Lattice engineering enables definition of molecular features allowing for potent small-molecule inhibition of HIV-1 entry

Yen-Ting Lai1, Tao Wang2, Sijy O’Dell1, Mark K. Louder1, Arne Schön3, Crystal S. F. Cheung1, Gwo-Yu Chuang1, Aliaksandr Druz1, Bob Lin1, Krisha McKee1, Dongjun Peng1, Yongping Yang1, Baoshan Zhang1, Alon Herschhorn4,5, Joseph Sodroski4, Robert T. Bailer1, Nicole A. Doria-Rose1, John R. Mascola1, David R. Langley6 & Peter D. Kwong1

Diverse entry inhibitors targeting the gp120 subunit of the HIV-1 envelope (Env) trimer have been developed including BMS-626529, also called temsavir, a prodrug version of which is currently in phase III clinical trials. Here we report the characterization of a panel of small-molecule inhibitors including BMS-818251, which we show to be >10-fold more potent than temsavir on a cross-clade panel of 208-HIV-1 strains, as well as the engineering of a crystal lattice to enable structure determination of the interaction between these inhibitors and the HIV-1 Env trimer at higher resolution. By altering crystallization lattice chaperones, we identify a lattice with both improved diffraction and robust co-crystallization of HIV-1 Env trimers from different clades complexed to entry inhibitors with a range of binding affinities. The improved diffraction reveals BMS-818251 to utilize functional groups that interact with gp120 residues from the conserved β20-β21 hairpin to improve potency.
The entry of HIV-1 into target cells is a critical event in the viral life cycle and a target for drug development. Viral entry is mediated by the HIV-1 envelope (Env) glycoprotein trimmer, a type 1 fusion machine comprised of three gp120 subunits and three noncovalently linked gp41 subunits, which binds to cell-surface receptors and fuses viral and host cell membranes. Entry inhibitors targeting the gp120 subunit have been developed, with a promising small-molecule lead, fostemsvir, the prodrug version of BMS-626529 (temsavir) currently in phase III clinical trials. Notwithstanding its advanced development and novel mode of action, next-generation inhibitors of fostemsvir have been sought to improve ADME (absorption, metabolism, distribution and elimination) profile, to overcome expected drug resistance, and to increase potency. We note that these qualities may be related; for example, increasing the potency of an inhibitor can be an effective way to counter drug resistance, as resistance mutations generally have only incremental effects on the dose—response of a drug.

X-ray crystallography is often instrumental in determining drug-binding mode and in facilitating structure-based drug design. However, structure-based drug design can only be reliably carried out with crystals that diffract with resolution sufficient to provide accurate structural models; unfortunately, this resolution prerequisite has been difficult to achieve for many drug targets, even with extensive screening of crystallization conditions and protein variants. Crystal engineering represents an alternative strategy for crystal improvement, whereby inspection of a lattice with poor diffraction identifies weak lattice contacts, which can then be altered through structure-based design. However, both of these strategies can inadvertently introduce modifications that change the properties of protein targets and even their structures.

Crystallization chaperones, such as antibody fragments, have also been used to facilitate formation of crystal lattice contacts for difficult protein targets. We recently reported the structure of BMS-626529 (temsavir) in complex with an HIV-1 Env trimer bound by crystallization chaperones comprising the antigen-binding fragments (Fabs) of antibodies 35022 and PGT122. We also reported the structure of BMS-378806 (ref. 18), the prototype small molecule for this class of compounds, in the same Env-35O22-PGT122 lattice. In both cases, the resolution was only 3.8 Å, and there was uncertainty in the positioning of small-molecule atoms and in the definition of side-chain interactions.

To obtain structural information of improved accuracy, we test a strategy involving the lattice-based engineering of crystallization chaperones. This strategy provides a way to improve a lattice without altering the protein target. We engineer crystallization chaperones to identify a crystal lattice suitable for determining high-resolution structures of inhibitors, spanning a range of >6000-fold neutralization potency, in complex with envelope trimers of clade A and B HIV-1 strains. We use this lattice to examine small-molecule inhibitors related to BMS-626529 and report structures of multiple small-molecule inhibitors, including that of BMS-818251, an HIV-1 entry inhibitors with >10-fold higher potency than BMS-626529, which reveal structural determinants of potent HIV-1 inhibition and provide insights into the design of better entry inhibitors for this class of HIV-1 drugs.

Results
BMS-818251 shows >10-fold increased potency over temsavir. By screening a library of temsavir derivatives, we identified two compounds, BMS-814508 and BMS-818251, which showed improved entry inhibition of the laboratory-adapted HIV-1 strain NL4-3. The EC50 for BMS-814508 and BMS-818251 was 0.495 ± 0.069 and 0.019 ± 0.003 nM, respectively, 4-fold and ~100-fold more potent than BMS-626529, which had an EC50 of 2.2 ± 0.6 nM against the same strain. Both of the improved compounds used a cyano alkenel to replace an amide group with different thiazole substituents replacing the triazole on the 6-azaindole core of BMS-626529 (temsavir) (Fig. 1a, Supplementary Fig. 1).

To verify neutralization potency seen in the neutralization-sensitive NL4-3 strain, we further tested the ability of these compounds to inhibit entry on a cross-clade panel of 30 HIV-1 strains derived from primary isolates (Fig. 1b). The 30 strains from all major HIV-1 clades were chosen to encompass the complete range of neutralization sensitivity towards BMS-626529 in a 208-strain panel (ref. 18). To provide comparative data, we assessed the well-studied entry inhibitors BMS-378806 and BMS-626529 as controls, and also included BMS-386150, which is very similar to BMS-626529, with a bromine modification in the compound to aid crystal structure determination (Fig. 1a). A recently reported compound 484, which exhibits broad reactivity despite its moderate micro-molar range potency, was also tested. Of note, these six compounds all derive from a piperazine-piperidine bridge, with various degrees of modification, representing a collection of diverse entry inhibitors targeting the gp120 subunit of the HIV-1 Env trimer (Fig. 1a and Supplementary Fig. 1).

Neutralization of the 30-strain panel generally recapitulated prior observations (Fig. 1b, c). BMS-626529 neutralized this panel of viruses with a geometric mean IC50 of ~0.040 μM, which was within a factor of three of the geometric mean IC50 of 0.015 μM that has been previously reported based on its inhibition of the 208-strain panel. Consistent with the literature, BMS-626529 (temsavir) was substantially more potent than BMS-378806, and compound 484 neutralized 11 of the 30 viruses with micro-molar potency. Despite an overall similarity with BMS-378806, BMS-386150 had a substantially weaker (~7-fold lower) neutralization potency, and BMS-814508 neutralized 28 of the 30 viruses with a mean inhibition similar to that of BMS-626529 (temsavir). Cytotoxicity was observed for BMS-814508 at the highest concentration (20 μM) in the neutralization assay (Supplementary Fig. 2). Strikingly, compound BMS-818251 (geometric mean IC50 ~ 0.002 μM) showed roughly 20-fold higher neutralization potency than BMS-626529 (geometric mean IC50 ~ 0.04 μM) on this 30-strain panel, with more than half of the viruses neutralized in the subnanomolar range.

The superior potency of BMS-818251 observed on the 30-strain panel prompted us to assess the neutralization potency of this compound on the larger 208-strain panel (Supplementary Table 1). The geometric mean IC50 of BMS-818251 on the 208-strain panel was 0.0015 μM, in close agreement with the geometric mean IC50 obtained from the 30-strain panel. When compared to the neutralization potency on the same 208-strain panel, BMS-818251 has a 10.6-fold higher geometric mean IC50 and 18-fold higher median IC50 compared to BMS-626529 (temsavir) (median IC50 was 0.0005 and 0.0090 μM for BMS-818251 and BMS-626529, respectively; geometric mean IC50s were 0.0015 and 0.0159 μM, respectively).

To assess whether the improved potency of BMS-818251 over temsavir was reflected by higher affinity to the HIV-1 Env trimer, we determined the binding affinity of BMS-818251 to HIV-1 Env trimer by isothermal calorimetry (ITC). ITC showed BMS-818251 to have an affinity of 0.047 μM to BG505 DS-SOSIP.664 Env trimer at 37 °C, which was substantially higher than the 0.402 μM affinity of BMS-626529 under the same condition (Supplementary Fig. 3). Thus, the >10-fold improved potency of BMS-818251 over temsavir in neutralization assays was associated with an 8.6-fold higher affinity, as determined by the ITC binding assay.
**Fig. 1** Diverse HIV-1 entry inhibitors span >6000-fold differences in neutralization potency, with BMS-818251 being >20-fold more potent than BMS-626529 (temasvir). a HIV-1 entry inhibitors with common functional groups shown in black and unique features in red. b Neutralization assay of entry inhibitors against thirty HIV-1 isolates from all major HIV-1 clades. c Neutralization data shown as a scatter plot, with the geometric mean shown as horizontal bars. Dotted lines show the detection limits of neutralization assay.
Lattice engineering enables higher-resolution structures. To provide a structure–function relationship between these small-molecule leads and their HIV-1 target, we sought to crystallize these compounds with the HIV-1 Env trimer. Previously we determined cocrystal structures of BMS-626529 and BMS-378806 at a resolution of 3.8 Å, bound to the BG505 SOSIP.664 trimer along with two crystallization chaperones, Fab 35O22 and PGT122 (ref.18). The BG505 SOSIP.664 Env trimer is a near-native Env trimer comprising fully cleaved gp120 and gp41 subunits, derived from strain BG505, with stabilized disulfide (SOS) and I559P (IP) mutation and truncated at residue 664 (ref. 21). These cocrystal structures defined the general pose of drug compounds in the gp120 binding pocket, but left uncertain the atomic position of chemical substituents. To improve the resolution of the diffraction from this lattice (Fig. 2a, b), which allowed determination of several Env structures in independent studies (Supplementary Table 2), we attempted to reduce lattice flexibility in the crystallization chaperones (Fab 35O22 and Fab PGT122). The constant region of the 35O22 Fab appeared hyper-flexible as judged by its high B-factor (Supplementary Fig. 4a). Removal of the constant region by converting the Fab 35O22 to an scFv led to crystals that retained the original lattice, with almost identical lattice parameters (Supplementary Table 3). In structures of the 35O22 scFv, we observed some regions to be more flexible than in the Fab context, mainly due to a lack of structural support of antibody constant regions. To mitigate this flexibility, we tested two replacement variants: (i) replacing residues 10ELK12 with β-strand-prefering residues 10TTT12; (ii) altering A16S and F84S to reduce exposed hydrophobic atoms; named 35O22_3T2S as in Supplementary Fig. 4b. Crystals generated with 35O22_3T2S has a substantially lower Wilson B factor of 28 Å² compared to 86 Å² for the original crystal (Fig. 2c). The refined structure of scFv 35O22_3T2S showed an average B-factor of 71 Å² vs. 202 Å² for the 35O22 Fab in the previous BMS-626529 complexed structure. In addition to mutations introduced into 35O22, two methionine mutations (MM) were introduced onto the Fab of 3H109L, a precursor of PGT122 found to crystallize better.22 These methionine mutations potentially interact with the base of gp41 molecules, one layer higher in the lattice (Fig. 2b, d). Finally, two potential glycosylation sites were introduced onto the 35O22 scFv, which enhanced its production yield by tenfold (Supplementary Fig. 4b).

To assess the impact of our lattice engineering, we determined the structure of BMS-378806 in the improved lattice with the modified antibody fragments (termed 35O22_3T2S and 3H109L_MM). We observed almost identical unit cell dimensions as the original lattice18, but with superior diffraction (Fig. 2c and Supplementary Fig. 4). The resulting electron density maps showed superior density for many regions of the Env trimer. For example, residues 65–73 in the constant region 1 (C1) of gp120 could now be accurately modeled with histidine residues at positions 66 and 72 forming an unusual stacking interaction, which was modeled differently in prior structures (Supplementary Fig. 5). Another region of improved structure modeling occurred at the α9 helix of gp41, where weak electron density previously made placement of amino acid side chains ambiguous,22,24,25, the clear density in the new structure allowed unambiguous placement of amino acid side chains and correct registry of backbone atoms (Supplementary Fig. 6). It is important to note that the C1 region and the α9 helix did not directly contact the small-molecule inhibitor; thus the improved resolution of these regions was a consequence of the improved lattice, not an effect of small-molecule stabilization. We also noted that for very dynamic regions of the Env (such as the inter-helical region of gp41), lattice engineering was unable to overcome the intrinsic flexibility of these regions, and they remained unresolved in our crystal structures. While the real space correlation for BMS-378806 only improved from 0.91 to 0.95, the improved electron density did allow for greater structural certainty in placement of small-molecule atoms such as the methyl on the piperase ring or the rotameric state of the oxocarboxylic moiety. Overall, the engineered lattice allowed for substantially greater certainty in the positioning of both small-molecule compound and interacting Env-trimer side chains.

Structural determination of compound 484-Env interactions. As a test of the engineered lattice, we chose to examine compound 484, which previously resisted cocrystalization, presumably due to its micro-molar range activity.20 A binding pose of compound 484 had previously been proposed based on computational docking.20 We used the engineered lattice to determine the cocrystal structure of 484 with the BG505 SOSIP.664 Env trimer and compared it to the docking simulation. The cocrystal structure (Fig. 3a) was similar to the overall pose of 484 in the docking study, though with notable differences. The thiophene ring essentially flipped 180° between cocrystal structure and docking pose, a difference supported both by the clear electron density defining 484 in the new lattice and by the strong scattering from sulfur, which served to identify this atom in the thiophene ring. Additionally, the piperazinone ring rotated ~45° leading the acetyl-cyclopropane and chloro-benzene to shift 1.5 Å toward the thiophene ring (Fig. 3a).

Overall, 484 interacted mainly with three gp120 fragments: the C-terminus of α1 helix (residues 107–117; colored cyan in Fig. 3b); the C-terminal part of the CD4-binding loop and subsequent strands (residues 369–385; colored magenta in Fig. 3b); and the β20–β21 hairpin (residues 423–436; colored green in Fig. 3b). Hydrophobic interactions dominated the interface between 484 and the BG505 SOSIP.664 Env trimer (Fig. 3c), with the piperazinone ring stacked between two tryptophan residues, W427 of the β20–β21 loop and W112 of the α1-helix. Two residues outside of the three major fragments, residues V255 and M475, also provided hydrophobic interaction with 484. The cocrystal structure of compound 484 thus demonstrated the ability of the engineered lattice to provide atomic-level detail on the interaction between a low potency, low affinity compound, and the HIV-1 Env trimer.

Tail region of BMS-818251 is critical for potent inhibition. We next investigated whether crystal structures derived from the engineered lattice could provide a structure–function explanation for the very high potency of BMS-818251 against HIV-1 BG505 compared to the closely related BMS-814508 (IC₅₀ values < 0.0003 vs. 0.008 μM for BMS-818251 and BMS-814508, respectively, against BG505; Fig. 1b). These analogs differ only in the chemistry of their respective tail functional groups, the moiety extending from the thiazole ring (Fig. 1a). Examination of the cocrystal structures of BMS-818251 and BMS-814508 with the BG505 SOSIP.664 Env trimer revealed the tail of BMS-818251 to make hydrophilic interactions with BG505 trimer (Fig. 4a), while the tail of BMS-814508 did not make defined contacts with BG505 trimer (Fig. 4b). Specifically, the tail of BMS-818251 interacts with the side chains of K117, R429 and Q432 of
BG505 trimer through the terminal hydroxyl group, with clear electron density defining the conformation of the BMS-818251 tail functional group (Fig. 4a; right panel). In addition, residue D113 of Env forms hydrogen bond with the amide nitrogen on the tail of BMS-818251. The lack of electron density surrounding the tail functional group of the highly similar BMS-814508 (Fig. 4b; right panel) suggested that this functional group can adopt multiple conformations and lacks stabilizing interactions with Env.

The difference in the interaction between Env trimer and the tails of BMS-818251 or BMS-814508 likely explains the difference in their neutralization potency. Interestingly, although the chemistry of BMS-818251 and BMS-814508 was identical beyond the tail functional group, we did observe additional interactions between Env residues F382 and E370 with BMS-818251, but not with BMS-814508 (Fig. 4a, b, middle panels). Finally, we observed the piperidine ring of BMS-818251 to adopt a twisted-boat conformation (Fig. 4a, b, left panels), with slightly higher energy than the chair conformation adopted by BMS-814508, suggesting a difference in tail interactions to affect the conformation of the entire compound.

BMS-818251 accommodates residue variations with high potency. BMS-818251 showed superior neutralization potency against the 30-strain panel compared with BMS-814508, ranging...
from ~2.7- to 90-fold higher (Supplementary Fig. 7a; excluding three strains, DJ263.8, T253-11, and CAP256.206.C9, that were below the detection limit of assay for both BMS-818251 and BMS-814508). Although the different tails of BMS-818251 and BMS-814508 induced subtle differences throughout their Env-bound conformation, we nonetheless focused on Env residues 117, 429, and 432 to provide an explanation for the wide range of improved potency of BMS-818251, because of the direct interaction between these residues and the BMS-818251 tail in the BG505 trimer cocrystal structure (Fig. 4a).

Residue 117 was Lys in all 30 tested HIV-1 strains (except for strain CNE56, where it is a Gln), with multiple residue types at residues 429 and 432; we therefore categorized the fold improvement based on the residue identity in residues 429 and 432 (Supplementary Fig. 7b). With the most prevalent combination of R429 and Q432 (9 out of 30 strains), the improvement in neutralization potency for BMS-818251 ranged from ~2.7 to ~50 fold. With G429 and a positively charged residue 432 (K or R; 5 out of 30 strains), the improvement ranged from ~3.3 to ~53.3-fold, similar to the R429 and Q432 combination. With negatively charged E429 and a positively charged residue 432 (K or R; 8 out of 30 strains), the fold improvements were similar compared to the combination of R429 and Q432 (ranging from ~3.3 to 90-fold). Finally, when the residues at 429 and 432 were both positively charged (R429/K432 for REJO.67 and K429/K432 for BL01.DK), the improvement were generally minimal (6.1-fold and 5.3-fold, respectively). Thus a wide range of neutralization sensitivity was observed even with identical or homologous residues at positions 429 and 432. However, we noted that the neutralization potency of BMS-818251 improved for most of the viruses tested. Overall, the virus neutralization data supported the general benefit of the tail functional group in BMS-818251 compared to its counterpart in BMS-814508, and revealed that improved potency could generally be achieved despite variations at positions 429 and 432, which directly interacted with the BMS-818251 tail functional groups.

**Structural basis of clade-dependent neutralization.** We sought to examine whether the new lattice could provide experimental feedback on the recognition of Env trimers from different HIV-1 clades. Clade-dependent sensitivity has been observed for the entry inhibitors in the BMS-626529 family. Among the three
major clades that had been well studied with this class of inhibitors, clade B strains were the most sensitive, followed by clade C strains with moderate sensitivity and clade A strains with least sensitivity. To investigate the structural attributes responsible for this clade dependency, we first determined the compound-free Env-trimer structure of the B41 strain, a clade B virus with only 78% sequence identity to the clade A BG505. Residues that differed in sequence between B41 and BG505 were observed to account for 28% of the Env-trimer surface (Fig. 5a). Surface differences were also present in the epitopes of the crystallization chaperones 35O22_3T2S and 3H109L_MM, but these modified antibodies were able to accommodate these changes, and crystals of the B41 SOSIP.664 Env trimer had unit cell dimensions similar to those of the BG505 SOSIP.664 Env trimer crystals (Fig. 5b).

We next sought to determine cocrystal structures of EnvS from HIV-1 strains B41 and BG505 with BMS-386150, a compound similar to the prototypic BMS-378806, though with a bromine modification to aid structure determination (Fig. 1a). Although there were differences in the three major binding fragments (Fig. 5c, colored in cyan, magenta and green respectively, as defined in Fig. 3b) interacting with BMS-386150, the interacting residues were mostly conserved between B41 and BG505 EnvS. Furthermore, the BMS-386150-interacting residues specific for B41 or BG505 (highlighted in black rectangles in Fig. 5c) were conserved between these two strains (Supplementary Fig. 8). This observation suggested that the clade dependency of entry inhibitors in the temsavir family was not directly caused by sequence variation in the binding pocket, but might rather be a consequence of various degrees of conformational pliability at the inhibitor binding site among the different clades. Indeed, a modest but significant shift of the β20−β21 hairpin loop was observed in the BMS-386150-complexed structures of B41 and BG505 Envs (Fig. 5d, upper panel), supporting the idea that conformational changes around the binding pocket, instead of the sequence identity of residues of the binding pocket, dictate the clade-dependent neutralization potency of this class of entry inhibitors. This conclusion was further strengthened by the fact that B41 and BG505 showed conformational differences around the drug-binding pocket in their compound-free structures (Fig. 5d, lower panel).

**Discussion**

Crystal engineering has been extensively explored for small molecules crystals, such as metal-organic frameworks,
which has led to crystals with extraordinary properties, including those with lattice cages, capable of holding diverse small molecules\textsuperscript{30,31}. Crystals of biological macromolecules may be more difficult to manipulate due to their greater complexity; however, recent advances have allowed crystal engineering of biological crystals of DNA, protein and peptides\textsuperscript{32–34}. We report here the engineering of a protein crystal composed of a highly glycosylated trimer and two crystallization chaperones, which were adapted antibody fragments. We focused on modifying the crystallization chaperones, to remove chaperone flexibility, while avoiding alterations to the target protein. Of note, the antibody fragments were resilient to alteration, and for one of the crystallization chaperones (35O22 Fab), we removed half of its weight, while increasing its capacity to facilitate crystallization. The resulting engineered lattice led to higher resolution and yielded twice as many unique measured reflections as the original crystal, while maintaining similar unit cell parameters; this allowed for improved electron density and more accurate

Fig. 5 Generalizable crystallization platform for diverse HIV-1 envelope and drug complexes. a B41 SOSIP.664 Env trimeric structure in surface representation shows gp120 in black and gp41 in gray; the residues that differ between the B41 SOSIP.664 and BG505 SOSIP.664 Env trimers are colored as red. b B41 SOSIP.664 Env trimeric structure is shown in the context of P6\textsubscript{4} lattice. c Structures of BMS-386150 in complex with the BG505 (top) or B41 (bottom) SOSIP.664 Env trimers. Residues interacting with BMS-386150 that are different between B41 and BG505 are highlighted in black box. d The backbone conformational differences in the structures of BMS-386150 complexed with BG505 (purple) or B41 (yellow) are shown as ribbon diagram (top), the same region in the apo-form BG505 and B41 are also shown (bottom)
modeling of small-molecule inhibitors and Env trimers. Overall, the lattice engineering we report is conceptually similar to a retro-engineering strategy reported recently for an organic molecular crystal\textsuperscript{[35]}, but has not to our knowledge been used for macro-molecular lattice improvement.

Although antibody fragments and other crystallization chaperones are popular tools for crystallizing difficult protein targets\textsuperscript{[36,37]}, the retrospective engineering of crystallization chaperones has not been extensively explored. More and more protein targets are being crystallized in the presence of crystallization chaperones\textsuperscript{[38,39]}, as many of these crystals diffract poorly due to suboptimal chaperone properties\textsuperscript{[40,41]}, the engineering strategy reported here may represent a straightforward option for improving their diffraction.

The antiviral potency of the fostemsavir inhibitor is known to display HIV-1 clade dependence on neutralization\textsuperscript{[18]}, but the molecular basis for this dependence has not been identified. Interestingly, we observed conformational differences in the interaction of BMS-386150 with Env trimers from clade A BG505 and clade B B41 HIV-1 strains, despite conservation of all Env residues that contacted BMS-386150. Differences between these interactions related to the positioning and orientation of contact residues, which differed despite being identical in sequence. Specifically, we observed the BG505 and B41 Env trimers to adopt slightly different conformations in their respective binding pockets around the β20–β21 hairpin loop region. This difference in binding pocket conformation led to different conserved residues interacting with the inhibitor, manifested as clade dependency in neutralization assays. Our results thus provide insight into the complexity of clade dependence and reveal the importance of obtaining experimental information on small molecules in complex with target Env trimers from diverse clades. The crystal structures of BMS-386150 also provided a basis for its reduced potency, as the bromine substitution, despite being beneficial for structure determination, abrogated the hydrophobic interactions between I424, M426 and M434 and the methoxy functional group that was replaced.

Despite the general importance of clade dependency, we note that the tail functional group that endowed high potency to BMS-818251 could accommodate variations on interacting residues at positions 429 and 432 of the HIV-1 Env. The effects of variation in these residues on the enhanced potency of BMS-818251 did not show clade dependency (Supplementary Fig. 7b). The flexibility of the tail functional group of BMS-818251, which can adopt different conformations through torsional angle freedom, may allow it to accommodate the different conformations of the β20–β21 structural element in different clades. Furthermore, we note that the BMS-818251 tail contains two functional groups that can be either hydrogen-bond donors or acceptors, potentially increasing its versatility to interact with HIV-1 Env, while the BMS-814508 tail contains two nitrogens that can only be hydrogen-bond acceptors. Thus, multiple ways may exist for small molecules to succumb to or to overcome clade dependency.

ITC binding data showed that BMS-818251 has a ~10-fold higher affinity to BG505 DS-SOSIP.664 trimmer compared to BMS-626529 binding, which is in good agreement with the overall >10-fold improved neutralization potency of BMS-818251. However, we note that the neutralization potency is higher than the binding affinity measured by ITC (for example, the BG505 virus was neutralized at 0.0003 μM by BMS-818251, while the ITC-determined binding affinity between BMS-818251 and BG505 DS-SOSIP.664 Env trimmer was 0.047 μM). This discrepancy is likely because BG505 DS-SOSIP.664 Env trimmer is not a perfect mimic of the native BG505 Env trimmer on the virion surface. Nonetheless, the fold improvement in affinity correlated well with the fold improvement in neutralization. Thus, BG505 DS-SOSIP.664 Env trimmer, despite being an imperfect mimic of native Env, can be a valuable surrogate for the evaluation of improved next-generation entry inhibitors.

It is noteworthy that BMS-818251 exhibited higher antiviral potency than BMS-814508, a related compound that differs only in the tail functional group, for all 30 viruses (excluding 3 viruses that were below the detection limit of assay) tested in this study (Supplementary Fig. 7). AE strains generally are less sensitive to temsavir but were sensitive to BMS-818251, suggesting that BMS-818251 could potentially be pan-reactive to HIV-1 Env. Indeed, we observed a general improvement for BMS-818251 across all major HIV-1 clades, likely enabled by its novel tail functional group. Further optimization of this tail functional group may lead to inhibitors with greater resistance to viral escape. We note that the prodrug version of temsavir is one of the few gp120-directed therapeutics being assessed clinically, with gp120-reactive antibodies like antibody VRC01 in clinical assessment showing lower breadth and potency than temsavir (Supplementary Fig. 9). Overall, the structure of the BMS-818251-Env complex enabled by BMS818251-based crystal engineering reveals molecular features explaining the superior neutralization potency of this compound and assisting in the rational design of next-generation inhibitors targeting the gp120 component of the HIV-1 Env.

Methods Synthesis of small-molecule entry inhibitors. The BMS entry inhibitors were synthesized based on Supplementary Figure 10. The oxalamides were prepared via acylation of indole with oxalyl chloride followed by coupling of the acid chloride with 1-benzoylpiperazine (ref. 42) or azaindole with methyl chlorooxoacetate in the presence of AICl3, followed by hydrolysis of the methyl ester and coupling with 1-benzoylpiperazine (ref. 45) or 2-pyrazinyl-3-(piperidin-4-ylidene)acetamide. The substitutions on C-7 of 6-azaindole were installed initially via either copper(0)-mediated radical reaction or palladium(0)-mediated Stille coupling, followed by structural modifications. Inhibitors were characterized by 1H-NMR, 13C-NMR and high-resolution mass spectrometry after synthesis; the representative spectra of BMS-818251, the most potent compound reported in this manuscript are shown in Supplementary Fig. 11. Compound 484 were prepared following published protocol\textsuperscript{[46]} and described here. N,N-disopropyl ethylamine (0.6 mmol), 4-dimethylaminopyridine (0.03 mmol) and N-(3-dimethylaminopropyl)-N-ethyl-carbamidene hydrochloride (0.45 mmol) were added to a solution of carboxylic acid (0.3 mmol) and amine (0.3 mmol) in anhydrous dichloromethane (2 ml). The solution was stirred at room temperature for 5 h and concentrated in vacuo. The resulting residue was purified by flash chromatography and the purity of Compound 484 was greater than 95% according to NMR analysis.

HIV-1 neutralization assays. Neutralization was measured using single-round-of-infection HIV-1 Env-pseudoviruses and TZM-bl target cells (ARP Cat No. #8129)\textsuperscript{[44,45]}. Pseudoviruses were prepared at 3000 M with 4% DMSO (dimethyl sulfoxide) in PBS (phosphate-buffered saline) and diluted 20 times as the starting concentration for serial dilution in the neutralization assays. Pseudoviruses of ten strains from HIV-1 clades A, B, and C were incubated with serial dilutions of drug compounds, then added to TZM-bl target cells which have a luciferase reporter gene. After incubation at 37 °C for 2 days, infection of target cells was measured by luciferase activity. Neutralization curves were fit by nonlinear regression using a five-parameter hill slope equation\textsuperscript{[47]}. Proteins expression and purification. BG505 SOSIP.664 is a construct with modifications that enabled the production of fully cleaved, prefusion-closed, soluble HIV-1 Env trimmer\textsuperscript{[48]}. Proteins were purified following published protocols\textsuperscript{[49]} with modifications described below. For 1 L of Gnti\textsuperscript{+} cell culture (ATCC Cat No. CRL-3022), 600ug of BG505 SOSIP.664 (with N332 glycan) expression plasmid and helper expression plasmid for Furin were diluted in 50 ml Opti-MEM medium (Life Technology). Three milliliters of Turbo293 transfection reagent (Speed Bio-Systems) was diluted into 50 ml Opti-MEM medium and incubated for 5 min. Diluted transfection reagent was then added into diluted DNA and incubated for 15 min. The transfection reaction and DNA complex were added into 800 ml of Gnti\textsuperscript{−}293 cell culture. On the second day of transfection, added 80 ml HyClone SFM4HEK293 medium to each flask of 11 cell cultures. Return flasks to shaker incubator at 120 rpm, 37 °C, 9% CO\textsubscript{2} for additional 6 days. Supernatant was affinity-purified over 2G12 column, eluted by 3 M MgCl\textsubscript{2}, followed by size exclusion chromatography in 5 mM Hepes and 150 mM NaCl. Variants of 35O22 and 3H109L antibody fragments were produced in Gnti\textsuperscript{−} and Expi cells respectively following published protocol\textsuperscript{[50]}. 3H109L variants were produced as Fab, with heavy chain truncated at the end of constant domain I and a purification tag containing...
thrombin site, Strept-tag and 6xHis-tag was appended. After tandem purification by Strept column and Ni-NTA column, the tag was removed by thrombin digestion at 4°C overnight, followed by size exclusion chromatography. 35O22 scFv has a single 6xHis-tag for purification and the tag is retained after purification. Two artificial glycans introduced to 35O22 were preserved by not treating with endoglycosidase.

**Crytalization and data collection.** Purified trimers were mixed with 3H109L Fab variant (with two methionine modifications) and 35O22 scFv variant (with 3T, 2S and 2 glycan modifications) in a 1:3:2:3 molar ratio (gp140 protomer: Fab-scFv) and incubated overnight at room temperature before being purified by SEC and concentrated to 10 M. The overall resolution was determined as the highest resolution for which the completeness was greater than 50% and the I/λ2 was >2.0. Data were then truncated by UCLA anisotropy server (https://services.mbi.ucla.edu/anisocalse/).

**Structure solution and refinement.** The structures of entry inhibitors in complex with BG505 or B41 SOSIP.664 in the presence of crystallization chaperone variants were solved by molecular replacement using Phaser in CCP4 using the nondrug-scFv:Fab:scFv) and incubated overnight at room temperature before being purified by SEC and concentrated to 10 M for crystallization. Glycans on the Env and antibody fragments were preserved without the treatment of glycosidase. Crystalization conditions were obtained by sparse chemical screening and optimized manually after initial hits were identified. Diffraction data were collected at APS ID22 and processed using HKL2000. The overall resolution was 6. Allen, R. C., Popat, R., Diggle, S. P. & Brown, S. P. Targeting virulence: can we make evolution-proof drugs? Nat. Rev. Microbiol. 12, 300–308 (2014).

**References**

1. Engelman, A. & Chereshpanov, P. The structural biology of HIV-1: mechanistic and therapeutic insights. Nat. Rev. Microbiol. 10, 279–290 (2012).

2. Brandenberg, O. F., Magnus, C., Ragoese, R. R. & Trkola, A. The HIV-1 entry process: a stoichiometric view. Trends Microbiol. 23, 763–774 (2015).

3. Lu, L., Yu, F., Cai, L., Dedhia, A. K. & Jiang, S. Development of small-molecule HIV entry inhibitors: a stoichiometric view. Curr. Top. Med. Chem. 16, 1074–1090 (2016).

4. Meanwell, N. A. et al. Inhibitors of HIV-1 attachment: the discovery and development of Tmsavir and its Prodrug Fostemsavir. Nat. Rev. Mol. Cell Biol. 19, 1115–1122 (2018).

5. Patel, S. B. et al. Lattice stabilization and enhanced diffraction in human p38 MAPK by Pymol. Nat. Chem. 4, 1389–1397 (2012).

6. Patel, S. B., Lima, R. & Dobbek, H. Crystallography and drug discovery: reorienting the interface of leucines. Protein Sci. 16, 803 e793 (2017).

7. Koide, S. Engineering of recombinant crystallization chaperones. Curr. Opin. Struct. Biol. 19, 449–457 (2009).

8. Hughes, D. & Andersson, D. I. Evolutionary consequences of drug resistance: shared principles across diverse targets and organisms. Nat. Rev. Genet. 16, 459–471 (2015).

9. Murray, C. W. & Rees, D. C. The rise of fragment-based drug discovery. Nat. Chem. 1, 187–192 (2009).

10. Thomas, S. E. et al. Structural biology and the design of new therapeutics: from HIV and cancer to mycobacterial infections: a paper dedicated to John Kendrew. J. Mol. Biol. 429, 2677–2693 (2017).

11. Blundell, T. L. Protein crystallography and drug discovery: recollections of knowledge exchange between academia and industry. EJCJ 4, 308–321 (2017).

12. Zheng, H., Hou, J., Zimmerman, M. D., Wlodawer, A. & Minor, W. The future of x-ray crystallography in drug discovery. Expert Opin. Drug Discov. 9, 125–137 (2014).

13. Bauman, J. D. et al. Crystal engineering of HIV-1 reverse transcriptase for structure-based drug design. Nucleic Acids Res. 36, 5083–5092 (2008).

14. Yamada, H. et al. Crystal lattice engineering, an approach to engineer protein crystal contacts by creating intermolecular symmetry: crystallization and structure determination of a mutant human RNase 1 with a hydrophobic interface of leucines. Protein Sci. 16, 1389–1397 (2007).

15. Patel, S. B. et al. Lattice stabilization and enhanced diffraction in human p38 alpha crystals by protein engineering. Biochim. Biophys. Acta 1696, 67–75 (2004).

16. Hvosmcu, S. et al. Effects of protein engineering and rational mutagenesis on crystal lattice of single chain antibody fragments. Proteins 82, 1884–1895 (2014).

17. Koide, S. Engineering of recombinant crystallization chaperones. Curr. Opin. Struct. Biol. 19, 449–457 (2009).

18. Pancera, M. et al. Crystal structures of trimeric HIV envelope with entry inhibitors BMS-378806 and BMS-626529. Nat. Chem. 11, 1115–1122 (2017).

19. Nowicka-Sans, B. et al. In vitro antiviral characteristics of HIV-1 entry inhibitor BMS-626529, the active component of the prodruk BMS-663068. Antimicrob. Agents Chemother. 56, 3496–3507 (2012).

20. Herschhorn, A. et al. The beta20-beta21 of gp120 is a regulatory switch for HIV-1 Env conformational transitions. Nat. Commun. 8, 1049 (2017).

21. Sanders, R. W. et al. A next-generation cleaved, soluble HIV-1 Env trimmer, BGS05 SOSIP.664gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. PLoS Pathog. 9, e1003618 (2013).

22. Garces, F. et al. Affinity maturation of a potent family of HIV antibodies is primarily focused on accommodating or avoiding glycans. Immunity 43, 1053–1063 (2015).

23. Strong, M. et al. Toward the structural genomics of complexes: crystal structures of a PE/EPE protein complex from Mycobacterium tuberculosis. Proc. Natl Acad. Sci. USA 103, 8060–8065 (2006).

24. Pancera, M. et al. Structure and immune recognition of trimeric pre-fusion HIV-1 Env. Nature 514, 455–461 (2014).

25. Guenaga, J. et al. Glycine substitution at helix-to-coil transitions facilitates the structural determination of a stabilized subtype C HIV envelope glycoprotein. Immunity 46, 792–803 e793 (2017).

26. Mastalerz, M. Crystalline engineering: covalent crystalline growth. Nat. Chem. 5, 810–811 (2013).

27. Hopp, J. T. Crystal engineering: towards artificial enzymes. Nat. Chem. 2, 432–433 (2010).

28. Thomas, J. M. Diffusionless reactions and crystal engineering. Nature 495, 461–466 (2013).

29. Zaworotko, M. J. Crystal engineering comes of age. Nat. Chem. 3, 653 (2011).

30. Lee, S., Kapustin, E. A. & Yaghi, O. M. Coordinative alignment of molecules in chiral metal-organic frameworks. Science 353, 808–811 (2016).

31. Inokuma, Y. et al. X-ray analysis on the nanogram to microgram scale using a Cryo electron tomogram. Proc. Natl Acad. Sci. USA 109, 7304–7309 (2012).

32. Lai, T. Y. et al. Structure of a designed protein cage that self-assembles into a highly porous cove. Nat. Chem. 6, 1065–1071 (2014).

33. Little, M. A. et al. Trapping virtual pores by crystal retro-engineering. Nat. Chem. 7, 153–159 (2015).

34. Ghosh, D., Kumari, P., Jaiman, D. & Shukla, A. K. Methodological advances: the unsung heroes of the GPCR structural revolution. Nat. Rev. Mol. Cell Biol. 16, 69–81 (2015).

35. Lieberman, R. L., Culver, J. A., Entminger, K. C., Pai, J. C. & Maynard, J. A. Crystalization chaperone strategies for membrane proteins. Methods 55, 293–302 (2011).

36. Dunbar, J. et al. SABDab: the structural antibody database. Nucleic Acids Res. 42, D1140–D1146 (2014).
39. Pardon, E. et al. A general protocol for the generation of Nanobodies for structural biology. Nat. Protoc. 9, 674–693 (2014).
40. Misasi, J. et al. Structural and molecular basis for Ebola virus neutralization by protective human antibodies. Science 351, 1343–1346 (2016).
41. Lee, P. S. et al. Receptor mimicry by antibody F045-092 facilitates universal binding to the H3 subtype of influenza virus. Nat. Commun. 5, 3614 (2014).
42. Meanwell, N. A. et al. Inhibitors of HIV-1 attachment. Part 2: an initial survey of indole substitution patterns. Bioorg. Med. Chem. Lett. 19, 1977–1981 (2009).
43. Wang, T. et al. Discovery of 4-benzoyl-1-[(4-methoxy-1H-pyrrolo[2,3-b]pyridin-3-yl]oxoacetyl]-2-(R)-methylpiperazine (BMS-378806): a novel HIV-1 attachment inhibitor that interferes with CD4-gp120 interactions. J. Med. Chem. 46, 4236–4239 (2003).
44. Li, M. et al. Human immunodeficiency virus Type 1 env clones from acute and early subtype B infections for standardized assessments of vaccine-elicited neutralizing antibodies. J. Virol. 79, 10108–10125 (2005).
45. Seaman, M. S. et al. Tiered categorization of a diverse panel of HIV-1 Env pseudoviruses for assessment of neutralizing antibodies. J. Virol. 84, 1439–1452 (2010).
46. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307–326 (1997).
47. Collaborative Computational Project, N.. The CCP4 suite: programs for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 50, 760–763 (1994).
48. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132 (2004).
49. Adams, P. D. et al. PHENIX: a comprehensive Python-based system formacromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221 (2010).
50. Delano, W. L. The PyMOL Molecular Graphics System (DeLano Scientific, San Carlos, CA, 2002).

Acknowledgements

We thank D. Burton and M. Fineberg for the sequences of NAC-identified antibodies, including the 3H109L member of the PGT121 lineage used in crystallization, M. Connors for the sequence of antibody 3SO22 used in crystallization, J. Stuckey for assistance with figures, and members of Structural Biology and Structural Bioinformatics Core Sections, Vaccine Research Center for discussions and comments on the manuscript and the WCMC/AMC/TSRI HIVRAD team for their contributions to the design and validation of near-native mimicry for soluble BG505 SOSIP.664 trimers. Funding was provided by the Intramural Research Program of the Vaccine Research Center, National Institute of Allergy and Infectious Diseases, the Intramural AIDS Targeted Antiretroviral Program, the National Institute of General Medical Sciences, National Institutes of Health. Use of sector 22 (Southeast Region Collaborative Access team) at the Advanced Photon Source was supported by the US Department of Energy, Basic Energy Sciences, Office of Science, under contract no. W-31-109-Eng-38.

Author contributions

Y.-T.L. determined the crystal structures with assistance from D.R.L. and P.D.K. S.O., M. K.N., B.L., K.M., R.T.B., N.A.D.-R. and J.R.M. performed virus neutralization assays. A.S. performed and analyzed ITC experiments. C.S.C., A.D., D.P., Y.Y. and B.Z. expressed and purified proteins. T.W., A.H. and J.S. synthesized small-molecule compounds. G.-Y. C. analyzed virus resistance data. Y.-T.L. and P.D.K. analyzed the data and wrote the paper.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-07851-1.

Competing interests: D.R.L. and T.W. own stock in Bristol-Myers Squibb. All other authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.