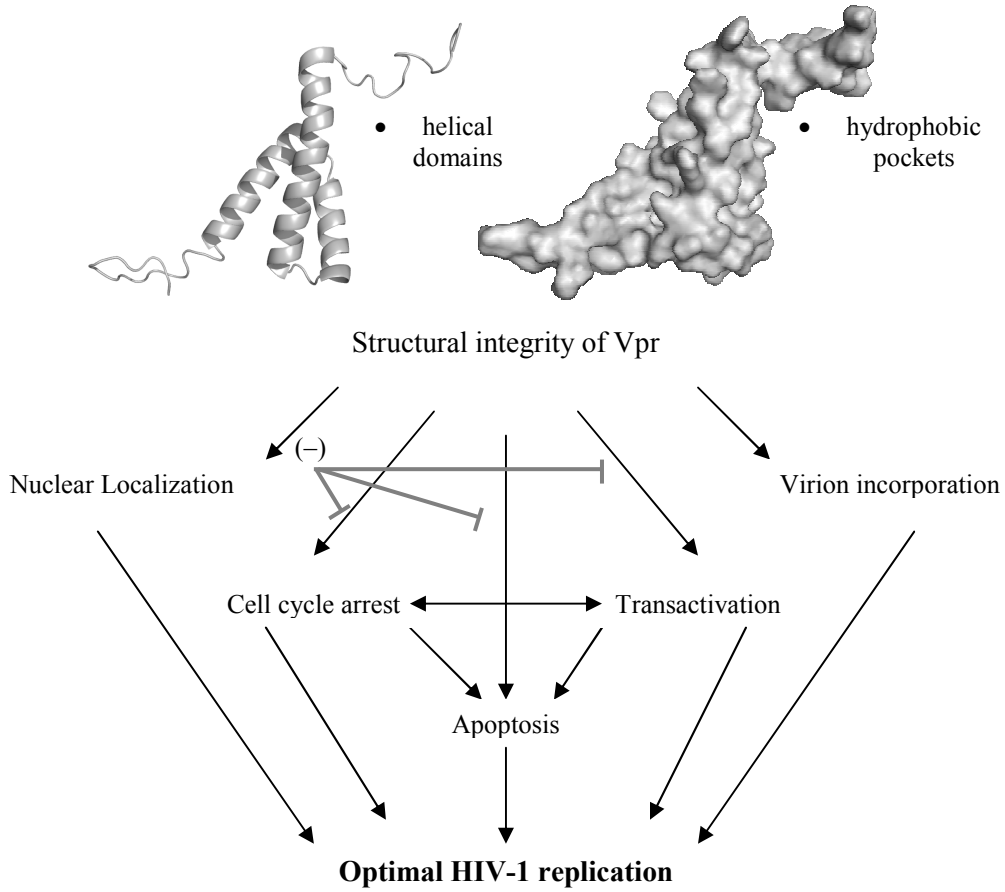


Figure 2.5 Interrelationship of the different functions of HIV-1 Vpr.

Individual functions of Vpr are dictated by the overall structural integrity of the protein. It is believed that the helical domains (ribbon diagram) and hydrophobic pockets (molecular surface diagram) of Vpr are the main contributors to its multiple roles during the viral life cycle. Inhibition of nuclear localization may decrease cell cycle arrest, transactivation and apoptosis. Altogether, these functions synergistically influence the optimal replication of HIV-1 (Pandey et al., 2009).

2.6 Vpr-induced Cell Cycle Arrest

Vpr was reported to be able to manipulate the cell cycle and induce arrest at the G2-to-M transition in 1995 by several groups independently (He et al., 1995, Jowett et al., 1995, Re et al., 1995, Rogel et al., 1995). Apart from mammalian cells, Vpr is also known to inhibit cell growth in the fission yeast, *Schizosaccharomyces pombe* (Zhao et al., 1996), and in the budding yeast, *Saccharomyces cerevisiae* (Macreadie et al., 1995). The growth inhibition in those eukaryotic systems implies that the cellular target of Vpr is a highly conserved protein.

G2 arrest usually takes place in response to DNA damage, in this phase the cell can repair existing DNA damage before undergoing mitosis. If there are too much DNA damage and ineffective DNA repair, it will lead to self-destruction of the cell by apoptosis (Orren et al., 1997). Similarly, Vpr-caused G2 arrest in cells normally leads to apoptosis (Chang et al., 2000, Shostak et al., 1999, Stewart et al., 1997, Watanabe et al., 2000, Zhu et al., 2001).

At first it was speculated that induction of cell cycle arrest and apoptosis would be disadvantageous to a retrovirus because during G2 arrest, cellular protein translation would be significantly down-regulated (Pyronnet and Sonenberg, 2001), and the induction of apoptosis would prevent the budding of progeny virions through cellular membranes. However, a recent model proposed for HIV-1 infection of activated lymphocytes suggests that after the virus establishes itself in the host cell, progeny virions are produced in a burst during the G2 phase of the cell cycle (Goh et al., 1998), and then apoptosis is induced as a late event (Zhu et al., 2001). Findings suggest that cell cycle arrest in G2 may actually serve the virus by allowing for more efficient viral replication via upregulation of viral transcription because the viral LTR is highly active in G2 (Goh et al., 1998, Hrimech et al., 1999, Zhu et al., 2001).

In 1994, a cellular protein found in association with Vpr by co-precipitation experiments was first described but at that time it of unknown function and was named Vpr-binding protein (VprBP) (Zhao et al., 1994a). Initial studies did not show VprBP to be involved in the cell cycle effects of Vpr, and it was only in recent years that researchers were able to gain more insights on the relationship between Vpr and VprBP through other studies conducted on the ubiquitin/proteasome system (Ciechanover, 1998). In order to fully understand the Vpr-VprBP interaction, it is essential to know the ubiquitin-proteasome system of protein degradation.

2.6.1 Vpr and the Ubiquitin Proteasome System

A few years ago, it was first reported that Vpr interacts with cullins 1 and 4 (Cul1, Cul4), components of modular ubiquitin ligases (Schrofelbauer et al., 2005) and this finding was followed by the identification of a family of proteins that were associated with the damaged-DNA specific binding protein 1 (DDB1), a Cul4 adaptor. This DDB1-associated family of proteins, which also includes VprBP, functions as substrate specificity modules in Cul4- and DDB1-based ubiquitin ligase complexes (Angers et al., 2006, He et al., 2006, Higa et al., 2006a, Jin et al., 2006).

Following this discovery, VprBP was named DDB1- and Cul4A-associated factor 1 (DCAF1); henceforth, 'VprBP' is used interchangeably with 'DCAF1' in this thesis. Next, it was shown that, through its interaction with VprBP, Vpr becomes a part of a larger complex (Figure 2.6) that includes Cul4A, DDB1, and, presumably, Ring box protein 1 (Rbx1) and a ubiquitin-conjugating enzyme or E2 (Belzile et al., 2007, DeHart et al., 2007, Hrecka et al., 2007, Le Rouzic et al., 2007, Schrofelbauer et al., 2007, Tan et al., 2007, Wen et al., 2007).

In vivo, DDB1 links Cul4 to substrate specificity subunits which are referred to as DCAFs. The ubiquitination substrates for several DCAFs have been identified and were found to be involved in cell cycle events such as – recruiting the origin of replication factors to prevent re-replication of DNA (Higa et al., 2003, Hu et al., 2004), promoting degradation of target proteins and histones as part of a response to DNA damage (Sugasawa et al., 2005, Wang et al., 2006), being responsible for destruction of cyclin-dependent kinase inhibitors (Higa et al., 2006b). Taken together, these reports show that the Cul4-DDB1 ubiquitin ligases are likely to be involved in genome stability, DNA replication, and cell cycle checkpoint control.

Based on the above reports, a model for the Vpr–VprBP complex was proposed in which Vpr binds to a Cul4-DDB1-DCAF1 ubiquitin ligase (Figure 2.6), to trigger polyubiquitination and subsequent degradation of a putative cellular protein, which results in activation of the G2 checkpoint (Dehart and Planelles, 2008). This model proposes that Vpr binds to VprBP and the putative ubiquitination substrate protein on two different interfaces.

Using site-directed mutagenesis, Zhao et al showed that Vpr binds to VprBP at a leucine-rich (LR) motif ^{60}LIRILQQLL^{68} located on the third α-helix of Vpr (Zhao et al., 1994a). And much recently, residue Q65 was shown to be essential for the Vpr–VprBP interaction as the Q65R Vpr mutant was unable to bind to VprBP; the Q65R Vpr mutant also failed to induce G2 arrest, indicating that the Vpr-VprBP interaction is required for Vpr's cell cycle arrest activity (Le Rouzic et al., 2007). However, truncation of 18 residues from the carboxy-terminal of Vpr (aa 79-96) or replacement of arginine at position 80 by alanine, Vpr(R80A), resulted in Vpr proteins that still bound to DCAF1 but were unable to induce G2 arrest (DeHart et al., 2007, Le Rouzic et al., 2007).

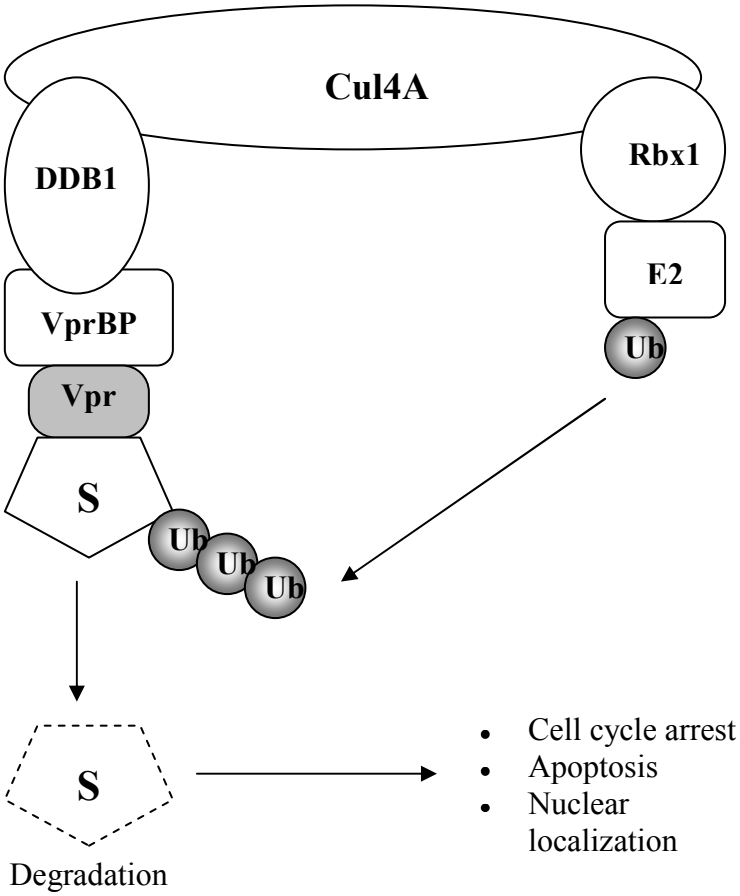


Figure 2.6 Proposed model for the interaction of Vpr with an E3 ubiquitin ligase.

Diagrammatic representation of the ubiquitin ligase complex that interacts with Vpr. The current model states that Vpr recruits an unknown cellular substrate (S) to the Cul4/DDB1/ VprBP ubiquitin ligase; it is thought that the cellular substrate protein subsequently becomes ubiquitinated and degraded by the 26 S proteasome, leading to the various cytopathic effects of Vpr (DeHart et al., 2007).

It was thus proposed that the interaction of Vpr with an E3 ubiquitin ligase would either inhibit or activate the ligase's enzymatic activity. If, Vpr acts as an inhibitor of the E3 ligase, then the depletion of DCAF1 would mimic the cell cycle arrest activity of Vpr. However, from several RNA interference (RNAi) experiments, it can be deduced that RNAi-mediated depletion of DCAF1 abolished Vpr‑mediated G2 arrest (DeHart et al., 2007, Hrecka et al., 2007, Le Rouzic et al., 2007, Wen et al., 2007), showing that Vpr actively recruits DCAF1 and promoted the activation of the E3 ligase, because depletion of DCAF1 were able to counteract the effect of Vpr. This finding supports the idea that Vpr requires VprBP to trigger the activity of the E3 ligase.

Overall, these observations indicate that binding of Vpr to DCAF1 is required, but not sufficient, for induction of G2 arrest, and the carboxy-terminal domain of Vpr is involved in the recruitment of a cellular protein, whose ubiquitination leads to G2 arrest (Figure 2.6).

Therefore, as previously mentioned, Vpr-induced G2 arrest is important for HIV assault and most likely contributes to the pathogenesis of HIV-1 by increasing the transcriptional activity of the viral promoter by several-fold during G2/M (Goh et al., 1998, Hrimech et al., 1999, Zhu et al., 2001), which leads to enhanced production of viral particles (Goh et al., 1998). It was also suggested that accumulation of infected cells in G2 may be favorable for the selective translation of viral products because of the presence of a putative internal ribosome entry site in the HIV-1 genome (Brasey et al., 2003). Ultimately, G2 arrest leads to the death of infected cells by apoptosis.