How We Responded to the Referees’ Comments

DexDesign: An OSPREY-based algorithm for designing de novo D-peptide inhibitors

Nathan Guerin, 1 Henry Childs, 3 Pei Zhou, 4 Bruce R. Donald 1,2,3,4,*

1Department of Computer Science, Duke University, Durham, NC 27708 USA
2Department of Mathematics, Duke University, Durham, NC 27708 USA
3Department of Chemistry, Duke University, Durham, NC 27708 USA
4Department of Biochemistry, Duke University School of Medicine, Durham, NC 22710 USA
*Correspondence: brd+peds23@cs.duke.edu

We’d like to thank the reviewers for their careful consideration, helpful comments, and feedback on our manuscript. This document contains our response to each referee comment, and highlights specific changes to the manuscript. For referee convenience, we have included a copy of DexDesign that highlights all revisions since the initial submission in red. Further, all references in this response, for example “1,4”, refer to sources cited in the bibliography at the end of this document.

Referee 1’s Comments:

“The authors report methodology and software to design D-peptide inhibitors of PDZ proteins, along with applications to two cases.

While the problem is important and a lot of work has been done, it is unfortunate that there is a basic error in the software and calculations, so that the results are incorrect.

Indeed, the authors try to produce D-amino acids by applying a mirror reflexion plane to an initial L-amino acid. Unfortunately, this is incorrect whenever the side chain lacks a symmetry plane, which is the case of Ile and Thr, which each contain an asymmetric carbon in their side chain. If one applies a mirror reflexion to an L-Thr amino acid, one does indeed obtain a D-amino acid, but it is not D-Thr. To be sure, the resulting side chain comes off the backbone in the desired D orientation. But the beta carbon has been switched from its normal S chirality to an unnatural R chirality: this is no longer a threonine side chain but an inverted version. The same problem occurs for Ile: its asymmetric beta carbon has lost its natural chirality. As a result, the software does not produce D-Thr or D-Ile, but two unnatural or "noncanonical" D-amino acids. Each of these makes interactions with its surroundings that differ from D-Thr and L-Ile. The authors perform peptide design, but not in the space they claim. Results obtained for any variants containing the mirrored Thr or Ile are not valid for D-Thr or D-Ile, but correspond instead to the D versions of two unnatural amino acids. Note that the software will also fail for the many noncanonical amino acids that have asymmetric carbons in their side chain, drastically reducing its generality...”

Answer: After much thought and consulting with the PEDS Editor-in-Chief, we believe there is a misunderstanding here, which we clarify in this RTR and the revised manuscript. We demonstrate
why our method produces D-Threonine and D-Isoleucine, as opposed to biologically uncommon allo- stereoisomers.

Our use of reflection is based on the mirror symmetry that enables Peter Kim’s experimental technique of Mirror-Image Phage Display\(^1\). Reflection has also previously been used by other research groups, such as Phillip Kim’s at the University of Toronto\(^2,3\), to design D-peptides. As shown in Supplementary Figure 1 (included below for convenience), starting with L-threonine, specifically \((2S,3R)-2\text{-amino}-3\text{-hydroxybutanoic acid}\), and reflecting across the \(x\)-\(y\) plane we obtain the correct enantiomer of L-threonine, namely D-threonine or \((2R,3S)-2\text{-amino}-3\text{-hydroxybutanoic acid}\). In Section 2.1 of the manuscript, we described the reflection function to invert chirality in the following way:

\[
\text{Let } r(s, a) \text{ be a function that reflects all atoms in protein structure } s \text{ across a plane } a. \\
\text{Without loss of generality, we let } a \text{ be the } x\text{-}y \text{ plane and define } r(s) = r(s, a) \text{ henceforth}...
\]

The reviewer’s claim that, upon reflection, “the beta carbon [of L-threonine] has been switched from its normal S chirality to an unnatural R chirality: this is no longer a threonine side chain but an inverted version…” is incorrect. As shown in Supplementary Table 1 (also included below for convenience), L-threonine has its \(\alpha\)-carbon in S configuration and \(\beta\)-carbon in R configuration. Analogously, D-threonine’s \(\alpha\)-carbon is in R configuration and \(\beta\)-carbon in S configuration. Supplementary Figure 1 demonstrates how reflection across a plane correctly flips the configuration of both stereogenic atoms of L-threonine, resulting in D-threonine.

<table>
<thead>
<tr>
<th>Stereoisomer</th>
<th>Configuration of Stereogenic Atoms</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-threonine</td>
<td>((2S,3R))</td>
<td>Pure and Applied Chemistry(^4)</td>
</tr>
<tr>
<td>D-threonine</td>
<td>((2R,3S))</td>
<td></td>
</tr>
<tr>
<td>D-allothreonine</td>
<td>((2R,3R))</td>
<td></td>
</tr>
<tr>
<td>L-isoleucine</td>
<td>((2S,3S))</td>
<td>PubChem(^5)</td>
</tr>
<tr>
<td>D-isoleucine</td>
<td>((2R,3R))</td>
<td>PubChem(^6)</td>
</tr>
<tr>
<td>D-alloisoleucine</td>
<td>((2R,3S))</td>
<td></td>
</tr>
</tbody>
</table>

Supplementary Table 1: The configurations of the stereogenic atoms of stereoisomers of threonine and isoleucine. Note that the configuration of both stereogenic atoms flips when converting from L to D, and vice versa.
Supplementary Figure 1: **Reflection across a plane converts L-peptides to their D-counterparts, and vice-versa.** Without loss of generality, here we reflect the L-tripeptide Ala-Thr-Ala (left) to its D-counterpart (right) across the x-y plane with the function $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ such that $f(x, y, z) = (x, y, -z)$. Particularly noteworthy is that the reflection operation, implemented previously by Garton et al. for generation of a D-peptide database, correctly converts even L-amino acids with more than one stereocenter, namely threonine and isoleucine, to their D-counterparts. The second residue in the L- and D-peptides above is threonine. The absolute configuration of L-threonine’s two carbon stereocenters is (2S,3R). In contrast, D-threonine’s two carbon stereocenters have an absolute configuration of (2R,3S). As shown in the figure and described above, reflecting L-threonine, specifically (2S,3R)-2-amino-3-hydroxybutanoic acid, across the x-y plane obtains the correct enantiomer, namely D-threonine or (2R,3S)-2-amino-3-hydroxybutanoic acid. Likewise, starting with L-isoleucine, specifically (2S,3S)-2-amino-3-methylpentanoic acid, and reflecting across the x-y plane obtains the correct enantiomer, namely D-isoleucine or (2R,3R)-2-amino-3-methylpentanoic acid. See Supplementary Table 1 for further specification of the stereogenic atom configurations.

When presented with the above arguments and new figure and table, PEDS editor in chief Roberto Chica stated:

*Dear Bruce,*

*Thank you for your email. I think your rebuttal is well written and convincing (specifically [Supplementary Table 1]). The reviewers’ comments suggest a misunderstanding on their part as to what the absolute stereocenter configurations are for the various Thr and Ile isomers. You should include this in your response to reviewers.*

*Sincerely,*

*Roberto*
We have attached the full letter and email exchange between the corresponding author, Bruce Donald, and Roberto Chica. We have also added Supplementary Figure 1 to the Supplementary Information Section D to illustrate the energy-equivariant reflection transformation and its application to peptides included in the DexDesign manuscript. We have also added Supplementary Table 1 to Supplementary Information Section C to specify the configuration of the stereogenic atoms. Additionally, we have attached to our response a PDB file containing CALP-DPR3 and MAST2-DPR3, two D-peptides generated with DexDesign. These contain D-Ile (2R, 3R) and D-Thr (2R, 3S), respectively.

“Calling a simulation an "experiment" (pg 13) is especially problematic in a paper that makes an enormous number of predictions but offers no experimental tests. Many journals consider this a fatal flaw in any computational protein design manuscript. At the leadst, the authors should make it a lot clearer that there are no experimental tests, something that is a bit glossed over in the current text.”

Answer: We appreciate this feedback, and have made adjustments to better reflect the type of data used as validation for the DexDesign algorithm. The word “experiment” in Section 4.2 has been changed to “computational experiment.” This terminology is standard in the computational structural biology community. Additionally, we have inserted the following sentence in manuscript Section 5: “Future work includes the important \textit{in vitro} experimental validation of the algorithm we have presented here.”

While DexDesign does not currently have experimental data for the proposed D-peptide redesigns, the OSPREY protein redesign software suite has previously used similar design approaches and rankings paradigms with considerable success. As reported\textsuperscript{8} by our lab in 2020, OSPREY K* scores analyzed using a Spearman rank correlation coefficient produced accurate retrospective and prospective predictions on the effect of 41 sets of mutations on the PPI of cRaf-RBD:KRas complex; KRas is a protein involved in cancers\textsuperscript{9}. OSPREY was also used to design a new variant of c-Raf-RBD, c-Raf-RBD(RKY), that bound roughly 5 times more tightly than the previous best known binder and 36 times more tightly than the wildtype sequence in a biolayer interferometry assay. Therefore, the positive results obtained by using a Spearman’s ranking paradigm for DexDesign bolsters these designs.

Lastly, we performed additional computational experiments and extensive data analysis with 13 new designs that address design questions from Referee 2 (below). This increased native sequence recovery to 44%, bolstering D-peptide designs. Recovery is a widely used metric for determining the quality of a computational tool (see RTR Table 1 below). These computational experiments are discussed in the revised manuscript Section 4.2.1 and visualized in Figure 10, with data in Supplementary Tables 6 and 7.

“Terminology on pg 1 is incorrect: "D" is not an abbreviation of dextrorotary, and "L" is not an abbreviation of levorotary. The corresponding abbreviations are "d" and "ld.
Indeed, not all D-amino acids are dextrorotary, same for L; it depends on the side chain; see any standard biochemistry textbook.”

**Answer:** We thank the referee for the opportunity to clarify our use of D/L notation for DexDesign. In our original manuscript, we adopted this language in its exact form from previous publications in the Public Library of Science and Proceedings of the National Academy of Sciences. However, based on further investigation in response to feedback from Referee 1, we now realize there is a misunderstanding of dextrorotatory/levorotatory in previous design work. Although previous authors claim that a reflection operation changes the direction of plane polarized light, no citation or experimental data is provided to support this claim. In fact, not all L-amino acids are levorotatory/- with regards to optical activity, and may instead exhibit dextrorotatory/+ clockwise plane polarization (ex. L (+)-alanine).

The misuse of stereoisomer nomenclature in past publications originates from conflating operational, systemic, and handedness naming systems for enantiomers. We understand that the operational labelling of molecules as (dextrorotatory/+) or (levorotatory/-) only relates to the rotation of the plane of polarized light with no direct indication of spatial arrangement. The systematic nomenclature (R, S) for mixed-chirality based design analyzes chiral carbons using the CIP priority convention. Lastly, our manuscript uses handedness notation (D, L) that evaluates the absolute configuration. Looking down the H-Cα bond, a molecule with CO, side R, and N in a clockwise manner is labeled as L (laevus, meaning left). A molecule from the same perspective with N, side chain R, and CO in a clockwise manner is labeled as D (dexter, meaning right). Nevertheless, we have removed all references to laevorotatory and dextrorotatory in the revised manuscript to prevent the propagation of ambiguous terminology. For clarity, we now note in Section 1 that DexDesign refers to DexterDesign, which indicates D-space (dexter-space, D-amino acids). In short, our manuscript follows the stereoisomer naming convention used in *Introduction to Protein Structure* by Carl Branden and John Tooze.

**Referee 2’s Comments:**

“**This article describes the development of a computational procedure for the design of D-peptide ligands for target proteins.** This method, called DexDesign, is implemented in the protein design software OSPREY. The method is cleverly implemented, and represents a potentially useful tool in the protein designer’s toolbox as it addresses an unmet need, namely the ability to design D-peptide binders. To demonstrate the utility of their method, the authors use it to design hypothetical D-peptide binders for two target proteins. They perform careful and insightful analyses of computational models and make comparisons with previously known L-peptide binders to showcase the improved binding interactions predicted for their D-peptide designs vs. the known L-peptide binders. However, there is no experimental validation of the designs, which is the biggest flaw of this study, since these types of computational designs cannot be confirmed without experimental validation. In the absence of this experimental validation, the authors perform a computational benchmark that aims to reproduce one known D-peptide/protein..."
interaction (GyGlanvdessG/streptavidin). Unfortunately, the results from this benchmark experiment are inconclusive, since the template D-peptide sequence GyGlanvdessG is not recovered (only 21% sequence similarity is achieved), likely due to important differences between the backbone conformations of the designed D-peptide and the crystallographic D-peptide, as pointed out by the authors…”

**Answer:** We thank the referee for the opportunity to discuss our confidence in our *de novo* D-peptide designs. As stated in our response to Referee 1’s comments, while DexDesign does not currently have experimental data for the proposed D-peptides redesigns, the OSPREY protein redesign software suite has previously used similar design approaches and rankings paradigms with significant success. As reported by our lab in 2020, OSPREY K* scores that use a Spearman rank correlation coefficient produced accurate retrospective and prospective predictions on the effect of 41 sets of mutations on the PPI of cRaf-RBD:KRas complex; KRas is a protein involved in cancers. OSPREY was also used to design a new variant of c-Raf-RBD, c-Raf-RBD(RKY), that bound roughly 5 times more tightly than the previous best known binder and 36 times more tightly than the wildtype sequence in a biolayer interferometry assay. We mentioned K*’s previous predictive accuracy and design successes because DexDesign uses the identical K*-based Spearman ranking paradigm to predict the *de novo* peptide inhibitors presented in this manuscript.

Furthermore, we would like to point out why high native sequence recovery is undesirable: a 25% to 50% native sequence recovery is considered excellent, and higher is considered overfitting. In the case of native sequence recovery, the community has converged on this desired percentage, which is empirically grounded, because computational predictions are being compared to experimental sequences. To back this claim, we have included RTR Table 1 below. This table is included in this response for referee convenience, and is not included in the revised manuscript. This table includes native sequence recovery rates for an array of computational tools, all of which generated experimentally successful designs.

<table>
<thead>
<tr>
<th>Method</th>
<th>Sequence Recovery (%)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rossetta Framework</td>
<td>22</td>
<td>Zhou et al. (PNAS)</td>
</tr>
<tr>
<td>ProteinMPNN</td>
<td>41.2 - 50.8</td>
<td>Dauparas et al. (Science)</td>
</tr>
<tr>
<td>ChromaDesign</td>
<td>35-80</td>
<td>Ingraham et al. (Nature)</td>
</tr>
<tr>
<td>dTERMen/TERMinator</td>
<td>24.32 - 42.22</td>
<td>Li et al. (arXiv)</td>
</tr>
</tbody>
</table>

RTR Table 1. **Native Sequence Recovery Rates for Various Computational Tools.** Method: indicates the tool or software suite for which native sequence recovery was analyzed. **Sequence Recovery (%):** the native sequence recovery as reported in the publication. **Source:** the publication where native sequence recovery was originally reported.

Therefore, DexDesign’s native sequence recovery of 21% for DRPV is comparable, albeit slightly underperforms, other tools in the computational protein design domain. However, using increased backbone sampling and remodeling as suggested by Referee 2, we have increased the native sequence recovery to 44%. This is reported in the revised manuscript (MS Section 4.2.1, Figure 10, Supplementary Tables 6 and 7). This is a positive result, since improved backbone sampling (as suggested by the referee) was expected to improve the model, and therefore the
prediction accuracy, which is consistent with our new results. Discussion and additional data for native sequence recovery and backbone flexibility is continued in the next answer below, where additional computational experiments improved native sequence recovery and K* scores.

“Overall, I cannot recommend publication of this manuscript in its current form because it lacks experimental validation and robust benchmarks to demonstrate the utility of the method (which is expected for these types of computational studies). Nevertheless, I believe that with additional computational benchmarks to robustly validate the method, this article has the potential to be of high interest to readers of PEDS. Therefore, I encourage the authors to perform benchmarks on a higher number of D-peptide/L-protein pairs available on the PDB, possibly with the following modifications to improve prediction accuracy:

1. Use more than one MASTER fragment as template for sequence design. Since the difference in sequence between computational design and benchmark peptide is likely caused by differences in backbone conformations, it may help to try other backbones with similar RMSDs.
2. Remodel the MASTER fragment backbone conformation prior to using for sequence design (but without introducing explicit biases towards the correct conformation). This may help to make it more similar to the benchmark template and thereby improve sequence recovery. Several methods exist to remodel backbone conformations (e.g., backrub, MD, kinematic closure, etc.).”
3. Allow more than one sequence to be produced by the search algorithm. Some lower-ranked but still favorable sequences may be more similar to the template sequence.”
4. Compare the computational model and crystal structure of the GyGlanvdessG/streptavidin complex (and other D-peptide/protein pairs used for benchmark) using structural metrics such as sidechain rmsd, percentage of peptide residues adopting the correct rotamer, recapitulation of important H-bonds and other interactions with protein target, etc. A structural comparison may provide complementary information to the sequence recovery analysis.”

**Answer:** We thank Referee 2 for their insightful comments concerning native sequence recovery and backbone flexibility. In response, we have performed additional computational experiments and extensive data analysis with 13 new designs that address suggestions 1-4. We have simplified the report of this analysis into two computational experiments: Experiments 1 and 2. As with the original DexDesign algorithm, computational Experiments 1 and 2 permit ligand translation and rotation when calculating structural molecular ensembles, partition functions, and binding affinity.

Experiment 1 incorporates additional sampling of MASTER-returned backbones (suggestions 1 and 3). Specifically, 10 substructures were used as a template for redesign. This computational experiment is designed to explore a larger backbone conformational space by performing DexDesign on sequences with backbone alignment RMSD ranging from 0.50-0.70 Å
(as returned by MASTER). This method produced the best improvements in native sequence recovery and predicted binding affinity.

Experiment 2 includes backbone minimization (via molecular dynamics\textsuperscript{20}) before design (suggestion 2). In this design strategy, substructures are minimized before any design techniques (Minimum Flexible Set, Inverse Alanine Scanning, K*-based Mutational Scanning) are performed. Because downstream design decisions are sensitive to MASTER-returned backbones (manuscript Section 4.2), remodeling fragments to obtain more chemically realistic energies before design improved native sequence recovery and predicted binding affinity.

Both Experiment 1 and Experiment 2 include additional analysis, as detailed in Supplementary Tables 6 and 7 (included below for convenience). A new visualization of this data, Figure 10, is also included. These analyses address the referee’s 4\textsuperscript{th} suggestion by providing a comprehensive structure and mutation selection comparison.

Sampling of additional backbone conformations and remodeling before design resulted in improved K* (log\textsubscript{10}) scoring and native sequence recovery (sequence similarity). A discussion of the methodology and results for these computational experiments is now detailed in a newly added manuscript Section 4.2.1, and is also summarized below.
<table>
<thead>
<tr>
<th>Match</th>
<th>Name</th>
<th>Original Sequence</th>
<th>bbRMSD (pre)</th>
<th>bbRMSD (post)</th>
<th>Source PDB</th>
<th>Final Sequence</th>
<th>Sequence Similarity (%)</th>
<th>K* Score (log_{10})</th>
<th>Inverse Alanine Scan Native Residue Rank</th>
<th>K*-based Mutational Scan Native Residue Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>YEWLNDDV</td>
<td>0.58</td>
<td>1.57</td>
<td>7NPA</td>
<td>WRAMLENIY</td>
<td>0.00</td>
<td>35.19</td>
<td>1</td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>10</td>
<td>YCDYALCSS</td>
<td>0.52</td>
<td>0.88</td>
<td>1T6V</td>
<td>DMANVYYSE</td>
<td>36.36</td>
<td>36.06</td>
<td>16</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td>11</td>
<td>FKOSODYA</td>
<td>0.52</td>
<td>0.31</td>
<td>3D89</td>
<td>HIRIDYSE</td>
<td>44.44</td>
<td>38.88</td>
<td>18</td>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td>12</td>
<td>TVDPYRVA</td>
<td>0.52</td>
<td>1.38</td>
<td>1ZSV</td>
<td>HNYVEENIE</td>
<td>27.27</td>
<td>29.68</td>
<td>1</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>16</td>
<td>SUPPVVIA</td>
<td>0.53</td>
<td>1.69</td>
<td>1V3U</td>
<td>WDELFEYIE</td>
<td>25.00</td>
<td>29.29</td>
<td>1</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>31</td>
<td>KSRPDKQN</td>
<td>0.56</td>
<td>0.69</td>
<td>5GZU</td>
<td>WERVDDQIE</td>
<td>40.00</td>
<td>31.59</td>
<td>8</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>70</td>
<td>CTPITDIE</td>
<td>0.59</td>
<td>0.39</td>
<td>6NUB</td>
<td>WHELDIEW</td>
<td>30.00</td>
<td>32.23</td>
<td>19</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td>500</td>
<td>TVDPYRVA</td>
<td>0.61</td>
<td>1.54</td>
<td>2Y9S</td>
<td>MVYKEEIL</td>
<td>5.88</td>
<td>27.64</td>
<td>1</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>295</td>
<td>TVDPYRVA</td>
<td>0.66</td>
<td>0.25</td>
<td>2Y9S</td>
<td>FWMNYDNNW</td>
<td>33.33</td>
<td>27.02</td>
<td>1</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>Mean</td>
<td>N/A</td>
<td>0.57</td>
<td>0.92</td>
<td>Consensus Sequence:</td>
<td>27.48</td>
<td>32.78</td>
<td>6.9</td>
<td>6.7</td>
<td>12.2</td>
<td>18.2</td>
</tr>
<tr>
<td>Median</td>
<td>N/A</td>
<td>0.54</td>
<td>0.78</td>
<td>MWE[R/L/K][L/I/M]/D/[W/Y]/IE</td>
<td>31.67</td>
<td>31.91</td>
<td>2</td>
<td>6.5</td>
<td>13.5</td>
<td>18.5</td>
</tr>
</tbody>
</table>

Supplementary Table 6. Experiment 1: Analysis of native sequence recovery and backbone flexibility for 10 MASTER-returned substructures with additional sampling. Name: the match number of the backbone as returned by MASTER. A lower match number correlates to a lower backbone RMSD (Å). See Supplementary Figure 2 for more information about the composition of the MASTER database. Original Sequence: the substructure sequence before design. bbRMSD (pre): the pre-design alignment RMSD (Å) as returned by MASTER. bbRMSD (post): the post-design backbone alignment RMSD (Å) of the D-peptide to the original query of GyGlanvdessG (PDB ID: 5n8j, see Section 4.2 for more information on the substructure search). All alignments were calculated using PyMOL\(^{31}\). Source PDB: the PDB file from which the backbone segment was sourced. Final Sequence: the sequence of the match after implementation of DexDesign. Sequence Similarity (%): the calculated sequence similarity between the matches’ final sequence and the GyGlanvdessG wildtype sequence. This measures the native sequence recovery. All calculations were performed using VectorBuilder’s sequence alignment tool\(^{32}\). K* Score (log_{10}): The predicted binding affinity of the designed D-peptide. This value is calculated using the K* algorithm (see Supplementary Section A). Inverse Alanine Scan Native Residue Rank: Metric for determining recovery of individual residues relative to GyGlanvdessG using Inverse Alanine Scanning (see Section 2.3.1.2). This metric is formatted for a single sequence (one match), where each entry indicates the native residue predicted affinity for each of the nine positions among the 20 D-amino acids. E.g., an entry of 3 indicates that the native residue was ranked as the third highest predicted binding affinity residue. An entry of “None” indicates that no K* (log_{10}) score was produced. K*-based Mutational Scan Native Residue Rank: Metric for determining recovery of individual residues relative to GyGlanvdessG using K*-based Mutational Scanning (see Section 2.3.1.3). This metric is formatted in the same manner as Inverse Alanine Scan Native Residue Rank.

http://www.peds.oupjournals.org
Supplementary Table 7. **Experiment 2: Analysis of native sequence recovery and backbone flexibility for 3 MASTER-returned substructures with pre-design minimization.** **Name:** the match number of the backbone as returned by MASTER. A lower match number correlates to a lower backbone RMSD (Å). See Supplementary Figure 2 for more information about the composition of the MASTER database. **Original Sequence:** the substructure sequence before design. **bbRMSD (pre):** the pre-design alignment RMSD (Å) as returned by MASTER. **bbRMSD (post):** the post-design backbone alignment RMSD (Å) of the D-peptide to the original query of GyGlanvdessG (PDB ID: 5n8j, see Section 4.2 for more information on the substructure search). All alignments were calculated using PyMOL. **Source PDB:** the PDB file from which the backbone segment was sourced. **Final Sequence:** the sequence of the match after implementation of DexDesign. **Sequence Similarity (%):** the calculated sequence similarity between the matches’ final sequence and the GyGlanvdessG wildtype sequence. This measures the native sequence recovery. All calculations were performed using VectorBuilder’s sequence alignment tool. **K* Score (log10):** The predicted binding affinity of the designed D-peptide. This value is calculated using the K* algorithm (see Supplementary Section A). **Inverse Alanine Scan Native Residue Rank:** Metric for determining recovery of individual residues relative to GyGlanvdessG using Inverse Alanine Scanning (see Section 2.3.1.2). This metric is formatted for a single sequence (one match), where each entry indicates the native residue predicted affinity for each of the nine positions among the 20 D-amino acids. E.g., an entry of 3 indicates that the native residue was ranked as the third highest predicted binding affinity residue. **K*-based Mutational Scan Native Residue Rank:** Metric for determining recovery of individual residues relative to GyGlanvdessG using K*-based Mutational Scanning (see Section 2.3.1.3). This metric is formatted in the same manner as Inverse Alanine Scan Native Residue Rank.

<table>
<thead>
<tr>
<th>Name</th>
<th>Original Sequence</th>
<th>bbRMSD (pre)</th>
<th>bbRMSD (post)</th>
<th>Source PDB</th>
<th>Final Sequence</th>
<th>Sequence Similarity (%)</th>
<th>K* Score (log10)</th>
<th>Inverse Alanine Scan Native Residue Rank</th>
<th>K*-based Mutational Scan Native Residue Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match 1</td>
<td>RYEGLFNN</td>
<td>0.48</td>
<td>0.55</td>
<td>3HOE</td>
<td>WIEYMEFND</td>
<td>21.43</td>
<td>35.13</td>
<td>1 9 10 9 4 2 16 2 13</td>
<td>19 7 11 12 7 2 10 7 13</td>
</tr>
<tr>
<td>Match 18</td>
<td>YCYVALC5S</td>
<td>0.52</td>
<td>1.72</td>
<td>1T6V</td>
<td>MMLWIECSQ</td>
<td>36.36</td>
<td>33.74</td>
<td>17 4 6 9 3 4 13 3 17</td>
<td>17 2 4 9 4 5 17 2 16</td>
</tr>
<tr>
<td>Match 295</td>
<td>TVDFYPMBVA</td>
<td>0.66</td>
<td>0.29</td>
<td>2Y05</td>
<td>QYMMRMRDIE</td>
<td>30.77</td>
<td>29.22</td>
<td>1 8 14 10 8 3 12 2 18</td>
<td>17 5 8 10 12 4 12 3 16</td>
</tr>
</tbody>
</table>

Mean N/A 0.56 0.85
Median N/A 0.52 0.55

Consensus Sequence: ([W/M/Q][I/L/Y]M[Y/W/R]ME[F/C/R][N/S/I][D/Q/E])
Figure 10. **Quantitative analysis of D-peptides with additional backbone sampling and remodeling.** Blue bars note the predicted binding affinity (K* score (log$_{10}$), see Supplementary Information A) while green bars note the native and DexDesign-predicted sequences’ similarity. The numbers on the y-axis denote both K* score (log$_{10}$) and sequence similarity (%). A horizontal red dashed line shows comparison of native sequence recovery to DRPV (Section 4.2). Experiment 1 and 2 data are separated by the vertical dashed black line. See Supplementary Information Tables 6 and 7 for the full dataset. As reported in Section 4.2.1, two in silico experiments were performed to evaluate DexDesign native sequence recovery and predicted binding affinity using additional backbone sampling and remodeling. These experiments improved native sequence recovery results for DPRV and GyGlanvdessG over the model using a single backbone conformation (Section 4.2). Sampling of additional backbone conformations (Experiment 1) resulted in the greatest increase in native sequence recovery from 21.43% (DPRV) to 44.44% (Match 11). This computational experiment also resulted in the greatest improvement in predicted binding affinity, with an improvement in K* score (log$_{10}$) from 32.8 (DPRV) to 39.54 (Match 500). Experiment 2 indicates the benefit of incorporating diverse scaffolds from homology modeling into DexDesign for native sequence recovery, though Match 1 suggests a potential trade-off between sequence recovery and predicted binding affinity. The median sequence similarity and K* score (log$_{10}$) for Experiment 1 were 31.7 and 31.9, respectively. The median sequence similarity and K* score (log$_{10}$) for Experiment 2 were 30.8 and 33.7, respectively. Overall, sampling of additional backbone substructures (Experiment 1) and remodeling of ligand backbone structure (Experiment 2) resulted in an almost universal increase in native sequence recovery and predicted binding affinity compared to the single backbone conformation DPRV native sequence recovery experiments (Section 4.2).
As included in Supplementary Table 6, sampling of additional backbone conformations results in improved K* \((\log_{10})\) scoring and native sequence recovery (sequence similarity). As discussed in our previous answer on native sequence recovery rates for DPRV, and as outlined in RTR Table 1, ideal native sequence recovery ranges from 25-50%. Of the 10 system designed in Experiment 1, only 2 peptides, Match 2 and Match 100, have sequence similarity outside of this range. Match 11 resulted in the highest sequence similarity of 44.44%, and the 2\textsuperscript{nd}-best K* score \((\log_{10})\) of 38.8. Match 500 reports the best K* score \((\log_{10})\) of 39.54 with the 3\textsuperscript{rd}-highest sequence similarity of 36.36%. These new designs demonstrate improvement over the DPRV peptide reported in the manuscript Section 4.2, which had a sequence similar of 21.43% and a K* score \((\log_{10})\) score of 32.8. Four of the ten designs have a predicted binding affinity greater than DRPV. Five systems have a predicted binding affinity greater than the wildtype D-peptide GyGlanvdessG, which has a K* score \((\log_{10})\) of 32.2. These results demonstrate the benefit of additional backbone sampling, outline valuable future steps to improve D-peptide design, and reports data that implement referee suggestions 1 and 3. This data table has been added to Supplementary Information Section C.

Supplementary Table 7 includes data pertaining to remodeling of MASTER-returned backbones before design without inclusion of explicit bias. Interestingly, minimization\(^{20}\) of Match 1 (the same match used to generate DPRV) before design did not change the sequence similarity of 21.43%, but did improve the K* score \((\log_{10})\) from 32.8 to 35.13. The conformational and energy landscape of the sampled backbones are improved, but perhaps do not optimize for wildtype recovery due to the presence of different D-space amino acids with higher affinity (an observation discussed in manuscript Section 4.2, see ST0929 scaffold). Minimization of Match 10 and 295 increased native sequence recovery by 14.93% and 9.34%, respectively, at the cost of worse K* scores \((\log_{10})\). Therefore, while predicted binding affinity is slightly worse, the native sequence recovery is comparable to computational tools with successful experimental results (see RTR Table 1). This data reports the benefit of incorporating improved backbone sampling and diverse scaffolds from homology modeling into DexDesign for native sequence recovery with a potential trade-off between sequence recovery and predicted binding affinity. This computational experiment reports data that implements referee suggestion 2. This data table has been added to Supplementary Information Section C.

The columns of Inverse Alanine Scan Native Residue Rank and K*-based Mutational Scan Native Residue Rank indicate the propensity of the K* algorithm (Supplementary Section A) and DexDesign design techniques (MS Section 2.3.1) to select wildtype residues at each of the nine residue positions of MASTER-returned backbones. Residues with a rank of “None” commonly indicate areas of significant steric clashes stemming from inadequate flexibility, and are an area of desirable improvement for later DexDesign iterations. For Experiment 1, the most conserved residues (by median) were at positions 1, 2, 5, 6, and 8. At these positions, selection of the GyGlanvdessG wildtype residue were predicted among the top 32.5% of possible selections or better. For Experiment 2, the most conserved residues (by median) were at positions 1, 5, 6, and 8. At these positions, selection of the wildtype residue were predicted among the top 20% of possible selections or better. Most notably, conserved regions between Experiment 1 and Experiment 2 capture most nonpolar, hydrophobic residues (GyGlanvdessG residue positions 1-3,
5), which are known to be highly conserved among high-affinity binders with intended folds\textsuperscript{23}. Furthermore, wet labs rarely test only the GMEC (best design based on energy). Therefore, computational methods that demonstrate selection within roughly the top third, or better, of highly conserved sequence motifs are a reasonable and effective filter for experimental screenings. These columns provide an analysis of the selection of D-amino acids relative to individual wildtype residues, effectively implementing additional structural metrics (referee suggestion 4).

Overall, sampling of additional backbone substructures (Experiment 1) and remodeling of ligands (Experiment 2) resulted in an almost universal increase in native sequence recovery and predicted binding affinity. As reported in manuscript Section 4.2, the native sequence recovery (sequence similarity) and K* (log\textsubscript{10}) score of DRPV without additional backbone sampling and remodeling was 21.43% and 32.8, respectively. The median K* (log\textsubscript{10}) score for Experiment 1 was 31.91 and for Experiment 2 was 33.74. The best native sequence recovery was seen in Experiment 1, Match 11 with an increase from 21.43% to 44.44%. The best predicted binding affinity resulted from Experiment 1, Match 500, with an improved K* (log\textsubscript{10}) score from 32.8 to 39.54. This addresses referee suggestions 1, 2, and 3. Additionally, evaluation of D-amino acid selections relative to design techniques (Inverse Alanine Scanning and K*-based Mutational Scanning) indicates that DexDesign maintains highly-conserved residues while acting as an effective pruning filter for experimental screenings. This addresses referee suggestion 4.

This new data is now included in our revised manuscript Section 4.2.1, Figure 10, and Supplementary Tables 6 and 7 (Supplementary Information Section C). Furthermore, because the provided feedback was highly effective for improving D-space designs, we have included a note thanking Reviewer 2 in the manuscript Acknowledgements.

“Please remove the word “new” when describing the method (title of article, abstract, introduction, section 2.2, section 2.2.1 and section 2.2.2 titles, Conclusions, Figure 3 title). It is understood that the method is new since the authors are reporting it for the first time. Yet, it will not be new a few years from now so it is best to avoid these superfluous words in scientific writing (ie their removal will not change the quality of the work).”

**Answer:** We thank the referee for their foresight, and have removed or altered the word “new” from the Title, Abstract, Background and Introduction (Sections 1.1-1.4), Section 2.2 (including 2.2.1 and 2.2.2), Section 2.3.1, Conclusions, and Figure 3. In Section 2.2, “new features in OSPREY” was changed to “Capabilities added to OSPREY” to distinguish the DexDesign specification options from coexisting OSPREY 3.0 (command line\textsuperscript{24-26} protocols. Section 2.2.1 was changed from “New feature: customize existing or add new conformation libraries” to “Customize existing or add new conformation libraries.” We retain the word “new” in this one case, because in the future, a user of our software could add new libraries at will. Section 2.2.2 was changed from “New feature: D-protein/peptide design” to “An algorithm for D-protein/peptide design.” In all other requested sections, the word “new” was removed.
In Section 1.3, please remove the text “(also reviewed in Donald 2011, Chapter 9)” and replace by a reference.

**Answer:** We have removed the requested phrase, leaving only a reference to Donald 2011.

Please remove the summary at the end of Section 1.4. This text is redundant.

**Answer:** We have confidence in the clarity of Section 1 to define the novel contributions of DexDesign, and have removed the summary of contributions from Section 1.4.

Section 2.3.1.3: The authors mention that they systematically mutate residues in the DPR scaffold to all 20 amino acids. Yet, one of these is not a mutation. It would be preferable to simply state that authors sample rotamers for all 20 amino acids at those positions (ie avoid the term mutation).

**Answer:** We thank the referee for their attention to detail, and have changed “A Mutational Scan uses K* to systematically mutate a residue in the DPR scaffold to all 20 amino acids” to “A Mutational Scan uses K* to systematically sample rotamers for all 20 amino acids at each DPR residue” in order to avoid the word “mutation.”

In section 4.1, the authors write “whereas in some cases we achieved improved binding”. This is not demonstrated, as experiments to evaluate binding affinity have not been performed. Please reword to clarify that these are computational predictions.

**Answer:** We agree that this language should be clarified to reflect that binding affinity is computational, as discussed in Supplementary Section A. We have removed “… achieved improved binding through…” in favor of “… achieved predicted improved binding through…”.

We anticipate the inclusion of the word “predicted” sufficiently clarifies that binding affinity predictions are computational.

In section 3.1, bullet point #2, the authors write that “we believe these facets to be sufficient, but not necessary”. This statement is confusing. Can the authors clarify?

**Answer:** While the interested reader is invited to view Section 4 on this point in the main text, we agree that this statement may cause confusion in isolation. We anticipate the problem may originate from the term “sufficient” when discussing predicted binding affinity fluctuations between restituted and replicated biophysical facets. We have changed the statement to “These facets have proven sufficient for the systems presented in this manuscript, however, there may exist D-peptide inhibitors that initiate novel interactions not following this paradigm.” We believe this change better reflects the dichotomy of replication and restitution in D-space design while remaining

http://www.peds.oupjournals.org
appropriately brief. Additionally, we have changed “see Section 4” to “see Section 4.1” in order to directly point the reader to the supporting information.

Section 3.2 (page 10, lines 23-27). The first sentence of this paragraph is redundant and should be removed. (“The three canonical PDZ domain binding motifs we used as criteria to further validate the CALP-PEPs were: 1) the presence of an H-bond network between the peptide's C-terminal carboxylate and the CBL; 2) the presence of peptide:β2 backbone interactions; and 3) whether the D-peptide filled the hydrophobic pocket typically filled…”)

**Answer:** We have removed the requested sentence and inserted a parenthetical statement both abbreviating the validation criteria and directing the reader to Section 3.1. Paragraph three of Section 3.2 now begins: “To quantify validation criteria (1) (the presence of H-bond network, see Section 3.1 bullet point 2)…” Similar statements were inserted when criteria (2) and (3) are referenced later in the paragraph.

Figure 1, panel d: Please show only backbone atoms (or cartoon) of MASTER fragments, for clarity (it is difficult to see with all the sidechains in the way).

**Answer:** We have revised Figure 1, panel D to only include backbones atoms for the multicolored wire representation of MASTER-return backbone fragments. We anticipate that this adjustment sufficiently provides clarity.

Figure 1, panel f: The image is clearly rotated in addition to being a mirror image. Yet, only the fact that the mirror image is shown is indicated by the arrow. Please clarify.

**Answer:** We thank the referee for their attention to detail. While panel F correctly indicates a mirror reflection, the reference frame positions the L-form protein slightly more into the page than panel E. We have corrected this, and anticipate the new panel clarifies that only a reflection operation occurs.

Figure 5: Please move x-axis label underneath x-axis instead of on the top of graph (where it gives the impression that it is the graph’s title).

**Answer:** We have made the recommended change, and the label for Figure 5 is now underneath the x-axis.

The DEXdesign method is based on generating the mirror image of amino acids to obtain side-chain rotamers for the D-enantiomer. While this is OK for most amino acids, it will not work for amino acids whose side chains have stereocenters, such as the beta-Carbons
of Thr and Ile. In those cases, the authors would obtain D-allo-threonine and D-allo-isoleucine instead of D-threonine and D-isoleucine. It would be good for authors to fix this issue (or at least acknowledge it).

**Answer:** We thank the referee for this feedback, and appreciate the opportunity to specify why our reflection method produces D-Ile (2R, 3R) and D-Thr (2R, 3S), as opposed to allo- stereoisomers. Referee 1 also provided this feedback; please refer to this answer (beginning on page 1) for an extensive explanation of our methodology and confirmation of correctness by PEDS editor in chief Roberto Chica.

Suggestion: The introduction would benefit from some streamlining. A lot of words are dedicated to inform us of what will be described later and to provide lots of background that is not essential to understand the purpose of the method. I believe the introduction could be made more concise without losing any of the crucial information for understanding the motivations, choice of targets and methodological approach.

**Answer:** We thank the referee for their feedback, and have made revisions throughout the Background and Introduction (Section 1). This section has been reduced from 1,249 words to 998 words, without loss of accuracy. Specifically, the following revisions were made:

**Section 1:**

**Original (147 words):**

“In Section 1.1, we describe the benefits of incorporating D-amino acids into therapeutic peptides. Section 1.2 provides background on PDZ domains in general and two PDZ domains in particular that researchers have investigated targeting with L-peptides for biomedical purposes. Section 1.3 describes previous computational protein redesign software and algorithms for designing proteins and peptides incorporating non-canonical and D-amino acids. Section 1.4 concludes with a summary of DexDesign, a new algorithm we developed and incorporated into the protein design software OSPREY.

With the necessary background covered, the rest of this article focuses on an application of the DexDesign algorithm to generate de novo D-peptide inhibitors of two biomedically important PDZ domains targets: CAL and MAST2. We then evaluate each computationally generated peptide using multiple structural criteria, including predicted binding affinity and whether a D-peptide mimics binding interactions previously shown to be important to L-peptide binding to PDZ domains.”

**Revised (68 words):**

“The substitution of D-amino acids for L-amino acids in peptides is one strategy medicinal chemists have used to address these shortcomings. The following details the background, benefits of D-peptides, previous work in noncanonical design, and contributions of DexDesign, an algorithm we developed and incorporated into the protein design software OSPREY to generate
and analyze de novo D-peptide inhibitors of two biomedically important PDZ domains targets: CAL and MAST2.”

Section 1.2.1

Original (48 words):

“Cystic fibrosis can cause serious pulmonary and respiratory problems in the lungs by causing the development of a thick mucus that promotes bacterial infection and inflammation. It is caused by many mutations in the cystic fibrosis transmembrane conductance regulator (CFTR), such as ΔF508, which causes destabilized, misfolded CFTR.”

Revised (19 words):

“Mutations in the cystic fibrosis transmembrane conductance regulator (CFTR), such as ΔF508, cause destabilized, misfolded, and less efficient CFTR.”

Original (removed, 19 words):

“A number of research groups have developed peptide stabilizers that bind to CALP, preventing CFTR lysosomal trafficking and degradation.”

Original (45 words):

“Cyclic peptides targeting CALP have also been developed—Dougherty et al. developed a highly selective stabilizing cyclic peptide that binds CALP with a KD of 6 nM. Competitive peptide inhibitors have also been developed with the goal of developing new methods of managing neurological disease.”

Revised (6 words):

“Cyclic peptides have also been developed.”

Section 1.2.2

Original (61 words):

“In a neuron, phosphatase and tensin homolog deleted on chromosome 10 (PTEN)’s SLiM interacts with the PDZ domain of microtubule-associated serine-threonine kinase 2 (MAST2) to regulate pathways inhibiting neuronal survival, regrowth, and regeneration. The rabies virus glycoprotein’s C-terminal residues interact with MAST2’s PDZ domain, disrupting the ability of MAST2 and PTEN to form a complex and inhibit neurite outgrowth and apoptosis.”

Revised (28 words):

“The rabies virus glycoprotein’s C-terminal residues interact with MAST2’s PDZ domain, disrupting the ability of MAST2 and PTEN to form a complex and inhibits neurite outgrowth and apoptosis.”

Original (removed, 14 words):
“In Section 3, we present 30 de novo D-peptide inhibitors targeting CALP and MAST2.”

Section 1.3

Original (52 words):

“OSPREY has been used to, among other things, predict resistance mutations ablating efficacy of antibiotics used to treat methicillin-resistant Staphylococcus Aureus and small molecule inhibitors used to treat melanoma, lung, stomach and colorectal cancers, design and structurally characterize peptide inhibitors of CALP for treating CFTR, and improve broadly neutralizing antibodies against HIV-1.”

Revised (34 words):

“OSPREY has been used to, among many other things, design and structurally characterize peptide inhibitors of CALP for treating CFTR, predict resistance mutations in bacteria and cancer, and improve broadly neutralizing antibodies against HIV-1.”

Section 1.4

Original (133 words):

“Given the biomedical importance of modulating CALP and MAST2 PDZ domain interactions, coupled with the advantages of D-peptide therapeutics, we use DexDesign to predict D-peptides inhibitors of these two protein targets.

In summary, this paper makes the following contributions:

1. A new computational protocol, DexDesign, for designing de novo D-peptide binders,
2. Three novel design techniques leveraging continuous flexibility and the computation of \( \varepsilon \)-accurate ratios of partition functions over molecular ensembles: the Minimum Flexible Set, Inverse Alanine Scan, and \( K^* \)-based Mutational Scan,
3. Application of DexDesign to predict D-peptide binders to the PDZ domains of CALP and MAST2,
4. Multi-criterion computational validation and structural analyses of the DexDesign-generated peptides,
5. OSPREY-generated structural ensembles of the D-peptide:PDZ domain complexes, and
6. An open source implementation of DexDesign in the computational protein redesign software OSPREY.”

Revised (13 words):

“We use DexDesign to predict D-peptides inhibitors of CALP and MAST2 PDZ domains.”
Bibliography


(2) Garton, M.; Sayadi, M.; Kim, P. M. A Computational Approach for Designing D-Proteins with Non-Canonical Amino Acid Optimised Binding Affinity. PLOS ONE 2017, 12 (11), e0187524. https://doi.org/10.1371/journal.pone.0187524.


(17) Robust deep learning–based protein sequence design using ProteinMPNN.
https://doi.org/10.1126/science.add2187.
(29) Cailliet-Saguy, C.; Maisonneuve, P.; Delhommel, F.; Terrien, E.; Babault, N.; Lafon, M.; Cordier, F.; Wolff, N. Strategies to Interfere with PDZ-Mediated Interactions in Neurons:
https://doi.org/10.1016/j.pbiomolbio.2015.02.007.

https://doi.org/10.1074/jbc.RA119.008238.

https://doi.org/10.1126/scisignal.2000510.

https://doi.org/10.1021/acs.jpcb.9b07278.


