nature biomedical engineering

ARTICLES

https://doi.org/10.1038/s41551-020-0556-3

Check for updates

De novo development of proteolytically resistant therapeutic peptides for oral administration

Xu-Dong Kong[1], Jun Moriya[1], Vanessa Carle[1], Florence Pojer[2], Luciano A. Abriata[2,3], Kaycie Deyle[1] and Christian Heinis[1]✉

The oral administration of peptide drugs is hampered by their metabolic instability and limited intestinal uptake. Here, we describe a method for the generation of small target-specific peptides (less than 1,600 Da in size) that resist gastrointestinal proteases. By using phage display to screen large libraries of genetically encoded double-bridged peptides on protease-resistant fd bacteriophages, we generated a peptide inhibitor of the coagulation Factor XIa with nanomolar affinity that resisted gastrointestinal proteases in all regions of the gastrointestinal tract of mice after oral administration, enabling more than 30% of the peptide to remain intact, and small quantities of it to reach the blood circulation. We also developed a gastrointestinal-protease-resistant peptide antagonist for the interleukin-23 receptor, which has a role in the pathogenesis of Crohn's disease and ulcerative colitis. The de novo generation of targeted peptides that resist proteolytic degradation in the gastrointestinal tract should help the development of effective peptides for oral delivery.

Orally available peptides are a major goal in the pharmaceutical industry that has not been reached for most targets; almost none of the more than 60 approved peptides are orally available[1,2]. One of the exceptions is the GLP-1 receptor antagonist semaglutide, which was recently approved for oral application[3]. Oral delivery is preferred for drug administration because of its convenient self-administration, wide range of available dosage adjustments, and quick termination if adverse effects occur. The limited oral availability of peptides is mainly due to degradation by proteases in the gastrointestinal tract, poor crossing of the epithelial layer due to large size and polar surface, and first-pass metabolism in the liver[4]. Fortunately, orally available peptide drugs derived from exceptionally stable natural peptides do exist and show that this is, in principle, possible[4–6]. For example, cyclosporine is a cyclic undecapeptide that is passively permeable due to N-methylated amide bonds and intra-molecular hydrogen bonds[7] (oral bioavailability of around 30%), and linaclotide is a tetradecapeptide constrained by three disulfide bridges that acts locally in the intestine and thus needs to only resist proteases and not be absorbed.

Through powerful in vitro display techniques such as phage display and mRNA display, peptide ligands for virtually any disease can be generated, but so far, no peptide isolated with such techniques has been sufficiently stable to survive oral administration. Peptides can be engineered for stability by altering the amino acid sequence to eliminate vulnerable sites, cyclizing[8] or stapling[9] the backbone to impose conformational constrains that hinder protease cleavage, grafting binding sequences onto highly stable protein scaffolds such as cysteine-knots[10], or by formulating the peptides[11]. While these approaches can yield substantial improvements, the engineered peptides are typically not sufficiently stable for gastrointestinal conditions. Instead of engineering peptide stability after the fact, a possible approach for discovering stable peptides is based on proteolytic phage display in which genetically encoded libraries are exposed to proteases, and only intact clones are identified by affinity selection[12–14]. In all studies that applied this principle, the proteolytic pressure used was much lower than that in the gastrointestinal tract, and the stabilities of the isolated peptides were far too low to survive oral administration. For example, we previously applied proteolytic phage display to isolate peptides constrained by a chemical linker that had a substantially enhanced stability[15]—although they were still not able to survive gastrointestinal conditions—with a half-life of only 30 min in 500-fold diluted mouse intestinal fluid. There are two probable reasons for this: (1) the chemically modified peptides were only semi-constrained, with inadequate numbers of protease-resistant library members to generate a sufficient sequence space from which to sample target-binding ligands, and (2) the phage particles themselves were rapidly degraded in relevant concentrations of intestinal fluid[15]; this instability was due to a mutant p3 coat protein lacking six cysteines, mutated to accommodate the chemical-modification technology[16].

In a recent study developing double-bridged peptides, conformationally constrained by two chemical linkers instead of only one, we discovered a peptide with a particularly high stability in human plasma[17]. In this peptide (sequence: RCCQGCRVLCY), Cys2 and 6 and Cys3 and 10 are bridged by chemical linkers, forming a tightly knotted structure that can explain the unusually high protease resistance. We speculated that screening full libraries of these doubled-bridged peptides would confer the additional constraints necessary to resist protease degradation during selection while still retaining sufficient sampling space to identify target binders—placing a library of peptides that could be administered orally within reach. To this end, we first developed a procedure to encode the chemical modifications required for double-bridged peptides without the need for the six-cysteine mutant coat protein to ensure proteolytically stable phage. Using these phage as the library backbone, we isolated target-specific peptides with nanomolar affinities that resisted proteases in undiluted simulated intestinal fluid (SIF). We also showed that they remain active in the gastrointestinal

[1]Institute of Chemical Sciences and Engineering, School of Basic Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. [2]Protein Production and Structure Core Facility, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. [3]Laboratory for Biomolecular Modeling, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL) and Swiss Institute of Bioinformatics, Lausanne, Switzerland. ✉e-mail: christian.heinis@epfl.ch

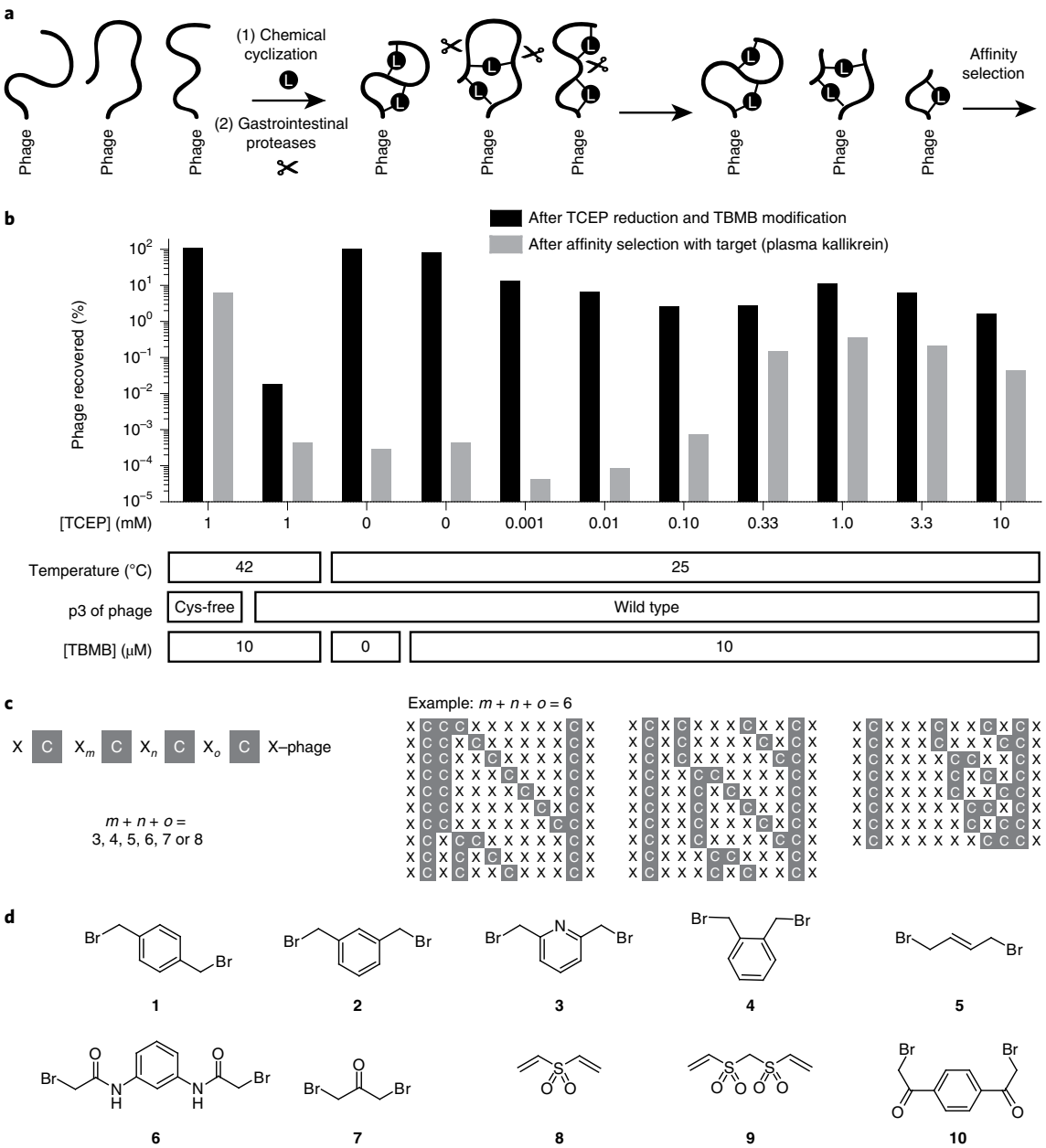


Fig. 1 | Genetic encoding of double-bridged peptides by protease-resistant phage. a, Phage display selection strategy. 'L' indicates chemical linkers that cross-link pairs of cysteines. Scissors represent gastrointestinal proteases. **b**, Test of different reaction conditions for cyclizing the peptide PK15 displayed on protease-resistant wild-type phage. The per cent of intact phage was estimated on the basis of the residual t.u. Correct peptide cyclization was assessed by quantifying the number of phage captured on immobilized plasma kallikrein to which PK15 cyclized by the alkylating agent TBMB binds. **c**, Format of phage peptide library. **d**, Bis-electrophilic cyclization reagents for the generation of double-bridged peptides.

tract of mice on oral administration and were able to somewhat penetrate the epithelial layer to mark the beginnings of orally available peptides identified from in vitro display. Analysis of the structure and study of the chemical bridges revealed that both the particular peptide format and the selected amino acid sequences are essential for the high stability, indicating the need for these constrained, protease-resistant libraries in identification of target-binding ligands.

Results

Encoding double-bridged peptides by protease-resistant phage. The in vitro evolution of proteolytically stable peptides (Fig. 1a) requires phage particles that resist gastrointestinal-relevant protease concentrations. The highest proteolytic pressure for peptides is encountered in the upper region of the small intestine, which contains high concentrations of trypsin, chymotrypsin and smaller quantities of other proteases, including exopeptidases[18]; this pressure is mimicked using SIF made up with 10 mg ml^{-1} of porcine pancreatin in phosphate buffer. As indicated above, the disulfide-free phage previously applied for the selection of chemically constrained peptides function in 500-fold diluted SIF (20 μg ml^{-1} pancreatin)[15], but are rapidly degraded in undiluted SIF. Wild-type fd bacteriophage resist undiluted SIF for at least 30 min at 37 °C, but contain six cysteines in the p3 phage coat protein (three disulfide bridges) that can interfere with the chemical cyclization reaction applied to displayed peptides[19]. Toward the use of wild-type phage, we

searched for thiol-reducing conditions that could reduce the cysteines in the displayed peptides to correctly cyclize them while leaving the six disulfide-forming cysteines in the wild-type p3 coat protein untouched. To follow the cyclization efficiency, we cloned a peptide onto the wild-type p3 that would function only when the correct reduction took place. To do this, we used a previously developed constrained peptide PK15 (ACSDRFRNCPADEALCG) that binds to plasma kallikrein only when cyclized via the three cysteines with 1,3,5-tris(bromomethyl)benzene (TBMB) (inhibition constant, $K_i = 1.5$ nM) but not in its linear or disulfide-cyclized form[19] ($K_i > 100$ μM, Fig. 1b and Supplementary Fig. 1). Wild-type phage displaying PK15 were first incubated with different concentrations (0 to 10 mM) of Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) to reduce the three peptide cysteines and then cyclized with TBMB using known conditions[19]. The reduction reaction was performed at 25 °C, a lower temperature than previously[19] (42 °C) to disfavour p3 unfolding and thus to prevent exposition of the coat protein's disulfide bridges. The number of phage that remained functional were measured by quantifying the transducing units (t.u.), and the number of phage displaying a correctly cyclized PK15 peptide were assessed by measuring the number that could be captured on immobilized plasma kallikrein (Fig. 1b). The TCEP reduction and TBMB cyclization reduced the number of infective phage by around tenfold for most conditions. The percentage of phage captured relative to the input increased with higher reducing agent concentrations and plateaued at 0.33 mM of TCEP. The percentage of phage captured maximally was around 0.4% of the input and thus lower than that achieved with the disulfide-free phage (around 6%). Nevertheless, the absolute number of wild-type phage was higher, because nearly all wild-type phage particles can infect bacterial cells, whereas less than 1% of disulfide-free phage particles are infective[16]. In a final test, we verified that wild-type phage treated with the optimized TCEP concentration (1 mM) and cyclization linker (20 μM) fully resisted the physiological pancreatin concentration (100% SIF; Supplementary Fig. 2a).

We expected that we would need a high diversity to find target-binding peptides that resist gastrointestinal conditions, so we generated the following particularly large phage display library. We displayed random peptides of the form $XCX_mCX_nCX_oCX$ ($m+n+o = 3, 4, 5, 6, 7$ or 8; where X indicates random amino acids) on phage, which gave a total of 155 different cysteine-spacing formats (Fig. 1c) and 2×10^{10} different peptide sequences. This library was cyclized with the bifunctional reagents (**1**) to (**10**) (Fig. 1d) using suitable concentrations (Supplementary Fig. 2b), and the four cysteines could pair in three different ways per peptide sequence to generate as many as 6×10^{11} double-bridged peptides.

Enrichment of target-specific, protease-resistant peptides. The sub-libraries cyclized with the 10 linkers were separately incubated with 0.1 mg ml^{-1} of porcine pancreatin (1% SIF) for 30 min at 37 °C, which was 100-fold lower than the protease pressure in the intestine and thus rather mild for the initial rounds of panning. The protease-digested library was subsequently panned against 0.5 μg of immobilized coagulation Factor XIa (FXIa) as a target (Fig. 2a). FXIa is an attractive therapeutic target because inhibition of the protease in animals prevents thrombosis without causing bleeding, which is a limitation of current anticoagulants[20]. After two and three rounds of phage selection, the DNA of 0.5 to 2×10^6 phage from each selection was sequenced, and the peptide sequences were compared using an alignment tool[21]. Consensus groups, indicative of the isolation of target-specific peptides, were identified in selections with nine of the ten linkers (Supplementary Fig. 3). The selection with linker (**7**) yielded particularly strong consensus groups defined by a large number of similar but different peptides per group (Fig. 2b). The most abundant peptides of two groups, F1 and F2, were synthesized, and both inhibited FXIa, showing that the affinity selections yielded target-specific ligands (K_i of 190 ± 30 nM and 1.0 ± 0.2 μM, respectively; Supplementary Fig. 4).

After the first three rounds of panning, the peptide diversity remained high, as seen by hundreds of different peptides that had remained in the pool and shared similar but different sequences (Fig. 2b,c). To enrich those peptides with the highest binding affinity, we performed a fourth round, reducing the amount of immobilized target protein (100, 20 or 4 ng) and holding the pancreatin pressure at 0.1 mg ml^{-1} (Fig. 2a). The number of phage captured with the smallest amount, 4 ng, was higher than in the control selection without target, indicating that the selection stringency could be further increased, and we therefore reduced the amount of immobilized target to 1 ng in the fifth and sixth selection rounds. In these latter two rounds, we performed selections using either 0, 0.1 or 1 mg ml^{-1} pancreatin in parallel, the highest of these concentrations being tenfold higher than in rounds 1–4 and representing 10% SIF (Fig. 2a). DNA sequencing of clones isolated in rounds 5 and 6 showed lower diversity and a marked change in the consensus groups (Supplementary Fig. 5), as illustrated in Fig. 2b for peptides isolated in round 6 with linker (**7**) under high proteolytic pressure. An analysis of the most abundant peptides showed that they were enriched to different extents at the three protease pressure conditions, as shown clearly for peptide F3, which was more enriched at 1 mg ml^{-1} pancreatin concentration (10% SIF) than at a tenfold lower concentration or no pancreatin (Fig. 2c; F3 is shown in red).

We synthesized several double-bridged peptides isolated in round 6 under the highest protease pressure (peptides F3–F12) along with two peptides selected in round 6 without pancreatin (F13 and F14) and two peptides from round 2 (F1 and F2), and we compared their binding affinities and stabilities (Fig. 2d). The linear peptides were cyclized with an excess of (**7**), the three isomers were separated by high-performance liquid chromatography, and the binding affinity was measured in FXIa inhibition assays (Supplementary Figs. 6 and 7). For all peptides, one of the isomers was far more active than the other two, as expected for phage-selected double-bridged peptides[17]. The stability of the most active isomer of each peptide was assessed by incubation in 10 mg ml^{-1} pancreatin (100% SIF) for different time periods at 37 °C and quantification of the intact peptide by liquid chromatography with mass spectrometry (LC–MS). A clear correlation between the selection pressure applied in the phage selections and the affinity and stability was found, as shown in Fig. 2d. Peptides isolated under high proteolytic pressure (blue) were more stable than those isolated in round 6 without pancreatin (light blue) or in round 2 (grey). In addition, peptides from round 6 had higher binding affinities than the two that were most abundant in round 2, showing that the increased pressure for binding affinity in rounds 4 to 6 enriched for high-affinity binders.

The efficient proteolytic selection became even more evident when the abundance of peptides displaying different levels of stability was retrospectively analysed for the different selection conditions of round 6 (Supplementary Fig. 8). F3, the most stable peptide, became more abundant in the selections with higher pancreatin concentration. Conversely, peptide F4, with a medium stability, was equally abundant under all three conditions, and the labile peptide F13 was diminished at higher pancreatin concentrations (Supplementary Fig. 8).

Characterization, size reduction and affinity enhancement. To determine the cysteine connectivity of the best peptides, we individually synthesized the three possible bridged isomers using orthogonal cysteine protecting groups triphenylmethyl and diphenylmethyl and measured their affinity and stability. This is exemplified for F3 in Fig. 3a–c, where it was determined that isomer 2 bridged at Cys2–Cys8 and Cys7–Cys10 inhibited FXIa with a K_i of 19 ± 2 nM, while the other two isomers showed 30- and 90-fold weaker inhibition, respectively (Fig. 3b). The three isomers had very different

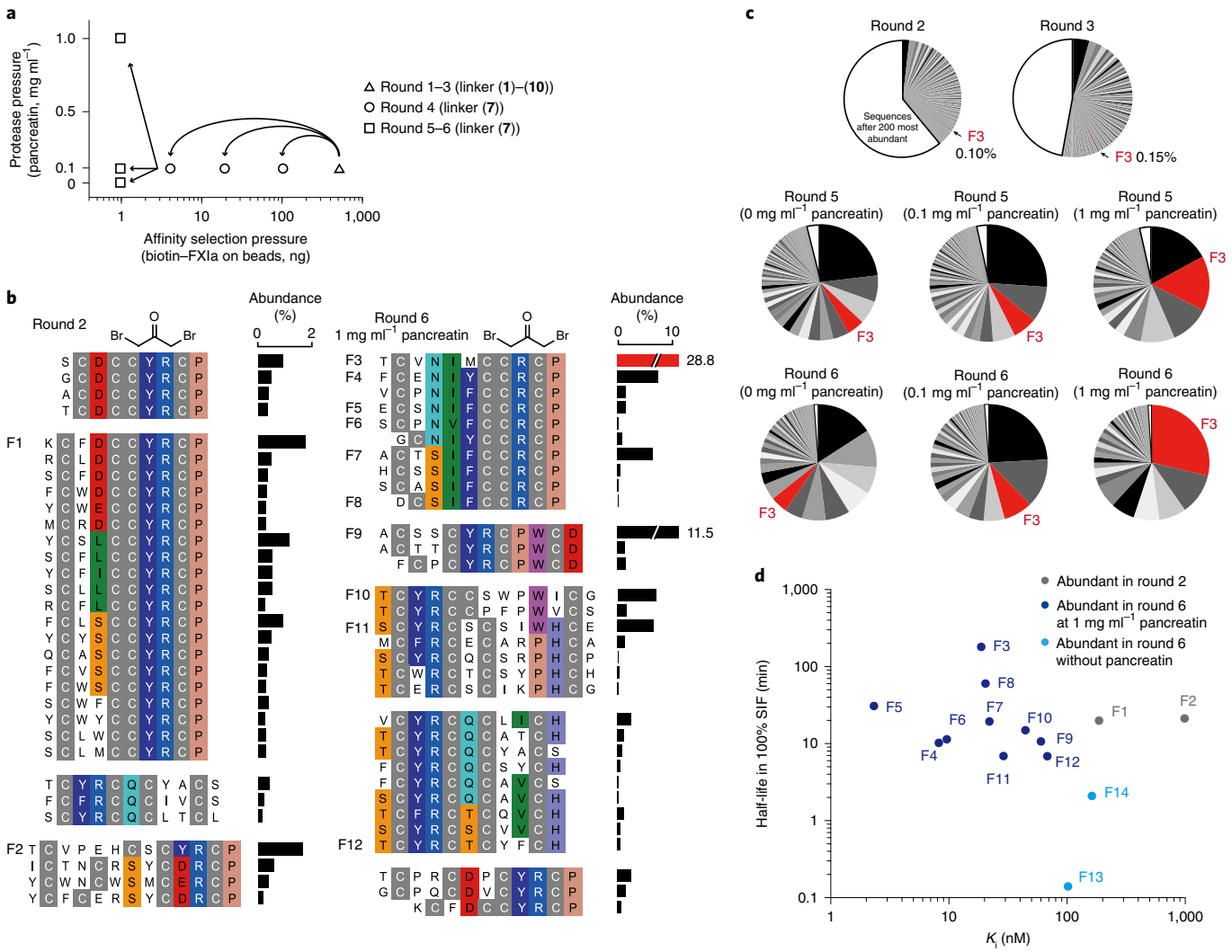
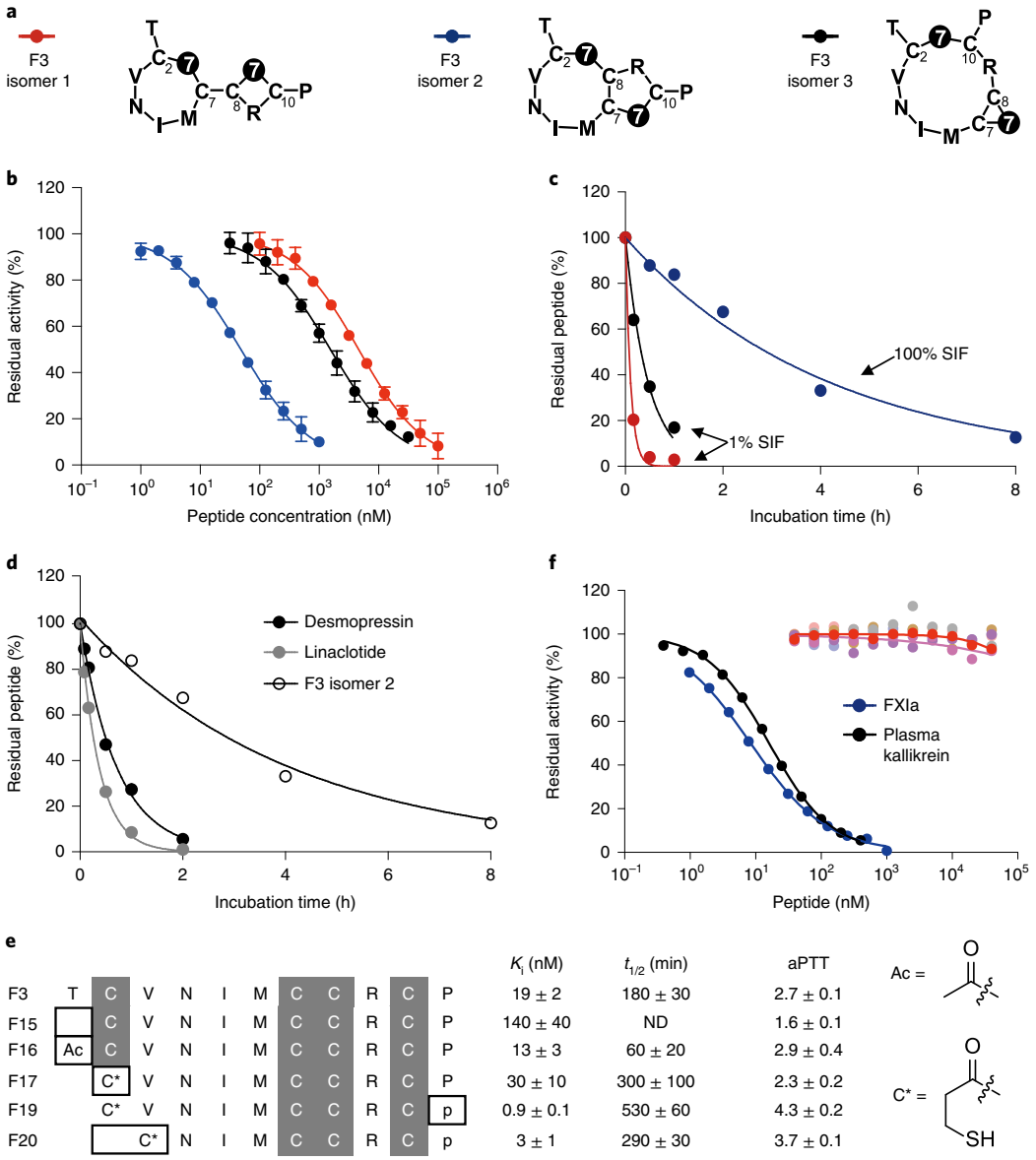


Fig. 2 | Panning of double-bridged peptide libraries under protease pressure. a, Conditions applied to control the selection stringency for binding affinity (amount of target protein FXIa) and proteolytic stability (concentration of porcine pancreatin). In rounds 4–6, 3 selections with different target concentrations (round 4) or pancreatin concentrations (rounds 5 and 6) were performed in parallel. **b**, Sequences of peptides isolated in round 2 and round 6 (1 mg ml^{-1} pancreatin condition) using cyclization reagent (**7**). Sequences isolated with other chemical linkers, in other rounds, or under other protease pressure are shown in Supplementary Figs. 3 and 5. Sequence similarities are highlighted in colour, and the abundance of each peptide is indicated. **c**, Relative abundance of the 200 most frequently found sequences from rounds 2, 3, 5 and 6. The abundance of the protease-resistant clone F3 is shown in red. **d**, Activity and stability of peptides identified in different rounds of selection and under different protease pressure. Data are shown for the most active isomers of each peptide sequence. Mean values of K_i are based on three independent measurements. Mean values of the half-lives are based on two independent measurements. The purity, K_i and $t_{1/2}$ values of the peptides are shown in Supplementary Fig. 7.

proteolytic stabilities (Fig. 3c), with the active isomer 2 demonstrating resistance to pancreatin concentrations of 10 mg ml^{-1} (100% SIF; half-life, $t_{1/2} = 180 \pm 30$ min), which was magnitudes better than the other two isomers with 41 and 9-fold shorter half-lives in 100-fold diluted pancreatin (0.1 mg ml^{-1} of pancreatin, 1% SIF). An LC–MS analysis of the two unstable isomers revealed cleavage at peptide bonds Cys8-Arg9 and Arg9-Cys10 (isomer 1) and Arg9-Cys10 and Met6-Cys7 (isomer 3; Supplementary Fig. 9). A comparison with two clinically used oral peptide drugs, desmopressin ($t_{1/2} = 30$ min) and linaclotide ($t_{1/2} = 16$ min), showed a 6- and 11-fold higher stability for F3 isomer 2 in 100% SIF (Fig. 3d).

To facilitate epithelial-layer crossing and increase the chances for oral availability, we reduced the size of the most stable peptide, F3 isomer 2. On the basis of the consensus group, the N-terminal ring was less important and truncation without much activity loss was more feasible from this end. Removing the N-terminal exocyclic amino acid (Thr1; F15) reduced the binding affinity sevenfold, but this loss was reverted when the N-terminus was acetylated (F16), suggesting that the positive charge could have caused the affinity loss (Fig. 3e). Complete elimination of the N-terminal amino group on Cys2 by replacing the cysteine with mercapto-propionic acid (Cys* or C*), eliminating the charge, yielded F17 with an activity comparable to F3 ($K_i = 30 \pm 10$ nM (F17) versus $K_i = 19 \pm 2$ nM (F3)). Introducing a Pro11→D-Pro mutation, identified in a parallel screen (F18; Supplementary Fig. 10), yielded F19 (K_i of 0.9 ± 0.1 nM), and further elimination of Val3 led to the truncated F20 (K_i of 3 ± 1 nM). F20 contained nine amino acids, counting the N-terminal amine-deficient cysteine (Cys*) and the three other cysteines, and had a molecular mass of 1,134.4 Da.

The various size-reduced peptides displayed half-lives in 100% SIF of between 60 min and several hours and efficiently blocked



		K_i (nM)	$t_{1/2}$ (min)	aPTT
F3	T C V N I M C C R C P	19 ± 2	180 ± 30	2.7 ± 0.1
F15	C V N I M C C R C P	140 ± 40	ND	1.6 ± 0.1
F16	Ac C V N I M C C R C P	13 ± 3	60 ± 20	2.9 ± 0.4
F17	C* V N I M C C R C P	30 ± 10	300 ± 100	2.3 ± 0.2
F19	C* V N I M C C R C p	0.9 ± 0.1	530 ± 60	4.3 ± 0.2
F20	C* N I M C C R C p	3 ± 1	290 ± 30	3.7 ± 0.1

Fig. 3 | Characterization and size reduction of F3, the most stable peptide. a, Chemical structures of the three isomers of F3. The linker is indicated with a black circle. **b**, Inhibition of FXIa by the three isomers. Data are mean ± s.d. of four independent measurements. **c**, Stability of F3 isomers in SIF (100% SIF used for isomer 2 and 1% SIF used for isomers 1 and 3). Mean values of two independent measurements are shown. **d**, Stability of the oral peptide drugs desmopressin and linaclotide in 100% SIF. The stability of F3 isomer 2 is shown for comparison (data from **c**). Mean values of two independent measurements are shown. **e**, Activity and stability of size-reduced derivatives of F3. The amino acids modified from peptide to peptide are highlighted with a black frame. Data are mean ± s.d. of four independent measurements for FXIa inhibition (K_i) and inhibition of the intrinsic coagulation pathway (fold prolongation of activated partial thromboplastin time (aPTT) at 30 μM peptide). Data are mean ± s.d. of three independent measurements for the half-life in 100% SIF. ND, not determined; p, D-proline; C*, mercaptopropionic acid. **f**, Specificity of F20 assessed in inhibition assays with nine trypsin-like serine proteases. The curves showing no inhibition correspond to thrombin (red), plasmin (pink), tissue plasminogen activator (purple), trypsin (grey), Factor Xa (brown), urokinase-type plasminogen activator (light blue) and Factor XIIa (light pink). Mean values of two independent measurements are shown.

activation of the intrinsic coagulation pathway, tested in human plasma ex vivo by assessing the prolongation of the activated partial thromboplastin time (Fig. 3e). A specificity profiling of proteases sharing a high sequence identity (56–84% in active site; Supplementary Fig. 11) revealed a high target selectivity for all peptides (Fig. 3f and Supplementary Fig. 12). Only plasma kallikrein, which shares the highest sequence and structure similarity with FXIa, was inhibited, although this may be an advantage in antithrombotic applications as the protease is also a coagulation factor in the intrinsic pathway[20].

Helical structure and molecular basis of proteolytic stability. The structure of the reduced-size peptide F19 bound to the FXIa catalytic domain was solved by X-ray crystallography, revealing a helical conformation (Fig. 4a; 2.9 Å resolution, Protein Data Bank (PDB) ID 6TWB; Supplementary Fig. 13 and Supplementary Tables 1 and 2). The amino acids Asn3 to Cys9 form three i → i + 3 hydrogen bonds and had φ and ψ dihedral angles characteristic of a 3_{10} helix (Supplementary Fig. 13c). The helix points towards the active site of FXIa with its C-terminal ring, sticking the Arg8 side chain into the S1 sub-site, and it covers a total surface area of 649 Å^2. Crystallization

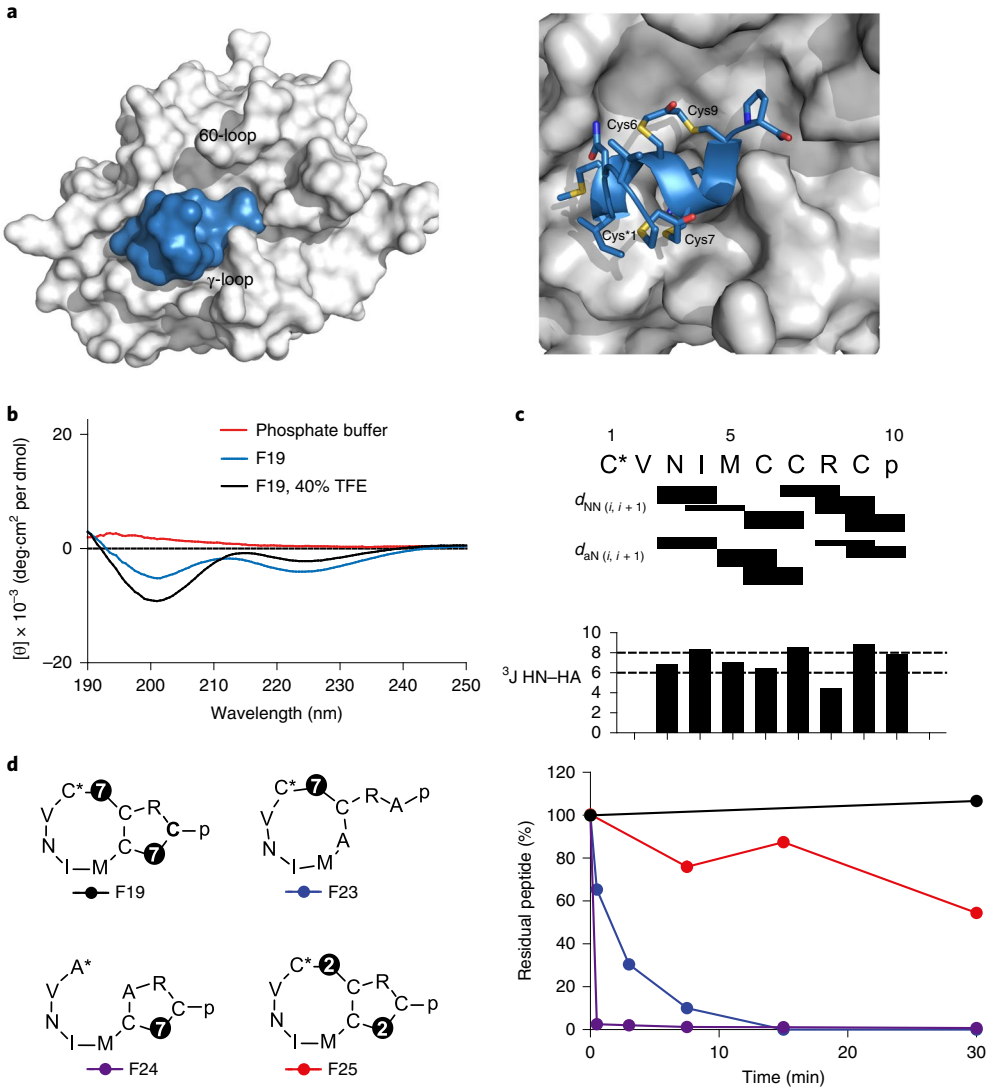


Fig. 4 | Structure analysis and impact of chemical bridges on stability. a, Crystal structure of double-bridged peptide F19 bound to the active site of FXIa (PDB ID 6TWB). Left: surface representation of the peptide (blue) and target (grey). Right: cartoon representation of the peptide. The four bridged cysteines are indicated. **b**, Circular dichroism spectra of F19 at a concentration of 200 μM recorded at 25 °C in phosphate buffer or buffer containing 40% v/v of the helix-promoting solvent trifluoroethanol (TFE). The experiment was measured three times with similar results. **c**, Top: strengths of H–H nuclear Overhauser effects in F19. Medium- or long-range signals were not observed. Bottom: ^{3}J HN–HA coupling in F19. For helical and β-sheet structures, values are typically below 6 Hz and above 8 Hz, respectively. **d**, Stability of F19 with one or two chemical bridges formed using (**7**) or with the two chemical bridges formed by (**2**) in 100% SIF. The linker is indicated with a black circle. The fraction of intact peptide is shown as a percentage. Mean values of two independent measurements are shown.

of F21, which was not reduced in size (Supplementary Fig. 14), showed the same helical structure (2.9 Å resolution, PDB ID 6TWC; details of the structure are described in the Supplementary Results).

We speculated that helix formation could hinder the access of proteases and explain the high proteolytic stability, as has been found for other peptides[9]; however, analysis of unbound F19 by circular dichroism spectroscopy (Fig. 4b) and NMR (Fig. 4c and Supplementary Fig. 15) indicated that the peptide did not form a helical structure in solution. We subsequently studied the extent to which the two chemical bridges contributed to the proteolytic stability by comparing the stability of linear, single-bridged and double-bridged peptide F19 (C*VNIMCCRCp) in 100% SIF (Fig. 4d and Supplementary Fig. 16). As expected, the linear peptide (with thiol groups removed to prevent oxidative cyclization; F22) was rapidly cleaved ($t_{1/2} < 1$ min). Peptides containing a single bridge were also much less stable than the double-bridged F19, with a half-life of 1.8 min for the Cys*1–Cys7 bridge (first ring; F23) and less than 1 min for the Cys6–Cys9 bridge peptide (second ring; F24). Of note, the double-bridged peptide cyclized with a different linker ((**2**) instead of (**7**); F25) had a half-life of around 12-fold shorter duration, suggesting that the linker structure is important for imposing the proper conformation for resisting proteases. Taken together, the peptide does not have a defined fold in solution, although the presence of the two linkers impose conformational constraints that hinder rapid proteolytic degradation.

Peptides remain intact in gastrointestinal tract of mice. The most stable peptides exhibited half-lives of more than 1 h in undiluted SIF (100% SIF), which suggested that many of them could survive in the gastrointestinal tract. We next tested the stability of the

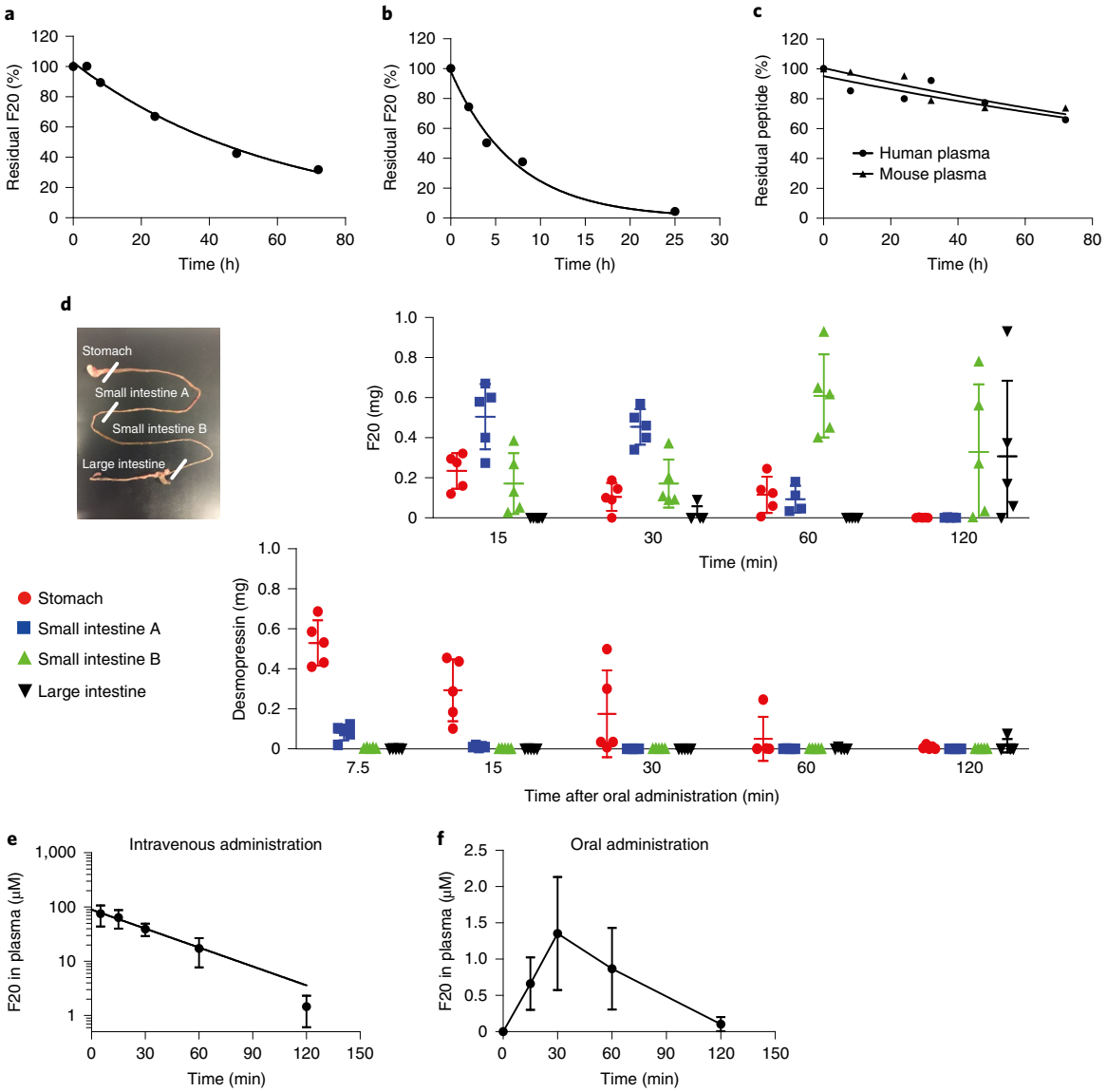


Fig. 5 | Stability and oral administration in mice. a–c, Stability of F20 in SGF (**a**), 100% SIF (**b**) and human and mouse plasma (**c**). Mean values of two independent measurements are shown. **d**, Quantity of intact F20 or desmopressin recovered from the gastrointestinal tract organs of mice administered orally with 100 mg kg^{-1} F20 or desmopressin. $n = 5$ per time point; data are mean ± s.d. The amount of peptide indicated was normalized to a mouse weighing 20 g. The small intestine was sliced into two equal pieces labelled small intestine A (includes duodenum and jejunum) and small intestine B (ileum), as indicated. **e**, Concentration of F20 in mouse plasma following intravenous administration. Dose, 10 mg kg^{-1}; $n = 5$. Data are mean ± s.d. **f**, Concentration of F20 in mouse plasma following oral administration. Dose, 100 mg kg^{-1}; $n = 5$ per time point. Data are mean ± s.d.

smallest peptide, F20, in the gastrointestinal tract of mice upon oral administration and whether it could be detected in blood, indicating an ability to traverse the epithelial layer. In ex vivo experiments, F20 had a half-life of 44 h in simulated gastric fluid (SGF), 4.8 h in 100% SIF and 141 h in mouse plasma (Fig. 5a–c). We administered 100 mg kg^{-1} F20 to 20 mice by gavage into the stomach, euthanized 5 mice after each time point (15, 30, 60 and 120 min) and analysed samples taken from the stomach, upper small intestine (duodenum and jejunum), lower small intestine (ileum), colon and blood (Fig. 5d) for the peptide by LC–MS. The relatively high dose was applied to reliably detect and quantify the peptide in case only a small fraction of it remained in the gastrointestinal tract or was absorbed into the blood stream. As a control, we performed the same experiment with the orally available peptide drug desmopressin (Fig. 5d). At all four time points analysed, between 32 and 45% of the administered F20 could be recovered from the gastrointestinal tract as fully intact peptide. At 15 and 30 min, most of the peptide was found in the upper part of the small intestine, at 60 min most of it was in the lower part of the small intestine, and at 120 min most of the peptide was in the large intestine. By contrast, the control peptide desmopressin was found intact in only small quantities, and mainly in the stomach and not in the intestines. An analysis of blood samples showed that small quantities of F20 reached the blood stream with a C_{max} (maximum plasma concentration) of 1.4 μM at 30 min and a calculated oral availability (%F) of 0.26%, which is probably too low for therapeutic use against FXIa (Fig. 5e,f).

Protease-resistant IL-23R antagonists. The result with the FXIa inhibitor, showing long peptide survival in all regions of the gastrointestinal tract and a relatively low uptake into the blood stream,

suggested that our method is particularly suited to address disease targets located in the gastrointestinal tract, such as proteins in tissues of the intestinal mucosa or submucosa. One such target is the interleukin-23 receptor (IL-23R), which has a key role in Crohn's disease and ulcerative colitis, two inflammatory conditions of the digestive tract[22]. The target has been validated using IL-23- and IL-23R-specific monoclonal antibodies, but no oral drugs are yet available. Protagonist Therapeutics and Janssen Biotech have recently revealed PTG-200 (compound C), a peptide-based IL-23R antagonist for oral administration[23,24]. The molecule was generated on the basis of linear peptide precursors that were degraded within seconds, even in strongly diluted SIF. They followed a classical peptide-engineering approach in which thousands of peptide variants containing unnatural amino acids were synthesized and activity-enhancing or stability-enhancing modifications were added in iterative rounds of engineering to reach a nanomolar affinity and high proteolytic stability[23]. Oral administration of PTG-200 to rats showed pharmacological activity, indicating that the peptides can cross the colon epithelium to reach and antagonize IL-23R. Following promising results of a phase 1 clinical trial, the two companies announced in 2019 that they would expand the PTG-200 collaboration and co-develop second-generation antagonists[25].

To demonstrate the ability of our technology to develop peptide drug candidates with higher initial proteolytic stability (before introducing unnatural amino acids), we panned our peptide phage display library cyclized with linkers (**1**) to (**8**) against human IL-23R (Supplementary Fig. 17). After three rounds without proteolytic pressure, phage were best enriched into clear consensus groups in the library cyclized with linker (**7**) (Fig. 6a), and we thus panned those isolated phage in two further rounds with proteolytic pressure. Phage were enriched at 1% SIF over background, but not at 10%, indicating that some binders survived the lower protease pressure but not the higher one. Sequencing showed that the peptides selected under protease pressure in round five converged to two consensus groups (Fig. 6a). We synthesized two peptides from this round, I1 and I2, and two peptides from round three, I3 and I4. We measured their binding to IL-23R by surface plasmon resonance (SPR) and their stability in 1% and 10% SIF. As found with the FXIa inhibitors above, peptides enriched under proteolytic pressure displayed a higher stability (Fig. 6b). The two peptides I1 and I2 had half-lives in 1% SIF of 23 min and 6.7 min, respectively, and thus a much higher stability than peptides that served as a starting point for the engineering of the Protagonist–Janssen compound. Synthesis of the three regioisomers of I1 and characterization revealed that the peptide bridged at Cys2–Cys11 and Cys4–Cys7 was the active isomer (isomer 3, termed I1; Fig. 6c). A competition experiment using SPR revealed that I1 competes with IL-23 for binding to IL-23R, thus indicating that the peptide binds to a region that is suited for therapeutic intervention (Fig. 6d).

Analysis of the degradation pathway of I1 showed that a peptide bond in the Trp-Trp-Leu loop connecting Cys7 and Cys11 was cleaved first, followed by a cascade of cleavage events (Supplementary Fig. 18). Replacing Leu10 with α-methyl-Leu hindered this ring opening and rendered the double-bridged peptide highly stable ($t_{1/2}$ in 100% SIF = 4.0 ± 0.9 h) while conserving the binding affinity (K_d = 210 ± 10 nM; Supplementary Fig. 19). This result indicated that peptide I1 had one major vulnerable site, and protecting it in the context of the double-bridge backbone topology was sufficient to generate a peptide, I5, that resists 100% SIF with a half-life of several hours. This is in contrast to the monocyclic Protagonist–Janssen peptide that required synthesis of more than one thousand peptide variants[23], inserting a total of three unnatural amino acids, an N-terminal acetyl group and a C-terminal amide group to reach a stability sufficiently high to attempt clinical trials. With its high binding affinity and good stability, I5 provides a starting point for developing an oral treatment of inflammatory disorders such as Crohn's disease on the basis of IL-23R inhibition.

Discussion

Despite decades of effort, the oral delivery of peptide-based drugs remains a major pharmaceutical challenge. Currently, biological drugs are generally delivered by intravenous or subcutaneous injection, which is effective but not desirable for patients, particularly for those with chronic conditions. Drugs based on peptides, being smaller than proteins and antibodies, have been the most amenable to oral delivery, but translating an existing bioactive peptide into an oral drug is extremely difficult. While the stability of peptides can be improved, typically by the stepwise introduction of stability-enhancing modifications, the process is cumbersome. For many peptides, the sufficiently high stability required to survive the gastrointestinal tract still cannot be reached. In this study, we report a strategy and a first example of a de novo generated peptide that binds a therapeutic target with high affinity and resists proteases in all regions of the gastrointestinal tract of a mouse.

We generated gastrointestinal-system-stable peptides by combining conformationally constrained double-bridged peptides with a phage display-based selection procedure with the aim of fishing protease-resistant peptides out of large combinatorial libraries. The combination of the double-bridged peptide format and proteolytic phage display was technically challenging, as chemically modifiable peptide libraries had to be encoded on protease-resistant phage, which was managed using conditions that specifically reduced and alkylated the cysteines in displayed peptides without affecting the three disulfide bridges in the wild-type phage p3 coat protein. This procedure did in fact enrich for target-specific peptides with a high protease stability, indicating that the library format still provided sufficient sequence space for discovering target binders after protease digestion.

Analysis of the vast sequence data and the enrichment factors during the extensive biopanning rounds revealed that some peptide formats, characterized by specific cysteine-spacing patterns and consensus sequences, are better suited for generating protease-stable peptides. This observation is in line with the general finding that some cyclic peptides derived from nature or in vitro selections can more easily be improved in terms of stability, whereas others represent evolutionary dead ends that seem to be intrinsically unstable. This suggests that peptides for oral administration may be best developed by simultaneously selecting for stability and target affinity rather than first selecting binders that are subsequently modified to improve stability.

The double-bridged peptide that best resisted pancreatin proteolysis in vitro also remained largely intact in all segments of the gastrointestinal tract of mice, including the highly acidic stomach and the small intestine, which exerts enormous proteolytic pressure. While many studies with orally applied peptides or proteins use specific formulations to prevent exposure to the acidic stomach fluid or co-apply protease inhibitors to counter the activity of pancreatic proteases, we gavaged the 'naked' peptide dissolved in PBS, enabling assessment of the stability independent of formulations. The oral peptide drug desmopressin, used as a positive control, was degraded rapidly in the intestine, underscoring the high stability achieved with the approach. An analysis of blood samples showed that less than 1% of the double-bridged peptide reached the blood stream, which was not surprising given the large polar surface of the peptide, which even contains a positive charge. Nevertheless, some of the peptide remained intact long enough to reach the blood stream and be detected, meaning that better oral availability could be reached in the future for other peptides with a smaller polar surface; this could be achieved by removing the charges and/or N-methylating some of the peptide bonds.

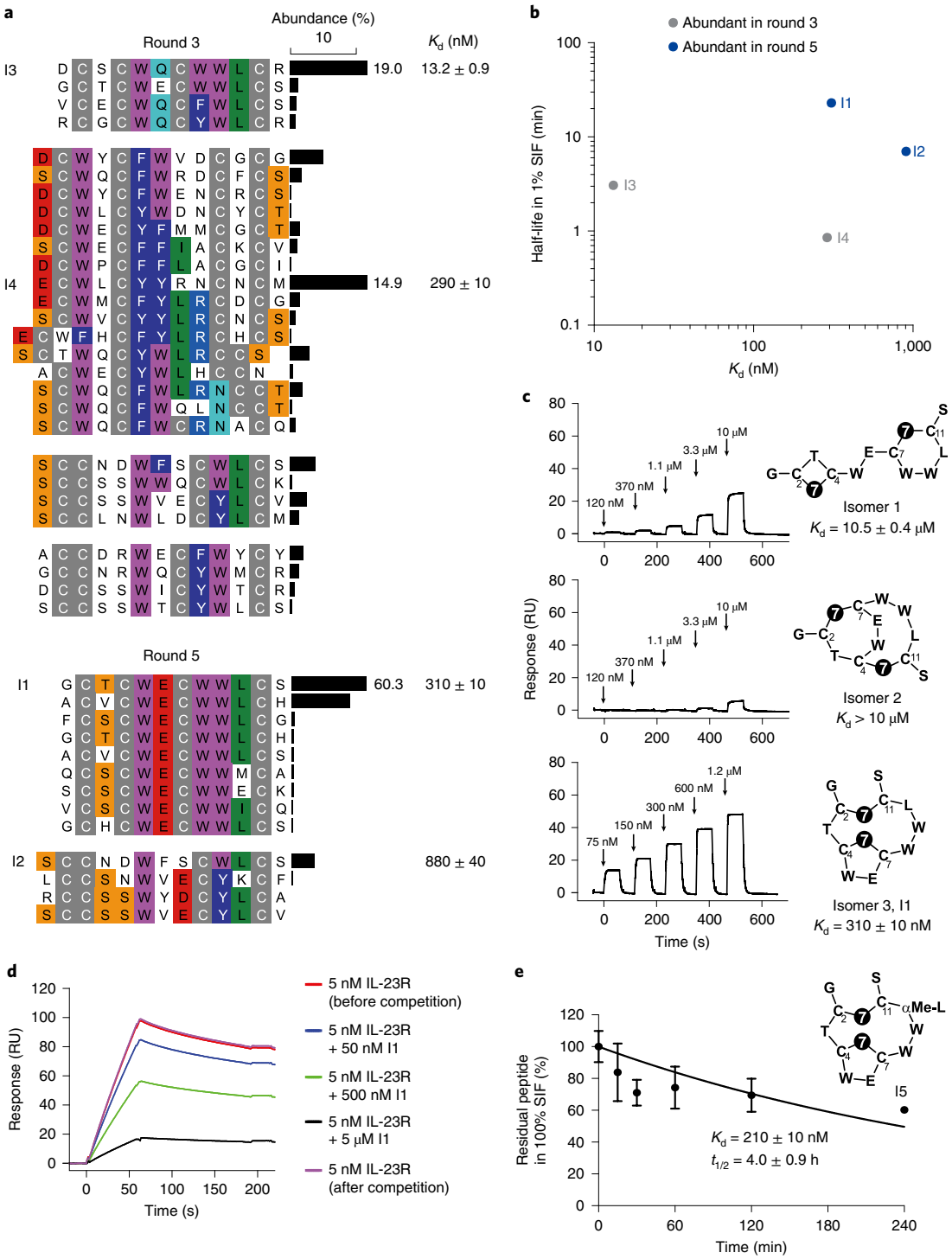


Fig. 6 | Stable IL-23R antagonists. a, Peptides enriched after three phage selection rounds without protease pressure (round 3) and after two additional rounds with 1% SIF exposure (round 5). Data are mean ± s.d. of the K_d for the active isomers based on three independent measurements. **b**, Binding affinity and half-life in 1% SIF of the two most abundant peptides identified in rounds 3 and 5. Data are the mean from two independent measurements. **c**, Single-cycle SPR sensorgrams for the three isomers of peptide I1. Data are mean ± s.d. of the K_d from three independent measurements. RU, response unit. **d**, Competition of immobilized IL-23 and peptide I1 for binding to IL-23R. The experiment was repeated three times independently with similar results. **e**, Residual intact peptide I5 after incubation in 100% SIF, determined by LC–MS. K_d values are mean ± s.d. from three independent measurements. Residual peptide quantities and $t_{1/2}$ values are mean ± s.d. from four independent measurements. In **c**,**e**, the linker is indicated with a black circle.

We have established a general approach for generating peptides from scratch, to bind almost any desired target, that could simultaneously bind the target of interest and resist proteolytic degradation in the gastrointestinal tract. While further work is required to achieve oral peptides that more efficiently reach the blood stream, the approach presented in this study can readily be used to develop oral peptide drugs that act locally on targets in the lumen of the gastrointestinal tract, in the same fashion as the recently approved,

nature-derived peptide drug linaclotide, which acts on a target in the lumen of the gastrointestinal tract.

Methods

Quantifying peptide cyclization efficiency on phage. Phage particles that display the peptide sequence of the plasma kallikrein-specific bicyclic peptide PK15 (ACSDRFRNCPADEALCG)[19] as a fusion of wild-type p3 coat protein were used to assess the reaction conditions that efficiently cyclize the peptide in the presence of the p3 disulfide bridges. A phage vector, named fd-tet-PK15, for producing these phage in *Escherichia coli* cells was cloned as described below. Disulfide-free phage displaying the PK15 peptide were used as a positive control in this experiment. Phage were produced in 0.5 l 2YT medium containing 10 μg ml^{-1} tetracycline (for fd-tet-PK15) or 30 μg ml^{-1} chloramphenicol (for fdg3p0ss21-PK15) that were inoculated with glycerol stocks of TG1 *E. coli* cells carrying either of the two plasmids, and grown at 30 °C overnight. The phage were purified by polyethylene glycol (PEG) precipitation and dissolved in reaction buffer (20 mM NH_4HCO_3, 5 mM EDTA, pH 8.0). From a 0.5 l culture, 1×10^{11} t.u. fdg3p0ss21-PK15 and 1×10^{13} t.u. fd-tet-PK15 were typically obtained. Phage were diluted with reaction buffer and incubated with various TCEP concentrations at 25 or 42 °C for 30 min in 1 ml reactions to reach final concentrations of 2×10^{9} t.u. disulfide-free phage or 2×10^{10} t.u. wild-type phage. Phage were then precipitated by addition of 0.25× volume of PEG solution (20% (w/v) PEG 6000 and 2.5 M NaCl), the pellet was dissolved in 0.9 ml reaction buffer and a small sample was removed for phage titre determination. TBMB in 0.1 ml of acetonitrile was added to reach a final concentration of 10 μM TBMB and 10% acetonitrile, and the phage reaction was incubated at 30 °C for 1 h. The phage were purified by PEG precipitation and dissolved in 0.5 ml of washing buffer (10 mM Tris-Cl, pH 7.4, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM $CaCl_2$) containing 1% (w/v) BSA and 0.1% (v/v) Tween-20. A sample was kept for titre determination and the remaining phage were incubated with 0.3 μg biotinylated plasma kallikrein (3 pmol) on 5 μl streptavidin beads (6.5–9.0 pmol μl^{-1} biotin binding sites) for 2.5 h. The beads were washed 3 times with 0.5 ml of washing buffer and 3 times with 0.5 ml of washing buffer containing 0.1% (v/v) Tween-20. The phage were eluted by incubation for 5 min in 0.2 ml of glycine buffer (20 mM, pH 2.2), the solution was transferred to a new tube, 0.2 ml of Tris-Cl buffer (1 M, pH 8.0) was added to neutralize the pH, and the titre of all samples was determined by *E. coli* infection and plating of dilutions on 2YT agar plates containing chloramphenicol (for fdg3p0ss21-PK15; 30 μg ml^{-1}) or tetracycline (for fd-tet-PK15; 10 μg ml^{-1}).

Displaying peptide PK15 on wild-type phage. The phage vector fd-tet-PK15 was cloned to produce phage that display the PK15 peptide as a fusion with the wild-type p3. The vector was constructed by inserting DNA coding for PK15 into the vector fd-tet using a whole-plasmid PCR strategy. The following primers were used, where PK15_forward anneals in the *p3* gene (residues in italic) and appends the PK15 DNA (residues in bold), and fd_reverse anneals in the leader sequence. In a 50 μl PCR reaction, PK15_forward (200 nM final concentration), fd_reverse (200 nM, final concentration), dNTP mix (200 μM each, final concentration), 20 ng of phage vector fd-tet as template, 10 μl of 5× HF Phusion buffer and two units of Phusion polymerase (Thermo Fisher) were used. The reaction was mixed by pipetting and then immediately incubated in a thermocycler with the following program: initial denaturation for 2 min at 95 °C, 25 cycles of 30 s at 95 °C, 45 s at 40 °C and 9 min at 72 °C, and final elongation for 7 min at 72 °C. The primer sequences are: PK15_forward, 5′-CCTTTCTATGCGGCCCAGCCGGCCATGG CA**GCATGTAGCGATCGTTTTCGTAATTGTCCGGCAGAT GAAGCAC TGTGTGGT**GGTTCTGGC*GCTGAAACTGTTGAAAGTTG*-3′; and fd_reverse, 5'-AGGCCGGCTGGGCCGCATAGAAAGGAAC-3'.

The PCR product was electrophoresed on an agarose gel, the DNA of the correct band was extracted, digested with SfiI (the cleavage site is underlined in the primers shown above) and ligated with T4 ligase. The ligation mixture was incubated at 70 °C for 10 min to inactivate the T4 ligase and transformed in *E. coli* DH5α by electroporation. The sequence of constructed phage vector fd-tet-PK15 was validated by Sanger sequencing.

Assessing maximal tolerated concentrations of cyclization reagent. Potential detrimental effects of the thiol-reactive cyclization reagents on phage were tested by incubating 10^{9} t.u. of phage with the chemical reagents and measuring the number of infective phage. All chemical reagents (linkers (**1**) to (**10**)) used in this study were commercially available. Phage were reduced by incubation with 1 mM of TCEP at 25 °C for 30 min in reaction buffer followed by precipitation of the phage with 0.25× volume of 20% (w/v) PEG 6000, 2.5 M NaCl to remove the excess of TCEP. Ten microlitres of thiol-reactive reagents at concentrations ranging from 0.1 to 6.4 mM in acetonitrile were added to 90 μl of 10^{9} t.u. phage in reaction buffer and incubated for 1 h at 30 °C. The number of infective phage was quantified by mixing 180 μl of exponentially growing *E. coli* TG1 cells with 20 μl of tenfold dilutions of the phage and plating the cells on tetracycline 2YT agar plates (10 μg ml^{-1}).

Assessing maximal tolerated concentrations of pancreatin. The resistance of phage particles (with or without chemical modification) to protease degradation was tested by incubating phage with porcine pancreatin solution and measuring the number of infective phage. Disulfide-free phage or wild-type phage displaying peptide PK15 (5×10^{9} t.u.) were treated with 0 or 1 mM of TCEP at 25 °C for 30 min in reaction buffer followed by precipitation of the phage with 0.25× volume of 20% (w/v) PEG 6000 containing 2.5 M of NaCl to remove the excess of TCEP. Next, 50 μl of linker (**1**) at concentration of 0 or 200 μM in acetonitrile were added to 450 μl of phage in reaction buffer, and incubated for 1 h at 30 °C. Phage were precipitated again by adding 125 μl of 20% (w/v) PEG 6000 containing 2.5 M of NaCl to remove the excess linker and dissolved in 500 μl of PBS buffer (pH 7.4). Then, 50 μl of this phage solution was added to 50 μl of porcine pancreatin in PBS buffer (pH 7.4) to reach final pancreatin concentrations of 0, 0.001, 0.01, 0.1, 1 and 10 mg/ml, and thus SIFs of 0, 0.01, 0.1, 1, 10 and 100%. The solutions were incubated at 37 °C for 30 min and the number of infective phage was quantified by mixing 180 μl of exponentially growing *E. coli* TG1 cells with 20 μl of tenfold dilutions of the phage and plating the cells on chloramphenicol (for disulfide-free phage) or tetracycline (for wild-type phage) 2YT agar plates.

Cloning of phage peptide libraries. DNA encoding the displayed random four-cysteine peptides of the form $XCX_mCX_nCX_oCX$ (where X is any amino acid, C is cysteine and *m*, *n* and *o* are numbers of random amino acids) were cloned into the phage vector fd-tet upstream of wild-type *p3* by PCR amplification of the entire plasmid using degenerate primers and subsequent connection of the PCR product ends. The numbers of amino acids *m*, *n* and *o* are equal or larger than 0, and $m+n+o$ was between 3 and 8 for the libraries XD9 to XD14. The phage vector fd-tet-PK15 was used as template for the PCR reaction. DNA encoding the peptides was appended to the plasmid DNA through forward primers that annealed in the N-terminal region of *p3*. The reverse primer annealed in the region coding for the *p3* leader sequence. All primers contained a SfiI cleavage site at the 5′ end, which allowed for the generation of complementary sticky ends for efficient DNA circularization by T4 ligase. The ligated DNA was purified and transformed into electrocompetent TG1 cells, plated on 2YT tetracycline plates (10 μg ml^{-1}), and incubated overnight at 37 °C. The cells on the plates were recovered in 2YT medium containing 10% (v/v) glycerol and stored at −80 °C. The number of transformed cells was estimated on the basis of dilutions and the total number for all libraries exceeded 2×10^{10}.

Cyclization of peptide phage display libraries. For the first round of selection, the six phage sub-libraries XD9 to XD14 were produced separately in 0.5 l of 2YT medium containing tetracycline (10 μg ml^{-1}), purified by PEG precipitation using 0.25× volume of 20% (w/v) PEG 6000, 2.5 M NaCl as previously described, and pooled together. PEG-purified phage (typically around 10^{14} t.u.) were reduced in 45 ml of reaction buffer with 1 mM TCEP at 25 °C for 30 min. The phage solution was chilled on ice, precipitated by adding 0.25× volume of ice-cold 20% (w/v) PEG 6000, 2.5 M NaCl, and centrifuged at 5,000*g* for 30 min at 4 °C. The supernatant was discarded and the phage pellets were resuspended in 50 ml of reaction buffer and distributed into 10 aliquots of 4.5 ml. The peptides in each aliquot were chemically modified by adding 500 μl of the chemical reagents (**1**) to (**10**) in acetonitrile (to reach final concentrations of 20 μM for (**1**), (**2**) and (**4**), and 40 μM for (**3**), (**5**), (**6**), (**7**), (**8**), (**9**) and (**10**)) and incubation at 30 °C for 1 h. The phage particles were precipitated by adding 1.25 ml of ice-cold 20% (w/v) PEG 6000, 2.5 M NaCl and centrifuged at 5,000*g* for 30 min at 4 °C. The supernatant was discarded and the phage pellets were resuspended in 4.5 ml of PBS (10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4, 137 mM NaCl, 2.7 mM KCl). For the selection rounds 2 to 6, phage were produced, purified and cyclized following the same procedure, but with the following modifications. Phage of each selection were produced in 25 ml 2YT culture, yielding 10^{11}–10^{12} t.u. for each selection. Phage were cyclized in volumes of 1 ml using the same solvent and reagent concentrations and resuspended in 0.45 ml PBS.

Proteolytic phage selection. For the first round of selection against FXIa, phage in 4.5 ml PBS were mixed with 500 μl of 1 mg ml^{-1} porcine pancreatin (Sigma-Aldrich, P3292) in PBS (pH 7.4) to reach a condition of 1% SIF, and incubated at 37 °C for 30 min. For the following rounds, phage in 0.45 ml PBS were incubated with 50 μl of 1 mg ml^{-1} pancreatin to reach 1% SIF. In rounds 5 and 6, phage in 0.45 ml PBS were alternatively incubated without pancreatin or with 50 μl of 10 mg ml^{-1} pancreatin to reach 10% SIF. The proteases were removed by precipitating the phage with ice-cold 20% (w/v) PEG 6000, 2.5 M NaCl, centrifugation at 5,000*g* for 30 min at 4 °C and discarding the supernatant. The phage pellets were washed twice with ice-cold 4% (w/v) PEG 6000, 0.5 M NaCl to remove the residual protease and resuspended in washing buffer containing 1% (w/v) BSA and 0.1% (v/v) Tween-20. The phage were incubated for 30 min at room temperature to allow blocking of the phage with BSA and Tween-20. For the phage selection against IL-23R, no protease pressure was used in the first 3 rounds and 1% SIF was used in round 4 and round 5.

Target immobilization. Human FXIa and IL-23R were immobilized on magnetic beads by random biotinylation of amino groups and addition to streptavidin or neutravidin beads. The two types of beads were used alternatingly in the six selection rounds of phage selection to disfavour enrichment of streptavidin- or neutravidin-specific peptides. Human FXIa (Molecular innovations, HFXIA) was

biotinylated by incubation of 200 µg protein (2.5 nmol) in 0.8 ml PBS (pH 7.4) with a 20-fold molar excess of EZ-Link Sulfo-NHS–LC–Biotin (Thermo Fisher Scientific) in 5 µl DMSO (10 mM) for 1 h at room temperature. IL-23R (purified as shown in Supplementary Information) was biotinylated by incubating 79 µg of protein (2.1 nmol) in 0.2 ml of 20 mM HEPES (pH 7.4) and 150 mM NaCl with a sixfold molar excess of EZ-Link Sulfo-NHS–LC–Biotin in 1.2 µl of DMSO (10 mM) for 2 h on ice. Excess biotin was removed by chromatography using a PD-10 or Superdex75 10/300 GL column and PBS. The percentage of biotinylated FXIa and IL-23R was assessed by capturing 2 µg of biotinylated protein on 20 µl magnetic streptavidin beads (Dynal, M-280 from Life Technologies) and subsequent analysis of immobilized protein by SDS-PAGE as previously described[26]. Neutravidin beads were prepared by reacting 6 mg of neutravidin (Pierce) with 10 ml of tosyl-activated magnetic beads (Dynal, M-280 from Invitrogen) according to the supplier's instructions. For the first to third round of phage selection against FXIa, 0.5 µg of target protein immobilized on 20 µl strepavidin beads (rounds 1 and 3) or on 10 µl neutravidin beads (round 2) were used for each selection performed with one of the 10 different cyclization reagents. For round 4 of selection, either 0.1 µg, 20 ng or 4 ng of biotinylated FXIa were immobilized on 10 µl neutravidin beads. For rounds 5 and 6 of selection, 1 ng of biotinylated FXIa was immobilized on 10 µl strepavidin beads (round 5) or on 5 µl neutravidin beads (round 6). For phage selection against IL-23R, 2.5, 1.25, 0.625, 0.3 and 0.1 µg of target protein were immobilized on 20 µl of streptavidin beads (rounds 1, 3 and 5) or on 10 µl of neutravidin beads (rounds 2 and 4), and were used for the selection performed with linker (**7**). Biotinylated protein was immobilized on magnetic beads by incubation of protein and pre-washed beads in 200 µl PBS for 10 min at room temperature on a rotating wheel (10 rpm). The beads were washed twice with 1 ml of washing buffer and resuspended in 0.5 ml washing buffer.

Phage display panning. For the phage panning experiments, magnetic beads with immobilized FXIa or IL-23R were blocked by incubation with washing buffer containing 1% (w/v) BSA and 0.1% (v/v) Tween-20 for 30 min at room temperature on a rotating wheel (10 rpm). In parallel, phage displaying double-bridged peptides were blocked in the same buffer. Blocked beads with target protein were mixed with blocked phage and incubated for 30 min on a rotating wheel (10 rpm) at room temperature. Beads were washed eight times with washing buffer containing 0.1% (v/v) Tween-20 and twice with washing buffer. The phage were eluted by incubating the beads with 100 µl glycine buffer (50 mM, pH 2.2) for 5 min and transferred to 50 µl of 1 M Tris-Cl buffer, pH 8. Phage were incubated with 5 ml of exponentially growing *E. coli* TG1 cells (optical density at 600 nm, 0.4 – 0.8) for 30 min at 37 °C and the cells were plated on a 14 cm 2YT plates containing tetracycline (10 µg ml^{-1}). The next day, bacterial cells were recovered from the plates in 4 ml of 2YT medium containing 10% (v/v) glycerol, and stored at −80 °C.

Determination of peptide stability in simulated gastric and intestinal fluid. SGF was prepared according to United States Pharmacopeia specifications (Test Solutions, United States Pharmacopeia 35, NF 30, 2012). Sodium chloride (0.2 g) was added to a 100 ml flask and dissolved in 50 ml of water. A volume of 0.7 ml of 10 M HCl was added to adjust the pH of the solution to 1.2. To this, 0.32 g of pepsin (Sigma-Aldrich, P7125) was added and dissolved with gentle shaking and the volume made up to 100 ml with water. Pepsin was added only after the pH was adjusted to 1.2. SIF was prepared according to United States Pharmacopeia specifications (Test Solutions, United States Pharmacopeia 35, NF 30, 2012). Monobasic potassium phosphate (0.68 g) was dissolved in 25 ml of water. A volume of 7.7 ml of 0.2 M NaOH was added to adjust the pH to 6.8. To this, 1 g of pancreatin (Sigma-Aldrich, P3292) was added and shaked gently until dissolved and the volume was adjusted to 100 ml with water. Pancreatin was added after adjusting the pH to 6.8 to avoid precipitation of the enzyme. The stability of peptides in SGF and SIF was tested as follows. Two microlitres of peptide stock solution (2 mM in DMSO) was added to 98 µl of SGF or SIF and incubated at 37 °C. Samples (10 µl) were withdrawn at different intervals and added to ice-cold stop reagent (30 µl, methanol for gastric fluid and 0.1 M HCl for intestinal fluid) containing 10 µM of internal standard peptide to inactivate the enzymes and enable quantitative determination of intact peptide remaining. All samples were centrifuged at 10,000*g* for 10 min, and the supernatant was analysed quantitatively by LC–MS. Linaclotide used in stability assays was purchased from Medchem (HY-17584). Desmopressin was synthesized by standard solid-phase peptide synthesis.

Determination of plasma stability. Two microlitres of peptide stock solution (2 mM in DMSO) was added to 98 µl of plasma (human or mouse, Innovative Research) at 37 °C. Samples of 10 µl were taken at different time points (0 min, 4, 8, 24, 48 and 72 h) and incubated at 65 °C for 10 min to inactivate the plasma proteases. A volume of 5 µl of 7 M guanidine hydrochloride (pH 2.0) and 5 µl of internal standard peptide (200 µM in MilliQ H_2O) were added. The plasma proteins were precipitated by addition of cold ethanol (400 µl, containing 0.1% v/v trifluoroacetic acid) and incubation on ice for 60 min. The samples were centrifuged at 10,000*g* for 20 min, the supernatant was transferred to new tubes and the solvent was evaporated under vacuum (speed vac) at 50 °C. The samples were dissolved in 20 µl of H_2O containing 0.1% v/v trifluoroacetic acid and analysed by LC–MS.

Animal experiments. All experiments in mice were conducted in accordance with the terms of the Swiss animal protection law and were approved by the animal experimentation committee of the cantonal veterinary service (Canton of Vaud, Switzerland).

Determination of pharmacokinetics in mice. Female BALB/cByJ mice were injected with 10 mg kg^{-1} of F20 in 100 µl PBS (pH 7.4) through the tail vein (five mice per group). Blood samples (40 µl) were collected at different time points from the tail vein into tubes containing EDTA (K3 EDTA; Sarstedt, 20.1278), centrifuged at 4,000*g* for 10 min at 4 °C and the plasma stored at −80 °C. The samples were processed as described above and analysed by LC–MS. Peptides were quantified by extracting ion chromatograms on the basis of the peptide mass, integrating the area under the peaks of the curve, and normalization was based on calibration curves generated with the same peptides. Calibration curves were generated by analysing mouse plasma to which defined quantities of peptide were added using LC–MS.

Bioavailability and quantification of peptide in gastrointestinal tract. Peptide F20 dissolved in 200 µl PBS (10 mg ml^{-1}) was applied by oral gavage (100 mg kg^{-1}) to five mice (20 g weight; fasted for 15 h before administration) for each time point. Female BALB/cByJ mice were anaesthetized with ketamine and xylazine (100 and 10 mg kg^{-1}, respectively) administered intraperitoneally. Blood (0.5 ml) was collected by cardiac puncture and transferred in plasma tubes containing EDTA (K3 EDTA; Sarstedt, 20.1341) and centrifuged at 4,000*g* for 10 min. Plasma samples (200 µl) were mixed with 50 µl of 7 M guanidine hydrochloride (pH 2.0) and 5 µl of internal standard peptide (200 µM in MilliQ H_2O). The samples were processed as described above and analysed by LC–MS. The absolute oral bioavailability (*F*) was calculated with the dose-corrected area under the curve for the oral route (AUC_{po}) divided by the area under the curve for the intravenous route (AUC_{iv}), using equation (1), where $Dose_{po}$ and $Dose_{iv}$ are oral and intravenous doses, respectively:

$$F = 100 \times \frac{AUC_{po}/Dose_{po}}{AUC_{iv}/Dose_{iv}} \quad (1)$$

Stomach, small intestine and colon of mice were removed, and the small intestine was divided into two parts (upper part A and lower part B) and homogenized in 1 ml of organ lysis buffer (50 mM Tris-Cl, pH 8.0, 150 mM NaCl, 5 mM EDTA, 0.01% Triton X-100) containing 0.5% (v/v) protease inhibitor cocktail (Sigma, P8340). After centrifugation at 10,000*g* for 30 min at 4 °C, 15 µl of supernatant was mixed with 5 µl of 7 M guanidine hydrochloride (pH 2.0) and 5 µl of internal standard peptide (200 µM in MilliQ H_2O). The samples were analysed by LC–MS and the peptide was quantified as described above.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The main data supporting the results of this study are available within the paper and its Supplementary Information. The data used to make the graphs in the figures are provided as Supplementary Information. Next-generation sequencing source data are available from figshare at https://doi.org/10.6084/m9.figshare.11921139.v2. Coordinates of the two X-ray structures have been deposited in the Protein Data Bank with accession numbers 6TWB and 6TWC.

Received: 31 May 2019; Accepted: 7 April 2020;
Published online: 11 May 2020

References

1. Goldberg, M. & Gomez-Orellana, I. Challenges for the oral delivery of macromolecules. *Nat. Rev. Drug Discov.* **2**, 289–295 (2003).
2. Moroz, E., Matoori, S. & Leroux, J. C. Oral delivery of macromolecular drugs: where we are after almost 100 years of attempts. *Adv. Drug Deliv. Rev.* **101**, 108–121 (2016).
3. Pratley, R. et al. Oral semaglutide versus subcutaneous liraglutide and placebo in type 2 diabetes (PIONEER 4): a randomised, double-blind, phase 3a trial. *Lancet* **394**, 39–50 (2019).
4. Nielsen, D. S. et al. Orally absorbed cyclic peptides. *Chem. Rev.* **117**, 8094–8128 (2017).
5. Aguirre, T. A. S. et al. Current status of selected oral peptide technologies in advanced preclinical development and in clinical trials. *Adv. Drug Deliv. Rev.* **106**, 223–241 (2016).
6. Räder, A. F. B. et al. Orally active peptides: Is there a magic bullet? *Angew. Chem. Int. Ed.* **57**, 14414–14438 (2018).
7. Naylor, M. R., Bockus, A. T., Blanco, M. J. & Lokey, R. S. Cyclic peptide natural products chart the frontier of oral bioavailability in the pursuit of undruggable targets. *Curr. Opin. Chem. Biol.* **38**, 141–147 (2017).
8. Gentilucci, L., De Marco, R. & Cerisoli, L. Chemical modifications designed to improve peptide stability: incorporation of non-natural amino acids, pseudo-peptide bonds, and cyclization. *Curr. Pharm. Des.* **16**, 3185–3203 (2010).

9. Perry, A. F. et al. Hydrocarbon double-stapling remedies the proteolytic instability of a lengthy peptide therapeutic. *Proc. Natl Acad. Sci. USA* **107**, 14093–14098 (2010).
10. Wang, C. K. & Craik, D. J. Designing macrocyclic disulfide-rich peptides for biotechnological applications perspective. *Nat. Chem. Biol.* **14**, 417–427 (2018).
11. Banga, A. K. *Therapeutic Peptides and Proteins: Formulation, Processing, and Delivery Systems* 3rd edn (CRC Press, 2015).
12. Kristensen, P. & Winter, G. Proteolytic selection for protein folding using filamentous bacteriophages. *Fold. Des.* **3**, 321–328 (1998).
13. Sieber, V., Plückthun, A. & Schmid, F. X. Selecting proteins with improved stability by a phage-based method. *Nat. Biotechnol.* **16**, 955–960 (1998).
14. Howell, S. M. et al. Serum stable natural peptides designed by mRNA display. *Sci. Rep.* **4**, 6008 (2015).
15. Baeriswyl, V. & Heinis, C. Phage selection of cyclic peptide antagonists with increased stability toward intestinal proteases. *Protein Eng. Des. Sel.* **26**, 81–89 (2013).
16. Kather, I., Bippes, C. A. & Schmid, F. X. A stable disulfide-free gene-3-protein of phage fd generated by in vitro evolution. *J. Mol. Biol.* **354**, 666–678 (2005).
17. Kale, S. S. et al. Cyclization of peptides with two chemical bridges affords large scaffold diversities. *Nat. Chem.* **10**, 715–723 (2018).
18. Wang, J., Yadav, V., Smart, A. L., Tajiri, S. & Basit, A. W. Toward oral delivery of biopharmaceuticals: an assessment of the gastrointestinal stability of 17 peptide drugs. *Mol. Pharm.* **12**, 966–973 (2015).
19. Heinis, C., Rutherford, T., Freund, S. & Winter, G. Phage-encoded combinatorial chemical libraries based on bicyclic peptides. *Nat. Chem. Biol.* **5**, 502–507 (2009).
20. Bane, C. E. & Gailani, D. Factor XI as a target for antithrombotic therapy. *Drug Discov. Today* **19**, 1454–1458 (2014).
21. Rentero Rebollo, I., Sabisz, M., Baeriswyl, V. & Heinis, C. Identification of target-binding peptide motifs by high-throughput sequencing of phage-selected peptides. *Nucleic Acids Res.* **42**, e169 (2014).
22. Gaffen, S. L., Jain, R., Garg, A. V. & Cua, D. J. The IL-23–IL-17 immune axis: from mechanisms to therapeutic testing. *Nat. Rev. Immunol.* **14**, 585–600 (2014).
23. Bourne, G. T. et al. Oral peptide inhibitors of interleukin-23 receptor and their use to treat inflammatory bowel diseases. US patent US9624268B2 (2015).
24. Sayago, C. et al. Deciphering binding interactions of IL-23R with HDX–MS: mapping protein and macrocyclic dodecapeptide ligands. *ACS Med. Chem. Lett.* **9**, 912–916 (2018).
25. Patel, D. V. Protagonist therapeutics expands PTG-200 collaboration agreement with Janssen to include second generation oral IL-23 receptor antagonists. *Bloomberg* (8 May 2019); https://www.bloomberg.com/press-releases/2019-05-08/protagonist-therapeutics-expands-ptg-200-collaboration-agreement-with-janssen-to-include-second-generation-oral-il-23-receptor
26. Rentero Rebollo, I. & Heinis, C. Phage selection of bicyclic peptides. *Methods* **60**, 46–54 (2013).

Acknowledgements

We thank B. Mangeat from the Ecole Polytechnique Fédérale de Lausanne (EPFL) Gene Expression Core Facility for help with next-generation sequencing, A. Reynaud, D. Hacker, L. Durrer and S. Quinche from the EPFL Protein Production and Structure Core Facility for help with protein expression and crystallization, and E. Simeoni, I. Desbaillets, G. Ferrand and C. Waldvogel of the EPFL Center of Phenogenomics (CPG) for help with mouse experiments. The financial contributions from the Swiss National Science Foundation grants (project grant 157842, NCCR Chemical Biology and project grant 169526) and the EPFL are gratefully acknowledged.

Author contributions

X.-D.K. and C.H. conceived the strategy for proteolytic phage display with double-bridged peptides. X.-D.K. established the phage selection procedure, cloned the libraries, performed the phage selections against FXIa, synthesized and characterized the peptides, expressed and purified FXIa and determined the X-ray structures. J.M. expressed and purified IL-23R, performed the phage selections against IL-23R and characterized the peptides. V.C. identified the D-proline mutant. F.P. collected X-ray data and analysed the structures. L.A.A. performed the NMR study. X.-D.K., K.D. and C.H. wrote the manuscript with help from all authors.

Competing interests

J.M., X.-D.K. and C.H. are inventors on a patent protecting the IL-23R antagonists.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41551-020-0556-3.

Correspondence and requests for materials should be addressed to C.H.

Reprints and permissions information is available at www.nature.com/reprints.

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

nature research

Corresponding author(s): Christian Heinis

Last updated by author(s): Mar 27, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Illumina NGS software, Biacore 8K Control Software.
Data analysis	Matlab, Matlab scripts (published in Rentero Rebollo, I. et al., Nucleic Acids Res. 42, e169 (2014)), Prism, Pymol 2.3.2, XDS, Coot, Refmac5, Phenix, TopSpin, CARA, CSI 3.0, Biacore 8K Evaluation Software.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The main data supporting the results of this study are available within the paper and its Supplementary Information. The data used to make the graphs in the figures are provided as Supplementary Information. NGS source data are available from figshare at https://doi.org/10.6084/m9.figshare.11921139.v2. Data of the two X-ray structures have been deposited in the PDB database (6TWB, 6TWC).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Next-generation sequencing: around one million reads per sample (reason for sample size: around 10,000 to 100,000 samples were estimated to be required for a detailed analysis, but more phage were sequenced, as this was easily possible by NGS). Mouse experiments: five animals per group (reason for sample size: arbitrary based on previous experience and variability found for this animal model; the sample size was found to be suited to reach sufficiently small s.d. values).
Data exclusions	None. All attempts at replication were successful.
Replication	Ki determination with FXIa activity assay: four independent measurements. Specificity profiling with FXIa activity assay: four idependent measurements for FXIa, one measurement for FXa, two independent measurements for all other seven proteases. Stability of F3 and derivatives in 100% SIF: three independent measurements. Stability of F20 in SIF, SGF, plasma: two independent measurements. Activated partial thromboplastin time (aPTT): three independent measurements for key peptides (main text), two independent measurements for peptides shown in the Supplementary Information. Stability of F19 and derivatives in 100% SIF: two independent measurements. Affinity for IL-23R by SPR: three independent measurements. Stability of I5 in 100% SIF: four independent measurements.
Randomization	Mice were randomly assigned to groups.
Blinding	No blinding was applied because the readout was performed by instruments (peptide quantification by LC-MS).

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HEK293 and CHO cells were obtained from Thermo Fisher Scientific.
Authentication	None of the cell lines used was authenticated in the laboratory.
Mycoplasma contamination	Cells were not tested for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	BALB/cByJ female mice, 15-week-old on average (range, 11 to 20 weeks); weight: 18–25 g; supplier: Charles River.
Wild animals	The study did not involve wild animals.
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	All experiments in mice were conducted in accordance with the terms of the Swiss animal protection law and were approved by the animal experimentation committee of the cantonal veterinary service (Canton of Vaud, Switzerland).

Note that full information on the approval of the study protocol must also be provided in the manuscript.