Identifying Chemicals with Potential Therapy of HIV Based on Protein-Protein and Protein-Chemical Interaction Network

Bi-Qing Li[2,9,¤], Bing Niu[3,9], Lei Chen[4], Ze-Jun Wei[5], Tao Huang[6], Min Jiang[5], Jing Lu[7], Ming-Yue Zheng[7], Xiang-Yin Kong[5]*, Yu-Dong Cai[1]*

1 Institute of Systems Biology, Shanghai University, Shanghai, P. R. China, 2 Key Laboratory of Systems Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, P. R. China, 3 College of Life Science, Shanghai University, Shanghai, P. R. China, 4 College of Information Engineering, Shanghai Maritime University, Shanghai, P. R. China, 5 The Key Laboratory of Stem Cell Biology, Institute of Health Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, P. R. China, 6 Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York, United States of America, 7 Drug Discovery and Design Center (DDDC), Shanghai Institute of Materia Medica, Shanghai, P. R. China

Abstract

Acquired immune deficiency syndrome (AIDS) is a severe infectious disease that causes a large number of deaths every year. Traditional anti-AIDS drugs directly targeting the HIV-1 encoded enzymes including reverse transcriptase (RT), protease (PR) and integrase (IN) usually suffer from drug resistance after a period of treatment and serious side effects. In recent years, the emergence of numerous useful information of protein-protein interactions (PPI) in the HIV life cycle and related inhibitors makes PPI a new way for antiviral drug intervention. In this study, we identified 26 core human proteins involved in PPI between HIV-1 and host, that have great potential for HIV therapy. In addition, 280 chemicals that interact with three HIV drugs targeting human proteins can also interact with these 26 core proteins. All these indicate that our method as presented in this paper is quite promising. The method may become a useful tool, or at least plays a complementary role to the existing method, for identifying novel anti-HIV drugs.

Introduction

Human immunodeficiency virus (HIV) is a lentivirus belonging to retrovirus family that causes acquired immunodeficiency syndrome (AIDS) [1,2]. The global HIV and AIDS pandemic has caused nearly 60 million infections. Experts estimate that more than 25 million people have died of AIDS, and more than 33 million presently are living with HIV infection or AIDS [3].

During the last decade, the specific functions of HIV-1 encoded genes and related proteins have been extensively studied, which facilitated the development of the effective approved anti-AIDS drugs directly targeting the HIV-1 encoded enzymes, including reverse transcriptase (RT), protease (PR) and integrase (IN) [4,5]. Despite the great efforts in developing new effective antiviral agents and the introduction of combination of these drugs, namely highly active antiretroviral therapy (HAART), the incidence of HIV infections continues to rise, because of the rapid emergence of drug-resistant HIV-1 mutants as well as the severe side effects. Therefore, there is an urgent need for further improvement of the existing anti-HIV drugs [6]

and the introduction of novel drug design strategies [7] or novel antiviral targets with therapeutic potential for HIV infection [8].

Recently, it has been reported that several human proteins that were involved in HIV-1 life cycle and interactions with HIV-1 encoded proteins emerged as novel anti-HIV drug targets, including TSG101 [9], NF-κB [10], positive elongation factor P-TEFb [11] and cellular factors related to nuclear import of pre-integration complex [12]. Besides, small-molecule inhibition of the direct protein-protein interactions (PPI) that mediate numerous critical biological processes is an emerging area in current drug discovery [13,14,15,16].

Multiple PPI involved in many biological processes in the HIV-1 life cycle have been identified by genomics, proteomics and biochemical approaches recently [17,18,19]. Although most of these interactions are complicated and have not yet been fully investigated, current knowledge on the molecular interactions has significantly broadened the understanding of the HIV-1 life cycle and paved an new way for the anti-HIV drug development. In fact, there is an increasing number of examples of both chemical and biological small molecular HIV inhibitors targeting PPI emerging nowadays [20].

In this study, we compiled all the PPI from HIV-1, Human Protein Interaction Database [17,18,19]. A PPI network was constructed with all these human proteins based on STRING [21] and 26 of them with a score greater than 1000 were selected according to their betweenness. Then, 280 chemicals in STITCH [22] that can interact with three HIV drugs targeting human protein were identified. It has been shown that these 280 chemicals can also interact with the 26 core human proteins. Therefore, the 280 chemicals and 26 human proteins may possess the potential for HIV therapy. Our method may open a new way for HIV drug design or at least plays a complementary role to the existing method.

Materials and Methods

HIV-1, Human Protein Interaction Data

All the protein-protein interactions (PPI) data were retrieved from the HIV-1, Human Protein Interaction Database (http://www.ncbi.nlm.nih.gov/RefSeq/HIVInteractions/) [17,18,19]. It includes 5,126 PPI and involves 19 HIV-1 proteins corresponding to 9 HIV-1 genes as well as 1,450 human proteins corresponding to 1,431 human genes. The PPI data was given in Additional File S1.

Protein-Protein Interaction (PPI) Network

There are two PPI database: STRING (http://string-db.org/) [21] and HPRD (http://hprd.org/) [23]. The reasons why we chose STRING over HPRD are as following:

1). The STRING database includes more PPIs than HPRD. So far HPRD only contains 41,327 experiment supported PPI, while STRING contains 1,640,707 PPI including both direct ones (physical interactions) and indirect ones (functional interactions). HRPD is more likely to be a subset of STRING, since STRING includes the PPIs from experiments, existing databases, text-mining and predicted results.

2). The possible problems that the predicted PPIs with low confidence in STRING would cause can be avoided in our method. Since we used the weighted PPIs of STRING rather than the binary ones, the confidence of each PPI is considered. If a PPI has low confidence, it will be less important in Dijkstra's algorithm during the shortest path analysis, and most likely to be eliminated.

3). The PPIs in HRPD is not weighted. Therefore, it is difficult to do quantitative network analysis. Overall, we selected STRING to construct the PPI network. Each interaction in STRING is evaluated by an interaction confidence score in range from 1 to 999 to quantify the likelihood that an interaction may occur. For clarity, let $Q(p1,p2)$ denote the interaction confidence score of two proteins p1 and p2. The constructed network took proteins as its nodes, and the edge between any two nodes existed if and only if the corresponding proteins can interact with each other. To reflect the difference of interactions, each edge with endpoints v1 and v2 in the network was labeled with a score as the edge weight as follows:

$$W(v1,v2) = 1,000 - Q(p1, p2) \qquad (1)$$

where p1 and p2 were corresponding proteins of nodes v1 and v2, respectively.

Chemical-chemical Interactions and Protein-chemical Interactions

The data of chemical-chemical interactions and protein-chemical interactions was retrieved from STITCH (version 3.0) (http://stitch.embl.de/) [22], a well-known database containing 1,430,424 known or predicted chemical-chemical interactions

between 89,617 chemicals as well as 1,221,559 protein-chemical interactions between 16,721 proteins and 234,826 chemicals deriving from experiments, literature or other reliable sources. Five scores with titles "Similarity", "Experimental", "Database", "Textmining" and "Combined_score" in range from 1 to 999 were used to indicate the interactivity of two chemicals or a protein-chemical pair. Since the last score combines the information of others, it was used as the final interaction score.

Shortest Path and Betweenness

For the given node in a network, its betweenness is related to the number of the shortest paths connecting all pair of nodes such that the node is the member of them [24]. For the node in PPI network, its betweenness accounts for direct and indirect influences of proteins at distant network [25]. Hence, betweenness has been used for study various natural and man-made networks [24,26,27,28]. However, it is not necessary to calculate each node's betweenness and consider all shortest paths. Here, we proposed a new kind of betweenness, named as betweenness related to A, where A was a node subset in a network. For this kind of betweenness, we only calculated the betweenness of the node in A not all nodes in the network. For a node d in A, its betweenness related to A, denoted by BA(d), was calculated by the following two steps: (1) Find shortest paths connecting all pair of nodes in A; (2) Count the number of the shortest paths such that d was the member of them.

Results and Discussion

26 Core Human Proteins Identified According to their Betweenness

In our work, a protein-protein interaction network was constructed for the 1,450 HIV interacting proteins based on STRING. For each of the 1,450 proteins, its betweenness can be calculated according to the method in "Materials and methods". In details, 1,050,525 shortest paths were found to calculate the betweenness related to 1,450 proteins. If a node appears in more than 0.1% of these shortest paths, it is deemed to be more important than other nodes. Thus we selected 26 proteins with betweenness greater than 1000, which were listed in Table 1. All the betweennesses for these 1,450 proteins were given in Additional File S2.

22 of these 26 proteins are well known to directly interact with HIV proteins in previous studies (Table 1). Their interactions include inhibition, activation, cleavage, degradation and so on (see Additional File S1), which should be deemed as causative. The rest four proteins may act as infection related, such as EGFR, which was upregulated by HIV-1 Gag protein. According to the roles played by these causative proteins during HIV life cycle, we briefly classified them into three groups, which respectively take part in receptor interaction, transaction and replication, and host immune response.

Within them, CD4 (cluster of differentiation 4) and CCR5 (C-C chemokine receptor type 5) are acting as co-receptors for HIV entry into targeting cells [29]. CD4 is a glycoprotein expressing on the surface of many kinds of immune cells such as macrophages, monocytes and T-help cells, and dendritic cells. It is recognized as the primary co-receptor of HIV targeting. It interacts with the viral envelope glycoprotein (Env) to trigger a structural alterations in Env and enable the virus to recruit other co-receptors, like CCR5 or CXCR4 [30]. The chemokine receptors CCR5, member of the seven-transmembrane G protein-coupled receptor superfamily, is one of the principal co-receptors for majority HIV isolates. It interacts with HIV protein gp120 so that HIV gp41

Table 1. 26 core human proteins identified by betweenness in shortest path.

Ensembl protein ID	Gene symbol	Betweenness
ENSP00000269305*	TP53	4611
ENSP00000263253*	EP300	4581
ENSP00000264657*	STAT3	3613
ENSP00000339007*	GRB2	2822
ENSP00000384273*	RELA	2762
ENSP00000226730*	IL2	2738
ENSP00000344818*	UBC	2518
ENSP00000360266*	JUN	2036
ENSP00000270202	AKT1	2032
ENSP00000275493	EGFR	1961
ENSP00000354394*	STAT1	1929
ENSP00000011653*	CD4	1796
ENSP00000229135*	IFNG	1434
ENSP00000353483*	MAPK8	1287
ENSP00000344456*	CTNNB1	1275
ENSP00000292303*	CCR5	1243
ENSP00000350941*	SRC	1161
ENSP00000341189*	PTK2	1151
ENSP00000348461*	RAC1	1147
ENSP00000329623*	BCL2	1123
ENSP00000329357*	SP1	1108
ENSP00000380227*	ITGA4	1076
ENSP00000226574*	NFKB1	1073
ENSP00000343204*	JAK1	1073
ENSP00000401303	SHC1	1063
ENSP00000228307	PXN	1053

*Directly interact with HIV proteins.
doi:10.1371/journal.pone.0065207.t001

protein's shape were changed to penetrate the cell membrane [31]. A natural mutant CCR5Δ32 (32 base pair deletion) can provide highly protection in HIV infected individuals in homozygous state [32,33]. Besides, the small guanosine triphosphate hydrolase (GTPase) Rac1 (ras-related C3 botulinum toxin substrate 1) is reported to positively regulates co-receptor CXCR4 function [34].

The proteins mainly related to HIV transaction activity and replication in our result include TP53, EP300, STAT1, STAT3, GRB2, NF-κB complex subunit, polyubiquitin-C, Akt-1, interferon gamma, MAPK8, beta-catenin, SRC1, SP1, Bcl-2. Cellular tumor antigen p53 (TP53), a tumor suppressor participating in multiple pathway like cell cycle arrest or apoptosis, interacts with HIV-1 viral infectivity factor (Vif) to mediate G2 cell cycle arrest with a positive effect on HIV-1 replication [35]. Histone acetyltransferase p300 and SP1 interact with HIV-1 Viral protein-R (Vpr) to mediate Vpr activity in virion assemble, nucleus locating, promoter activation, cell cycle arrest or apoptosis induction [36]. In addition, Histone acetyltransferase p300, GRB2, Polyubiquitin-C, Akt-1, MAPK8 are all involved in the HIV trans- activating protein Tat mediated transactivation of HIV-1 LTR and viral replication, respectively [37,38,39,40,41]. Several proteins are to function by the similar pathway, such as the NF-κB signaling pathway or JAK-STAT pathway. Nuclear factor (NF)-κB complex is a master regulator of pro-inflammatory genes and is upregulated in HIV-1 infection. It plays a key role in the adaptive immune responses mounted against viruses, however, in addition to the protective effect, NF-κB may also contribute to viruses' replication, survival and spread [42,43,44]. The JAK-STAT pathway usually transmits information from chemical signals outside the cell and involved in regulation of the immune system. Here, it includes Interferon gamma, STAT3, beta-catenin to regulate the HIV replication in astrocytes [45]. This also explains that the key proteins in these pathways ranking higher in our result.

The rest proteins are associated with the immune response against HIV infection. Interleukin-2 (IL-2), a secreted cytokine, is observed increasing in early CD4+ T-cell response for HIV-1 infection to control viral replication [46], though this response will lose function with the disease processing [47]. Transcription factor AP-1 is recruited by HIV Nef protein to MHC-I cytoplasmic tail to disrupt the presentation of HIV-1 epitopes to anti-HIV cytotoxic T lymphocytes [48]. The aberrant changes in pp125FAK expression block the beta1 integrin-mediated protection effect for aberrant cell death in patients with AIDS [49]. Integrin alpha-4 expressed on NK cells is bound by HIV gp120 to suppressing NK cells [50]. JAK-STAT pathway is also responsible for the antiretroviral effect of IFN-gamma, and the Jak/STAT deficiency may contribute to the dysfunction of CD4 T cell responses to a cytokine like IL-2 by HIV [51,52].

Chemicals Related to 26 Core Human Proteins

Three approved HIV drugs targeting human proteins in Drugbank were collected (Table 2).

For each of the three HIV related drugs, its interactive chemicals in STITCH can be found. After collecting these chemicals and combining with the three drugs, we obtained 280 chemicals (Additional File S3). For each of 26 proteins, we can count the number of the interactive chemicals among these 280 chemicals. Figure 1 shows the number of interactive chemicals among 280 chemicals. From Fig. 1, we can see that the number of chemicals related to ENSP00000011653 (CD4) is the largest, followed by that of ENSP00000292303 (CCR5). The chemicals related to each of the 26 proteins were ranked according to their interaction score (see Additional File S4).

Chemicals Targeting Interaction between HIV and Human Co-receptors

CCR5, a membrane protein, is an important target of anti-HIV therapy as it is one of the major co-receptors for HIV-1infection. There are seven trans-membrane helix structures in CCR5, which formed a "pocket" structure (Fig. 2) [53]. In this "pocket", aromatic amino acid residuals, hydrophobic amino acid residuals, polar amino acid residuals and hydrophilicity amino acid residuals could bind with chemicals by π-π stacking interaction, hydrophobic interaction, hydrogen bonding interaction and salt-bridge interaction [54,55,56,57].

Table 2. Three HIV drugs targeting human proteins in DrugBank.

| Drugbank ID	Drug name	STITCH ID	Target gene
DB00900	Didanosine	CID100003043	PNP
DB00419	Miglustat	CID100051634	UGCG
DB04835	Maraviroc	CID100483407	CCR5

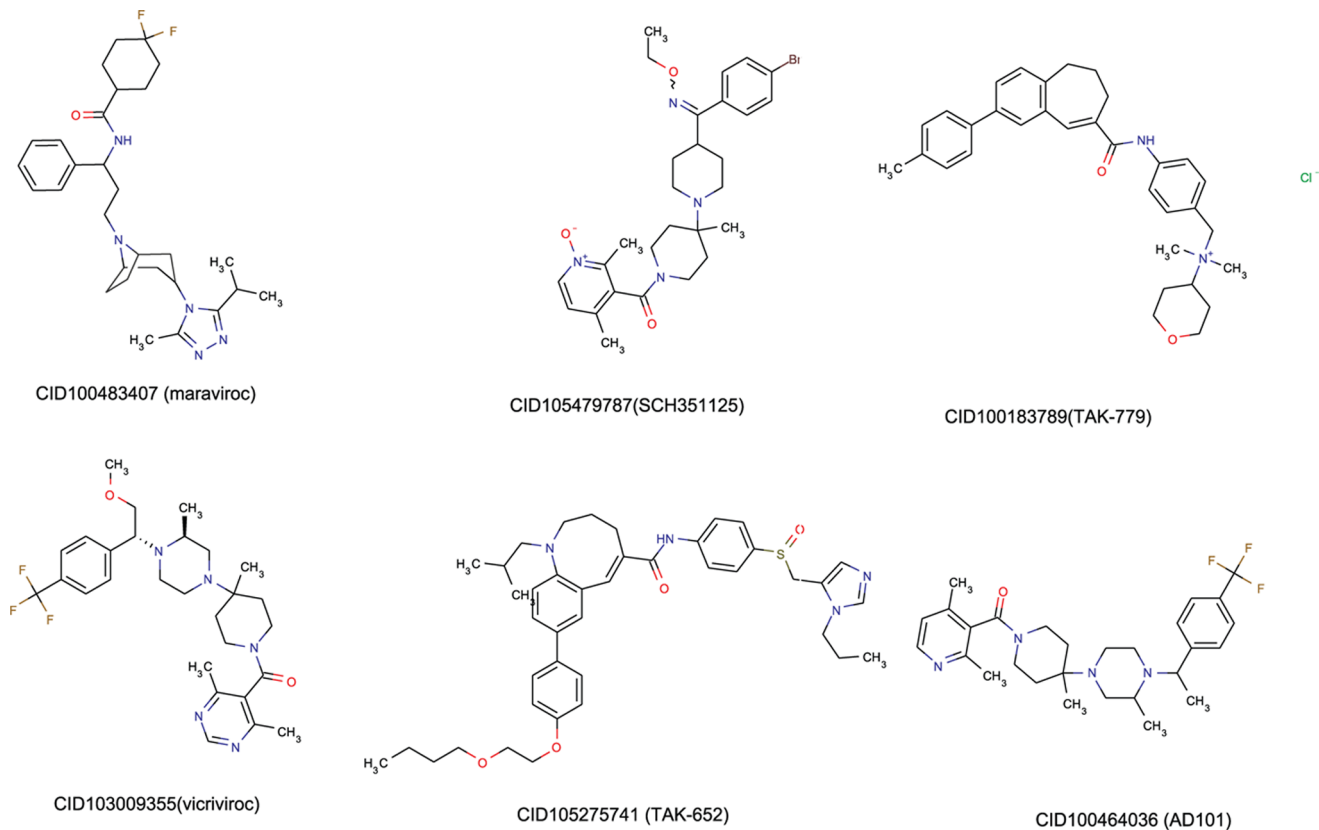doi:10.1371/journal.pone.0065207.t002
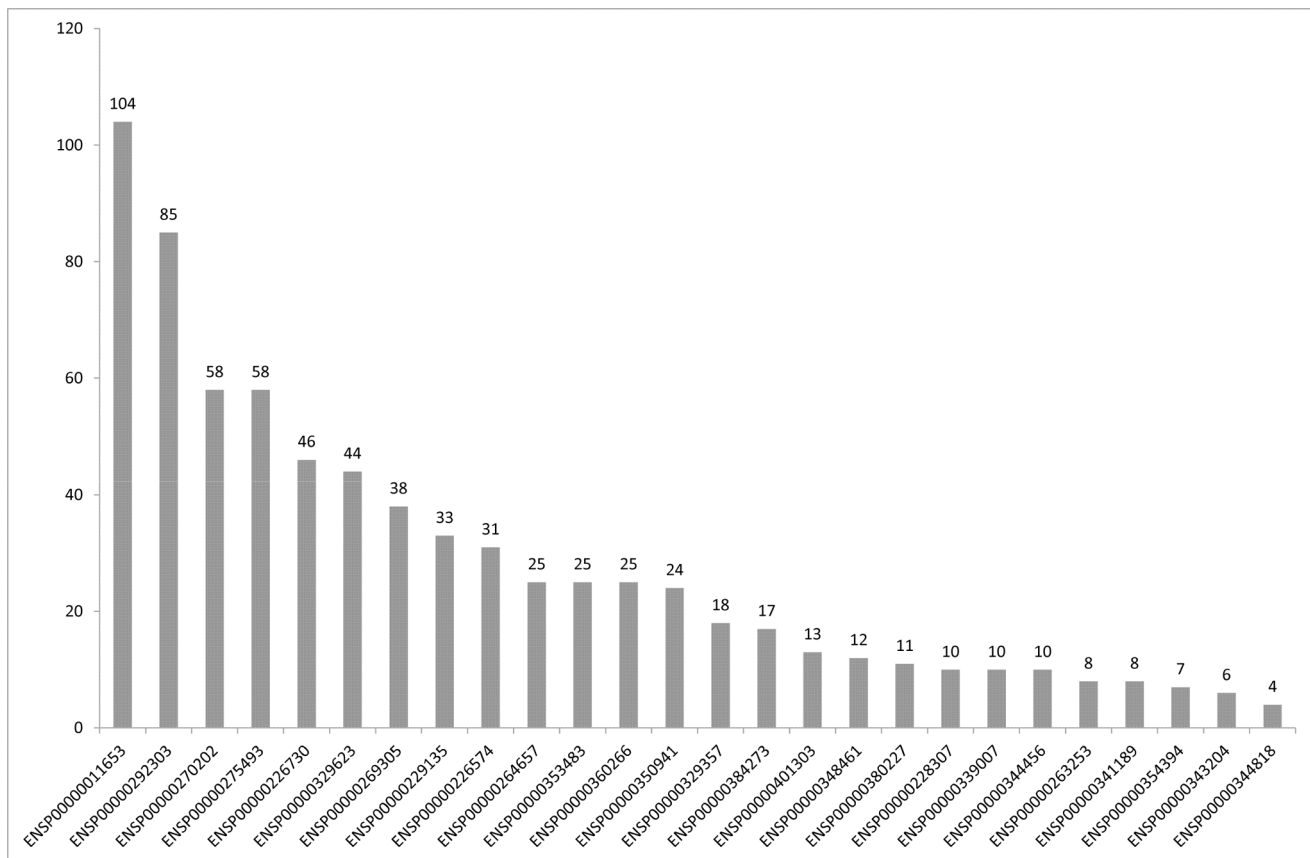
Figure 1. The number of interactive chemicals among 280 chemicals for each of the 26 core proteins.
doi:10.1371/journal.pone.0065207.g001

For ENSP00000292303 (CCR5), it can be found that some chemicals with interaction score higher than 900 are very similar in sub-structure (Fig. 3). Among these chemicals, CID100483407 (maraviroc) (Table 2) is a known anti-HIV drug which could bind with CCR5. As for chemicals CID105479787 (SCH 351125), CID100183789 (TAK-779), CID103009355 (vicriviroc), CID105275741 (TAK-652), CID100464036 (AD101), they are very similar to CID100483407 (maraviroc) in sub-structure. We speculated these chemicals may also have the same target CCR5.
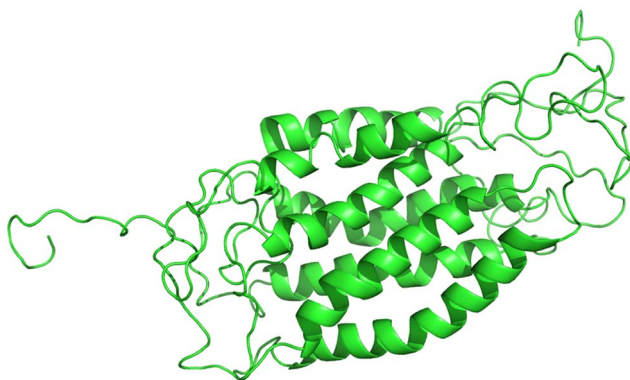
From Figure 3, it can be seen that all these chemicals have three hydrophobic structures containing basic nitrogen atoms (nitrogen atoms in bridge chain, piperidine ring, ammonium salt) and amide group. The three hydrophobic structures can interact with hydrophobic amino acid or hydrophobic structures of trans-membrane by hydrophobic interaction or π-π stacking interaction like CID100483407 (maraviroc). Benzene ring and Triazole ring of CID100483407 (maraviroc) can insert into hydrophobic pocket and form T shape π-π stacking by interacting with Tyrosine (Tyr108) and tryptophan (Trp86). In addition, cyclohexyl group of CID100483407 (maraviroc) can interact with isoleucine (Ile198) to form hydrophobic interaction. Furthermore, the interaction between basic nitrogen atoms in bridge chain of CID100483407 (maraviroc) and hydrophilicity amino acid residuals of CCR5 forms salt-bridge which is the major binding modes. Similar to CID100483407 (maraviroc), basic nitrogen atoms in bridge chain of CID105479787 (SCH 351125) can also form salt-bridge with glutamic acid (Glu283). Methylbenzene and formyl pyridine ring at nitrogen atom could form π-π stacking structure with tryptophan 86 and 248 (Trp86, Trp248), respectively. Meanwhile, acetyl group of piperidine ring and isoleucine (Ile198) could form hydrophobic structure. As for the other chemicals like CID100183789 (TAK-779), CID103009355 (vicriviroc), CID105275741 (TAK-652), CID100464036 (AD101), they are similar to CID100483407 (maraviroc). Therefore, they may be also considered as potential anti-HIV drugs targeting CCR5.
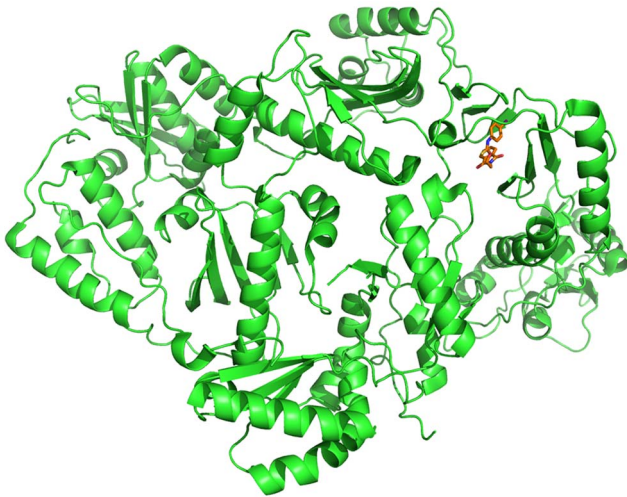

Figure 2. Three-dimensional structure of CCR5 based on PDB structure 1ND8 drawn with software Pymol. CCR5 is in green.
doi:10.1371/journal.pone.0065207.g002

Figure 3. Structures of 6 chemicals whose interaction scores with ENSP00000292303 (CCR5) are greater than 900. The figure was generated using ChemAxon. The 6 chemicals are CID100483407 (maraviroc), CID105479787 (SCH 351125), CID100183789 (TAK-779), CID103009355 (vicriviroc), CID105275741 (TAK-652) and CID100464036 (AD101), which can also be found in PubChem with the IDs 483407, 5479787, 183789, 3009355, 5275741, and 464036, respectively.
doi:10.1371/journal.pone.0065207.g003

Chemicals Targeting Interactions Involving HIV-1 Reverse Transcriptase

It can be found that some chemicals whose interaction score are higher than 740 for ENSP00000011653 (CD4) are also similar in sub-structure. Most of these chemicals are related to HIV-1 reverse transcriptase (HIV-1 RT) and HIV-1 Protease (HIV-1 PR). HIV-1 RT is a hetero-dimeric enzyme which is composed of two distinct subunits P66 and P51 [58,59]. The peptide sequence of P51 is identical to the first 440 amino acids of P66, and they form the two subunits of polymerases domain. The subunit looks like human's right hand which contains the finger, palm, thumb, and connection subdomains (see Fig. 4). The finger subdomain includes β-sheets and three α-helices, and the palm subdomain contains five α-helices. These α-helices and β-sheets of finger and palm subdomains could form hydrogen bonding structure with four β-sheets of thumb subdomain. The hand of the domain and the RNase H domain is connected by connection subdomain which is composed of a big β-sheet and two α-helices [58,59,60]. P66 also looks like a right hand, and it makes up a large template-primer binding cleft of polymerase. The 3′-OH terminus of the primer is positioned close to active site of polymerase (three catalytic amino acid residuals: Asp110, Asp185 and Asp186). P51 is processed by proteolytic cleavage of P66, which is different from P66 in structure although their amino acid sequences are similar [61]. The finger of P51 is close to the palm, and there is no template-primer binding cleft. As the active sites are buried, there is no catalytic activity for P51. Hence each P66/P51 dimer has only one active site which is located in P66. When HIV infects the host cell, HIV-1 RT creates single-stranded DNA from the RNA template. First, RT binds to RNA. Then the corresponding DNA nucleoside of host cell binds to phosphate group as substrate, and copy RNA nucleotide [61]. As RT is a essential enzyme during the replication of HIV-1, lack of HIV-1 RT could block the HIV-1 replication cycle, thus preventing HIV reproduction. Therefore, RT is regarded as an important anti-HIV target. At present, RT inhibitors could be classified to nucleoside analog reverse-transcriptase inhibitors (NARTIs) and non-nucleoside reverse-transcriptase inhibitors (NNRTIs).

Among the chemicals with score greater than 740, CID100003043 (didanosine) (Table 2) is a known anti-HIV drug targeting HIV-1 RT. Intriguingly, we found that CID100005726 (zidovudine) and CID100005155 (stavudine) may also bind to HIV-1 RT. These three chemicals are nucleoside analog which are very similar to RNA and DNA in structure (Fig. 5). Nucleoside analogs could be phosphorylated when they enter the cells. Then they compete with natural deoxynucleotides for binding with RT, thus inhibit the usage of nucleoside substrates by RT, arrest the growing of viral DNA and prevent viruses' reproduction [62,63,64]. In this study, CID100005726 (zidovudine), CID100005155 (stavudine), CID100003043 (didanosine) are phosphorylated to nucleoside 5′-monophosphate analog, nucleoside 5′-diphosphate analog, and nucleoside 5′-triphosphate analog, respectively, after the three chemicals enter the cells. Then the three analogs could bind with RT instead of natural nucleoside phosphate substrates (dTTP, dCTP, dATP, dGTP). As

Figure 4. Three-dimensional structure of HIV-1 RT based on PDB structure 2VG7 drawn with software Pymol. HIV-1 RT is in green. Ligand is in orange, red and blue.
doi:10.1371/journal.pone.0065207.g004

a result, the binding between natural nucleoside substrates and HIV-1 RT is blocked, and the HIV-1 RT is competitively inhibited. On the other hand, as there is no 3′-OH in these three chemicals, viral DNA could not grow after binding with the three chemicals. This could also prevent the HIV viruses' reproduction.

Other chemicals targeting HIV-1 RT including CID100060847 (BHAP), CID100004463 (nevirapine), CID105495818 (BMS-378806) could be classified to NNRTIs. Different to NARTIs, NNRTIs have two symmetrical aromatic rings, which show special butterfly-like shape (Fig. 6) [65,66]. Five β-sheets of P66 form a pocket which is composed of hydrophobic amino acids. Based on hydrophobic, hydrogen bonding and π-π stacking interaction derived from aromatic ring of aromatic amino acids, the aromatic rings from one side of these three chemicals can interact with aromatic amino acids including Tyr181, Tyr188, Phe227, Trp229 and those from another side can interact with hydrophobic amino acids including Val179, Val106 and Leul00. Therefore, a small hydrophobic pocket is formed by Tyr181, Tyr183 and Tyr188. As the three amino acids rotate outside, the entrance of the pocket will be exposed where HIV-1 RT can bind to these three chemicals. In this case, the relative locations of β4, β7 and β8 sheets will change. Complementary rearrangement of the confor-

mation of RT and CID100060847 (BHAP), CID100004463 (nevirapine), CID105495818 (BMS-378806) result in hydrophobic interactions [64,67,68]. As a result, the conformation of the newly located catalytic active site is similar to that of P51 [16]. Therefore, the new conformation is inactive. This is the reason why non-nucleoside analog has the ability to inhibit the RT by changing the conformation of catalytic site.

Chemicals Targeting Interactions Involving HIV-1 Protease

HIV-1 protease is a C2-symmetric homodimer including two monomers which have the identical polypeptide sequence with 99 residues (see Fig. 7) [69]. There is an active site (Asp-Thr-Gly) in the region between P25 and P27. The two subunits are connected by four β anti-parallel strands containing glycine, and each strand contains N-terminal domain and C-terminal domain. Both the monomers have a long "cavity" structure, on the bottom of which lie the catalytic aspartyl residues with planar configuration [69,70]. Due to the special structure of HIV-1 PR, the substrate peptide binds to the enzyme in an extended anti-parallel β sheet through the amino acid side chains from completely opposite directions [71]. It should be noted that the two subunits of enzyme are not completely identical, although they are symmetrical. Both the monomers have a "flap" structure which is made up of antiparallel β strands extending to subsite P1 and P2' [72]. Due to the different conformation of 'flap' of the two subunits, such a symmetrical conformation has the ability to recognize particular amino acid residues to control the substrates/inhibitors' access.

Both CID100003706 (indinavir) and CID100005076 (ritonavir) are peptidic chemicals which could compete with natural substrates as the substrates of HIV-1 PR. During the process of hydrolysis of peptide bonds of these two chemicals, one active water molecule is polarized by carboxyl group of Aspartyl residue. Nucleophilic O-atom of water attacks the carbonyl of the substrate's scissile bond to form a tetrahedral intermediate which could further become amino and carbonyl chemical. Then the hydroxy and carboxyl group form hydrogen interacts with Aspartate as hydrogen bond acceptor and donor, respectively. Meanwhile, aromatic rings of side chain of these two chemicals firmly bond to active region of protease through electrostatic and steric interactions. As a result, the conformation of "flap" changes and a tunnel structure which go through the dimmer obliquely forms in the active region. Then the symmetrical conformation of protease is broken and the flexible region closes. Eventually, HIV-1 PR is inhibited by damaging its activity.

CID100003043 (didanosine)   CID100005726 (zidovudine)   CID100005155 (stavudine)
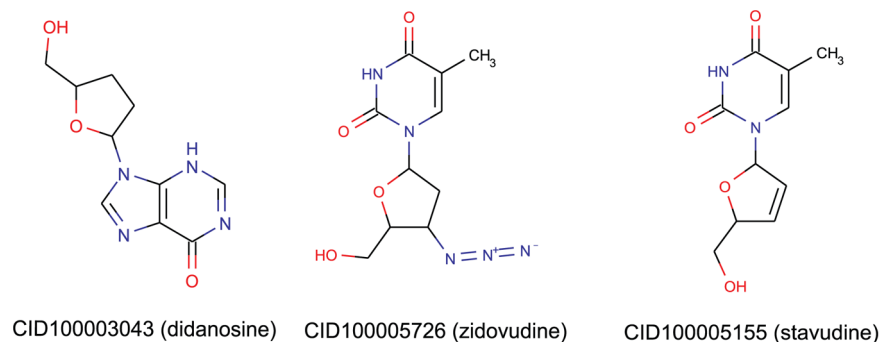
Figure 5. Structures for 3 chemicals whose interaction scores with ENSP00000011653 (CD4) are greater than 740. The figure was generated using ChemAxon. The 3 chemicals are CID100003043 (didanosine), CID100005726 (zidovudine) and CID100005155 (stavudine), which can also be found in PubChem with the IDs 3043, 5726 and 5155, respectively.
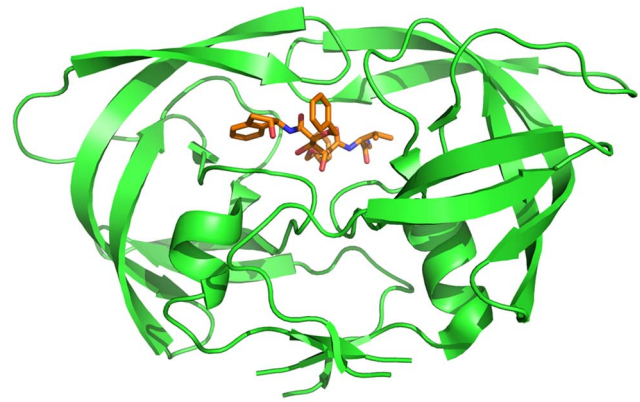doi:10.1371/journal.pone.0065207.g005
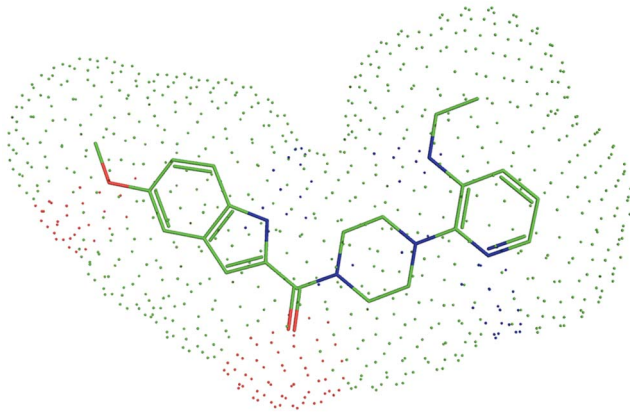
Figure 6. Non-nucleoside analog's butterfly-like shape drawn with software chemoffice and Pymol. The structure was retrieved from PubChem with the ID 60847. Dot surface is colored according to atom types.
doi:10.1371/journal.pone.0065207.g006

Miglustat may also Target HIV-1 Protease

Miglustat is another approved drug for HIV in Drugbank with the identity of CID100051634 (Table 2). It was revealed by our result that CID100051634 is also related to ENSP00000011653 (CD4), though interaction score of CID100051634 is only 359. However, it can be found that the structure of CID100051634 is still similar to the sub-structure of peptidic chemical. And the related studies show that CID100051634 also have anti-HIV activity in experiment. CID100051634 is an N-alkylated imino sugar. Clinical trials show that CID100051634 alter the glycosylation of envelope glycoproteins and decrease the infectivity in certain viral diseases such as HIV [73]. Carefully study CID100051634, we can found that CID100051634 has three hydroxyl groups which can interacts with Aspartate as hydrogen bond acceptor and donor, respectively. Therefore, we presume CID100051634 may also target HIV-1 PR.

New Combinations of HAART Proposed by Chemical-chemical Interaction

A combination of HAART generally includes two NARTIs and one drug in the following classes: NNRTI, protease inhibitor (PI), integrase strand transfer inhibitor (INSTI), or a CCR5 antagonist. According to the presumption that "two interactive chemicals are more likely to share similar biological functions [74,75,76]", we attempted to propose some new combinations of HAART through substituting the original components for their interactive chemicals. One way is replaced by inhibitors in the same class. NIH proposed some preferred regimens of HAART with optimal, durable efficacy, favorable tolerability and toxicity profile [77], such as atazanavir/ritonavir+tenofovir disoproxil fumarate/emtricitabine (ATV/r+TDF/FTC). TDF/FTC is often used as a backbone for boosted PI-based regimens in the initial treatment of HIV-1 infection [78]. Therefore, we attempted to substitute ATV for other PIs. We found 166 interactive chemicals of atazanavir from STITCH [22,79]. Except sulfate, eight interactive chemicals with the highest confidence score are PIs, including lopinavir (LPV), darunavir (DRV), ritonavir (RTV), saquinavir (SQV), fosamprenavir (FPV), nelfinavir (NFV), amprenavir (APV), and indinavir (IDV). All of these eight interactive chemicals related to ENSP00000292303 (CCR5), which is also shared by ATV. The combinations of one of the first five PIs with TDF/FTC are recognized regimens for anti-HIV-1 therapeutics by NIH [77].

Figure 7. Three-dimensional structure of HIV-1 PR based on PDB structure 1EBZ drawn with software Pymol. HIV-1 PR is in green. Ligand is in orange, red and blue.
doi:10.1371/journal.pone.0065207.g007

NFV+TDF/FTC is also used for clinical AIDS treatment, although this medication may cause life-threatening lactic acidosis [80]. APV combined with TDF/FTC have been observed to have additive synergistic effects for antiretroviral therapy [81]. Thus, it is reasonable to assume the validity of IDV+TDF/FTC for the treatment of AIDS, but it need the safety assessment.

Another way is substituted for inhibitor in different classes. For example, Trizivir (abacavir+lamivudine+zidovudine, ABC+3TC+AZT) is recommended as an initial antiretroviral therapy [77]. Here we took 3TC+AZT as a backbone, and substituted ABC for inhibitors in different classes. Three inhibitors with the highest interaction score, efavirenz (EFV), nevirapine (NVP) and delavirdine (DLV), are NNRTIs. ABC and its three interactive chemicals are associated with ENSP00000011653 (CD4) and ENSP00000292303 (CCR5). EFV or NVP with 3TC+AZT are recognized by NIH [77], so we thought it was feasible to use DLV+3TC+AZT for antiretroviral therapy.

Conclusion

At present, there is a great need for alternative way of inhibition for the design of anti-HIV therapeutics, because of the increased resistance of HIV to already approved drugs. Recently, inhibition of protein-protein interactions in the HIV life cycle is increasingly recognized as a valuable new avenue in drug design. In this work, we identified 26 core human proteins which play important roles in the HIV life cycles by interacting with HIV encoded proteins. In addition, 280 chemicals that interact with three HIV drugs targeting human proteins can also interact with these 26 core proteins. Therefore, the 280 chemicals may possess the potential for HIV therapy through intervention of PPI between 26 core human proteins and HIV encoded proteins. Our method may open a new way for HIV drug design or at least plays a complementary role to the existing method.

Supporting Information

Additional File S1 Protein-protein interaction information between HIV and human.
(XLS)

Additional File S2 All the betweennesses for the 1450 proteins involved in interaction with HIV encoded proteins.
(XLSX)

Additional File S3 The information of 280 chemicals.
(XLSX)

Additional File S4 The interacting chemicals for each of the 26 core proteins in 280 chemicals.
(XLSX)

Acknowledgments

Author Contributions

Conceived and designed the experiments: XYK YDC. Performed the experiments: BQL LC. Analyzed the data: BN LC MJ JL. Contributed reagents/materials/analysis tools: ZJW TH MYZ. Wrote the paper: BQL BN LC.

References

1. Weiss RA (1993) How does HIV cause AIDS? Science 260: 1273–1279.
2. Douek DC, Roederer M, Koup RA (2009) Emerging concepts in the immunopathogenesis of AIDS. Annu Rev Med 60: 471–484.
3. Dieffenbach CW, Fauci AS (2011) Thirty years of HIV and AIDS: future challenges and opportunities. Ann Intern Med 154: 766–771.
4. De Clercq E (2009) The history of antiretrovirals: key discoveries over the past 25 years. Rev Med Virol 19: 287–299.
5. De Clercq E (2009) Anti-HIV drugs: 25 compounds approved within 25 years after the discovery of HIV. Int J Antimicrob Agents 33: 307–320.
6. Christ F, Voet A, Marchand A, Nicolet S, Desimmie BA, et al. (2010) Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. Nat Chem Biol 6: 442–448.
7. Zhan P, Liu X, Li Z, Pannecouque C, De Clercq E (2009) Design strategies of novel NNRTIs to overcome drug resistance. Curr Med Chem 16: 3903–3917.
8. Greene WC, Debyser Z, Ikeda Y, Freed EO, Stephens E, et al. (2008) Novel targets for HIV therapy. Antiviral Res 80: 251–265.
9. Chen H, Liu X, Li Z, Zhan P, De Clercq E (2010) TSG101: a novel anti-HIV-1 drug target. Curr Med Chem 17: 750–758.
10. Mingyan Y, Xinyong L, De Clercq E (2009) NF-kappaB: the inducible factors of HIV-1 transcription and their inhibitors. Mini Rev Med Chem 9: 60–69.
11. Wang Y, Liu XY, De Clercq E (2009) Role of the HIV-1 positive elongation factor P-TEFb and inhibitors thereof. Mini Rev Med Chem 9: 379–385.
12. Zhan P, Liu X, De Clercq E (2010) Blocking nuclear import of pre-integration complex: an emerging anti-HIV-1 drug discovery paradigm. Curr Med Chem 17: 495–503.
13. Gerrard JA, Hutton CA, Perugini MA (2007) Inhibiting protein-protein interactions as an emerging paradigm for drug discovery. Mini Rev Med Chem 7: 151–157.
14. White AW, Westwell AD, Brahemi G (2008) Protein-protein interactions as targets for small-molecule therapeutics in cancer. Expert Rev Mol Med 10: e8.
15. Berg T (2008) Small-molecule inhibitors of protein-protein interactions. Curr Opin Drug Discov Devel 11: 666–674.
16. Busschots K, De Rijck J, Christ F, Debyser Z (2009) In search of small molecules blocking interactions between HIV proteins and intracellular cofactors. Mol Biosyst 5: 21–31.
17. Fu W, Sanders-Beer BE, Katz KS, Maglott DR, Pruitt KD, et al. (2009) Human immunodeficiency virus type 1, human protein interaction database at NCBI. Nucleic Acids Res 37: D417–422.
18. Ptak RG, Fu W, Sanders-Beer BE, Dickerson JE, Pinney JW, et al. (2008) Cataloguing the HIV type 1 human protein interaction network. AIDS Res Hum Retroviruses 24: 1497–1502.
19. Pinney JW, Dickerson JE, Fu W, Sanders-Beer BE, Ptak RG, et al. (2009) HIV-host interactions: a map of viral perturbation of the host system. AIDS 23: 549–554.
20. Rice AP, Sutton RE (2007) Targeting protein–protein interactions for HIV therapeutics. Future HIV Therapy 1: 369–385.
21. Jensen LJ, Kuhn M, Stark M, Chaffron S, Creevey C, et al. (2009) STRING 8-a global view on proteins and their functional interactions in 630 organisms. Nucleic acids research 37: D412–416.
22. Kuhn M, von Mering C, Campillos M, Jensen LJ, Bork P (2008) STITCH: interaction networks of chemicals and proteins. Nucleic Acids Res 36: D684–688.
23. Keshava Prasad TS, Goel R, Kandasamy K, Keerthikumar S, Kumar S, et al. (2009) Human Protein Reference Database–2009 update. Nucleic Acids Res 37: D767–772.
24. Joy MP, Brock A, Ingber DE, Huang S (2005) High-betweenness proteins in the yeast protein interaction network. Journal of Biomedicine and Biotechnology 2005: 96–103.
25. Freeman LC (1977) A set of measures of centrality based on betweenness. Sociometry 40: 35–41.
26. Li B-Q, Huang T, Liu L, Cai Y-D, Chou K-C (2012) Identification of Colorectal Cancer Related Genes with mRMR and Shortest Path in Protein-Protein Interaction Network. PloS one 7: e33393.
27. Girvan M, Newman MEJ (2002) Community structure in social and biological networks. Proceedings of the National Academy of Sciences 99: 7821–7826.
28. Li B-Q, Zhang J, Huang T, Zhang L, Cai Y-D (2012) Identification of retinoblastoma related genes with shortest path in a protein–protein interaction network. Biochimie 94: 1910–1917.
29. Berger EA, Murphy PM, Farber JM (1999) Chemokine receptors as HIV-1 coreceptors: Roles in viral entry, tropism, and disease. Annual Review of Immunology 17: 657–700.
30. Ryu SE, Truneh A, Sweet RW, Hendrickson WA (1994) Structures of an HIV and MHC binding fragment from human CD4 as refined in two crystal lattices. Structure (London, England : 1993) 2: 59–74.
31. Allen SJ, Crown SE, Handel TM (2007) Chemokine: Receptor structure, interactions, and antagonism. Annual Review of Immunology. 787–820.
32. Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, et al. (1996) Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. Cell 86: 367–377.
33. Dean M, Carrington M, Winkler C, Huttley GA, Smith MW, et al. (1996) Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Science 273: 1856–1862.
34. Zoughlami Y, Voermans C, Brussen K, van Dort KA, Kootstra NA, et al. (2012) Regulation of CXCR4 conformation by the small GTPase Rac1: implications for HIV infection. Blood 119: 2024–2032.
35. Izumi T, Io K, Matsui M, Shirakawa K, Shinohara M, et al. (2010) HIV-1 viral infectivity factor interacts with TP53 to induce G2 cell cycle arrest and positively regulate viral replication. Proceedings of the National Academy of Sciences of the United States of America 107: 20798–20803.
36. Kino T, Pavlakis GN (2004) Partner molecules of accessory protein Vpr of the human immunodeficiency virus type 1. DNA and Cell Biology 23: 193–205.
37. Kiernan RE, Vanhulle C, Schiltz L, Adam E, Xiao H, et al. (1999) HIV-1 Tat transcriptional activity is regulated by acetylation. Embo Journal 18: 6106–6118.
38. Rom S, Pacifici M, Passiatore G, Aprea S, Waligorska A, et al. (2011) HIV-1 Tat binds to SH3 domains: Cellular and viral outcome of Tat/Grb2 interaction. Biochimica Et Biophysica Acta-Molecular Cell Research 1813: 1836–1844.
39. Bres V, Kiernan RE, Linares LK, Chable-Bessia C, Plechakova O, et al. (2003) A non-proteolytic role for ubiquitin in Tat-mediated transactivation of the HIV-1 promoter. Nature Cell Biology 5: 754–761.
40. Deregibus MC, Cantaluppi V, Doublier S, Brizzi MF, Deambrosis I, et al. (2002) HIV-1-Tat protein activates phosphatidylinositol 3-kinase/AKT-dependent survival pathways in Kaposi's sarcoma cells. Journal of Biological Chemistry 277: 25195–25202.
41. Ju SM, Goh AR, Kwon D-J, Youn GS, Kwon H-J, et al. (2012) Extracellular HIV-1 Tat induces human beta-defensin-2 production via NF-kappaB/AP-1 dependent pathways in human B cells. Molecules and Cells 33: 335–341.
42. Chan JKL, Greene WC (2011) NF-kappa B/Rel: agonist and antagonist roles in HIV-1 latency. Current Opinion in Hiv and Aids 6: 12–18.
43. Fiume G, Vecchio E, De Laurentiis A, Trimboli F, Palmieri C, et al. (2012) Human immunodeficiency virus-1 Tat activates NF-kappa B via physical interaction with I kappa B-alpha and p65. Nucleic Acids Research 40: 3548–3562.
44. Chan JK, Greene WC (2012) Dynamic roles for NF-kappa B in HTLV-I and HIV-1 retroviral pathogenesis. Immunological Reviews 246: 286–310.
45. Li W, Henderson LJ, Major EO, Al-Harthi L (2011) IFN-γ Mediates Enhancement of HIV Replication in Astrocytes by Inducing an Antagonist of the β-Catenin Pathway (DKK1) in a STAT 3-Dependent Manner. The Journal of Immunology: 6771–6778.
46. Boaz MJ, Waters A, Murad S, Easterbrook PJ, Vyakarnam A (2002) Presence of HIV-1 gag-specific IFN-gamma+IL-2(+) and CD28(+)IL-2(+) CD4 T cell responses is associated with nonprogression in HIV-1 infection. Journal of Immunology 169: 6376–6385.
47. Jones RB, Yue F-Y, Gu XXJ, Hunter DV, Mujib S, et al. (2009) Human Immunodeficiency Virus Type 1 Escapes from Interleukin-2-Producing CD4(+) T-Cell Responses without High-Frequency Fixation of Mutations. Journal of Virology 83: 8722–8732.
48. Wonderlich ER, Leonard JA, Kulpa DA, Leopold KE, Norman JM, et al. (2011) ADP Ribosylation Factor 1 Activity Is Required To Recruit AP-1 to the Major Histocompatibility Complex Class I (MHC-I) Cytoplasmic Tail and Disrupt MHC-I Trafficking in HIV-1-Infected Primary T Cells. Journal of Virology 85: 12216–12226.
49. Ng TTC, Kanner SB, Humphries MJ, Wickremasinghe RG, Nye KE, et al. (1997) The integrin-triggered rescue of T lymphocyte apoptosis is blocked in HIV-1-infected individuals. Journal of Immunology 158: 2984–2999.

50. Reitano KN, Kottilil S, Gille CM, Zhang X, Yan M, et al. (2009) Defective Plasmacytoid Dendritic Cell-NK Cell Cross-Talk in HIV Infection. Aids Research and Human Retroviruses 25: 1029–1037.

51. Bovolenta C, Lorini AL, Mantelli B, Camorali L, Novelli F, et al. (1999) A selective defect of IFN-gamma- but not of IFN-alpha-induced JAK/STAT pathway in a subset of U937 clones prevents the antiretroviral effect of IFN-gamma against HIV-1. Journal of Immunology 162: 323–330.

52. Kryworuchko M, Pasquier V, Theze J (2003) Human immunodeficiency virus-1 envelope glycoproteins and anti-CD4 antibodies inhibit interleukin-2-induced Jak/STAT signalling in human CD4 T lymphocytes. Clinical and Experimental Immunology 131: 422–427.

53. Luster AD (1998) Chemokines - Chemotactic cytokines that mediate inflammation. New England Journal of Medicine 338: 436–445.

54. Kondru R, Zhang J, Ji C, Mirzadegan T, Rotstein D, et al. (2008) Molecular interactions of CCR5 with major classes of small-molecule anti-HIV CCR5 antagonists. Molecular Pharmacology 73: 789–800.

55. Tsamis F, Gavrilov S, Kajumo F, Seibert C, Kuhmann S, et al. (2003) Analysis of the mechanism by which the small-molecule CCR5 antagonists SCH-351125 and SCH-350581 inhibit human immunodeficiency virus type 1 entry. Journal of Virology 77: 5201–5208.

56. Nishikawa M, Takashima K, Nishi T, Furuta RA, Kanzaki N, et al. (2005) Analysis of binding sites for the new small-molecule CCR5 antagonist TAK-220 on human CCR5. Antimicrobial Agents and Chemotherapy 49: 4708–4715.

57. Maeda K, Das D, Ogata-Aoki H, Nakata H, Miyakawa T, et al. (2006) Structural and molecular interactions of CCR5 inhibitors with CCR5. Journal of Biological Chemistry 281: 12688–12698.

58. Spallarossa A, Cesarini S, Ranise A, Ponassi M, Unge T, et al. (2008) Crystal structures of HIV-1 reverse transcriptase complexes with thiocarbamate non-nucleoside inhibitors. Biochemical and Biophysical Research Communications 365: 764–770.

59. Spallarossa A, Cesarini S, Ranise A, Bruno O, Schenone S, et al. (2009) Novel modifications in the series of O-(2-phthalimidoethyl)-N-substituted thiocarbamates and their ring-opened congeners as non-nucleoside HIV-1 reverse transcriptase inhibitors. European Journal of Medicinal Chemistry 44: 1650–1663.

60. Cesarini S, Spallarossa A, Ranise A, Fossa P, La Colla P, et al. (2008) Thiocarbamates as non-nucleoside HIV-1 reverse transcriptase inhibitors. Part 1: Parallel synthesis, molecular modelling and structure-activity relationship studies on O-[2-(hetero)arylethyl]-N-phenylthiocarbamates. Bioorganic & Medicinal Chemistry 16: 4160–4172.

61. Jacobo-Molina A, Ding J, Nanni RG, Clark AD, Jr., Lu X, et al. (1993) Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 A resolution shows bent DNA. Proceedings of the National Academy of Sciences of the United States of America 90: 6320–6324.

62. Mitsuya H, Weinhold KJ, Furman PA, St Clair MH, Lehrman SN, et al. (1985) 3′-Azido-3′-deoxythymidine (BW A509U): an antiviral agent that inhibits the infectivity and cytopathic effect of human T-lymphotropic virus type III/lymphadenopathy-associated virus in vitro. Proceedings of the National Academy of Sciences of the United States of America 82: 7096–7100.

63. Ren J, Esnouf R, Garman E, Somers D, Ross C, et al. (1995) High resolution structures of HIV-1 RT from four RT-inhibitor complexes. Nature structural biology 2: 293–302.

64. Ren J, Esnouf R, Hopkins A, Ross C, Jones Y, et al. (1995) The structure of HIV-1 reverse transcriptase complexed with 9-chloro-TIBO: lessons for inhibitor design. Structure (London, England : 1993) 3: 915–926.

65. Ding J, Das K, Tantillo C, Zhang W, Clark AD, Jr., et al. (1995) Structure of HIV-1 reverse transcriptase in a complex with the non-nucleoside inhibitor alpha-APA R 95845 at 2.8 A resolution. Structure (London, England : 1993) 3: 365–379.

66. Ding J, Das K, Moereels H, Koymans L, Andries K, et al. (1995) Structure of HIV-1 RT/TIBO R 86183 complex reveals similarity in the binding of diverse nonnucleoside inhibitors. Nature structural biology 2: 407–415.

67. Esnouf R, Ren J, Ross C, Jones Y, Stammers D, et al. (1995) Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors. Nature structural biology 2: 303–308.

68. DeClercq E (1996) Non-nucleoside reverse transcriptase inhibitors (NNRTIs) for the treatment of human immunodeficiency virus type 1 (HIV-1) infections: Strategies to overcome drug resistance development. Medicinal Research Reviews 16: 125–157.

69. Ho DD, Toyoshima T, Mo H, Kempf DJ, Norbeck D, et al. (1994) Characterization of human immunodeficiency virus type 1 variants with increased resistance to a C2-symmetric protease inhibitor. Journal of Virology 68: 2016–2020.

70. Danley DE, Geoghegan KF, Scheld KG, Lee SE, Merson JR, et al. (1989) Crystallizable HIV-1 protease derived from expression of the viral pol gene in Escherichia coli. Biochemical and Biophysical Research Communications 165: 1043–1050.

71. Krzysztof A (1993) Crystal structures of HIV-1 protease-inhibitor complexes. Perspectives in Drug Discovery and Design 1: 23–48.

72. Shao W, Everitt L, Manchester M, Loeb DD, Hutchison CA, et al. (1997) Sequence requirements of the HIV-1 protease flap region determined by saturation mutagenesis and kinetic analysis of flap mutants. Proceedings of the National Academy of Sciences of the United States of America 94: 2243–2248.

73. Tierney M, Pottage J, Kessler H, Fischl M, Richman D, et al. (1995) The tolerability and pharmacokinetics of N-butyl-deoxynojirimycin in patients with advanced HIV disease (ACTG 100). The AIDS Clinical Trials Group (ACTG) of the National Institute of Allergy and Infectious Diseases. Journal of acquired immune deficiency syndromes and human retrovirology : official publication of the International Retrovirology Association 10: 549–553.

74. Hu LL, Chen C, Huang T, Cai YD, Chou KC (2011) Predicting biological functions of compounds based on chemical-chemical interactions. PLoS One 6: e29491.

75. Chen L, Zeng WM, Cai YD, Feng KY, Chou KC (2012) Predicting Anatomical Therapeutic Chemical (ATC) classification of drugs by integrating chemical-chemical interactions and similarities. PLoS One 7: e35254.

76. Chen L, Lu J, Zhang J, Feng KR, Zheng MY, et al. (2013) Predicting chemical toxicity effects based on chemical-chemical interactions. PLoS One 8: e56517.

77. Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. Department of Health and Human Services. Available: http://aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL.pdf. Accessed April 5, 2013.

78. Perry CM (2009) Emtricitabine/tenofovir disoproxil fumarate: in combination with a protease inhibitor in HIV-1 infection. Drugs 69: 843–857.

79. Kuhn M, Szklarczyk D, Franceschini A, Campillos M, von Mering C, et al. (2010) STITCH 2: an interaction network database for small molecules and proteins. Nucleic Acids Res 38: D552–556.

80. Emtricitabine/nelfinavir/tenofovir Side Effects. Available: http://www.drugs.com/sfx/emtricitabine-nelfinavir-tenofovir-side-effects.html. Accessed April 5, 2013.

81. TRUVADA. Available: http://dailymed.nlm.nih.gov/dailymed/drugInfo.cfm?id=17398. Accessed April 6, 2013.