To Study Molecular Determinants of HIV-1 Neuro Invasion and Prevalence Rate of HIV Associated Dementia in HIV Infected Pakistani Population.

By

SHEHLA ABBASI

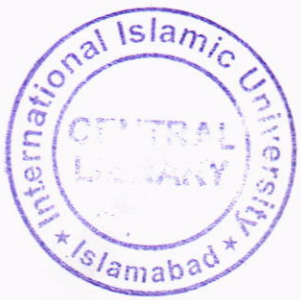
18-MSBI/FBAS/F08

Department of Bioinformatics and Biotechnology

Faculty of Basic & Applied Sciences

International Islamic University Islamabad

(2010)



① HIV (Viruses) medical microbiology

بسم الله الرحمن الرحيم

Department of Bioinformatics and Biotechnology

International Islamic University Islamabad

Dated: July 27, 2011

FINAL APPROVAL

It is certificate that we have read the thesis submitted by Ms. Shehla Abbasi and it is our judgment that this project is of sufficient standard to warrant its acceptance by the International Islamic University, Islamabad for the M.S Degree in Bioinformatics

COMMITTEE

External Examiner
Dr. Mahmood Akhter Kayani
Head of Department of Biosciences,
CIIT, Islamabad

Internal Examiner
Dr. Naveeda Riaz
Assistant Professor

Dr. Asma Gul
Assistant Professor

Supervisor
Dr. Shaheen Shahzad
Chairperson
Department of Bioinformatics and Biotechnology,
International Islamic University, Islamabad

Co-Supervisor
Dr. Muhammad Mukhtar
Vice Chancellor
Islamia University, Bahawalpur

Co-Supervisor
Ms. Saima Kalsoom
Quaid -e-Azam University, Islamabad

Dean ~~BI&BT~~ FBAS
Dr. Irfan Khan
International Islamic University, Islamabad

A thesis submitted to Department of Bioinformatics & Biotechnology, International Islamic University, Islamabad as a partial fulfillment of requirement for the award of the degree of M.S in Bioinformatics

DEDICATION

This humble effort

The fruit of studies and thoughts

Dedicated

To the Holy Prophet

Hazart Muhammad (P.B.U.H)

The ocean of knowledge and

The greatest social reformer,

My worthy father

Taj Mohammad Abbasi

Who always inspired and encouraged

Me for what I what and grafted in the

University to get on the higher Ideas of life,

My Beloved

Mother

Which is strongly believed that

Her prayers are always with me,

My Loving Brothers, Supporting Sister

For their love, encouragement

And their moral support and guidance

That inspired me to accomplish

This humble effort

And

Dedicated

To All Victims of HIV

Who fight courageously with this disease

And

Live with high spirits

DECLARATION

I hereby declare that the work present in the following thesis is my own effort, except where otherwise acknowledged and that the thesis is my own composition. No part of the thesis has been previously presented for any other degree.

Date July 27, 2021

Shehla Abbasi

(Shehla Abbasi)

ACKNOWLEDGEMENTS

I am indebted to Almighty Allah, The propitious, the benevolent and sovereign whose blessing and glory flourished my thoughts and thrived my ambitions, giving me talented teachers, affectionate parents, sweat siblings and unique friends. Crembling lips and wet eyes praise for Holy Prophet Muhammad (P.B.U.H) for enlightening our conscience with the essence of faith in Allah, converging all His kindness and mercy upon him.

The research work presented in this dissertation was accomplished under the sympathetic attitude, animate directions, observant pursuit, scholarly criticism, cheering perspective and enlightened supervision of Dr. Mohammad Mukhtar, Vice Chancellor, Islamia University Bahawalpur. His through analysis and rigorous critique improved not only the quality of this dissertation, but also my overall understanding of research concepts. I am grateful to his ever inspiring guidance, keen interest, scholarly comments and constructive suggestions through out this research work. I am indebted to him more than he knows.

I dream it my utmost pleasure in expressing my cardise gratitude with the profound benedictions to my supervisor Dr. Shaheen Shahzad, Assistant Professor, Department of Bioinformatics and Biotechnology, International Islamic University, Islamabad. Her generous and expert guidance, keen interest at every step and continuous encouragement throughout this research work enabled me to achieve this goal. Moreover, I am very thankful to her for critically reading this dissertation and providing me guidance to improve it without which this work would have been impossible.

I feel highly obliged to express my sincere gratitude to my co-supervisor Ms. Saima Kulsoom, lecturer who helped me to cross every obstacle that came in my way to achieve this goal. I want to extend my deep emotions of appreciations, gratitude and indebtedness for her valuable guidance.

I would also like to acknowledge International Islamic University Islamabad which not only provided me the platform to undertake this research work and accomplish it within due course time but also provided me a thought of 'life with purpose' along with the great Islamic and moral values.

A special thanks to bioinformatics faculty for equipping me with the latest knowledge and providing me academic base to take up this work. It would be very unjust if I do not pay my sincere gratitude to Sir Imran Saeed, Sir Hafiz Bashir, Sir Sajjad, Ms. Zareen Saba, Mrs. Samreena, Mrs. Khan, and many more for enlightening my thoughts and broadening my vision towards life.

I shall be failing in my duty if I do not put across my gratitude to all lab members of Dr. Mohammad Mukhtar in PMAS-Arid Agriculture University Rawalpindi for their great support, warm welcome and knowledge sharing, without them achieving this goal would have been very thorny. I am also very grateful to National AIDS Control Program (NACP) staff for providing me all the facilities, moral and financial help to accomplish wet lab part of my work especially Dr. Nadeem, Mr. Mohammad Qasim, Mr. Usman Waheed and Mr. Tariq Mir.

No acknowledgments could ever adequately repay my obligations to my affectionate and adoring Parents, Family Members and my very dear friends Javeria Irum, Pakeeza Akram, Faiza Naeem, Ammara Khalid, Maimoona Ali, Mehwish Huma, Anisa Zia, Shekoh Waraich and Atiya Mehmood Ahmed whose hands always rose in prayers for me and without their encouragement moral and financial support; the present distinction would have merely been a dream. May Allah Almighty infuse me with the energy to fulfill their noble inspirations and expectations and further edify my competence.

Finally, I apologize if I have caused anger or offence to anybody and the errors that remain in the manuscript are my responsibility alone.

May Allah bless all these people with long, happy and peaceful lives (Ameen)

Shehla Abbasi

CONTENTS

LIST OF ABBREVIATIONS

AIDS	Acquired immunodeficiency syndrome
BBB	Blood brain barrier
CNS	Central nervous system
HAD	HIV-associated dementia
ADC	AIDS dementia complex
HAART	Highly active antiretroviral therapy
APOBEC3G	Apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3
HIV	Human immunodeficiency virus type-1
ZO-1	Zonula occluden-1
NSE	Neuron Specific Enolase
ART	Anti retroviral therapy
ELISA	Enzyme-linked immunosorbent assay
HANDs	HIV-associated Neurocognitive Disorders
MAGUK	Membrane-associated guanylate kinase
PDZ	PSD-95/ disc-large/Zonula occludens-1
SH3	Src homology-3
GUK	Guanylate kinase
3D	Three-dimensional
°A	° Angstroms
VMD	Visual Molecular Dynamics

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

HIV- associated dementia (HAD) largely depends on continuous seeding of central nervous system with HIV infected immune system cells and direct invasion of HIV or HIV proteins mainly gp120 and tat from the periphery. It has been reported that HIV envelop glycoprotein gp120 is responsible for proteasome-mediated degradation of tight junction protein zonula occluden-1(ZO-1), but detailed molecular mechanism still needs to be figured out. HIV-associated dementia represents the most severe manifestation of HIV-associated Neurocognitive Disorders (HANDs) that leaves it's life lasting impacts on patient's cognitive, behavioral and motor functions therefore patients are consistently required to be monitored for initial symptoms of HAD but in Pakistan no studies have been persuaded to figure out prevalence rate of HAD in non- ART HIV patients.

This research work is focused on two main domains; understanding molecular mechanism lying in proteasome-mediated degradation of ZO-1 protein via gp120 protein and studying prevalence rate of HAD in non- ART HIV infected Pakistani population. Using bioinformatics techniques it is inferred that gp120 protein initially marks ZO-1 protein which then lead to proteasome mediated degradation. Residues of gp120 protein involved in this molecular mechanism are thr 22, thr 23 and ser 11 that resides in V3 loop of gp120 protein while that of PDZ-1 domain are gln 70 and asn 72. There exists a great probability that gp120 protein marks ZO-1 protein via interacting at multiple positions as Ile 1 of first highlighted peptide of V3 region shares the distances of 4.36 and 3.95 angstrom with ile 18 and trp 19 of PDZ-1 domain respectively. Enzyme-linked immunosorbent assay (ELISA) revealed that prevalence rate of

HAD in Pakistan is 12.5% which is far greater than neighboring country India's prevalence rate of 1-2% and meets with westerns countries 10-24%. So during the treatment of HIV patient physician should also focus in detail on mental health of patient due to high prevalence rate of HAD in Pakistan.

INTRODUCTION

Human immunodeficiency virus type-1 (HIV) is a causative agent of acquired immunodeficiency syndrome (AIDS), a condition in which human immune system is undermined to extent that it is not able to resist any opportunistic infection causing organism and this conditions of total immune system failure is termed as AIDS . AIDS is a multisystem disorder; it affects human lungs, gastrointestinal tract, spleen, genital tract and brain, to cause neuroAIDS (Ghafouri *et al.*, 2006; Clarke *et al.*, 2000). Since brain is a vital human organ so it is provided with supplementary protection by nature in the form of blood brain barrier (BBB) which is a continuous mono layer of tightly packed endothelial cells; it functions as a filter and separate central nervous system (CNS) from periphery. BBB is selectively permeable that does not allows water-soluble molecules, ions and immune sytem cells to move across it (Gras *et al.*, 2010; Sharma *et al.*, 2009). Hence any blood –prone disease causing organism should disrupt it to invade CNS. However, the exact mechanism for neuroinvasion of HIV is still unclear while the most compelling evidences are available for ***trojan horse hypothesis***; according to this model it is believed that HIV-1 and other lentiviruses hijacks the immune system cells (most important are macrophages) , hides in them and cross blood brain barrier with these infected immune system cells. Another mode of neuronvasion by HIV under research and

discussion is disruption of blood brain barrier either by HIV it self or via HIV proteins namely, Tat protein and GP120 protein (Engelhardt *et al.*, 2009; Weiss 1993).

Neuro invasion by HIV leads to ***HIV-associated dementia (HAD)*** often also known as AIDS dementia complex (ADC), HIV encephalopathy and HIV-associated neurocognitive disorder. In the absence of opportunistic infections, major clinical symptoms of HAD includes impaired short term-memory coupled with reduced ability of mental concentration, leg weakness, slowness of hand movement and gait as well as depression. These symptoms are often accompanied by behavioral symptoms such as personality changes, apathy and social withdrawal (Reger *et al.*, 2002; Janssen *et al.*, 1992). During years when AIDS was appearing as an epidemic, neuroAIDS appeared only in patients with low CD4 cell count and high viral load. But with the advent of highly active antiretroviral therapy (HAART) in 1995 neuroAIDS was first suppressed; as HAART prolonged the progress of AIDS, HAD has become an independent death factor due to AIDS (Kaul 2009). HIV associated dementia is prevailed in 10-24% HIV-infected patients in western countries (Grant *et al.*, 2005) while in India only 1-2% of HIV patients faces deadly outcomes of HIV-associated dementia (Wadia *et al.*, 2001; Satishchandra *et al.*, 2000). In Pakistan no study has been carried out till date to determine the prevalence of HAD in Pakistani population of HIV & AIDS patients.

HIV Gp120 is an envelope glycoprotein that is meant to seek target site for attachment of HIV on target's cell surface (Clapham *et al.*, 2001). It's contribution in neuroAIDS is well suspected as in one study it is implicated as reason for disruption of tight junctions between endothelial cells of BBB however, disrupting molecular mechanisms are not known (Kaul 2009). In recent study it is identified that blood brain barrier is breached by gp120 mediated proteasomal degradation of Zonula occluden-1 (ZO-1) protein (Nakamuta *et al.*, 2008). ZO-1 is a multi-domain scaffolding protein that joins and holds transmembrane proteins of tight junctions with actin filaments of cytoskeleton therefore is responsible for BBB's integrity (Alan *et al.*, 2009). However, till date no detailed mechanism has been elucidated for proteasome-mediated degradation of ZO-1 protein (Nakamuta *et al.*, 2008).

Human apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3 (APOBEC3G) is an innate immune system protein that like other antiviral cytidine deaminases exerts its anti-HIV-1 effect through modification of viral genome involving deamination of cytidine to uridine (C → U). HIV-1 Vif protein counteracts this deamination by deactivating APOBEC3G through unique binding that result in limited encapsidation of antiviral APOBEC3G protein followed by proteosomal degradation (Huthoff *et al* 2007).

Molecular mechanism for proteasomal degradation of APOBEC3G (A3G) by HIV-1 vif is very well understood. The initial step of this mechanism involves interaction of HIV-1 Vif with huAPOBEC3G involving three main residues; Aspartic acid (128), Proline (129) and Aspartic acid (130). This interaction results in recruitment of a cullin 5-elongin B/C E3 ubiquitin ligase which polyubiquitinates APOBEC3G and finally A3G is targeted to proteasome for degradation. Hence the initial step, i.e., interaction of huAPOBEC3G with HIV-1 vif, is required to mark this innate immune system protein as a target for proteasomal degradation (Huthoff *et al.*, 2007).

Based on these grounds we hypothesize that tight junction protein zonula occluden-1 is first marked by HIV-1 gp120 protein to further proceed towards proteasomal degradation. Hence this research work is carried out to reveal whether Zonula occluden-1 interacts with HIV-1 gp120 protein or not, if it does then which residues of both ZO-1 and gp120 is important for this explicit interaction. Success in this project will lead to expounding of molecular mechanism behind HIV neuroinvasion, a foremost rationale of neuroAIDS.

Neuron Specific Enolase (NSE) is a type of enolase, enzymes responsible for catalysis of 2-phosphoglycerate (2-PG) to phosphoenolpyruvate (PEP) which is the ninth step of glycolysis. NSE is found in abundance in neurons and neuroendocrine

origin and in case of any injury to neurons it's level is increased in serum and cerebrospinal fluid therefore it act as a biomarker (Lee *et al.*, 2010).

Prevalence of HIV-associated dementia in HIV infected patients of Pakistani population is not reported till date. This research work will also report prevalence of HAD in Pakistani population of non-ART HIV patients using wet lab technique ELISA and NSE as a biomarker. This study will impart it's effects on determining whether the strain of HIV in Pakistan patients is strong enough to evade Blood Brain Barrier security or not, statistics would help in medication of patients and would help to follow mental health of patients.

Following are the objectives of this study;

1. Determining whether zonula occluden-1 protein interact with gp120 or not for proteasome mediated degradation of ZO-1.
2. If it does then determining residues of both ZO-1 and gp120 involved in this interaction.
3. Studying prevalence of HIV-associated dementia in non-ART HIV patients of Pakistani population.

LITERATURE REVIEW

Human immunodeficiency virus (HIV), member of retrovirus family attacks human immune system that begins to fail and self invited opportunistic infection causing agents produces life threatening diseases, this condition is called acquired immunodeficiency syndrome (AIDS). HIV enters in the body via four main routes: unsafe sex, contaminated needles, breast milk, and transmission from an infected mother to her baby at birth (Weiss 1993; Douek *et al.*, 2009).

2.1 Stages of HIV Infection & AIDS

HIV after entry in the body of uninfected person from infected patient remains asymptomatic for two to four weeks; this very initial stage of infection is called ***incubation period*** (Grabar *et al.*, 2009). During and after this period HIV attacks immune system cells mainly CD4+ T cells, macrophages and dendritic cells. This leads to rapid fall of circulating CD4+ T cells and marked increase of viral load (approaching several million viruses per mL) within bodily fluids (blood, semen, breast milk, vaginal fluid and pre-ejaculate) where HIV can remain as either free particle or reside inside infected immune system cells. This is ***acute HIV infection*** stage (Piatak *et al.*, 1993). During this period which lasts for usually about 2 to 4 weeks, 80 to 90% patients shows most common symptoms of which may include fever, lymphadenopathy, pharyngitis, rash,

myalgia, malaise, mouth and esophageal sores, and may also include, but less commonly, headache, nausea and vomiting, enlarged liver/spleen, weight loss, thrush, and neurological symptoms. Infected individuals may experience all, some, or none of these symptoms. Since these symptoms are not very specific therefore often patients remains undiagnosed (Pantaleo *et al.*, 1997).

Recognition of foreign particles inside body evokes immune response and results into the activation of CD8+ killer T cells, which kills infected immune system cells either by antibody production or seroconversion. This response of body resumes CD4+ T cells count and viral load within bodily fluids is minimized. A good CD8+ killer T cells response leads to slow progression of disease and better prognosis, although HIV is not completely eliminated from peripheral blood (Reger *et al.*, 2002).

Because of strong immune system response body apparently becomes healthy and enters in infection's ***clinically latency*** stage, which lasts from two weeks to twenty years depending upon strain of attacking HIV, patient's strength of immune response, treatment during course of infection and several others. During this phase of disease HIV silently keeps replicating in lymphoid organs, where large amounts of virus become trapped in the follicular dendritic cells (FDC) network and continuously weakens immune system by three main mechanisms: First, direct viral killing of infected cells; second, increased rates of apoptosis in infected cells; and third, killing of infected $CD4^+$ T cells by CD8 cytotoxic lymphocytes that recognize infected cells (Kahn *et al.*, 1998).Tissues in the

surroundings are rich in CD4+ T cells which gradually and rapidly become infected, at this still infectious stage $CD4^+$ and $CD45RO^+$ T cells carry most of the proviral load (Burton *et al.*, 2002).

This constant decline in the quantity of CD4+ T cells results in the CD count below critical level, i.e., 200 cells per μL as a result cell-mediated immunity is lost, variety of microbes that are usually resisted successfully by a robust CD4+ - mediated immunity in healthy human, attacks facing no resistant from target (Clapham *et al.*, 2001). This stage of disease is called ***Acquired immune deficiency syndrome (AIDS)***.The first symptoms that often appears in AIDS patients includes moderate and unexplained weight loss, prostatitis, skin rashes, recurring respiratory tract infections and oral ulcerations (Decker *et al.*, 2000). Early on the microbes against which resistance is lost are *Mycobacterium tuberculosis* and oral Candida species (Feldman *et al.*, 2005). Later, herpes simplex eruptions, Epstein-Barr virus-induced B-cell lymphomas, or Kaposi's sarcoma are responsible for the worsening condition of infected body. Pneumonia is common among AIDS patients and often become fatal (Clapham *et al.*, 2001). The final stage of AIDS is accompanied by attack of cytomegalovirus (another herpes virus) or Mycobacterium avium complex (Zaidi *et al.*, 2002).

2.2 HIV Compartmentalization

Human immunodeficiency virus targets different human tissues, a concept called ***HIV compartmentalization***. Targets for trafficking of HIV are lungs, blood, spleen, gastro-intestinal tract, genital tract and brain. Although CD4+ T cells are major infected cells during the course of disease as they accounts for over 99% of replicating virus within host but, macrophages (<1% of total infected cells burden) are considered responsible for disseminating HIV throughout the host. Figure 2.1 depicted the dissemination of HIV when it is transmitted via sexual course (Clarke *et al.*, 2000);

2.3 HIV Entry in to Host Cell

Variation in theV3 loop sequence of Gp120 envelope protein of HIV is most importantly responsible for tissue specific HIV infection (Clarke *et al.*, 2000). Gp120 protein is a glycoprotein exposed on the surface of virus, encoded by *env* gene and is approximately 500 residues long. Virus enters in host cell by seeking specific cell type receptors, heterodimers of three Gp120 protein forms an envelope spike which binds to target cell via head while it's tail is bound to a transmembrane glycoprotein gp41. While HIV infects a cell; Gp120 seeks a target site, attaches itself to it and then detaches. This seeking, attachment and detachment procedure is followed by fusion of gp41 protein to target cell surface. Since HIV belongs to genus *Lentivirus,* therefore like other members of same genus it transfers it's positive sense single stranded RNA in to host cell which is

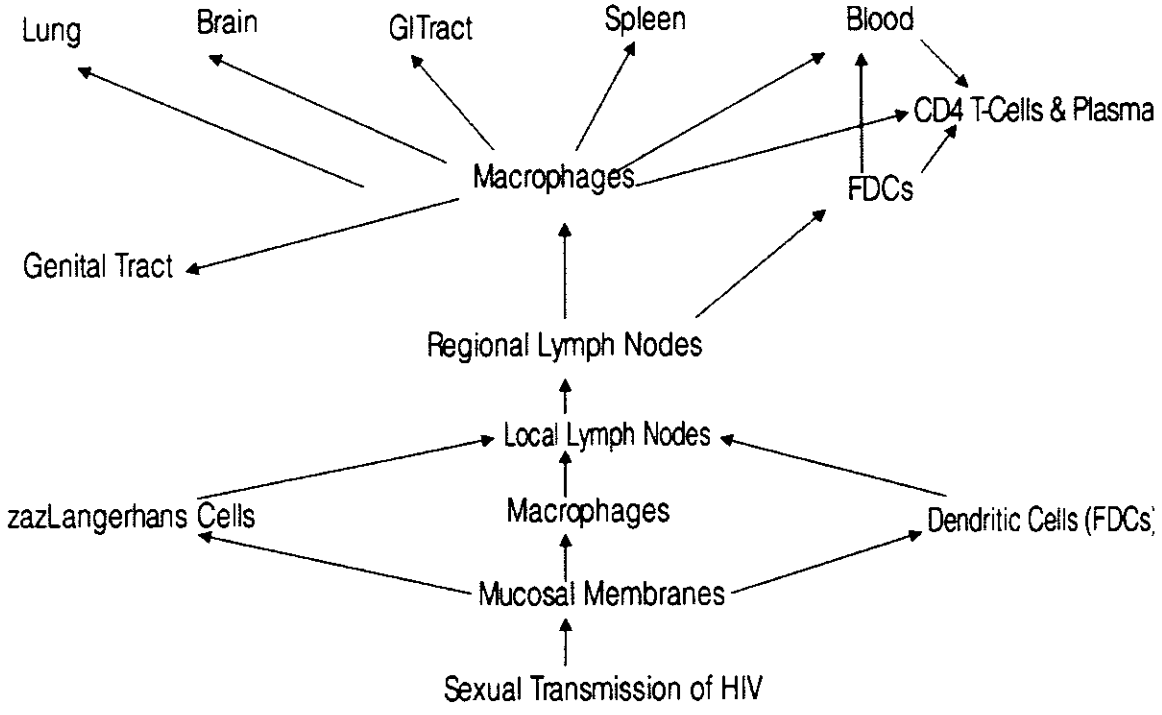


Fig.2.1 Dissemination of HIV when viral mode of entry inside the human body is sexual.

then converted to double stranded DNA by virus encoded protein reverse transcriptase. This double stranded DNA is incorporated into host cell DNA via viral enzyme integrase and HIV is ultimately transcribed using host machinery (Smith *et al.*, 2006). Initially it was believed that HIV only infects CD4 + T-cells and therefore, is mainly involved in immune system suppression. But, in 1985 HIV was isolated from brain tissues, spinal fluid and peripheral nerves of HIV- infected patients (Sharma *et al.*, 2009). This observation provided the evidence that HIV enters the brain in early stage of infection most probably via infected macrophages and lymphocytes then reside there primarily in perivascular macrophages and microglia (Gras *et al.* 2010). Brain is a vital organ that regulates the body's action and reactions (Murre *et al.*, 1995). HIV infection to brain cells induces neurocognitive complications defined as HIV-associated Neurocognitive Disorders (HANDs) (Kaul *et al.*, 2009).

2.4 HIV-Associated Neurocognitive Disorders (HANDs)

HAND is further divided into three categories according to the standardized measure of dysfunction. These includes; asymptomatic neurocognitive impairment, mild neurocognitive disorder (MND) and HIV-associated dementia (HAD) (Kaul *et al.*, 2009). In the first stage of neurocognitive complications, i.e., asymptomatic neurocognitive impairment, no symptoms for brain specific disorders appears however, patient is carrier and he/she has potential to develop severe life threatening brain disorders so requires

examination by physician on regular basis. When patient reaches to the stage of Mild neurocognitive disorder (MND), he/she faces slowed extremity movements however; patient is able to perform daily normal functions, normal gate and strength is retained (Pollok *et al.*, 2001).

HIV-associated dementia represents the most severe manifestation of HANDs that leaves it's life lasting impacts on patient's cognitive, behavioral and motor functions. Neuropathological correlates of HAD comprises of microglial nodules, multinucleated giant cells, and activated resident microglia, infiltration predominantly by monocytoid cells, including blood-derived macrophages, widespread reactive astrocytosis, myelin pallor, and decreased synaptic and dendritic density along with prominent neuronal loss. This damage and loss of neuron occurs in different parts of brain including cerebellum, putamen, frontal cortex and substantia nigra (Gras *et al.* 2010).

AIDS has become pandemic after it's successful spread being epidemic for more than four decades. Three years back in 2007, Survey showed that approximately 33.2 million people in the world are living with HIV & AIDS, including 3,30,000 children (UNAIDS, WHO 2007). Considering HIV as one of the biggest life threatening disease world wide, scientists have put forward all possible resources and efforts to discover the cure for this disease. However, till date no success has been marked in vaccine development for HIV & AIDS. Advent of highly active antiretroviral therapy

(HAART) in mid 1990's was a major success in drug development of HIV infection; it does not cure the disease but at least delays the progression of AIDS (Kaul *et al.*, 2009).

Patients with HIV associated dementia suffer with the problems of lack of concentration, memory loss, difficulty in understanding and responding to question, poor coordination, weakness in legs, loss of bladder or bowl control, mood swings and impaired judgment (Gray *et al.*, 2001). In short patient is not even able to take care of his/her self. In the beginning of AIDS HAD occurred to patients with very low CD4 count and high viral load, HAART at first reduced the incidence of HAD but now as disease progresses and life span of AIDS patient is increased due to treatment, HAD has became independent risk factor for death due to AIDS (Kaul *et al.*, 2009).

2.5 Blood Brain Barrier

The brain monitors and regulates the body's actions and reactions. It continuously receives sensory information, and rapidly analyzes this data and then responds, controlling bodily actions and functions (Murre *et al.*, 1995). Therefore, brain is a power hub of all body functions and due to it's highly imperative status it is provided by extra protection from Allah (SWT) in the form of blood brain barrier (BBB).

Blood Brain barrier is tightly packed continuous mono layer of endothelial cells localized at central nervous system (CNS) microvessels. Blood brain barrier does not allow free paracellular diffusion of water-soluble molecules and cells, between peripheral

blood and cerebrospinal fluid. This selective permeability is achieved by the presence of adherens junctions and tight junctions between two adjacent endothelial cells of BBB and by low pinocytic activity of these cells (Engelhardt *et al.*, 2009). The adherens junctions are considered responsible for mediating cell-cell adhesion and signaling pathways that control cell morphology, growth and differentiation and have been considered as important regulators of vascular permeability. The inter-endothelial tight junction (TJ) regulates the movement of ions, macromolecules and immune cells through the paracellular space and is critical for ion transport (Alan *et al.* 2009). TJs are composed of intricate complex of several transmembrane and cytoplasmic proteins ultimately linked to actin filaments of cytoskeleton (Engelhardt *et al.*, 2009) as shown in fig 2.2.

Among integral membrane proteins, occludin was first protein discovered to be localized within tight junctions including that of BBB. Any null mutation in occludin does not interfere with normal morphology of tight junction found in BBB therefore, it is not considered to play an important role in tight junction assembly. Whereas claudins, integral membrane tight junction protein family comprises of 24 members, have been proved to be sufficient for the configuration of tight junctions. Member of claudins family are not randomly distributed among all tissues rather they show cell-specificity; claudin-5, claudin-3 and claudin-12 are localized in BBB's tight junctions. Recently discovered endothelial cell-selective adhesion molecule (ESAM)-1 and JAM-A, member of Ig supergene family also comprises tight junctions but their contribution to it is still

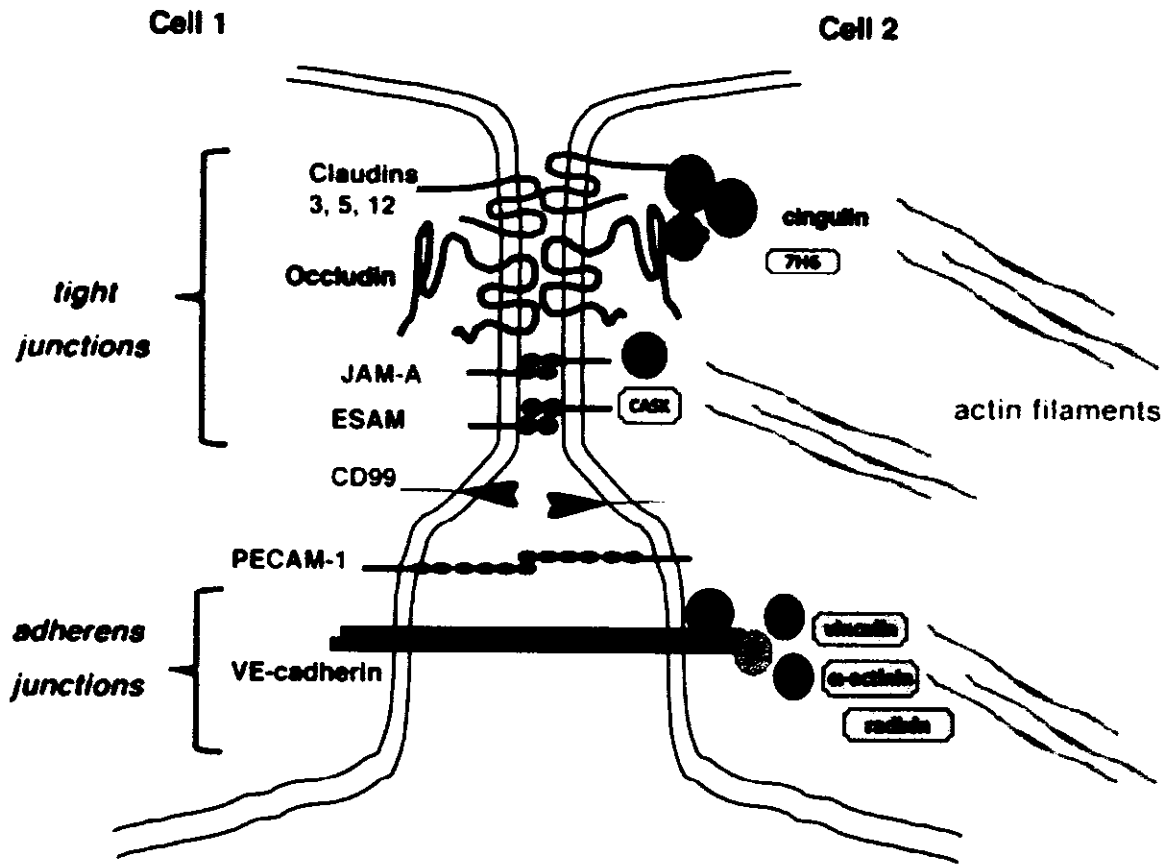


Fig.2.2 Organization of tight junctions, adherens junctions and outside of these junctions between two adjacent endothelial cells of blood brain barrier.

unknown. These integral protein are connected to actin filaments via cytoplasmic proteins zonula occludens (ZO)-1, ZO-2, cingulin, AF-6 and 7H6 (Engelhardt *et al.*, 2009).

2.6 HIV Neuroinvasion

Through this extensive network of transmembrane and cytoplasmic proteins, blood brain barrier secures the brain from organisms causing CNS infections and neurodegenerative diseases. So if an organism succeeds in causing any brain disease it has to adopt certain mechanisms to make neuroinvasion possible. Human immuno deficiency virus invades immune system during initial phase of infection and then it's presence is maintained throughout patient's life. A multistep model of HIV neuroinvasion is demonstrated in figure 2.3; digits are encircled to illustrate proposed multistep process of neuroinvasion by HIV.

(1) Expression of chemokine receptors on brain cells which includes soluble fracktaline(Fkn) and CXCL12, supports a slow but continuous entry of monocytes and macrophages into the CNS. Due to expression of CX_3 CR1, CD16 + activated monocytes, which are main reservoirs of HIV, are slowly and continuously crossing BBB and therefore considered as most prominent cell type involved in neuroinvasion by HIV to cause HIV-associated dementia.

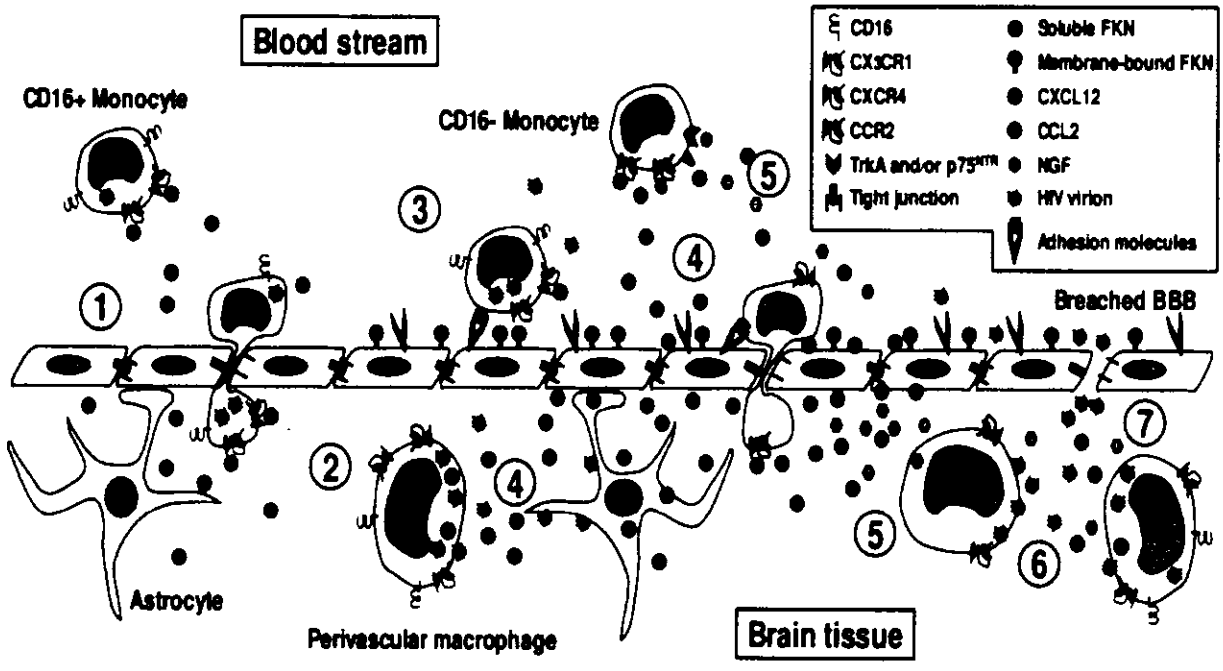


Fig.2.3 Mechanistic model for HIV neuroinvasion in humans.

(2) These infiltrated HIV infected monocytes produce HIV inside CNS along with inflammatory mediators in perivascular areas; in turn it activates neighboring astrocytes and BBB endothelium.

(3) In response to this monocyte conscription is enhanced.

(4) CCL2 which is over expressed by infected macrophages and astrocytes attract CD16 – and CCR2 + monocytes towards perivascular areas.

(5) Both nerve growth factor (NGF) and CXCL12 that are over expressed in brain with inflammatory responses, NGF increases CXCR4 expression and enhances attraction for uninfected monocytes by CXCL12.

(6) Locally produced HIV may infect activated but uninfected perivascular macrophages, amplifying the activation-attraction-infection cycle.

(7) HIV products as well as local inflammation induce the disruption of tight junctions in BBB, enhancing neuroinvasion by HIV (Gras *et al.*, 2010).

Human immunodeficiency virus is responsible for cognitive and behavioral dysfunctions collectively termed as HIV-1 associated dementia. These worst manifestations are caused by HIV-1 permeation, toxic virus products and blood derived activated macrophages passage across BBB. Disruption in the integrity of BBB is central to adverse pathological outcomes of HAD since this BBB's disruption is most common among patients with dementia when compared to patients having AIDS but no dementia (Engelhardt *et al.*, 2009). There exist strong evidences that show two HIV neurotoxic

viral proteins, i.e., Gp120 and Tat increases blood brain barrier's permeability for HIV while Tat protein has also been reported to bind with neurons to cause neuronal degeneration (Gray *et al.*, 2001).

2.7 Tight Junction Protein, Zonula Occluden-1

Recently it has been established that Gp120 protein is involved in degradation of zonula occluden-1 protein by poteasome mediated degradation while the underlying steps involves in proteasomal degradation are not elucidated. ZO-1 protein is considered to play central role in proper organization of tight junctions as it act as a connector between transmembrane proteins and actin cytoskeleton (Gray *et al.*,2001). Zonula occluden-1 is a member of membrane-associated guanylate kinase (MAGUK) homologue protein family (Nakamuta *et al.*, 2008)]. It is a multidomain protein as shown in figure 2.4. The N-terminal half of ZO-1 protein comprises of three PSD-95/ disc-large/Zonula occludens-1 (PDZ) domains, a Src homology-3 (SH3) domain and so-called unique motif lies left to a region of homology to guanylate kinase (GUK) domain. Many of cytosolic and transmembrane proteins localized to tight junction binds with protein binding motifs found in N-terminal. Like, claudins binds to first PDZ domain (PDZ-1), PDZ-3 domain binds to JAM proteins and occludins and tricellulin binds to U5+GUK domain. While the C-terminal region of protein binds to F-actin or with other actin-binding proteins, for example cortactin and protein 4.1. C-terminal region also contains the binding sites of AF-6 and cingulin (Nakamuta *et al.*, 2008).

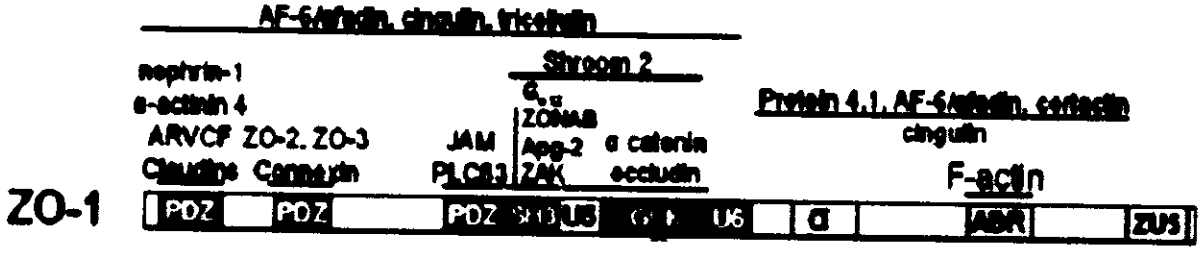


Fig.2.4 Domain organization of Zonula occluden-1 (ZO-1) protein. (On top are proteins which bind to protein binding motifs of respective domains)

2.8 Proteasome-mediated Degradation of APOBEC3G Protein

Since the mechanism involved in proteasome mediated degradation of ZO-1 protein is not known but, another HIV protein Vif mediates proteasomal degradation of human innate immune system protein apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3 (APOBEC3G) and it's mode of action is well elucidated (Alan *et al.*, 2009). Human APOBEC3G is an innate immune system protein which gets packed in HIV virion particles, causes deamination of cytosine to uridine and results in to loss of genetic integrity of HIV particle. HIV counters this innate immune system protein by Vif -mediated proteasomal degradation. The major steps followed are;

1. Hiv-1 Vif interacts with huAPOBEC3G. The residues of APOBEC3G required for interaction with Vif are: Aspartic acid (128), Proline (129) and Aspartic acid (130).
2. Thereby recruits a cullin 5-elongin B/C E3 ubiquitin ligase.
3. A3G is poly-ubiquitinated.
4. Finally targeted to proteasome for degradation.

Hence, HIV Vif protein ***marks*** human APOBEC3G for proteasomal degradation (Huthoff *et al.*, 2007). It is possible that Gp120 protein might also follow same path for proteasome-mediated degradation of ZO-1 protein.

2.9 Enzyme-linked immunosorbent assay (ELISA)

Enzyme-linked immunosorbent assay (ELISA) or enzyme immunoassay (EIA) is a biochemical technique intended to detect the presence of antibody or antigen in sample. Technique of EIA is based on three basic steps;

1. Antigen with unknown amount is affixed to some solid support.
2. Specific antibody is added, so that it binds to antigen.
3. And finally, to detect the level of binding some substance is added that the enzyme converts to detectable signal.

Since ELISA is performed to evaluate either presence of antigen or antibody in the sample therefore it is applied successfully for determining serum anitbody concentrations as with the HIV test. Food industry also shares application of ELISA mainly in detecting potential food allergens like milk, peanuts, walnuts, almonds, and eggs. EIA is also applied in toxicology as a rapid presumptive screen for certain drug classes (Lequin *et al.*, 2005). Because of it's high sensitivity ELISA was first test used for the HIV screening (Kamimura M *et al.*, 2009).

2.10 Neuron Specific Enolase

Enolase or phosphopyruvate dehydratase is a mettaloenzyme responsible for the catalysis of 2-phosphoglycerate (2-PG) to phosphoenolpyruvate (PEP) which is the ninth and penultimate step of glycolysis. Depending on environmental

concentrations of substrate, Enolase can also catalyse the reverse reaction. It is found in all organisms and tissues capable of glycolysis or fermentation, after 1934 when it was first discovered it has been isolated from variety of sources including human muscles and erythrocytes (Kustrzeba-Wojcicka *et al.*, 2000).

There exist three subunits of enolase, α, β, and γ, that combine to form five different isoenzymes including αα, αβ, αγ, ββ, and γγ (Piast *et al.*, 2005). Out of these, three are more commonly found more in adult human cells than the others. These are;

1. αα or non-neuronal enolase (NNE), it is isolated from variety of tissues including liver, brain, kidney, adipose and spleen. NNE is also termed as Enolase 1.
2. ββ or muscle specific enolase (MSE). As name indicates it is found only in muscles and also termed as enolase 3.
3. γγ or neuron-specific enolase (NSE). This type is found in abundance in neurons and cells of neuroendocrine origin. Also termed as enolase 2.

Since enolase 2 is specific to neurons and is measurable in blood and cerebrospinal fluid (Lee *et al.*, 2010) therefore, it is used as a biomarker of ischemic brain damage and has already been evaluated in traumatic brain injury, stroke and anoxic encephalopathy after cardiac arrest (Selakovic *et al.*, 2005). As in case of HIV-associated dementia neuro-degenaration also occurs involving damage of neurons as a result of which level of neuron specific enolase is increased in the blood

and cerebrospinal fluid. Therefore, NSE can be used as a biomarker to detect level of neuron damage in HIV & AIDS patients.

Till date no study has been published that reports the prevalence of HIV-associated dementia in HIV affected patients of Pakistani population. Studying prevalence of HAD in HIV & AIDS patients that are still not on ART, could be very important in determining whether the strain of HIV in Pakistan patients is strong enough to evade Blood Brain Barrier security or not, statistics would help in medication of patients and would help to follow mental health of patients.

MATERIALS & METHODS

3.1 Retrieving 3-D Structures and Sequences

Genbank available at National Center of Biotechnology Information (NCBI) was scanned to retrieve complete amino acid sequence of Zonula occluden-1 protein (also termed as tight junction protein), HIV-1 envelope glycoprotein Gp120, and V3 region of gp120 protein.

A Resource for Studying Biological Macromolecules (RCSB) which is a member of Protein Data Bank (PDB) that contains information about experimentally-determined structures of proteins, nucleic acids, and complex assemblies; was use to retrieve the available structures of Zonula occluden-1 protein, its PDZ-1, PDZ-2 and SH3-GUK domains, HIV-1 gp120 protein and that of V3 loop.

3.2 Basic Local Alignment

BLAST tool was used to scan whole Protein Data Bank (PDB) database via protein-protein BLAST (blastp) algorithm to retrieve a protein whose three dimensional structure is experimentally predicted and shows maximum homology with V3 region of viral gp120 protein.

3.3 3-Dimensional Structure Prediction

Modeller9v1 is software that made prediction of three-dimensional (3D) structure easy and relatively more reliable. It is based on homology or comparative modeling techniques of bioinformatics and involve following major steps;

1. An alignment file for query sequence is compiled and saved with *.ali* extension.
2. Python file of modeller named align2d is executed using *"modeller9v1 align2d.py"* command, in order to align query and template sequence.
3. On the basis of query-template sequence homology, for 3-D structure prediction *"model-single.py"* file is executed.
4. Best model is selected on basis of least objective function value.

3.4 3-D Structure Evaluation

Predicted structure is evaluated using two platforms:

3.4.1 Modeller9v1

Modeller9v1 provides the facility to evaluate stability of template and predicted structures. Since the 3-D structure of template sequence is experimentally predicted so, there was no requirement to evaluated it using any computational means.

Hence only the predicted structure with least objective function value is evaluated by executing modeller's python file *"evaluate-model.py"*

3.4.2 PROCHECK

PROCHECK is another web based server that provides the facility to evaluate stereochemical quality of a protein structure.

3.5 Analyzing 3-D Structures

Raswin provides both graphic user interface and command line window to explore different dimensions of a molecule. This standalone software was explored to analyze the docked structures, to roughly estimate distances thereby, for identifying interacting regions of two docked proteins. Predicted 3-D structures were also visualized via RASWIN.

3.6 Macro Molecular Docking

Quatenary structure of complexes is computationally modelled, formed by either two or more interacting biological macromolecules and this approach is termed as "macromolecular docking". Both protein-protein complexes and protein-nucleic acid complexes respectively are first and second most frequent attempted targets of such modelling. Macromolecular docking is a challenging domain of bioinformatics as proteins are complex macromolecule to be handled. However, there are certain programs

that provide macromolecular docking facility with certain limitations. Hex5.1 is available for 12 years in market with some important preferable features;

1. It is still the only docking and superpostion program to use spherical polar Fourier (SPF) correlations to accelerate the calculations.
2. Its still one of the few docking programs which has built-in graphics to view the results.
3. It is the first protein docking program to be able to use modern graphics processor units (GPUs) to accelerate the calculations.

It is reported to be used for establishing the molecular interaction between two macromolecules namely, Mycoplasma pneumoniea enolase and human plasminogen. Hex5.1 has a major limitation for docking macromolecules that the radial shape functions used in Hex decay exponentially beyond about 35 ° Angstroms from the chosen origin Hence it does not support efficiently the structures whose distance from the centroid is more than 35 ° angstrom. On following basis full length zonula occluden-1 and gp120 proteins were not docked.

1. No 3-dimensonal structure was available for both proteins in any of structural databases by using either of techniques, i.e., x-ray crystallography or NMR-studies.

2. Sequence of ZO-1 protein with swiss-prot accession number *Q07157.3* is as shown in table 3.1.

Tertiary structures of all domains except PDZ-3 are available but as obvious from above highlighted and non-highlighted regions comparison and general PDB BLAST of ZO-1 clearly depict that homology is not to the mark where homology modeling is possible and as domains are functionally independent so, domains of ZO-1 are docked individually.

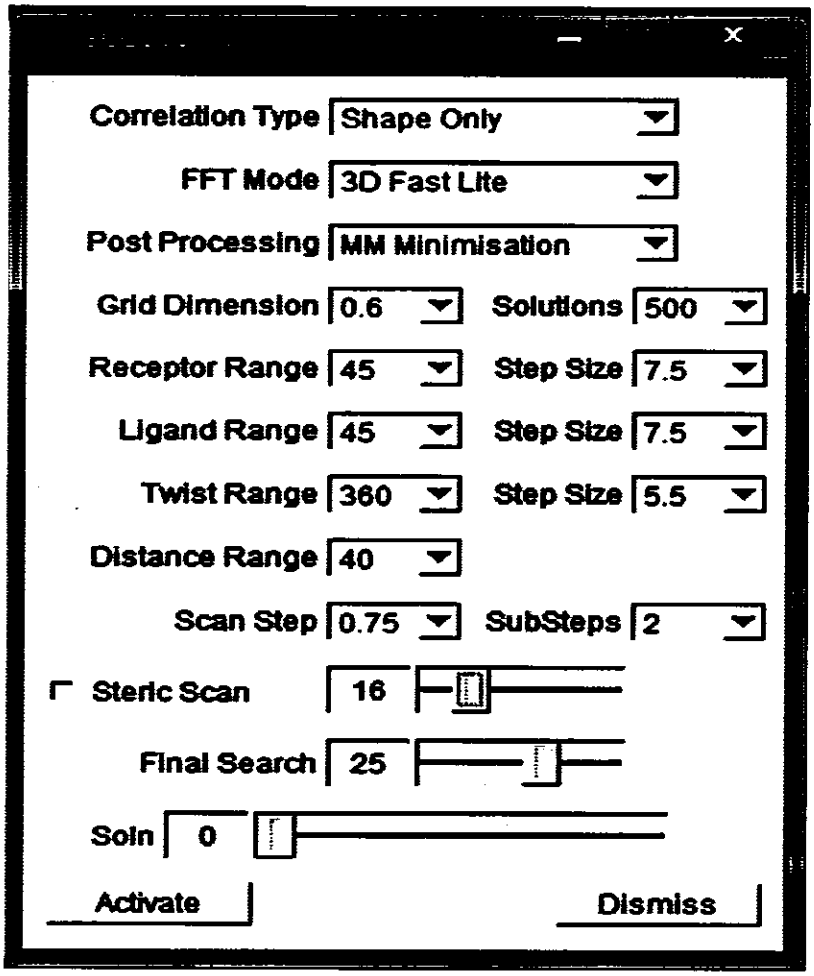
3. Structure of Gp120 protein is also not available and for our purpose it is not required as well since, residues in the V3 region or its immediate vicinity form the binding site for the gp120 co-receptors on T cells and macrophages. Also many antibodies recognize the V3 loop of gp120, which is a major neutralizing determinant of the virus. Therefore, each domain of ZO-1 is docked with V3 loop of gp120 protein.

Parameters used for docking via Hex are shown in table 3.2; these values are reported in manual to produce best docking results.

Table 3.1 Sequence of Zonula occluden-1 protein with accession no. *Q07157.3.* (in the table Yellow colored highlighted region is PDZ-1 domain, Turquoise colored highlighted region is PDZ-2, Green colored highlighted region is PDZ-3, Light gray colored highlighted region is SH3 and Magenta colored highlighted region is GUK domain.)

MSARAAAAKSTAMEETA[redacted]KVQIPVSRPDPEPVSDNEE
DSYDEEIHDPRSGRSGVVNRRSEKIWPRDRSASRERSLSPRSDRRSVASSQPAKPT[redacted]
[redacted]RATLLNVPDLSDSIHSANASERDDISEIQSLASDHSGRSHDRPPRRSRSRSPDQRSEPSDHSR
HSPQQPSNGSLRSRDEERISKPGAVSTPVKHADDHTPKTVEEVTVERNEKQTPSLPEPKPVYAQV
GQPDVDLPVSPSDGVLPNSTHEDGILRP[redacted]
[redacted]KKDVYRRIVESDV[redacted]
[redacted]SVQ
YTLPKTAGGDRADFWRFRGLRSSKRNLRKSREDLSAQPVQTKFPAYERVVLREAGFLRPVTIFGP
IADVAREKLAREEPDIYQIAKSEPRDAGTDQRSSGII[redacted]
[redacted]QQQNQLVWVSEGKADGATSDDLDLHDDRLSYLSAPGSEYSMYSTDSRHTSDYED
TDTEGGAYTDQELDETLNDEVGTPPESAITRSSEPVREDSSGMHHENQTYPPYSPQAQPQPIHRID
SPGFKPASQQKAEASSPVPYLSPETNPASSTSAVNHNVNLTNVRLEEPTPAPSTSYSPQADSLRTPS
TEAAHIMLRDQEPSLSSHVDPTKVYRKDPYPEEMMRQNHVLKQPAVSHPGHRPDKEPNLTYEPQ
LPYVEKQASRDLEQPTYRYESSSYTDQFSRNYEHRLRYEDRVPMYEEQWSYYDDKQPYPSRPPF
DNQHSQDLDSRQHPEESSERGYFPRFEEPAPLSYDSRPRYEQAPRASALRHEEQPAPGYDTHGRL
RPEAQPHPSAGPKPAESKQYFEQYSRSYEQVPPQGFTSRAGHFEPLHGAAAVPPLIPSSQHKPEAL
PSNTKPLPPPPTQTEEEEDPAMKPQSVLTRVKMFENKRSASLETKKDVNDTGSFKPPEVASKPSG
APIIGPKPTSQNQFSEHDKTLYRIPEPQKPQLKPPEDIVRSNHYDPEEDEEYYRKQLSYFDRRSFEN
KPPAHIAASHLSEPAKPAHSQNQSNFSSYSSKGKPPEADGVDRSFGEKRYEPIQATPPPPPLPSQYA
QPSQPVTSASLHIHSKGAHGEGNSVSLDFQNSLVSKPDPPPSQNKPATFRPPNREDTAQAAFYPQK
SFPDKAPVNGTEQTQKTVTPAYNRFTPKPYTSSARPFERKFESPKFNHNLLPSETAHKPDLSSKTP
TSPKTLVKSHSLAQPPEFDSGVETFSIHAEKPKYQINNISTVPKAIPVSPSAVEEDEDEDGHTVVAT
ARGIFNSNGGVLSSIETGVSIIIPQGAIPEGVEQEIYFKVCRDNSILPPLDKEKGETLLSPLVMCGPH
GLKFLKPVELRLPHCDPKTWQNKCLPGDPNYLVGANCVSVLIDHF

Table 3.2 Parameters used for docking.



3.7 Distance Calculation between Receptor and ligand Residues

VMD is a molecular visualization program for displaying, animating, and analyzing large biomolecular systems using 3-D graphics and built-in scripting, in our studies it was utilized;

1. To identify and select the residues within 5.0 ° angstrom of possible interacting regions within docked complex.
2. To calculate distances between possible interacting residues within 5.0 ° angstrom of apparently interacting regions of receptor (in our case any domain of ZO-1) identified via RASWIN.

3.8 Serum Collection

To study prevalence of HIV-1 associated dementia (HAD) in HIV infected patients of Pakistani population who are not on anti retroviral therapy (ART), samples of serum were collected from forty non-ART HIV patients of Pakistani population. Since HIV screening test in twin cities of Pakistan occurs in HIV referral laboratory of National AIDS Control Program (NACP) therefore, all samples were collected from NACP. In table 3.3 the fake ID of HIV patient (because of privacy perspectives) and their mode of transmission of HIV is mentioned, this is to make sure that samples from different group of HIV patients is collected.

Table 3.3 Fake IDs and mode of transmission of HIV in fourty non-ART HIV patients.

Patient ID	Mode of Transmission	Patient ID	Mode of Transmission
4	ID user + Sex history	24	ID user
5	ID user + Sex history	25	ID user
6	Blood transfusion	26	ID user
7	Not known	27	Blood transfusion
8	Blood transfusion	28	ID user + Sex history
9	ID user + Sex history	29	ID user + Sex history
10	Not known	30	ID user + Sex history
11	Not known	31	ID user + Sex history
12	Not known	32	ID user + Sex history
13	Not known	33	ID user + Sex history
14	Not known	34	ID user + Sex history
15	Not known	35	ID user + Sex history
16	ID user + Sex history	36	ID user + Sex history
17	ID user	37	Blood transfusion
18	ID user	38	Blood transfusion
19	Blood transfusion	39	Not known
20	ID user + Sex history	40	ID user + Sex history
21	ID user + Sex history	41	ID user + Sex history
22	ID user	42	ID user + Sex history
23	ID user + Sex history	43	ID user + Sex history

To verify the correct functioning of ELISA kit, three positive controls of non-HIV dementia patient's serum were collected with the courtesy of Doctor Mazhar badshah.

3.9 Enzyme-Linked Immunosorbent Assay (ELISA)

3.9.1 Preparation of Reagents

3.9.1.1 NSE Standards

0.75mL of distilled water was added to each vial and was mixed gently afterwards, it was allowed to reconstitute by allowing it to stand for 15 minutes.

3.9.1.2 Wash Solution

50mL of washed concentrate was poured in to the clean container and diluted 25 folds by adding 1200mL of distilled water to give a buffered wash solution.

3.9.1.3 Antibody Solution

Since numbers of strips used were 12 therefore, 600 micro litter of HRP anti-NSE tracer was mixed with 12mL 0f Biotin Anti-NSE

3.9.2 Assay Procedure

The neuron specific Enolase kit of DRG Company is imported by NACP Pakistan for quantitative determination of NSE in human serum. As mentioned in protocol of DRG NSE kit each serum sample was poured in two consecutive wells. To successfully determine the quantity of serum in HIV patients and in dementia patients most importantly following steps were followed;

1. 25 micro litter of NSE standards and HIV & non-HIV dementia patient's serum was poured into the strip wells in scheme shown in table 3.4.
2. 100 micro litter of antibody solution was added to each well using 100 micro litter precision pipette.
3. The plate was then incubated for 1 hour at room temperature with constant shaking of plate using microplate shaker.
4. After incubation each strip was aspirated and washed six times.
5. 100 micro litter of TMB HRP-Substrate solution was added to each well.
6. Afterwards, again incubated with constant shaking for 30 minutes at room temperature.
7. In last, for reading results 100 micro litter of stop solution was added , mixed and absorbance was read at 405nm in a microplate spectrophotometer.

Table 3.4 HIV ELISA-work sheet (Standards A to E were available with kit, while 1, 2 and 3 are ID's allotted to non-HIV dementia patient's serum and rests are fake IDs of non-ART HIV patients)

	1	2	3	4	5	6	7	8	9	10	11	12
A	Standard A	Standard E	4	8	12	16	20	24	28	32	36	40
B	Standard A	Standard E	4	8	12	16	20	24	28	32	36	40
C	Standard B	1	5	9	13	17	21	25	29	33	37	41
D	Standard B	1	5	9	13	17	21	25	29	33	37	41
E	Standard C	2	6	10	14	18	22	26	30	34	38	42
F	Standard C	2	6	10	14	18	22	26	30	34	38	42
G	Standard D	3	7	11	15	19	23	27	31	35	39	43
H	Standard D	3	7	11	15	19	23	27	31	35	39	43

EIA KIT Vironostika HIV Ag/Ab **PLATE NO.:** 1

DATE: 16-8-2010 **KIT LOT NO.:** A60MA

RESULTS

4.1 Sequences and Three Dimensional Structures

Searching Genbank for human Zonula occluden-1 (tight junction1) protein retrieved 1,754 amino acids long protein with Swiss-Prot accession number **Q07157.3**, search of PDB database retrieved null result for complete structure of ZO-1. 3-Dimensional structure prediction is possible via exploring structural bioinformatics but it has certain limitations applicable in predicting structure for ZO-1 like;

1. Ab-inito 3-D structure prediction which can be done from scratch does not provide very reliable results and for such a long protein this is not a preferred option.
2. Next option could be a homology modeling using several web-based servers and standalone softwares, this requires a template whose 3-D structure is predicted with certain minimum query coverage and homology with query sequence. To proceed for this template searching PDB BLAST was performed; top results are shown in figure 4.1.

Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value	Links
2UF5_A	Chain A, Crystal Structure Of The Sh3-Guanylate Kinase Core Dom	515	516	16%	4e-146	S
2KT_A	Chain A, Crystal Structure Of The Sh3-Gmase Fragment Of Tight Jun	334	334	16%	2e-82	S
2H3C_A	Chain A, Crystal Structure Of Zo-1 Pdz1 Bound To A Phage-Derivec	226	228	9%	8e-50	S
2H3B_A	Chain A, Crystal Structure Of Zo-1 Pdz1 Bound To A Phage-Derivec	225	226	9%	2e-49	S
2H2W_A	Chain A, Crystal Structure Of Zo-1 Pdz1	221	222	9%	1e-48	S
3CO1_A	Chain A, The Crystal Structure Of Zo-1 Pdz2 In Complex With The (	174	174	9%	3e-43	S
2DM5_A	Chain A, Solution Structure Of The Second Pdz Domain From Huma	137	137	8%	2e-38	S
2RCZ_A	Chain A, Structure Of The Second Pdz Domain Of Zo-1 >pdb|2RCZ	136	136	8%	6e-38	S
2CSI_A	Chain A, Solution Structure Of N-Terminal Pdz Domain From Mouse	154	154	8%	5e-37	S
3E17_A	Chain A, Crystal Structure Of The Second Pdz Domain From Human	118	118	4%	2e-26	S
2DSG_A	Chain A, Solution Structure And Binding Property Of The Domain- S	118	118	4%	2e-26	S
1DXM_A	Chain A, Crystal Structure Of The Gmp Bound Sh3-Hook-Gk Fragm	92.4	92.4	16%	1e-18	S
1KJA_A	Chain A, Sh3-Guanylate Kinase Module From Psd-95	90.9	90.9	16%	4e-18	S
3GSS_A	Chain A, The Structure Of Unc5b Cytoplasmic Domain	72.8	72.8	4%	1e-12	S
1QLC_A	Chain A, Pdz1 Domain Of Synapse Associated Protein 97	93.2	93.2	8%	2e-09	S
2KA9_A	Chain A, Solution Structure Of Psd-95 Pdz12 Complexed With Cypr	147	147	8%	6e-09	S
1BE9_A	Chain A, Cyclic Peptides Targeting Pdz Domains Of Psd-95: Structur	60.5	60.5	5%	7e-09	S
1QEF_A	Chain A, Pdz1 Of Sap90	60.5	60.5	5%	7e-09	S
1TLP_A	Chain A, The First Pdz Domain Of Psd-95 >pdb|1TU2|A Chain A Thr	58.9	58.9	5%	2e-08	S
1UJT_A	Chain A, Crystal Structure Of The Fsd93 Pdz1 Domain	93.2	93.2	8%	4e-08	S
1TP5_A	Chain A, Pdz3 Domain Of Psd-95 Protein Complexed With Kkeqtpv Pe	124	124	10%	1e-07	S
1TLH_A	Chain A, Crystal Structure Of The Third Pdz Domain Of Psd-95 >pdb	123	123	10%	1e-07	S
1BE9_A	Chain A, The Third Pdz Domain From The Synaptic Protein Psd-95 I	123	123	10%	1e-07	S
1D5T_A	Chain A, Solution Structures Of The Pdz Domain Of Human Unrarre	89.7	89.7	8%	2e-07	S

Fig. 4.1 PDB BLAST results for complete Zonula occluden-1 protein structure.

Maximum coverage of query sequence is 16% which is less for template used for reliable 3-D structure prediction. Hence homology modeling with single template is not a suitable option.

3. However there is another option of homology modeling using multiple templates covering different portions of query sequence, but as shown in figure 4.2 query sequence coverage is only for N-terminal region for residues less than 800, for rest of 800 residues ab-initio modelling is not an appropriate choice.

Domains inside protein are functionally independent, i.e., they are able to perform their function independent of remaining protein part, so PDB protein database was then scanned for individual domains of ZO-1. As shown clearly in figure 4.2 that 3-D structure for C-terminal domains are not predicted by either way, i.e., NMR studies or X-ray crystallography so search was made only for five domains namely; PDZ-1, PDZ-2, PDZ-3, SH3 and GUK domains. It is reported that SH3 and GUK domains, a conserved feature of MAGUK family, are dependent on each other for their correct folding so they are treated as single module instead of two separate domains (Lye *et al.*, 2010). Table 4.1 shows accession numbers of ZO-1's available structure of domains.

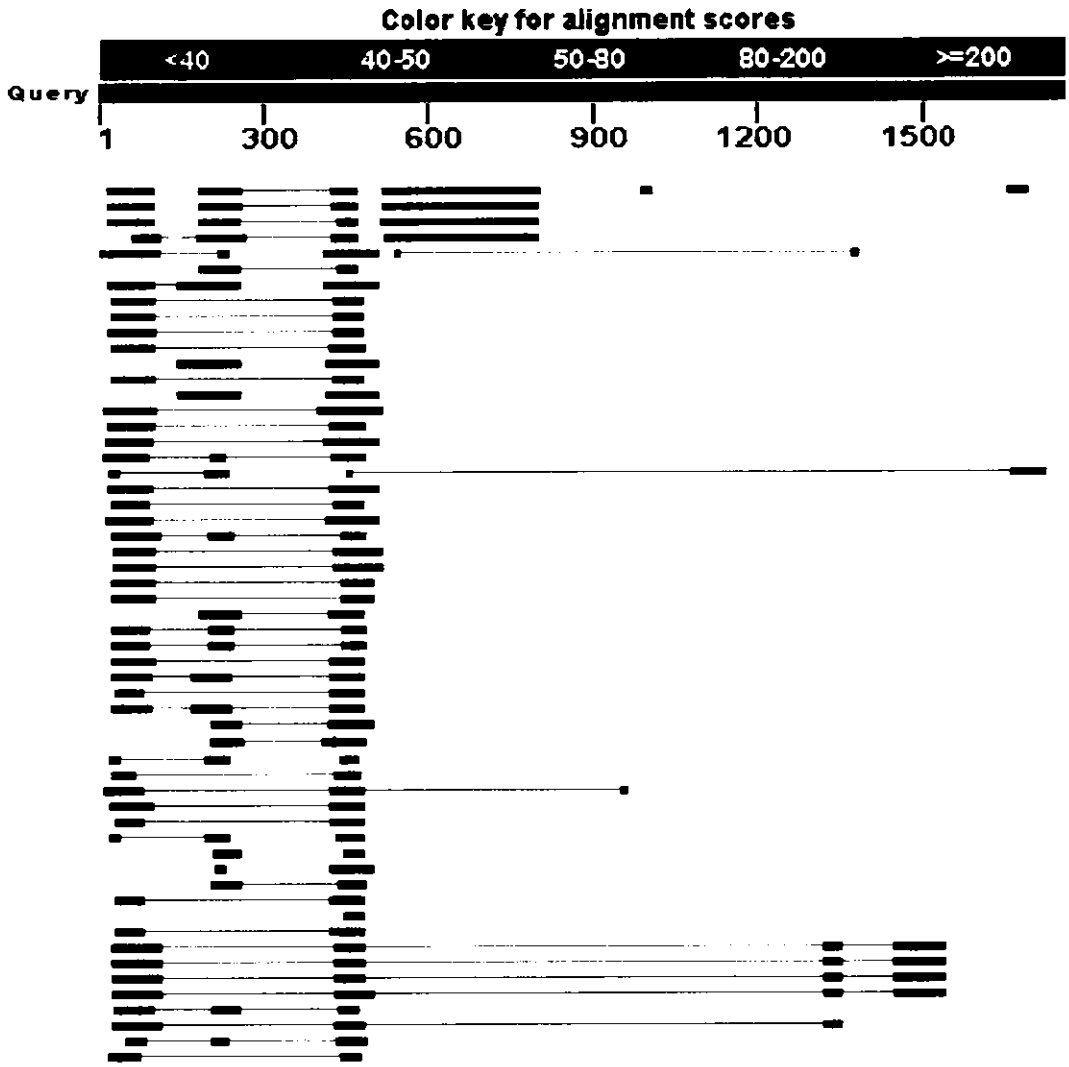


Fig. 4.2 Results for protein sequence BLAST for Zonula occluden-1 protein.

Table 4.1 Names and PDB-ID of ZO-1 protein domains and V3 loop of viral envelop protein gp120.

Domain Name	**PDB-ID**
PDZ-1	2H3M
PDZ-2	2RCZ
SH3-GUK module	3LHS
Consensus V3 loop of Gp120 protein	1CE4

Three dimensional structure of PDZ-3 domain is not predicted so efforts were made to predict it via computational means. Basic local alignment via scanning PDB protein database for PDZ-3 domain sequence retrieved 104 BLAST hits, out of which top hits are shown in figure 4.3. First hit showed significant query coverage of 79% with minimum E-value therefore, selected as a template for 3-dimensional structure prediction.

Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value	Links
1MFG_A	Chain A, Solution Structure Of The Pdz Domain From Riken Cdna 2	52.0	52.0	79%	6e-08	S
1TP3_A	Chain A, Pdz3 Domain Of Psd-95 Protein Complexed With Kketpv P	49.7	49.7	95%	3e-07	S
1UEL_A	Chain A, Solution Structure Of The Second Pdz Domain Of Human K	47.8	47.8	70%	1e-06	S
1BE9_A	Chain A, Crystal Structure Of The Third Pdz Domain Of Psd-95	47.0	47.0	96%	2e-06	S
1TP4_A	Chain A, Crystal Structure Of The Third Pdz Domain Of Psd-95 >pdb	47.0	47.0	96%	2e-06	S
2DM9_A	Chain A, Solution Structure Of The Eighth Pdz Domain Of Human In	46.2	46.2	84%	4e-06	S
1HE9_A	Chain A, Crystal Structure Of The 3rd Pdz Domain Of Human Discs	45.4	45.4	96%	6e-06	S
1UEZ_A	Chain A, Solution Structure Of The First Pdz Domain Of Human Kiaa	44.7	44.7	80%	1e-05	S
2EH6_A	Chain A, Solution Structure Of The Sixth Pdz Domain Of Human Ina	44.3	44.3	70%	1e-05	S
1WHT_A	Chain A, Solution Structure Of The Third Pdz Domain Of Syntrophin- A	44.3	44.3	96%	1e-05	S
2PDF_A	Chain A, The Crystal Structure Of The 7th Pdz Domain Of Mpdz (Mu	43.9	43.9	67%	2e-05	S
2DM0_A	Chain A, 7th Pdz Domain Of Multiple Pdz Domain Protein Mpdz	43.9	43.9	67%	2e-05	S
2AEG_A	Chain A, Solution Structure Of Harmonin Pdz2 In Complex With The	43.1	43.1	50%	3e-05	S
1WEI_A	Chain A, Solution Structure Of Rsgi Ruh-006, The Third Pdz Domain	43.1	43.1	96%	3e-05	S
2DEN_A	Chain A, Solution Structure Of The Second Pdz Domain Of Harmonin	43.1	43.1	53%	3e-05	S
2AHF_A	Chain A, Crystal Structure Of The First Pdz Domain Of Human Scrib	42.7	42.7	80%	4e-05	S
2A5Q_A	Chain A, Solution Structure Of The First Pdz Domain Of Scribble Hom	42.7	42.7	80%	4e-05	S
2DPG_A	Chain A, The Crystal Structure Of The 10th Pdz Domain Of Mpdz >pdb	42.4	42.4	95%	5e-05	S
2ZXK_A	Chain A, Crystal Structure Of Pdz Domain Of Synaptojanin-2 Bindin	42.0	42.0	73%	6e-05	S
2ZXN_A	Chain A, Crystal Structure Of Pdz Domain Of Synaptojanin-2 Bindin	42.0	42.0	53%	7e-05	S
2ENC_A	Chain A, Solution Structure Of The Pdz Domain From Human Synap	41.6	41.6	53%	8e-05	S
3EPL_A	Chain A, Crystal Structure Of The 3rd Pdz Domain Of Human Memb	41.2	41.2	60%	1e-04	S
2ZTZ_A	Chain A, Crystal Structure Of The Sap97 Pdz2 I342w C378a Mutant	41.2	41.2	80%	1e-04	S
2A3E_A	Chain A, Solution Structure Of The Second Pdz Domain From Huma	40.8	40.8	58%	2e-04	S
2OQS_A	Chain A, Structure Of The Hdlg/SAP97 PDZ2 IN COMPLEX WITH HPV	40.8	40.8	80%	2e-04	S
1WHC_A	Chain A, Solution Structure Of The Fourth Pdz Domain Of Kiaa1095	40.4	40.4	80%	2e-04	S
2RCZ_A	Chain A, Structure Of The Second Pdz Domain Of Zo-1 >pdb|2RCZ	40.4	40.4	58%	2e-04	S

Fig. 4.3 Three dimensional structure alignment results for PDZ-3 domain of ZO-1 protein.

4.2 Predicted Three Dimensional Structures

Chain A, Solution Structure of the PDZ domain from Riken Cdna (PDB ID =1WI2) which shares maximum of 42% identity with PDZ-3 domain of Zonula occluden-1 protein is used as template for homology modeling using modeller9v1. Ten protein structures with different conformation and objective function values were generated, predicted 3-dimensional structure of PDZ-3 domain with least objective function value is shown in figure 4.4.

Two peptides in V3 region (with Genbank accession number AAB05002.1) nearby V3 loop were subjected to homology modeling as well due to their possible binding with PDZ-1 domain. From total ten structures for each peptide, final model selection was again based on least objective function value. Figure 4.5 and figure 4.6 shows predicted 3- dimensional structure of first and second peptide respectively.

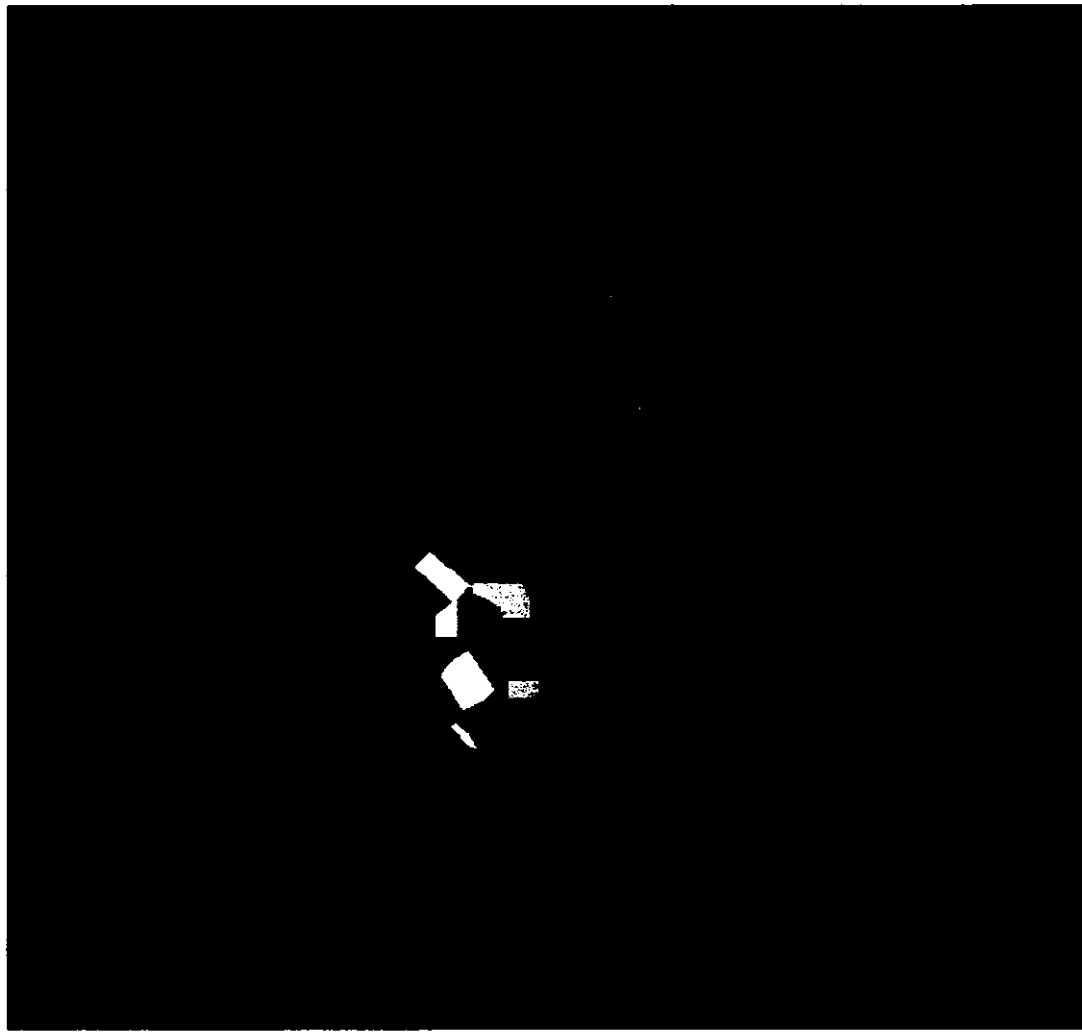


Fig. 4.4 Three dimensional structure of PDZ-3 domain with least objective function value

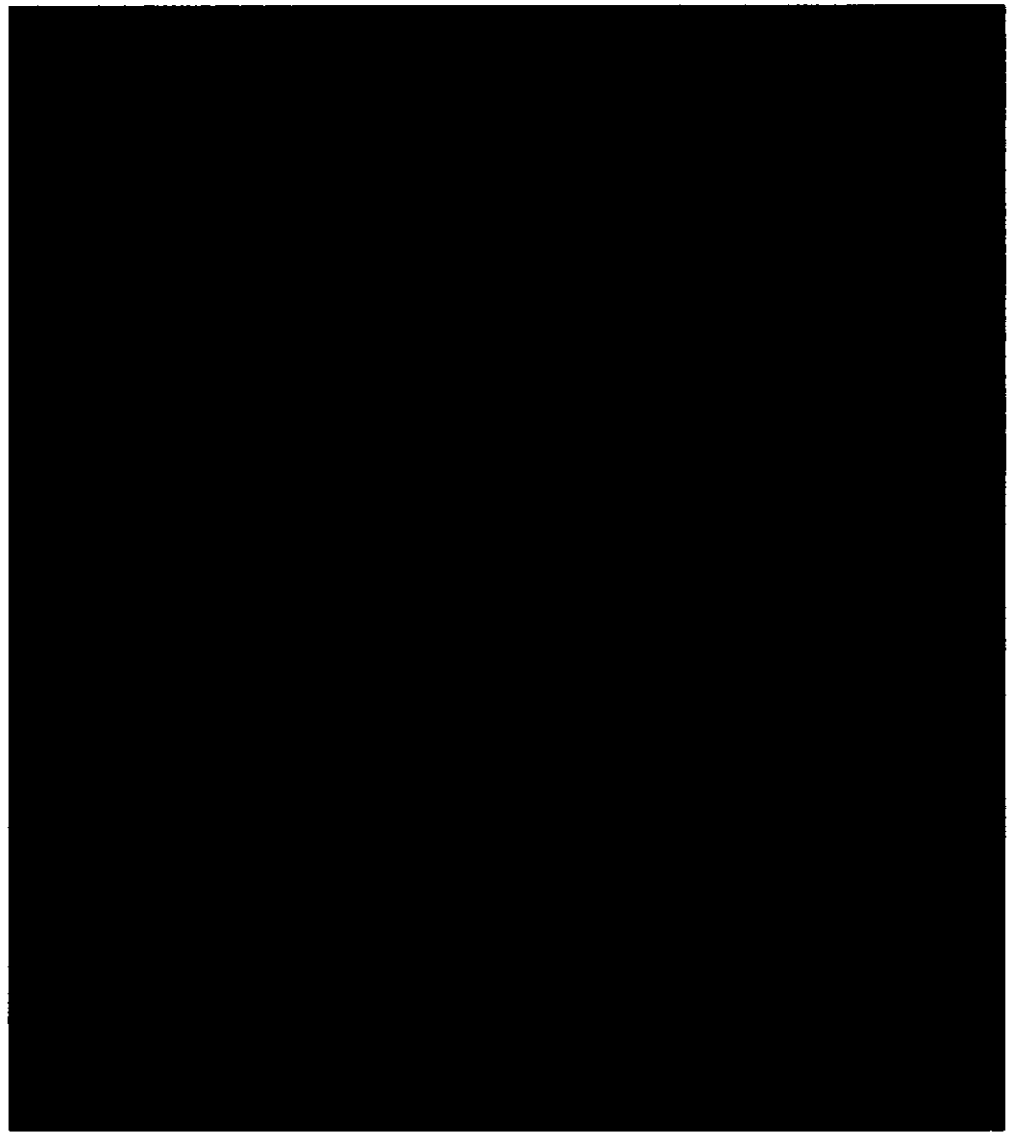


Fig. 4.5 Predicted three dimensional structure of first highlighted peptide.

Fig. 4.6 Predicted three dimensional structure of second highlighted peptide.

4.3 Evaluation of 3-Dimensional Structures

4.3.1 Modeller9v1 Evaluation

Modeller9v1's provided facility to evaluate quality of 3-D structures was utilized; all three predicted 3-D structures that include ZO-1's PDZ-3 domain, first highlighted peptide and second highlighted peptide of V3 region of viral envelope protein Gp120 were evaluated using *"evaluate-model.py"*.

The sum of all numbers for energy values of each residue of PDZ-3 domain, first highlighted peptide and second highlighted peptide are -2.8457, -1.2287 and -1.4577 respectively. This shows that all predicted three structures are stable.

4.3.2 Ramachandaran Plot

Procheck web based server was used to build ramachandaran plots for all three predicted 3-D structures, as shown in figures (4.7, 4.8 and 4.9). In figure 4.7 Plot statistics showed that 85.5% of residues lied in most favoured region, although best results are with more than 90% of residues within most favoured region but as 0% residues lied in disallowed region therefore, this structure is taken as acceptable structure for further processing. As shown in both figures 4.8 & 4.9 100% residues lies in most favoured region, this proves the credibility of both structures.

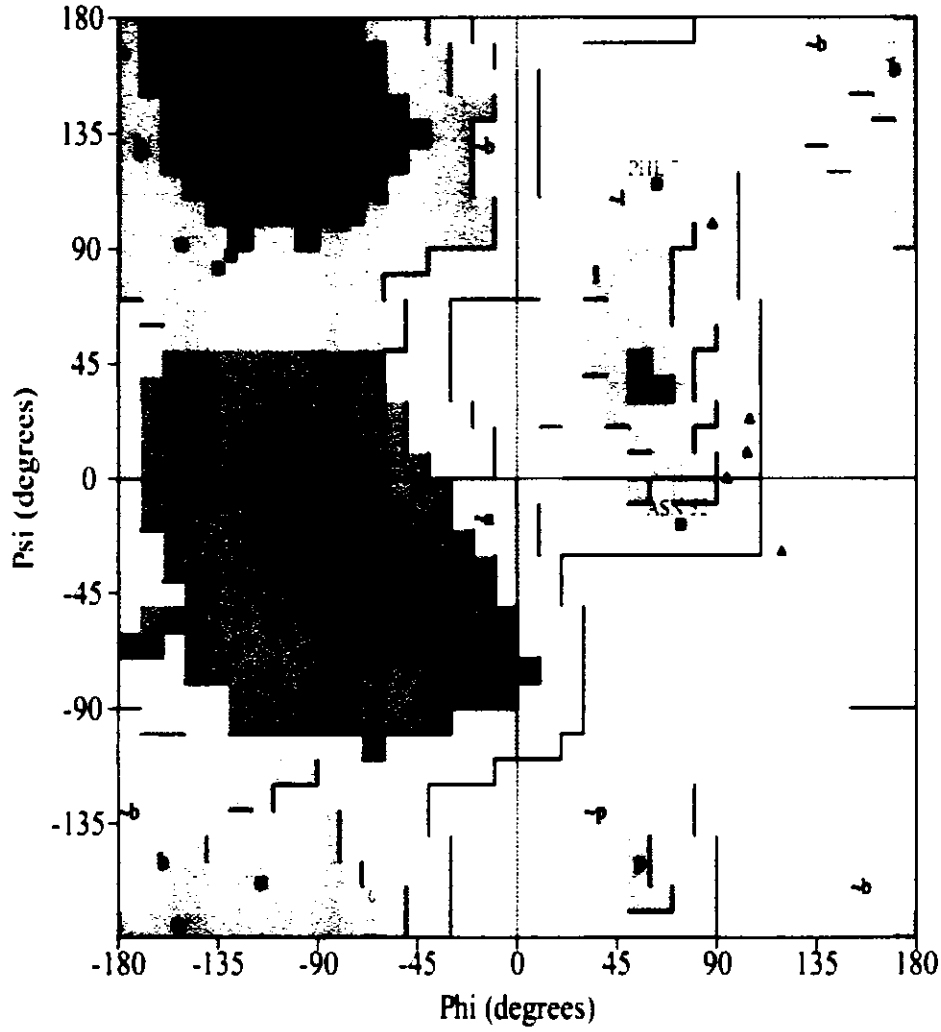


Fig. 4.7 Ramachandaran plot for predicted 3-D structure of PDZ-3 domain.

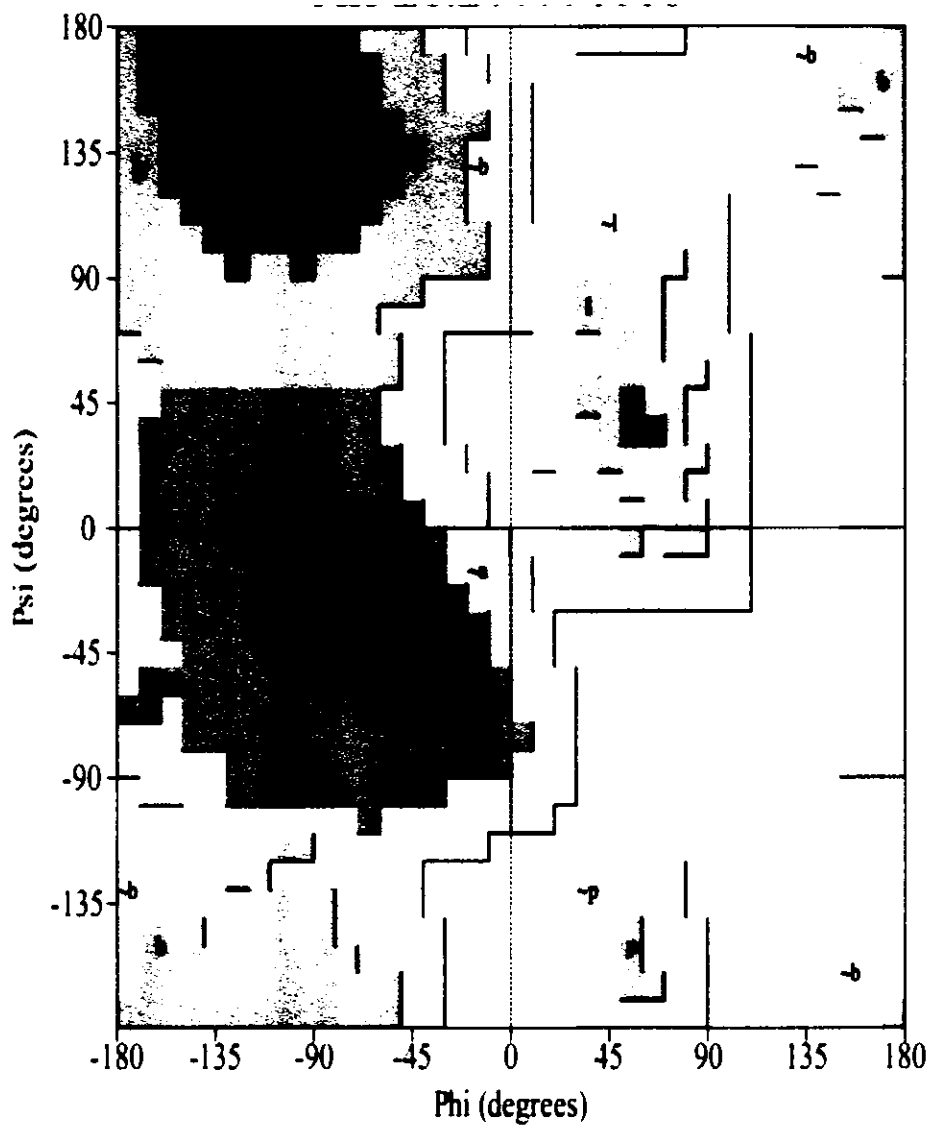


Fig. 4.8 Ramachandaran plot for first peptide of V3 region.

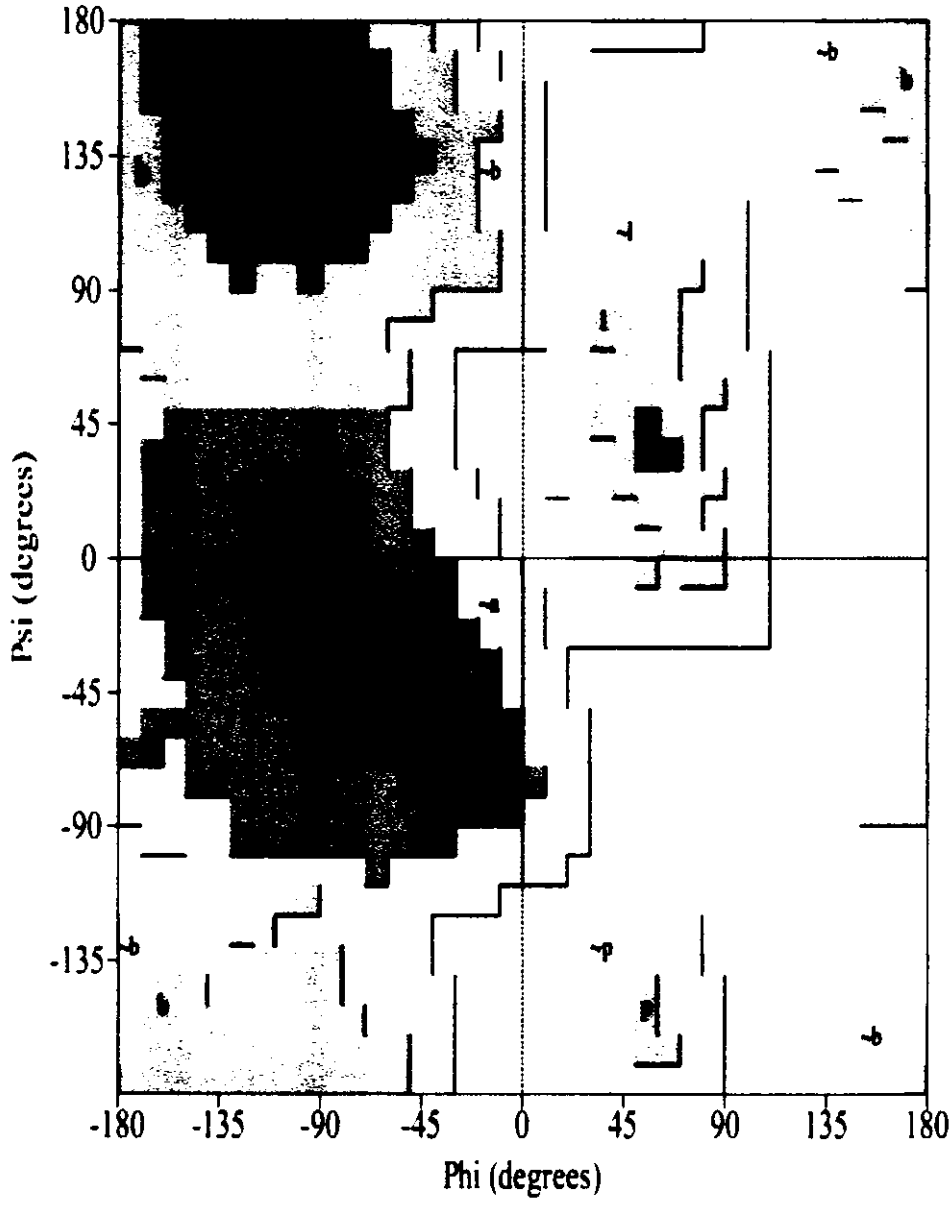


Fig. 4.9 Ramachandaran plot for second peptide of V3 region.

4.4 Macro Molecular Docking

Before docking two important points were taken under consideration;

1. Since Hex 5.1 do not perform well for structures that shares distance from their centroid more than 35 angstrom, so for each structure for which apparently there could be the possibility of greater size, distance from centroid was measured using VMD. Figures (4.10, 4.11, 4.12 and 4.13) shows distance from centroid of domains PDZ-1, PDZ-2, PDZ-3 and SH3-GUK module respectively.

As shown in figures (4.10, 4.11, 4.12, 4.13) only SH3-GUK modules exceeds limit of 35 angstrom, to resolve this issue docking was done twice using centroid of each domain, i.e., SH3 and GUK but it produced same docking results each time so for further work only single docked structure was used.

Since V3 loop is only 19 residues long so there is no possibility of it's structure exceeding the optimum distance length of 35 angstrom, same as case with first and second highlighted peptide of V3 region of gp120 protein where each is 21 residues long hence, their distances from centroid were not calculated.

2. PDB structures of PDZ-2 and SH3-GUK contains HOH hetero atoms in their PDB file with accession numbers "2rcz" and "3lh5" respectively. For docking

purposes these hetero atoms were removed manually from PDB files of each domain.

5.1 version of Hex standalone software was explored for docking domains without hetero atoms with consensus V3 loop of Gp120 protein. One hundred results were generated for each docking result, out of which result with minimum energy value was selected. Docking outputs with minimum energy values are shown in figures 4.14, 4.15, 4.16 and 4.17.

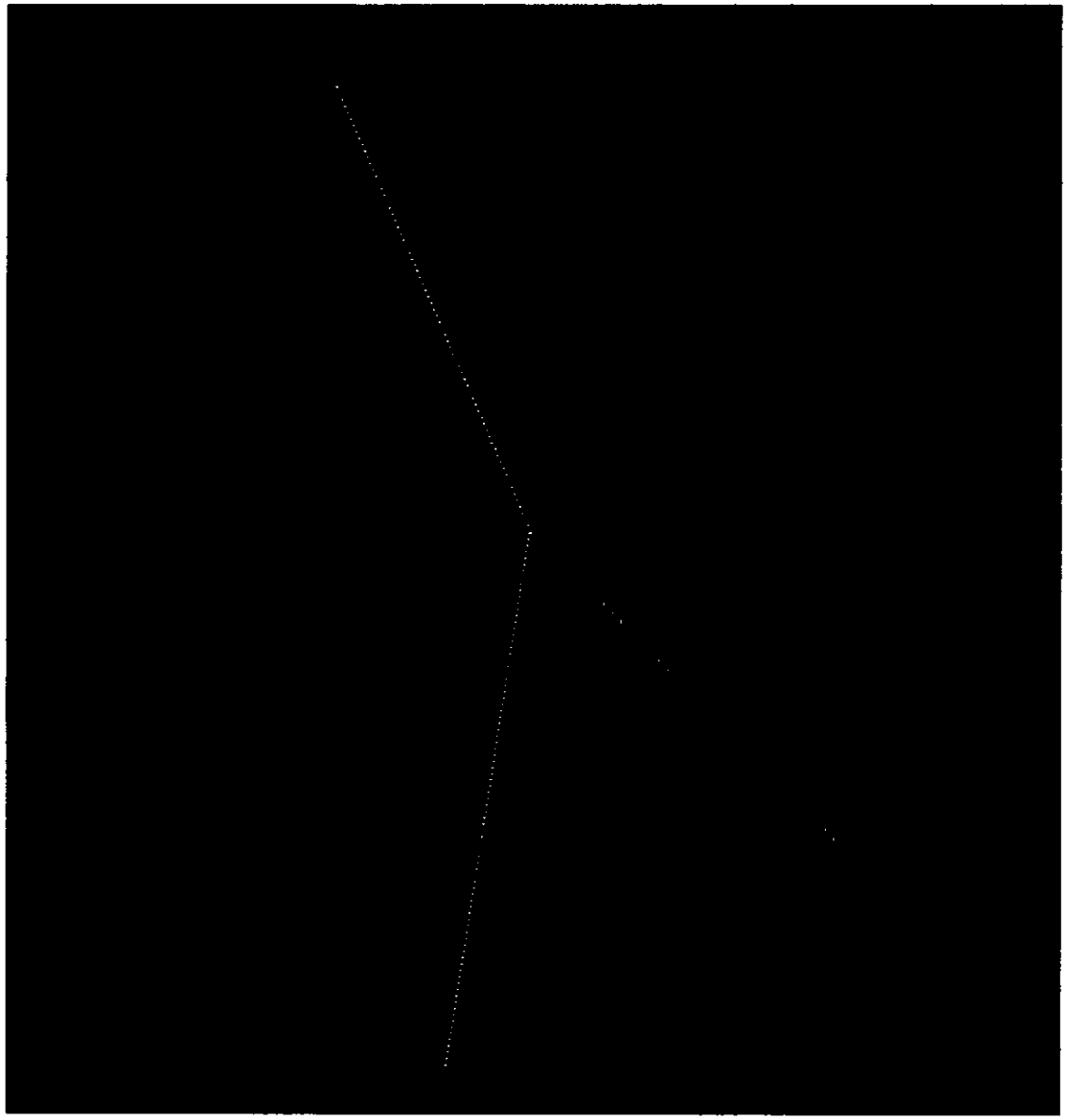


Fig. 4.10 Distances from centroid for PDZ-1 domain.

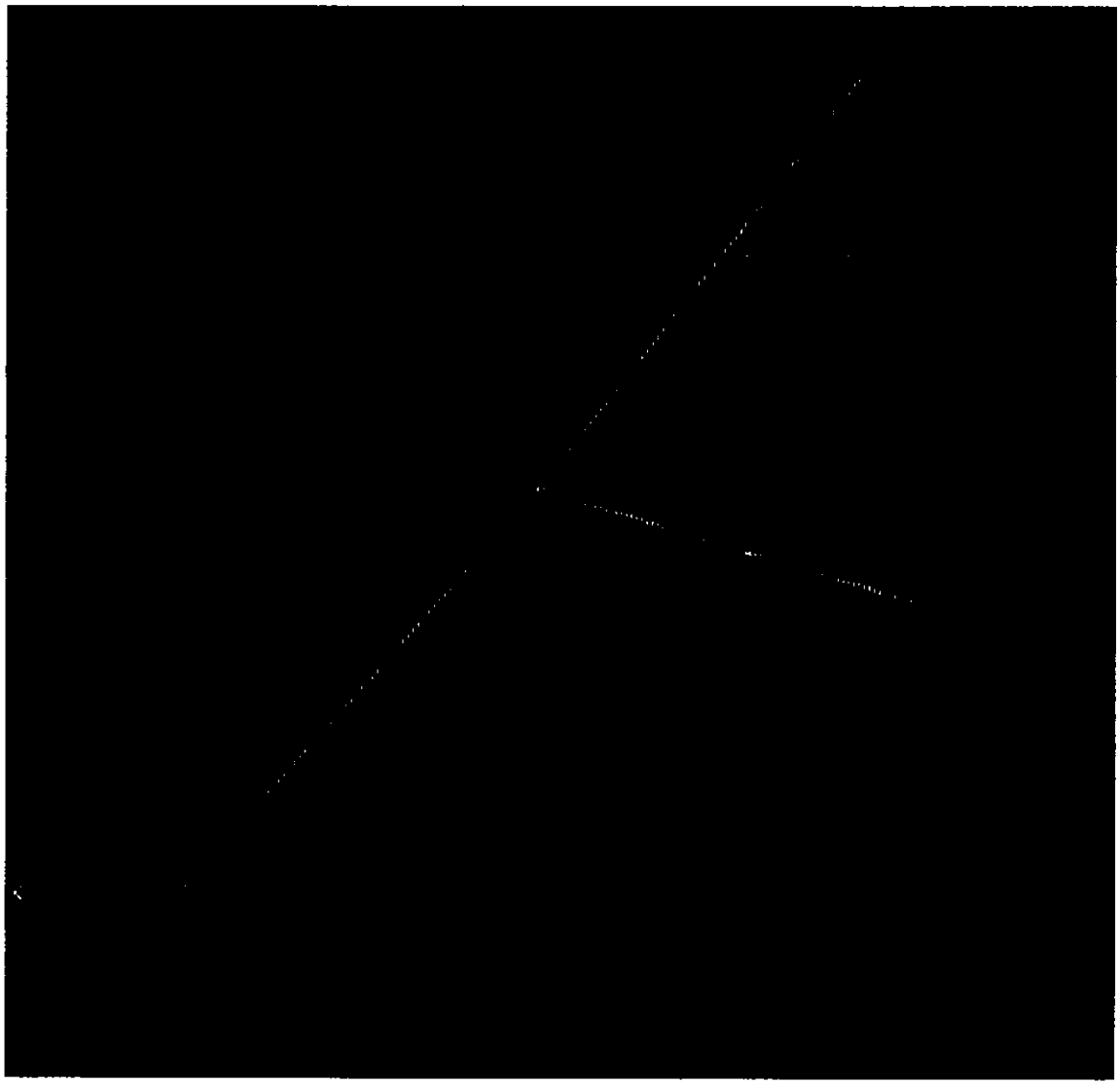


Fig. 4.11 Distances from the centroid of PDZ-2 domain, showing all distances within acceptable range of 35 angstrom.

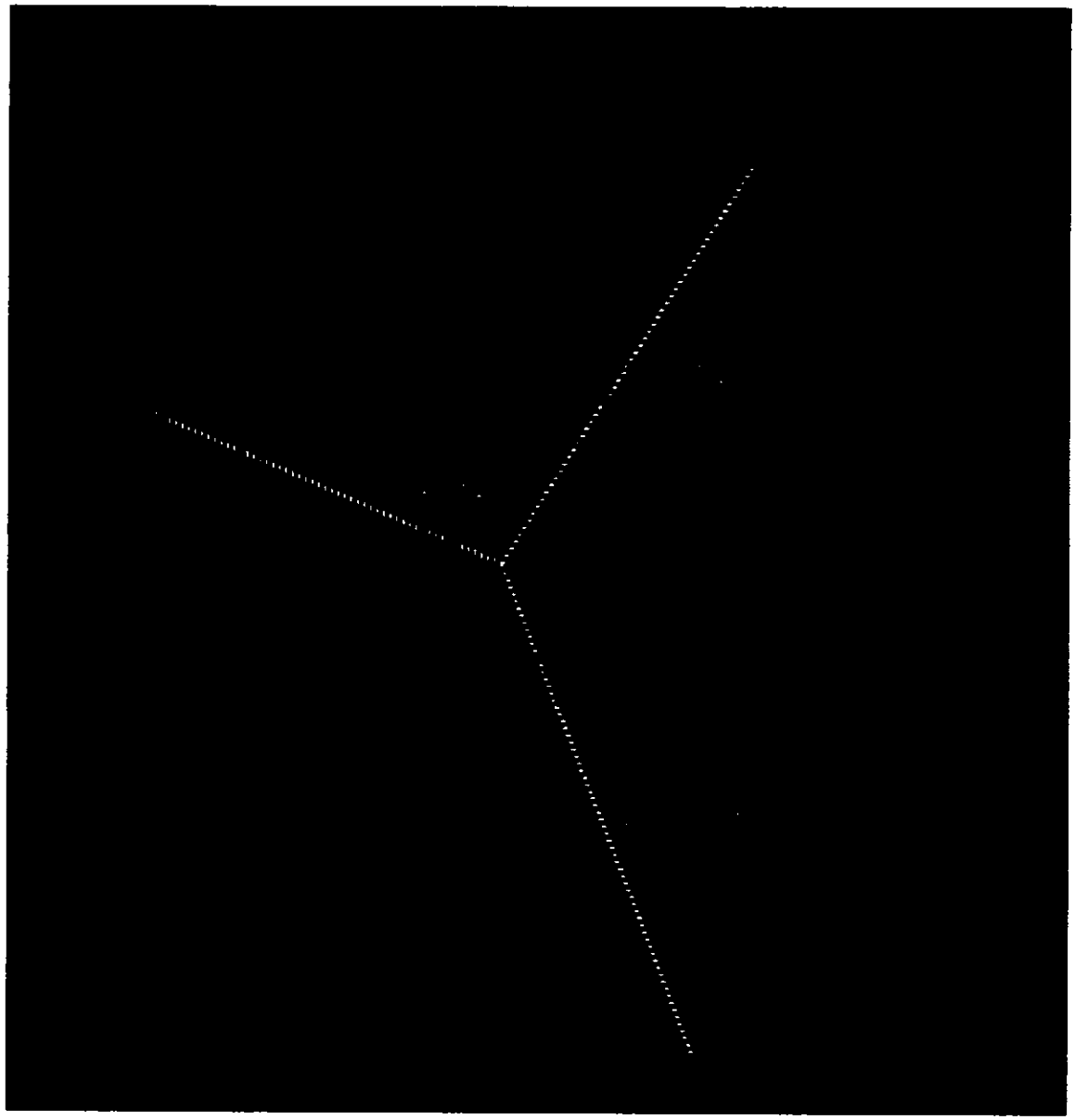


Fig. 4.12 Distances calculated via VMD from centroid of PDZ-3 domain and results shows all distance within 22 angstrom.

Fig. 4.13 Distances from the centroid of SH3-GUK module shows excision from the limit of 35 angstrom.

Fig. 4.14 Docking output of PDZ-1domain with 1ce4.

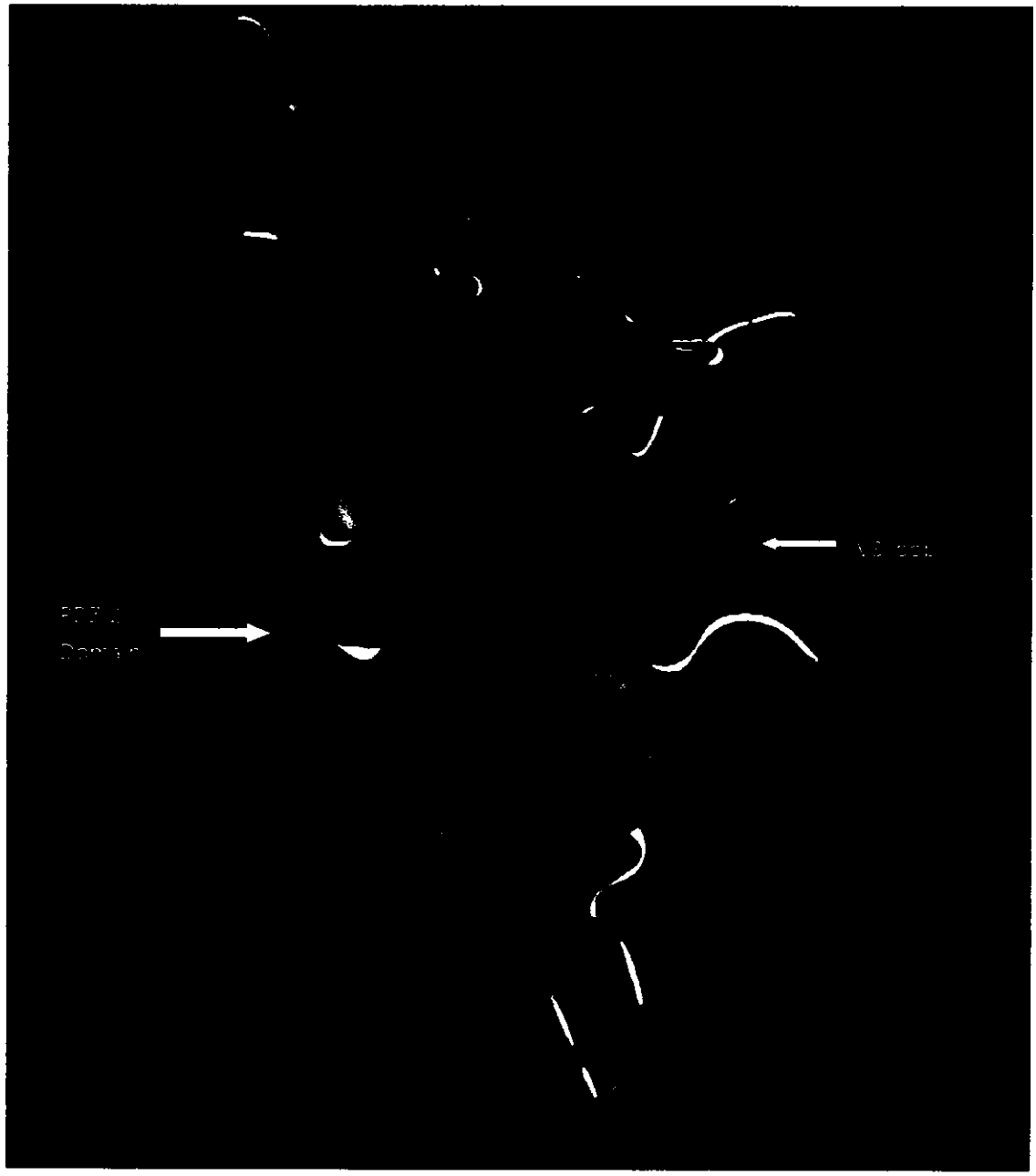


Fig. 4.15 Docking result of 2rcz with 1ce4 loop.

Fig. 4.16 Docking yield of PDZ-3 domain with 1ce4.

Fig. 4.17 Docking of SH3-GUK module with V3 loop of Gp120 protein.

Output of all work done to identify possible interacting domain showed PDZ-1 domain being very critical for interacting with 1CE4, therefore binding profile of PDZ-1 was subjected to critical analysis.

The binding profile for ZO-1 PDZ-1 is (Appleton *et al.*, 2006):

[R/K/S/T][T/S][W/Y][V/I/L]COOH

V3 region of gp120 and whole gp160 precursor protein (accession number AAK09421) with length of 790 residues were scanned maually to find any binding profile for PDZ-1. No appropriate result appeared in either case however two binding motifs of PDZ-1 domain with difference at third position (numbering 1 to left reidue) were observed at two locations in V3 region of gp20 protein. Total twenty one residue long peptide, ten from left and seven from right of this binding motif were selected for docking purposes in order to verify if region nearby V3 loop is involved in binding with PDZ-1.

VV[redacted]AINCTRPNNN[redacted]IGNIRQ

AHCNLSRTKWENT[redacted]

Red highlighted residues are of V3 loop, gray are 21 long peptides with yellow colored residues are binding motifs of PDZ-1.

Docking results for first and second highlighted peptides with PDZ-1 domain are shown in figure 4.18 and 4.19.

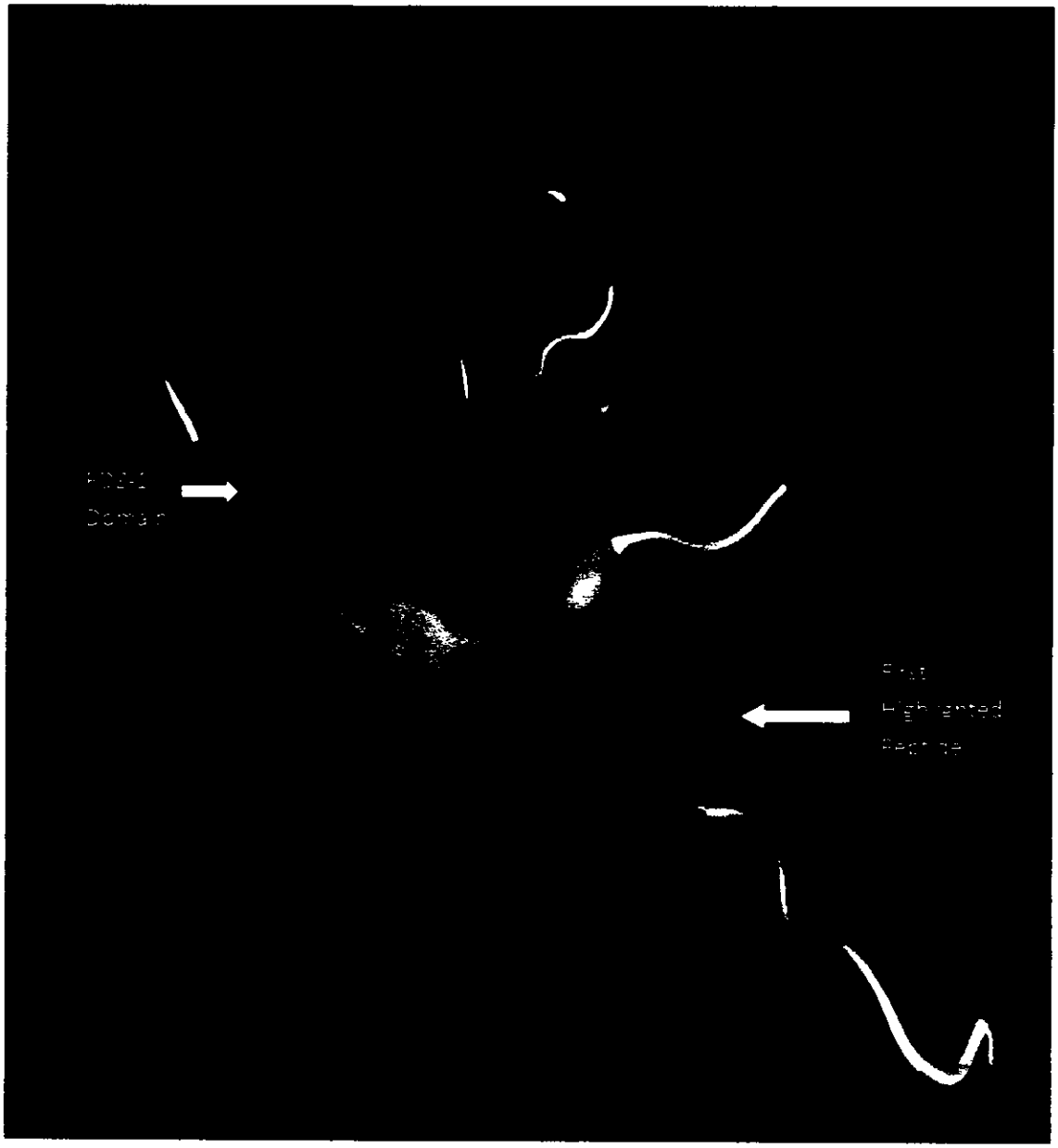


Fig. 4.18 Docking result of first highlighted peptide with PDZ-1 domain.

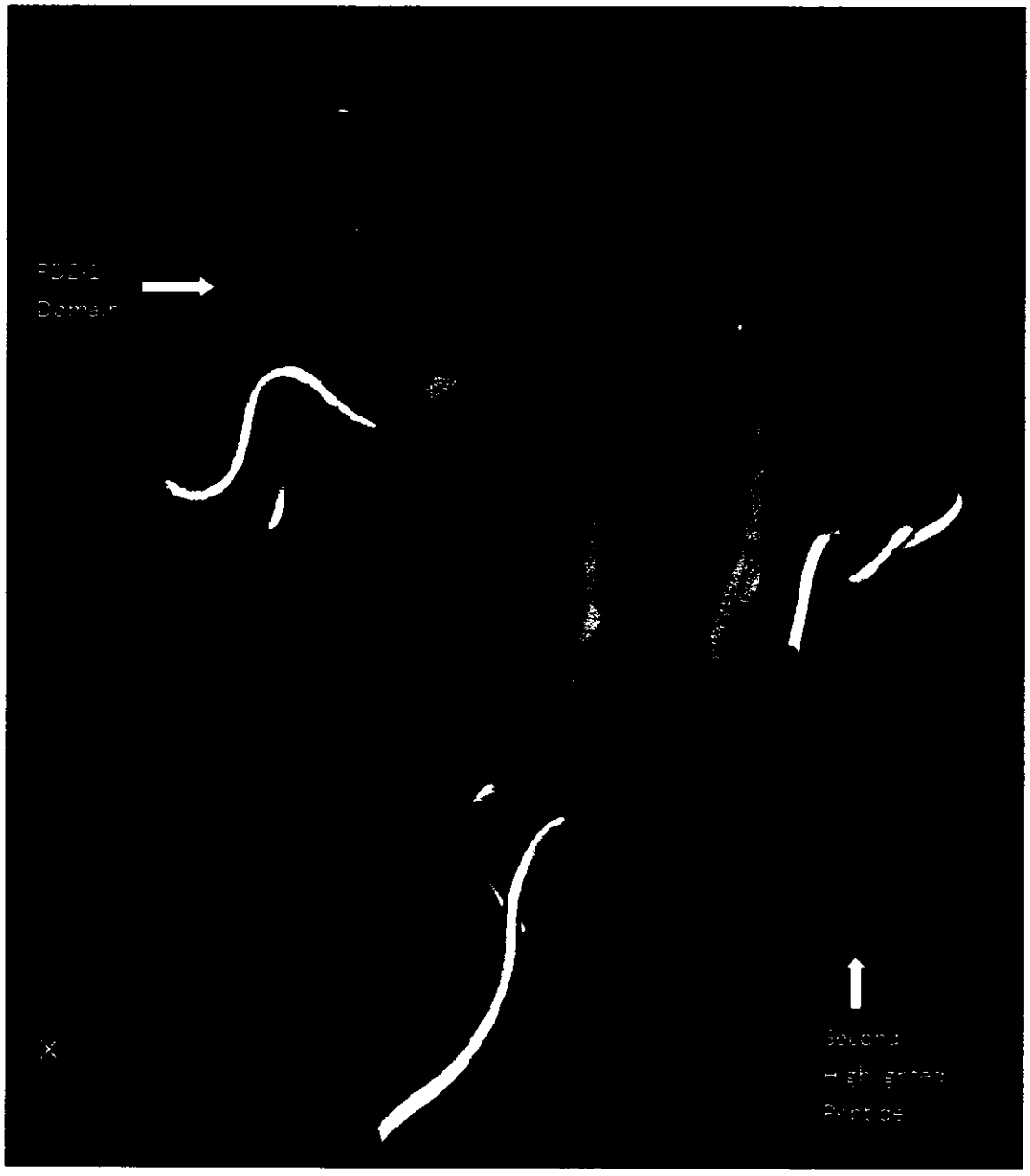


Fig. 4.19 Docking output of second highlighted peptide with PDZ-1 domain.

4.5 Analyzing Docking Outputs

RasWin is visualization tool for 3-D protein structures; possible interacting residues between receptor and the ligand were roughly estimated using it.

White circle in figure 4.20 shows interacting region of PDZ-2 domain with it's receptor, i.e., V3 loop of gp120 protein (blue colored). Identification of this region was followed by picking these residues and roughly estimating distances between them to limit further search by VMD tool. Via RasWin residues were identified for receptor molecule only because while processing with VMD identified receptor residues will limit the search only to ligand residues, which made this work more feasible and controlled. Following are the estimated interacting residues of PDZ-2 domain (numbering is done according to PDB file);

208, 209, 210, 211, 212, 213, 214 and 215 (CHAIN A)

193 TO 205 (CHAIN B)

Following are residues of PDZ-1 domain that possible interact with V3 loop shown within white circle in figure 4.21;

17, 18 19, 20, 21, 53, 54, 55, 108, 109, and 110

Interacting region of PDZ-3 domain and V3 loop are shown in figure 4.22. Possible interacting residues of PDZ-3 domain are;

36, 37, 38, 32, 33, 34, 81 and 82.

Same work on SH3-GUK module retrieved region shown in figure 4.23, residues of SH3-GUK module in close vicinity to ligand are;

762, 763, 764, 765, 766, 770, 771, 772, 795, 796 and 797.

Fig. 4.20 Interacting region of PDZ-2 domain with 1ce4.

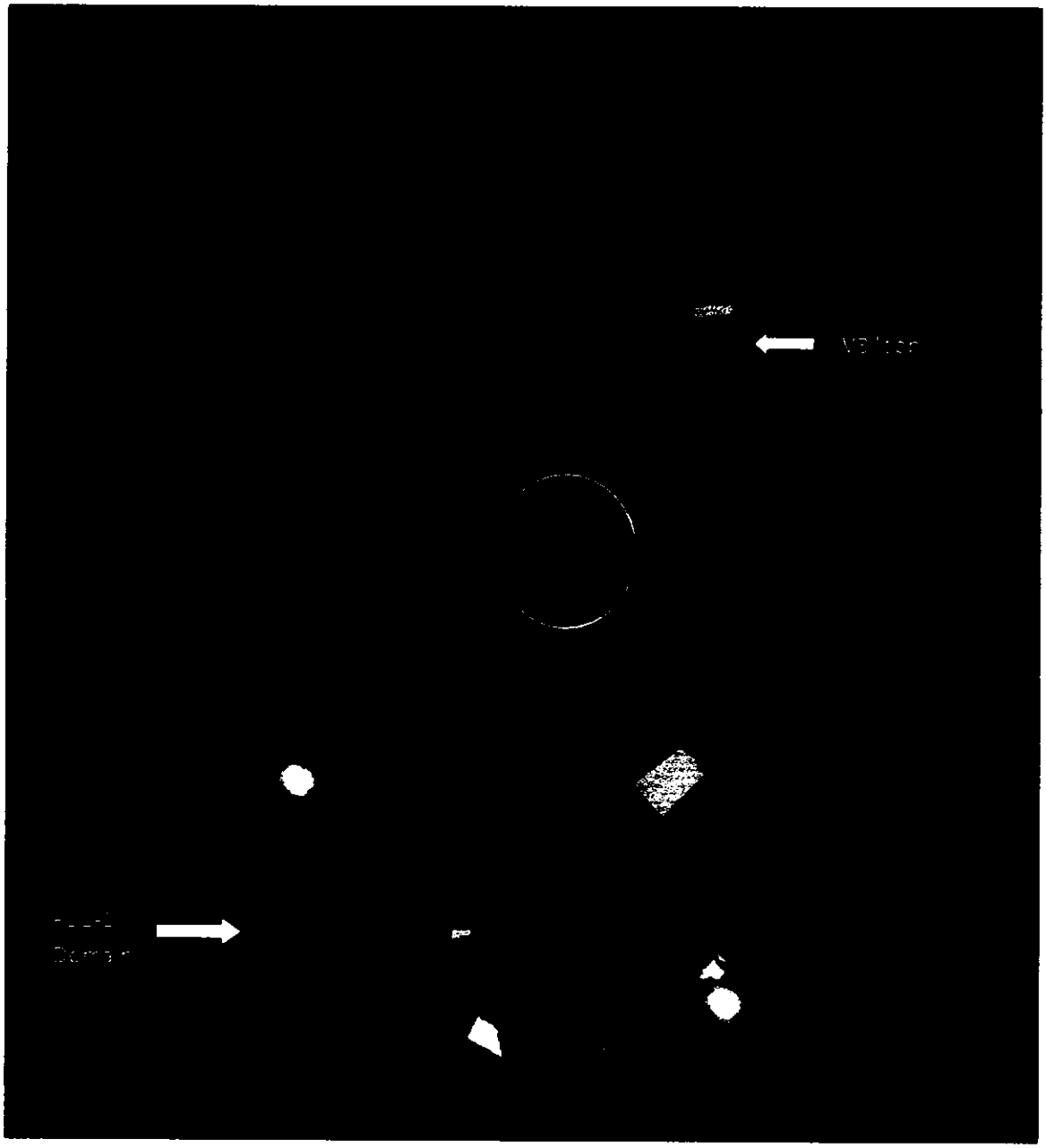


Fig. 4.21 Interacting region of PDZ-1 domain with V3 loop.

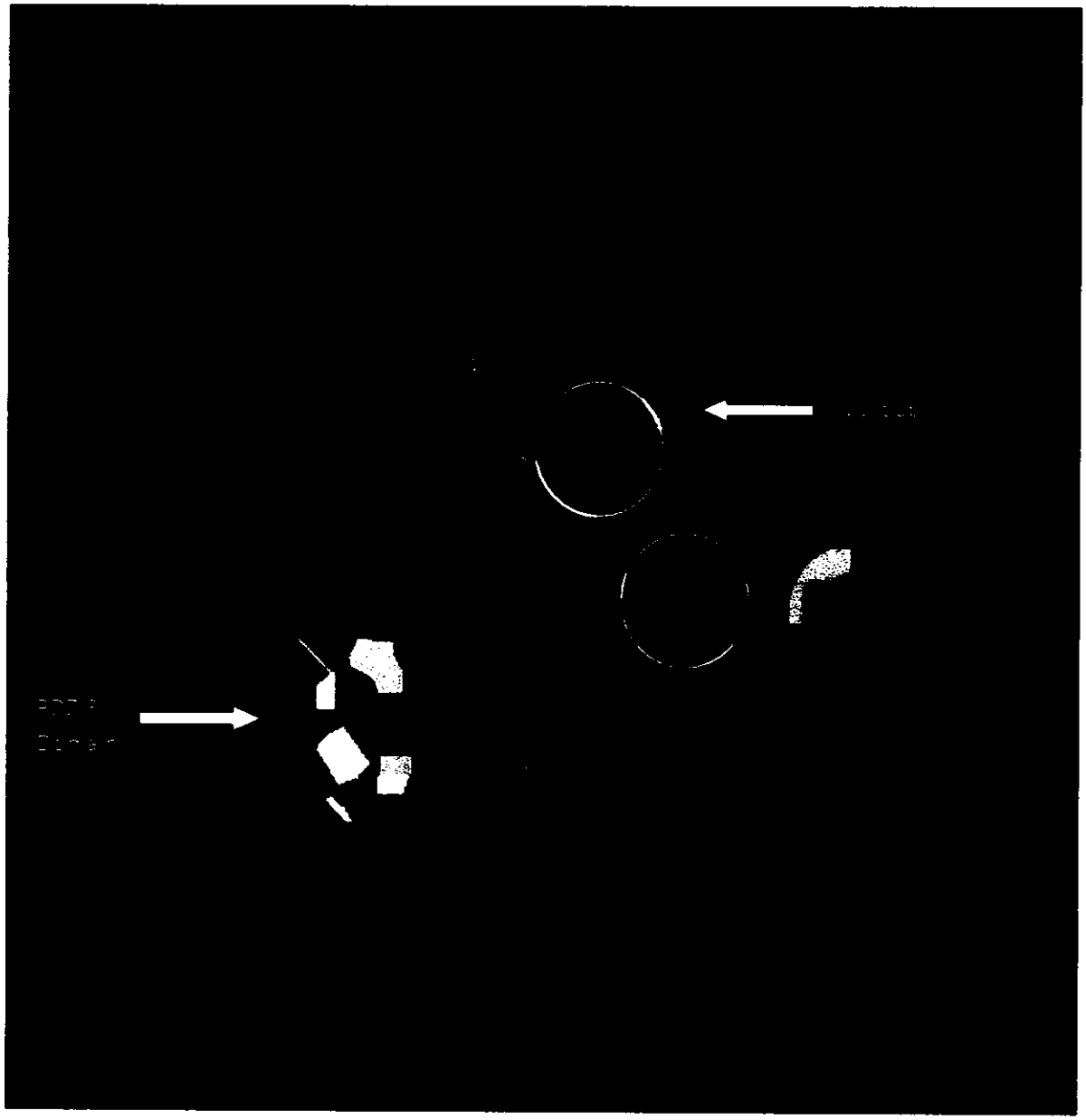


Fig. 4.22 Encircled is the interacting region of PDZ-3 domain with V3 loop.

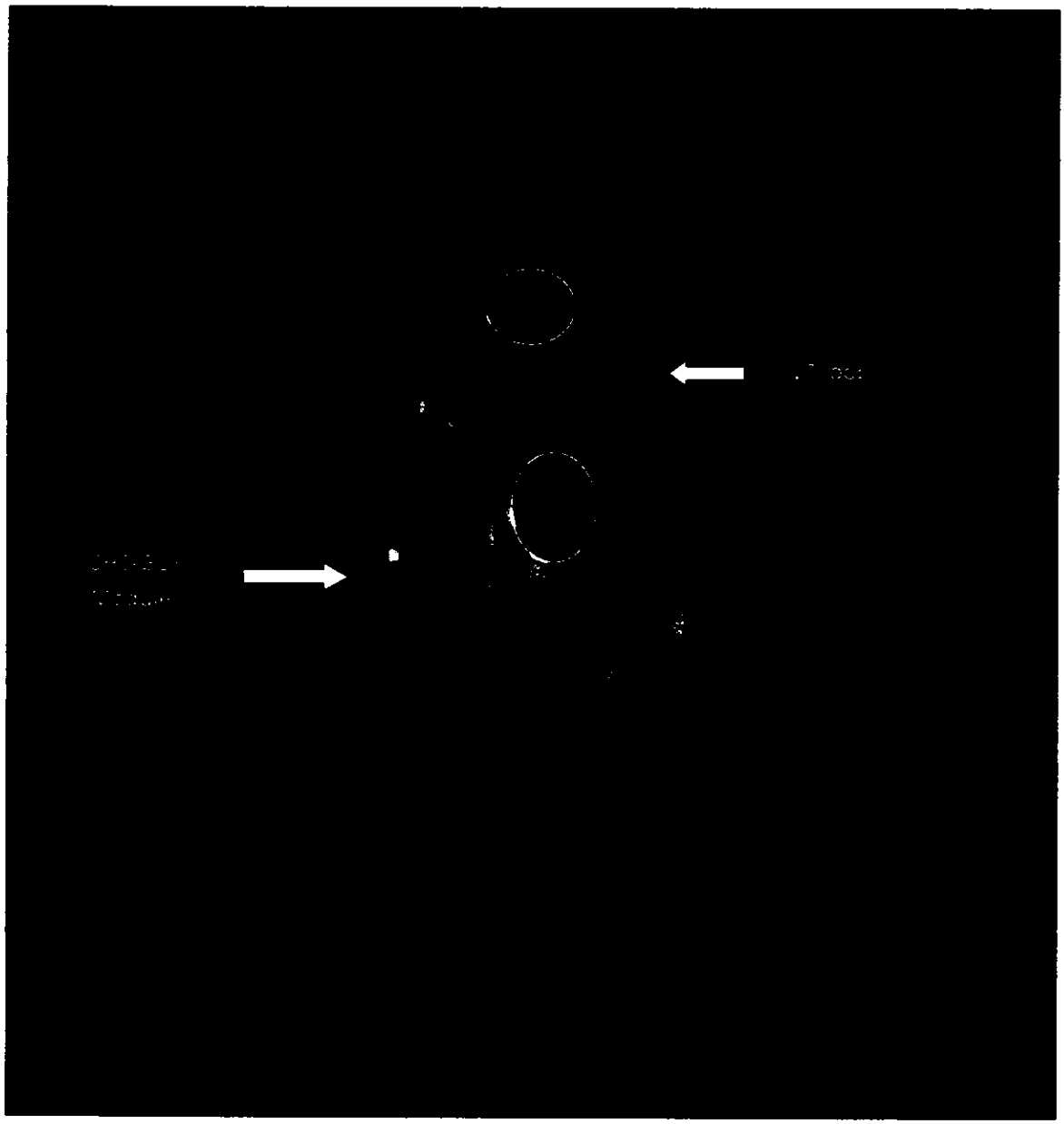


Fig. 4.23 White circles shows possible interacting region of SH3-GUK module with V3 loop.

Fig. 4.24 Possible interacting region of PDZ-1 domain and first highlighted peptide is shown within white circles.

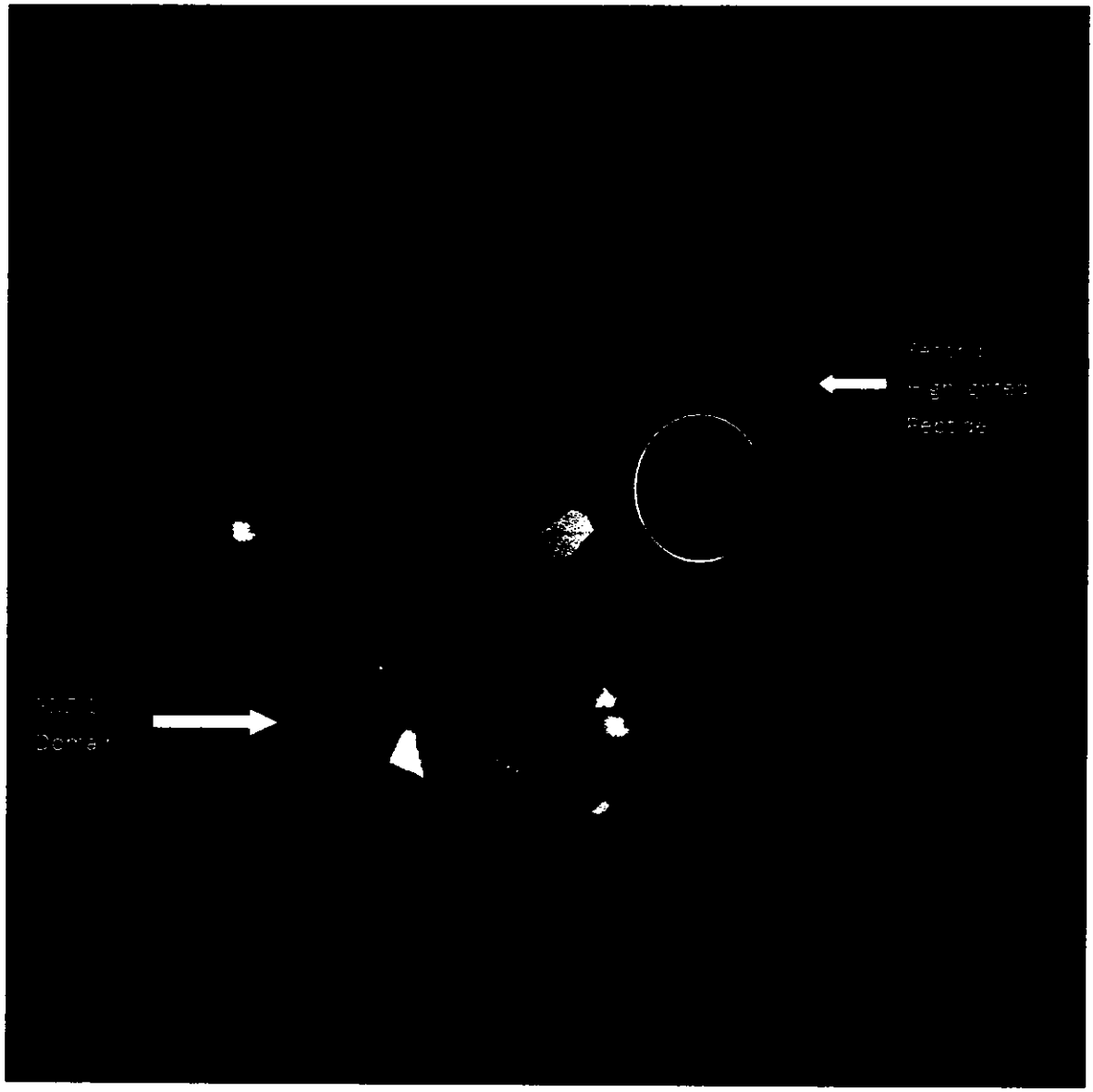


Fig. 4.25 White circle shows possible region of interaction between PDZ-1 domain and second highlighted peptide of V3 region.

Figure 4.24 shows possible interacting region of PDZ-1 domain with first peptide of V3 region. Following are the possible interacting residues;

18, 19, 20, 21, 22, 76, 77, 106, 107, 108, 109 and 110.

Figure 4.25 shows possible region of PDZ-1 domain of zonula occludens-1 protein interacting with second highlighted peptide of V3 region and this includes following residues; 21, 22, 23, 24, 66, 67 and 68.

4.6 Binding Interactions between Receptor and Ligand Residues

The above mentioned possible interacting residues for each docked result was then subjected to identify the residues within 5.0 angstrom of most probable interacting residues, this was done because the possibility of interaction between two residues exist only if the distance between them is not more than 5 angstrom. This was achieved by residue selection using Visual Molecular Dynamics (VMD) commands, output of this command was then labeled and residues of ligands were identified to calculate distance between residue of ligand and receptor.

4.6.1 Most Probable Interacting Residues of each Docked Result

Tables (4.2, 4.3 4.4, 4.5, 4.6 and 4.7) shows separately the residues with within 5.0 angstrom of possible interacting residues of ligand and receptor for each docked results.

4.6.2 Interactions between Receptor and Ligand Residues

After identifying the residues within close vicinity of receptor, residues of ligands were alienated and both residues of receptor and ligand were subjected to evaluate their possible interactions. Table 4.8 shows side chains of residues that can interact with each other via ionic interactions, hydrogen bonding, hydrophobic interactions etc.

Tables 4.9, 4.10, 4.11 and 4.12 shows the distances between residues of PDZ-1, PDZ-2, PDZ-3 and SH3-GUK domains of ZO-1 protein, identified in previous step with residues of V3 loop that lies inside 5.0 distance of possibly interacting residues of ZO-1 protein, respectively. While these distances between residues of PDZ-1 domain and first highlighted peptide of V3 region and second highlighted peptide of V3 region are mentioned in table 4.13 and 4.14 respectively.

Table 4.2 Residues within 5.0 angstrom of amino acids 17, 18 19, 20, 21, 70, 71, 72, 108, 109, and 110 of PDZ-1 domain.

RESIDUES	LIGAND/RECEPTOR	RESIDUES	LIGAND/RECEPTOR
Thr23	L	Thr105	R
Gln21	R	His22	R
Ile106	R	Arg107	R
Arg108	R	Arg74	R
Glu51	R	Gly24	L
Lys109	R	Lys110	R
Arg9	L	Trp19	R
Ile18	R	Thr8	L
Met17	R	Glu20	R
Thr105	R	Lys10	L
Thr8	R	Asn72	R
Asp73	R	Leu69	R
Gln68	R	Gln70	R
Thr22	L	Ser11	R
Glu25	L	Asn5	R
Asn6	R	Thr8	R
Asn7	L	Arg31	R
Arg74	R	Phe47	R
His46	R	Val55	R
Pro16	L	Asp58	R
Ile56	R	Phe20	L
Val50	R	Leu69	R
Arg74	R	Ser11	L
Arg9	L	Ile14	L
Asp73	R	Thr22	L
Thr23	L	Arg108	R
Ile26	L		

(R stands for receptor while L shows that residue belongs to ligand.)

Table 4.3 Residues within 5.0 angstrom of amino acids 193, 194, 195, 106, 197, 198, 199, 200, 201, 202, 203, 204, 205, 208, 209, 210, 211, 212, 213, 214 and 215 of PDZ-2 domain.

RESIDUES	LIGAND/RECEPTOR	RESIDUES	LIGAND/RECEPTOR
ARG 193	R	LYS 194	R
ASN 195	R	GLU 196	R
GLU 197	R	TYR 198	R
ARG 201	R	LEU 200	R
GLY 199	R	LEU 202	R
ALA 203	R	SER 204	R
HIS 205	R	SER 212	R
ILE 211	R	ILE 249	R
ILE 223	R	MET 258	R
GLU 210	R	ILE 232	R
VAL 229	R	VAL 208	R
ALA 245	R	LYS 209	R
PHE 207	R	ILE 206	R
HIS 13	L	SER 11	L
ILE 12	L	LYS 246	R
ASN 239	R	GLU 238	R
MET 240	R	SER 241	R
SER 252	R	ASN 233	R
LYS 153	R	GLY 254	R
LYS 255	R	LEU 256	R
LYS 191	R	LEU 189	R
SER 215	R	LEU 216	R
VAL 190	R	ARG 219	R
ALA 218	R	ALA 217	R
THR 22	L	GLY 24	L
THR 23	L	ARG 9	L
LEU 230	R	THR 237	R
GLN 261	R	VAL 228	R
PHE 20	L	SER 192	R
LEU 242	R	ILE 14	L

RESIDUES	LIGAND/RECEPTOR	RESIDUES	LIGAND/RECEPTOR
TYR 21	L	Ile 206	R
Thr 8	L	Ser 11	L
Leu 200	R	Gly 199	R
Leu 202	R	Val 229	R
Glu 197	R	Arg 9	L
Gly 24	L	Glu 25	L
Ile 26	L	Thr 22	L
Arg 201	R	His 205	R
Phe 207	R	Ala 203	R
Thr 23	L	Gly226	R
Val 228	R	Glu 225	R
Arg 219	R	Ala 218	R
Leu 216	R	Ala 217	R
Gln 224	R	Ile 223	R
Ile 211	R	Tyr 198	R
Phe 20	L	Gly 17	L
Ile 14	L	Pro 16	L
ARG762	R	ASN711	R
LEU768	R	LYS539	R
HIS767	R	ASN538	R
ASN764	R	ASN765	R
PRO710	R	LYS760	R
ARG757	R	LEU761	R
ASN764	R	HIS766	R
LYS763	R	PHE769	R
HIS759	R	LYS10	L
ASN7	L	THR8	L
ARG762	R	PRO4	L
ASN5	L	TYR755	R
SER758	R	PHE727	R
PRO730	R	ASN729	R

(R stands for receptor while L shows that residue belongs to ligand.)

Table 4.4 Residues within 5.0 angstrom of amino acids 206, 207,208, 209, 210,214, 215, 216, 239, 240 and 241 of SH3-GUK module.

RESIDUES	LIGAND/RECEPTOR	RESIDUES	LIGAND/RECEPTOR
ARG762	R	ASN711	R
LEU768	R	LYS539	R
HIS767	R	ASN538	R
ASN764	R	ASN765	R
PRO710	R	LYS760	R
ARG757	R	LEU761	R
ASN764	R	HIS766	R
LYS763	R	PHE769	R
HIS759	R	LYS10	L
ASN7	L	THR8	L
ARG762	R	PRO4	L
ASN5	L	TYR755	R
SER758	R	PHE727	R
PRO730	R	ASN729	R
THR770	R	PHE769	R
LYS10	L	ASN5	L
ARG645	R	GLN793	R
LYS564	R	GLU25	L
ARG9	L	THR771	R
THR772	R	THR8	L
VAL725	R	ILE724	R
PHE727	R	ASN774	R
ASN729	R	VAL726	R
ALA789	R	ILE773	R
LYS10	L	PHE769	R
THR8	L	GLN793	R
ARG645	R	GLN793	R
ASN565	R	HIS566	R
LEU638	R	LEU797	R
LEU796	R	GLN792	R
GLN791	R	ASN795	R
ALA641	R	GLU640	R
VAL798	R	GLN794	R

(R stands for receptor while L shows that residue belongs to ligand.)

Table 4.5 Residues within 5.0 angstrom of amino acids 18, 19, 20, 21, 22, 76, 77, 106, 107, 108, 109, 110 of PDZ-1 domain when docked with first highlighted peptide of V3 region.

RESIDUES	LIGAND/RECEPTOR	RESIDUES	LIGAND/RECEPTOR
ASN9	L	THR23	R
VAL24	R	LYS11	L
GLN21	R	HIS22	R
GLU20	R	ARG108	R
GLN68	R	ILE106	R
THR105	R	ARG107	R
ARG108	R	ARG74	R
GLU51	R	GLU20	R
LYS110	R	ILE18	R
TRP19	R	PHE6	L
MET17	R	ILE1	L
ILE13	L	ARG107	R
ARG74	R	THR105	R
ILE104	R	LYS11	L
ASN9	L	HIS22	R
GLN21	R	ARG108	R
GLU20	R	ILE106	R
GLN68	R	ARG74	R
ARG107	R	ILE18	R
LYS110	R	MET17	R
LYS109	R	ILE104	R
ILE13	L	LEU69	R
GLN70	R	ILE56	R

RESIDUES	LIGAND/RECEPTOR	RESIDUES	LIGAND/RECEPTOR
ASP73	R	SER3	L
VAL75	R	ALA76	R
THR105	R	PHE6	L
TRP19	R	ASN5	L
THR7	L	ALA10	L
ASP8	L	ARG74	R
ARG74	R	LYS11	L
ASN9	L	HIS22	R
GLN21	R	ARG108	R
GLU20	R	ILE106	R
GLN68	R	ARG74	R
ARG107	R	ILE18	R
LYS110	R	MET17	R
LYS109	R	ILE13	L
THR105	R	GLN70	R
ILE104	R	ASP73	R
LEU69	R	VAL75	R
ILE56	R	ALA76	R
SER3	L	PHE6	L
TRP19	R	ASN5	L
THR7	L	ALA10	L
ASP8	L		

(R stands for receptor while L shows that residue belongs to ligand.)

Table 4.6 Residues within 5.0 angstrom of amino acids 21, 22, 23, 24, 66, 67 and 68 of PDZ-1 when docked with second highlighted peptide of V3 region.

RESIDUES	LIGAND/RECEPTOR	RESIDUES	LIGAND/RECEPTOR
GLN21	R	HIS22	R
THR23	R	VAL24	R
LEU26	R	GLN68	R
ILE104	R	LEU69	R
ARG108	R	ILE106	R
ARG107	R	TRP19	R
GLU20	R	LYS103	R
ALA102	R	VAL24	R
LEU26	R	THR25	R
ILE106	R	THR105	R
GLU20	R	GLN21	R
HIS22	R	HIS21	L
ASN20	L	PHE19	L
ILE17	L	VAL18	L
THR105	R	ARG107	R
ILE4	L	LEU8	L
ILE104	R	THR25	R
GLU66	R	HIS22	R
VAL24	R	ARG108	R
THR25	R	GLN68	R
LEU69	R	GLY67	R
GLN70	R	VAL59	R
LYS61	R	ALA65	R
PRO64	R	LEU26	R

RESIDUES	LIGAND/RECEPTOR	RESIDUES	LIGAND/RECEPTOR
HIS27	R	LEU8	L
PHE12	L	GLN11	L
LYS7	L	GLN70	R
GLY62	R	LYS61	R
GLY63	R	ALA65	R
GLU66	R	ILE4	L
VAL59	R	HIS27	R

(R stands for receptor while L shows that residue belongs to ligand.)

Table 4.7 Within 5.0 of ala36, ala37, lys38, gln81, lys82, glu32, asp33 and ser34 of PDZ-3 domain when docked with V3 loop of viral gp120 protein.

RESIDUES	LIGAND/RECEPTOR	RESIDUES	LIGAND/RECEPTOR
VAL13	R	GLY24	L
VAL5	R	LYS38	R
GLY14	R	THR22	L
LEU15	R	GLU25	L
ILE27	L	TYR21	L
1LE26	L	ASP33	R
GLY28	L	GLU32	L
ASP29	L	SER34	R
THR23	L	ALA37	R
GLU39	R	GLY40	R
GLU25	L	ILE14	L
HIS13	L	VAL30	R
LEU31	R	LEU41	R
PRO35	R	LEU79	R
ASN5	L	ASN6	L
ARG3	L	ALA80	R
GLN81	R	LYS82	R

(R stands for receptor while L shows that residue belongs to ligand.)

Table 4.8 Possible interactions between side chains of residues along with type of interactions.

Side Chain group	**Side Chain group**	**Type of interaction**
CO_2^-	NH_3^+	Ionic
HO	SH	Hydrogen bonding
HO	NH	
HO	OH	
C	C	Hydrophobic

Table 4.9 Binding interactions between residues of PDZ-1 domain and V3 loop after docking.

RECEPTOR →

LIGAND ↓

	17(met) S	18(ile) CG2	19(trp) CB	20(glu) OE1	21(gln) OE1	108(arg) 2HH1	109(lys) 1HZ	110(lys) 2HZ	70(gln) NE2	71(glu) OE2	72(asn) ND2
7(asn)	-	-	-	-	-	-	-	-	-	-	-
8(thr)	Y 5.28 OH	-	-	-	Y 17.19 HG21	-	-	-	Y 17.66 OG1	-	Y 12.91 OG1
9(arg)	-	-	-	Y 7.79 NH1	-	-	-	-	-	Y 9.37 HH22	-
10(lys)	-	-	-	Y 15.49 HB3	-	-	-	-	-	Y 12.73 HB3	-
11(ser)	Y 10.28 OH	-	-	-	Y 18.63 HB2	-	-	-	Y 10.67 OG	-	Y 4.65 OG
14(ile)	-	Y 16.95 CD1	Y 14.93 CD1	-	-	-	-	-	-	-	-
16(pro)	-	-	-	-	-	-	-	-	-	-	-

	17(met) S	18(ile) CG2	19(trp) CB	20(glu) OE1	21(gln) OE1	108(arg) 2HH1	109(lys) 1HZ	110(lys) 2HZ	70(gln) NE2	71(glu) OE2	72(asn) ND2
20(phe)	-	Y 18.98 CB	Y 17.62 CB	-	-	-	-	-	-	-	-
22(thr)	Y 15.33 HA	-	-	-	Y 16.88 OG1	-	-	-	Y 2.98 HA	-	Y 9.81 HA
23(thr)	Y 15.21 OH	-	-	-	Y 11.40 OG1	-	-	-	Y 4.42 OG1	-	Y 12.64 OG1
24(gly)	-	-	-	-	-	-	-	-	-	-	-
25(glu)	-	-	-	-	-	Y 9.36 OE1	Y 12.20 OE1	Y 9.48 OE1	-	-	-
26(ile)	-	Y 15.00 CG2	Y 15.58 CG1	-	-	-	-	-	-	-	-

(Number on right side of 3-letter amino acid code is PDB ID of residue. In each cell (Y) shows possibility of existence of interaction between receptor and ligand's residue followed by calculated distance which is then followed by atom used while distance is calculated. (-)shows that there is no possibility of existence of interaction or side chains exceeds 5.0 angstrom bound.)

To Study Mol. Determinants of HIV-1 Neuro Invasion and prevalence Rate of HAD in HIV infected Pakistani Population. 86

Chapter 4 Results

Table 4.10 (a) Binding interactions between residues of PDZ-2 domain and V3 loop after docking.

RECEPTOR ⟶

LIGAND ↓

	193 (arg)	**194 (lys)**	**195 (asn) ND2**	**196 (glu) OE2**	**197 (glu) OE1**	**198 (tyr) OH**	**199 (gly) N**	**200 (leu) CB**	**201 (arg)**	**202 (leu) CB**	**203 (ala) CB**	**204 (ser) OG**	**205 (his)**
14 (ile)	-	-	-	-	-	-	-	Y 12.03 CD1	-	Y 13.18 CD1	Y 12.38 CD1	-	-
21 (tyr) OH	-	-	Y 27.63 OH	-	-	Y 28.38 OH	Y 20.61 OH	Y 14.45 CB	-	Y 15.59 CG	Y 11.70 CB	Y 18.88 HH	-
11 (ser) OG	-	-	Y 16.36 OG	-	-	Y 17.61 OG	Y 12.13 OG	-	-	-	-	Y 10.72 HB2	-
20 (phe)	-	-	-	-	-	-	-	Y 10.99 CB	-	Y 14.10 CE1	Y 12.92 CD1	-	-
9 (arg)	-	-	-	Y 23.92 NH1	Y 15.14 NH1	Y 23.25 NE	-	-	-	-	-	Y 6.70 HG3	-
23 (thr) OG1	-	-	Y 24.02 OG1	-	-	Y 20.80 OG1	Y 15.24 OG1	-	-	-	-	Y 10.46 HG1	-
24 (gly)	-	-	-	-	-	Y 25.11 HA2	-	-	-	-	-	Y 11.79 HA2	-

	193 (arg)	**194 (lys)**	**195 (asn) ND2**	**196 (glu) OE2**	**197 (glu) OE1**	**198 (tyr) OH**	**199 (gly) N**	**200 (leu) CB**	**201 (arg)**	**202 (leu) CB**	**203 (ala) CB**	**204 (ser) OG**	**205 (his)**
22 (thr) OG1	-	-	Y 22.19 OG1	-	-	Y 21.78 OG1	Y 15.28 OG1	-	-	-	-	Y 10.69 HB	-
13 (his)	-	-	-	Y 20.28 NE2	Y 10.65 NE2	Y 20.68 NE2	-	-	-	-	-	Y 17.00 HD2	-
12 (ile)	-	-	-	-	-	-	-	Y 16.75 CG1	-	Y 14.31 CG1	Y 10.94 CG1	-	-

(Number on right side of 3-letter amino acid code is PDB ID of residue. In each cell (Y) shows possibility of existence of interaction between receptor and ligand's residue followed by calculated distance which is then followed by atom using which distance is calculated. (-) shows that there is no possibility of existence of interaction between receptor and ligand residue or side chains are out of 5.0 angstrom bound . First row and first column shows residues of receptor and ligand respectively.)

Table 4.10 (b) Binding interactions between residues of PDZ-2 domain and V3 loop after docking.

RECEPTOR →

LIGAND ↓

	208(val) CB	**209(lys) CB NZ**	**210(glu) OE1**	**211(ile) CB**	**212(ser) OG**	**213(gln) NE2**	**214(asp) OD1**	**215(ser) OG**
14 (ile)	Y CD1 13.38	Y CG1 9.62	-	Y 10.80 CG2	-	-	-	-
11 (ser)	-	Y 8.73 OG	-	-	Y 9.33 HG	Y 13.99 OG	-	Y HG 13.56
20 (phe)	Y CB 9.75	Y CB 8.38	-	Y 6.90 CD2	-	-	-	-
9 (arg)	-	-	Y NH1 7.71	-	-	-	Y NH1 22.35	-
23 (thr)	-	Y OG1 3.84	-	-	Y 10.98 HG22	Y 14.95 OG1	-	Y 12.79 HG22
24 (gly)	-	-	-	-	Y 16.70 HA2	-	-	Y 19.30 HA2
22(thr)	-	Y 4.02 HG1	-	-	Y 8.84 HG21	Y 13.17 OG1	-	Y 12.63 HG21

	208(val) **CB**	**209(lys)** **CB** **NZ**	**210(glu)** **OE1**	**211(ile)** **CB**	**212(ser)** **OG**	**213(gln)** **NE2**	**214(asp)** **OD1**	**215(ser)** **OG**
16 (pro)	-	-	-	-	Y 10.55 HD2	-	-	Y HB2 13.55
17 (gly)	-	-	-	Y 10.16 HA2	-	-	-	-
26 (ile)	Y CD1 9.76	Y CD1 7.60	-	Y 8.34 HD12	-	-	-	-
25 (glu)	-	-	-	-	Y 13.14 HE2	-	-	Y 17.51 HE21
8 (thr)	-	Y 5.20 OG1	-	-	Y 16.64 HG22	Y 19.28 OG1	-	Y 20.59 HG22

(Number on right side of 3-letter amino acid code is PDB ID of residue. In each cell (Y) shows possibility of existence of interaction between receptor and ligand's residue followed by calculated distance which is then followed by atom using which distance is calculated. (-) shows that there is no possibility of existence of interaction or side chains are out of 5.0 angstrom bound. First row and first column shows residues of receptor and ligand respectively. These tables (4.8 (a) and (b)) show distance between residues of PDZ-2 domain when docked with V3 loop of Gp120.)

Table 4.11 (a) Binding interactions between residues of PDZ-3 domain and V3 loop after docking.

RECEPTOR →

LIGAND ↓

	Ala 36 (CH3)	**Ala 37 (CH3)**	**Lys38 (NH3)**	**Gln 81 (NH2-C=O)**	**Lys 82 (NH3-)**
Cys1 (SH)	-	-	-	Y 17.18	Y 17.04
Thr2 (OH)	-	-	Y 11.63	Y 15.68	Y 17.73
Arg3 (NH2)	-	-	-	Y 13.42	-
Pro4 (CH2)	-	-	-	-	-
Asn5 (NH2-C=O)	-	-	Y 18.59	-	-
Asn6 (NH2-C=O)	-	-	Y 19.67	Y 9.00	Y 6.31
Asn7 (NH2-C=O)	-	-	Y 17.26	-	-
Thr8 (OH)	-	-	Y 14.34	-	-
Arg9 (NH2)	-	-	-	Y 10.83	-
Lys10 (NH3)	-	-	-	Y 17.62	-
Ser11 (OH)	-	-	Y 13.45	Y 17.43	Y 20.55
Ile12 (CH3)	Y 12.17	Y 12.42	-	-	-
His13 (NH)	-	-	-	-	-
Ile14 (CH3)	-	-	-	-	-
Gly15	-	-	-	-	-

	Ala 36 (CH3)	**Ala 37 (CH3)**	**Lys38 (NH3)**	**Gln 81 (NH2-C=O)**	**Lys 82 (NH3-)**
Pro16 (CH2)	Y 19.28	Y 16.65	-	-	-
Gly17	-	-	-	-	-
Arg18 (NH2)	-	-	-	-	-
Ala19 (CH3)	Y 17.35	Y 13.66	-	-	-
Phe20 (Benzene ring)	Y 18.04	Y 13.93	-	-	-
Tyr21 (Phenyl)	-	-	Y 13.99	Y 19.31	Y 22.70
Thr22 (OH)	-	-	Y 5.94	Y 23.96	Y 27.56
Thr23 (OH)	-	-	Y 10.40	Y 30.01	Y 32.77
Gly24	-	-	-	-	-
Glu25 (CO2)	-	-	Y 7.20	-	-
Ile26 (CH3)	Y 13.90	Y 9.47	-	-	-
Ile27 (CH3)	-		-	-	-
Gly28	-	-	-	-	-
Asp29 (CO2)	-	-		-	-
Ile30 (CH3)	Y 15.70	Y 12.81	-	-	-
Arg31 (NH2)	-	-	-	-	-
Gln32 (O=C-NH2) | CH2	-	-	Y 11.09	Y 19.80	Y 18.32
Ala33 (CH3)	Y 16.01	Y 14.87	-	-	-

	Ala 36 (CH3)	**Ala 37 (CH3)**	**Lys38 (NH3)**	**Gln 81 (NH2-C=O)**	**Lys 82 (NH3-)**
His34 (NH)	-	-	-	-	-
Cys35 (SH)	-	-	-	Y 19.84	Y 18.68

(Number on right side of 3-letter amino acid code is PDB ID of residue. In each cell (Y) shows possibility of existence of interaction between receptor and ligand's residue followed by calculated distance. (-) shows hat there is no possibility of existence of interaction or functional group lies outside the distance limit of 5.0 angstrom. First row and first column shows residues of receptor and ligand respectively. These tables (4.8 (a) and (b)) show distance between residues of PDZ-2 domain when docked with V3 loop of Gp120.)

Table 4.11 (b) Binding interactions between residues of PDZ-3 domain and V3 loop after docking.

RECEPTOR →

LIGAND ↓

	Glu 32 (CO2)	**Asp 33 (CO2)**	**Ser 34 (OH)**
Cys1 (SH)	-	-	-
Thr2 (OH)	Y 18.86	Y 18.31	Y 13.10
Arg3 (NH2)	Y 19.90	Y 21.26	Y 15.38
Pro4 (CH2)	-	-	-
Asn5 (NH2-C=O)	Y 20.46	Y 20.54	Y 17.42
Asn6 (NH2-C=O)	-	-	-
Asn7 (NH2-C=O)	Y 19.97	Y 21.04	Y 19.52
Thr8 (OH)	-	-	-
Arg9 (NH2)	Y 16.88	Y 16.98	Y 13.07
Lys10 (NH3)	Y 12.12	Y 11.67	Y 11.50
Ser11 (OH)	Y 12.42	Y 14.81	Y 13.50

	Glu 32 (CO2)	**Asp 33 (CO2)**	**Ser 34 (OH)**
Ile12 (CH3)	-	-	-
His13 (NH)	Y 8.47	Y 8.89	Y 9.84
Ile14 (CH3)	-	-	-
Gly15	-	-	-
Pro16 (CH2)	-	-	-
Gly17	-	-	-
Arg18 (NH2)	-	-	-
Ala19 (CH3)	-	-	-
Phe20 (Benzene ring)	-	-	-
Tyr21 (Phenyl)	Y 12.58	Y 12.83	Y 15.02
Thr22 (OH)	Y 8.55	Y 7.36	Y 6.41
Thr23 (OH)	-	-	-
Gly24	-	-	-
Glu25 (CO2)	-	-	Y 9.05
Ile26 (CH3)	-	-	-
Ile27 (CH3	-	-	-
Gly28	-	-	-
Asp29 (CO2)	-	-	Y 11.76

	Glu 32 (CO2)	**Asp 33 (CO2)**	**Ser 34 (OH)**
Ile30 (CH3)	-	-	-
Arg31 (NH2)	-	-	-
Gln32 (O=C-NH2) | CH2	Y 18.49	Y 15.98	Y 11.62
Ala33 (CH3)	-	-	-
His34 (NH)	-	-	-
Cys35 (SH)	-	-	-

(First row shows possible interacting residue of PDZ-3 domain while first column shows each residue of V3 loop, within brackets are interacting side chains of receptor and ligand. (Y) shows possibility of existence of interactions and (-) shows that there exist no possibility of interaction b/w receptor and ligand or side chains are out of 5.0 angstrom bound.)

Table 4.12 Binding interactions between residues of SH3-GUK module and V3 loop after docking.

RECEPTOR →

LIGANDS ↓

	762(arg)	763(lys) NZ	764 (asn) ND2	765(asn) ND2	766(his)	770 (thr) OG1	771 (thr) OG1	772 (thr) OG1
4 (pro)	-	-	-	-	-	-	-	-
5 (asn)	-	-	-	-	-	Y 14.14 HB2	Y 10.96 HB2	Y 5.28 HB2
7 (asn)	-	-	-	-	-	-	-	-
8 (thr)	-	Y 12.53 OG1	Y 10.60 OG1	Y 9.55 OG1	-	Y 10.35 HG21	Y 8.25 HG21	Y 2.58 HG21
10 (lys)	-	-	-	-	-	-	-	-
25 (glu)	-	-	-	-	-	-	-	-
9 (arg)	-	-	-	-	-	-	-	-

(Number on left side of 3-letter amino acid code is PDB ID of residue. In each cell (Y) shows possibility of existence of interaction between receptor and ligand's residue followed by calculated distance. (-) shows hat there is no possibility of existence of interaction or functional group lies outside the distance limit of 5.0 angstrom.

Table 4.13 Binding interactions between residues of PDZ-1 domain and first highlighted peptide of V3 region after docking.

RECEPTOR →

LIGAND ↓

	106 (ile) CB	**107 (arg) NH1**	**108 (arg) NH2**	**109 (lys) NZ**	**110 (lys) NZ**	**76 (ala) CB**	**77 (met) SD**	**18 (ile) CB**	**19 (trp) CB**	**20 (glu) 0E2**	**21 (gln)**	**22 (his)**
11 (lys)	-	-	-	-	-	-	-	-	-	Y NZ 21.28	-	-
13 (ile)	Y CG2 12.00	-	-	-	-	-	-	Y CG2 8.14	Y CG2 5.83	-	-	-
3 (ser)		-	-	-	-	-	-	-	-	-	-	-
6 (phe)	Y CD2 10.64	-	-	-	-	Y CB 7.57	-	Y CE2 7.98	Y CE2 3.81	-	-	-
7 (thr)	-	-	-	-	-	-	Y OG1 13.19	-	-	-	-	-
5 (asn)	-	-	-	-	-	-	-	-	-	-	-	-
8 (asp)	-	Y OD1 6.39	Y OD1 18.65	Y OD1 18.04	Y OD1 20.76	-	-	-	-	-	-	-
9 (asn)	-	-	-	-	-	-	-	-	-	-	-	-
10 (ala)	Y CB 13.96	-	-	-	-	Y CB 7.58	-	-	-	-	-	-
1 (ile)	-	-	-	-	-	-	-	Y CG1 4.36	Y CG1 3.95	-	-	-

(First row shows possible interacting residue of PDZ-1 domain while first column shows residue of first highlighted peptide of V3 region, within brackets are interacting sidchains of receptor and ligand. 'Y' shows possibility of existence of interactions and '-' shows that there exist no possibility of interaction b/w receptor and ligand or functional group lies outside 5.0 angstrom bound.)

Table 4.14 Binding interactions between residues of PDZ-1 domain and second highlighted peptide of V3 region after docking.

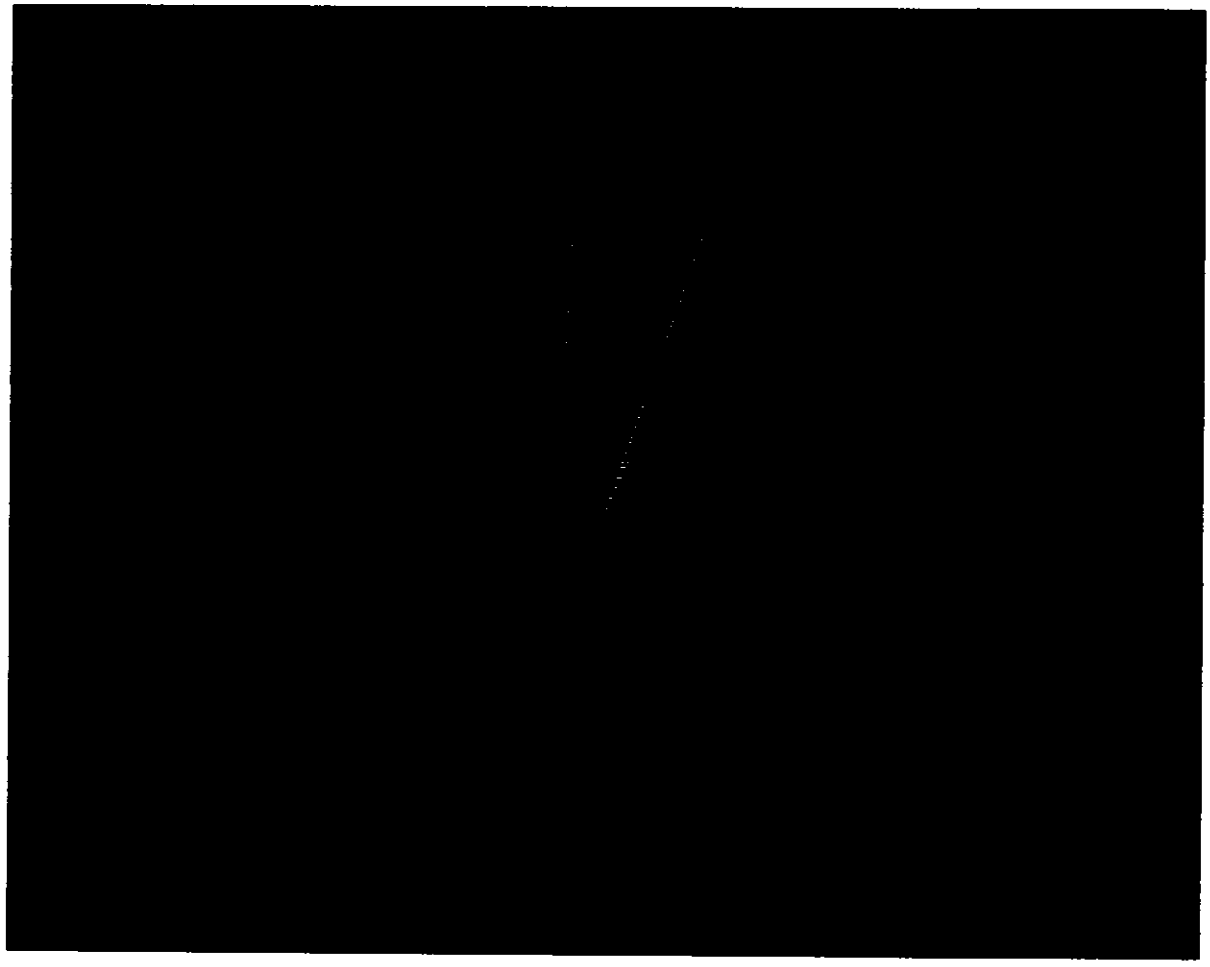
RECEPTOR →

LIGAND ↓

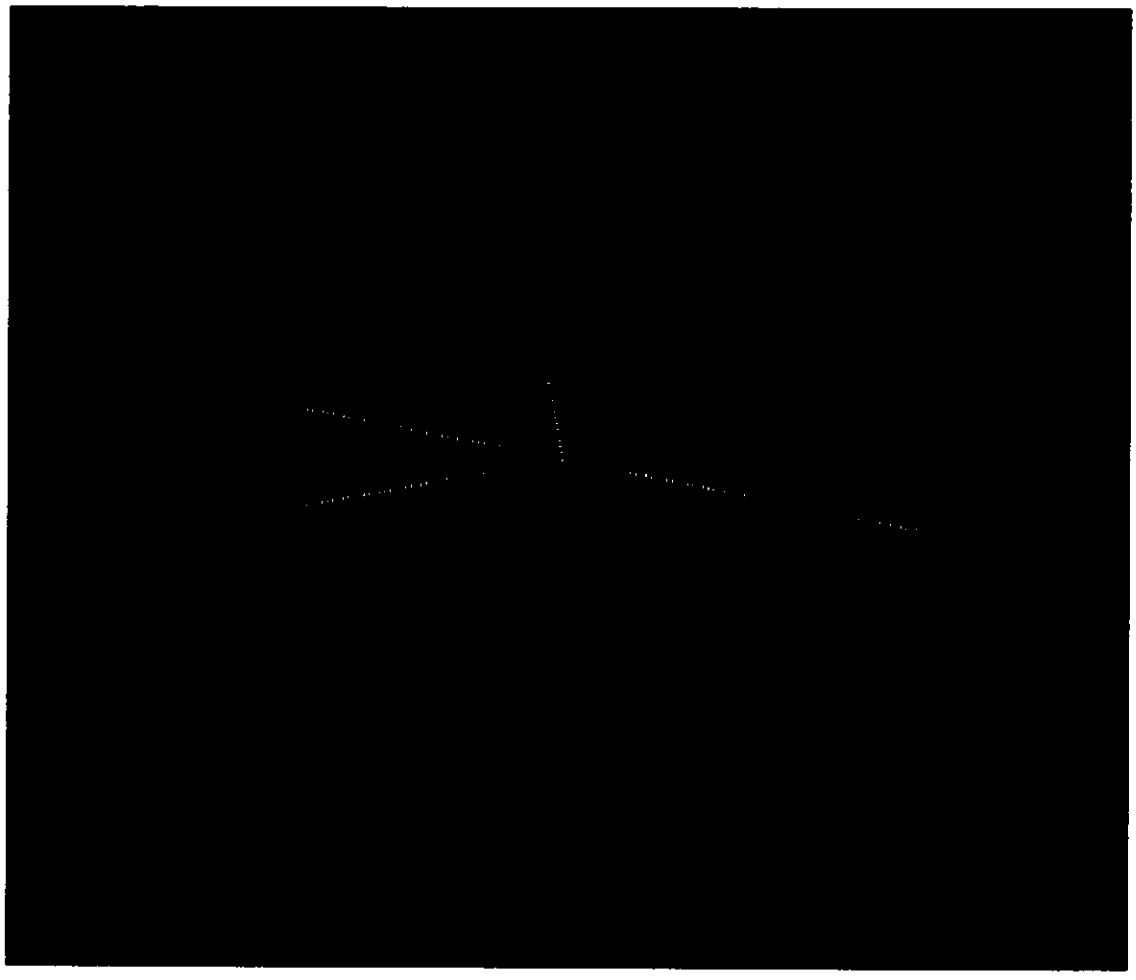
	GLN21	HIS22	THR23	VAL24	GLU66	GLY67	GLN68
ile4 CG2	-	-	-	Y CG1 6.04	-	-	-
leu8 CD1	-	-	-	Y CG1 3.14	-	-	-
ile17 CG2	-	-	-	Y CB 5.57	-	-	-
val18 CG2	-	-	-	Y CB 8.37	-	-	-
phe19 CZ	-	-	-	Y CB 6.82	-	-	-
asn20 ND2	-	-	Y HG1 6.36	-	-	-	-
his21	-	-		-	-	-	-
phe12	-	-	-	-	-	-	-
gln11	-	-	-	-	-	-	-
lys7 NZ	-	-	-	-	Y OE1 7.05	-	-

(First row shows possible interacting residue of PDZ-1 domain while first column shows residue of second highlighted peptide of V3 region, for receptor residues interacting side chains are mentioned in each cell while that for ligand are mentioned in first column followed by residue name and no. (Y) shows possibility of existence of interactions and (-) shows that there is exist no possibility of interaction b/w receptor and ligand or functional group lies outside 5.0 angstrom bound.)

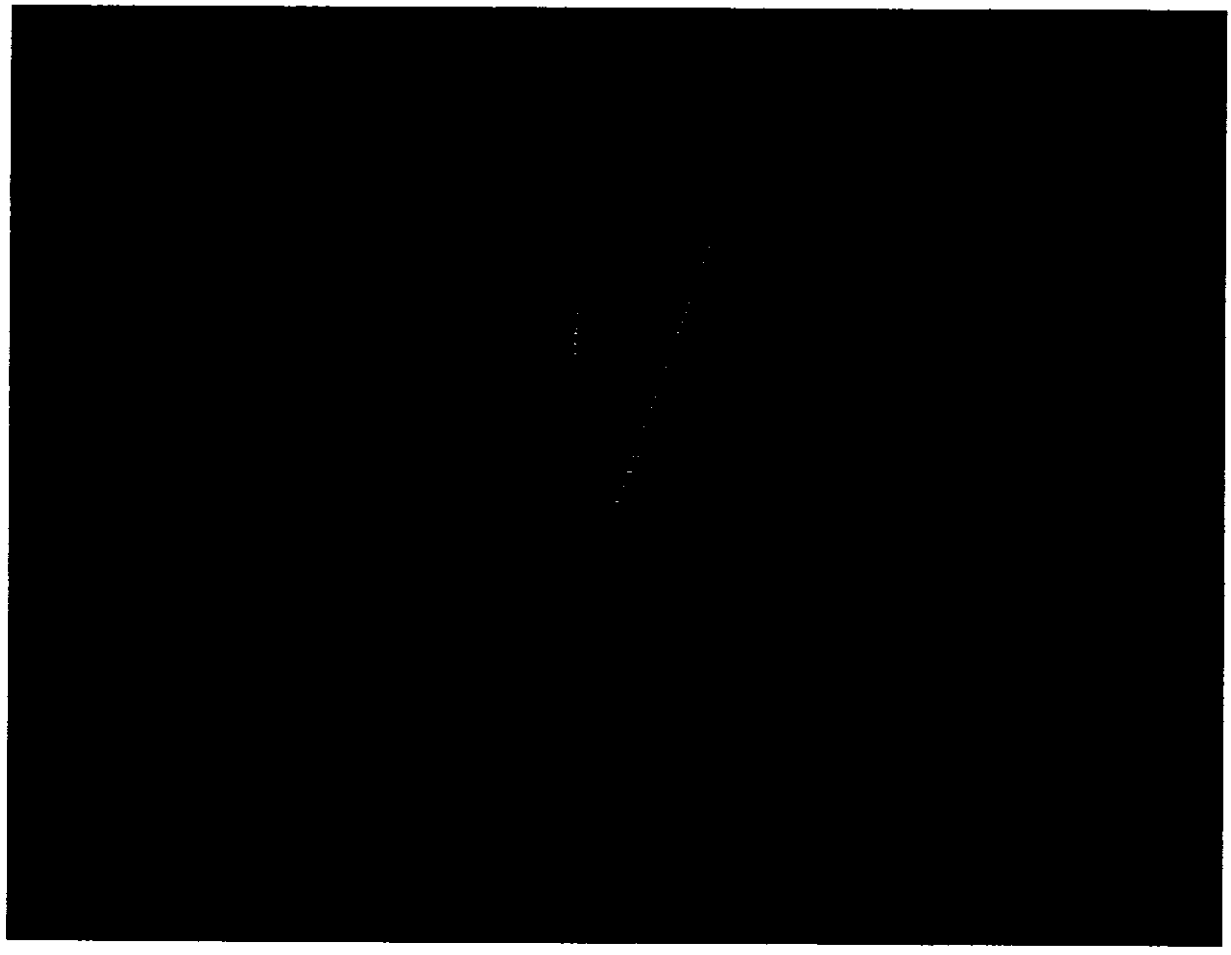
After analyzing tables 4.9 to 4.14, it is revealed that at point where V3 loop of gp120 protein shows distances less than 5.0 with ZO-1 protein residues lies in domain PDZ-1, PDZ-2, SH3-GUK domain, first and second highlighted peptide of V3 region. Gln 70 and asn 72 of PDZ-1 domain may interact with thr 22, thr 23 and ser 11 of V3 loop sharing distances 2.98, 4.42 and 4.65 respectively (shown in figures 4.26 & 4.27). Lys 209 of PDZ-2 shares distance of 3.84 angstrom with thr 23 of V3 loop (shown in figure 4.28). Whereas, thr 772 of SH3-GUK module share distance of 2.58 with thr 8 of V3 loop (shown in figure 4.29). Ile 1 of first highlighted peptide of V3 region share the distance of 4.36 and 3.95 angstrom with ile 18 and trp 19 of PDZ-1 domain respectively (shown in figure 4.30). Single residue leu8 of second highlighted peptide in V3 region shares the distance of 3.14 angstrom with val24 of PDZ-1 domain (shown in figure 4.31).



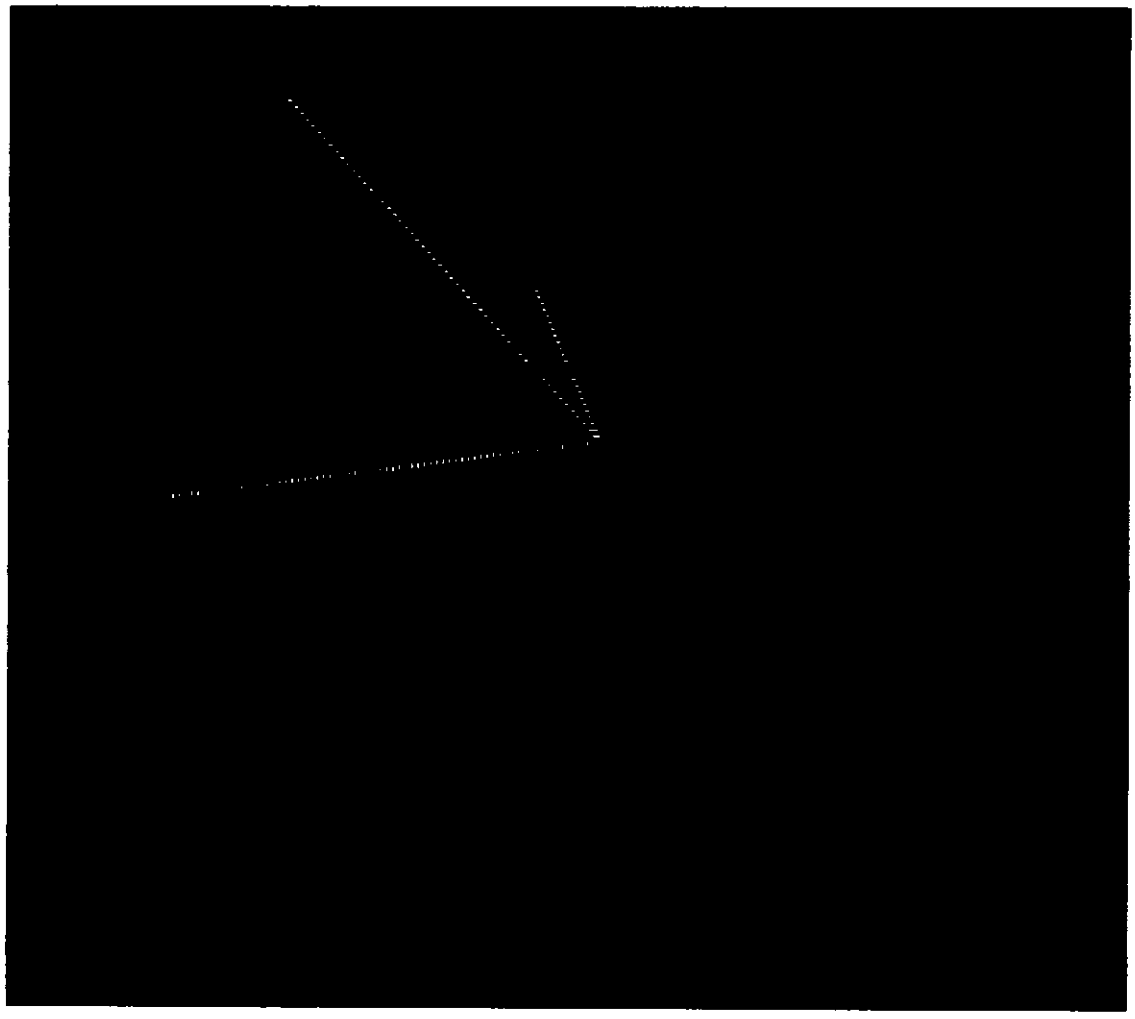
4.26 Distances between residues of PDZ-1 and V3 loop.



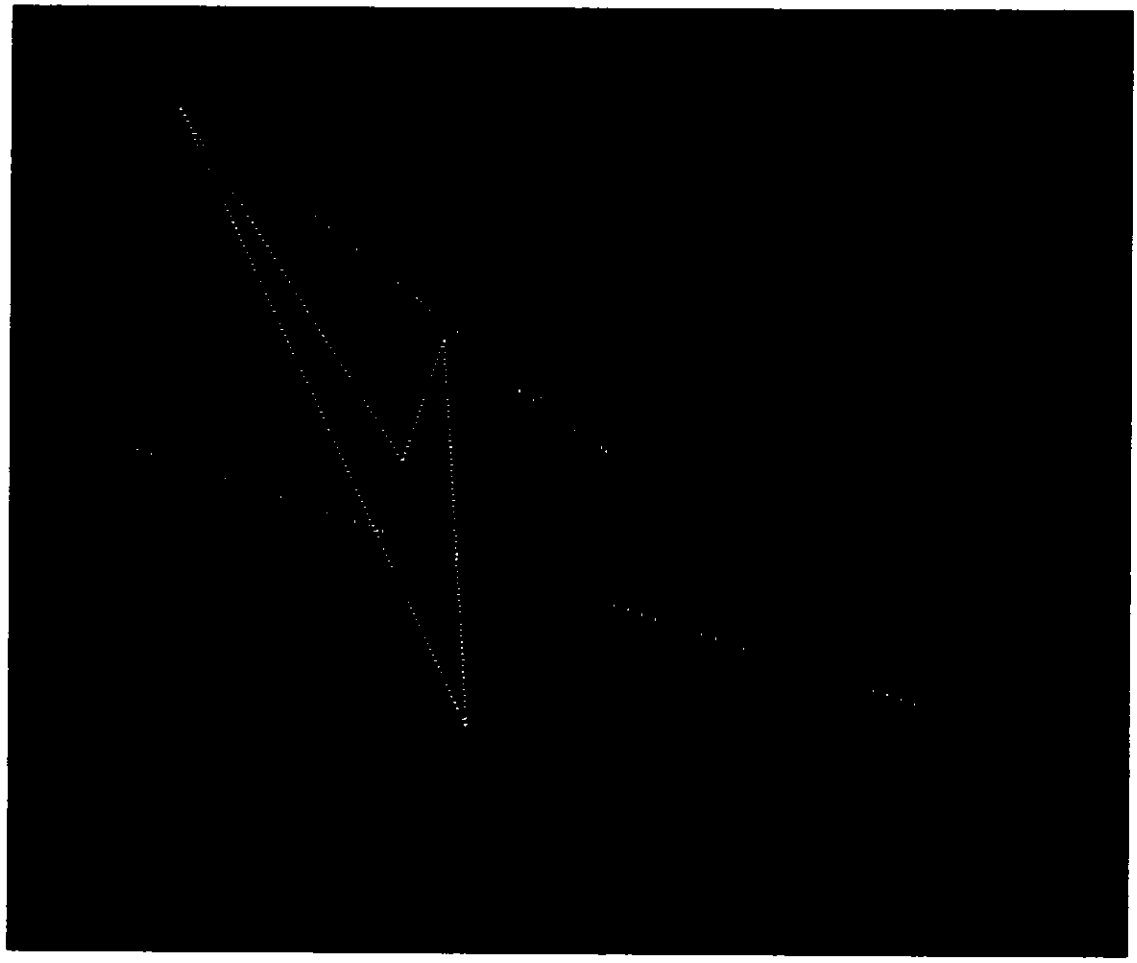
4.27 Distances between residues of PDZ-1 and V3 loop.



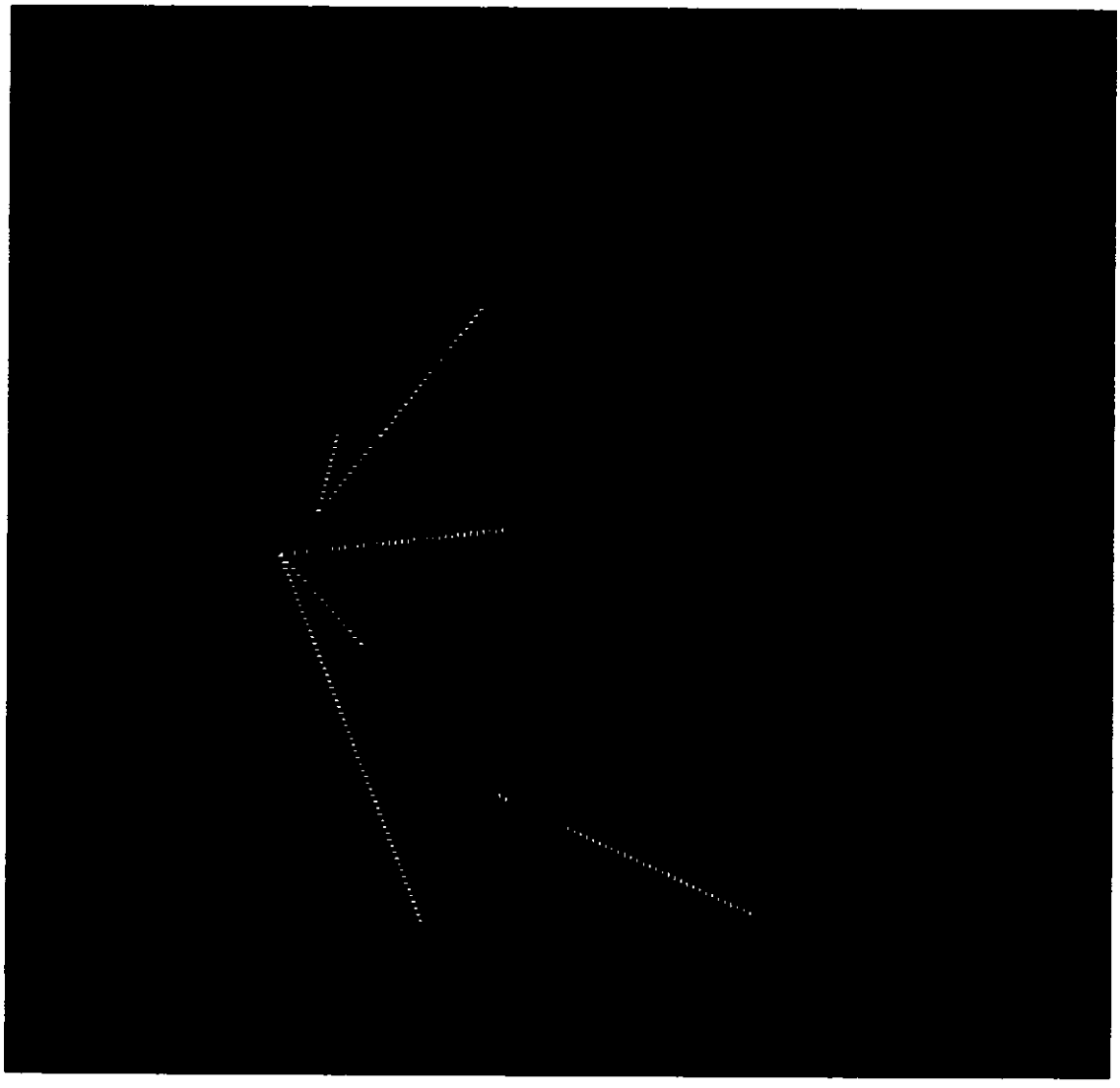
4.28 Distances between residues of PDZ-2 and V3 loop.



4.29 Distances between residues of SH3-GUK module and V3 loop.



4.30 Distances between residues of PDZ-1 and first highlighted peptide.



4.31 Distances between residues of PDZ-1 and second highlighted peptide.

4.7 Enzyme-Linked Immunosorbent Assay (ELISA)

The pictorial results of ELISA test performed to study the prevalence of HIV-associated dementia in HIV- infected population of Pakistani population is shown in figure. As obvious from visual results shown in figure 4.26, first column and two rows of second column shows gradual increase in yellow color that depicts the presence neuron specific enolase in sample. Since row 3 to 8 of second column contained the sample of non-HIV dementia patients, presence of yellow color in samples of these wells verifies the accuracy of results. Out of total 40 HIV patient's serum eight shows increased concentration of NSE, which means in Pakistan HAD prevails in non-ART HIV patients. But to deduce exact prevalence percentage statistical results were considered (shown in table 4.15 and figure 4.32). Exact statistical output pf ELISA in shown in page no. 110. In the last section of output page, values are showing concentration of neuron specific enolase in each sample. Total no of well columns of plate are shown on the top of table ranging from 1 to 12 while rows of plate are labeled using alphabets starting from A to H.

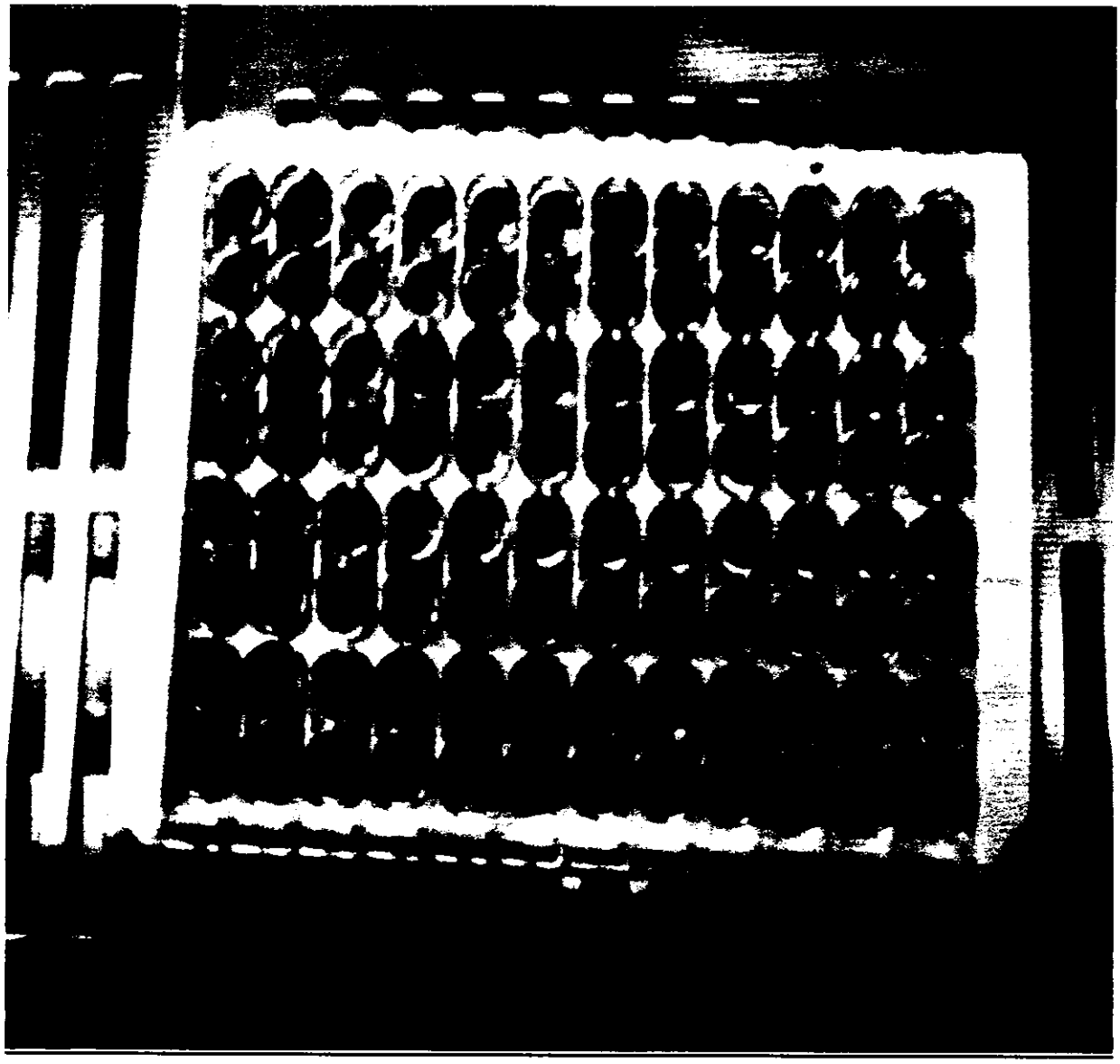


Fig 4.32 96 wells plate used to run ELISA test, yellow colored wells shows presence of neuron specific enolase in samples.

Table 4.15 Patient's ID and average concentration of NSE is mentioned.

Patient ID	Average Concentration	Patient ID	Average Concentration
4	-	24	25.56
5	0.1365	25	42.45
6	16.645	26	6.2495
7	0.039	27	0.4295
8	-	28	12.755
9	25.025	29	-
10	3.2425	30	0.4685
11	3.047	31	0.039
12	0.156	32	0.391
13	0.938	33	0.078
14	-	34	0.234
15	-	35	2.187
16	-	36	0.078
17	0.391	37	0.547
18	0.43	38	-
19	0.859	39	0.2345
20	11.15	40	0.703
21	0.8985	41	0.7815
22	21.245	42	0.000
23	0.586	43	0.3905

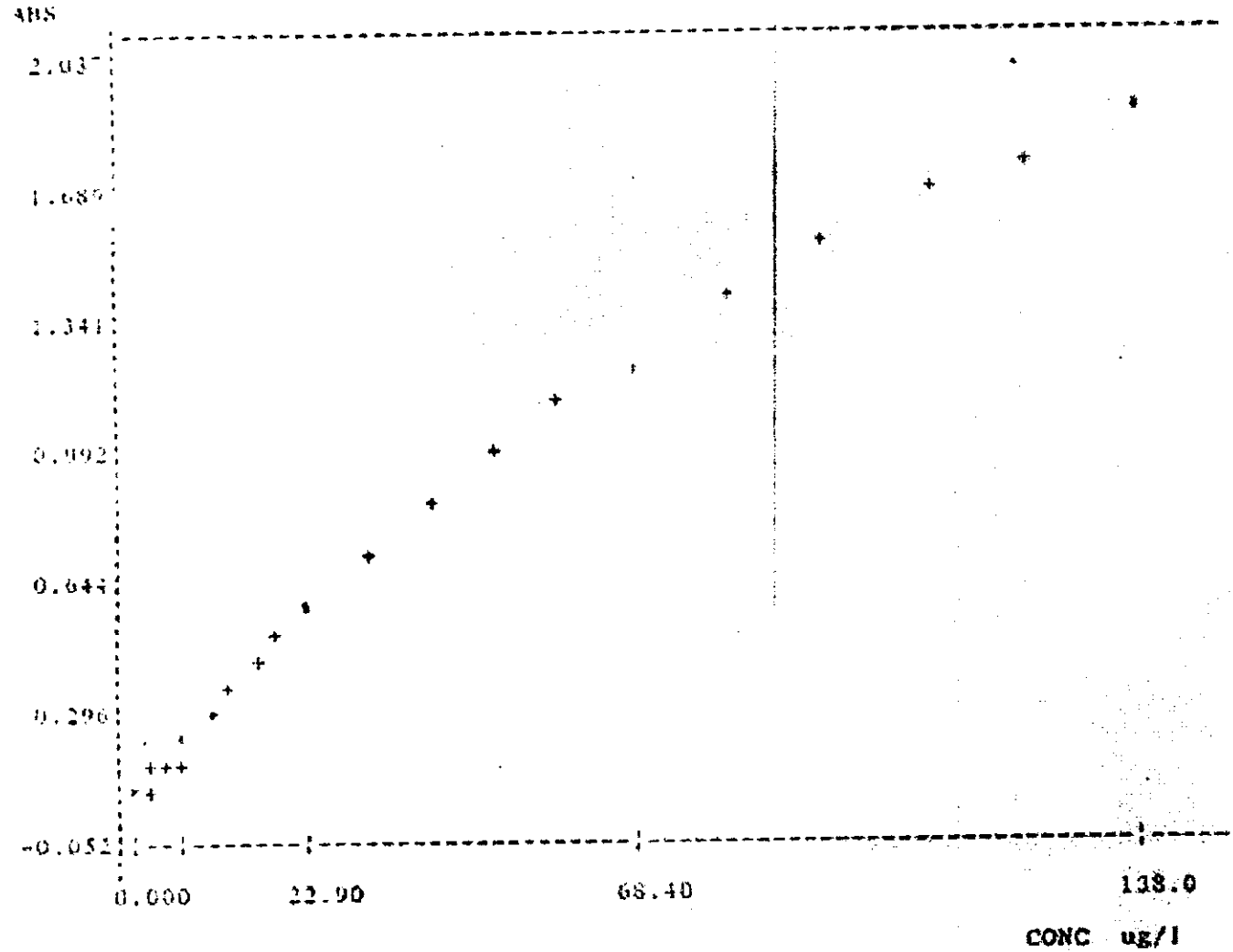
MULTISKAN EX CUBIC SPLINE V. 2.3

ABSORBANCE MODE
CONTINUOUS MOVEMENT
FILTER 405

CUBIC SPLINE CALCULATION

	ABS	CONC	UNIT	WELLS	
STD 1	0.071	0.000	ug/l	A1	B1
STD 2	0.218	7.500	ug/l	C1	D1
STD 3	0.541	22.90	ug/l	E1	F1
STD 4	1.236	68.40	ug/l	G1	H1
STD 5	1.914	138.0	ug/l	A2	B2

AXIS SCALE LIN/LIN



* = STANDARD POINT
+ = CURVE POINT

CONCENTRATIONS
16. AUG 2010 13:22:17

	1	2	3	4	5	6	7	8	9	10	11	12
A	****	****	****	****	****	****	10.07	19.45	14.16	****	0.078	****
B	0.312	136.9	****	****	0.156	****	12.23	31.67	11.35	0.391	0.078	0.703
C	7.266	12.79	0.234	21.94	****	****	0.078	39.96	****	0.156	0.547	0.938
D	7.741	10.71	0.039	28.11	0.938	0.391	1.719	44.94	****	0.000	****	0.625
E	22.96	9.024	15.84	2.344	****	0.469	22.96	7.187	0.312	0.312	****	0.000
F	22.82	11.83	17.43	4.141	****	0.391	19.53	5.312	0.625	0.156	****	****
G	65.56	3.516	0.039	3.125	****	****	0.625	0.625	0.039	****	0.078	0.469
H	71.30	3.750	****	2.969	****	0.859	0.547	0.234	****	2.187	0.391	0.312

DISCUSSION

Human immunodeficiency virus type 1(HIV-1) invade into central nervous system even in presence of intricate security in the form of blood brain barrier (BBB), a tightly packed mono layer of endothelial cells that separates peripheral blood from cerebro spinal fluid thus does not allows blood borne disease causing agents to enter inside brain (Kaul *et al.*, 2009). If some blood borne disease causing agent is able to infect brain then it must adopt some mechanism to breach this convoluted barrier. Incase of neuroAIDS, HIV-1 is able to breach blood brain barrier, two most supported theories with strong evidences are;

1. Trojan horse theory.
2. HIV-1 products, i.e., Gp120 and Tat proteins degrades blood brain barrier and leaves an open path for HIV-1 virions to enter freely inside brain (Sharma *et al.*, 2009).

Recently in a study by Nakamuta *et al.*, 2008, it is revealed that HIV-1 envelop protein Gp120 is involved in proteasomal degradation of Zonula Occluden-1 protein, a scaffold tight junction protein considered to play central role in sustenance of tight junctions in blood brain barrier via unknown mechanism. However, another HIV protein Vif is involved in proteasomal degradation of an innate immune system protein called

apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3 (APOBEC3G) and it's mode of action is well elucidated. It comprises of following major steps;

1. Hiv-1 Vif interacts with huAPOBEC3G. The residues of APOBEC3G required for interaction with Vif are: Aspartic acid (128), Proline (129) and Aspartic acid (130).
2. Thereby recruits a cullin 5-elongin B/C E3 ubiquitin ligase.
3. A3G is poly-ubiquitinated.
4. Finally targeted to proteasome for degradation.

Thereby, the first step for proteasomal degradation is marking of APOBEC3G protein by HIV-1 Vif protein (Huthoff *et al.*, 2007).

Macromolecular docking is a technique used in bioinformatics to computationally model the quaternary structure of complexes formed by two or more interacting biological macromolecules. Protein–protein complexes are the most commonly attempted targets of such modelling, followed by protein–nucleic acid complexes. Considering the complexity of process only few sofwares provides the facility of macromolecular docking, Hex 5.1 has been reported in a research paper by Chumchua *et al.*, 2008 to successfully dock *Mycoplasma pneumoniae enolase* with human *plasminogen*. Therefore, in present study Hex 5.1 is explored to dock ZO-1 protein with Gp120 protein.

Detailed study of docking results revealed that Gp120 protein of HIV-1 may bind or interact with Zonula occluden-1 protein. For strong binding the distance between two atoms should be less than 4.0 angstrom and the distance ranging from 4.0 to 5.0 angstrom shows weak binding, beyond this range does not exist the chances of interaction. Since in tables ranging from 7 to 12 distances of all possible interacting residues were mentioned but for further evaluation only distances ranging within 5.0 angstrom are considered. This is because of possibility that V3 loop of Gp120 protein may not bind with ZO-1 but only interact to mark it for degradation as it happens incase when Vif protein of HIV-1 lead to the proteasomal degradation of APOBEC3G protein of innate immune system.

Out of four domains namely PDZ-1, PDZ-2, PDZ-3, SH3-GUK and two peptides of V3 region labeled as first highlighted peptide and second highlighted peptide, both two peptides and PDZ-1, PDZ-2 and SH3-GUK module's residue shows close interaction with PDZ-1 domain and V3 loop respectively. For establishing either presence of binding or interaction it is feasible to have two or more than two residues to interact with target molecule since only single residue of PDZ-2, SH3-GUK module and second highlighted peptide shows distance less than 5.0 angstrom with residue of V3 loop and PDZ-1 domain respectively therefore, PDZ-2, SH3-GUK module and second highlighted peptide does not interact or bind with their respective targets.

Two residues of PDZ-1 domain namely gln70 and asn72 interact with three residues of V3 loop, i.e., thr22, thr23 and ser11. Since Gln70 is at a distance of 2.98 angstrom from thr22 and 4.42 angstrom from thr23 after docking therefore gln70 forms a strong hydrogen bond with thr22 while weak hydrogen bond exist between gln70 and thr23. As asn72 shares the distance of 4.65 angstrom with ser11 of V3 loop after docking so it means that there exist a weak hydrogen bond between them.

Zonula occluden-1's PDZ-1 domain is a protein binding domain with established binding profile (Appleton *et al.*, 2006)

[R/K/S/T][T/S][W/Y][V/I/L]COOH

So if PDZ-1 domain has to bind with V3 loop then target must possess this motif, whole neurotropic gp120 with accession number AAK09421.1 was manualy analysed for this motif and it showed zero existance of whole exact motif however interestingly at several places this motif was found with the exception of precense of mutation at tird position from left of motif as shown in figure 33. To evaluate the importance of third residue of motif two peptides from V3 region with accession number AAB05002.1 were modeled, evaluated and docked these peptides contained these so called binding motifs in it as shown below;

VV[redacted]AINCTRPNNN[redacted]IGNIRQAHCNLSRTKWE
NT[redacted]

Grey highlighted regions are first and second highlighted peptides in sequence with yellow residues showing so called binding motifs and red highlighted peptides are V3 loop.

Results showed that this third residue in peptides must be either tryptophan or lysine for the binding with PDZ-1 domain as none of the peptide bonded with PDZ-1 domain, thus this study confirms the previously reported studies illustrating the importance of third residue from left of binding motif in binding (Appleton *et al.*, 2006).

Since V3 loop does not contain binding profile of PDZ-1 domain and distances are showing two week hydrogen bonds and one strong hydrogen bond therefore this study propose that PDZ-1 domain interact not bind with V3 loop and via this mechanism ZO-1 protein gets marked with gp120 protein which ultimately leads to proteasomal degradation. Residues gln 70 and asn 72 of PDZ-1 domain interact with thr 22, thr 23 and ser 11 of V3 loop sharing distances 2.98, 4.42 and 4.65 respectively. Also there are chances that gp120 protein marks ZO-1 protein via interacting at multiple positions as first highlighted peptide in V3 region also shows distances less than 5.0 angstrom after docking with PDZ-1 domain. Ile 1 of first highlighted peptide of V3 region shares the distances of 4.36 and 3.95 angstrom with ile 18 and trp 19 of PDZ-1 domain respectively.

The wet lab part of whole study was successfully completed to reveal the statistics of HIV & AIDS patients having HIV associated dementia in Pakistan population. Fig. 4.27 shows complete statistical results for ELISA test performed at National AIDS Control Program's referral laboratory. 13.000 is taken as cut-off value, samples having values above this are considered to possess greater than normal concentration of neuron specific enolase in their serum. Since each serum sample was poured twice therefore average value of concentrations for each patients ID is shown in table 4.15.

Since from total forty samples five have average concentration greater than cut-off value 13 therefore, it is estimated that prevalence of HIV-associated dementia in HIV & AIDS patients of Pakistani population is 12.5% which is far greater than neighboring country India's prevalence rate of 1-2% and meets with westerns countries 10-24%. So during the treatment of HIV patient physician should also focus in detail on mental health of patient due to high prevalence rate of HAD in Pakistan.

CONCLUSION & FUTURE WORK

This study was focused on understanding molecular mechanism of HIV neuroinvasion and in determining prevalence rate of HIV-associated dementia in non-ART HIV infected Pakistani population. It is concluded that residues gln 70 and asn 72 of PDZ-1 domain interact with thr 22, thr 23 and ser 11 of V3 loop sharing distances 2.98, 4.42 and 4.65 respectively therefore, it marks ZO-1 protein which then leads to proteasome-mediated degradation. There is a strong probability that there are multiple points of interactions since Ile 1 of first highlighted peptide of V3 region shares the distances of 4.36 and 3.95 angstrom with ile 18 and trp 19 of PDZ-1 domain respectively.

It is also derived that prevalence rate of HAD in Pakistan is 12.5% which is far greater than neighboring country India's prevalence rate of 1-2% and meets with westerns countries 10-24%. So during the treatment of HIV patient physician should also focus in detail on mental health of patient due to high prevalence rate of HAD in Pakistan.

Dry lab part of work should be subjected to site directed mutagenesis technique of wet lab for further confirmation of results; this would be great contribution in understanding molecular mechanisms lying within neuroinvasion of HIV.

REFERENCES

1. Alan SF and James MA (2009). Zonula Occludens-1 and -2 Are Cytosolic Scaffolds That Regulate the Assembly of Cellular Junctions. Molecular Structure and Function of the Tight Junction. Annals of the New York Academy of Sciences 1165: 113–120
2. Appleton BA, Zhang, Wu P, Yin JP, Hunziker W, Skelton NJ, Sidhu SS and Wiesmann C (2006). Comparative Structural Analysis of the Erbin PDZ Domain and the First PDZ Domain of ZO-1. The Journal of Biological Chemistry 281: 22312–22320
3. Burton GF, Keele BF, Estes JD, Thacker TC and Gartner S. (2002). Follicular dendritic cell contributions to HIV pathogenesis. Seminars in Immunology 14 (4): 275–284
4. Clapham PR and McKnight A (2001). HIV-1 receptors and cell tropism. British Medical Bulletin 58 (4): 43–59
5. Clarke JR, White NC, and Weber JN (2000). HIV Compartmentalization: Pathogenesis and Clinical Implications. AIDS Reviews 2: 15-22
6. Decker CF and Lazarus A (2000). Tuberculosis and HIV infection. How to safely treat both disorders concurrently. Postgraduate Medicine 108 (2): 57–60, 65–68

7. Douek DC, Roederer M, Koup RA (2009). Emerging concepts in the immunopathogenesis of AIDS. Annual Review of Medicine 60: 471–84
8. Engelhardt B and Sorokin L (2009). The blood-brain and the blood cerebrospinal fluid barrier: function and dysfunction. Seminars in Immunopathology 31:497–511
9. Feldman C (2005). Pneumonia associated with HIV infection. Current Opinion in Infectious Diseases 18 (2): 165–170
10. Ghafouri M, Amini S, Khalili K and Sawaya BE (2006). HIV-1 associated dementia: symptoms and causes. Retrovirology 3(28)
11. Grabar S, Selinger-Leneman H, Abgrall S, Pialoux G, Weiss L and Costagliola D. (2009). Prevalence and comparative characteristics of long-term nonprogressors and HIV controller patients in the French Hospital Database on HIV. AIDS 23 (9): 1163–1169
12. Grant I, Sacktor H, and McArthur J (2005). HIV neurocognitive disorders. The Neurology of AIDS (2nd Ed.). London, UK: Oxford University Press 357–373
13. Gras G and Kaul M (2010). Molecular mechanisms of neuroinvasion by monocytes-macrophages in HIV-1 infection. Retrovirology 7:30
14. Gray F, Adle-Biassette H, Chrétien F, Lorin de la Grandmaison G, Force G and Keohane C. (2001). Neuropathology and neurodegeneration in human immunodeficiency virus infection. Pathogenesis of HIV-induced lesions of the

brain, correlations with HIV-associated disorders and modifications according to treatments. Clinical Neuropathology 20 (4): 146–155

15. Huthoff H and Malim M H (2007). Identification of amino acids residues in APOBEC3G required for regulation by human immunodeficiency virus type 1 vif and virion encapsidation. Journal of Virology 81:3807-3815
16. Janssen RS, Nwanyanwu OC, Selik RM and Stehr-Green JK (1992). Epidemiology of human immunodeficiency virus encephalopathy in the United States. Neurology 42(8):1472-1476
17. Kaul M (2009). HIV-1 associated dementia: update on pathological mechanisms and therapeutic approaches. Current Opinion in Neurology. 22: 315–320
18. Kahn JO and Walker BD (1998). Acute Human Immunodeficiency Virus type 1 infection. New England Journal of Medicine 331 (1): 33–39
19. Kustrzeba-Wojcicka I, Golczak M (2000). Enolase from *Candida albicans* purification and characterization. Comparative Biochemistry and Physiology Part B 126: 109–120
20. Lee SY, Kim YC, Kim WJ (2010). Serum neuron-specific enolase level as a biomarker in differential diagnosis of seizure and syncope. Journal of Neurology 257(10): 1708-1712
21. Lequin RM (2005). Enzyme immunoassay (EIA)/enzyme-linked immunosorbent assay (ELISA). Clinical Chemistry 51(12):2415-8

22. Lye MF, Fanning AS, Sul Y, Anderson JM, and Lavie A (2010). Insights into regulated ligand binding sites from the structure of zo-1 sh3-guanylate kinase module. JBC Papers in Press, Manuscript M109.093674

23. Murre, Sturdy JM, DP (1995). The connectivity of the brain: multi-level quantitative analysis. Biological cybernetics 73 (6): 529–45

24. Nakamuta S , Endo H , Higashi Y , Kousaka A , Yamada H, Yano M and Hiroshi K (2008). Human immunodefeciency virus type1 gp120-mediated disruption of tight junction proteins by induction of proteasome-mediated degradation of zonula occludens-1 and zonula occludens-2 in human bain microvascular endothelial cells. Journal of Neurovirolog 14: 186-195

25. Pantaleo G, Demarest JF, Schacker T, Vaccarezza M, Cohen OJ, Daucher M, Graziosi C, Schnittman SS, Quinn TC, Shaw GM, Perrin L, Tambussi G, Lazzarin A, Sekaly RP, Soudeyns H, Corey L and Fauci AS. (1997). The qualitative nature of the primary immune response to HIV infection is a prognosticator of disease progression independent of the initial level of plasma viremia. Proceedings of the National Academy of Sciences of the USA 94: 254–258

26. Piatak M, Jr, Saag, MS, Yang LC, Clark S J, Kappes JC, Luk KC, Hahn BH, Shaw GM and Lifson JD (1993). High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. Science 259 (5102): 1749–1754

27. Piast M, Kustrzeba-Wojcicka I, Matusiewicz M and Banas T (2005). Molecular evolution of enolase. Acta biochimica polonica 52(2) : 507–513

28. Pollok RC (2001). Viruses causing diarrhoea in AIDS. Novartis Foundation Symposium 238: 276–83

29. Reger M, Welsh R, Razani J, Martin DJ and Boone KB (2002). A meta-analysis of the neuropsychological sequelae of HIV infection. Journal of International Neuropsychological Society 8(3):410-424

30. Satishchandra P, Nalini A, Gourie-Devi M, Khanna N, Santosh V, Ravi V, Desai A, Chandramuki A, Jayakumar PN, and Shankar SK (2000). Profile of neurologic disorders associated with HIV/AIDS from Bangalore, South India (1989-96). Indian Journal of Medical Research 11: 14–23

31. Selakovic V, Raicevic R and Radenovic L (2005). The increase of neuron-specific enolase in cerebrospinal fluid and plasma as a marker of neuronal damage patients with acute brain infarction. Journal of Clinical Neuroscience 12(5): 542–547

32. Sharma D and Bhattacharya J (2009). Cellular and molecular basis of HIV-associated neuropathogenesis. Indian Journal of Medical Research 637-651

33. Smith, Johanna A, Daniel, and Rene (2006). Following the path of the virus: the exploitation of host DNA repair mechanisms by retroviruses. ACS Chemical Biology 1 (4): 217–26

34. UNAIDS, WHO (2007). 2007 AIDS epidemic update

35. Wadia RS, Pujari SN, Kothari S, Udhar M, Kulkarni S, Bhagat S, and Nanivadekar A (2001). Neurological manifestations of HIV disease. Journal of Association of Physicians of India 49: 343–348

36. Weiss RA (1993). How does HIV cause AIDS? Science 260 (5112): 1273–1279

37. Zaidi SA and Cervia JS (2002). Diagnosis and management of infectious esophagitis associated with human immunodeficiency virus infection. Journal of the International Association of Physicians in AIDS Care 1(2): 53–62