Evolutionary Analysis of HIV-1 Protease Inhibitors: Methods for Design of Inhibitors That Evade Resistance

Daniel Stoffler, Michel F. Sanner, Garrett M. Morris, Arthur J. Olson, and David S. Goodsell*
Department of Molecular Biology, The Scripps Research Institute, La Jolla, California

ABSTRACT **Drug-resistant strains are rapidly selected during AIDS therapy because of the high rate of mutation in HIV. In this report, we present an evolutionary simulation method for analysis of viral mutation and its use for optimization of HIV-1 protease drugs to improve their robustness in the face of resistance mutation. We first present an analysis of the range of resistant mutants that produce viable viruses by using a volume-based viral fitness model. Then, we analyze how this range of mutant proteases allows development of resistance to an optimal inhibitor previously designed by computational coevolution techniques. Finally, we evaluate the resistance patterns of commercially available drugs, and we discuss how resistance might be overcome by optimizing the size of specific side-chains of these inhibitors. Proteins 2002;48:63–74.**

Key words: HIV-1 protease; drug resistance; computer-aided drug design; coevolution; AIDS therapy

INTRODUCTION

Traditional drug discovery techniques seek compounds that maximally inhibit a single target enzyme. In most cases, this is an effective approach. Most enzymes are highly specific for their substrate, so an inhibitor of similar shape and chemical nature will bind tightly to the active site. These inhibitors will also be relatively robust to resistance mutation in the enzyme, because highly specific enzymes have little range for remodeling their active sites without compromising binding and catalysis. This approach has been widely used to design inhibitors for diverse enzymatic targets, including HIV-1 protease. Protease-inhibiting drugs currently in testing and in use have been optimized to bind tightly and specifically to the wild-type protease (reviewed in Ref. 1). HIV-1 protease, however, is somewhat different from most typical enzymes: it binds to and cleaves a collection of different peptides. This broader specificity makes resistance mutation possible, because the active site is remodeled to reject inhibitors while retaining the ability to bind to the appropriate peptide sequences needed for viral maturation. This broad specificity, combined with the low fidelity of reverse transcriptase[2,3] and the high replication rate of the virus,[4] make drug resistance a major problem. When patients are treated with the wild-type-optimized drugs currently used in therapy, drug resistant strains of virus are rapidly selected.

In this report, we describe an evolutionary method to analyze the performance of a given inhibitor against a mutating target. The evolutionary approach to inhibitor design quantifies the susceptibility of a trial inhibitor to resistance as the virus mutates and reveals the relative importance and roles of individual protease mutants in this evolution. On the basis of these simulations, we predict optimal combinations of drugs for use in AIDS therapy and suggest alterations of currently-used inhibitors to improve their efficiency and ultimately evade drug resistance.

MATERIALS AND METHODS

The evolutionary simulations presented here require the evaluation of fitness for millions of mutant strains of the protease. To make this computation feasible, we have used a simple fitness evaluation model based on Michaelis-Menten kinetics using binding energies estimated by a simple measure of volume complementarity between inhibitors, substrates, and the mutant proteases. The simplifications used in this model, described below, limit the applicability of the results. Mutations distant from the active site and the effects of flexibility are not modeled. Atomic-level predictions are beyond the reach of these simulations, but the simulations do provide valuable information on optimal sizing of individual side-chains in inhibitors and general concepts for the design of resistance-evading drugs.

Modeling of Viral Fitness

Viral fitness, the likelihood that a given virus may reproduce, is evaluated for each mutated protease when challenged by a given inhibitor. As in previous work,[5,6] the fitness of a mutant virus is proportional to the rate of polyprotein processing. For a complete description, please see the previous reports; only a short summary is provided here. The reaction velocity of a mutant protease cleaving a given polyprotein processing site in the presence of an inhibitor is estimated by using Michaelis-Menten kinetics. To evaluate the overall fitness of the virus, reaction

Grant sponsor: National Institutes of Health; Grant number: PO1 GM48870; Grant sponsor: Swiss National Science Foundation: Grant number: 823A-061225.

*Correspondence to: David S. Goodsell, Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, E-mail: goodsell@scripps.edu

Received 25 September 2001; Accepted 8 February 2002

velocities are estimated for nine processing sites, and the slowest rate is taken to be the rate-limiting step, defining the viability. Kinetic constants are evaluated by using a number of simplifications. The Michaelis constant is assumed to be equal to the dissociation constant of the inhibitor to the mutant protease, which is estimated by using the volume-based binding energy described below, and a constant V_{max} is used throughout. In the current work, $[I] = 10[S] = K_M(wt)$, where $K_M(wt)$ is the estimated Michaelis constant of the uninhibited wild-type protease. Choice of different values generally changes the absolute value of the fitness but retains a similar ranking of particular mutant proteases relative to one another. It has been estimated that the reduction of protease activity to ~2% that of wild type is sufficient to block viral replication,[7] whereas restoration of protease activity to ~26% that of wild type will yield a viable resistant strain.[8]

Binding energies are estimated by using an empirically derived binding model based on the complementarity of protease subsite and inhibitor side-chain volumes. Potentials are calibrated against a set of 74 substrate sequences that are known to be cleaved, including the 63 used in previous work[5,6] and adding 11 sequences isolated by phage display techniques (personal communication from Z. Beck and J. Elder, Department of Molecular Biology at the Scripps Research Institute). A potential of mean force is created by assuming Boltzmann-type statistics, which postulates that side-chains that are commonly found in a given protease subsite will interact with favorable free energy. The potentials are parameterized on the basis of the volume of the subsite by using protein side-chain volumes evaluated from a sum of atomic volumes. The potentials were tested on a set of 55 cleaved peptides (not used in the parameterization) and 1727 uncleaved peptides,[5,6] correctly predicting 87% of the cases.

The effects of protease mutation are evaluated by using an ad hoc method. The energetic consequences of these mutations are estimated by shifting the wild-type potentials according to the magnitude of the volume change caused by the mutation. For example, a V32L mutation increases the size of the protease residue by about 18 Å^3, decreasing the size of the S2 and S2′ subsites. Thus, the S2 and S2′ potentials for this mutant are shifted by 18 Å^3 toward lower values, shifting the minimum and favoring the binding of smaller inhibitor side-chains.

Protease residues are allowed to mutate conservatively because the volume-based model does not evaluate chemical complementary. D29 and D30 are allowed to mutate to E, N, or Q; the other eight residues are allowed to mutate to any uncharged amino acid. In addition, mutations are limited to those reachable by a point mutation, for example, A mutates only to G, P, S, T, and V. Based on these restrictions, all possible one-site HIV-1 protease mutants are computed at each generation, and the viabilities (see below) of the mutants probed against the inhibitor are determined. A recursive algorithm is used to generate n-site mutations. Here, we explore a maximum depth of three-site mutations.

As in our previous work,[5] we limited protease mutations to those that directly contact the inhibitor in the active site [Fig. 1(a) and (b)]. These residues are determined by searching for contacts between protease Cβ atoms and substrate Cβ atoms that are <6 Å apart. Ten protease residues satisfy this criterion: G27, A28, D29, D30, V32, I47, G48, G49, I50, and I84. The catalytic aspartate residue, D25, also meets these criteria but is not allowed to mutate. The commonly observed mutant V82 is omitted from the list by this distance cutoff; however, this should have little effect on the results because I84 provides remodeling of the S1/S1′ pockets similar to that of V82[6] and the volume-based model used here does not evaluate directionality in the interaction between side-chains and subsites.

Subsite-Volume Computation for Inhibitors

The volume of nonstandard side-chains in each inhibitor was evaluated in two steps: (i) side-chain atoms were assigned to a given subsite by comparison to a modeled octapeptide substrate, and (ii) volumes were estimated by using a numerical approximation based on molecular surfaces. A generic, non-native substrate $(Ala)_8$ octapeptide bound to HIV-1 protease was built by using the structure of the substrate-form of LP-149 complexed with FIV D30N protease, PDB accession code 3fiv.[9] These FIV-PR structures provided valuable experimental data about the orientation of the scissile peptide bond. The backbone torsion angles of the substrate form of LP-149 were used to define the conformation of the octapeptide. The conserved residues of the active site region of the FIV D30N mutant, LP-149 complex, were superimposed on the corresponding residues of HIV-1 protease. Once constructed, the sequence of the $(Ala)_8$ octapeptide was modified and energy minimized by using the AMBER force field and the Discover module of the InsightII 97.0 package (Molecular Simulations, Inc.). Two water molecules were included in the model: (i) the water that bridges the tips of the flaps and donates two hydrogen bonds to two backbone carbonyls of the substrate ("water 301") and (ii) the lytic water that binds between the catalytic aspartic acids.

The model octapeptide complex was superimposed with the crystallographic structures of the inhibitor complexes. Atoms of the drug molecule were assigned to the subsite corresponding to the nearest Cβ atom in the modeled octapeptide. In cases in which side-chains were particularly large, this resulted in the splitting of a single functional group into two adjacent subsites. To calculate the volume of side-chains in each subsite, a solvent-excluded surface[10] was computed for each of the subsets of atoms assigned to a given binding site. These surfaces were computed by using MSMS.[10] MSMS reports a numerical approximation of the volume inside such a closed surface.[11]

Data Visualization

The evolutionary search method presented here produces large numbers of protease/inhibitor interactions to be evaluated. Visualizing these large data sets is a difficult

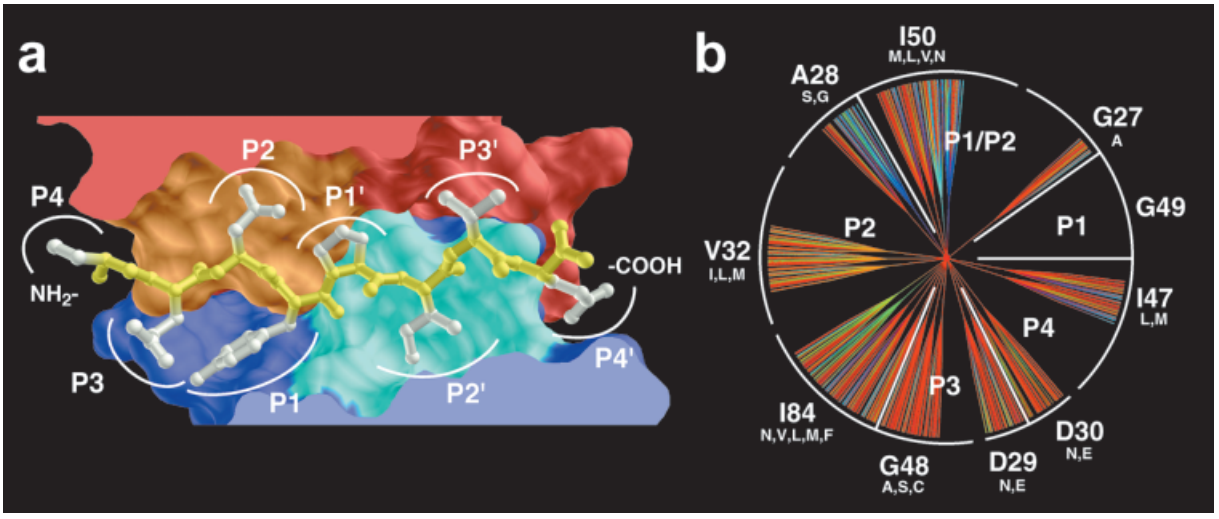

Fig. 1. Active site of HIV-1 protease. HIV-1 protease is represented by molecular surfaces and ribbon diagrams with subunit A in red and B in blue. The 11 protease residues that were allowed to mutate in this study (D25, G27, A28, D29, D30, V32, I47, G48, G49, I50, and I84) are colored orange and cyan, respectively. A ball-and-stick representation of a natural polyprotein substrate (SQNYPIVQ) is bound in the active site, with the backbone in yellow and side chains in white. View of the entire complex. **a:** Close-up of the active site, with the eight subsites (denoted S4–S1 and S1′–S4′) labeled. Peptides bind to HIV-1 protease in extended form with eight contiguous residues on the peptide, labeled P4–P4′, making contact with the eight enzyme subsites S4–S4′. Images were created with the Python Molecular Viewer.[12] **b:** Evolutionary diagram of the set of all viable HIV-1 protease mutants that cleave the polyprotein-processing sites. For ease in comparison, HIV-1 protease mutants are ordered by location within the active site, counterclockwise from S1 to S4 starting at the 3 o'clock position. For each site of mutation, the possible mutants are sorted by amino acid size, starting with the smallest, glycine, moving counterclockwise to the largest, tryptophan. In this diagram, each mutated protease is assigned a unique position, which is used in all subsequent diagrams allowing easy comparison with the diagrams in other figures. Mutated proteases with a viability $>$ 5% are represented by a dot that is connected by lines to viable parents and children. Protease mutants with a viability $<$ 5% are not drawn, leaving "holes" in the diagram. The colors represent the viabilities of the proteases: dark-blue for low viabilities of 5–10%, light-blue for 11–25%; green for 26–50%, yellow for 51–75%, orange for 76–99%, and red for 100% viability and higher. The single letter amino acid codes, such as I50 or A28, refer to the one-site mutants at the first step of each branch.

task. We have developed an interactive evolutionary tree diagram to sort and display the predicted drug resistance patterns and navigate through the data. Visualization of the evolutionary tree diagram is done with PMV (Python Molecular Viewer,[12] available on the www at http://www.scripps.edu/~sanner/python/), an extensible and platform-independent molecular viewing and modeling environment based on the object-oriented programming language Python (http://www.python.org). Evolutionary tree diagrams are computed by a recursive algorithm written in Python and visualized in PMV. Each "parent" protease, displayed as a small dot, is connected by a line to its "children" proteases. All proteases are "objects" containing individual information such as viability and mutated amino acid sequence. The data may be accessed by mouse-clicking on a particular protease; thus, the evolutionary tree diagram serves as an interactive visual database. A graphical user interface allows users to reshape the evolutionary tree diagram interactively by modifying viability threshold and other criteria.

Two types of diagram have proven useful. The first is a uniform format tree, with each specific mutation in a standard location. These are used in all of the figures, allowing comparison of evolutionary trees from inhibitor to inhibitor. The second is an expanded version, spreading the viable branches across the entire circle as included in Figure 2. This representation allows individual branches to be explored in additional detail. Further information on this graphical tool will be presented separately.

Subsite Volume Optimization

Each inhibitor side-chain volume was optimized individually while holding the remaining subsites constant. Evolutionary diagrams were calculated for incremental increases and reductions in the volume. Volumes were selected first that reduced the number of branches within the diagram, then that reduced the viability of individual strains within the remaining branches. This optimization was performed manually, because the search space was relatively simple. Because the protease is limited to changes between the 20 natural amino acids, most often forcing a large change in volume, changes in the evolutionary diagrams were abrupt and severe as the inhibitor side-chain volumes were changed, making determination of optimal volumes unambiguous.

RESULTS AND DISCUSSION

We separate this report into three general sections. First, we analyze the range of HIV-1 protease mutations that yield viable viruses, quantifying the mutational plasticity of the virus. Then, we study the structural basis of a robust inhibitor developed in previous studies, exploring how this inhibitor covers most of the range of viable mutants. Finally, we analyze the susceptibility of existing drugs used for treatment of AIDS and make suggestions for modifying these drugs to improve their robustness in the face of viral resistance mutation.

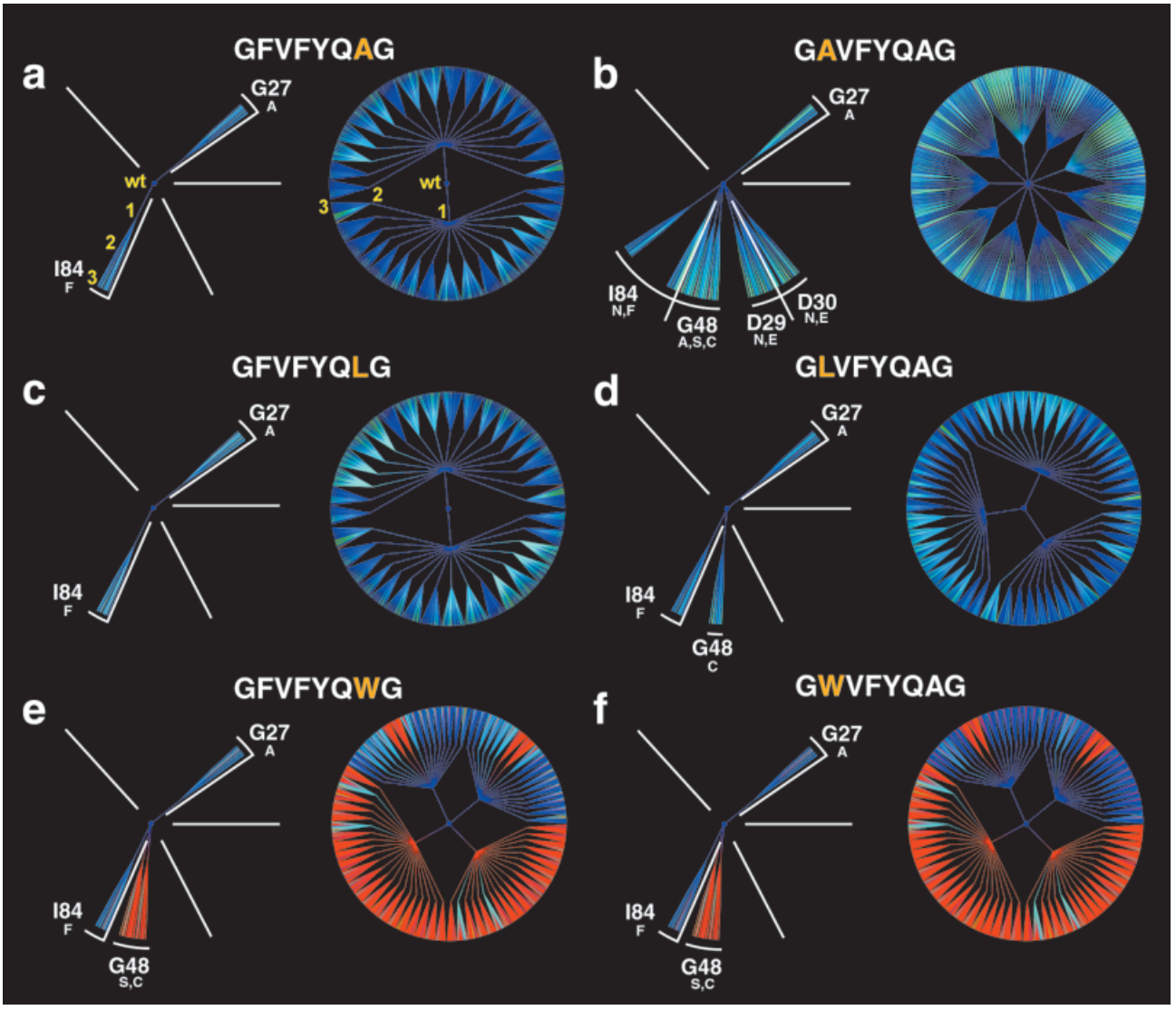

Fig. 2. Resistance susceptibility of the coevolutionary inhibitor (GFVFYQAG) with modifications at P3 and P3′. **a:** Evolutionary diagram of viable mutants when inhibited by the coevolutionary inhibitor. The diagram on the left is arranged and colored as in Figure 1(c), and the diagram on the right shows the same branches, but spread equidistantly around the circle to allow inspection of individual mutants. **b:** Alanine at position P3 induces multiple resistances, yet with low viabilities. **c**. Leucine at P3′ yields almost no effect, but **(d)** leucine at P3 induces a G48C resistance. Tryptophan at phenylalanine at P3′ (**e**) or P3 (**f**) induces strong resistances with viabilities of 100% for G48 mutants.

Limits of Mutational Freedom in HIV-1 Protease

The first step toward isolating the structural features that are important in developing resistance-evading inhibitors is to quantify the range of mutations that are possible within each subsite,[6] evaluating the mutational flexibility of protease residues in each individual subsite. The range of mutations that are available to HIV-1 protease is limited by the need to bind and cleave the nine polyprotein processing sites. Unlike most enzymes, HIV-1 protease features a broad specificity, cleaving a variety of different peptide sequences within the viral polyprotein.[13] This gives the protease more latitude for mutation than a typical enzyme. Successful mutants will remodel the active site to reduce binding of inhibitors while still binding and cleaving the processing sites. The eight subsites show different mutational plasticity, dependent on the size of the subsite and the range of sizes of the substrate sidechains that must be accommodated.

The range of viable mutant proteases is shown in Figure 1(c). In this graph, no inhibitors are included in the fitness calculation, so the graph includes the entire set of proteases that retain the ability to cleave the polyprotein processing sites in the absence of inhibitor. Looking at the single site mutants, we can get a measure of the mutational plasticity of the different protease subsites. Two sites that remodel the S1 and S1′ sites may mutate over large ranges: G27 can increase by 26 Å^3 to alanine, and I84 can mutate over the range from 57 to 102 Å^3, from asparagine to phenylalanine. The I84 mutation (and V82 mutation, with similar contacts but not modeled in this

study) are commonly observed in patients treated with protease inhibitors (Stanford HIV Database: http://hivdb.stanford.edu/hiv/). Similarly, the G48 position, which contacts S3 and S3′, is free to increase in size by 44 Å^3 to cysteine, a mutation also observed in clinical isolates. The S2/S2′ sites, however, are more limited: I47 mutates only to leucine or methionine, a change of only 1 Å^3, and A28, I50, and V32 are also limited to a range of mutants with volumes not much larger than the wild type.

Optimal Coevolutionary Inhibitor

When a virus is challenged with an inhibitor, it is limited to the range of viable mutants shown in Figure 1(c) when developing a drug-resistant strain. All other mutants are unviable, even in the absence of inhibitor. The peptidomimetic inhibitor GFVFYQAG, the most robust inhibitor found by computational coevolution techniques,[5] is very successful against these viable strains. As depicted in Figure 2(a), most of the viable mutant strains are effectively inhibited. Challenging the set of allowed protease mutants with the coevolutionary inhibitor results in only two weakly resistant protease mutant strains, G27A and I84F.

These evolutionary tree diagrams may be used to evaluate the importance of each side-chain in the inhibitor for the development of resistance. In Figure 2, we modify the P3 and P3′ side-chains and evaluate the resistance profiles of the modified inhibitors. The P3/P3′ position is interesting because inhibitors with large side-chains at P3 or P3′ are particularly susceptible to viral resistance.[14,15] The native protease favors large side-chains at this position. However, the polyprotein processing sites all contain amino acids with smaller side-chains, so the protease is free to mutate, reducing the size of the S3/S3′ subsites. This excludes inhibitors with large side-chains at P3/P3′ but still allows binding and processing of the viral polyprotein.

This behavior is reflected in the evolutionary analysis, presented in Figure 2. Replacing the large phenylalanine at P3 with a much smaller alanine induces multiple resistances, as expected, yet with overall low viabilities [Fig. 2(b)]. However, alanine at P3′ [Fig. 2(a)], which is identical with the coevolutionary inhibitor, is an excellent inhibitor. Replacing phenylalanine at P3′ with a medium-sized leucine has almost no effect [Fig. 2(c)]; however, replacing alanine at P3 with leucine induces a G48C resistance mutation [Fig. 2(d)]. Finally, as predicted, replacing phenylalanine at P3 or alanine at P3′ with a much larger tryptophan [Fig. 2(e) and (f)] induces strong resistances with wild-type viabilities for G48 mutants. These diagrams show that an optimal inhibitor side-chain should be large, but not too large, at position P3, and that position P3′ should be small or medium-sized.

Resistance to Clinically Applied Drugs

Many of the most successful inhibitors are designed on the basis of rapidly cleaved polyprotein-processing sites, placing a transition-state mimetic at the P1-P1′ scissile bond.[16] Side-chains are then optimized for maximal fit to the wild-type virus, creating nanomolar-level inhibitors but often opening avenues for viral drug resistance. Our evolutionary analysis identifies the features on the inhibitor that are susceptible to resistance mutation, which may be used (as described in the next section) to optimize the drugs for performance against probable mutants that may develop during treatment. Note that these results are based on a simple volume-based energy model, so the results must be taken as suggestions and not as precise atomic-level predictions.

The evolutionary analysis of resistance against saquinavir is shown in Figure 3(b). G48ASC mutants are the most resistant, followed by the highly resistant mutants D29NE and I84N. In addition, G27A, A28S, D30NE, V32ILM, I47LM, I50ML, and I84VLMF mutants are found, yet with a much lower viability. However, if these one-site mutants are allowed to mutate further, many two-site and three-site mutants possess a considerably higher viability. These results are consistent with increased side-chain volume of G48V mutations observed in laboratory isolates[17] and clinical isolates,[18] and common observation of the I84V mutation.[19]

The evolutionary graph for ritonavir is shown in Figure 3(e). The ritonavir diagram has the largest range of viable mutant strains of the four drugs analyzed here. Our evolutionary analysis predicts highly viable D30E, V32ILM, and G48C mutants when probed against ritonavir, and slightly weaker G27A, D30N, and I84NF mutants. This is consistent with experimental observations: resistance to ritonavir is higher than resistance to any other inhibitor.[20] The major resistance mutations observed in laboratory isolates are V32I, G48V, and I84V[21,22]; and G48V and I84V in clinical isolates.[18]

Indinavir resistance mutations are presented in Figure 4(b). Our viral fitness model predicts several weakly viable mutants when probed against indinavir, including G27A, D29N, D30NE, G48ASC, and I84VF. Only the D29E mutant exhibits a slightly raised viability. Commonly detected in laboratory isolates are G48V, I84V, and L90M[23] and clinical isolates contain G48V, I84V mutants.[18] V32I mutants with lower viability are also observed in both laboratory and clinical isolates. The observed V32I mutation is absent in our predicted model, revealing the limitations of a volume-based approach. Recent studies report that all indinavir resistance mutations reported are also selected by ritonavir, except I64V, and most of them are also selected by saquinavir.[24] Comparing the various diagrams presented in this study reveals a consistent pattern: the indinavir diagram is a subset of the saquinavir [Fig. 3(b)] and ritonavir [Fig. 3(e)] diagrams.

Nelfinavir resistance mutations are presented in Figure 4(e). The evolutionary analysis predicts G48ASC mutants as the most resistant when inhibited by nelfinavir, followed by G27A and D29NE mutants. In addition, low-viability mutants D30NE, I47LM, I50ML, and I84NVLMF are detected. The mutations observed in laboratory isolates and clinical isolates are D30N, G48V, and I84V.[22,25]

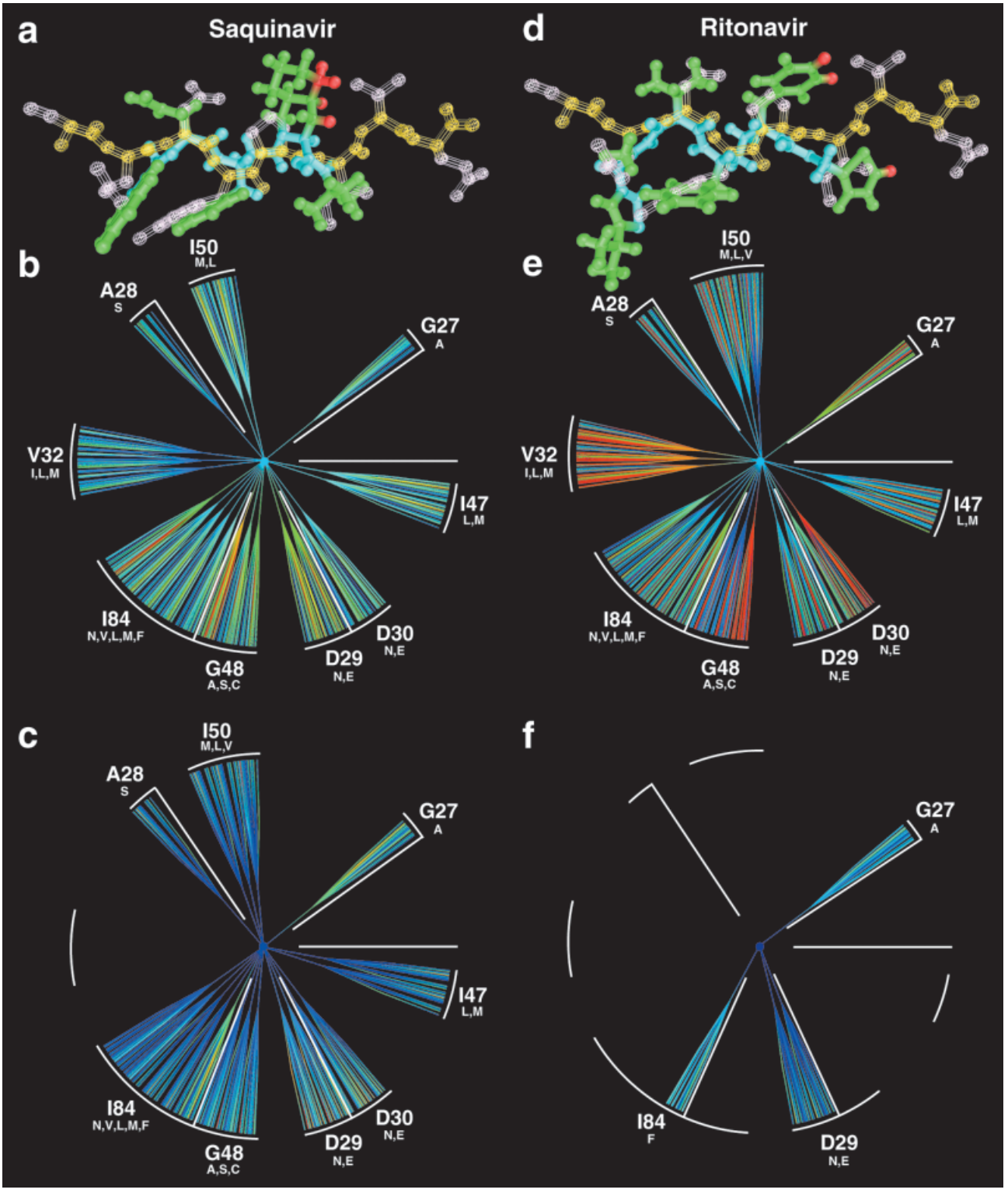

Fig. 3. Drug resistance patterns for saquinavir and ritonavir. **a:** Ball-and-stick representations of saquinavir (filled bonds) superimposed on HIV-1 protease natural substrate (SQNYPVIQ, wireframe bonds). The saquinavir backbone is cyan and side-chains are green, and side-chain atoms, which reach into a neighboring protease subsite, are red. Saquinavir occupies subsites S3 through S3′, leaving S4 and S4′ empty. **b:** Drug resistance pattern of saquinavir depicted as evolutionary tree diagram. D29NE, G48ASC, and I84N mutants are the most resistant when inhibited by saquinavir. **c:** Drug resistance pattern for an optimized version of saquinavir. This theoretical compound lowers viabilities of almost all saquinavir-resistant mutants and completely eliminates the V32ILM resistance mutants. However, it triggers a weakly viable I50V mutant and increases viabilities of G27A and I84F mutants. **d:** Ball-and-stick representation of ritonavir, which occupies subsites S3 through S4′, leaving S4 empty. **e:** Drug resistance pattern of ritonavir. Ritonavir is the most susceptible of the five inhibitors used for this study, allowing development of many resistant strains of the virus, including G27A, D30EN, V32ILM, G48C, and 184NF. **f:** Drug resistance pattern of an optimized version of ritonavir. The optimized ritonavir is more efficient than the optimized saquinavir because it offers an additional side-chain (P4) to optimize.

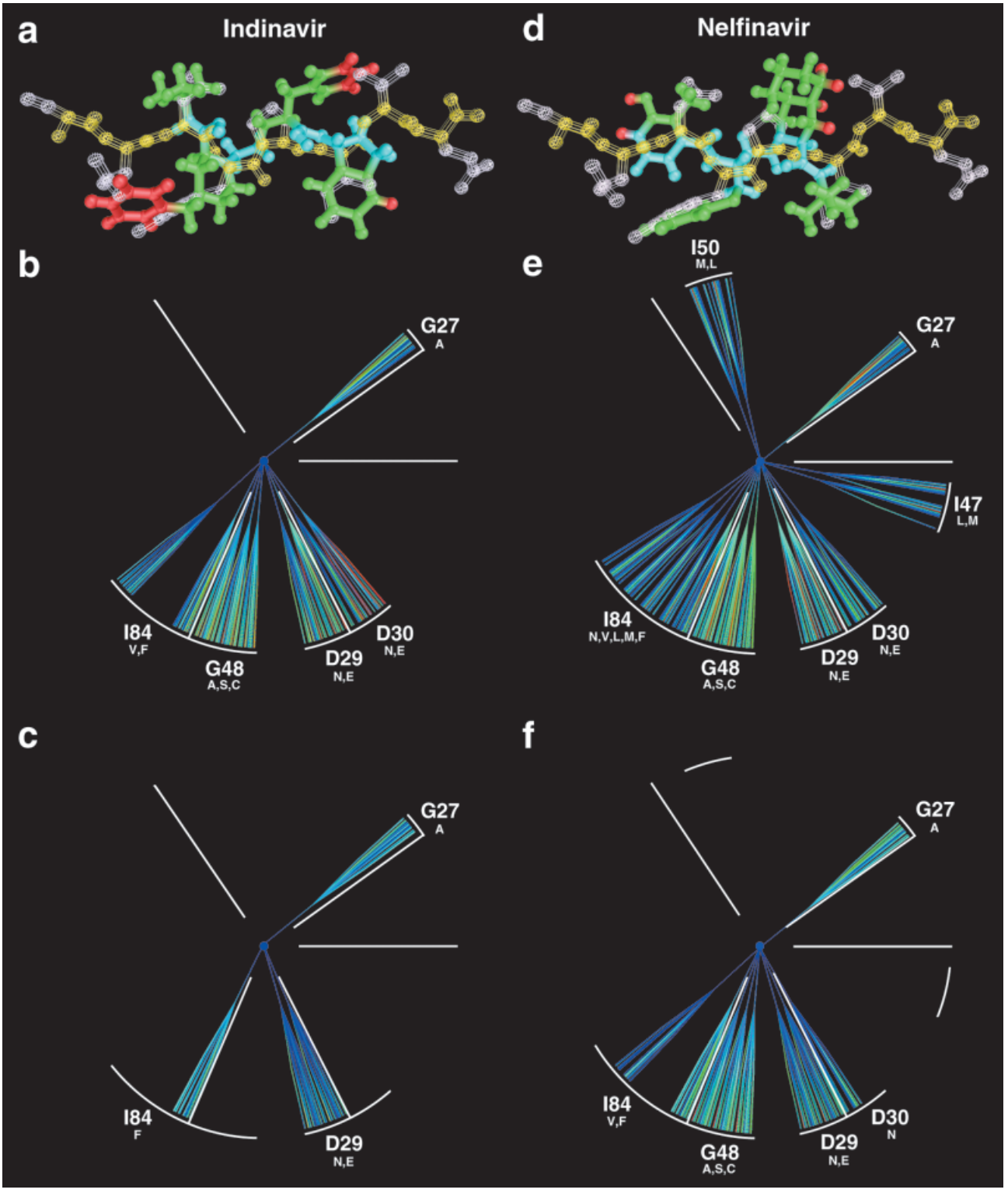

Fig. 4. Drug resistance patterns for indinavir and nelfinavir. Representation and coloration as in Figure 3. **a:** Indinavir side-chains occupy enzyme binding pockets S3 through S4', leaving S4 empty. **b:** Drug resistance pattern of indinavir. Weakly viable one-site mutants, such as G27A, D29NE, D30NE, G48ASC, and I84V, are selected when inhibited by indinavir. **c:** Drug resistance pattern for an optimized indinavir, which is the same as optimized ritonavir [Fig. 3(f)], because they both occupy the same protease subsites. **d:** Nelfinavir occupies enzyme-binding sites S4, and S2 through S3′, leaving S3 and S4' empty. **e:** Drug resistance pattern of nelfinavir. G27A, D29NE, and G48ASC mutants are the most resistant when inhibited by nelfinavir. **f:** The drug resistance pattern of optimized nelfinavir, with nine weakly viable resistant mutant strains.

TABLE I. Side-Chain Volumes for HIV-1 Protease Inhibitors

	Side-Chain Volumes (Å^3)							
Inhibitor	P4	P3	P2	P1	P1′	P2′	P3′	P4′
Saquinavir	−1	109	57	83	119	79	46	−1
Ritonavir	−1	118	62	101	97	87	15	0
Indinavir	−1	83	77	102	101	81	51	5
Nelfinavir	15	−1	58	107	129	80	31	−1
TL-3	37	26	62	98	93	62	29	13
Optimal	**0**	**115**	**60**	**109**	**100**	**60**	**30**	**0**

Values represent side-chain volumes; −1 indicates absent side-chains. Note that the value of zero in the optimized P4 and P4′ positions refer to glycine, with a side-chain of zero size, not the absence of a side-chain. The optimized drug depicts the most efficient side-chain volumes for each listed inhibitor, omitting drug side-chains marked with −1, for example, optimized saquinavir side-chain volumes are −1, 115, 60, 109, 100, 60, 30, and −1.

Evolutionary Optimization of Clinically Applied Drugs

Using evolutionary analysis, we have optimized each of the clinically used drugs, identifying and modifying those features of the drugs that are used to compromise binding in resistant strains. The results are included in Figures 3(c), 3(f), 4(c), and 4(f), with numerical values in Table I. In all cases, the improved versions reduce the viabilities of almost all drug-resistant mutants and completely eliminate the common V32ILM resistance mutants. The results reveal three important structural features that are important in the design of robust inhibitors: (i) compensatory changes in P1/P1′ and P2/P2′, (ii) proper sizing of P3/P3′, and (iii) the advantages of including P4 and P4′ residues.

Optimizing the center four subsites is a delicate problem, requiring a trade-off between optimization at P1/P1′ and P2/P2′. Looking at the saquinavir optimization (similar effects are seen in the other drugs), size changes in P1 and P1′ alone, either larger or smaller, do not lower or remove the G27A resistance, which even persists in the inhibitor found by a coevolutionary approach. However, the viability of resistance mutants in the neighboring S2 and S2′ subsites are lowered by optimizing the P1 and P1′ side-chains by slightly increasing the size of the saquinavir side-chain P1 and decreasing the large saquinavir P1′ side-chain to a medium-sized amino acid. A small increase in the saquinavir P2 side-chain considerably lowers the resistance of A28S, V32ILM, I50V, and I84NVF mutants. In contrast, the saquinavir P2′ side-chain is much larger than the optimal size suggested by the S2′ energy potentials. Decreasing P2′ to an optimal volume has the same effect as increasing P2; it lowers the resistance of the mutants described for P2. However, it is the combined effect of both optimizing P1/P1′ and P2/P2′ that completely eliminates V32ILM resistance, but it induces the weakly resistant I50V mutant in the process.

Optimal sizing of the P3/P3′ side-chains is also essential. As shown in Figure 3(c), the most resistant mutation observed against saquinavir, G48A [red lines in Fig. 3(b)], may be significantly reduced by increasing the P3 side-chain and decreasing the P3′ side-chain. Crystal structure data of an HIV-1 protease G48V mutant complexed with saquinavir recently revealed new insights into the mechanism of drug resistance at position G48.[26] Because of an increased gap between the mutated protease V48 and the P3 saquinavir quinoline ring, the van der Waals interactions between saquinavir and the protease flap are diminished by >18 contacts compared to the wild-type complex. In addition, a hydrogen bond between the G48 carbonyl oxygen of the wild-type enzyme and the backbone amino group at the P2 subsite of saquinavir is absent in the mutant structure. This observation is in agreement with our suggestion to increase slightly the size of the saquinavir P3 side-chain, to fill the described gap to evade drug resistance. However, as discussed above, there is an upper limit to this P3 size increase, because side-chains that are too large face strong resistance by direct steric exclusion.

Finally, the ideal inhibitor should fill all eight subsites. Saquinavir occupies subsites S3 through S3′ but leaves S4 and S4′ empty. Hence, the virtual drug also leaves S4/S4′ empty, enhancing the development of resistance mutants. At a glance, one might expect that adding P4 and P4′ would give the virus new targets for resistance mutations; however, as discussed above, S4 and S4′ are restrained in mutational freedom, binding a wide range of amino acids in the processing sites. Thus, the virus is not free to mutate the protease in S4 and S4′, because it must continue to accommodate this wide range of P4 and P4′ side-chains in its substrates. Adding P4 and P4′ substituents to saquinavir would dramatically improve drug efficiency by adding additional binding strength with little chance of resistance development, resulting in a drug with only two remaining resistance strains, G27A and I84F, as in the coevolutionary inhibitor depicted in Figure 2(a).

Inhibitor TL-3

TL-3 [Fig. 5(a)] is a novel C2-symmetric inhibitor with a high potency *in vitro* against simian, feline, and human immunodeficiency viruses.[15] It was specifically designed to inhibit FIV protease, yet tests with clinical isolates have shown it is also efficacious against HIV-1 protease mutants that are resistant to commercial drugs.[27] Common to four isolates resistant to TL-3 is a V82A mutation increasing the IC50 by fourfold compared with wild-type protease. In addition, familiar three-site mutations (M46I/F53L/V82A) increase resistance 10-fold and six-site mutations (L24I/M46I/F53F/L63P/V77I/V82A) increase resistance 29-

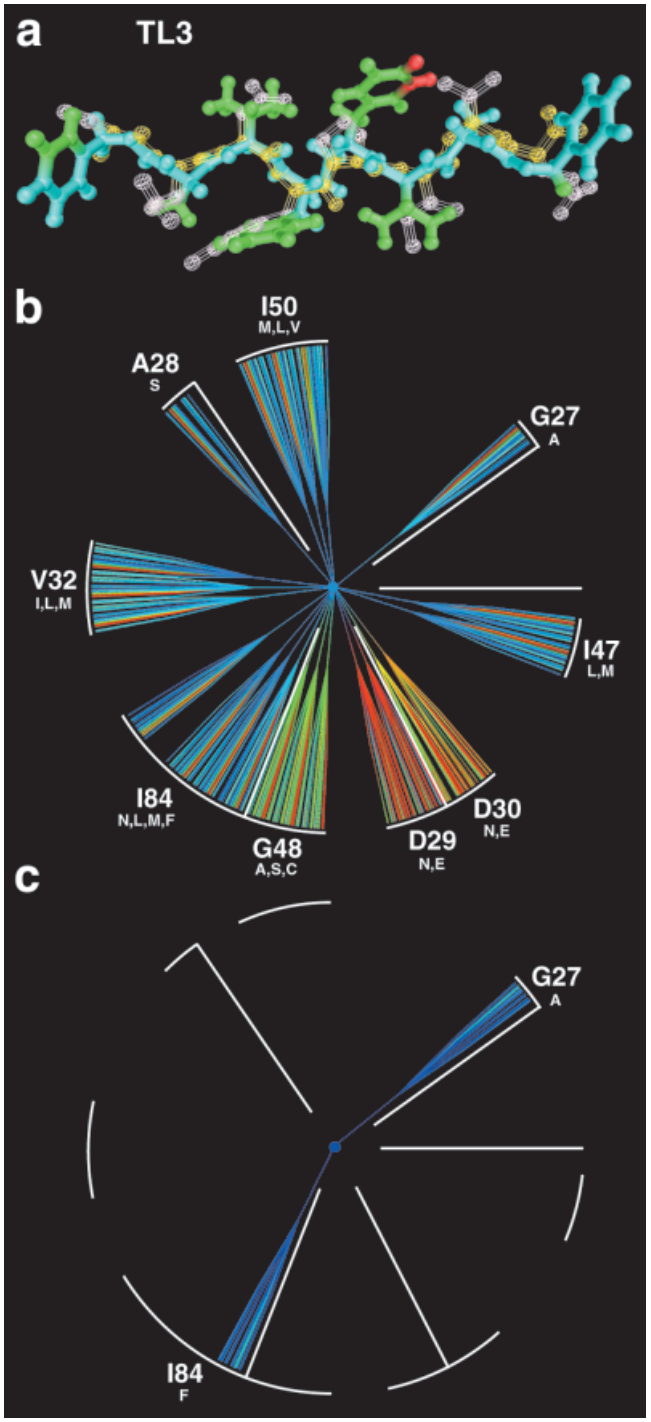

Fig. 5. Drug resistance pattern for TL-3. **a:** TL-3 occupies all eight enzyme binding pockets, S4 through S4′. Representation and coloration as in Figure 3. **b:** Drug resistance pattern of TL-3 depicted as evolutionary tree diagram. TL-3 cross-resistance to saquinavir, ritonavir, and nelfinavir, but not to indinavir, occurs in highly viable D29NE, D30NE, and G48ASC mutants, which are seen in all of these evolutionary trees. Most one-site mutations exhibit a low viability; however, two- and three-site mutants significantly increase protease resistance. **c:** Optimized TL-3 removes all resistant protease mutants except I84F and G27A, which cannot be overcome with any of the drugs used in this study.

fold, providing cross-resistance to saquinavir, ritonavir, and nelfinavir.

This behavior is reflected in our predicted resistance diagram [Fig. 5(b)]. Most one-site mutations exhibit a low viability if probed against TL-3 (G27A, A28S, V32ILM, I47LM, and I84NLMF). In agreement with Buehler's observation, two- and three-site mutants significantly increase protease resistance. Cross-resistance to saquinavir, ritonavir, and nelfinavir, but not to indinavir occurs in highly viable D29NE, D30NE, and G48ASC mutants, mainly induced by the two large TL-3 Cbz groups at P4 and P4′, respectively (see below).

TL-3 occupies all eight binding pockets; thus, the optimized virtual drug (Table I) is expected to be highly potent and to show a similar pattern as that observed with the best inhibitor from coevolutionary studies. Indeed, the optimized virtual inhibitor [Fig. 5(c)] is as efficient as the coevolution inhibitor [Fig. 2(a)]. In fact, a closer inspection reveals even a slightly higher efficiency of the virtual inhibitor because the volume optimization is not restricted by amino acid volumes. Individual subsite optimization for P3 through P3′ is performed as described for saquinavir, yielding similar results. Decreasing the large Cbz group at P4 to the size of a glycine (Table I) removes the D29NE resistance strain from the diagram. The energy potentials also favor a small side-chain at P4′, but the Cbz group at this position protrudes in a different direction and is not treated as the S4 side-chain [Fig. 5(a)]. Thus, our optimization does not suggest reduction of this group.

After optimization, I84F and G27A resistance mutants remain, which cannot be removed by modification of any of the drugs used in this study. Both G27 and I84 are caught in the tug-of-war between P1/P1′ optimization and P2/P2′ optimization described in the saquinavir section above, and allowance of the weak resistance provided by these mutations seems to be the best compromise possible for inhibitor design.

Multidrug Therapy

Ideally, multidrug therapy should combine drugs that inhibit different branches of the evolutionary tree, with each drug covering any weaknesses of the other. Figure 6 shows evolutionary diagrams for HIV-1 protease for treatment with the six possible pairs of the drugs tested here. Evolutionary diagrams for multidrug therapy are evaluated by testing mutant viruses with each drug separately and then taking the lowest viability. Thus, the technique does not take into account any synergy between the drugs when acting on a single mutant virus; only the coverage of the set of drugs over the entire evolutionary diagram are considered. The results are consistent in most cases with the recommendations for therapy presented by the AIDS Treatment Information Service (ATIS) published by the Department of Health and Human Services at the National Library of Medicine (http://www.hivatis.org).

In all of the evolutionary diagrams, combination therapy with two drugs decreases the number of viable mutant strains and reduces the viability of individual strains relative to monotherapy. Combinations with indinavir are

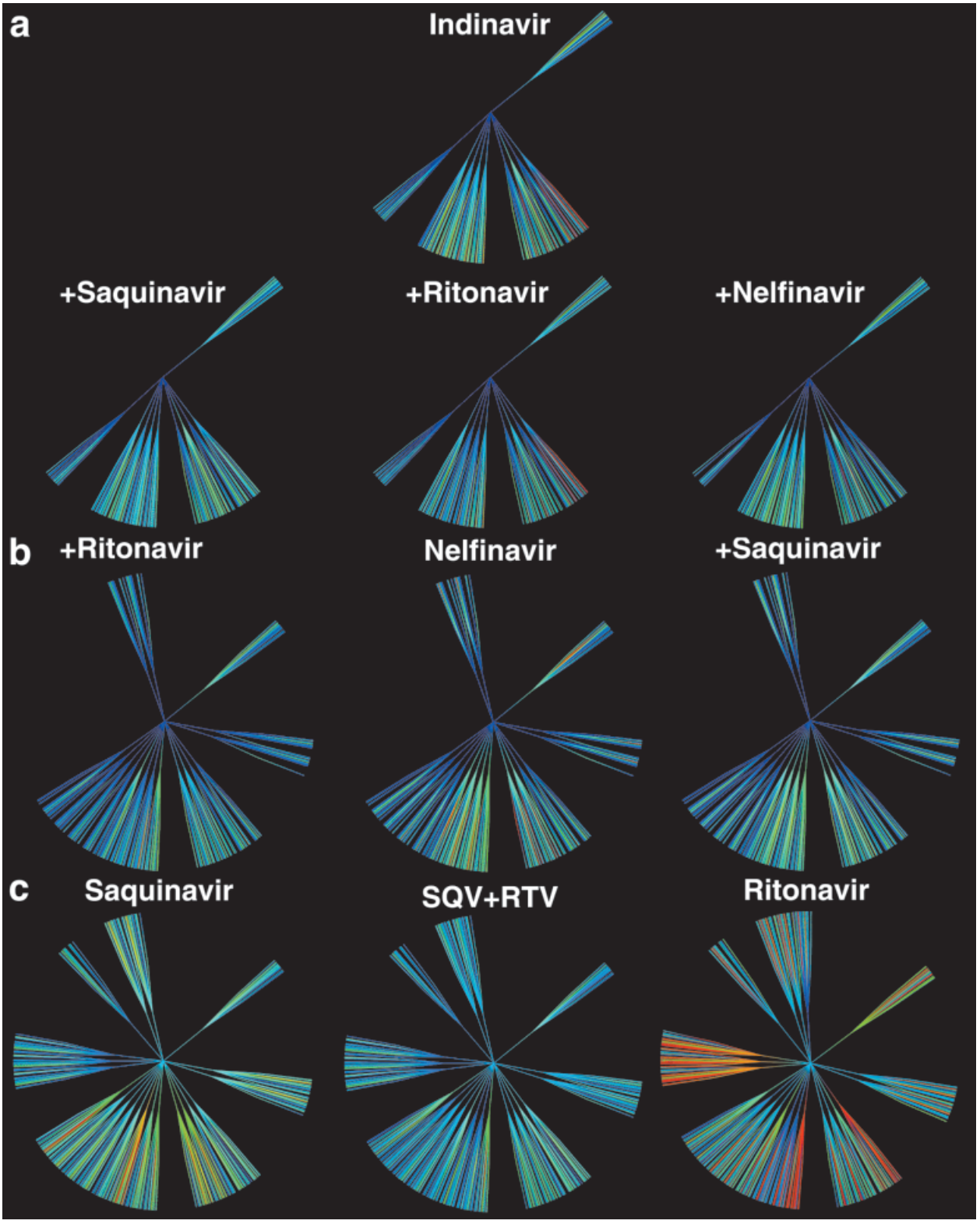

Fig. 6. Drug resistance patterns for combination therapy with two protease inhibitors. Representation and coloration as in Figure 3. **a:** Combinations with indinavir show slight improvement over treatment with indinavir alone, showing the same pattern of resistant strains, but reducing the viability of each strain. **b:** Combinations with nelfinavir also show reduction in the viability of resistant strains. **c:** Combination of saquinavir with ritonavir shows marked reduction of the number of resistant mutants relative to when the drugs are used separately.

the best choices based on the diagrams, shown in Figure 6(a). The best combination is saquinavir and indinavir, which reduces the number and viability of mutants to minimal values. The indinavir/nelfinavir and indinavir/ritonavir combinations also show a good coverage in the diagrams, but with a few more viable mutants than the saquinavir/indinavir combination. Of the three indinavir combinations, only the indinavir/ritonavir combination is included in the ATIS recommended list because indinavir alone is nearly as effective as the combinations of two protease inhibitors (when in combination with RT inhibitors). The indinavir/ritonavir combination is "strongly recommended" because addition of ritonavir has been shown to reduce the metabolism of other drugs administered concurrently, an effect that is not modeled in our work.

Two combinations with nelfinavir also perform well, reducing the viability of resistant strains relative to therapy with nelfinavir alone, as shown in Figure 6(b). In the ATIS recommendations, saquinavir/nelfinavir is an alternate recommendation and ritonavir/nelfinavir is given no recommendation because of insufficient data.

Finally, the combination of saquinavir with ritonavir shows the poorest of the six possible combinations, as shown in Figure 6(c). The saquinavir/ritonavir combination shows significant reduction of the many resistant strains seen for the drugs alone, but still has the highest number of resistant strains of all of the combinations. In contrast to these results, this combination has been shown to be effective in clinical studies[28] and is strongly recommended by the ATIS treatment guidelines. This may be attributed to the synergistic activity of ritonavir with saquinavir and other drugs, as mentioned above.

CONCLUSIONS

Robust inhibitors may be designed by optimizing structures in the context of the mutant proteases that we might expect a patient to encounter or develop. These side-chain optimization studies highlight a few sites for easy resistance development in HIV-1 protease, which provide quick ways to improve the robustness of a given inhibitor. The P3 side-chain is particularly easy to see. Ritonavir was designed to fill the S3 site with a very large P3 side-chain. This allows the virus to develop resistance easily with G48 mutations, by simple constriction of the S3 subsite. Knowing that this mutation is probable, we can reduce the ritonavir P3 side-chain, lowering G48 resistance from 100% to 5%, and evading this resistant mutant before it develops. Looking at the inhibitors tested here, we also see the advantages of including P4 and P4′ side-chains. These positions add binding strength, and as long as they are kept small, they are free of susceptibility to resistance mutation. Of course, these advantages must be weighed against any pharmacological disadvantages of larger compounds.

ACKNOWLEDGMENTS

The authors thank Sophie Coon for her continuous programming support. This work was supported by grant PO1 GM48870 from the National Institutes of Health and Swiss National Science Foundation grant 823A-061225 to DS. This is manuscript 13938-MB from the Scripps Research Institute.

REFERENCES

1. Tomasselli A, Heinrikson RL. Targeting the HIV-protease in AIDS therapy: a current clinical perspective. Biochim Biophys Acta 2000;1477:189–214.
2. Preston BD, Poiesz BJ, Loeb LA. Fidelity of HIV-1 reverse transcriptase. Science 1988;242:1168–1171.
3. Roberts JD, Bebenek K, Kunkel TA. The accuracy of reverse transcriptase from HIV-1. Science 1988;242:171–173.
4. Coffin JM. HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. Science 1995;267: 483–489.
5. Rosin CD, Belew RK, Morris GM, Olson AJ, Goodsell DS. Coevolutionary analysis of resistance-evading peptidomimetic inhibitors of HIV-1 protease. Proc Natl Acad Sci USA 1999;96:1369–1374.
6. Rosin CD, Belew RK, Walker WL, Morris GM, Olson AJ, Goodsell DS. Coevolution and subsite decomposition for the design of resistance-evading HIV-1 protease inhibitors. J Mol Biol 1999;287: 77–92.
7. Rose JR, Babe LM, Craik CS. Defining the level of HIV-1 protease activity required for HIV-1 particle maturation and infectivity. J Virol 1995;69:2751–2758.
8. Tang J, Hartsuck JA. A kinetic model for comparing proteolytic processing activity and inhibitor resistance potential of mutant HIV-1 proteases. FEBS Lett 1995;367:112–116.
9. Laco GS, Schalk-Hihi C, Lubkowski J, Morris G, Zdanov A, Olson AJ, Elder JH, Wlodawer A, Gustchina A. Crystal structures of the inactive D30N mutant of feline immunodeficiency virus protease complexed with a substrate and an inhibitor. Biochemistry 1997; 36:10696–10708.
10. Sanner MF, Olson AJ, Spehner J-C. Reduced surface: an efficient way to compute molecular surfaces. Biopolymers 1996;38:305–320.
11. Koenderink JJ. Stocks-Green theorem: solid shape. Cambridge, MA: The MIT Press; 1990.
12. Sanner MF. Python: a programming language for software integration and development. J Mol Graphics Model 1999;17:57–61.
13. Tomasselli AG, Heinrikson RL. Specificity of retroviral proteases: an analysis of viral and nonviral protein substrates. Methods Enzymol 1994;241:279–301.
14. Lee T, Le V-D, Lin Y-C, Wong AL, Olson AJ, Elder JH, Wong C-H. Development of a new type of protease inhibitors, efficacious against FIV and HIV variants. J Am Chem Soc 1999;121:1145–1155.
15. Lee T, Laco GS, Torbett BE, Fox HS, Lerner DL, Elder JH, Wong C-H. Analysis of the S3 and S3′ subsite specificities of FIV protease; development of a broad-based protease inhibitor efficacious against FIV, SIV and HIV in vitro and ex vivo. Proc Natl Acad Sci USA 1998;95:939–944.
16. Roberts NA, Martin JA, Kinchington D, Broadhurst AV, Craig JC, Duncan IB, Galpin SA, Handa BK, Kay J, Krohn A, Lambert RW, Merrett JH, Mills JS, Parks KEB, Redshaw S, Taylor DL, Thomas GJ, Machin PJ. Rational design of peptide-based HIV proteinase inhibitors. Science 1990;248:358–361.
17. Jacobsen H, Yasargil K, Winslow DL, Craig JC, Krohn A, Duncan IB, Mous J. Characterization of human immunodeficiency virus type 1 mutants with decreased sensitivity to proteinase inhibitor Ro 31–8959. J Virol 1995;206:527–534.
18. Hertogs K, deBethune MP, Miller V, Ivens T, Schel P, Cauwenberge AV, Eynde CVD, Gerwen VV, Azijn H, Houtte MV, Peeters F, Staszewski S, Conant M, Bloor S, Kemp S, Larder B, Pauwels R. A rapid method for simultaneous detection of phenotypic resistance to inhibitors of protease and reverse transcriptase in recombinant human immunodeficiency virus type 1 isolates from patients treated with antiretroviral drugs. Antimicrob Agents Chemother 1998;42:269–276.
19. Vaillancourt M, Irlbeck D, Smith T, Coombs RW, Swanstrom R. The HIV type 1 protease inhibitor saquinavir can select for multiple mutations that confer increasing resistance. AIDS Res Hum Retrovir 1999;15:355–363.
20. Molla A, Korneyeva M, Gao Q, Vasavanonda S, Schipper PJ, Mo

HM, Markowitz M, Chernyavskiy T, Niu P, Lyons N, Hsu A, Granneman GR, Ho DD, Boucher CA, Leonard JM, Norbeck DW, Kempf DJ. Ordered accumulation of mutations in HIV protease confers resistance to ritonavir. Nat Med 1996;2:760–766.
21. Markowitz M, Mo H, Kempf DJ, Norbeck DW, Bhat TN, Erickson JW, Ho DD. Selection and analysis of human immunodeficiency virus type 1 variants with increased resistance to ABT-538, a novel protease inhibitor. J Virol 1995;69:701–706.
22. Gong YF, Robinson BS, Rose RE, Deminie C, Spicer TP, Stock D, Colonno RJ, Lin PF. In vitro resistance profile of the human immunodeficiency virus type 1 protease inhibitor BMS-232632. Antimicrob Agents Chemother 2000;44:2319–2326.
23. Tisdale M, Myers RE, Maschera B, Parry NR, Oliver NM, Blair ED. Cross-resistance analysis of human immunodeficiency virus type 1 variants individually selected for resistance to five different protease inhibitors. Antimicrob Agents Chemother 1995;39:1704–1710.
24. Condra JH, Schleif WA, Blahy OM, Gabryelski LJ, Graham DJ, Quintero JC, Rhodes A, Robbins HL, Roth E, Shivaprakash M, Titus D, Yang T, Teppler H, Squires KE, Deutsch PJ, Emini EA. In vivo emergence of HIV-1 variants resistant to multiple protease inhibitors. Nature 1995;374:569–571.
25. Winters MA, Schapiro JM, Lawrence J, Merigan TC. Human immunodeficiency virus type 1 protease genotypes and in vitro protease inhibitor susceptibilities of isolates from individuals who were switched to other protease inhibitors after long-term saquinavir treatment. J Virol 1998;72:5303–5306.
26. Hong L, Zhang XC, Hartsuck JA, Tang J. Crystal structure of an in vivo HIV-1 protease mutant in complex with saquinavir: insights into the mechanisms of drug resistance. Protein Sci 2000;9:1898–1904.
27. Buehler B, Lin YC, Morris GM, Wong CH, Richman DD, Elder JH, Torbett BE. Viral evolution in response to the broad-based retroviral protease inhibitor TL-3. J Virol 2001. Submitted for publication.
28. Cameron DW, Japour AJ, Xu Y, Hsu A, Mellors J, Farthing C, Cohen C, Poretz D, Markowitz M, Follansbee S, Angel JB, McMahon D, Ho D, Devanarayan V, Rode R, Salgo MP, Kempf DJ, Granneman R, Leonard JM, Sun E. Ritonavir and saquinavir combination therapy for the treatment of HIV infection. AIDS 1999;13:213–224.