Method to generate highly stable D-amino acid analogs of bioactive helical peptides using a mirror image of the entire PDB

Michael Garton[a], Satra Nim[a], Tracy A. Stone[b,c], Kyle Ethan Wang[d], Charles M. Deber[b,c], and Philip M. Kim[a,d,e,1]

[a]Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto M5S 3E1, Canada; [b]Department of Biochemistry, University of Toronto, Toronto M5S 1A8, Canada; [c]Division of Molecular Medicine, Research Institute, Hospital for Sick Children, Toronto M5G 0A4, Canada; [d]Department of Molecular Genetics, University of Toronto, Toronto M5S 1A8, Canada; and [e]Department of Computer Science, University of Toronto, Toronto M5S 2E4, Canada

Edited by William F. DeGrado, University of California, San Francisco, CA, and approved December 21, 2017 (received for review July 4, 2017)

Biologics are a rapidly growing class of therapeutics with many advantages over traditional small molecule drugs. A major obstacle to their development is that proteins and peptides are easily destroyed by proteases and, thus, typically have prohibitively short half-lives in human gut, plasma, and cells. One of the most effective ways to prevent degradation is to engineer analogs from dextrorotary (D)-amino acids, with up to 10^5-fold improvements in potency reported. We here propose a general peptide-engineering platform that overcomes limitations of previous methods. By creating a mirror image of every structure in the Protein Data Bank (PDB), we generate a database of ∼2.8 million D-peptides. To obtain a D-analog of a given peptide, we search the (D)-PDB for similar configurations of its critical—"hotspot"—residues. As a proof of concept, we apply our method to two peptides that are Food and Drug Administration approved as therapeutics for diabetes and osteoporosis, respectively. We obtain D-analogs that activate the GLP1 and PTH1 receptors with the same efficacy as their natural counterparts and show greatly increased half-life.

peptide | D-amino acid | biologics | D-analog | peptidomimetics

Proteins and peptides have a number of properties that make them highly effective as therapeutic agents. These include very precise specificity, high binding affinity, low toxicity, and low risk of drug–drug interactions. Their diversity also provides very broad coverage of disease targets. Despite this, there are relatively few peptide drugs approved—around 60—compared with around 1,500 small molecule drugs. One major reason for this is their susceptibility to degradation by proteases and rapid renal clearance (1). Consequently, they often have prohibitively short gut, blood plasma, and intracellular half-lives. Peptides therefore tend to have low i.v. bioavailability and especially poor oral bioavailability, requiring frequent injections and severely limiting their use. Many peptide drug candidates struggle to progress beyond preclinical experiments due to bioavailability considerations. An array of techniques designed to stabilize peptides and increase their half-life has emerged and is currently driving a rapid expansion in drug candidates (2). These include PEGylation, backbone modifications, cyclization, stapling, and lipidation (3). One of the most effective approaches is the incorporation of dextrorotary (D)-amino acids (4, 5).

All amino acids except glycine exhibit chirality and, therefore, can exist in one of either (D) or levorotary (L) forms—so-called because of their influence on plane-polarized light. (D)-amino acids are occasionally found in nature (e.g., in some venoms, antibiotics, and peptidoglycan cell walls); however, this is extremely rare (6). Biology is peculiarly homochiral and constructed almost exclusively from the (L)-enantiomer. A useful consequence of this is that (D)-proteins are highly resistant to degradation and have low immunogenicity (7). The fundamental change in backbone–side-chain connectivity and geometry means they are not recognized as proteins by many (L)-proteins—including proteases. Consequently, (D)-proteins are reported to have greatly increased gut, blood plasma, and intracellular half-lives (8). Better cell penetration has also been reported in some cases (9, 10). This behavior can impart potency improvements of up to five orders of magnitude compared with their (L)-counterparts (11).

There are currently two approaches to engineering proteins with (D)-amino acids. Both approaches have significant limitations that preclude application to the majority of known or putative therapeutic peptides and drug targets (12–14). Simply replacing (L) for (D)-amino acids is generally ineffective as side-chain orientations with respect to the target are completely altered (15). Fig. 1*A* shows the consequence of simple (D) replacement in helical peptides. Change in side-chain orientation prevents correct binding geometry and destroys target binding.

An elegant solution to this problem in unstructured peptides is retro-inversion (RI). RI involves reversing the (D)-peptide sequence—flipping the termini and, thus, restoring the (L)-amino side chain angles. It has been used with some success on unstructured peptides (16, 17). The extended (D)-peptides assume side chain topology similar to their parent molecule but with inverted amide peptide bonds. However, RI usually fails if the peptide has secondary structure, owing largely to the topological properties of helices. (D)-peptides always adopt left-handed helices (18, 19), while (L)-peptide helices are always right-handed.

Significance

Using D-amino acids as the building blocks for bioactive peptides can dramatically increase their potency. However, simply swapping regular levorotary amino acids for dextrorotary (D)-amino acids alters the peptide surface topology and function is lost. Current methods to overcome this are not generally applicable and exclude the majority of therapeutic targets. By creating a mirror image of all 111,867 protein structures in the Protein Data Bank (PDB), we convert this repository into a D-peptide database with 2.8 million D-peptide structures. This D-PDB can be searched to find therapeutically active topologies, demonstrated here by the discovery of D-peptide GLP1R and PTH1R agonists. Evaluation of D-PDB coverage suggests that it holds candidates for most therapeutic targets and, thus, potentially contains hundreds of potent drug leads.

Author contributions: M.G., C.M.D., and P.M.K. designed research; M.G., S.N., T.A.S., K.E.W., C.M.D., and P.M.K. performed research; S.N., T.A.S., and C.M.D. contributed new reagents/analytic tools; M.G., S.N., K.E.W., and P.M.K. analyzed data; and M.G. and P.M.K. wrote the paper.

Conflict of interest statement: The University of Toronto is in the process of filing for a patent on this method.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: pi@kimlab.org.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1711837115/-/DCSupplemental.

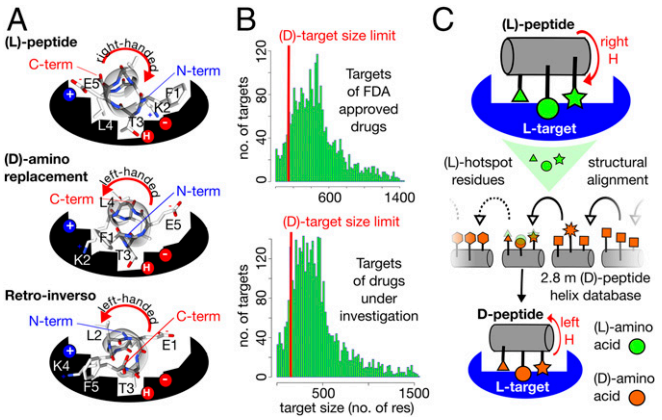

Fig. 1. Limitations of current (D)-protein engineering techniques and a solution. (*A*) Loss of specific peptide–target interactions as a consequence of direct conversion to (D)-amino acids in helical peptides. Charged groups are shown on the target (black) as a white "plus" or "minus" signs on blue or red (respectively) spots. Peptide charges are shown as blue or red plus and minus signs. Target hydrogen bond participating groups are shown as a white "H" on a red spot. Red curly arrows highlight the change in helix handedness from right to left upon conversion to (D). (D)-helix left-handedness means that helical peptide–target interactions fail to be restored, even when subject to RI. (*B*) Histograms showing the distribution of protein target sizes (in residues) for FDA-approved drugs (*Upper*) and drug candidates being investigated (*Lower*). Targets that meet current commercial size limits for (D)-target synthesis are dwarfed by those precluded from MIPD. (*C*) Schematic overview of an approach presented in this work. Hotspot residues constellations are used to search a (D) amino acid version of the PDB comprising ~2.8 million helices. Matches bind to the (L)-target with comparable affinity.

Left-handed (D)-helices remain left-handed even when the sequence is reversed in RI (Fig. 1*A*). The resulting topological differences destroy binding (12, 15). Approximately 62% of protein–protein recognition in the Protein Data Bank (PDB) is mediated by helical elements (20) and 80% of Food and Drug Administration (FDA)-approved peptide drugs are helical (14). This means that the majority of therapeutically interesting peptides are inaccessible to the RI technique.

The only alternative to RI for engineering (D)-amino peptides is mirror image phage display (MIPD). In MIPD, targets are synthesized in (D)-space and used as bait for a randomized (L)-amino peptide library (21). Successful candidate peptides/proteins subsequently made with (D)-amino acids bind the native (L)-protein target with the same affinity as their reverse. Histograms in Fig. 1*B* show drug target sizes for FDA-approved drugs and for targets of drugs subject to preclinical testing or clinical trial. (D)-protein synthesis is currently limited to a target size of ~150 residues by commercial techniques, although synthesis of up to 312 residues has been reported in an exceptional case (22). This means that MIPD is limited to only a small subset of known targets. Importantly, this target size limitation largely precludes membrane proteins, which comprise ~60% of all therapeutic targets (23). Isolated extracellular domains can be made; however, these usually fail to adopt the correct conformation without constraint by the full protein. Transmembrane regions, which are difficult to produce recombinantly, are also often involved in the ligand interaction. Furthermore, agonistic activity requires more than simple binding, making agonist selection difficult with MIPD. In addition to size limitations, many targets require chaperones or obligate heterodimeric partners to fold. (L)-chaperones are highly unlikely to specifically recognize a (D)-protein substrate because their topology is very different. Folding is therefore usually precluded (24) although an exception has been demonstrated for DapA folding by GroEL/ES (22)—thought to proceed using nonspecific hydrophobic interactions.

RI and MIPD limitations mean that the majority of known and putative therapeutic targets are inaccessible to current (D)-peptide engineering techniques. Here, we propose an approach that overcomes these limitations and enables the design of helical (D)-peptides to a much broader range of targets. A schematic of our approach is shown in Fig. 1*C*.

The PDB contains over 110,000 naturally occurring and engineered structures. It is therefore a very rich source of information for the rational design of proteins. Our approach exploits this resource by creating a mirror image version of the entire repository—thereby rendering every structure in (D)-amino acids. We further compartmentalize the structures into ~2.8 million helices and call this database the "(D)-PDB." The (D)-PDB is then scanned—using structural alignment—for residue configurations that match the hotspot residue configurations of therapeutically interesting (L)-peptides (Fig. 1*C*). Hotspot residues are those identified as contributing significantly to target recognition, binding, and receptor activation. They are a small subset of the full peptide—typically no more than three or four residues. Finding a structurally equivalent set in the (D)-PDB is therefore highly probable.

Using the glycogen-like-peptide (GLP1) and parathyroid hormone (PTH) as proof-of-concept test cases, we successfully generated (D)-helix agonists of the GLP1 and PTH1 receptors using matches discovered in the (D)-PDB.

Results

(D)-PDB Construction. Internal interactions of a protein are identical in its mirror image. This allowed the creation of a parallel protein database composed of (D)-proteins simply by flipping structure files with Cartesian coordinates along the x axis. Each flipped structure is composed entirely of (D)-amino acids and should fold exactly as the in silico structure shows when synthesized. A schematic showing (D)-PDB construction is shown in Fig. 2.

After removing any nonprotein molecules such as DNA, solvent, and ions, each file in the PDB is flipped along the x axis to create a mirror image version. Nonhelical parts of the protein were then removed and each helix was put into a separate file, totaling more than 2.8 million helix files. This separation ensured that hotspot alignments would only occur on relatively short, contiguous peptide regions. Redundancy was allowed, as even small differences—such as different side chain rotamers—increase the method power. Since protein regions without secondary structure can effectively be converted to (D) experimentally using RI, we removed such regions from the (D)-PDB. Beta-sheet/strand structures were also removed for simplicity, and because therapeutic peptides tend to be either helical or unstructured.

Query Preparation. It is first necessary to identify or make a crystal structure—or NMR solution structure—of the functional (L)-peptide. A homology model could also be used. It should be noted that homology model effectiveness will likely be highly dependent on the degree of conservation with known structures. Residues critical to target binding and activity can then be identified—often from the literature—by alanine scanning mutagenesis. Ideally this is done experimentally, but with a target bound structure, it can also be carried out computationally using techniques such as thermodynamic integration or free energy perturbation.

Once the hotspot residues are identified, various atom sets are designated within each residue. Usually these are pairs of atoms, but in the case of ring-containing amino acids such as Phe, Tyr, and Trp, a set may include three. Each set is ranked according to its importance to target interaction, with level 1 being the highest. Level 1 usually means the atom pair or triplet furthest from the backbone and, thus, closest to the target. We assume that if level 1 can be matched, the remaining side chain atoms need not match to be effective. This greatly increases the chance of a match in the (D)-PDB. The other levels are used if level 1 atom pairs do not produce any suitable matches. Lower atom level matches can be used because one of the residue's rotamers—above this level—will usually correctly position the level 1 atoms.

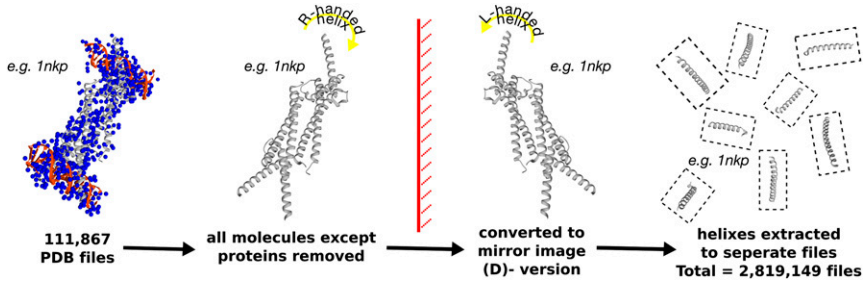

Fig. 2. Schematic illustrating (D)-PDB construction. Every PDB file is retrieved, some containing various nonpeptide molecules such as nucleic acids (orange) and solvent (blue). These are removed before creating a mirror image of the remaining protein molecule Cartesian coordinates—resulting in helix handedness change from right to left (yellow arrows). More than 2.8 million (D)-helices are extracted into separate files. Example PDB file used is 1NKP (Myc–DNA complex).

Intramolecular clash can occur between these rotamers and nonhotspot residues in the match. This is identified by full reconstruction of the match, with rotamers that allow correct level 1 positioning. Matches thereby considered nonviable are discarded. Any three atoms of a ring (or rings) can be used to ensure that the correct planarity is represented. Default atom levels for each standard amino acid are shown in Fig. S1.

Another way to increase the likelihood of a (D)-PDB match is to group residues by similarity. For example, if a query hotspot is Arg, then matches with both Arg and Lys could be allowed. A (D)-peptide Lys match may be effectively used in the final design, or mutated to (D)-Arg with little effect on helix integrity. Similarity residue groupings are shown in Fig. S1 and, with atom renaming, can be used in combination with atom levels to maximize (D)-match likelihood.

In many cases, (L)-peptides of interest have both helical and unstructured elements. Only hotspots in the helical region are used, on the basis that unstructured peptide can be generated by RI, and added in postprocessing. To facilitate RI linkage, the last helical residue immediately adjacent to the unstructured region is designated a "junction" residue and included in the alignment (D)-PDB scan. Only backbone atoms (N, O, C, and CA) are used for junctions unless the junction is also a hotspot. Using backbone atoms ensures the postmatch-added RI unstructured peptide will be oriented in the same direction as the (L)-equivalent was. This ensures that correct arrangement of unstructured hotspots in relation to structured hotspots is possible when the unstructured RI region is attached.

For the RI peptide sections to be attached, D-matches need to have opposite sequence direction to the L-query. To ensure matches have this reversed directionality, junction query backbone atoms are rotated 180° about the CA–CB bond axis. A rotation of 108.5° about the CA—such that CA–R and CA–H exchange positions—is also performed to precisely recapitulate backbone direction. This is necessary to facilitate extension of the reversed sequence of adjoined unstructured RI regions, where N and C termini are switched. Junction residues thereby allow an RI version of unstructured regions to be attached to (D)-PDB matches. Implementing the 180° rotation step means that D-matches always have the correct sequence direction for RI extension.

Test Case: GLP-1. GLP-1 is currently of interest as a diabetes mellitus and obesity treatment (25) and was chosen as a good proof-of-concept test case. It is a particularly challenging case that tests the limits—and demonstrates the full utility—of this method. It involves multiple helices, multiple unstructured regions, negatively charged hotspots, positively charged hotspots, hydrophobic hotspots, ringed hotspots, and a junction–hotspot residue. GLP-1 is a helical GPCR agonist, and this makes engineering a (D)-analog very difficult using conventional methods. There is good availability of structures and hotspot residue information (25) together with a structure for the ligand bound to the extracellular domain of the B-class GPCR (26).

Fig. 3 shows the process of preparing GLP1 structures to query the (D)-PDB. Unbound NMR solution structures (PDB ID code: 4GZM) and the receptor bound crystal structure (PDB ID code: 3IOL) were used as starting points. GLP-1 is composed of two helices joined by a four-residue flexible linker. Each helix was set up separately with a view to relinking two matches using the retro inverted linker sequence. Helix 1 runs from T7 to Y13. Helix 2 runs from A18 to K28. In helix 1, T7 and D9 are identified as hotspot residues, while T7 and Y13 act as junction residues. F17, I18, and L20 are the hotspots for helix 2, while A18 and K28 are the junction residues. Fig. 3*B* shows how hotspot and junction residues are prepared following extraction from their structure. First, the junction residue backbone atoms are rotated about the CA–CB axis by 180° and then by 108.5° about the CA along a defined plane. This ensures that a (D)-match can accept correctly orientated RI linker and terminal tail sequence in postmatch processing. Following this, six query structures are generated for helix 1, and 27 for helix 2, one for each combination of atom levels (Fig. 3*C*). In the event of no good match, each of the 33 query structures can be rerun using chemically similar residues. Fig. 3*D* delineates the order in which each of these was prepared, together with the combinations of atom levels used in each case. K34, while not a definitive hotspot, has been shown to contribute slightly. For this reason, both Lys and Arg were queried before "any," as a positive charge was slightly preferred.

(D)-Match Output Processing. After running each query variant sequentially as outlined in Fig. 3, a number of matches were located in the (D)-PDB. Match quality was measured using the root-mean-square deviation (rmsd) of every atom level combination with corresponding level combinations—if they exist—in every (D)-PDB file. The rmsd cutoff was set for <1.5 Å, although <1.0 Å is ideal if possible. The best match for helix 1 was found in 3s6d.pdb at 0.5 Å, and in 4rzf.pdb at 0.9 Å for helix 2. Fig. 4*A* shows query–match structural alignments and—together with colored dots—indicates the successful query variants. Match sequences are reverse ordered due to the junction backbone 180° rotation, allowing RI peptide extension as planned. Both sequences are substantially different to their (L)-query. These matches were then combined with RI unstructured regions to construct the full (D)-analog of GLP-1 (Fig. 4*B*). A full (D)-analog structure was constructed and docked to the GLP1 receptor (GLP1R) ECD structure (Fig. 4*C*). Residues R12 and Q13 were found to clash with the receptor and were therefore mutated to alanine. W3 and H23 were also provisionally mutated to original query residue types—subject to checks on helix integrity.

The full (D)-analog sequence was checked for helix integrity using PSI-pred (27) (Fig. 3*D*). In addition to the mutations, this was to check that secondary structure is preserved when matched helices are removed from their full protein context. It also provides assurance that the full (D)-analog can fold in the same way as its components. This is necessary in order that the (D)-peptide configuration of hotspot residues closely resembles that of the (L)-peptide, and is presented as such to the target. It also

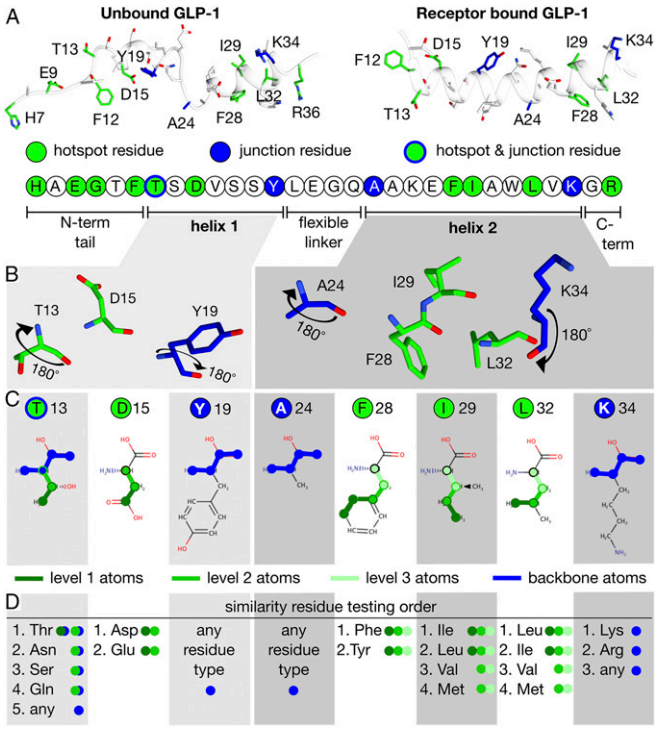

Fig. 3. Preparation of GLP-1 queries for scanning the (D)-PDB. (*A*) GLP-1 structures and sequence, including the free peptide in solution (*Left*) and receptor ECD bound structure (*Right*). Free GLP-1 was solved using NMR and reveals a central unstructured linker region in contrast to GLP-1R bound. Hotspot and junction residues are annotated in green and blue, respectively. (*B*) Hotspot are extracted separately for helix one and two, together with junction residues that have their backbone atoms rotated 180°. Rotation ensures (D)-peptide matches have reversed sequence order—a requirement for RI linker and tail attachment. (*C*) Levels (1–3) are assigned to hotspot atom pairs or triplets according to estimated import to target binding. (*D*) Atom levels are combined with similar residues. A combination order of decreasing quality—to sequentially test the (D)-PDB until close matches are identified—is thereby established.

highlights any influence that unstructured RI regions may have on the helix or vice versa—such as unwanted structure induced into an RI region by adjoining helix. Fig. 4*D* shows that (D)-GLP1 has approximately the same secondary structure profile as (L)-GLP1. An isoelectric point prediction of pH 4.66, and net charge evaluation of −1.9 at pH 7 for (D)-GLP1 using pepcalc (28), indicates that it has good solubility and, thus, is suitable for experimental validation.

Experimental Validation of (D)-GLP1. The best candidate was then synthesized from (D)-amino acids and tested for its capacity to activate GLP1R. Binding of GLP-1 to the GLP1R has previously been shown to activate adenylyl cyclase (AC) with consequent production of cAMP, which, in turn, activates protein kinase A (PKA) to phosphorylate and activate cAMP response element-binding protein (CREB). In the present study, we investigated the ability of (D)-GLP1 peptide to induce activation of GLP1R and compared the response with native (L)-GLP1 peptide. We generated a stable GLP1 receptor/CRE-luciferase expressing HEK293 cell line and observed a cAMP-inducible luciferase expression following treatment with Forskolin (Fig. 5*A*). (L)-GLP1 peptide increased luciferase expression in GLP1 receptor expressing HEK293 cells, but was inactive in pCDNA3.1 HEK293 cells. (L)-GLP1 peptide displayed an EC_{50} value of 59.6 nM with 67.2% efficacy relative to maximum stimulation by Forskolin (Fig. 5*A*). (D)-GLP1 peptide also increased luciferase expression in GLP1 receptor expressing HEK293 cells (Fig. 5*A*). (D)-GLP1 peptide displayed an EC_{50} value of 2.2 μM with a similar efficacy as the (L)-GLP1 peptide. A scrambled version of (D)-GLP1 was simultaneously tested as a negative control—to account for any nonspecific effects—and showed no activity.

To investigate the mechanisms underlying the effects of (D)-GLP1 peptide on GLP1R, we studied downstream effects of activating GLP1R with (D)-GLP1 peptide. We looked to find if activation of GLP1R with (D)-GLP1 peptide would induce phosphorylation of ERK1/2 and AKT. In HEK293 cells expressing GLP1R, 10 μM (L)-GLP1 peptide evoked a robust increase in ERK activation as assessed by the increase in phospho-ERK1/2 (Fig. S4). The maximum level of phospho-ERK1/2 was achieved around 60 min after stimulation. (D)-GLP1 peptide at a concentration of 10 μM also activated ERK1/2, evoking a maximum increase of phospho-ERK1/2 around 60 min after stimulation. The level of phospho-ERK1/2 was sustained after 120 min following (D)-GLP1 treatment while the signal decreased after 60 min with (L)-GLP1.

Resistance to protease degradation is one of the most useful properties of D-peptides generally. We carried out quantitative analysis of the (D)-GLP1 Proteinase K (ProtK) resistance and compared with (L)-GLP1. Fig. 5*C* shows total loss of (L)-GLP1 in <1 h, while 80% of (D)-GLP1 can still be detected after 6 h exposure to ProtK.

Test Case 2: Parathyroid Hormone. To demonstrate the general applicability of this technique, another test case was selected. Parathyroid hormone (PTH) is an FDA-approved treatment for osteoporosis delivered by daily s.c. injection. Osteoporosis affects ∼200 million people worldwide but only a fraction receive

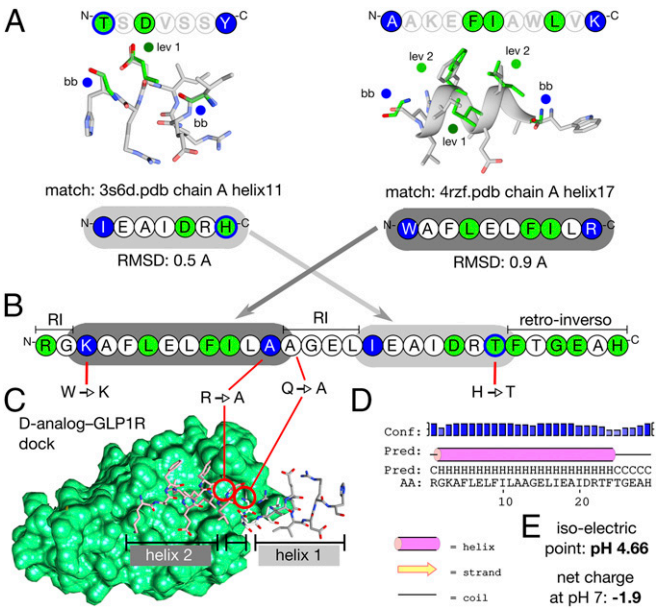

Fig. 4. GLP-1 Best (D) match results and full (D)-peptide construction. (*A*) (L)-query sequences (*Upper*) showing hotspots (green), junction residues (blue), and remaining original sequence (gray). Closest matching (D) structures are shown with atom levels annotated with dots corresponding to colors from Fig. 3. Match sequences are significantly different to query sequences. Helix 1 is highlighted light gray and helix 2 is dark gray (*Lower*). (*B*) Full (D)-analog construction from best (D) match helix sequences juxtaposed with RI linker and terminal tail sequences. (*C*) Construction of D-analog structure from match helices and modeled linker. Docking to GLP-1R ECD identifies potential steric clashes, circumvented by mutation to alanine. Reintroduction of native peptide side chains at two junction positions is also judged prudent. (*D*) PSI-PRED predicts that correct secondary structure is maintained in the (D)-analog, with medium to high confidence (blue bars). (*E*) Solubility check results predict good solubility.

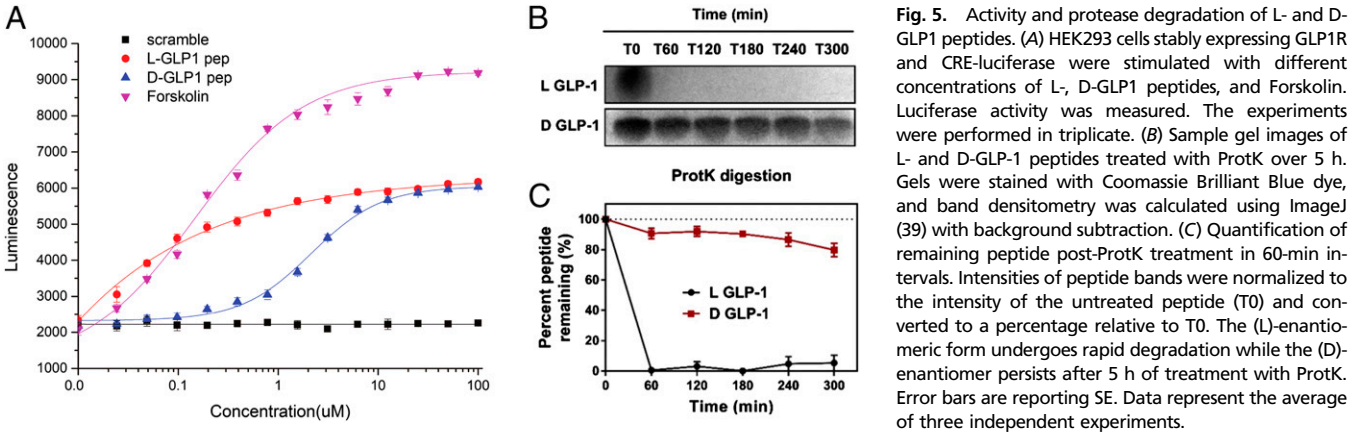

Fig. 5. Activity and protease degradation of L- and D-GLP1 peptides. (*A*) HEK293 cells stably expressing GLP1R and CRE-luciferase were stimulated with different concentrations of L-, D-GLP1 peptides, and Forskolin. Luciferase activity was measured. The experiments were performed in triplicate. (*B*) Sample gel images of L- and D-GLP-1 peptides treated with ProtK over 5 h. Gels were stained with Coomassie Brilliant Blue dye, and band densitometry was calculated using ImageJ (39) with background subtraction. (*C*) Quantification of remaining peptide post-ProtK treatment in 60-min intervals. Intensities of peptide bands were normalized to the intensity of the untreated peptide (T0) and converted to a percentage relative to T0. The (L)-enantiomeric form undergoes rapid degradation while the (D)-enantiomer persists after 5 h of treatment with ProtK. Error bars are reporting SE. Data represent the average of three independent experiments.

PTH, partly due to the lack of an oral delivery option. (D)-peptides have shown some oral bioavailability in human trial (29, 30) and, thus, a (D)-analog of PTH could be of interest. PTH is also of interest for treating hyperparathyroidism (31) and to promote bone growth following fracture (32).

The same process as described for GLP-1 was repeated: Crystal structures and hotspot residues were identified from the literature (33, 34). Fig. S5*A* shows PTH (1–34) with hotspot residues colored green and junctions in blue, again split into two helices. Helix one hotspots + junctions found a closest (D)-PDB match of 0.95 Å, while helix two was 0.82 Å (Fig. S5*B*). Reconstruction of the full (D)-peptide using RI for the linker and terminal tails is shown in Fig. S5*C*. A structural model of the (D)-PTH was constructed and positioned on the receptor to align with hotspot residues. Several mutations were introduced to remove clash and enhance similarity to (L)-PTH (M1–M5).

As with GLP-1 and GLP1R, binding of PTH (residues 1–34) to the parathyroid receptor (PTH1R) has also been shown to activate AC, triggering cAMP production. This activates PKA to phosphorylate and activate CREB. Fig. 6*A* shows that the (D)-PTH designed here activates PTH1R with a potency and efficacy comparable to (L)-PTH and Forskolin. Protease stability was also calculated and again showed a dramatic difference in degradation rate between the (L) and (D) versions (Fig. 6*B* and Fig. S5*D*). All of the (L)-PTH is degraded in under 1 h, while more than 85% of the (D)-analog is still detectable at 6 h.

Discussion

We have developed a method for converting (L)-peptides to highly stable (D)-analogs by exploiting a mirror image version of the PDB. We used it to engineer (D)-peptide analogs of the agonists GLP-1 and PTH that successfully activate GLP1R and PTH1R, respectively. It is a simple and inexpensive method to implement, especially if a starting structure is already available—or can be obtained with reasonable confidence by homology modeling. Otherwise, helical (L)-peptide structures are mostly straightforward to obtain using X-ray crystallography or NMR. Information on hotspot residues can often be sourced from the literature, or otherwise obtained by straightforward alanine scanning mutagenesis experiments.

An interesting point emerging from this study is a powerful confirmation of the notion of hotspot residues. We have shown that a peptide can be totally changed—except for the position a small subset of residues—and function is still retained.

While (D)-PTH was comparable to (L)-PTH potency and efficacy, (D)-GLP1 potency was ∼40-fold lower than the native peptide, albeit with similar efficacy. Further optimization could involve testing multiple candidates—as only one of seven (D)-GLP1 candidates produced by the method was tested—and refining with mutagenesis experiments. However, as a proof-of-concept study, we were primarily concerned with the less trivial problem of finding functional (D)-scaffolds. Affinity is a common victim of methods to engineer stability. However, increased stability and longer plasma half-life means that even a large affinity loss can still yield a net improvement in potency. Phospho-ERK experiments indicated that half-life was increased by ∼fivefold. It is well established that GLP1R has rapid internalization and desensitization (35, 36); it is thus likely that this process is responsible for the relatively modest improvement in activity duration, rather than degradation. This was confirmed by proteinase K degradation experiments, which showed dramatic improvement in stability for both GLP-1 and PTH. It should be noted that our study compared the activity of (D)-analogs to the native hormones, and not to the many available analogs that may have higher potency (particularly in the case of GLP1). Therefore, depending on the application, much additional work may be required to optimize (D)-analogs and fully assess their therapeutic potential compared with currently approved solutions. This may involve introducing noncanonical amino acids or chemical modifications.

Nonhotspot residues can vary greatly between the original (L)-peptide and (D)-analogs engineered this way. While not contributing

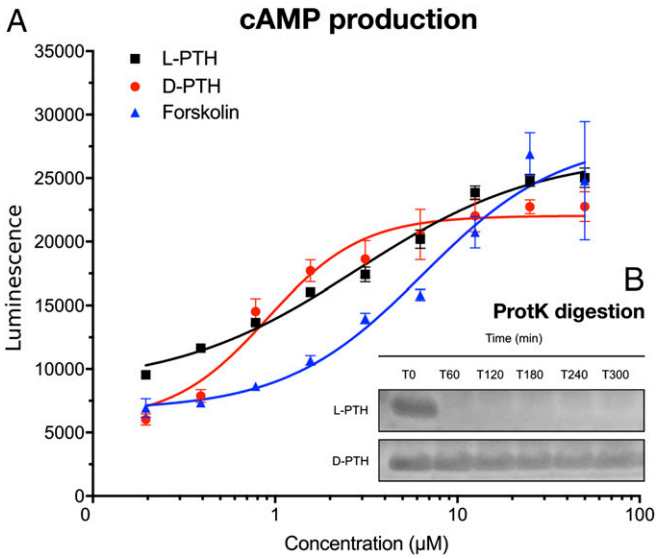

Fig. 6. Activity and protease degradation of (L)- and (D)-PTH peptides. (*A*) HEK293 cells stably expressing PTH1R and CRE-luciferase were stimulated with different concentrations of (L)-, (D)-PTH peptides and Forskolin. Luciferase activity was measured. The experiments were performed in triplicate. (*B*) Sample gel images of (L)- and (D)-PTH peptides treated with ProtK over 5 h. Gels were stained with Coomassie Brilliant Blue dye, and band densitometry was calculated using ImageJ (39) with background subtraction.

significantly to the interaction, these differences may still adversely affect binding. For instance, bulky or charged (D)-peptide residues may interact with the target in a disruptive manner, especially if that space in the (L)-version is occupied by small or uncharged residues. Mutagenesis could be used to resolve this. GLP-1 was one of the more challenging cases: an agonist consisting of two helices connected by a flexible linker. Reconstructing the full peptide from two different (D) matches, such that both helices retain correct relative orientations to each other, is difficult. Cases with a single helix, fewer hotspots, or simpler antagonistic action, will likely be more tractable. There is one RI limitation that our method does not overcome. Target interactions with the (L)-peptide backbone may be lost upon conversion to a (D)-peptide. Binding is likely to be adversely affected if recognition involves hydrogen bonding with the peptide backbone.

(D)-analogs generally avoid some of the limitations of stabilizing methods such as stapling, lipidation, PEGylation (2). These approaches can lead to significant conformational change that can adversely affect their activity. Reduced solubility is another common drawback associated with such approaches. In certain cases, where these limitations are not catastrophic, (D)-analogs could potentially be enhanced using these techniques. Combining approaches is likely to be additive or synergistic in terms of increasing half-life. As such, (D)-PDB matching can be seen as complementing other techniques, rather than competing with them.

Peptide therapeutics are currently undergoing a huge expansion, and the market size is predicted to continue its dramatic increase over the next few years (37). The most recent published estimate for the number of peptides in clinical and preclinical development is 140 and 500, respectively (3). With ∼80% of these likely to be helical (14), this means that over 500 of these are potentially immediately applicable for use with the (D)-PDB method. The majority of these are at present prohibited by the limitations of current methodologies. It should be noted that this estimate was published in January 2015 and, therefore, the current number of peptides in development is likely to now be significantly higher. Several (D)-amino acid containing peptide therapeutics have been approved for use, thus far indicating no inherent toxicity to humans (37).

Our work demonstrates proof of concept and leaves scope for further development into, for example, β-strand peptides, and the engineering of larger (D)-proteins. We anticipate that (D)-PDB matching will become another key tool for finding stable lead molecules in early stage drug discovery.

Materials and Methods

The latest protein database was downloaded and converted to the (D)-PDB by inverting the *x* coordinate sign in each structure. Hotspot information for GLP-1 and PTH was readily available in the literature. Structural alignments were carried out with Click (38), while PSI-PRED (27) was used to predict the likely secondary structure, and PepCalc (28) was used to predict peptide solubility. Both (L)- and (D)-peptides were synthesized by Lifetein LLC. The HEK293 cell line was obtained from the American Type Culture Collection (ATCC), and cells stably expressing hGLP1R and reporter CRE-Gaussia Luciferase construct were read using a luminometer with a 480-nm filter. ProtK (Bioshop) digestions were repeated three times, and densitometry of bands was determined using ImageJ software (39).

For a more detailed description, please refer to the *Supporting Information*.

ACKNOWLEDGMENTS. We thank Daniel Drucker and his laboratory for providing the BHK cells expressing rat and human GLP-1R. This work was supported by grants of the Canadian Institute of Health Research Grants MOP-123526 and PJT-153279 (to P.M.K.) and a Natural Sciences and Engineering Research Council of Canada grant (to C.M.D.).

1. Bruno BJ, Miller GD, Lim CS (2013) Basics and recent advances in peptide and protein drug delivery. *Ther Deliv* 4:1443–1467.
2. Corbi-Verge C, Garton M, Nim S, Kim PM (2017) Strategies to develop inhibitors of motif-mediated protein-protein interactions as drug leads. *Annu Rev Pharmacol Toxicol* 57:39–60.
3. Fosgerau K, Hoffmann T (2015) Peptide therapeutics: Current status and future directions. *Drug Discov Today* 20:122–128.
4. Welch BD, VanDemark AP, Heroux A, Hill CP, Kay MS (2007) Potent D-peptide inhibitors of HIV-1 entry. *Proc Natl Acad Sci USA* 104:16828–16833.
5. Liu M, et al. (2010) D-peptide inhibitors of the p53-MDM2 interaction for targeted molecular therapy of malignant neoplasms. *Proc Natl Acad Sci USA* 107:14321–14326.
6. Kreil G (1997) D-amino acids in animal peptides. *Annu Rev Biochem* 66:337–345.
7. Uppalapati M, et al. (2016) A potent D-protein antagonist of VEGF-A is nonimmunogenic, metabolically stable, and longer-circulating in vivo. *ACS Chem Biol* 11:1058–1065.
8. Rabideau AE, Pentelute BL (2015) A D-amino acid at the N-terminus of a protein abrogates its degradation by the N-end rule pathway. *ACS Cent Sci* 1:423–430.
9. Nickl CK, et al. (2010) (D)-Amino acid analogues of DT-2 as highly selective and superior inhibitors of cGMP-dependent protein kinase Ialpha. *Biochim Biophys Acta* 1804:524–532.
10. Brugidou J, Legrand C, Méry J, Rabié A (1995) The retro-inverso form of a homeobox-derived short peptide is rapidly internalised by cultured neurones: A new basis for an efficient intracellular delivery system. *Biochem Biophys Res Commun* 214:685–693.
11. Veine DM, Yao H, Stafford DR, Fay KS, Livant DL (2014) A D-amino acid containing peptide as a potent, noncovalent inhibitor of α5β1 integrin in human prostate cancer invasion and lung colonization. *Clin Exp Metastasis* 31:379–393.
12. Li C, et al. (2010) Limitations of peptide retro-inverso isomerization in molecular mimicry. *J Biol Chem* 285:19572–19581.
13. Raibaut L, Ollivier N, Melnyk O (2012) Sequential native peptide ligation strategies for total chemical protein synthesis. *Chem Soc Rev* 41:7001–7015.
14. Law V, et al. (2014) DrugBank 4.0: Shedding new light on drug metabolism. *Nucleic Acids Res* 42:D1091–D1097.
15. Li C, et al. (2013) Functional consequences of retro-inverso isomerization of a miniature protein inhibitor of the p53-MDM2 interaction. *Bioorg Med Chem* 21:4045–4050.
16. Li H, et al. (2015) Novel retro-inverso peptide inhibitor reverses angiotensin receptor autoantibody-induced hypertension in the rabbit. *Hypertension* 65:793–799.
17. Ben-Yedidia T, Beignon AS, Partidos CD, Muller S, Arnon R (2002) A retro-inverso peptide analogue of influenza virus hemagglutinin B-cell epitope 91-108 induces a strong mucosal and systemic immune response and confers protection in mice after intranasal immunization. *Mol Immunol* 39:323–331.
18. Hung LW, Kohmura M, Ariyoshi Y, Kim SH (1998) Structure of an enantiomeric protein, D-monellin at 1.8 A resolution. *Acta Crystallogr D Biol Crystallogr* 54:494–500.
19. Novotny M, Kleywegt GJ (2005) A survey of left-handed helices in protein structures. *J Mol Biol* 347:231–241.
20. Jochim AL, Arora PS (2010) Systematic analysis of helical protein interfaces reveals targets for synthetic inhibitors. *ACS Chem Biol* 5:919–923.
21. Schumacher TN, et al. (1996) Identification of D-peptide ligands through mirror-image phage display. *Science* 271:1854–1857.
22. Weinstock MT, Jacobsen MT, Kay MS (2014) Synthesis and folding of a mirror-image enzyme reveals ambidextrous chaperone activity. *Proc Natl Acad Sci USA* 111:11679–11684.
23. Overington JP, Al-Lazikani B, Hopkins AL (2006) How many drug targets are there? *Nat Rev Drug Discov* 5:993–996.
24. Rüdiger S, Schneider-Mergener J, Bukau B (2001) Its substrate specificity characterizes the DnaJ co-chaperone as a scanning factor for the DnaK chaperone. *EMBO J* 20:1042–1050.
25. Manandhar B, Ahn JM (2015) Glucagon-like peptide-1 (GLP-1) analogs: Recent advances, new possibilities, and therapeutic implications. *J Med Chem* 58:1020–1037.
26. Underwood CR, et al. (2010) Crystal structure of glucagon-like peptide-1 in complex with the extracellular domain of the glucagon-like peptide-1 receptor. *J Biol Chem* 285:723–730.
27. Buchan DW, Minneci F, Nugent TC, Bryson K, Jones DT (2013) Scalable web services for the PSIPRED protein analysis workbench. *Nucleic Acids Res* 41:W349–W357.
28. Lear S, Cobb SL (2016) Pep-Calc.com: A set of web utilities for the calculation of peptide and peptoid properties and automatic mass spectral peak assignment. *J Comput Aided Mol Des* 30:271–277.
29. Dunbar RL, et al. (2017) Oral apolipoprotein A-I mimetic D-4F lowers HDL-inflammatory index in high-risk patients: A first-in-human multiple-dose, randomized controlled trial. *Clin Transl Sci* 10:455–469.
30. Bloedon LT, et al. (2008) Safety, pharmacokinetics, and pharmacodynamics of oral apoA-I mimetic peptide D-4F in high-risk cardiovascular patients. *J Lipid Res* 49:1344–1352.
31. Rubin MR, et al. (2016) Therapy of hypoparathyroidism with PTH(1-84): A prospective six year investigation of efficacy and safety. *J Clin Endocrinol Metab* 101:2742–2750.
32. Della Rocca GJ, Crist BD, Murtha YM (2010) Parathyroid hormone: Is there a role in fracture healing? *J Orthop Trauma* 24(Suppl 1):S31–S35.
33. Dean T, Khatri A, Potetinova Z, Willick GE, Gardella TJ (2006) Role of amino acid side chains in region 17-31 of parathyroid hormone (PTH) in binding to the PTH receptor. *J Biol Chem* 281:32485–32495.
34. Pioszak AA, Xu HE (2008) Molecular recognition of parathyroid hormone by its G protein-coupled receptor. *Proc Natl Acad Sci USA* 105:5034–5039.
35. Widmann C, Dolci W, Thorens B (1995) Agonist-induced internalization and recycling of the glucagon-like peptide-1 receptor in transfected fibroblasts and in insulinomas. *Biochem J* 310(Pt. 1):203–214.
36. Roed SN, et al. (2015) Functional consequences of glucagon-like peptide-1 receptor cross-talk and trafficking. *J Biol Chem* 290:1233–1243.
37. Qvit N, Rubin SJ, Urban TJ, Mochly-Rosen D, Gross ER (2017) Peptidomimetic therapeutics: Scientific approaches and opportunities. *Drug Discov Today* 22:454–462.
38. Nguyen MN, Tan KP, Madhusudhan MS (2011) CLICK–topology-independent comparison of biomolecular 3D structures. *Nucleic Acids Res* 39:W24–W28.
39. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH image to ImageJ: 25 years of image analysis. *Nat Methods* 9:671–675.