JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

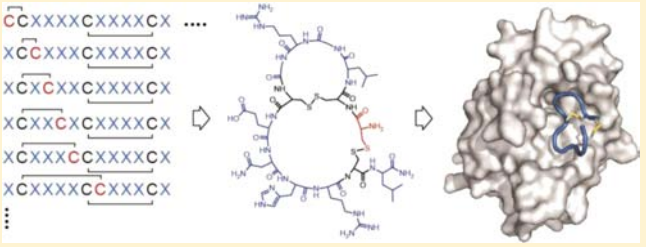
Bicyclic Peptide Ligands Pulled out of Cysteine-Rich Peptide Libraries

Shiyu Chen,[†] Inmaculada Rentero Rebollo,[†] Sergey A. Buth,[‡] Julia Morales-Sanfrutos,[†,§] Jeremy Touati,[†] Petr G. Leiman,[‡] and Christian Heinis*[,†]

[†]Institute of Chemical Sciences and Engineering and [‡]Institute of Physics and Biological Systems, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

Supporting Information

ABSTRACT: Bicyclic peptide ligands were found to have good binding affinity and target specificity. However, the method applied to generate bicyclic ligands based on phage-peptide alkylation is technically complex and limits its application to specialized laboratories. Herein, we report a method that involves a simpler and more robust procedure that additionally allows screening of structurally more diverse bicyclic peptide libraries. In brief, phage-encoded combinatorial peptide libraries of the format $X_mCX_nCX_oCX_p$ are oxidized to connect two pairs of cysteines (C). This allows the generation of $3 \times (m + n + o + p)$ different peptide topologies because the fourth cysteine can appear in any of the $(m + n + o + p)$ randomized amino acid positions (X). Panning of such libraries enriched strongly peptides with four cysteines and yielded tight binders to protein targets. X-ray structure analysis revealed an important structural role of the disulfide bridges. In summary, the presented approach offers facile access to bicyclic peptide ligands with good binding affinities.



INTRODUCTION

Cyclic peptides can bind to their protein targets with high affinity and specificity. Many nature-derived cyclic peptides or derivatives thereof are used today as therapeutics, for example, the immunosuppressant drug cyclosporine, the growth-inhibiting hormone analogue ocreotide and the antibiotic vancomycin.[1,2] A first phage-selected cyclic peptide, a mimetic of erythropoietin, was recently clinically approved.[3] Cyclic peptide ligands or inhibitors represent also a valuable research tool.[4] Connecting the amino and carboxy termini or amino acid side chains in linear peptides reduces the conformational flexibility of the peptide backbone and leads to a higher affinity due to a smaller entropic penalty upon binding as well as to a better target specificity due to a smaller number of possible conformers.[5] Additionally, the cyclization increases the proteolytic stability of peptides.[5] These favorable binding and stability characteristics are even more pronounced in polycyclic peptides in which two or more macrocyclic rings constrain the conformation of the peptide chain.[6,7]

Monocyclic peptides cyclized by a disulfide bridge are routinely isolated by phage display or other display techniques.[8,9] Many of these peptides bind with micromolar and a few with nanomolar affinity to their targets.[10,11] Ligands with nanomolar or even picomolar affinities were developed based on bicyclic peptides.[12−15] The bicyclic structures were obtained by cyclizing phage-encoded linear peptides containing three cysteine residues with tris-(bromomethyl) benzene (TBMB)[16] prior to affinity selection (Figure 1A, lower reaction).[14] Screening peptide libraries of the format CX_mCX_nC-phage (m, n = 3−6) cyclized with TBMB yielded potent antagonists of plasma kallikrein,[13,14] cathepsin G,[14] and urokinase-type plasminogen activator (uPA).[12,15] A limitation of the strategy is the technically complex procedure required to chemically cyclize the peptides. The phage need to be purified, the phage-displayed peptides reduced, and the reducing agent removed by filtration under conditions that prevent reoxidation. Finally, the cysteines need to be reacted with TBMB at a narrow window of conditions that ensures quantitative cyclization without impairing the phage functionality. Because of this delicate procedure, laborious quality checks are applied at each of these experimental steps to ensure quantitative bicyclization of the phage peptides.[17]

In biopanning experiments recently performed in our laboratory using bicyclic peptide libraries generated by cyclization with TBMB, we observed that the number of cysteine residues in the isolated peptides varied significantly. In most selections, we found mainly peptides with three cysteines, but occasionally there was a strong enrichment of peptides with four cysteines. The statistical probability that peptides in the most frequently applied 6×6 library (CX_6CX_6C) contain three and four cysteines is 68% and 26%, respectively. We speculated that peptides containing four cysteines were isolated as bicyclic peptides with two disulfide bonds being formed by air oxidation (as a result of imperfect cysteine reduction and TBMB cyclization). If this was true, phage-encoded bicyclic peptides could be generated simply by connecting two pairs of cysteines on phage. This would be unexpected as peptides with four cysteines being displayed in multiple copies on phage could easily cross-link with each other or the four cysteines could

Received: January 24, 2013
Published: April 5, 2013

   dx.doi.org/10.1021/ja400461h | J. Am. Chem. Soc. 2013, 135, 6562−6569

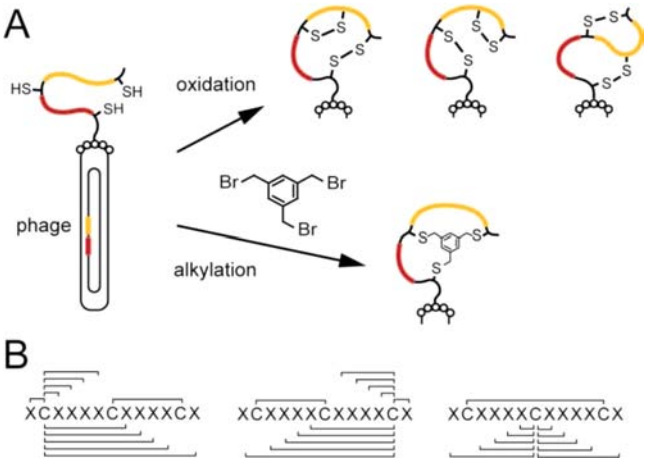


Figure 1. Generation of phage-encoded bicyclic peptides by disulfide bridge formation. (A) Peptides containing three cysteines in fixed positions and random sequences (red and yellow) are displayed on phage. In the upper reaction, two pairs of cysteines are air oxidized to form disulfide bridges. To obtain bicyclic peptides, an additional cysteine residue needs to be present in the random region of the peptide. Three different regioisomers can be formed. In the lower reaction, bicyclic peptides are formed by reacting the three cysteine residues with an alkylating reagent, such as TBMB. The single product isomer contains three thioether bonds. (B) A large number of topologically diverse bicyclic structures can be generated by placing the fourth cysteine in any of the randomized positions and by forming disulfide bridges between two pairs of cysteines.

connect differently to form regioisomers, complicating or hindering the phage selection.

In this work, we tested if bicyclic peptide ligands could indeed be isolated from libraries of random peptides containing cysteines in three fixed positions, such as the above-mentioned 6 × 6 phage display library. We reasoned that structurally highly diverse bicyclic peptides could be formed if linear peptides of the format $X_mCX_nCX_oCX_p$ are oxidized: A first disulfide bridge could form between any two of the three cysteines and a second disulfide bridge could form between the third cysteine and a fourth cysteine appearing in any of the $(m + n + o + p)$ randomized positions (Figure 1A, upper reaction, Figure 1B). In total, as many as $3 \times (m + n + o + p)$ different peptide topologies could form. Such a large structural diversity should increase the chance of finding bicyclic peptides that have shapes complementary to surfaces of protein targets and bind with high affinity. In addition, the procedure to cyclize peptides by oxidation of cysteines would be significantly easier and more reliable than cyclization by alkylation.

EXPERIMENTAL SECTION

Phage Affinity Selection of Bicyclic Peptides. Phage were produced and purified as described previously.[14] Peptides on phage were cyclized by reaction with TBMB or allowed to oxidize in buffered saline overnight. Human uPA (UPA-LMW, 33 kDa; Innovative Research, Novi, MI, USA) was biotinylated as described in the Supporting Information. Biotinylated protein (2 μg) was immobilized on 50 μL of magnetic streptavidin beads (Dynabeads M-280, Invitrogen Dynal Biotech AS, Oslo, Norway), unbounded uPA removed, and the beads blocked in 0.5 mL 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 for 30 min. For the selection of binders to streptavidin, the streptavidin beads (50 μL) were directly blocked. At the same time, the phage displaying cyclized peptides (about 10^{11} t.u. dissolved in 2 mL of washing buffer) were blocked by addition of 1 mL of washing buffer containing 3% w/v BSA and 0.3% v/v Tween 20 for 30 min. Blocked beads with antigen (0.5 mL) and phage (2 mL) were mixed together and incubated for 30 min on a rotating wheel at room temperature (rt). The beads were washed eight times with washing buffer containing 0.1% v/v Tween 20 and twice with washing buffer. Phage were eluted by incubation with 100 μL of 50 mM glycine, pH 2.2, for 5 min. Eluted phage were transferred to 50 μL of 1 M Tris-Cl, pH 8.0 for neutralization, and incubated with 50 mL of exponentially growing TG1 cells (OD_{600} = 0.4) for 90 min at 37 °C. The phage-infected cells were pelleted by centrifugation, resuspended in 2 mL 2YT medium, and plated on large 2YT/chloramphenicol (30 μg/mL chloramphenicol) plates. A second round of panning was performed following the same procedure. For the uPA selections, neutravidin-coated magnetic beads were used to immobilize the biotinylated protease. Neutravidin beads were prepared by reacting 0.8 mg of neutravidin (Pierce, Thermo Fisher Scientific, Waltham, MA, USA) with 2×10^9 tosyl-activated magnetic beads (2×10^9 beads/mL; Dynabeads M-280, Invitrogen Dynal Biotech AS) according to the supplier's instructions.

Chemical Synthesis of Bicyclic Peptides. Linear peptides were synthesized by solid-phase peptide synthesis (SPPS) as described in the Supporting Information. For cyclization by oxidation, the fully deprotected linear peptides were purified by preparative reversed-phase HPLC (RP-HPLC) using a Vydac C18 (218TP) column (22 × 250 mm) (Grace and Vydac, Hesperia, USA). A linear gradient with a mobile phase composed of eluent A (99.9% v/v H_2O, 0.1% v/v TFA) and eluent B (94.9% v/v acetonitrile, 5% v/v H_2O, and 0.1% v/v TFA) from 5% to 50% over 30 min at a flow rate of 20 mL/min was applied. The masses of peptides in different fractions were determined by MALDI-TOF mass spectrometry as described in the Supporting Information. The fractions with desired product were lyophilized. Purified peptides were oxidized by incubation in 100 mM NH_4Ac, pH 8.0, buffer containing 10% DMSO for 24 h at rt. Individual regioisomers of the bicyclic peptides were produced as illustrated in Figure S1. Linear peptides were synthesized by SPPS wherein one pair of cysteines was protected with Trt (trityl, removed under high concentration of TFA in cleavage cocktail) and the second pair with Acm (acetamidomethyl, removed by iodine oxidation). After cleavage from the rink amide resin, 10 mg of crude linear peptide was oxidized in 10 mL of 100 mM NH_4Ac, pH 8.0, buffer containing 10% DMSO for 24 h at rt. The peptide with a first disulfide bridge formed was purified by RP-HPLC. One mg of monocyclic peptide having a pair of cysteines protected with Acm was dissolved in a mixture of 500 μL of AcOH and 100 μL of H_2O. Ten equiv of I_2 in 15 μL of MeOH were added, and the reaction incubated for 2 h at rt. The mass of the peptide was determined by ESI mass spectrometry and in case of incomplete cyclization, further oxidized with another 10 equiv of I_2. After 2 h oxidation at rt, the reaction mixture was quenched by stepwise adding 0.1 M ascorbic acid in H_2O until the reaction mixture was colorless. The deprotected and oxidized peptides were lyophilized and purified by RP-HPLC. For the cyclization with TBMB, crude peptides (1 mM) in 50% v/v 60 mM NH_4HCO_3, pH 8.0, and 50% v/v acetonitrile were reacted with TBMB (1.5 mM) for 1 h at 30 °C. The reaction products were lyophilized and purified by RP-HPLC as described above. The purified peptides were lyophilized and dissolved in H_2O for activity measurements. The purity of the peptides was assessed by analytical RP-HPLC on an Agilent 1260 system, using a C18 column and the same buffer system as for preparative RP-HPLC. The molecular mass was determined by MALDI-TOF mass spectrometry as described in the Supporting Information.

Determination of Inhibitory Activity of Bicyclic Peptides. The inhibitory activity of bicyclic peptides was determined by incubating human uPA (1.5 nM) with different concentrations of inhibitor and quantification of the residual activity with a fluorogenic substrate (Z-Gly-Gly-Arg-AMC, 50 μM; Bachem, Switzerland) as described before.[12] The inhibitory constant K_i was calculated using equations described in the Supporting Information.

Crystallization and Structure Determination. The catalytic domain of human uPA with two mutations (C122A and N145Q) eliminating the surface-exposed free cysteine residue and the

dx.doi.org/10.1021/ja400461h | J. Am. Chem. Soc. 2013, 135, 6562−6569

glycosylation site was expressed and purified as described in the Supporting Information. The crystallization was carried out employing the method of hanging drop vapor diffusion. Crystals were obtained by mixing 1.25 μL of LMW human uPA-C122A/N145Q protein (15 mg/mL, 528 μM) in 50 mM HEPES, pH 7.0, and 100 mM NaCl with 1.25 μL of 3 mM oxidized and a purified regioisomer of peptide UK504 which was allowed to equilibrate against 1 mL of 50 mM sodium citrate, pH 4.9, 5% v/v PEG400, 1.8 M $(NH_4)_2SO_4$, and 0.05% NaN_3. Crystals appeared at 20 °C within 3 days (Figure S2). For data collection, the crystals were dipped into the LV CryoOil (Jena Biosciences, DE) and flash frozen in a vaporized nitrogen stream at 100 K. X-ray diffraction data were collected at the beamline PXIII of the Swiss Light Source at the Paul Scherrer Institute (SLS, Villigen, Switzerland) at a wavelength of 1 Å. Best crystals diffracted to 1.5 Å resolution. Crystals belong to the *R*3 space group with cell dimensions of a = 120.809 Å and c = 43.061 Å in the hexagonal setting. The diffraction data were indexed, integrated, and scaled with iMOSFLM.[18] The structure was solved by molecular replacement using the uPA protein chain from the 2NWN structure as the search model. ARP/wARP was used to automatically build the uPA and the peptide atomic models. These models were manually corrected with Coot.[19] Coot, Refmac5, and Phenix were used for refinement.[20−22] The refinement statistics are summarized in Table S1. The atomic coordinates and the structure factors were deposited in the Protein Databank under the accession number 4GLY.

RESULTS

Phage Panning of Disulfide-Cyclized Peptides. Phage of the 6 × 6 library (library size: 4 × 10^9)[14] were produced, cyclized by oxidation or by reaction with TBMB, and subjected to affinity selections with the model target streptavidin. Panning phage display peptide libraries against this target had previously yielded peptides with the short amino acid consensus sequence 'HPQ' (Figure 2).[23−25] Peptides containing the 'HPQ' motif bind particularly well to streptavidin when cyclized by a disulfide bridge.[26] In two consecutive rounds of panning, 2.4 × 10^4 and 1.5 × 10^9 phage (rounds 1 and 2, respectively) were isolated out of the disulfide-cyclized library and 3.4 × 10^4 and 8 × 10^8 phage (rounds 1 and 2, respectively) out of the TBMB-cyclized library.

Similarly, phage of a second peptide library having the format XCX_4CX_4CX (4 × 4 library, library size: 4 × 10^8)[27] were produced, cyclized by oxidation or by reaction with TBMB and subjected to affinity selections with human uPA as target. This serine protease is overexpressed in different tumor cell lines and implicated in tumor growth and invasion.[28] In previous phage selections with peptide libraries cyclized by TBMB, we were able to isolate potent and selective inhibitors of this target.[14,15] The titers of phage isolated herein were 3.1 × 10^5 and 1.1 × 10^8 (rounds 1 and 2, respectively) from the disulfide-cyclized library and 5.6 × 10^5 and 1.3 × 10^8 (rounds 1 and 2, respectively) from the TBMB-cyclized library. The strong increase of the phage titers in the second round and the presence of consensus sequences indicated an enrichment of specific binders to uPA (Figure 3A,B).

Prevalence of Peptides with Four Cysteines. The random amino acid positions in both libraries were encoded by 'NNK' codons resulting in a 1:32 chance of finding a cysteine residue in each of the randomized positions. The theoretical probability of finding 0, 1, or 2 additional cysteines in these peptides was 68%, 26%, and 5% for the 6 × 6 library (Figure 2C) and 73%, 24%, and 3% for the 4 × 4 library (Figure 3C). Sequencing of at least 10 clones from each of the 2 naïve libraries showed that the peptides contained the expected number of cysteines (Figures 2D and 3D).

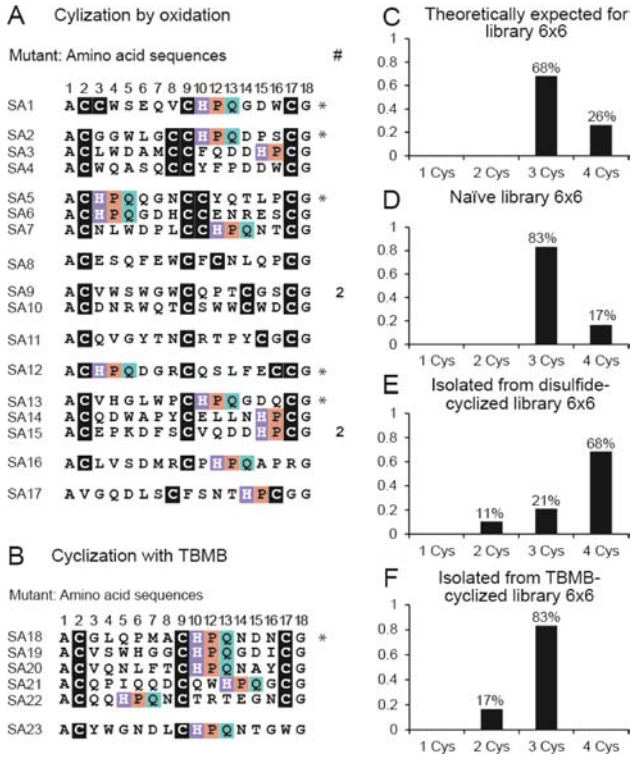


Figure 2. Phage selection of bicyclic peptides against streptavidin. (A,B) Sequences of peptides isolated from the 6 × 6 phage peptide library cyclized by oxidation (A) or by reaction with TBMB (B). For peptides found more than once, the frequency is indicated next to the sequence. The 'HPQ' motif, which was described previously to bind to the biotin-binding site of streptavidin is highlighted. The six peptides marked with an asterisk were synthesized. All these peptides bound to streptavidin (Figure S3). (C) Bar diagram showing the fractions of peptides in the 6 × 6 phage peptide library that theoretically contain 1, 2, 3, or 4 cysteines. These numbers were calculated based on the number of amino acids randomized and the type of codons used in the construction of the phage peptide library. (D−F) Fractions of peptides that were found to contain the indicated number of cysteine residues in the naïve library (D), after two selection rounds, cyclizing the peptides by oxidation (E), or by reaction with TBMB (F).

After two rounds of panning, the number of cysteines per peptide was found to be strongly influenced by the cyclization method. In selections using the oxidatively cyclized peptide libraries, most of the isolated clones contained an even number of cysteines. In the panning against streptavidin, 68% and 11% of the isolated peptides contained four or two cysteines, respectively (Figure 2E). Similarly, in the panning against uPA, most peptides contained four (62%) or two cysteines (23%) (Figure 3E). Peptides with only two cysteines are theoretically not expected as the library is designed to contain three cysteines in constant positions. However, inefficient chemical coupling of nucleotides in the synthesis of degenerated primers used for library cloning leads to peptides with shorter or longer sequences and peptides with fewer than three cysteines. Such peptides occur rarely in the library, and their appearance in the selection indicates a strong selective advantage. In contrast, peptides with three cysteines, representing a majority in both naïve libraries, underwent a strong negative selection.

In selections with phage libraries cyclized by TBMB, most of the isolated peptides contained exactly three cysteine residues (83% in the selection with streptavidin and 97% in the selection with uPA) (Figures 2F and 3F). The rest of the peptides

dx.doi.org/10.1021/ja400461h | J. Am. Chem. Soc. 2013, 135, 6562−6569

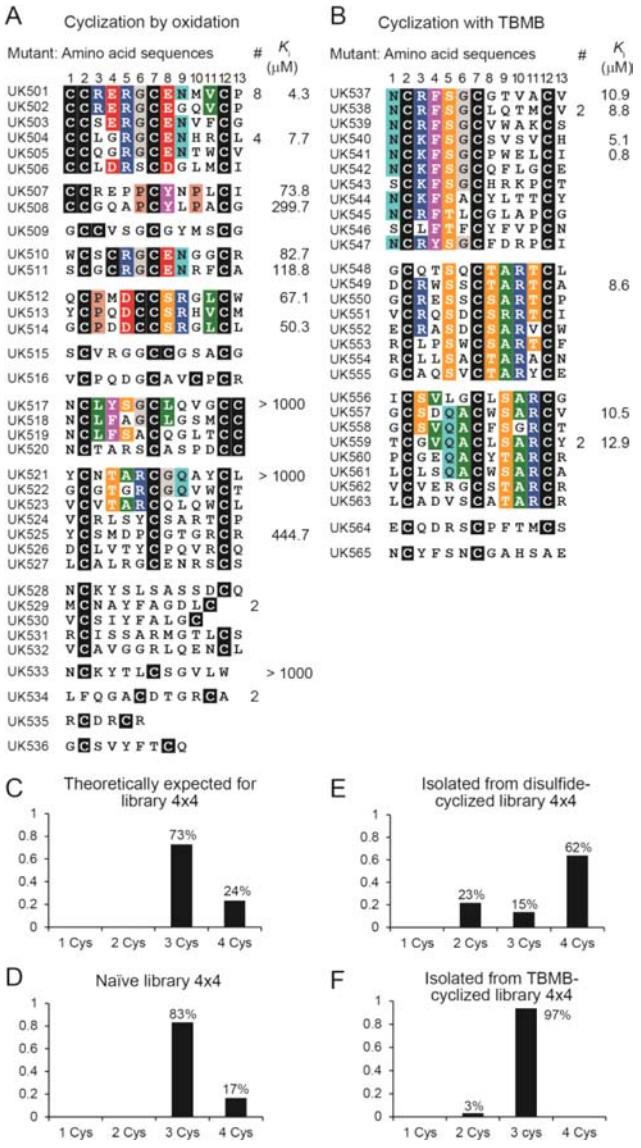


Figure 3. Phage selection of bicyclic peptides against the serine protease uPA. (A,B) Peptides isolated from the 4 × 4 library cyclized by oxidation (A) or by reaction with TBMB (B). Conserved amino acids in the peptides are highlighted in color. For peptides found more than once, the frequency is indicated next to the sequence. Inhibitory activities of some peptides are indicated with their K_i. For the bicyclic peptides UK501 and UK504, the K_i of the most active isomers is indicated. (C) Theoretically expected fraction of peptides with the indicated total number of cysteines. (B–D) Bar diagram showing the fractions of peptides in the 4 × 4 phage peptide library that theoretically contain 1, 2, 3, or 4 cysteines. These numbers were calculated based on the number of amino acids randomized and the type of codons used in the construction of the phage peptide library. (D–F) Fractions of peptides that were found to contain the indicated number of cysteine residues in the naïve library (D), after two selection rounds, cyclizing the peptides by oxidation (E), or by reaction with TBMB (F).

contained only two cysteines, and no peptide from these selections contained four cysteines. This pattern was diametrically opposite to the result found for selections with disulfide-cyclized peptides. The reasons for the strong enrichment of peptides with specific numbers of cysteines in both cyclization strategies were not clear at this point, but they were found out later as described below. We speculated at this point that unpaired cysteines might form disulfide bridges with cysteines of neighboring peptides and thus impair infection or that a cyclized configuration would bring a strong advantage in the binding of the peptides.

Sequence Diversity of Binders. In all selections, consensus sequences were found, indicating that the isolated peptides formed specific interactions with the targets. In the selection with streptavidin, the motif 'HPQ' appeared in most peptides (Figure 2A,B). The peptides isolated from the disulfide-cyclized phage library contained a fourth cysteine in 7 of the 12 randomized positions, resulting in a large diversity of bicyclic peptide topologies (Figure 2A). The 'HPQ' motif was present in different positions of the peptides, preferentially following one of the cysteine residues. In the selections with TBMB-cyclized peptides, the 'HPQ' motif was also found in different positions and also in both of the peptide loops (Figure 2B). It is likely that more stringent selection conditions would yield peptides with an extended consensus sequence. Six of the peptides were synthesized, and all were found to bind to streptavidin (Figure S3).

In the selections against human uPA, stronger consensus sequences were observed after the two iterative rounds of panning (Figure 3A,B). The selection with disulfide-cyclized 4 × 4 peptide library yielded six different consensus sequences (Figure 3A). Previously, one or two, and rarely three, consensus sequences were found in selections with the TBMB-cyclized 6 × 6 peptide library against several targets.[12−14] In the selection with the TBMB-cyclized peptide library performed in this work, also only three different consensus sequences were found, and two of them were similar. The higher sequence diversity found for the disulfide-based bicyclic peptides most likely results from the larger topological diversity of this library: The fourth cysteine can be placed in 10 positions, and the four cysteines can pair in three different ways, resulting in 30 different peptide topologies. All preferred positions for the fourth cysteine were adjacent to one of the fixed cysteines (positions 1, 6, and 13). The consensus sequences of bicyclic peptides isolated with the two different cyclization methods were different. An exception is the short consensus sequence 'TAR' that was found in peptides of both selections. Peptides with this latter sequence likely bind similarly to uPA.

Cysteine Pairing in Disulfide-Cyclized Peptides. Twelve peptides isolated against uPA in selections with oxidative cyclization were chemically synthesized, reduced, and purified by RP-HPLC chromatography. Oxidative cyclization of peptides containing four cysteines yielded in most cases three products that eluted with different retention times. The 4-cysteine peptides UK501 and UK504 that were most frequently isolated (8 and 4 times, respectively) eluted in two or three peaks (Figure 4A,B). The masses of all these peaks corresponded to a peptide configuration with two disulfide bridges. Some of the peaks were larger than others showing that certain regioisomers were preferentially formed. A peptide with three cysteines (UK525) yielded a complex mixture of products upon oxidation (Figure 4C). This peptide had most likely formed dimers via the third cysteine that could not pair intramolecularly, as discussed above. A peptide with only two cysteines (UK533) was also synthesized and yielded a single product upon oxidation, as expected (Figure 4D).

The regioisomers of the bicyclic peptides UK501 and UK504 were individually prepared by synthesizing the peptides with cysteine pairs having orthogonal protection groups that could be removed and oxidized stepwise (trityl, Trt, and acetoami-

dx.doi.org/10.1021/ja400461h | J. Am. Chem. Soc. 2013, 135, 6562−6569

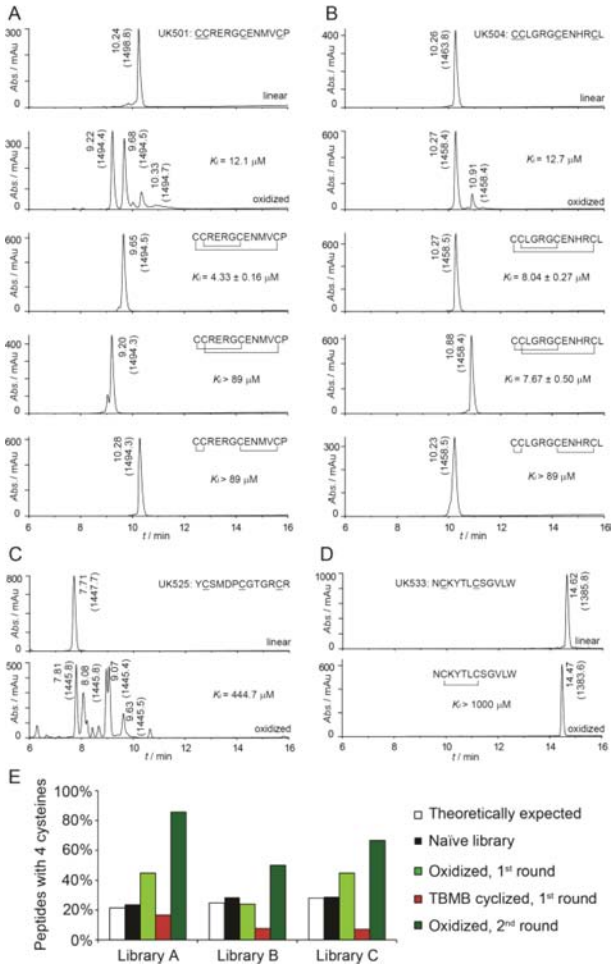


Figure 4. Formation of disulfide bridges in cysteine-rich peptides. (A–D) Four peptides containing either four (UK501 and UK504), three (UK525), or two (UK533) cysteines were synthesized, purified, and analyzed by analytical HPLC before (top panels) and after oxidation (panels below). Before oxidation, all peptides showed a single peak with the expected mass. After oxidation, some of the peptide eluted of at multiple retention times, indicating that different products were formed through the pairing of cysteines. Oxidation of peptides with four cysteines yielded at least two of the three possible regioisomers. Oxidation of the peptide with three cysteines yielded a complex product mixture, and oxidation of the peptide with two cysteines gave a single product. In case of the peptides UK501 and UK504 having four cysteines, the three bicyclic peptide regioisomers were also synthesized applying orthogonal thiol protection groups to direct the pairing of specific cysteines (Figure S1). The retention times (minutes) and the masses (Da, in parentheses) of the peaks are indicated. (E) Propagation of phage displaying cysteine-rich peptides. Indicated is the fraction of phage containing peptides with four cysteines, before and after bacterial infection. Phage DNA of three different naïve libraries (libraries A–C; CX_mCX_nC, m, n = 3–6)[15] was sequenced and the percentage of peptides with four cysteines determined (naïve library). Phage of these libraries were either allowed to form disulfide bridges by oxidation or modified with TBMB. Phage DNA of infected cells was sequenced, and the fraction of peptides with four cysteines determined (first round). A second cycle of phage production and reinfection was performed for the phage cyclized by oxidation (second round).

nomethyl, Acm; Figure S1). The three regioisomers of UK501 had different retention times in RP-HPLC (Figure 4A). Those of UK504 eluted at two different times. The characteristic retention times allowed identification of the regioisomers that formed in the previous experiment in which all four cysteines were allowed to oxidize simultaneously. The regioisomer of UK501 that was formed in a smaller quantity upon simultaneous oxidation of the four cysteines turned out to be the one having a disulfide bridge between the neighboring cysteines 1 and 2. This result is in line with the observation that vicinal cysteines in proteins prefer not to form disulfide bridges.[29]

The most potent disulfide-based bicyclic peptide inhibitor of uPA, a regioisomer of UK501 (cysteine pairing: Cys1-Cys12 and Cys2-Cys7), had a K_i of 4.33 ± 0.16 μM in buffered saline at physiologic pH. This activity is 7-fold better than the best monocyclic peptide upain-1, which was isolated from a disulfide-constrained monocyclic peptide phage library (CSWRGLENHRMC; K_i = 29.9 μM at physiological pH).[30] Most peptides isolated by TBMB-cyclization showed inhibitory activities in the same range. One showed an exceptionally good K_i of 0.78 μM (UK541, Figure 3B).

Phage Displaying Peptides with Unpaired Cysteines Are Less Efficiently Propagated and Have a Lower Infectivity. Phage displaying peptides with a single cysteine residue were previously found to undergo negative selection during phage amplification.[24,31] The presence of unpaired cysteine residues has been associated with reduced periplasmic expression[32] and interference with the cysteine residues present in phage coat proteins.[24,31,33] To investigate this phenomenon in the context of our cysteine-rich peptide libraries, which are displayed on cysteine-free domains D1 and D2 of the phage p3, we analyzed the number of cysteine residues after iterative rounds of phage production and infection, without performing any affinity selection. We used three different phage libraries (libraries A–C) displaying peptides with three cysteine residues spaced by variable numbers of random amino acids (CX_mCX_nC-phage; m, n = 3–6).[15] Sequencing of more than 50 clones of these libraries showed that 23% of the peptides contained 4 cysteines (average of all three libraries) (Figure 4E). After phage production, half of the phage were modified with TBMB, and half were let oxidize, and cells were infected with the two populations. In the case of the oxidized half, the fraction of clones containing 4 cysteines rose to 38%. In contrast, in the case of the TBMB-modified half, the percentage of clones containing 4 cysteines dropped to 7%. This finding showed that the presence of unpaired cysteines can significantly influence the efficiency of phage propagation. To further confirm the trend, another round of phage production and infection was performed. This time the percentage of clones containing four cysteines rose in average to 67% of the whole phage population. The results of these phage propagation experiments suggested that peptides isolated before in affinity selections with streptavidin and uPA had enriched specific cysteine numbers primarily due to the higher infectivity of the phage carrying these peptides rather than the cyclic peptide format that improved the binding affinity.

To confirm the influence of the cysteine number on phage infectivity, 21 phage clones displaying peptides with either 2, 3, or 4 cysteines were produced individually and analyzed. Quantification of the purified phage by SDS-PAGE (analysis of the major coat protein pVIII) showed that all phage were produced in comparable quantities. Phage displaying peptides with 3 cysteines showed an ~5-fold lower infectivity than those with peptides having 2 or 4 cysteines (Figure S4).

Crystal Structure of Bicyclic Peptide UK504 Bound to uPA. The structure of a disulfide-cyclized peptide regioisomer

dx.doi.org/10.1021/ja400461h | J. Am. Chem. Soc. 2013, 135, 6562–6569

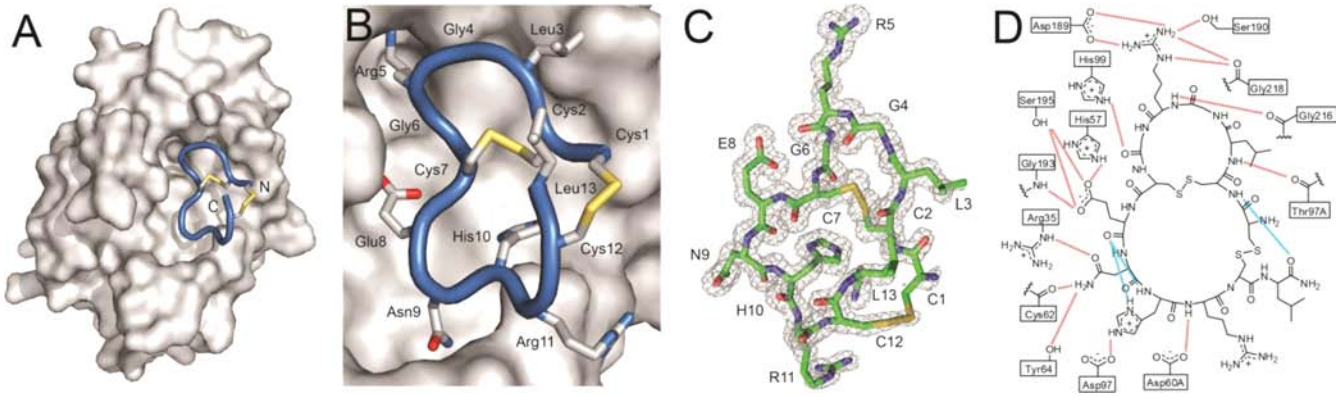


Figure 5. Crystal structure of bicyclic peptide UK504 bound to the active site of human uPA. (A) UK504 (green) binds to the substrate-binding pocket of uPA. The disulfide-linked cysteine side chains are shown in yellow. (B) A close-up view of the uPA substrate-binding pocket. The atoms of the amino acid side chains are shown in gray (carbon), blue (nitrogen), red (oxygen), and yellow (sulfur). (C) The $2F_o - F_c$ electron density map contoured at 1.0 Å is shown with the UK504 atomic model. (D) Intramolecular (blue) and intermolecular (red) H-bond interactions of UK504 with uPA (Tables S2 and S3).

of UK504 (CCLGRGCENHRCL-NH_2) in complex with the catalytic domain of human uPA (GenBank ID CAI13969, residues 179−423) was crystallized (Figure S2). The structure was determined at a resolution of 1.5 Å (Figure 5A,B and Table S1) (PDB entry 4GLY). Clone UK504 was chosen as a representative disulfide-cyclized bicyclic peptide due to its high similarity to other peptides in the same consensus sequence. The electron density of the peptide had a high quality which was comparable to that of the protein (Figure 5C). It was therefore possible to build the complete atomic model of the peptide. The electron density for the terminal residues Cys1 and Leu13 is weaker than that for the rest of the peptide, as these two amino acids are fully solvent exposed and do not participate in binding to the protein target. The crystal structure confirmed that Cys1 and Cys2 are linked to Cys12 and Cys7, respectively.

The first peptide ring connected by Cys2 and Cys7 (CLGRGC) contains Arg5, which occupies the S1 specificity pocket of uPA (Figure 5D). The guanidinium group of Arg5 forms two H-bonds and a complementary charge interaction with Asp189 at the bottom of the S1 pocket as well as three H-bonds with Ser190 and Gly218 (chymotrypsin numbering is used for all uPA residues). The second ring is formed by the vicinal Cys1 and Cys2 residues that pair with Cys12 and Cys7, respectively. Glu8 of the second ring forms H-bonds with Ser195 and His57 of the catalytic triad, potentially inactivating the protease.

Peptide Inhibitors with Different Scaffolds Share a Common Paratope. The structure of bicyclic peptide UK504 in complex with uPA was compared with those of the phage-selected monocyclic peptide upain-1 (PDB entry 2NWN)[30,34] and the bicyclic peptide UK18 (PDB entry 3QN7)[12] that were previously determined by X-ray crystallography (Figure 6). The protein backbones of the three uPA structures are similar, and they can be superimposed onto each other with the root-mean-square deviation (RMSD) of 0.6 Å or less, showing that inhibitor binding has little influence on the conformation of the protein. In UK504 and upain-1, the C-terminal halves of the peptide starting from the conserved arginine (UK504: RGCENHRCL, upain-1: RGLENHRMC; identical amino acids underlined) have a similar conformation and occupy the same site in the substrate-binding pocket of uPA. In upain-1, this region is cyclized by one disulfide bridge, and in UK504 the same region is linked via two disulfide bridges to the N-terminus. The N-terminal regions of the two inhibitors adopt different conformations. The N-terminus of UK504 is longer and fills a cavity in the substrate-binding pocket of uPA. UK504 and UK18 have completely different backbone conformations. The trajectories of the two peptides overlap in a region of around five amino acids, but the peptide backbones run in opposite directions.

While the three inhibitors have different structures, similar amino acids form the key interactions with uPA. Namely, all the inhibitors have an arginine residue that occupies the S1 pocket of uPA (highlighted in blue in Figure 6). Furthermore, the three inhibitors contain either a glutamate (Glu8 in UK504 and Glu 7 in upain-1) or aspartate residue (Asp8 in UK18) that forms two H-bonds with Ser195 and His57 (highlighted in red in Figure 6). Retrospective analysis of all the isolated peptides showed that two other consensus motifs contained a pair of conserved arginine and aspartate/glutamate. The first of these consensus sequences formed by the clones UK510 and UK511 shares a high similarity with the crystallized bicyclic peptide UK504 and likely binds in the same orientation. The second consensus formed by clones UK512, UK513, and UK514 contains the conserved arginine and glutamate residues in different positions. If these residues form similar contacts with uPA as the arginine and aspartate of UK504, it is likely that the peptides have an entirely different topology.

■ DISCUSSION

By choosing a peptide library format with cysteines in three constant positions and random intervening and flanking amino acids, we could generate libraries that contain a large fraction of disulfide-linked bicyclic peptides. Our concerns that topoisomers resulting from differently connecting pairs of cysteines could hinder the phage selections, proved to be not true. Similarly, oxidative cross-linking of cysteine-rich peptides on the same or different phage particles did not hinder the isolation of bicyclic peptides. The lower infectivity of phage with unpaired cysteines and their less efficient propagation most likely supported the powerful selection of disulfide-circularized bicyclic peptides against streptavidin and uPA, allowing isolation of good binders in only two rounds of affinity selection.

dx.doi.org/10.1021/ja400461h | J. Am. Chem. Soc. 2013, 135, 6562−6569

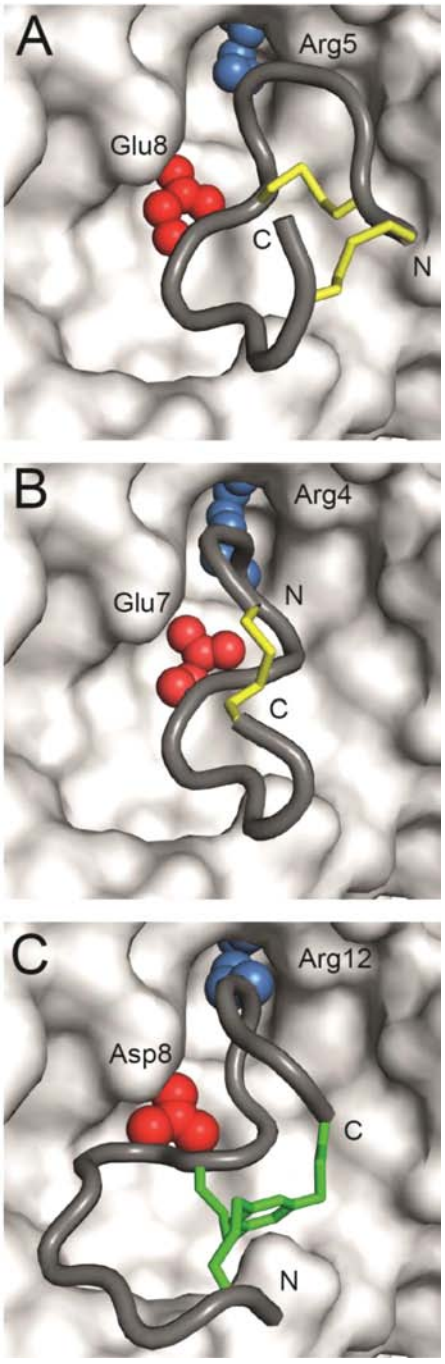


Figure 6. Structures of monocyclic and bicyclic peptide inhibitors of human uPA. (A) Bicyclic peptide UK504 based on two disulfide bridges (yellow). (B) Monocyclic peptide upain-1 having one disulfide bridge (yellow).[30,34] (C) Bicyclic peptide UK18 having three cysteines linked via thioether bonds to a mesitylene core (green).[12] All the three inhibitors have a common paratope formed by Arg (blue) which binds to the S1 site and Glu or Asp (red) forming H-bonds with the active site residues Ser195 and His57 of uPA.

The bicyclic peptide libraries have a high topological diversity because the fourth cysteines can occur in any one of the randomized amino acid positions. This allows the facile generation of differently sized bicyclic peptides. In the previously developed approach based on TBMB-cyclized peptides, only a single bicyclic peptide format is generated per type of library.[14,15] The ability of the peptides to pair the four cysteines in three different ways further increases the library diversity. The high structural diversity of the disulfide-cyclized peptide library was substantiated by the large number of consensus sequences that was found in the selection against the human serine protease uPA.

Disulfide-based bicyclic peptides are good binders as demonstrated by isolating in a single attempt and after only two iterative rounds of panning an inhibitor that binds with a 7-fold higher affinity than the best monocyclic human uPA inhibitor. While cyclization with TBMB gave the best binder in this study, it is not possible to state at this stage if one of the two cyclization methods yields better binders in general. Future phage selections with many different targets will allow answering this question. Considering the large sequence diversity of the peptides isolated in two rounds of panning, it is likely that inhibitors with even higher affinities could be enriched in a third or fourth round of panning. It has been shown in several studies that the library size and diversity correlate with the affinity of isolated ligands.[35,36]

Although the oxidative cyclization of the isolated peptides can yield up to three topoisomers, their synthesis and purification are technically simple. If a specific isomer is desired, the active regioisomer can be produced using orthogonal thiol protecting groups as demonstrated. Alternatively, a mixture of the three possible regioisomers is produced, and the desired isomer separated chromatographically. The latter strategy is particularly efficient if the desired isomer is preferentially formed as found for bicyclic peptides UK501 and UK504. For many biotechnological or research applications, isomer mixtures may be produced and applied without prior purification of the active isomer.

The crystal structure of the bicyclic peptide UK504 showed that the peptide bound to uPA has a rigid structure and fully occupies the substrate-binding site of its target. In this example, the disulfide bonds constrained the 13 amino acids peptide in two rings of nearly equal size. Interestingly, the crystallized bicyclic peptide interacts with similar key amino acids with uPA as previously developed peptidic uPA inhibitors that have different core structures.

■ CONCLUSIONS

We show that disulfide-cyclized bicyclic peptides with good affinities for targets of interest can be pulled out of phage-encoded cysteine-rich peptide libraries. The libraries have a high topological diversity because the fourth cysteines can occur in any one of the randomized amino acid positions and the four cysteines can pair in three different ways. The most important advantage of the presented strategy is its technically simple experimental procedures. The cysteine-rich peptides are automatically air oxidized during the production of the phage and can readily be subjected to affinity selections. The oxidation reaction is robust and does not require verification of the cyclization efficiency. Given this technical advantage and the high structural diversity of the library, the presented strategy and library format should find broad application for the generation of peptide ligands.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Additional information about the crystal structure, detailed experimental procedures, figures showing the synthetic strategy, the binding of peptides to streptavidin, and the expression and infectivity phage, and a picture of the crystals are provided. This

dx.doi.org/10.1021/ja400461h | J. Am. Chem. Soc. 2013, 135, 6562−6569

material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

christian.heinis@epfl.ch

Present Address

[§]Department of Chemistry, University of Oxford, Oxford, United Kingdom.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The financial contribution from the Swiss National Science Foundation (SNSF Professorship PP00P3_123524/1 to C.H.) is gratefully acknowledged. J.M.S. was supported by a fellowship of the Ministerio de Educacion del Gobierno de España. We thank the staff of the Swiss Light Source (SLS) of the Paul Scherer Institute for help with X-ray data collection.

REFERENCES

(1) Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K. *Nat. Rev. Drug Discovery* **2008**, *7*, 608.

(2) Loffet, A. *J. Pept. Sci.* **2002**, *8*, 1.

(3) Macdougall, I. C.; Rossert, J.; Casadevall, N.; Stead, R. B.; Duliege, A. M.; Froissart, M.; Eckardt, K. U. *The N. Engl. J. Med.* **2009**, *361*, 1848.

(4) Gentilucci, L.; Tolomelli, A.; Squassabia, F. *Curr. Med. Chem.* **2006**, *13*, 2449.

(5) Li, P.; Roller, P. P. *Curr. Top. Med. Chem.* **2002**, *2*, 325.

(6) Daly, N. L.; Craik, D. J. *Curr. Opin. Chem. Biol.* **2011**, *15*, 362.

(7) Angelini, A.; Morales-Sanfrutos, J.; Diderich, P.; Chen, S.; Heinis, C. *J. Med. Chem.* **2012**, *55*, 10187.

(8) Ladner, R. C. *Trends Biotechnol.* **1995**, *13*, 426.

(9) Szardenings, M. *J. Recept. Signal Transduction Res.* **2003**, *23*, 307.

(10) Katz, B. A. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 27.

(11) Ladner, R. C.; Sato, A. K.; Gorzelany, J.; de Souza, M. *Drug Discovery Today* **2004**, *9*, 525.

(12) Angelini, A.; Cendron, L.; Chen, S.; Touati, J.; Winter, G.; Zanotti, G.; Heinis, C. *ACS Chem. Biol.* **2012**, *7*, 817.

(13) Baeriswyl, V.; Rapley, H.; Pollaro, L.; Stace, C.; Teufel, D.; Walker, E.; Chen, S.; Winter, G.; Tite, J.; Heinis, C. *ChemMedChem* **2012**, *7*, 1173.

(14) Heinis, C.; Rutherford, T.; Freund, S.; Winter, G. *Nat. Chem. Biol.* **2009**, *5*, 502.

(15) Rentero Rebollo, I.; Angelini, A.; Heinis, C. *MedChemComm* **2012**, *4*, 145.

(16) Timmerman, P.; Beld, J.; Puijk, W. C.; Meloen, R. H. *ChemBioChem* **2005**, *6*, 821.

(17) Rentero Rebollo, I.; Heinis, C. *Methods* **2013**, DOI: 10.1016/j.ymeth.2012.12.008.

(18) Leslie, A. G. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2006**, *62*, 48.

(19) Emsley, P.; Cowtan, K. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126.

(20) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 235.

(21) Murshudov, G. N.; Skubak, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. *AActa Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 355.

(22) Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Echols, N.; Headd, J. J.; Hung, L. W.; Jain, S.; Kapral, G. J.; Grosse Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R. D.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. *Methods* **2011**, *55*, 94.

(23) Devlin, J. J.; Panganiban, L. C.; Devlin, P. E. *Science* **1990**, *249*, 404.

(24) Kay, B. K.; Adey, N. B.; He, Y. S.; Manfredi, J. P.; Mataragnon, A. H.; Fowlkes, D. M. *Gene* **1993**, *128*, 59.

(25) McLafferty, M. A.; Kent, R. B.; Ladner, R. C.; Markland, W. *Gene* **1993**, *128*, 29.

(26) Giebel, L. B.; Cass, R. T.; Milligan, D. L.; Young, D. C.; Arze, R.; Johnson, C. R. *Biochemistry* **1995**, *34*, 15430.

(27) Baeriswyl, V.; Heinis, C. *Protein Eng., Des. Sel.* **2013**, *26*, 81.

(28) Andreasen, P. A.; Egelund, R.; Petersen, H. H. *Cell. Mol. Life Sci.* **2000**, *57*, 25.

(29) Carugo, O.; Cemazar, M.; Zahariev, S.; Hudaky, I.; Gaspari, Z.; Perczel, A.; Pongor, S. *Protein Eng.* **2003**, *16*, 637.

(30) Hansen, M.; Wind, T.; Blouse, G. E.; Christensen, A.; Petersen, H. H.; Kjelgaard, S.; Mathiasen, L.; Holtet, T. L.; Andreasen, P. A. *J. Biol. Chem.* **2005**, *280*, 38424.

(31) McConnell, S. J.; Uveges, A. J.; Fowlkes, D. M.; Spinella, D. G. *Mol. Diversity* **1996**, *1*, 165.

(32) Schmiedl, A.; Breitling, F.; Winter, C. H.; Queitsch, I.; Dübel, S. *J. Immunol. Methods* **2000**, *242*, 101.

(33) Wind, T.; Kjær, S.; Clark, B. F. C. *Biochimie* **1999**, *81*, 1079.

(34) Zhao, G.; Yuan, C.; Wind, T.; Huang, Z.; Andreasen, P. A.; Huang, M. *J. Struct. Biol.* **2007**, *160*, 1.

(35) Griffiths, A. D.; Williams, S. C.; Hartley, O.; Tomlinson, I. M.; Waterhouse, P.; Crosby, W. L.; Kontermann, R. E.; Jones, P. T.; Low, N. M.; Allison, T. J.; et al. *EMBO J.* **1994**, *13*, 3245.

(36) Sheets, M. D.; Amersdorfer, P.; Finnern, R.; Sargent, P.; Lindquist, E.; Schier, R.; Hemingsen, G.; Wong, C.; Gerhart, J. C.; Marks, J. D. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6157.