Biophysical Chemistry 196 (2015) 100–108

Contents lists available at ScienceDirect

Biophysical Chemistry

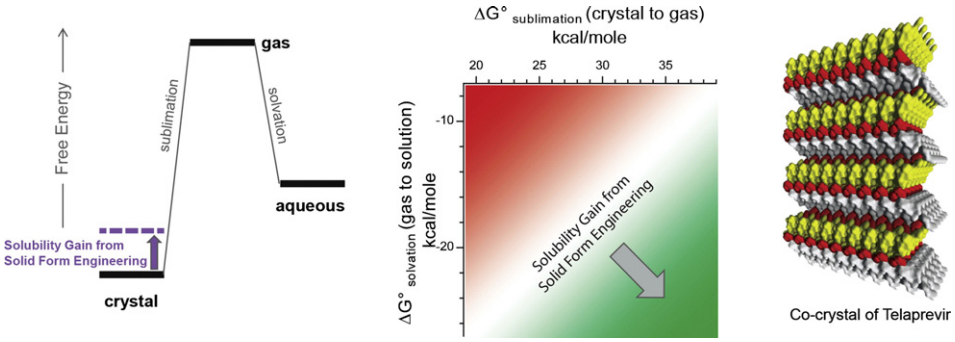
journal homepage: http://www.elsevier.com/locate/biophyschem

The potency–insolubility conundrum in pharmaceuticals: Mechanism and solution for hepatitis C protease inhibitors

Patrick R. Connelly *, Phillip W. Snyder, Yuegang Zhang, Brian McClain, Brian P. Quinn, Steve Johnston, Ales Medek, Jerry Tanoury, James Griffith, W. Patrick Walters, Eleni Dokou, Dragutin Knezic, Philip Bransford

Vertex Pharmaceuticals Incorporated, 50 Northern Ave., Boston, MA 02210, United States

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 19 June 2014
Received in revised form 4 August 2014
Accepted 4 August 2014
Available online 18 September 2014

Keywords:
Physical-organic
Drug design
Bioavailability
Solvation thermodynamics
Crystallography

ABSTRACT

As compounds are optimized for greater potency during pharmaceutical discovery, their aqueous solubility often decreases, making them less viable as orally-administered drugs. To investigate whether potency and insolubility share a common origin, we examined the structural and thermodynamic properties of telaprevir, a sparingly soluble inhibitor of hepatitis C virus protease. Comparison of the hydrogen bond motifs in crystalline telaprevir with those present in the protease–telaprevir complex revealed striking similarities. Additionally, the thermodynamics of telaprevir dissolution closely resembles those of protein–ligand dissociation. Together, these findings point to a common origin of potency and insolubility rooted in particular amide–amide hydrogen bond patterns. The insolubility of telaprevir is shown by computational analysis to be caused by interactions in the crystal, not unfavorable hydrophobic hydration. Accordingly, competing out the particular amide–amide hydrogen bond motifs in crystalline telaprevir with 4-hydroxybenzoic acid yielded a co-crystalline solid with excellent aqueous dissolution and oral absorption. The analysis suggests a generalizable approach for identifying drug candidate compounds that either can or cannot be rendered orally bioavailable by alteration of their crystalline solid phases, in an approach that provides a pragmatic way to attain substantial enhancements in the success rate of drug discovery and development.

1. Introduction

Pharmaceutical R&D, even today, suffers from a notoriously low success rate. A sizeable fraction of drug candidates fail, many because,

* Corresponding author. Tel.: +1 781 389 0682.
E-mail address: patrick_connelly@vrtx.com (P.R. Connelly).

http://dx.doi.org/10.1016/j.bpc.2014.08.008

although they are potent in vitro, their physical properties are poorly suited to becoming medicines [1,2]. These failures impose a severe cost on society, resulting both in more expensive health care and in otherwise preventable suffering [3]. Of these physical properties, aqueous solubility may cause the most trouble, owing to a conundrum widely known in the pharmaceutical industry: as compounds are optimized for higher potency, they tend to become less soluble in water [4]. Compounds with poor aqueous solubility generally fail to reach a concentration in the gastrointestinal tract that is sufficient for absorption into the systemic circulation, which makes them unusable as orally-administered drugs. The potency–insolubility conundrum has prompted speculation, concern, and attempts to taxonomize the problem or manage it with trial-and-error empiricism (including correlations with octanol water partition coefficients, Lipinski's Rule of 5 and the Biopharmaceutical Classification System) [2,5]. However, a fundamental molecular level thermodynamic connection between these two properties has yet to be elucidated.

Given how often drug discovery results in compounds that are both potent and insoluble, we wondered whether these properties might share some common origin. On a general level, potency and insolubility both reflect the stability of supramolecular structures formed by the drug. A potent drug forms a stable protein–drug complex, i.e., a complex with an unfavorable free energy of dissociation. Similarly, an insoluble drug forms a stable crystal, i.e., a crystal with an unfavorable free energy of dissolution. The processes of dissociation (for protein–drug complexes) and dissolution (for crystals of drug) are expressed formally by the following equilibria:

$$\mathrm{PX(aq)} \xrightarrow{K_D,\,\Delta J_D} \mathrm{P(aq)} + \mathrm{X(aq)} \tag{1}$$

$$\mathrm{X(s)} \xrightarrow{K_S,\,\Delta J_S} \mathrm{X(aq)}. \tag{2}$$

For both processes, general thermodynamic changes (ΔJ_D [dissociation] or (ΔJ_S [dissolution]) reflect the breakdown of noncovalent interactions in the supramolecular structures from their initial states (PX(aq) for the protein–drug complex and X(s) for the crystalline solid drug) and the formation of new interactions in their final states (unliganded protein, P(aq), and free drug in solution, X(aq)). With these broad structural and thermodynamic similarities in mind, we sought a system in which the conundrum manifested itself so that we could explore it in detail at the atomic level.

The consequences of the potency–insolubility conundrum became clear during the discovery and development of telaprevir (Fig. 1a), a small-molecule inhibitor of hepatitis C virus (HCV) protease [6,7]. Telaprevir binds tightly to HCV protease (K_D = ~30 nM at 37 ° C) but has low aqueous solubility. When stabilized as a high-energy amorphous form, the compound becomes much more soluble; this form was recently approved (as Incivek) to treat HCV infection. However, telaprevir required 20 challenging and costly years of R&D. In order to better understand the conundrum, and to solve it in a way that may be useful to streamline future work on potent and insoluble compounds, we studied the structural and thermodynamic properties of telaprevir and its target, the NS3 protease of HCV.

2. Results

The NS3 protease, like many proteases, has an active site composed substantially of apolar residues, with 355.7 Å of apolar surface area (52.5% of the total). Thus, successful competitive inhibitors of NS3 are likely to display an apolar surface as well. Apolar surfaces interact stably with each other as dehydration and packing effects lower the free energy of interaction. In addition, they create a felicitous environment for hydrogen bonds to form between amides on the inhibitor and backbone amides on the protease. The hydrophobic environment protects hydrogen bonds from the surrounding water and lowers the local dielectric, which results in tighter hydrogen bonds and a more stable protein–drug complex [8–10]. Amide–amide hydrogen bonds, in particular, are known to be crucial to many protein–drug interactions, protein–protein interactions, and protein folding [9].

We investigated the patterns of hydrogen bonds between NS3 protease and telaprevir in the manner of Etter et al. [11]. Strikingly, we determined that the very same pattern of amide–amide hydrogen bonds that composes the crystalline drug also forms in the protein–drug complex (Fig. 1). In crystalline telaprevir, these bonds create two fused supramolecular ring systems comprising 10 and 12 ring atoms, respectively, each ring system employing two hydrogen bond donors and two acceptors, as depicted in Fig. 1B. In graph set notation, these interactions are referred to as R2,2(10) and R2,2(12) to indicate rings containing 2 hydrogen bond donors and 2 hydrogen bond acceptors with 10 atoms and 12 atoms in the rings, respectively. Hydrogen bonds form between i) the amide oxygen and the proton on the terminal amide nitrogen, ii) the proton on the amide nitrogen and the adjacent activated ketone oxygen, and iii) the amide oxygen and the proton on the amide nitrogen. The same amide groups participate in the bond between the telaprevir and the NS3 protease, creating an identical 10 atom ring structure constructed from two hydrogen bond donors and two acceptors (Fig. 1C). In addition, other aspects of the supramolecular architectures of crystalline telaprevir (viz., a 12-atom ring motif constructed from two hydrogen bond donors and two acceptors) and the NS3–telaprevir complex (a 14 member ring motif constructed from two hydrogen bond donors and two hydrogen bond acceptors — R2,2(14)) are not only similar to each other, but also isostructural to parallel and anti-parallel β-sheets, respectively, in folded proteins. Further, in crystalline telaprevir, dimers joined by 10-member ring motifs form the basic unit of a structural hierarchy: the dimers assemble into rods, the rods into sheets, and the sheets into stacks (as shown in Fig. 1D–F). This stacked structure of crystalline telaprevir has a packing density of 0.7, close to the average packing density of the NS3–telaprevir complex (0.8) and of folded proteins generally (0.75) [12–13]. Taken together, these similarities emphasize that crystalline telaprevir and the telaprevir–protease complex share several structural features in common, structural features that are also shared by folded proteins. This commonality suggests a preliminary explanation for the potency–insolubility conundrum: a drug that binds tightly to NS3 protease, like telaprevir, may also bind tightly to itself, resulting in a stable, insoluble crystal. Thus, optimizing potency may also mean, unwittingly, optimizing the stability of the crystalline drug.

To investigate the energetic relationship between potency and insolubility, we studied the thermodynamics of transfer of crystalline telaprevir to water. The very low aqueous solubility (~7 μmol) and the rate of dissolution of crystalline telaprevir made direct calorimetric measurements for heats of solution inaccessible by experiment (i.e., below the detection limits of even the most sensitive calorimeters available). Therefore, to characterize the thermodynamics of dissolution, we determined the solubility of telaprevir in water across the temperature range of 5°–65 °C to high precision (Fig. 2a). To analyze the dissolution thermodynamics of telaprevir, it is crucial to take into account all the species that form in solution from three reversible reactions as follows: i) epimerization of the stereocenter adjacent to the ketone, ii) hydration of the ketone, and iii) hydration of the ketone of the epimer. The Supplementary information details the experimental approach used to deconvolute the thermodynamic parameters of dissolution from those for hydration of the ketone (which implemented cryo-probe NMR spectroscopy) and epimerization (which implemented liquid chromatography).

In three aspects, telaprevir's thermodynamics of dissolution (or, formally, the transfer from the crystal to aqueous solution, $N_{xtal} \rightarrow N_{aq}$ in Table 1) closely resembles the processes of protein unfolding and protein–ligand dissociation. First, the van't Hoff plot in Fig. 2a shows distinct curvature as dissolution takes place with a large positive heat

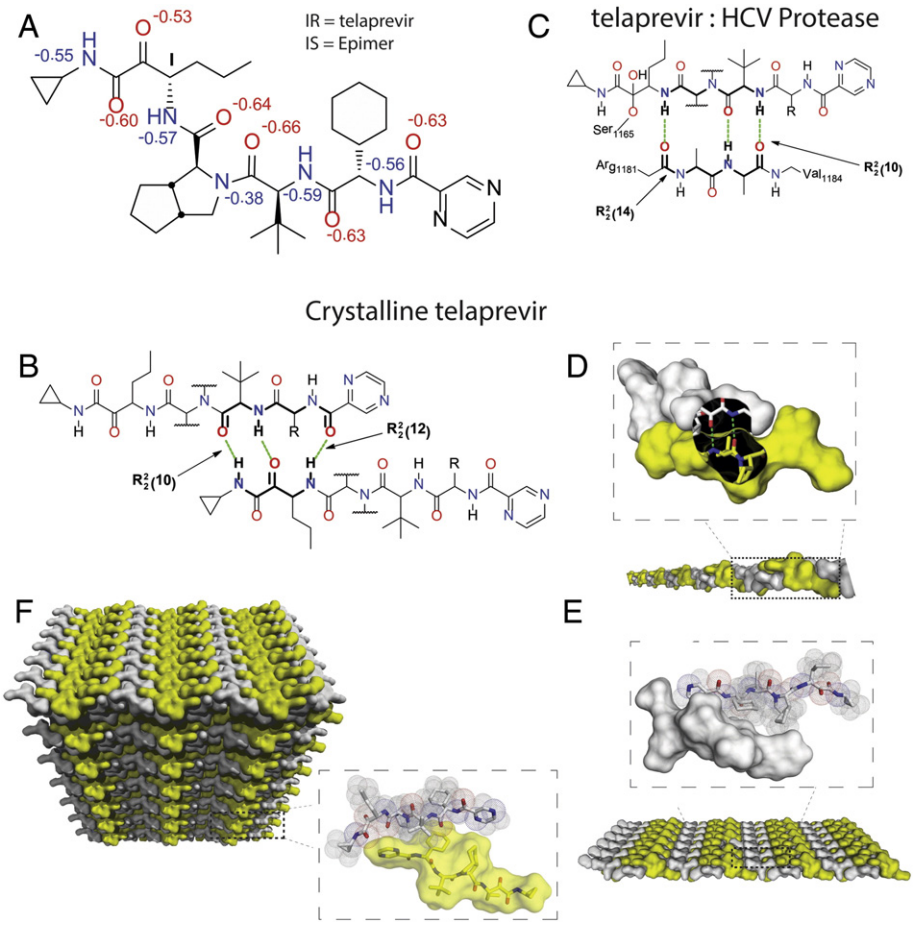


Fig. 1. Structural features of telaprevir. Nitrogen atoms are highlighted in blue and oxygen atoms in red. A) Telaprevir structure, showing natural charges (positive, blue; negative, red) as calculated by NBO 5.9 [24]. B) The parallel telaprevir dimer forms ring motifs of R2,2(10) and R2,2(12) graph sets. C) Telaprevir binds similarly to NS3 protease, forming motifs of R2,2(10) and R2,2(14) graph sets. D–F) Structural hierarchy of telaprevir: telaprevir dimers assemble into rods (D), the rods into sheets (E), and the sheets into stacks (F).

capacity change ($\Delta C_p = 260 \pm 70$ cal/K-mol), which is accounted for by the changes in solvent-accessible polar and nonpolar surface areas from crystalline to solution states. (The value calculated using the method of surface area coefficients [14] is 330 cal/K-mol.) A large positive heat capacity change has been observed in the aqueous dissolution of other crystalline systems and, importantly, is a common feature of other noncovalent supramolecular disassembly processes including protein unfolding, protein–peptide dissociation, and the dissociation of many protein–drug complexes [15–17]. (ΔC_p for the dissociation of telaprevir and NS3 protease is estimated to be 106 cal/K-mol [see Supplement].) Second, the ratio of the entropy change to the heat capacity change ($\Delta S/\Delta C_p = 0.2$ at 25 °C) is remarkably close to that observed for protein unfolding and the aqueous dissolution of nonpolar gases, liquid hydrocarbons, and hydrophobic crystalline cyclic dipeptides (see inset in Fig. 2b). Lastly, the stability of both crystalline telaprevir and the telaprevir–NS3 complex, as given by the free energy of dissolution/dissociation in water, has a maximum in temperature of 87 °C (Fig. 2B). Other processes that exhibit this stability phenomenon include the cold and heat denaturation of proteins, dissociation of protein–peptide complexes and telaprevir–NS3 binding (see Supplement) [18–20]. In sum, while the protein–drug interaction of NS3 and telaprevir is structurally and energetically comparable to other assembly processes involving proteins (protein–protein interactions and protein folding), it is also structurally and energetically comparable to the drug–drug interaction that stabilizes crystalline telaprevir. Thus, potency and insolubility seem to derive from a common origin that includes the same amide–amide hydrogen bond patterns, sheltered by a hydrophobic microenvironment.

To relate the observed thermodynamics of dissolution of telaprevir to the structural features that stabilize its crystal, and to determine why the crystalline form of the drug is so insoluble in water, we calculated the thermodynamics of dissolution in a two-step process. First, we used molecular mechanics and normal mode analysis to calculate the thermodynamic parameters (ΔG, ΔH, ΔS, and ΔC_p) for the transfer of telaprevir from its crystal to the vapor phase; then we used a molecular dynamics/free-energy perturbation technique to calculate the thermodynamics for transfer from the vapor phase to solution [21–23]. Table 1 details the net results for comparison to the experimental values of thermodynamic parameters, and the supplementary information details the computational methods employed. The magnitudes of the free energies of transfer from crystal to vapor are large and positive, while those from vapor to water are large and negative. These findings clearly demonstrate that the stability of the crystal lattice ($\Delta G_{xtal \rightarrow vapor}$), rather than the compound's aversion to water ($\Delta G_{vapor \rightarrow aqueous}$), is responsible for the insolubility of telaprevir. Of the structural factors that contribute to insolubility, electrostatic and dispersion interactions among molecules of telaprevir in the crystal lattice are the largest (i.e., values of $\Delta G = 44$ kcal/mol and $\Delta G = 22$ kcal/mol for electrostatic and dispersion interactions for the transfer of crystal to vapor, respectively, Table S3).

Having concluded that the interactions in crystalline telaprevir are principally responsible for its insolubility, we hypothesized, as a practical corollary, that interrupting the hydrogen bonding and packing that

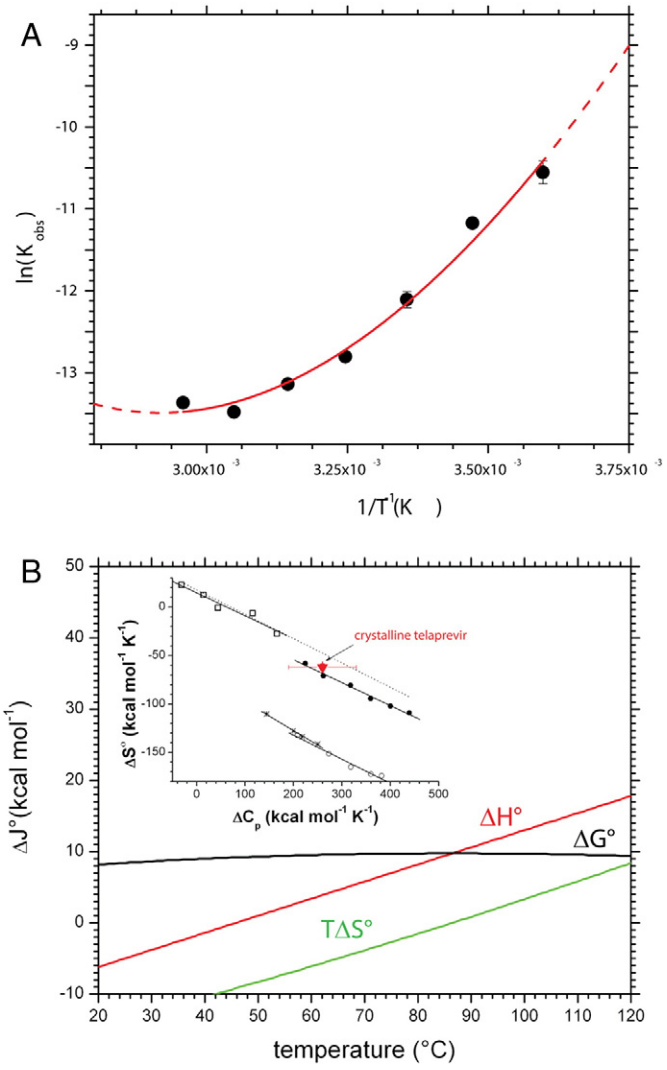


Fig. 2. Thermodynamics of telaprevir dissolution in water. A) van't Hoff plot showing the retrograde temperature-dependent solubility for telaprevir in water. Curve is fitted to experimental results, quantified by HPLC, with parameters reported in Table 1. See Supplement for experimental details. B) Thermodynamic functions for the transfer of crystalline telaprevir to water, showing behavior similar to that for protein unfolding and protein–ligand dissociation. The labels indicate the functions for enthalpy (◎H°), entropy (T◎S°), and free energy (◎G°) of dissolution. The inset shows relationships between change in entropy and change in heat capacity at constant pressure for the dissolution in water of rare gases (x), saturated hydrocarbons (○), liquid hydrocarbons (●), solid cyclic peptides (□), and crystalline telaprevir (red). The dotted line represents a linear fit of data for the unfolding of 11 proteins [17].

stabilize the crystal could result in a higher-energy solid form thereby enhancing the effective aqueous solubility of the compound. We focused on the common hydrogen bond motif – the ten atom ring system constructed from hydrogen bonds formed between the proton of the nitrogen and the oxygen of the amides straddling the *tert*-butyl group – found in both the crystal of telaprevir and the NS3–telaprevir complex. We evaluated the natural charge on all amide units using NBO 5.9 and found that the oxygen adjacent to the octahydrocyclopenta[*b*]pyrrole ring had the most negative natural charge. Correspondingly, the nitrogen of the same amide bond was overwhelmingly more electropositive than the other N atoms that could participate in hydrogen bonds (Fig. 1A). This result is consistent with Etter's rules [11] and points to this bond as the likely strongest hydrogen bond stabilizing both the crystal of telaprevir and, perhaps, the NS3–telaprevir complex.

The above analysis suggests that using another molecule to interrupt the crucial hydrogen bond and form a co-crystal may lead to a higher-energy, more soluble solid form. To that end, we tested a range of amide- and carboxylic acid-containing compounds, which have the ability to form ring motifs mimicking, and competing energetically with, those formed in crystalline telaprevir. 4-Hydroxybenzoic acid (4-HBA) was found to form a 1:1 co-crystal with telaprevir that contained the expected similar supermolecular ring structure in place of the previous O–HN interaction (Fig. 3a). (Experimental details are provided in the Supplement.) As in the neat crystal, telaprevir dimers assemble into rows, then sheets (Fig. 3b–c); however, in this case the stacking of the sheets is interrupted by rows of 4-HBA that have bonded to individual molecules of telaprevir (Fig. 3d). This co-crystal form displayed a 100-fold enhancement in effective solubility[1] over that of neat crystalline telaprevir (Fig. 3e). Crucially, this increase in in vitro solubility translates to an increase in in vivo exposure. When the telaprevir:4-HBA co-crystal was dosed at 375 mg in dogs, it achieved an ~8-fold increase in oral exposure over a suspension of neat crystalline telaprevir (Fig. 3e). In fact, the integrated area under the curve of concentration versus time demonstrates that the 4-HBA co-crystal achieves the same exposure as the commercial tablet of telaprevir which contains an amorphous form of the drug. Analysis of the pharmacokinetics (assuming the measured effective solubility for the materials in simulated intestinal fluid at 37 °C, and assuming permeability estimates scaled from Caco-2 measurements) with an ACAT (advanced compartmental absorption and transit) model reveals that the improved oral exposure is a direct consequence of the enhanced effective aqueous solubility of the 4-HBA co-crystal (see solid lines in Fig. 3e; detail in SI Appendix) [26]. In the model, all disposition parameters (e.g., clearance, two compartment kinetic parameters, and permeability) are all constrained to be the same between amorphous and co-crystalline telaprevir as described in Materials & methods.

3. Discussion

Thus, the case of telaprevir is one example of how a potent, insoluble drug can be rendered as a viable solid dosage form: we identified the supramolecular structural similarity between the protein–drug complex and the crystalline drug (the ten atom ring system comprised of 2 hydrogen bond donors and 2 acceptors), determined that the insolubility of telaprevir was due to the strength of its crystal lattice – most notably its hydrogen bonds – and interrupted these bonds with a competing molecule. This approach may not be the only way to optimize both potency and solubility: in some cases, it may be feasible to re-design the covalent structure of the molecule so that its potency and solubility are not driven by the same atoms. Further, in cases where a protein–drug crystal structure is not available, structure–activity analyses may still reveal the atoms responsible for potency. However, this solution has the same prerequisite: understanding the structural origin of the compound's insolubility, as demonstrated above.

Although our approach to telaprevir was successful, we recognize that it is useful mainly insofar as it can be generalized to other compounds. Further investigation revealed that telaprevir is far from unique. We examined the structures of several other drug compounds which showed the same striking pattern, forming similar hydrogen bonding

[1] This study examined what is often termed "kinetic solubility" or "effective solubility" of the co-crystal, rather than the true thermodynamic solubility of the co-crystal in water. Formally, the solution produced by dissolving the co-crystal in water is supersaturated with respect to telaprevir and the co-crystal must be stabilized by adding polymers to the dosing suspension (see SI Appendix). At equilibrium the co-crystal will, over time, undergo a solvent-mediated phase transition to neat telaprevir and the amounts of compound in solution will decrease accordingly to that of neat telaprevir. This transition is slow so that the effective solubility of telaprevir in aqueous solution produced by dissolving the co-crystal is high. In vitro experiments in which excess co-crystalline solid is administered to simulated gastric fluid at 37 °C show that the concentration of drug in solution for up to 3 h is enhanced many 100-fold over that of neat crystalline telaprevir. Stabilization of supersaturation of aqueous co-crystal suspensions by the addition of polymers is a commonly observed effect; however, the precise mechanism of such stabilization requires further investigation.

Table 1
Measured, derived, and calculated values of the thermodynamic parameters of dissolution of telaprevir in aqueous solution. K_{eq}, equilibrium constant; $\Delta G°$, change in free energy; $\Delta H°$, change in enthalpy; $\Delta S°$, change in entropy; $\Delta C_p°$, change in heat capacity at constant pressure.

Equilibrium[a]	K_{eq}	$\Delta G°$ (kcal/mol)	$\Delta H°$ (kcal/mol)	$\Delta S°$ (cal/K-mol)	ΔC_p (cal/K-mol)
Measured values[b]					
$N_{(xtl)} \rightleftharpoons^{K_{obs}} N_{(aq)} + H_{(aq)}$	8.7 ± 0.8 μM	6.91 ± 0.09	−11.5 ± 0.8	−62 ± 3	260 ± 70
$N_{(aq)} \rightleftharpoons^{K_H} H_{(aq)}$	8.0 ± 0.3	−1.23 ± 0.04	−5.5 ± 0.4	−14 ± 1	NA[c]
Derived values[d]					
$N_{(xtl)} \rightleftharpoons^{K_{sol}} N_{(aq)}$	1.0 ± 0.1 μM	8.2 ± 0.1	−5 ± 1	−45 ± 3	240 ± 70
Calculated values[e]					
$N_{(xtl)} \rightleftharpoons^{K_{sol}} N_{(aq)}$	0.1 μM	9.4 ± 0.3	−8 ± 6	−55 Desmond	330[f]
	250 μM	4.9	26.9	73.8 Cosmo	
	4.3E7 M	−10.4	11.6	73.8 ZAP	
$N_{(g)} \rightleftharpoons^{K_{solv}} N_{(aq)}$	1.4E16	−18.9 ± 0.3	−67 ± 6	−146 Desmond	
	1.6E17	−23.4	−31.7	−27.8 Cosmo	
	3.0E28	−38.7	−47.0	−27.8 ZAP	
$N_{(xtl)} \rightleftharpoons^{K_{vap}} N_{(g)}$	1.5E−21	28.3	58.6	101	

[a] N represents keto form, while H represents hydrated form. All thermodynamic parameters represent estimates from van't Hoff analysis at reference temperature of 298.15 °C.
[b] Data and fits appear in supporting information.
[c] Data did not support fitting with a heat capacity term.
[d] Equations used to derive these values and their derivations appear in supporting information. Errors represent the values determined by propagation.
[e] Values for free energy, enthalpy, and entropy were calculated using the computational methods described in the supplemental information.
[f] The value of the heat capacity was estimated using the relationship determined by Murphy and Gill.

motifs in crystals and when bound to their targets. First, we examined the three other HCV protease inhibitors. In each case, identical structural similarities among their neat forms, complexes with HCV protease, and co-crystals formed with carboxylic acids as those found with telaprevir were revealed (see SI Appendix). In addition to HCV protease inhibitors, we extended the analysis to the targets of classic insoluble drugs (viz., cyclosporine, celecoxib, tacrolimus and fluconazole), to the HIV protease inhibitor ritonavir, and to the non-nucleoside HIV reverse transcriptase inhibitor, efavirenz. In each case, similar hydrogen bond patterns formed between neat crystalline and protein-bound structures (see SI Appendix). Efavirenz forms nearly identical hydrogen-bonding patterns in the protein–drug complex as it does in the thermodynamically most stable crystalline form (Fig. 3e): the protein–drug complex is connected by a 9 member supermolecular ring system comprised of two hydrogen bond donors and two hydrogen bond acceptors (R2,2(9)); the neat efavirenz crystal is constructed from interacting 8 member rings utilizing the very same hydrogen bond donor and acceptor atoms. Moreover, in two co-crystals of efavirenz, co-formers compete with the key hydrogen bond-forming unit to interrupt the principle structural motif in the neat crystal — the same phenomenon that allowed us to create the co-crystal of telaprevir with 4-hydroxybenzoic acid. The prevalence of these hydrogen bonding trends across investigational drugs remains to be determined, as does their link with the thermodynamics of their binding and aqueous dissolution. However, the trend provided here suggests that the approach developed here for telaprevir may be applicable beyond protease inhibitors — or even peptidomimetic drugs in general.

Regardless of the structural origin of the potency–insolubility conundrum, one may generally leverage elements of the analysis detailed here for its practical value in determining which insoluble compounds can be improved by manipulation of their solid forms and which cannot. Computationally, one may determine if the insolubility is due to the compound "liking its crystal" (i.e., favoring the crystalline state energetically) or due to the compound "disliking water" (i.e., disfavoring the aqueous state energetically) by calculating the sublimation and hydration free energies, respectively. By examining the relative values of the solvation and sublimation free energies, it may be feasible to determine if the origin of the insolubility is driven largely by the stability of the compound's solid state or by unfavorable solvation in water. Compounds that are deemed insoluble by virtue of their neat crystalline solid states can be subjected to solid form investigations targeting the identification of higher energy/higher solubility solids such as co-crystals and amorphous materials. Compounds with insolubility driven by less favorable solvation in water can be spared from fruitless attempts to identify higher energy/higher solubility solids. In this way, computation can cut down on costly preclinical and clinical development efforts that will lead to dead ends.

Fig. 4. provides a precise illustration of how sublimation and solvation free energies taken together with an understanding of the intestinal permeability of a compound can be used to evaluate the potential for a solid to achieve reasonable plasma concentrations upon oral administration. Fig. 4A is a contour plot showing the fraction of a dose that is absorbed into plasma as functions of the intestinal permeability (as given by the permeation number) and solubility (as given by the dose number). For compound X, with a permeation number of 1, the solubility of a neat crystalline form is too low to achieve any significant absorption upon dosing. However, when the compound is made more soluble (e.g., by rendering as a higher energy solid form such as an amorphous form), it achieves substantial oral absorption (>30% bioavailability). For compound Y, with permeation number = 10, a boost in solubility doesn't provide for appreciable oral absorption (<10% bioavailability). By estimating a compound's permeability through, for instance, a standard Caco-2 cell assay, the critical questions that can be answered by calculations are as follows: 1) what effective solubility is necessary to enable appreciable oral absorption?; and 2) can solubility be enhanced by rendering the compound as a high energy solid, such as an amorphous or co-crystalline material? Simple oral absorption calculations as depicted in Fig. 4A provide insight into the first question. The calculation of the sublimation and solvation energies as in Fig. 4B helps answer the second question together with experiments that test the solubility advantage achieved from rendering a compound amorphous. The free energy of solution is given by the sum of solvation and sublimation free energies, and it can be assumed to first approximation that the solubility advantage gained from rendering a compound amorphous applies to the sublimation free energy term. For example, Compound X in Fig. 4B has 1 μM aqueous solubility under physiological pH and temperature; its insolubility is largely due to its high sublimation energy (51 kcal/mol) relative to its high solvation energy. If rendering the compounds as an amorphous solid can lower the sublimation energy by 5 kcal/mol, it will enhance the solubility by 100 fold, leading to an appreciable fraction absorbed (Fig. 4 A). In contrast, compound Y has

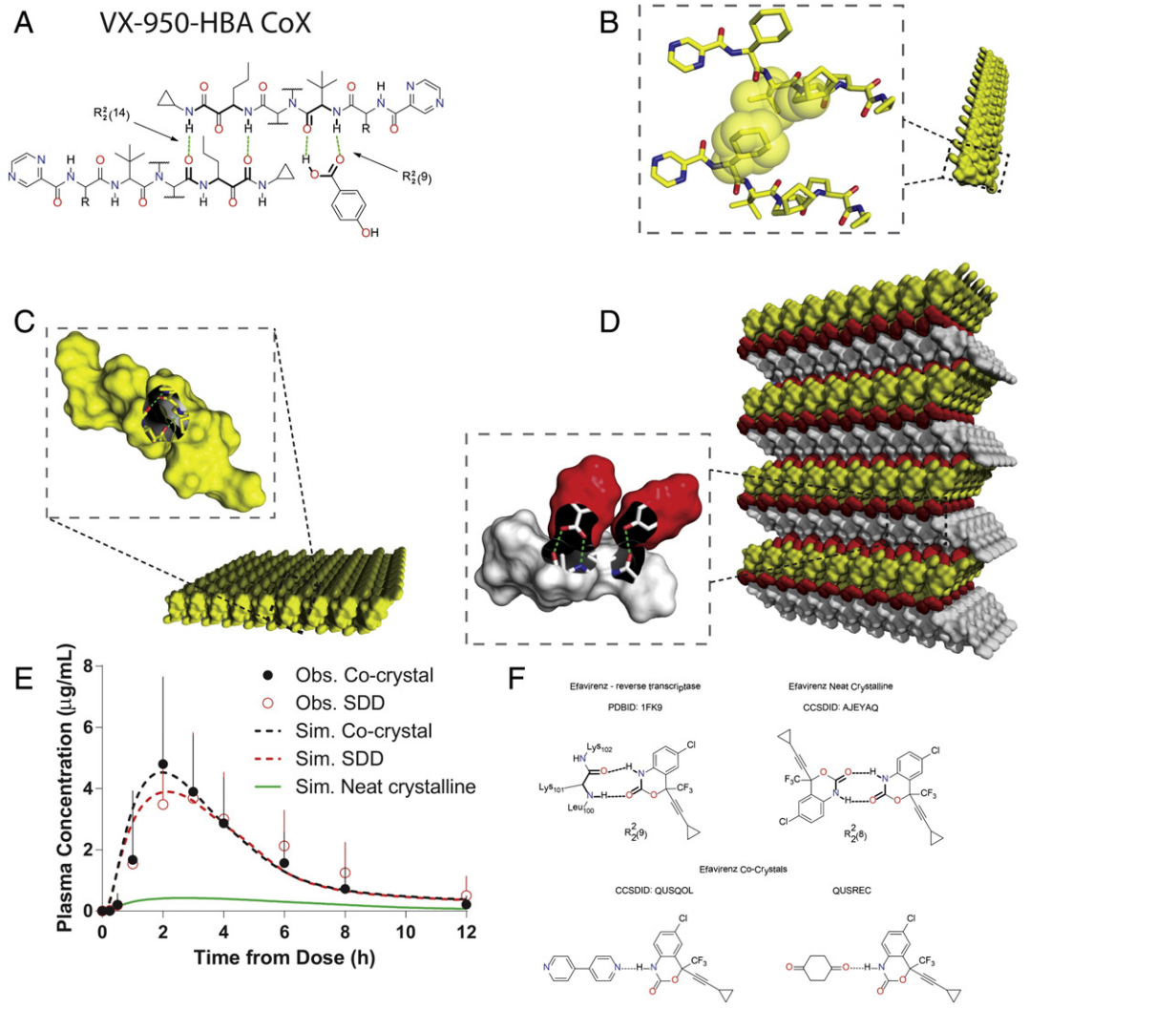


Fig. 3. Structure, function, and comparisons of the 1:1 telaprevir/4-HBA co-crystal. A) Telaprevir/4-HBA co-crystal structure, depicting ring motifs with graph sets similar to telaprevir crystals and NS3–telaprevir complexes — indicating that 4-HBA competes out telaprevir molecules during crystal formation. Nitrogen atoms shown in blue, and oxygen in red. B–D) Structural hierarchy of telaprevir/4-HBA: as in the neat crystal, telaprevir dimers assemble into rods (B), then sheets (C), but here, sheets are interrupted by rows of 4-HBA molecules (D). E) The telaprevir/4-HBA co-crystal (black squares) showed 100-fold higher aqueous solubility (in simulated intestinal fluid) than the commercial telaprevir tablet (purple squares). Solid lines represent pharmacokinetic models that use solubilities in simulated gastric and intestinal fluid as measured in vitro (black, co-crystal; blue, commercial tablet; violet, neat crystalline telaprevir). F) For efavirenz, protein–drug and drug–drug complexes display nearly identical hydrogen-bonding patterns.

nearly a 1 nM solubility which is largely due to its weak solvation energy (−23 kcal/mol) relative to its sublimation energy (44 kcal/mol). Rendering this compound as an amorphous solid to decrease the sublimation energy barrier by ~2 kcal/mol fails to improve its solubility enough (Fig. 4B); the consequence is that its anticipated bioavailability is less than 10% (Fig. 4A). Operationally, one can utilize estimates of the intestinal permeability from an in vitro Caco-2 assay, estimates of the solubility in the intestinal lumen from determinations of solubility in simulated intestinal fluids for crystalline and amorphous forms of a compound, and/or estimates of the sublimation and solvation energies from calculation, to help make a critical drug development decision: to suspend further preclinical and clinical studies and avoid fruitless expensive and time consuming work that will not be successful, or to advance a compound with a high probability of a successful outcome.

When these computational approaches are combined with recent advances in the ability to predict crystal structures, a path opens up to compute aqueous solubility *as well as* its aqueous solvation and sublimation contributions *prior* to making compounds. Having this knowledge for compounds at the medicinal chemistry design stage *prior* to the actual synthesis allows one to evaluate the relative merits of pursuing the costly synthesis and biological characterization of vast sets of compounds, only to find out later that they cannot be rendered as solids with sufficient effective aqueous solubility to be developed as oral drugs. It enables focus on compounds that are likely to succeed by improvements in aqueous solubility attained through solid phase manipulations. In these ways, the overall approach outlined here could potentially increase the speed and success rate of drug discovery and development dramatically; however, rigorous validation of computational methods, and sufficient precision of the methods, will be necessary to achieve the potential success rate improvements.

The potency–insolubility conundrum has led to countless powerful, promising medicines being abandoned long before a patient ever takes them. It has also led to lengthy and costly failed efforts to formulate insoluble compounds through solid form manipulations that had little to no chance of succeeding in the first place. Overcoming this Catch-22, in at least some cases, could lead to less costly drugs that reach patients more quickly. A recent study calculated the cost of developing a new drug at $1.9 billion, 10-fold higher than in the 1970s [25]. This growth in costs is not sustainable indefinitely; the need for smarter, more efficient drug discovery is immediate and intense. We hope that our approach will be one of many contributions to this end.

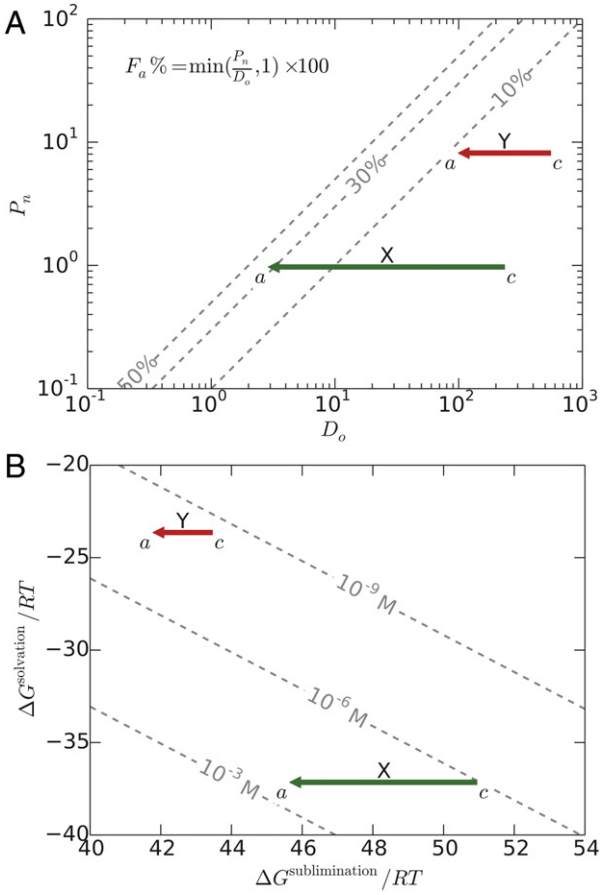


Fig. 4. Computational evaluation of the potential for improvements insolubility and oral absorption of compounds by altering their solid forms. A) Contour plot of bioavailability, F_a, versus dose number, D_o as a scaled measure of solubility and permeation number, P_n as a scaled measure of gut permeability. D_o is the ratio of the mass of administered dose divided by volume of fluid intake relative to the thermodynamic solubility. P_n is the ratio of permeation rate through the gut wall relative to the transit rate through the small intestine. Arrows depict the change in dose number and bioavailability due to transformation from crystalline (c) to amorphous material (a). Compound X's bioavailability improves sufficiently to warrant pharmaceutical development while Y's improves negligibly and does not warrant development. B) Contours of solubility versus sublimation and solvation free energy. As in A, arrows depict the transformation from crystalline to amorphous material. Compound X is more soluble than Y, and its solubility increases more when transformed from a crystalline (c) to an amorphous (a) material.

4. Materials & methods

The SI text details the procedures employed for single molecule crystallography, determination of the thermodynamic parameters of dissolution, computational approaches, and pharmacokinetic assays in vivo. The procedures are briefly described here.

5. Crystallography

Data were collected on a Bruker APEX II CCD diffractometer (Bruker AXS, Madison, Wisconsin). The instrument is equipped with a Cu Kα sealed-tube X-ray generator. Data are collected at 100 K with a nitrogen open flow system or at room temperature. Cu Kα seal tube radiation was used at 40 kV, 35 mA. Camera distance is set at 5.25 cm. Oscillation photos around ω and ϕ angles were collected. Data were integrated and scaled by the APEX software. The structures are solved and refined with the SHELXTL package.

6. Determination of the thermodynamic parameters of dissolution

We quantified the solubility of telaprevir by determining, with high-performance liquid chromatography (HPLC), the concentration of the compound in the aqueous portion of saturated suspensions of the crystalline material in water. van't Hoff analysis (i.e., the measurement of solubility at varying temperature) provided the apparent enthalpy, entropy, and heat capacity of dissolution. The apparent concentration of telaprevir at equilibrium at a given temperature, however, depends on the equilibria depicted in Fig. 1B, namely the hydration of the α-ketoamide and epimerization of the chiral center at C-3. To derive the actual values of the equilibrium constant between the crystalline and solution forms of telaprevir, we measured the equilibrium constants for i) hydration of the α-ketoamide (by solution phase nuclear magnetic resonance experiments) and ii) epimerization of the chiral center at C-3 (by HPLC using a method that provided separation of the two diastereomers). Measuring the temperature dependence of these equilibrium constants provided estimates of the enthalpy and entropy of hydration and epimerization, again, by van't Hoff analysis. Below we detail the experimental procedures and derive the expressions for estimating the thermodynamic parameters that appear in Table 1 and the corresponding errors in those estimates.

7 . Gas to solution transfer calculations

Three approaches were used to calculate the solvation free energies of telaprevir and are presented for comparison. The methods are referred to as COSMO/SA [22], ZAP/SA [23], and DESMOND [21]. The first 2 methods, COSMO/SA and ZAP/SA, split the solvation free energy into two components — an entropic surface area term and an enthalpic electrostatic term. The surface area term is the same in both methods and is calculated using the method of Rizzo et al. which uses an atom type weighted surface area function [28]. The enthalpy for the COSMO/SA approach was calculated using the density functional theory module DMol3 with the Cosmo solvation model from within Materials Studio [27]. The enthalpy for the ZAP/SA approach used OpenEye's Poisson Boltzmann solver Zap with AM1BCC charges [23]. To account for multiple conformations in the gas state, ten low energy conformers generated with OpenEye's Omega2 were used in the COSMO/SA and ZAP/SA calculations to obtain an average result. The energy window for the ten conformers was 2.7 kcal/mol. The third method used the Desmond molecular dynamics/free energy perturbation module of Maestro. These calculations followed the approach described by Shivakumar and coworkers [21]. For each Desmond run, the default parameters were selected except for temperature (which varied between 278 and 318 K) and water model (TIP4pEW). The starting structure of a single molecule of telaprevir was identical to that used for the crystal-to-gas calculations. The simulations comprised a 2 ns simulation of telaprevir solvated in TIP4pEW water molecules in a box that extended 10 Å from telaprevir in each direction. The key output of the simulation was the "solvation energy," which corresponds to the free energy of transferring a single molecule of telaprevir from aqueous solution to the gas phase. This value was corrected to a standard state of 1 M using the following procedure. The output values for the volume of the simulation box at each time step were averaged to determine the volume that contained a single molecule of telaprevir, and this volume was used to calculate the average concentration during the simulation. The free energy at the reference state of 1 M was calculated using Eq. (3):

$$\Delta G = -RT \ln (c/c_{ref}). \quad (3)$$

The error in the values of ΔG_{solv} is that determined by propagation from the single molecule free energy and the variation in the volume of the box and the simulated temperature over the course of the simulation. The values of the solvation free energy estimated at different temperatures were plotted as a function of the inverse of the simulation

temperature (T_{sim}). The simulation temperature was also determined from the output data from each run (i.e., values of T_{sim} represent the average temperature over each time step of the simulation). The values of the enthalpy of solvation were determined by van't Hoff analysis of the computed free energies at different temperatures, while the data were fit by a nonlinear least squares method using three adjustable parameters: K_{sol}, ΔH, and ΔC_p at the reference temperature of 298.15 K. We found that the solvation results from DESMOND, when combined with our calculated gas to crystal results, best matched the experimental data for the neat crystal form of telaprevir.

8. Gas to crystal transfer calculations

Gas to crystal transfer energies were calculated with the AMOEBA (atomic multipole optimized energetics for biomolecular applications) force field [29] — available in the Tinker molecular modeling package. AMOEBA attempts to approach quantum mechanical accuracy by using an atomic fixed charge model of monopole up to quadrupole approximation along with a polarization term which allows for charge redistribution in response to an external electric field. To calculate equilibrium charge distributions, multiple rounds of self-induction of dipole centers are carried out until convergence is achieved using the method of Thole [30].

For non-trivial molecules, the electronic/torsional parameters required by AMOEBA must be pre-calculated using quantum mechanical methods. For this study we used Gaussian 09 [31]. The full parameter generating process can now be carried out automatically by use of the program Poltype [32] — a python script which runs the QM calculations, derives the electronic/torsional parameters, and obtains the final van der Waals and remaining bonded parameters from a lookup table. The large number of atoms in telaprevir – 102 – made running the QM calculations on the whole molecule unfeasible. Thus we fragmented telaprevir and processed each fragment with Poltype. The resulting parameters for each fragment were then combined to create the final telaprevir parameter file. The fragments overlapped to some extent and ranged in size from 15 to 45 atoms. To obtain gas phase conformations, each of the two independent molecules in the crystallographic asymmetric unit was minimized in vacuum using Tinker's minimize program with an energy convergence cutoff of 0.001 kcal/mol/A. For gas phase, calculations were performed on both molecules and the average was used. The crystallographic unit cell was minimized with Tinker's xtalmin program using particle mesh Ewald summation for electrostatic interactions. Due to the greater difficulty in obtaining convergence with mutual polarization, an energy convergence cutoff of 0.3 kcal/mol/A was used.

The transfer of a molecule from gas to the crystal involves a loss of translational and rotational degrees for freedom, a change in the vibrational modes and frequencies, and the introduction of intermolecular non-bonded interactions. The calculation of the translational, rotational and vibrational transfer energies followed the general protocol laid out by Brady and Sharp [20]. All calculations were at T = 298 K.

9. Pharmacokinetics

To determine whether the higher solubility of the telaprevir/4-HBA co-crystal in vitro translated to increased absorption in vivo, we dosed suspensions of crystalline telaprevir and of the co-crystal in methyl cellulose or hydroxymethyl cellulose acetate succinate in rats at a dose level of 30 mg/kg and compared the resulting pharmacokinetic profiles. Then, to compare the absorption of the co-crystal with that of commercially marketed telaprevir, we processed the co-crystal into tablets and dosed these and commercial telaprevir tablets in male beagle dogs. Finally, to determine whether the increased exposure observed for the co-crystal in dogs is actually due to the co-crystal's greater solubility (as opposed to, e.g., increased permeability or decreased clearance), we constructed a model using GastroPlus (version 8.0, Simulations Plus, Lancaster, CA) with the following input parameters: permeability (4.73×10^{-4} cm/s), solubility for the co-crystalline telaprevir in stomach is 0.15 mg/mL and the intestine is 0.1 mg/ml, while for the commercial tablet it is 0.1 mg/mL for both the stomach and the intestine. The dissolution rates for the co-crystal and amorphous materials are .24 and .48 ml/mg/min, respectively. Conventional physiological parameters for dogs (e.g., intestinal surface areas and transit times) were used. The disposition parameters were constrained to be the same for both co-crystalline and amorphous telaprevir with clearance of 0.64 L/h/kg, first pass extraction of 41%, plasma volume of 0.29 L, and K12, K21, K13, and K31 as 2.25, 3.42, 0.94 and 0.16 h^{-1} respectively.

Acknowledgments

We wish to acknowledge discussions with John Thomson, Youssef Bennani, Patricia Hurter, Mariusz Krawiec, Mark Namchuk, Kumkum Saxena, William Taylor, Chong-Hui Gu and William Nugent.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bpc.2014.08.008.

References

[1] R.A. Lipper, E pluribus product, Mod. Drug Discov. 2 (1999) 55–60.
[2] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, Adv. Drug Discov. Rev. 46 (1) (2001) 3–26.
[3] J.A. DiMasi, R.W. Hansen, H.G. Grabowski, The price of innovation: new estimates of drug development costs, J. Health Econ. 22 (2) (2003) 151–185.
[4] G.L. Amidon, H. Lennernäs, V.P. Shah, J.R. Crison, A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability, Pharm. Res. 12 (3) (1995) 413–420.
[5] I.M. Jacobson, J.G. McHutchison, G. Dusheiko, et al., Telaprevir for previously untreated chronic hepatitis C virus infection, N. Engl. J. Med. 364 (25) (2011) 2405–2416 (et).
[6] A.D. Kwong, R.S. Kauffman, P. Hurter, P. Mueller, Discovery and development of telaprevir: an NS3-4A protease inhibitor for treating genotype 1 chronic hepatitis C virus, Nat. Biotechnol. 29 (11) (2011) 993–1003.
[7] J.L. Kim, K.A. Morgenstern, C. Lin, et al., Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide, Cell 87 (2) (1996) 343–355.
[8] J. Gao, D.A. Bosco, E.T. Powers, J.W. Kelly, Localized thermodynamic coupling between hydrogen bonding and microenvironment polarity substantially stabilizes proteins, Nat. Struct. Mol. Biol. 16 (7) (2009) 684–690.
[9] D.W. Bolen, G.D. Rose, Structure and energetics of the hydrogen-bonded backbone in protein folding, Annu. Rev. Biochem. 77 (2008) 339–362.
[10] P.R. Connelly, R.A. Aldape, F.J. Bruzzese, et al., Enthalpy of hydrogen bond formation in a protein–ligand binding reaction, Proc. Natl. Acad. Sci. 91 (5) (1994) 1964–1968.
[11] M.C. Etter, J.C. MacDonald, J. Bernstein, Graph-set analysis of hydrogen-bond patterns in organic crystals, Acta Crystallogr. B46 (1990) 256–262.
[12] F.M. Richards, The interpretation of protein structures: total volume, group volume distributions and packing density, J. Mol. Biol. 82 (1) (1974) 1–14.
[13] K.P. Murphy, S.J. Gill, Calorimetric measurement of the enthalpy of dissolution of diketopiperazine in water as a function of temperature, Thermochim. Acta 139 (1989) 279–290.
[14] S.M. Habermann, K.P. Murphy, Energetics of hydrogen bonding in proteins: a model compound study, Protein Sci. 5 (7) (1996) 1229–1239.
[15] P.R. Connelly, J.A. Thomson, Heat capacity changes and hydrophobic interactions in the binding of FK506 and rapamycin to the FK506 binding protein, Proc. Natl. Acad. Sci. 89 (11) (1992) 4781–4785.
[16] R.S. Spolar, M.T. Record Jr., Coupling of local folding to site-specific binding of proteins to DNA, Science 263 (5148) (1994) 777–784.
[17] K.P. Murphy, P.L. Privalov, S.J. Gill, Common features of protein unfolding and dissolution of hydrophobic compounds, Science 247 (4942) (1990) 559–561 (2).
[18] P.L. Privalov, Thermodynamic problems of protein structure, Annu. Rev. Biophys. Chem. 18 (1989) 47–69.
[19] P.R. Connelly, R. Varadarajan, J.M. Sturtevant, F.M. Richards, Thermodynamics of protein–peptide interactions in the ribonuclease S system studied by titration calorimetry, Biochemistry 29 (25) (1990) 6108–6114.
[20] G.P. Brady, K.A. Sharp, Energetics of cyclic dipeptide crystal packing and solvation, Biophys. J. 72 (2 Pt 1) (1997) 913–927.
[21] D. Shivakumar, E. Harder, W. Damm, R. Friesner, W. Sherman, J. Chem. Theory Comput. 8 (2012) 2553.
[22] A. Klamt, G. Schüümann, COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient, J. Chem. Soc. Perkin Trans. 2 (1993) 799–805.

[23] SPICOLI version 1.1.3, OMEGA version 2.5.0, ZAP version 2.1.4., QUACPAC version 1.6.0, OpenEye Scientific Software, Inc., Santa Fe, NM, 2012.
[24] NBO 5.9., E.D. Glendening, J.K. Badenhoop, A.E. Reed, et al., Theoretical Chemistry Institute, University of Madison, Madison, WI, 2009. (http://www.chem.wisc.edu/~nbo5.5).
[25] J. Mestre-Ferrandiz, J. Sussex, A. Towse, The R&D Cost of a New Medicine, Office of Health Economics, UK, 2012.
[26] F. Maltais, Y.C. Jung, M. Chen, J. Tanoury, et al., In vitro and in vivo isotope effects with hepatitis C protease inhibitors: enhanced plasma exposure of deuterated telaprevir versus telaprevir in rats, J. Med. Chem. 52 (24) (2009) 7993–8001.
[27] Materials Studio Release Notes, Release 5.5, Accelrys Software Inc., San Diego, 2010.
[28] R.C. Rizzo, T. Aynechi, D.A. Case, I.D. Kuntz, Estimation of absolute free energies of hydration using continuum methods: accuracy of partial charge models and optimization of nonpolar contributions, J. Chem. Theory Comput. 2 (1) (2006) 128–139.
[29] J.W. Ponder, C. Wu, P. Ren, et al., Current status of the AMOEBA polarizable force field, J. Phys. Chem. B 114 (8) (2010) 2549–2564.
[30] B.T. Thole, Molecular polarizabilities calculated with a modified dipole interaction, Chem. Phys. 59 (3) (1981) 341–350.
[31] M.J. Frisch, G.W. Trucks, H.B. Schlegel, et al., Gaussian 09. Revision B.01, Gaussian, Inc., Wallingford CT, 2012.
[32] J.C. Wu, G. Chattree, P. Ren, Theor. Chem. Acc. 131 (2012) 1138 (Poltype available from: http://water.bme.utexas.edu/wiki/index.php/Software:Poltype).