

Tackling the problem of HIV drug resistance

ABSTRACT

The virally-encoded HIV-1 protease is an effective target for antiviral drugs, however, treatment for HIV infections is limited by the prevalence of drug resistant viral mutants. In this review, we describe our three-pronged approach to analyze and combat drug resistance. Understanding the molecular basis for resistance due to protease inhibitors is a key initial step in this approach. This knowledge is being employed for the design of new, improved inhibitors with high affinity for resistant mutants as well as wild type enzyme. In parallel with experimental studies of diverse mutants and inhibitory compounds, we are developing efficient algorithms to predict drug resistance phenotype from genotype data. This approach has important practical applications in the clinic where genotyping is recommended for individuals with new infections.

INTRODUCTION

Human immunodeficiency virus (HIV) is the infectious agent for the HIV/AIDS pandemic with over 35 million people infected worldwide and an estimated 34 million deaths since 1981 [1]. Antiretroviral drugs remain the primary treatment and preventative options in the absence of an effective vaccine [2]. The standard therapeutic intervention uses combinations of drugs that inhibit different stages in the viral lifecycle: cell entry and fusion, reverse transcription, integration, and maturation [3]. Individuals with high risk of infection can benefit from pre-exposure prophylaxis or PrEP [4]. The genetic diversity of the virus, which arises from the high error rate of the reverse transcriptase and the rapid replication of the virus, is the major challenge for both therapy and PrEP [5]. The viral genomes can be classified into the more common HIV type 1 and the rarer type 2, while HIV-1 comprises four groups: M (subdivided into 9 subtypes A-J), O, N, and P with distinct geographical distributions. An infected individual may contain a mixture of evolving viral strains with a total estimated population of about 10^{10} HIV virions. This genetic diversity and rapid selection of variants enable the virus to escape the immune system and the existence of resistant mutants limits the effectiveness of drugs [6]. Due to this serious problem, resistance testing is recommended for newly diagnosed individuals, or those failing therapy [7]. A list of current mutations associated with resistance to the antiviral drugs is available and updated regularly [8]. In addition to the problem of drug resistance, treatment of HIV infections is hampered by the inaccessible reservoirs of latent virus due to its integration into the host cell genome.

The virally-encoded protease has proved a valuable target for antiviral drugs and a model for structure-guided design of inhibitors. The mature HIV protease is an aspartic protease comprised of two 99-amino acid subunits [3]. The protease acts during viral maturation to cleave the viral Gag and Gag-Pol polyprotein precursors into the separate enzymes and structural proteins [9,10]. This process must occur in an ordered fashion and the initial step is release of the mature protease by autocatalytic proteolysis from the Gag-Pol precursor. Hence, the protease is essential for viral replication and an excellent target for antiviral inhibitors. Inhibitors of the protease will act to decrease production of the other viral enzymes, reverse transcriptase and integrase, as well as the matrix, capsid and nucleocapsid structural proteins. Most protease inhibitors are based on transition state analogs of the peptide substrates. Historically, Alexander Wlodawer's group was a pioneer in crystallographic analysis of HIV protease with inhibitors and reported the first crystal structure of HIV protease in complex with a peptidic inhibitor [11]. Over the next years, a number of other co-crystal structures were described during the search for antiviral inhibitors [12]. Drugs that inhibit the HIV protease were first introduced in the clinic in 1995, and improved the survival of HIV-infected individuals over therapy with only reverse transcriptase inhib-

Irene T. Weber^{1,✉}

Robert W. Harrison²

¹Department of Biology, Georgia State University, Atlanta, GA, USA

²Department of Computer Science, Georgia State University, Atlanta, GA, USA

✉Department of Biology, Georgia State University, PO Box 4010, Atlanta, GA 30302-4010, USA; e-mail: iweber@gsu.edu

Received: May 26, 2016

Accepted: May 27, 2016

Key words: HIV/AIDS, X-ray crystallography, enzymes, bioinformatics machine learning

Abbreviations: HIV - human immunodeficiency virus; PrEP - pre-exposure prophylaxis; THF - tetrahydrofuran; PR20 - HIV protease mutant with 20 substitutions relative to standard wild type protease

Acknowledgements: We are indebted to Alexander Wlodawer for introducing us to the important and interesting problem of HIV protease. We thank Yuan-Fang Wang for help in preparing the figures. I.T.W. and R.W.H. are grateful for financial support for this research from the National Institutes of Health Award GM062920.

itors. Protease inhibitors are potent antiretroviral agents with effects on multiple stages of the lifecycle, including viral entry as well as maturation [13].

Due to the small size of the enzyme and relative ease of structural analysis, the majority of the clinical protease inhibitors have been designed with the aid of X-ray crystal structures of the protease-inhibitor complexes, as reviewed in [14]. Currently, more than a thousand crystal structures have been reported for HIV protease or its mutants in the apo form and in the presence of inhibitors. This plethora of structural information has revealed how the protease binds substrates or inhibitors and guided the design of antiviral inhibitors. In addition, the structural data have improved our understanding of the reaction mechanism of the enzyme. Reaction intermediates have been trapped in the crystals for several steps in the reaction pathway, including the quasi-stable tetrahedral intermediate [15]. More recently, our neutron crystal structures of per-deuterated HIV protease with clinical inhibitors have given important insights into the detailed geometry of protons, hydrogen bonds and water molecules in the active site [16,17].

To date, nine protease inhibitors have been approved for clinical use. The inhibitors and their clinical effects and resistance mutations are reviewed in [18]. The majority of these drugs were designed to inhibit the wild type enzyme, however, the newest drugs, tipranavir and darunavir, were designed to target drug resistant protease variants. As measured by isothermal titration calorimetry, darunavir and tipranavir show high binding affinity of 5–10 and 20 picomolar, respectively, for wild type protease, while the earlier drugs, saquinavir, indinavir, nelfinavir and amprenavir exhibit poorer binding affinities of 0.2–0.4 nM [19,20]. The most recently approved drug, darunavir, has a broader clinical application due to fewer adverse side effects compared to treatment with tipranavir [18]. Moreover, several studies have demonstrated that physiological concentrations of darunavir and saquinavir also inhibit the autocatalytic processing of the protease from its precursor [21–23]. Although alternate binding sites in the protease have been proposed as drug targets, most design efforts have focused on competitive inhibitors that bind in the active site cavity of the protease. Current strategies tackle the challenges of drug resistance and poorly accessible reservoirs of virus.

MOLECULAR BASIS FOR RESISTANCE TO PROTEASE INHIBITORS

HIV resistance to drugs poses a critical challenge for effective therapy. The virus evolves resistant variants to protease inhibitors by balancing the need for the protease to recognize diverse cleavage sites while rejecting the inhibitors. Since the mutations are independent results of a stochastic process, coordinated mutation of both the cleavage sites and protease can only arise slowly. Analysis of the structures and activities of protease mutants has proved valuable for understanding the molecular basis for drug resistance. This knowledge can also be applied to optimize the design strategy for new inhibitors targeting resistant variants. Clinical resistance to protease inhibitors is associated with mutations in the protease gene and different drugs elicit distinct mutations [8]. The structural location of major non-polymorphic mutations associated with drug resistance is illustrated in figure 1A. The scale of the sequence variation is demonstrated by a recent comprehensive analysis of >100,000 HIV-infected individuals in 143 countries [24]. In the protease sequences, 47% of positions had one or more types of amino acid residues occurring with at least 1% prevalence. Clusters of variation were observed between residues in the surface loops of the protease and in the short helix (Fig. 1A). These variations include polymorphisms existing in different subtypes of HIV as well as non-polymorphic mutations selected by drug treatment. This report identified a total of 111 non-polymorphic mutations associated with drug resistance. It is worth noting that the majority of drug resistance mutations involve conservative substitutions of hydrophobic amino acids, such as valine to isoleucine or leucine to methionine [8]. Substitutions of four polar residues, D30N, Q58E, N83D and N88S/D, occur as major resistance mutations associated with nelfinavir or tipranavir treatment, and can alter the charge of the protein.

Many of the initial “major” mutations alter the protease binding site for inhibitors and substrates, and often produce defects in protease activity and viral replication. Consequently, the virus evolves additional “minor” or “accessory” mutations that may act to increase the protease activity and viral fitness [25]. These mutations occur in regions distal to the inhibitor binding site. Short insertions of amino acids have been seen and may act to increase viral fitness [26]. As the virus evolves towards higher resistance, multiple mutations accumulate in the protease and also in its

Table 1. Highly resistant protease variants and their binding affinity for darunavir (DRV).

Protease	K_d DRV (nM)	Relative K_d	Amino Acid Substitutions
Wild Type	0.005	1.0	
^a PR20	41.000	8200	Q7K, L10F, I13V, I15V, D30N , V32I , L33F , E35D, M36I, S37N, I47V , I54L , Q58E, I62V, L63P, A71V, I84V , N88D , L89T, L90M
^a P51*	37.000	7400	L10I, I15V, K20R, L24I, V32I , I33F , M36I, M46L , I54M , I63P, K70Q, V82I, I84V , L89M
^b PRdrv4	35.000	7000	L10F, I13V, K14R, V32I , L33F , K45T, M46I , I47V , I54L , I62V, L63P, A71T, I72T, G73T, V77I, P79S, I84V , L90M
^c PR _{S17}	50.000	10000	L10I, K20R, E35D, M36I, S37D, M46L , G48V , I54V , D60E, I62V, L63P, A71V, I72V, V77I, V82S , L90M , I93L

Underlined bold indicates major resistance mutations from <http://hivdb.stanford.edu/DR/PIResiNote.html>. *Mutant selected in laboratory [33]; other mutants were identified in clinical isolates. Binding affinity measurement from ^a[34], ^b[35], ^c[36].

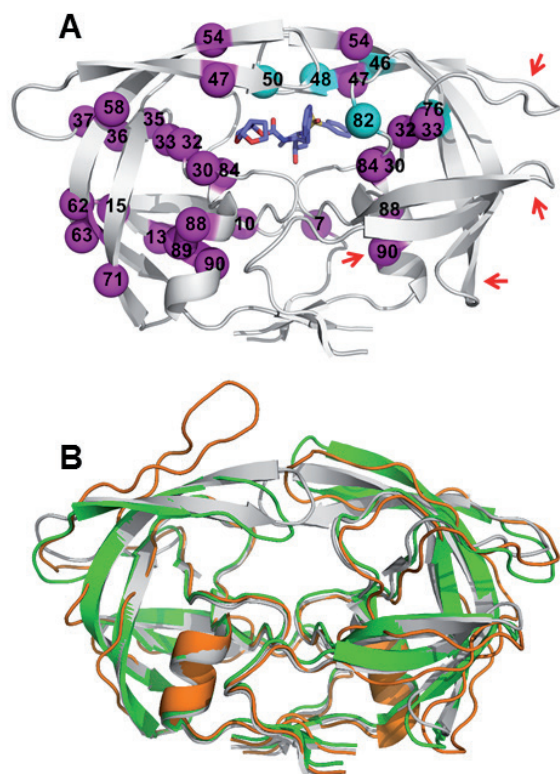


Figure 1. Dimer structure of HIV-1 protease. A) Protease dimer in complex with darunavir. The protease is shown in a gray ribbon representation with darunavir in blue sticks. Mutations in PR20 are shown as magenta spheres labeled with the residue number in the left subunit. Major mutations associated with resistance are shown in the right subunit in magenta spheres when the same site is mutated in PR20 and in cyan for mutations not present in PR20. The red arrows indicate the three surface loops and helix with high sequence variability. B) PR20 dimer with flap conformations seen in different crystal structures: the closed form in the presence of bound inhibitor (gray ribbons), widely separated open conformation flaps in the apo form (orange ribbons), and an asymmetric dimer with one flap wide open and the other flap tucked into the active site (green ribbons).

cleavage sites in the precursor polyproteins [25]. The most highly resistant proteases show changes in 20 or more amino acids in different regions of the structure, as reviewed in [27]. A variety of different sets of mutations occur in protease variants from drug resistant clinical isolates. Additional mutations in HIV Gag and Pol proteins co-evolve to compensate for loss of fitness due to protease mutations [28-31]. The massive number of potential combinations of mutations poses a severe problem for predictions of resistance from sequence, as discussed in a later section.

Structural and biochemical analyses have demonstrated the changes due to individual mutations. The structural effects of the major mutations in the inhibitor binding site have been described, however, the roles of many of the accessory mutations are poorly understood. Three general effects have been observed for protease bearing a single mutation, as reviewed in [32]. 1) Mutation of residues in the binding site (D30, V32, I47, I50, V82, I84) can directly alter the protease interactions with inhibitors. 2) Mutation of residues at the dimer interface (L24I, I50V, F53L) or in the hydrophobic core (L76V) can decrease the protease stability. 3) Mutation of residues in the flexible flaps can influence the dynamics of flap opening and closing. Recent experimental studies from several groups have examined the structural and en-

zymatic properties of a handful of highly resistant protease variants with up to 22 mutations, as reviewed in [27]. These proteases demonstrate several orders of magnitude lower affinity for the drugs. Selected examples of protease variants with poor affinity for the clinical inhibitor, darunavir, are listed in table 1. These highly resistant proteases from clinical isolates or a laboratory selected strain (P51) [33] exhibit 35–50 nM binding affinity for darunavir or ~10,000-fold worse than the value of 5 pM for wild type enzyme [34-36]. The protease sequences contain 14 to 20 mutations, including 5 to 8 major mutations. Only mutations of L10I/F and I54V/L/M occur in all four examples, suggesting multiple evolutionary pathways lead to high level resistance.

Our studies have focused on the extremely resistant variant PR20 with 20 substitutions relative to the wild type sequence. In contrast to wild type enzyme, autoprocessing of the precursor bearing the PR20 mutations is not significantly inhibited by darunavir and saquinavir [34]. Analysis of the crystal structures of PR20 and other multiply mutated variants showed two general changes compared to the wild type enzyme. 1) The highly resistant mutants often lose interactions with inhibitors due to direct and indirect changes in the binding site [35,37]. 2) The extreme mutants exhibit highly variable flap conformations in the absence of inhibitor [37,38]. In the absence of inhibitor, dimers of PR20 have been observed in diverse symmetric and asymmetric conformations as illustrated in figure 1B [37,39]. The mutations alter the dynamics of the transition between open and closed flap conformations, shifting toward the open conformation in the absence of inhibitor [39-42]. These structural and dynamic changes can be targeted in the designs of improved inhibitors.

STRATEGY FOR DEVELOPMENT OF NEW INHIBITORS TARGETING RESISTANT MUTANTS

A comprehensive description of diverse compounds, including natural products, which inhibit HIV-1 protease, is presented in [43]. The general design strategy pursued by our colleague, Arun Ghosh, is to introduce new interactions of the inhibitors with conserved regions of the protease dimer, in particular the main chain [44]. The potent antiviral inhibitor, darunavir, was designed based on this strategy with a *bis*-tetrahydrofuran (*bis*-THF) group at P2, and confirmed to form additional hydrogen bond interactions with the protease main chain [45,46]. Following this success, evaluation of a variety of new chemical substituents in the darunavir scaffold resulted in a series of antiviral inhibitors, described in [44]. The most potent of these inhibitors have been assessed against protease variants bearing single and multiple substitutions, as summarized in [27]. Crystal structures of inhibitor-protease complexes were solved and analyzed to understand the enzyme inhibition data and improve the design strategy.

Knowledge of the structural changes in the highly resistant variants such as PR20 has led to new insights for inhibitor designs [37]. Compounds were designed with 1) large groups at P2 to better fit in the enlarged S2 subsite, and 2) substituents providing interactions with the flexible flaps. Examples of inhibitors showing improved

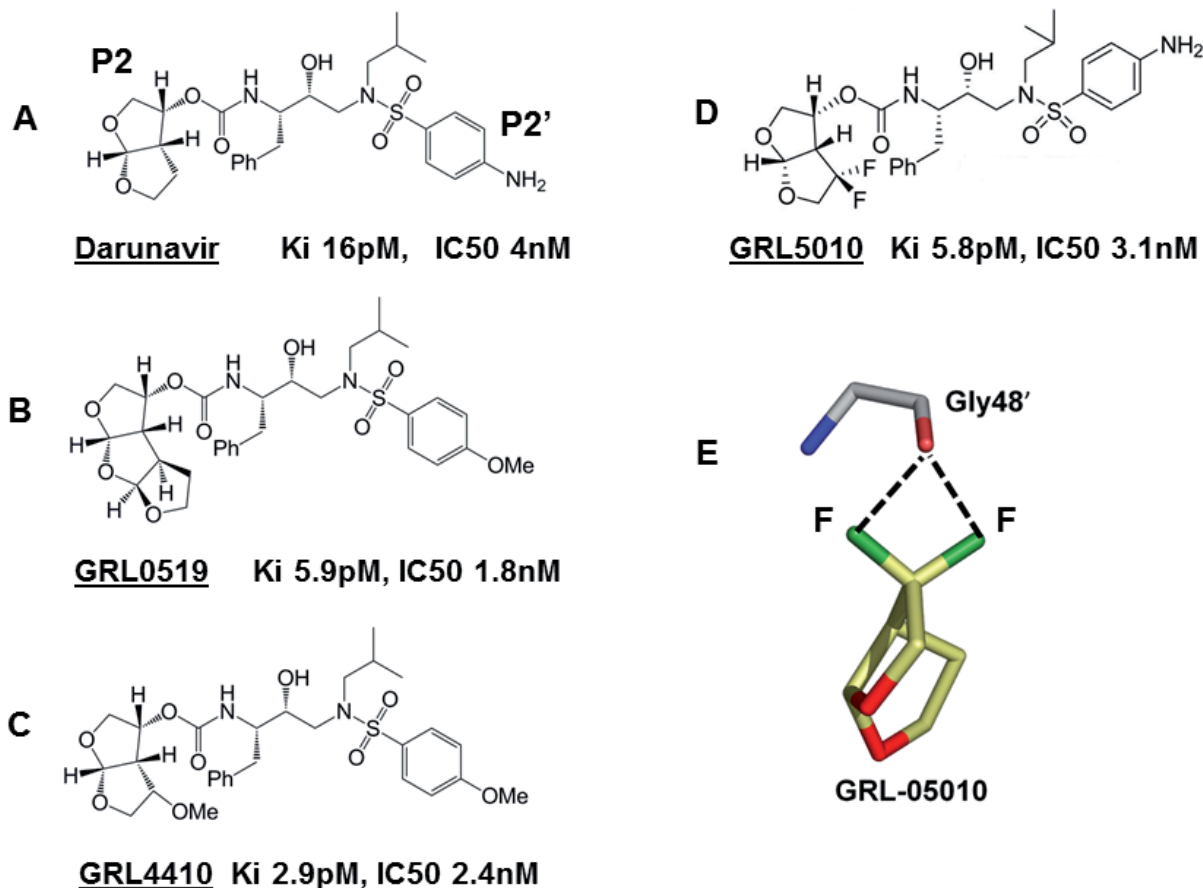


Figure 2. Chemical structure, inhibition value (K_i) and antiviral effect (IC_{50}) of darunavir and selected new inhibitors. A-D) darunavir, GRL0519, GRL4410 and GRL5010. E) Halide interactions of P2 group of GRL5010 (yellow sticks) with Gly48 in protease flap (gray sticks).

efficacy on resistant virus are shown in figure 2. These compounds were designed with larger substituents at P2 compared to *bis*-THF in darunavir and introduce new interactions with the wild type protease. We have employed the highly resistant PR20 variant to assess the binding of inhibitors [47,48]. Structural and calorimetric studies have identified antiviral inhibitors with higher binding affinity than darunavir for PR20 (Tab. 2). GRL0519 with *tris*-THF at P2 instead of the *bis*-THF of darunavir, has similar ~40 nM binding affinity for PR20. However, GRL04410 with an oxymethyl substituent on the *bis*-THF moiety and GRL5010 with *gem*-difluoro modification exhibit binding affinity for PR20 of, respectively, 10 and 20-fold better than darunavir [48]. Crystallographic analysis demonstrated that the hydrogen bond interactions of these inhibitors with protease are conserved in wild type enzyme and the PR20 structure. Inhibitors introduce new interactions with the flexible flaps, such as the unusual halide

Table 2. Binding affinity (K_i , nM) of selected antiviral inhibitors to PR20 and wild type protease.

	DRV	GRL0519	GRL04410	GRL5010
PR	0.005	0.005	0.003	0.006
PR20	41	39.5	4.3	1.7

Information in table 2 was taken from [34,47,48].

interactions with the carbonyl oxygen of Gly48 seen for GRL5010 (Fig. 2E). These novel antiviral inhibitors are promising candidates for future pharmacological development targeting highly resistant viral strains. Moreover, incorporation of fluorine in GRL5010 increases the lipid solubility and shows improved penetration of the blood-brain barrier [49]. Development of a drug based on this compound would help to eradicate the viral reservoirs in the central nervous system.

PREDICTION OF RESISTANCE FROM SEQUENCE

Currently, genotyping is recommended for new HIV infections or for individuals failing therapy in order to identify the presence of resistant mutations and guide the choice of drugs. Genotype assays for drug resistance are preferred over phenotype assays due to the advantages in terms of speed and cost [50]. The ability to rapidly and affordably sequence HIV from infected individuals opens the door to both the practical question of predicting drug resistance prior to therapy and the theoretical problem of understanding sequence and structural evolution of the virus under drug selection. Machine learning is a good method for a computational approach to these problems. Drug resistance can be predicted from genotype data by two general techniques: rule-based genotype interpretation systems and machine learning algorithms [51].

Supervised machine learning takes a labeled set of features, such as a sequence and activity data, and builds a computational model that reproduces the correlation between the features and labels [52]. The art of machine learning is in developing approaches that select meaningful features while excluding spurious features. With drug resistance due to HIV protease variants, the obvious set of features is the sequence of the protease coupled either to the relative drug resistance or a thresholded resistant/non-resistant label.

In the absence of structural data, machine learning and rule based approaches typically have predictive accuracies in the range of 60–70% [51,53]. While the overall accuracy is less than ideal, the individual tools can be remarkably consistent, for example classifying the same set of mutations as resistant with a 95% reproducibility [51]. This shows that the problem of predicting is well posed for machine learning, but that the set of features is insufficient when only sequence data are used. When structural data are included as features in the training set, the accuracy jumps into the 90–95% range independent of the type of machine learning algorithm used [53].

The limitation of using sequence data on its own is immediately obvious to a structural biologist. Proteins are not a linear set of letters, but are folded into complex and beautiful three-dimensional structures. Residues that are far apart in the linear sequence may actually be close together in space. The three dimensional context of the mutations is lost if only the sequence is used as a representation. A single mutation might contribute to drug resistance when its spatial neighbors are of one kind, but could be neutral or increase sensitivity with other kinds of neighbors. This introduces an apparent non-linearity when only sequence data are used.

Bose *et al.* [54] studied representations that could efficiently encode structure and sequence for machine learning. The encoding of the structure must be translationally and rotationally invariant because there is no privileged reference frame in the biological system. This immediately leads to using either distance measures or graph representations for the structure. Distances as features are problematic because it is difficult to define an automatic way to select which distances to use. Additionally, experimental or modeling errors in the distances mean that the data are inherently fuzzy. Graphs, which simply state that two atoms or residues are in contact with each other by some criterion, are less sensitive to errors, and can be rigorously defined. Bose *et al.* [54] examined several types of graphs and algorithms to reduce the graphs to a constant-sized data point. They found that Delaunay triangulation was the best graph for the problems they studied. They also found the graph could be reduced to an upper triangular matrix by summing over the kinds of amino acids on each end of an arc.

Yu *et al.* [53,55] applied these encodings to genotype/phenotype data for drug resistance of HIV protease and reverse transcriptase mutants. These results showed significantly higher classification accuracy than the purely sequence based approaches. In addition, the unified encoding

of sequence and structure can be used with regression analysis to predict the magnitude of resistance with high accuracy [55,56].

In the next stage of this study, mutants representing common features of high level resistance were selected from the data [56]. This selection was designed to give a tractable number of mutants for further analysis by biochemical and biophysical experiments. Combining mutants that represented high resistance to more than a single inhibitor resulted in a single sequence with high resistance to 6 drugs. The mutant with this sequence was designated PR_{S17} due to the presence of 17 substitutions relative to a standard wild type protease. PR_{S17} was verified to exhibit poor binding to six tested clinical inhibitors, 50 to >10,000-fold worse than wild type enzyme [36]. Therefore, PR_{S17} will be added to our list of extremely resistant mutants for evaluation of new antiviral inhibitors.

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

HIV infections and drug resistance are likely to continue as a problem in the absence of an effective vaccine, due to the high genetic variation, rapid turnover and existence of poorly accessible reservoirs of virus. This emphasizes the critical need for the development of targeted treatment based on genotype data and new antiretroviral drugs for both therapy and pre-exposure prophylaxis. Our studies of HIV protease tackle the challenge of drug resistance on several fronts by: 1) developing new algorithms to predict resistance from genotype data, 2) elucidating the molecular basis for resistance, and 3) incorporating this knowledge in the design strategy for novel antiviral inhibitors. Based on our studies of HIV protease and those of other groups, we propose a new paradigm for drug resistance. High level resistance to drugs does not rely on a handful of major mutations, but rather requires the coordinated effects of multiple substitutions to remodel the protease and its substrates. This new paradigm must be addressed in the interpretation of genotype data and in the design of antiviral inhibitors.

REFERENCES

1. World Health Organization <http://www.who.int/mediacentre/factsheets/fs360/en/>
2. Safrit JT, Fast PE, Giebler L, Kuipers H, Dean HJ, Koff WC (2016) Status of vaccine research and development of vaccines for HIV-1. *Vaccine* 34: 2921-2925
3. Menéndez-Arias L (2013) Molecular basis of human immunodeficiency virus type 1 drug resistance: overview and recent developments. *Antiviral Res* 98: 93-120
4. Wilton J, Senn H, Sharma M, Tan DH (2015) Pre-exposure prophylaxis for sexually-acquired HIV risk management: a review. *HIV AIDS (Auckl)* 7: 125-136
5. Santoro MM, Perno, CF (2013) HIV-1 Genetic variability and clinical implications. *ISRN Microbiol* 2013: 481314
6. Lloyd SB, Kent SJ, Winnall WR (2014) The high cost of fidelity. *AIDS Res Hum Retroviruses* 30: 8-16
7. Tang MW, Shafer RW (2012) HIV-1 Antiretroviral resistance: scientific principles and clinical applications. *Drugs* 72: e1-e25
8. Wensing AM, Calvez V, Günthard HF, Johnson VA, Paredes R, Pillay D, Shafer RW, Richman DD (2015) 2015 update of the drug resistance mutations in HIV-1. *Top Antivir Med* 23: 132-141

9. Konvalinka J, Krausslich HG, Muller B (2015) Retroviral proteases and their roles in virion maturation. *Virology* 479-480C, 403-417
10. Lee SK, Potempa M, Swanstrom R (2012) The choreography of HIV-1 proteolytic processing and virion assembly. *J Biol Chem* 287: 40867-40874
11. Miller M, Schneider J, Sathyanarayana BK, Toth MV, Marshall GR, Clawson L, Selk L, Kent SBH, Wlodawer A (1989) Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 Å resolution. *Science* 246: 1149-1152
12. Wlodawer A, Erickson JW (1993) Structure-based inhibitors of HIV-1 protease. *Annu Rev Biochem* 62: 543-85
13. Rabi SA, Laird GM, Durand CM, Laskey S, Shan L, Bailey JR, Chioma S, Moore RD, Siliciano RF. (2013) Multi-step inhibition explains HIV-1 protease inhibitor pharmacodynamics and resistance. *J Clin Invest* 123: 3848-3860
14. Wlodawer A, Vondrasek J (1998) Inhibitors of HIV-1 protease: a major success of structure-assisted drug design. *Annu Rev Biophys Biomol Struct* 27: 249-284
15. Shen CH, Tie Y, Yu X, Wang YF, Kovalevsky AY, Harrison RW, Weber IT (2012) Capturing the reaction pathway in near-atomic resolution crystal structures of HIV-1 protease. *Biochemistry* 51: 7726-7732
16. Weber IT, Waltman MJ, Mustyakimov M, Blakeley MP, Keen DA, Ghosh AK, Langan P, Kovalevsky AY (2013) Joint X-ray/neutron crystallographic study of HIV-1 protease with clinical inhibitor amprevir - insights for drug design. *J Med Chem* 56: 5631-5635
17. Gerlits O, Wymore T, Das A, Shen CH, Parks JM, Smith JC, Weiss KL, Keen DA, Blakeley MP, Louis JM, Langan P, Weber IT, Kovalevsky A. (2016) Long-range electrostatics-induced two-proton transfer captured by neutron crystallography in an enzyme catalytic site. *Angew Chem Int Ed Engl* 55: 4924-4927
18. Wensing AM, van Maarseveen NM, Nijhuis M (2010) Fifteen years of HIV protease inhibitors: raising the barrier to resistance. *Antiviral Res* 85: 59-74
19. Muzammil S, Armstrong AA, Kang LW, Jakalian A, Bonneau PR, Schmelmer V, Amzel LM, Freire E (2007) Unique thermodynamic response of tipranavir to human immunodeficiency virus type 1 protease drug resistance mutations. *J Virol* 81: 5144-5154
20. King NM, Prabu-Jeyabalan M, Bandaranayake RM, Nalam MN, Nalivaika EA, Özen A, Haliloglu T, Yilmaz NK, Schiffer CA (2012) Extreme entropy-enthalpy compensation in a drug-resistant variant of HIV-1 protease. *ACS Chem Biol* 7: 1536-1546
21. Louis JM, Aniana A, Weber IT, Sayer JM (2011) Insights into the inhibition of autoprocessing of natural variants and multidrug resistant mutant precursors of HIV-1 protease by clinical inhibitors. *Proc Natl Acad Sci USA* 108: 9072-9077
22. Davis DA, Soule EE, Davidoff KS, Daniels SI, Naiman NE, Yarchoan R (2012) Activity of human immunodeficiency virus type 1 protease inhibitors against the initial autocleavage in Gag-Pol polyprotein processing. *Antimicrob Agents Chemother* 56: 3620-3628
23. Huang L, Li Y, Chen C (2011) Flexible catalytic site conformations implicated in modulation of HIV-1 protease autoprocessing reactions. *Retrovirology* 8: 79
24. Rhee SY, Sankaran K, Varghese V, Winters M, Hurt CB, Eron JJ, Parkin N, Holmes SP, Holodniy M, Shafer RW (2016) HIV-1 protease, reverse transcriptase, and integrase variation. *J Virol* 90: 6058-6070
25. Fun A, Wensing AM, Verheyen J, Nijhuis M (2012) Human immunodeficiency virus Gag and protease: partners in resistance. *Retrovirology* 9: 63
26. Kožíšek M, Sasková KG, Rezacová P, Brynda J, van Maarseveen NM, De Jong D, Boucher CA, Kagan RM, Nijhuis M, Konvalinka J (2008) Ninety-nine is not enough: molecular characterization of inhibitor-resistant human immunodeficiency virus type 1 protease mutants with insertions in the flap region. *J Virol* 82: 5869-5878
27. Weber IT, Kneller DW, Wong-Sam A (2015) Highly resistant HIV-1 proteases and strategies for their inhibition. *Future Med Chem* 7: 1023-1038
28. Dam E, Quercia R, Glass B, Descamps D, Launay O, Duval X, Krausslich HG, Hance AJ, Clavel F (2009) Gag mutations strongly contribute to HIV-1 resistance to protease inhibitors in highly drug-experienced patients besides compensating for fitness loss. *PLoS Pathog* 5: e1000345
29. Kožíšek M, Henke S, Sasková KG, Jacobs GB, Schuch A, Buchholz B, Muller V, Krausslich HG, Rezacová P, Konvalinka J, Bodem J (2012) Mutations in HIV-1 gag and pol compensate for the loss of viral fitness caused by a highly mutated protease. *Antimicrob Agents Chemother* 56: 4320-4330
30. Özen A, Lin KH, Kurt Yilmaz N, Schiffer CA (2014) Structural basis and distal effects of Gag substrate coevolution in drug resistance to HIV-1 protease. *Proc Natl Acad Sci USA* 111: 15993-15998
31. Louis JM, Deshmukh L, Sayer JM, Aniana A, Clore GM (2015) Mutations proximal to sites of autoproteolysis and the alpha-helix that co-evolve under drug pressure modulate the autoprocessing and vitality of HIV-1 protease. *Biochemistry* 54: 5414-5424
32. Weber IT, Agniswamy J (2009) HIV-1 Protease: Structural Perspectives on Drug Resistance. *Viruses* 1: 1110-1136
33. Koh Y, Amano M, Towata T, Danish M, Leshchenko-Yashchuk S, Das D, Nakayama M, Tojo Y, Ghosh AK, Mitsuya H (2010) *In vitro* selection of highly darunavir-resistant and replication-competent HIV-1 variants by using a mixture of clinical HIV-1 isolates resistant to multiple conventional protease inhibitors. *J Virol* 84: 11961-11969
34. Louis JM, Aniana A, Weber IT, Sayer JM (2011) Insights into the inhibition of autoprocessing of natural variants and multidrug resistant mutant precursors of HIV-1 protease by clinical inhibitors. *Proc Natl Acad Sci USA* 108: 9072-9077
35. Kožíšek M, Lepšík M, Grantz Šašková K, Brynda J, Konvalinka J, Rezacová P (2014) Thermodynamic and structural analysis of HIV protease resistance to DRV - analysis of heavily mutated patient-derived HIV-1 proteases. *FEBS J* 281: 1834-1847
36. Park JH, Sayer JM, Annie Aniana A, Yu X, Weber IT, Harrison RW, Louis JM (2016) Binding of clinical inhibitors to a model precursor of a rationally selected multidrug resistant HIV-1 protease is significantly weaker than that to the released mature enzyme. *Biochem* 55: 2390-2400
37. Agniswamy J, Shen C-H, Aniana A, Sayer JM, Louis JM, Weber IT (2012) HIV-1 protease with 20 mutations exhibits extreme resistance to clinical inhibitors through coordinated structural rearrangements. *Biochemistry* 51: 2819-2828
38. Zhang Y, Chang YC, Louis JM, Wang YF, Harrison RW, Weber IT (2014) Structures of DRV-resistant HIV-1 protease mutant reveal atypical binding of DRV to wide open flaps. *ACS Chem Biol* 9: 1351-1358
39. Shen CH, Chang YC, Agniswamy J, Harrison RW, Weber IT (2015) Conformational variation of an extreme drug resistant mutant of HIV protease. *J Mol Graphics Model* 62: 87-96
40. Roche J, Louis JM, Bax A (2015) Conformation of inhibitor-free HIV-1 protease derived from NMR spectroscopy in a weakly oriented solution. *Chem Bio Chem* 16: 214-218
41. Nakashima M, Ode H, Suzuki K, Fujino M, Maejima M, Kimura Y, Masaoka T, Hattori J, Matsuda M, Hachiya A, Yokomaku Y, Suzuki A, Watanabe N, Sugiura W, Iwatani Y (2016) Unique flap conformation in an HIV-1 protease with high-level darunavir resistance. *Front Microbiol* 7: 61
42. Louis JM, Roche J (2016) Evolution under drug pressure remodels the folding free-energy landscape of mature HIV-1 protease. *J Mol Biol* 428: 2780-2792
43. Ghosh AK, Osswald HL, Prato G (2016) Recent Progress in the Development of HIV-1 Protease Inhibitors for the Treatment of HIV/AIDS. *J Med Chem* 59: 5172-5208
44. Ghosh AK, Anderson DD, Weber IT, Mitsuya H (2012) Enhancing protein backbone binding - A fruitful concept for combating drug-resistant HIV. *Angew Chem Int Ed Engl* 51: 1778-1802
45. Koh Y, Nakata H, Maeda K, Ogata H, Bilcer G, Devasamundam T, Kincaid JF, Boross P, Wang YF, Tie Y, Volarath P, Gaddis L, Harrison RW, Weber IT, Ghosh AK, Mitsuya H (2003) Novel bis-tetrahydrofuran-lyrethane-containing nonpeptidic protease inhibitor (PI) UIC-94017

- (TMC114) potent against multi-PI-resistant HIV *in vitro*. *Antimicrob Agents Chemother* 47: 3123-3129
46. Tie Y, Boross PI, Wang YF, Gaddis L, Hussain AK, Leshchenko S, Ghosh AK, Louis JM, Harrison RW, Weber IT (2004) High resolution crystal structures of HIV-1 protease with a potent non-peptide inhibitor (UIC-94017) active against multi-drug resistant clinical strains. *J Mol Biol* 338: 341-352
 47. Agniswamy J, Shen CH, Wang YF, Ghosh AK, Rao KV, Xu CX, Sayer JM, Louis JM, Weber IT (2013) Extreme multidrug resistant HIV-1 protease with 20 mutations is resistant to novel protease inhibitors with P1'-pyrrolidinone or P2-tris-tetrahydrofuran. *J Med Chem* 56: 4017-4027
 48. Agniswamy J., Louis JM, Shen CH, Yashchuk S, Ghosh AK, Weber IT (2015) Substituted bis-THF protease inhibitors with improved potency against highly resistant mature HIV-1 protease PR20. *J Med Chem* 58: 5088-5095
 49. Ghosh AK, Yashchuk S, Mizuno A, Chakraborty N, Agniswamy J, Wang YF, Aoki M, Gomez PM, Amano M, Weber IT, Mitsuya H (2015) Design of gem-difluoro-bis-tetrahydrofuran as P2 ligand for HIV-1 protease inhibitors to improve brain penetration: synthesis, X-ray studies, and biological evaluation. *Chem Med Chem* 10: 107-115
 50. Van Laethem K, Theys K, Vandamme AM (2015) HIV-1 genotypic drug resistance testing: digging deep, reaching wide? *Curr Opin Virol* 14: 16-23
 51. Prospero MC, De Luca A (2012) Computational models for prediction of response to antiretroviral therapies. *AIDS Rev* 14: 145-153
 52. Tarca AL, Carey VJ, Chen XW, Romero R, Drăghici S (2007) Machine learning and its applications to biology. *PLoS Comput Biol* 3: e116
 53. Yu X, Weber IT, Harrison RW (2013) Sparse representation for HIV-1 protease drug resistance prediction. *SIAM International Conference on Data Mining. SDM2013 Conference proceedings*, pp 342-349
 54. Bose P, Yu X, Harrison RW (2011) Encoding protein structure with functions on graphs. In *Bioinformatics and Biomedicine Workshops (BIBMW)*, 2011 IEEE International Conference, pp 338-344
 55. Yu X, Weber IT, Harrison RW (2014) Prediction of HIV drug resistance from genotype with encoded three-dimensional protein structure. *BMC Genomics* 15 (Suppl 5): S1
 56. Yu X, Weber IT, Harrison RW (2015) Identifying representative drug resistant mutants of HIV. *BMC Bioinformatics* 16 (Suppl 17): S1

Pokonać problem lekooporności wirusa HIV

Irene T. Weber^{1,✉}, Robert W. Harrison²

¹Department of Biology, Georgia State University, Atlanta, GA, USA

²Department of Computer Science, Georgia State University, Atlanta, GA, USA

✉e-mail: iweber@gsu.edu

Słowa kluczowe: HIV/AIDS, krystalografia rentgenowska, enzymy, bioinformatyczne systemy uczące się

STRESZCZENIE

Kodowana przez wirusowy materiał genetyczny proteaza HIV-1 jest skutecznym celem dla leków antywirusowych. Leczenie zakażeń HIV jest jednak znacznie utrudnione przez powszechne występowanie lekoopornych mutantów wirusa. W niniejszym artykule przeglądowym autorzy opisują stosowane przez nich kompleksowe podejście mające na celu analizę i zwalczanie lekooporności. Zrozumienie molekularnego podłoża oporności na inhibitory proteaz to kluczowy, pierwszy krok tego procesu. Następnie, zdobyta wiedza jest wykorzystywana podczas projektowania nowych, udoskonalonych inhibitorów o wysokim powinowactwie do enzymu występującego u opornych mutantów, jak i u dzikiego typu wirusa. Równoległe do badań nad różnorodnymi mutantami i inhibitorami trwa opracowywanie skutecznych algorytmów umożliwiających przewidywanie fenotypu oporności wirusa na podstawie informacji o jego genotypie. Podejście to ma ważne zastosowanie praktyczne w klinicyście, gdzie zaleca się przeprowadzenie genotypowania wirusa u pacjentów z nowymi zakażeniami.