Physicochemical Property Determinants of Oral Absorption for PROTAC Protein Degraders

Keith R. Hornberger* and Erika M. V. Araujo

Cite This: https://doi.org/10.1021/acs.jmedchem.3c00740

ABSTRACT: Heterobifunctional PROTAC degraders are gaining attention as a differentiated therapeutic modality with the potential for oral dosing in the clinic. Belonging to the beyond Rule of Five domain of physicochemical property space, we have sought to understand the determinants of oral absorption for this class of molecules for the rapid development of novel oral agents. We have collected a large data set from PROTAC molecules that have been dosed orally and intravenously in rats to estimate the fraction absorbed from oral dosing. Through this estimation, effects from differential hepatic clearance are normalized, allowing for a better assessment of the absorption. We demonstrate that rats are less permissive to PROTAC absorption than mice. The physicochemical properties of the molecules are then evaluated once compounds are rank-ordered by the fraction absorbed. We derive suggested design constraints on physicochemical properties for PROTAC molecules that are associated with higher probability of being orally absorbed.

INTRODUCTION

PROteolysis TArgeting Chimera (PROTAC) molecules are heterobifunctional compounds consisting of a protein targeting moiety (PTM) joined by a linker to a ubiquitin ligase moiety (ULM) (Scheme 1). They function by bringing a target protein in physical proximity to an ubiquitin ligase, leading to ubiquitin transfer and subsequent degradation of the target protein by the ubiquitin proteasome system. Over the past decade, there has been explosive growth in interrogating the potential of PROTAC protein degraders as new therapeutic agents, with interest from both academia and the pharmaceutical industry.1,2 Because PROTACs possess inherent functional activity and degrade a target protein rather than simply occupying a binding site, they are well suited for the treatment of diseases where achieving functional activity with small molecules is challenging (classically undruggable targets, proteins with scaffolding functions, or protein aggregates), where mechanisms of drug resistance are driven by over-expression or mutation of a target protein, or where isoform selectivity is desirable. Collectively, we refer to these use cases as the “Tenets of PROTAC targets”.1

Perhaps the greatest barrier to harnessing the potential of the PROTAC platform to deliver new drug molecules to patients is achieving suitable oral absorption. By their nature, PROTAC molecules occupy a space beyond that traditionally associated with oral absorption, as enumerated in the Rule of Five (Ro5)3,4 by Lipinski et al., or other analyses that have followed.3 More recently, there has been a renewed interest in understanding the behavior of molecules beyond the borders of Ro5 in extended Rule of Five (eRo5) or beyond Rule of Five (bRo5) space. A series of seminal publications by Kihlberg et al. in particular have laid out determinants of these spaces, where the criteria shift from “probable to be orally absorbed” to “possible to be orally absorbed”6–8

Received: April 25, 2023
hepatic clearance (10% $Q_h$). Because $F_{\text{max}}$ for compound A is only 10%, $f_a \times f_g = 0.50$, and this compound is relatively well absorbed despite its apparent low bioavailability. In contrast, for compound B, $F_{\text{max}}$ is 90%, $f_a \times f_g = 0.06$, and this compound is relatively poorly absorbed. In short, the readily calculated apparent value of $f_a \times f_g$ is a better measure of absorption than $F_{\text{obs}}$, because it appropriately factors in and normalizes for the hepatic clearance of the compound.

For our analysis of PROTAC molecules, we considered that to have desirable human pharmacokinetics, a compound should have a minimum oral bioavailability of 20% and low hepatic clearance (<25% $Q_h$), leading to a minimal $f_a \times f_g = 0.20/(1 - 0.25) \approx 0.26$. Therefore, we therefore set $f_a \times f_g = 0.25$ as a threshold of absorption for PROTACs to be minimally drug-like. It should be noted, however, that this cutoff is arbitrarily chosen based on practical experience, and there is little agreement on metrics constituting drug-likeness. A key assumption in our analysis is that the degree of absorption in preclinical species, specifically rats, will be predictive of human absorption. In contrast, it has been remarked upon many times in the literature that oral bioavailability in individual preclinical species (or aggregates thereof) is not well correlated with human bioavailability. This lack of correlation may be at least partially related to cross-species differences in hepatic clearance, which is better accounted for in the $f_a \times f_g$ term. However, species differences in $f_g$ driven by factors such as the complement of gut metabolic enzymes mean that $f_a \times f_g$ is an improved, but imperfect, measure of absorption.

### RESULTS

To analyze the Arvinas PK database, we selected rats as the preclinical species of interest because we have a large ($n = 1806$), diverse data set spanning multiple programs and chemotypes with the necessary measured intravenous and oral PK data. Rats are also a species that represents a well-established preclinical toxicology model for regulatory agencies. Attainment of exposure multiples above that required for efficacy in rats is thus a critical component for advancement to clinical trials. Furthermore, we have observed when comparing rodent species that rats are less permissive to oral absorption of PROTACs than mice (vide infra). Collectively, obtaining oral absorption in rats represents a higher barrier to compound progression than mice.

Although arguably similar analyses in higher species (e.g., dog or monkey) might be more useful or predictive of human outcomes, these data sets were more limited in scope and therefore deprioritized for analysis of physicochemical property trends. Furthermore, because compound progression to higher species PK studies is typically gated on PK performance in preclinical species, specifically rats, will be predictive of human absorption. In contrast, it has been remarked upon many times in the literature that oral bioavailability in individual preclinical species (or aggregates thereof) is not well correlated with human bioavailability. This lack of correlation may be at least partially related to cross-species differences in hepatic clearance, which is better accounted for in the $f_a \times f_g$ term. However, species differences in $f_g$ driven by factors such as the complement of gut metabolic enzymes mean that $f_a \times f_g$ is an improved, but imperfect, measure of absorption.

### Table 1. Hypothetical Compounds and Their PK Parameters, Illustrating Potential Disconnects between Oral Bioavailability and Fraction Absorbed

<table>
<thead>
<tr>
<th>Property</th>
<th>Compound A</th>
<th>Compound B</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI (%) $Q_h$</td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td>$F$ (%)</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>$f_a \times f_g$</td>
<td>0.50</td>
<td>0.06</td>
</tr>
<tr>
<td>PK diagnosis</td>
<td>high clearance</td>
<td>low absorption</td>
</tr>
</tbody>
</table>

To analyze the Arvinas PK database, we selected rats as the preclinical species of interest because we have a large ($n = 1806$), diverse data set spanning multiple programs and chemotypes with the necessary measured intravenous and oral PK data. Rats are also a species that represents a well-established preclinical toxicology model for regulatory agencies. Attainment of exposure multiples above that required for efficacy in rats is thus a critical component for advancement to clinical trials. Furthermore, we have observed when comparing rodent species that rats are less permissive to oral absorption of PROTACs than mice (vide infra). Collectively, obtaining oral absorption in rats represents a higher barrier to compound progression than mice.

Although arguably similar analyses in higher species (e.g., dog or monkey) might be more useful or predictive of human outcomes, these data sets were more limited in scope and therefore deprioritized for analysis of physicochemical property trends. Furthermore, because compound progression to higher species PK studies is typically gated on PK performance in rodents, there is an inherent selection bias in the higher species PK data set.

For curation, all 1806 PROTAC compounds with both rat iv and po PK data were considered and $f_a \times f_g$ calculated according to eq 3. This calculation assumes only hepatic clearance and ignores potential contributions from extrahepatic clearance mechanisms, which are difficult to determine systematically for large compound sets. Any apparent $f_a \times f_g$ values > 1 were set equal to 1 as the theoretical maximum. 75% of our compound collection displayed low or moderate clearance in rats (Figure S1), indicating that clearance is not a limiting factor in attaining oral bioavailability. A separate analysis of rat clearance vs physicochemical descriptors (Figure S2) did not identify any noteworthy
A total of 233 compounds (\(\sim 13\%\) of the data set, leaving \(n = 1573\) compounds in the final analysis) with rat CL \(\geq Q_h\) (55 mL/min/kg)\(^{17,18}\) were excluded because the theoretical \(F_{\text{max}} = 0\) if clearance is entirely hepatic, and thus, no meaningful assessment of \(f_a \times f_g\) can be made. Indeed, 79\% of the 233 high clearance compounds displayed \(F < 5\%\), as expected within a typical experimental error in PK studies; of the remainder, assessing \(f_a \times f_g\) is complicated by the potential contribution of extrahepatic clearance mechanisms. It should also be noted that the error in the apparent value of \(f_a \times f_g\) becomes large as CL approaches \(Q_h\) and \(F\) approaches zero, so such compounds should be treated with appropriate caution. Nonetheless, their inclusion in the analysis does not skew the overall interpretation. Where available (\(n = 1016\)), the same computations were done on the same compound set using mouse PK data using a standard \(Q_h\) value of 90 mL/min/kg.\(^{19}\)

Although a standard formulation was not used throughout the data set, our general practice is to dose solution formulations rather than suspensions, irrespective of the formulation chosen. Of the 1573 remaining compounds analyzed, there are 2395...
We observed upper limits where they are able to adapt their conformation, which is a composite of the like. 2102/2395 (88%) of those experiments were on the same compound assessing dose, schedule, vehicle, and the like. 2102/2395 (88%) of those experiments were conducted at a dose ≤ 30 mg/kg, reflecting normal “screening PK” conditions. For simplicity, compounds with >1 unique experiment (n = 352, 22% of the total) are represented in the figures by their average bioavailability. While this simplification may have the effect of depressing the reported success rates in attaining suitable \( f_a \times f_i \) particularly when dose escalations are included, it does not change the overall conclusions of the property analysis.

We began by comparing the calculated value of \( f_a \times f_i \) in rats to that in mice. As shown in Figure 1, \( f_a \times f_i \) values in rats are lower than those calculated from PK experiments in mice in 861/1016 (85%) of cases where comparator data in both rats and mice was available. All further analyses thus focused on rat data, as it is ultimately the most restrictive to compound advancement.

The results of the analysis are summarized in Figure 2. A clear cutoff was noted for HBD, with HBD ≤ 4 being an apparent upper limit for achieving \( f_a \times f_i \geq 0.25 \). Upon closer inspection of structures, however, nearly all (97%) of the three HBD and all the four HBD compounds with \( f_a \times f_i \geq 0.25 \) had one or two internally satisfied HBD, respectively. The cutoff is thus better stated as unsatisfied HBD ≤ 2. This observation potentially affords the chemist a greater degree of design flexibility with respect to HBD count, provided that an internal acceptor can be provided for additional HBD beyond the first two. A similar trend held for HBA, with ≈15 being an upper limit. Two compounds with HBA ≥ 15 achieved the threshold, but the limited sample size in this space (n = 13) may not be suitably powered to draw a strong conclusion. Previously, Veber et al. defined cutoffs for rat oral bioavailability based on RB (≤ 10) and TPSA (≤ 140 Å²). We observed upper limits of 14 RB and TPSA 200 Å² for \( f_a \times f_i \geq 0.25 \), which represent an expanded range vs the Veber rules. We found MW and c log P/D to have significantly expanded ranges vs the Ro5, with \( f_a \times f_i \geq 0.25 \) achievable up to a MW of 950, a log P range from 1 to 7, and a c log D range from 0 to 6. The upper limit on NAr was 5, although only a small fraction (9%) of compounds at NAr = 5 were well absorbed and a more realistic cutoff may be NAr = 4. Last, we calculated the previously reported AB-MPS score, which is a composite of c log D, NAr, and RB (Figure 3), and found the upper limit to be 20. The AB-MPS score was not more predictive of risk for low oral absorption than any of its individual composite parameters alone.

In general, we have observed that acceptable oral absorption in rats is possible in a window of HBA, RB, TPSA, MW, and c log P that is expanded vs traditional Ro5 guidance, but only MW, HBA, and c log P approach the bRo5 limit. This observation may be in part because these computed physicochemical descriptors are not completely independent variables, e.g., there is a positive correlation between MW and TPSA, MW and RB, etc. It has also been noted that there can be significant discrepancies between calculated and experimental log P/D values. Limits may be softer for properties like c log P and TPSA due to the chameleonic nature of larger molecules, where they are able to adapt their conformation, and therefore effective physicochemical properties, in response to their environment. In contrast, however, there are hard limits on HBD which hew closely to the Ro5 space. HBD have previously been noted as the main restriction to oral absorption even within the traditional Ro5. This latter constraint, therefore, becomes the most restrictive to compound design.

The practical consequences of these constraints can be seen when considering the additive physicochemical properties in the context of a PROTAC molecule. We calculated the properties for representative ULMs for cereblon and VHL (Figure 4), as well as the average properties for a linker derived from the 20 most represented linkers in our PROTAC collection. We then derive a “budget” remaining for a PTM for an orally bioavailable PROTAC degrader (Figure 5) and compare that budget to the properties of an average oral drug

---

**Table 2. Upper Limits of Physicochemical Properties for Oral Absorption for Ro5 + Veber, bRo5, and PROTACs**

<table>
<thead>
<tr>
<th>property</th>
<th>Ro5/ Veber</th>
<th>bRo5</th>
<th>PROTAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBD</td>
<td>5</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>HBA</td>
<td>10</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>RB</td>
<td>10</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>MW</td>
<td>500</td>
<td>1000</td>
<td>950</td>
</tr>
<tr>
<td>TPSA</td>
<td>140</td>
<td>250</td>
<td>200</td>
</tr>
<tr>
<td>NAr</td>
<td>ND</td>
<td>ND</td>
<td>5</td>
</tr>
<tr>
<td>c log P</td>
<td>5</td>
<td>7.5</td>
<td>7</td>
</tr>
<tr>
<td>c log D</td>
<td>ND</td>
<td>ND</td>
<td>6</td>
</tr>
</tbody>
</table>

---

**Figure 3.** Estimated fraction absorbed in rats as a function of AB-MPS score of the PROTAC molecules in the Arvinas database.

**Figure 4.** Representative ULMs used in physicochemical property budget analysis.
molecule (HBD 1.6, HBA 5.3, TPSA 74, RB 4.5, and MW 343).

**DISCUSSION AND CONCLUSIONS**

 Several conclusions are immediate from this analysis. First, only a judiciously chosen ULM (e.g., a lenalidomide-derived compound targeting CRBN) that is relatively compact will be compatible with rat oral absorption. Barring major changes, the opportunity for rat oral absorption sufficient to enable advancement toward preclinical development of PROTACs derived from larger ULMs, e.g., the prototypical VHL ligand described in Figure 4 and other reported ubiquitin ligase binders (IAP, MDM2, etc.), is limited.

Second, even with an efficient ULM, significant constraints are still in operation on the PTM. Particularly in the era of molecular property inflation, finding suitable PTM starting points with 1 unsatisfied HBD, 9 HBA, 120 TPSA, 10 RB, and 488 MW can be a challenging exercise. These guidelines should influence the selection of PTMs for oral PROTAC design. Furthermore, chemists working in the PROTAC space would be well served to commit part of their medicinal chemistry program to making PTM modifications to bring these parts of a PROTAC molecule in line with the properties needed for oral absorption. Historical PTMs may be suitable starting points, but in many cases, they need further modification. Our analysis also assumes an optimized, more druglike linker, and the derived property budget will be further shrunk with unoptimized PEG- or alkyl-type linkers.

Last, we have shown that PROTAC oral absorption in rats is generally more restricted than in mice. There is a dearth of published pharmacokinetic data for PROTACs in species other than mice, but mice may provide an incomplete picture of PROTAC absorption properties. Indeed, a few studies have reported useful levels of oral exposure in mice for challenging classes of PROTACs such as those based on recruitment of the VHL E3 ligase. Based on the relative permissiveness of mice toward oral absorption, such studies merit caution in their interpretation. Greater integration of pharmacokinetic data across multiple species is needed to support eventual clinical progression.

A potential limitation of this analysis is that the threshold for acceptable rat oral absorption—and by inference human oral absorption—is empirical. Oral absorption must ultimately be viewed through the lens of pharmacodynamics: “acceptable” is sufficient human in vivo free drug exposure at the site of action relative to potency on the target. Indeed, the original formulation of the Rule of 5 is derived in such fashion rather than prescribing an absorption or bioavailability floor. The lack of larger numbers of clinically advanced PROTAC agents precludes such an analysis for PROTACs, but the strong bias toward CRBN-derived heterobifunctional degraders in clinical trials and the alignment of our physicochemical property limits with a recent IQ consortium survey supports the trends we report here. It is possible that a sufficiently potent PROTAC could have clinical exposure suitable to drive efficacy despite having rat $f_a \times f_g < 0.25$. Nonetheless, such compounds may encounter other challenges to their preclinical and clinical progression, such as achieving efficacious exposure multiples in preclinical toxicology studies, which require dose escalation beyond the projected efficacious dose.

Looking to the future, several needs are apparent to expand the potential of PROTACs as a platform technology. First, there is a strong need for more screening work, both to identify new, compact ULMs, but also new PTMs. Indeed, an entirely new definition of a PTM may be required. Modern screening campaigns to discover PTMs that have functional activity, coupled with more challenging target space such as protein–protein interactions, have contributed significantly to molecular property inflation over the last two decades. An advantage of PROTAC degraders, however, is that the downstream
functional pharmacology is now event driven rather than occupancy driven, and because of this change, PTMs no longer need to have function—they simply need to bind their target, in some cases with only micromolar affinity. The heterobifunctional PROTAC mechanism engages the ubiquitin proteasome system and supplies the missing function. The built-in functionality of the PROTAC mechanism invites reinvestigation of old screening hits that were discarded because they were silent target binders that lacked functional activity. In addition, the physicochemical properties of compounds in a PTM screening library also merit reconsideration, given that PTMs derived from screening are now no longer the end point structures, but rather beginning points of a PROTAC medicinal chemistry campaign. Second, there is an ongoing need for more predictive in vitro ADME assays for PROTACs. It has been noted as a general trend across multiple companies, as well as at Arvinas, that such assays as configured for more traditional small molecules are often poorly predictive of in vivo PROTAC behavior. Third, greater attention to the potential of parenterally administered PROTACs is warranted. While the early clinical stage degraders have been biased toward oral administration, there may be cases where an intravenous or subcutaneous agent may be appropriate based on the target, its resynthesis rate, and the needs of the patient population.

The final and greatest need is for more detailed analysis of the determinants of successful oral absorption for prospective use. The present analysis is retrospective and furthermore is exclusive rather than inclusive. Our work lays out the boundaries of the PROTAC oral absorption space, beyond which the probability of failure is high. Comparing an early year of our data set to the most recent one, awareness of these limits coupled with other workflow enhancements has enriched our success in obtaining viable oral absorption in rats by ~6-fold. Further enrichment of this success rate is desirable for the continued growth of the oral PROTAC platform.

**EXPERIMENTAL SECTION**

ChemAxon JWS version 21.9 was used to calculate HBA, HBA, RB, NAr, MW, TPSA, c log P, and c log D. Data analysis and visualization were performed in D360 version 22.5.5 (Certara).

**ASSOCIATED CONTENT**

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jmedchem.3c00740.

Complete listing of experimental rat and mouse PK parameters, calculated physicochemical descriptors, and calculated $f_1 \times f_2$ for all compounds included in the analysis (CSV)

Rat clearance distribution and comparison of rat clearance to physicochemical descriptors (PDF)

**AUTHOR INFORMATION**

**Corresponding Author**

Keith R. Hornberger – Arvinas Operations, Inc., New Haven, Connecticut 06511, United States; orcid.org/0000-0002-6127-2814; Email: keith.hornberger@arvinas.com

**Author**


Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jmedchem.3c00740

**Author Contributions**

The manuscript was written through contributions of all authors. The authors have given approval to the final version of the manuscript.

**Notes**

The authors declare the following competing financial interest(s): Research was conducted at Arvinas Operations, Inc. where the authors are employees and shareholders.

**ACKNOWLEDGMENTS**

We thank Jordan Clark for assistance with the property calculations and Bradley Pearce, Song Yang, Larry Snyder, Andy Crew, and the Arvinas medicinal chemistry team for helpful discussions.

**ABBREVIATIONS**

$\log D$, calculated log octanol–water partition coefficient at pH 7.4; $\log P$, calculated log octanol–water partition coefficient; CRBN, cereblon; $f_1$, bioavailability; HBA, hydrogen bond acceptors; HBD, hydrogen bond donors; MW, molecular weight; NAr, number of aromatic rings; PK, pharmacokinetics; PTM, protein targeting moiety; RB, rotatable bonds; TPSA, topological polar surface area; ULM, ubiquitin ligase moiety; VHL, von Hippel–Lindau

**REFERENCES**


(13) AstraZeneca (ref 12) uses a 30% F cutoff but their analysis is based on mouse PK, vs our analysis based on rat PK. We have determined mouse to be a more permissive species than rat for PROTAC oral absorption (Figure 1).


