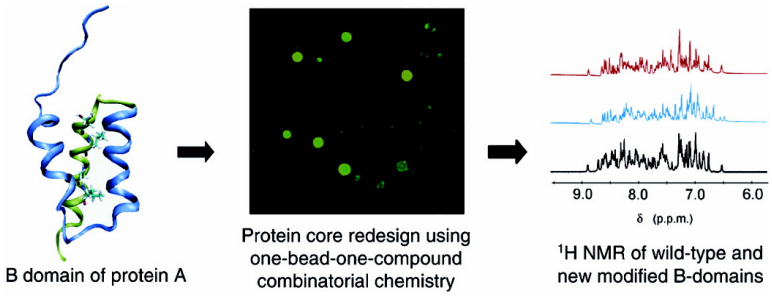
Article

Redesign of Protein Domains Using One-Bead-One-Compound Combinatorial Chemistry

Jose J. Pastor, Giovanna Granados, Natlia Carulla, Francesc Rabanal, and Ernest Giralt

J. Am. Chem. Soc., **2007**, 129 (48), 14922-14932 • DOI: 10.1021/ja073969x



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 11/09/2007

Redesign of Protein Domains Using One-Bead-One-Compound Combinatorial Chemistry

Jose J. Pastor,[†] Giovanna Granados,[†] Natàlia Carulla,[†] Francesc Rabanal,[‡] and Ernest Giralt*,[†,‡]

Contribution from the Institute for Biomedical Research, Barcelona Science Park, 08028 Barcelona, Spain, and Department of Organic Chemistry, University of Barcelona, 08028 Barcelona Spain

Received June 1, 2007; E-mail: egiralt@pcb.ub.cat

Abstract: A novel combinatorial strategy for the redesign of proteins based on the strength and specificity of intra- and interprotein interactions is described. The strategy has been used to redesign the hydrophobic core of the B domain of protein A. Using one-bead-one-compound combinatorial chemistry, 300 analogues of the C-terminal α-helix of the B domain, **H3**$_x$, have been synthesized using a biocompatible resin and the HMFS linker, allowing the screening to occur while the peptides were bound to the resin. The screening was based on the ability of the **H3**$_x$ analogues to interact with the N-terminal helices of the B domain, **H1**-**H2**, and retain the native B domain activity, that is binding to IgG. Eight different analogues containing some nonconservative mutations were obtained from the library, the two most frequent of which, **H3**$_{P1}$ and **H3**$_{P2}$, were studied in detail. CD analysis revealed that the active analogues interact with **H1**-**H2**. To validate the redesign strategy the covalent modified domains **H1**-**H2**-**H3**$_{P1}$ and **H1**-**H2**-**H3**$_{P2}$ were synthesized and characterized. CD and NMR analysis revealed that they had a unique, stable, and well-defined three-dimensional structure similar to that for the wild-type B domain. This combinatorial strategy allows us to select for redesigned proteins with the desired activity or the desired physicochemical properties provided the right screening test is used. Furthermore, it is rich in potential for the chemical modification of proteins overcoming the drawbacks associated with the total synthesis of large protein domains.

Introduction

Proteins show great potential in industrial, chemical, and pharmaceutical applications, since they can efficiently catalyze difficult reactions under mild conditions, with high regio- and stereospecificity and minor byproduct formation.[1−4] Thus, a considerable amount of work has been put toward the application of proteins to new chemical processes on both laboratory and industrial scales.[3,5] However, further efforts may be necessary to avoid problems, such as the sluggish catalysis of non-natural substrates, poor activity in nonaqueous media, and little tolerance for changes in operating parameters (e.g., temperature and pH).[6]

Proteins also offer great promise as therapeutic agents for various indications, due to the high affinity and specificity of their *in vivo* interactions with other proteins.[7] Owing to the advances in recombinant DNA techniques, the number of natural and non-natural peptides used in pharmaceutical purposes has surged in the past few years.[8−10] Nonetheless, peptide and protein drugs still suffer from poor biostability and pharmacokinetics and, to a lesser extent, poor selectivity for receptor subtypes.[11] The aforementioned limitations both in the use of enzymes in industrial, chemical, and pharmaceutical applications and in the use of proteins as therapeutic agents have sparked research on the modification of natural proteins so that they can be used for the intended purpose.

The function of a given protein (molecular recognition, catalysis...) is usually encoded in its three-dimensional structure, responsible for the three-dimensional setting of the amino acids in the space necessary to develop the activity. Thus, for a protein to function correctly it must fold into a single, defined tertiary structure.[12,13] The forces involved in the *folding process* are simple and known, and it is a balance among them, reached through different interactions among the amino acids of the sequence, that leads to a protein structure.[12,14,15] However, this balance is extremely complex, delicate, and, above all, poorly

[†] Institute for Biomedical Research.
[‡] University of Barcelona.

(1) Bolon, D. N.; Voigt, C. A.; Mayo, S. L. *Curr. Opin. Chem. Biol.* **2002**, *6* (2), 125−129.
(2) Ward, T. R. *Chem.—Eur. J.* **2005**, *11*, 3798−3804.
(3) Schmid, A.; Dordick, J. S.; Hauer, B.; Kiener, A.; Wubbolts, M.; Witholt, B. *Nature* **2001**, *409*, 258−268.
(4) Arnold, F. H. *Nature* **2001**, *409*, 253−257.
(5) DeSantis, G.; Jones, J. B. *Curr. Opin. Biotechnol.* **1999**, *10*, 324−330.
(6) Farinas, E. T.; Bulter, T.; Arnold, F. H. *Curr. Opin. Biotechnol.* **2001**, *12*, 545−551.
(7) a) Morishita, M.; Peppas, N. A. *Drug Discovery Today* **2006**, *11*, 905−910; b) Reichman, D.; Rahat, O.; Cohen, M.; Neuvirth, H.; Schreiber, G. *Curr. Opin. Struct. Biol.* **2007**, *17*, 67−76.
(8) Pastor, M. T.; Perez-Paya, E. *Mol. Divers.* **2003**, *6*, 149−155.
(9) Hey, T.; Fiedler, E.; Rudolph, R.; Fiedler, M. *Trends Biotechnol.* **2005**, *23*, 514−522.
(10) Frokjaer, S.; Otzen, D. E. *Nat. Rev. Drug Discov.* **2005**, *4* (4), 298−306.
(11) Harris, J. M.; Chess, R. B. *Nat. Rev. Drug Discov.* **2003**, *2* (3), 214−221.
(12) Dobson, C. M. *Nature* **2003**, *426*, 884−890.
(13) Damodharan, L.; Pattabhi, V. *Biochem. Biophys. Res. Commun.* **2004**, *323*, 996−1002.
(14) DeGrado, W. F.; Wasserman, Z. R.; Lear, J. D. *Nature* **1989**, *243*, 622−628.

10.1021/ja073969x CCC: $37.00

understood, meaning that most man-made mutations in natural proteins are either neutral or disruptive, due to our lack of knowledge about the whole protein folding process.[16−18] Modification of natural proteins is hence at odds with the protein's inherent need for a stable structure. In spite of this, there have been several reports demonstrating that modification of natural proteins[19−21] and the development of new proteins[22−26] are indeed possible, although extraordinarily difficult and time-consuming.

Another important factor to be considered in protein modification is the size of the target systems. For a protein with 100 amino acids, permutations with the 20 natural amino acids at all positions will yield 10^{130} different sequences, a number of proteins impossible to synthesize. Furthermore, only a small portion of these will be expected to have a stable structure.[27,28] The potential for success of single point modifications is thus limited, due to the numerous factors involved in protein activity and folding. The use of libraries of modified proteins is a more realistic approach, although the astronomical amount of potential modifications demands certain restrictions in its application.[29] A library of modified proteins must incorporate enough diversity to cover a significant part of the sequence space while simultaneously incorporating enough rational design to limit the exploration to only those regions most likely to yield sequences with the desired properties.[27] This balance is key to reach our objective of modifying natural proteins.

Several combinatorial approaches have been applied to the modification of natural proteins: (i) laboratory or directed evolution;[16,30,31] (ii) point mutagenesis;[20] (iii) partial domain swapping;[32,33] (iv) chemical methods.[34−36] Chemical methods are especially interesting as they facilitate the introduction of chemical diversity into a protein with respect to genetic methodologies (i.e., biosynthesis of proteins based on mutated RNA or DNA). Thus, for example, incorporation of D-amino acids into a protein sequence is particularly appealing as the resulting D-analogue has greater stability to proteolysis. This possibility is of great interest when proteins are to be used as therapeutic agents.[37−40] However, existing peptide synthetic methodologies cannot access the synthesis of large proteins in an easy, fast, and efficient manner, thus limiting their application to the combinatorial modification of proteins.[41]

Our laboratory has recently reported a new methodology[42] that allows the synthesis and selective detachment of large unprotected peptides bound to the biocompatible PEGA resin,[43] via the HMFS linker.[44] This methodology enabled the synthesis of the wild type C-terminal α-helix of the B domain of staphylococcal protein A (**$H3_{wt}$**, see below) and a scrambled version of the same peptide (**$H3_s$**). Very importantly, the purity of the peptides obtained revealed that it was feasible to perform a screening test while the peptide was still attached to the biocompatible resin.[45,46] This feature made this methodology suitable for the development of one-bead-one-compound combinatorial peptide libraries.[47−49]

Here we describe combinatorial strategy for natural protein redesign. The strategy is based on the hypothesis that globular proteins and protein domains can be disconnected through a covalent bond and that the two resulting moieties maintain their ability to form a globular structure very similar to that of the native protein in terms of both structure and function. This principle has been reported for different cases[50] including our recent work on the c*oupe du roi* bisection of uteroglobin.[51] This strategy offers two advantages. First, since the protein to be redesigned is cut in two, one of the fragments can be modified, with a wide range of non-natural elements, overcoming the problem derived from the synthesis of large proteins. Second, since the screening test can be based on the ability of the two noncovalent protein fragments to carry out the function of the native protein, the screening is selecting active, well-folded redesigned proteins.

One of the most difficult things to attain in the design of *de novo* proteins is the formation of a well folded core which ultimately results in a protein with a well-defined tertiary

(15) Bryson, J. W.; Betz, S. F.; Lu, H. S.; Suich, D. J.; Zhou, H. X.; O'Neil, K. T.; DeGrado, W. F. *Science* **1995**, *270*, 935−941.
(16) Endelman, J. B.; Silberg, J. J.; Wang, Z.-G.; Arnold, F. H. *Design Select.* **2004**, *17*, 589−594.
(17) Guo, H. H.; Choe, J.; Loeb, L. A. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9205−9210.
(18) Chen, Y.; Jin, R.; Dong, H.-Y.; Feng, Y.-M. *J. Biol. Chem.* **2004**, *279*, 55224−55233.
(19) Toma, S.; Campagnoli, S.; Margarit, I.; Gianna, R.; Grandi, G.; Bolognesi, M.; Filippis, V. D.; Fontana, A. *Biochemistry* **1991**, *30*, 97−106.
(20) Penning, T. M. *Chem. Rev.* **2001**, *101* (10), 3027−3046.
(21) Daugherty, P. S.; Chen, G.; Iverson, B. L.; Georgiou, G. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2029−2034.
(22) Raleigh, D. P.; Betz, S. F.; DeGrado, W. F. *J. Am. Chem. Soc.* **1995**, *117*, 7558−7559.
(23) Kortemme, T.; Ramirez-Alvarado, M.; Serrano, L. *Science* **1998**, *281*, 253−256.
(24) Obataya, I.; Kotaki, T.; Sakamoto, S.; Ueno, A.; Mihara, H. *Bioorgan. Med. Chem. Lett.* **2000**, *10*, 2719−2722.
(25) Ho, S. P.; DeGrado, W. F. *J. Am. Chem. Soc.* **1987**, *109*, 6751−6758.
(26) Struthers, M. D.; Cheng, R. P.; Imperiali, B. *Science* **1996**, *271*, 342−345.
(27) Moffet, D. A.; Hecht, M. H. *Chem. Rev.* **2001**, *101* (10), 3191−3204.
(28) Beasley, J. R.; Hecht, M. H. *J. Biol. Chem.* **1997**, *272*, 2031−2034.
(29) Matsuura, T.; Ernst, A.; Zechel, D. L.; Plückthum, A. *ChemBioChem* **2004**, *5*, 177−182.
(30) Neylon, C. *Nucleic Acids Res.* **2004**, *32*, 1448−1459.
(31) Taylo, S. V.; Kast, P.; Hilvert, D. *Angew. Chem., Int. Ed.* **2001**, *40*, 3310−3335.
(32) Ostermeier, M.; Nixon, A. E.; Shim, J. H.; Benkovic, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3562−3567.
(33) Riechmann, L.; Winter, G. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 10068−10073.
(34) Viljanen, J.; Tegler, L.; Broo, K. S. *Bioconjugate Chem.* **2004**, *15*, 718−727.
(35) DeGrado, W. F. *Chem. Rev.* **2001**, *101* (10), 3025−3026.
(36) Qi, D.; Tann, C.-M.; Haring, D.; Distefano, M. D. *Chem. Rev.* **2001**, *101* (10), 3081−3111.
(37) Zawadzke, L. E.; Berg, J. M. *J. Am. Chem. Soc.* **1992**, *114*, 4002−4003.
(38) Milton, R. C.; Milton, S. C.; Kent, S. B. *Science* **1992**, *256*, 1445−1448.
(39) Guichard, G.; Muller, S.; Regenmortel, M.; Briand, J. P.; Mascagni, P.; Giralt, E. *TIBTECH* **1996**, *14*, 44−45.
(40) Wiesehan, K.; Willbold, D. *ChemBioChem* **2003**, *4*, 811−815.
(41) Nilsson, B. L.; Soellner, M. B.; Raines, R. T. *Annu. Rev. Bioph. Biom. Struct.* **2005**, *34*, 91−118.
(42) Pastor, J. J.; Fernandez, I.; Rabanal, F.; Giralt, E. *Org. Lett.* **2002**, *4*, 3831−3833.
(43) Meldal, M. *Tetrahedron Lett.* **1992**, *33*, 3077−3080.
(44) (a) Rabanal, F.; Giralt, E.; Albericio, F. *Tetrahedron* **1995**, *51*, 1449−58. (b) Albericio, F.; Cruz, M.; Debethune, L.; Eritja, R.; Giralt, E.; Grandas, A.; Marchan, V.; Pastor, J. J.; Pedroso, E.; Rabanal, F.; Royo, M. *Synth. Commun.* **2001**, *31*, 225−232.
(45) Leister, W.; Strauss, K.; Wisnoski, D.; Zhao, Z.; Lindsley, C. *J. Comb. Chem.* **2003**, *5*, 322−329.
(46) Yurek, D. A.; Branch, D. L.; Kuo, M.-S. *J. Comb. Chem.* **2002**, *4*, 138−148.
(47) Lam, K. S.; Lebl, M.; Krchnák, V. *Chem. Rev.* **1997**, *97*, 411−448.
(48) Lebl, M.; Krchnák, V.; Sepetov, N. F.; Seligmann, B.; Strop, P.; Felder, S.; Lam, K. S. *Biopolymers* **1995**, *37*, 177−198.
(49) (a) Halkes, K. M.; St. Hilaire, P. M.; Crocker, P. R.; Meldal, M. *J. Comb. Chem.* **2003**, *5*, 18−27. (b) Yingyongnarongkul, B.; Howarth, M.; Elliott, T.; Bradley, M. *J. Comb. Chem.* **2004**, *6*, 753−760.
(50) (a) Tasayco, M. L.; Carey, J. *Science* **1992**, *255*, 594−597. (b) Shiba, K.; Schimmel, P. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 1880−1884. (c) de Prat Gay, G.; Fersht, A. R. *Biochemistry* **1994**, *33*, 7957−7963. (d) Yu, H.; Kono, M.; McKee, T. D.; Oprian, D. D. *Biochemistry* **1995**, *34*, 14963−14969. (e) Hamachi, I.; Hiraoka, T.; Yamada, Y.; Shinkai, S. *Chem. Lett.* **1998**, *6*, 537−538. (f) Ghosh, I.; Hamilton, A. D.; Regan, L. *J. Am. Chem. Soc.* **2000**, *122*, 5658−5659. (g) Magliery, T. J.; Wilson, C. G. M.; Pan, W.; Mishler, D.; Ghosh, I.; Hamilton, A. D.; Regan, L. *J. Am. Chem. Soc.* **2005**, *127*, 146−157.
(51) Nicolas, E.; Ferrer, C.; Taboada, L.; Giralt, E. *J. Am. Chem. Soc.* **2005**, *127*, 17719−17733.

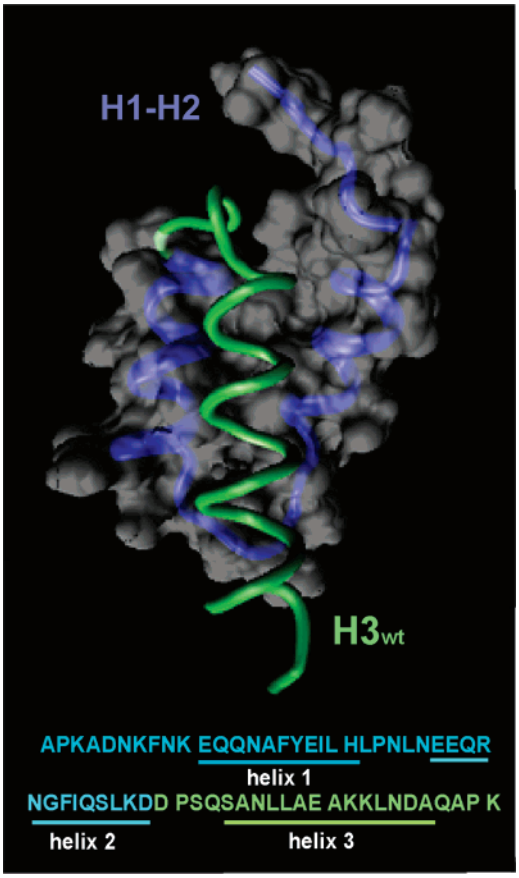


Figure 1. Structure of the B domain of staphylococcal protein A using data of Gouda et al. (PDB 1BDD).[56,57] The three-helix bundle is represented by ribbons. $\mathbf{H3_{wt}}$ is shown in green, the two N-terminal α-helices (**H1**−**H2**), in blue, and the van der Waals surface, in gray. The image was created using VMD.[53]

structure.[52] It is for this reason that we aimed at testing our strategy by redesigning the hydrophobic core of the B domain of staphylococcal protein A.

Results and Discussion

Design and Optimization of the Strategy. We chose the B domain of staphylococcal protein A as a model to develop our strategy.[54] The structure of the B domain is a three-helix bundle that comprises 61 amino acids (see Figure 1). The main known activity of the B domain is its ability to complex the Fc constant region of mammalian IgG with high affinity. This interaction is centered on one of the faces that make up the two N-terminal α-helices of the domain (denoted here as **H1**−**H2**). The C-terminal helix (i.e., third helix of the bundle denoted here as $\mathbf{H3_{wt}}$) though not directly involved in the interaction with IgG is required to maintain the folding and the activity of the domain.[55−57] The final goal of this work was to redesign the hydrophobic core of this protein.

General approaches to this problem would imply the synthesis and screening of libraries of entire modified domains.[58] Due to the large size of the target protein (61 amino acids), the use of chemical methods would require the individual synthesis of every modified domain, in order to guarantee the purity of the compounds synthesized and tested. This approach would therefore impose great restrictions on the diversity to be explored, as preparation of each library member would require a large amount of time. To overcome this problem we designed the strategy detailed in Figure 2, focused on a screening test based on noncovalent intra- and interprotein interactions.

According to this strategy, one-bead-one-compound libraries of $\mathbf{H3_{wt}}$ analogues would be synthesized on HMFS-PEGA resin and incubated with a solution that contained the peptide comprising the two natural N-terminal α-helices of the B domain (peptide **H1**−**H2**). When the interaction between $\mathbf{H3_{wt}}$ analogues and **H1**−**H2** would be strong and specific enough, noncovalent *pseudo*-B domains would get formed on the resin. Beads containing peptides folded in an active conformation could then be easily identified using a solution of fluorescently labeled IgG, since those modified *pseudo*-B domains that bind labeled IgG would generate fluorescence. Finally, manual separation of fluorescent beads followed by peptide sequencing by MALDI-TOF PSD[59] of the peptide contained on individual beads would reveal the identity of active $\mathbf{H3_{wt}}$ mimics (i.e., the peptides able to bind the natural N-terminal helices of the B domain, fold like the natural protein, and bind the immunoglobulin).

While optimizing the conditions for a suitable screening test, the first step in developing our strategy was to validate our hypothesis that the noncovalent interactions established in our model system were strong and specific enough. We thus started by incubating samples which contained 10 resin beads bearing either $\mathbf{H3_{wt}}$ or $\mathbf{H3_s}$ (scrambled sequence) with solutions of **H1**−**H2** and FITC-IgG, at different concentrations and different times, in pH = 7.4−7.5 phosphate buffer.[56,57,60−62] The **H1**−**H2** concentration ranged from 10 to 100 μM, whereas the FITC-IgG dilutions ranged from 1/1000 to 1/100 000. Time was kept constant for the incubation with the **H1**−**H2** solution and then varied from 1 to 6 h for the incubation with the FITC-IgG solution.

As shown in Figure 3, the beads bearing $\mathbf{H3_{wt}}$ showed high-intensity fluorescence, while the beads with nonactive peptide had only background levels of fluorescence. Optimal response was obtained for beads first incubated with a 100 μM solution of **H1**−**H2** and then with a 1/1000 solution of FITC-IgG for 6 h. These results validated our hypothesis that intradomain noncovalent interactions were sufficient to allow formation of *pseudo*-B domains with conserved IgG binding activity. We confirmed this mechanism of activity by performing competition experiments using increasing amounts of $\mathbf{H3_{wt}}$ in solution during the incubation of the peptide-resin with **H1**−**H2**. As expected, we observed a marked decrease in the fluorescence intensity of the positive beads that was directly proportional to the concen-

(52) Baltzer, L.; Nilsson, H.; Nilsson, J. *Chem. Rev.* **2001**, *101*, 3153−3163.
(53) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33−38.
(54) Björk, I.; Petersson, B.-A.; Sjöquist, J. *Eur. J. Biochem.* **1972**, *29*, 579−584.
(55) Moks, T.; Abrahmsén, L.; Nilsson, B.; Hellman, U.; Sjöquist, J.; Uhlén, M. *Eur. J. Biochem.* **1986**, *156*, 637−643.
(56) Gouda, H.; Torigoe, H.; Saito, A.; Sato, M.; Arata, Y.; Shimada, I. *Biochemistry* **1992**, *31*, 9665−9672.
(57) Gouda, H.; Shiraishi, M.; Takahashi, H.; Kato, K.; Torigoe, H.; Arata, Y.; Shimada, I. *Biochemistry* **1998**, *37*, 129−136.
(58) Hoess, R. H. *Chem. Rev.* **2001**, *101* (10), 3205−3218.
(59) Spengler, B. *J. Mass Spectrom.* **1997**, *32*, 1019−1036.
(60) Torigoe, H.; Shimada, I.; Saito, A.; Sato, M.; Arata, Y. *Biochemistry* **1990**, *29*, 8787−8793.
(61) Cedergren, L.; Andersson, R.; Jansson, B.; Uhlén, M.; Nilsson, B. *Protein Eng.* **1993**, *6*, 441−448.
(62) Kato, K.; Gouda, H.; Tanaha, W.; Yoshino, A.; Matsunaga, C.; Arata, Y. *FEBS Lett.* **1993**, *328*, 49−54.

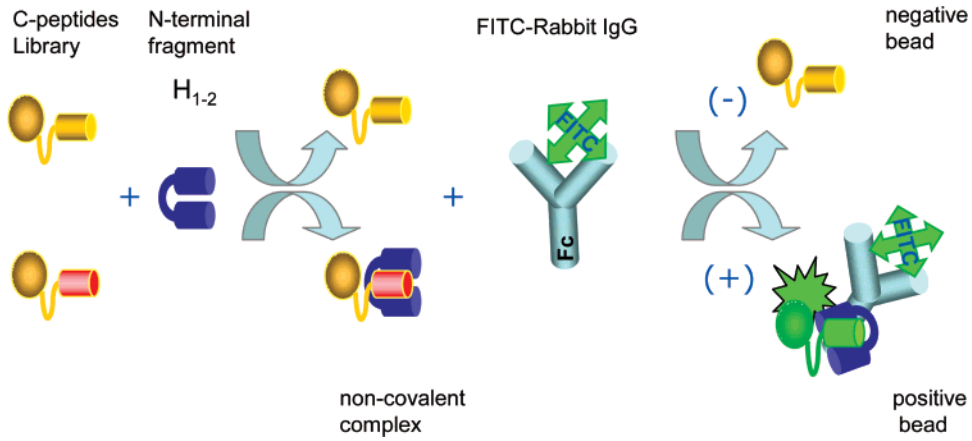


Figure 2. Schematic diagram depicting the designed strategy to screen a one-bead-one-compound library of $\mathbf{H3_{wt}}$ analogues. The strategy is based on incubation of the library beads with the N-terminal fragment of the B domain and fluorescein labeled rabbit IgG (FITC-IgG). Positively stained beads would be detected by visualization under a fluorescence bifocal magnifier.

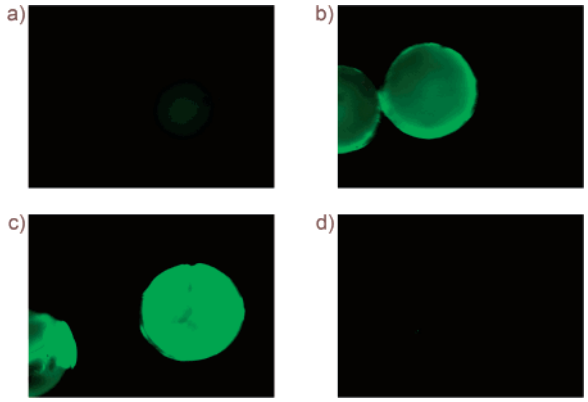


Figure 3. Fluorescence microscopy images of beads containing $\mathbf{H3_{wt}}$ and $\mathbf{H3_s}$ and submitted to different assay conditions. Pictures (a) to (c) correspond to beads bearing $\mathbf{H3_{wt}}$ and tested under the following conditions: (a) 10 μM of **H1**−**H2** and 1/1000 of FITC-IgG (1 h), (b) 100 μM of **H1**−**H2** and 1/1000 of FITC-IgG (1 h), (c) 100 μM of **H1**−**H2** and 1/1000 of FITC-IgG (6 h). Picture (d) correspond to beads bearing the negative control $\mathbf{H3_s}$ and tested under the same conditions as those for (c). Microscope camera acquisition parameters for these images were 100% gain, 0.3 s exposition time, and ×100 magnification.

tration of $\mathbf{H3_{wt}}$ in solution. Final corroboration was obtained by carrying out the experiment in the absence of **H1**−**H2**. No fluorescence was detected from any bead containing $\mathbf{H3_{wt}}$ or $\mathbf{H3_s}$, meaning that **H1**−**H2** was required for the observation of fluorescence and that the interaction of the N-terminal helices with the resin bound natural C-terminal helix (i.e., formation of noncovalent wild-type B domains) was responsible for the immobilization of IgG on the positive beads.

To determine if these conditions could be extended to the screening of libraries of $\mathbf{H3_{wt}}$ analogues, we use them to test four different samples of *ca.* 1000 beads each containing a 1:100 ratio of beads bearing $\mathbf{H3_{wt}}$ to beads bearing $\mathbf{H3_s}$. We observed intense fluorescence on 7 to 12 beads from each sample. These positive beads were then manually isolated from the mixture by micropipette, washed, and transferred to individual positions of a MALDI-TOF sample plate. The peptide contained in each bead was then directly released by treatment with 20% morpholine in DMF and sequenced by MALDI-TOF PSD. For all samples good fragmentation spectra were recorded and the peak pattern corresponded to $\mathbf{H3_{wt}}$. The results clearly supported the suitability of our strategy for the modification of the C-terminal α-helix of the B domain using one-bead-one-compound libraries of $\mathbf{H3_{wt}}$ analogues.

Design of L300 Library. Modification of $\mathbf{H3_{wt}}$ using our strategy was attempted with a library of 300 compounds of the general structure shown in Figure 4. The library was designed with a double purpose: first, to study the interactions of the amino acids located in the hydrophobic core of the protein and the tolerance to their modification; and second, to obtain new proteins based on the three-helix bundle motif with similar structure and activity to those for the natural B domain. The sequence of $\mathbf{H3_{wt}}$ was also included as a positive control to validate the methodology.

Besides keeping the natural residues of $\mathbf{H3_{wt}}$, the selection of the type of residue at the four variable positions was based on the three-dimensional structure of the natural B domain and the following general criteria. In position X44 (Ser in the natural B domain), Thr and Phe were included to assess whether the hydroxyl functionality or a hydrophobic group was preferred in stabilizing the three-helix bundle motif. In positions X47, X51, and X54 (Leu, Ala, and Leu, respectively, in the natural protein) we decided to map standard natural amino acid functionalities such as hydrophobic (Val or Phe), polar (Ser), and charged side chains (Lys or Glu). Pro was also included at positions X47 and X54 to limit the number of hypothetical possible positive sequences due to its well-known role of α-helix breaker.

The one-bead-one-compound designed library was synthesized on 300 mg of HMFS-PEGA resin[40] (0.2 mmol/g, ∼37 000 beads)[63,64] using standard split and mix procedures. The amino acids were incorporated via the Boc/Bn protection scheme, with the coupling conditions already described. Final side-chain deprotection was performed by HF treatment without magnetic stirring, to avoid peptide-resin damage.

The synthetic efficiency of the library was evaluated by (i) amino acid analysis of a sample of the peptide-resin containing all the members of the library; (ii) MS analysis of the peptide mixture released from a sample of the library containing all the members; and (iii) individual MALDI-TOF PSD sequencing of the peptide contained on each of a representative number of

(63) Zhao, P.-L.; Zambias, R.; Bolognese, J. A.; Boulton, D.; Chapman, K. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 10121−10126.
(64) Zhao, P.-L.; Nachbar, R. B.; Bolognese, J. A.; Chapman, K. *J. Med. Chem.* **1996**, *39*, 350−352.

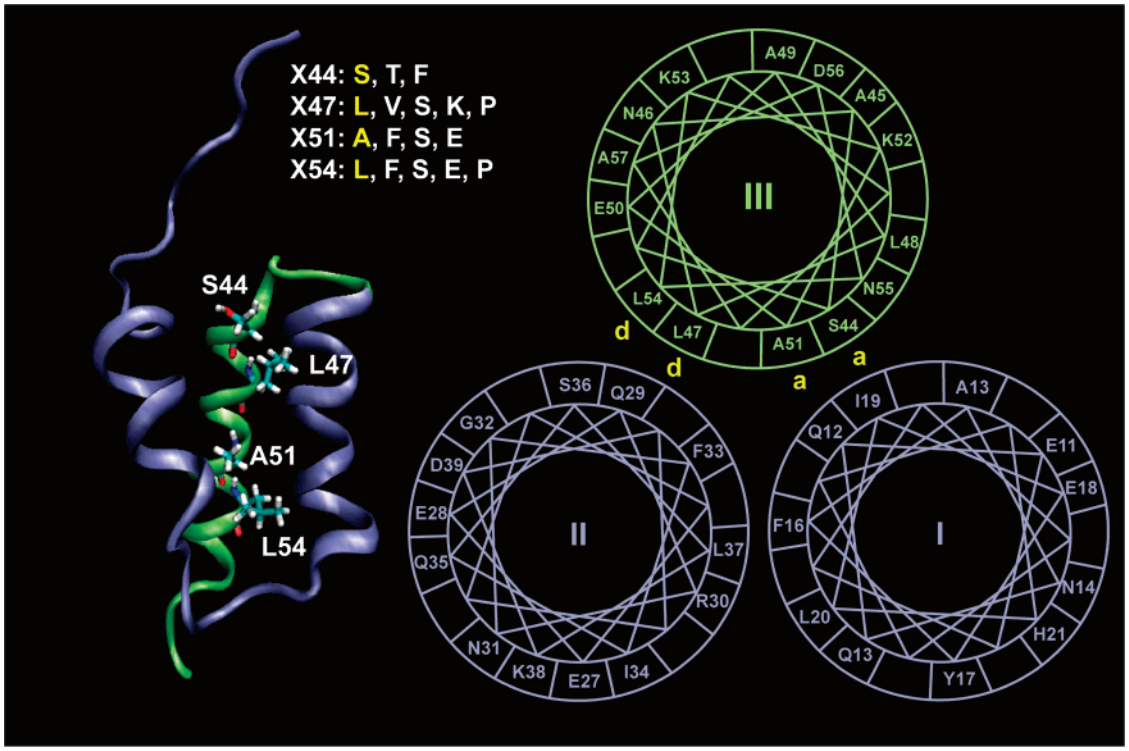
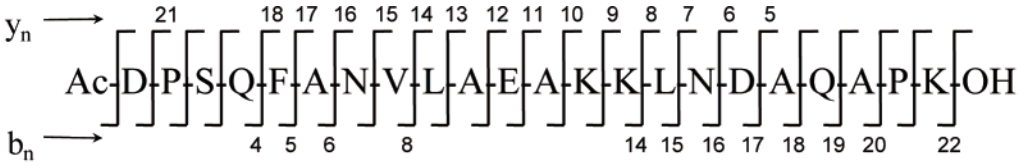


Figure 4. Design of the L300 library. The C-terminal helix (**H3$_{wt}$**) was subjected to a combinatorial design to construct a library consisting of 300 different peptides of 22 amino acids each. The variations were at four positions of the hydrophobic core of the helix bundle as shown above (positions: X44, X47, X51, and X54, corresponding to amino acids located in positions **a** and **d** of to the Edmunson wheel representation of the helix, shown on the right). The residues present in the wild type protein are shown in yellow. The design was based on the NMR structure of the B domain (shown on the left).



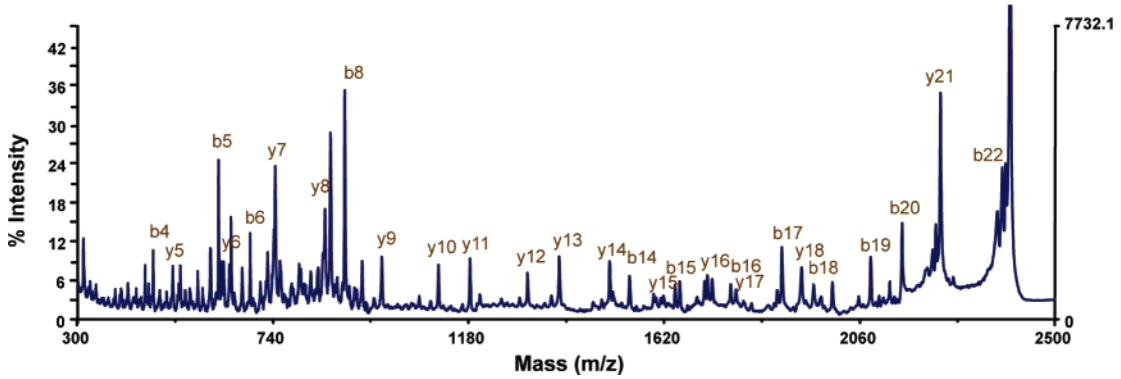


Figure 5. Example of a MALDI-TOF PSD spectrum recorded during the analysis of a peptide obtained from an isolated positive bead from library L300. The sequence deduced from the fragmentation pattern as well as the identified fragments are shown on the top.

library beads (20 beads). The amino acid composition of the library was in good agreement with the theoretical composition, and the shape and mass limits of the experimental and simulated MS spectra for the library were almost identical (see Supporting Information). Finally, the 20 beads analyzed by MALDI-TOF PSD showed a single main product of high purity and an experimental sequence matching a single theoretical member of the library (Figure 5). These results, supported by the purity of crude peptides **H3$_{wt}$** and **H3$_s$** previously synthesized by this methodology (see Supporting Information), clearly indicated correct synthesis of the library and ensured that each of the expected members of the library was present in equimolar proportion and in good purity.

Four portions of the library (*ca.* 1000 beads per sample) were independently screened under the optimized conditions previously described. Between 1 and 2% of the beads showed varied high-intensity fluorescence upon FITC-IgG incubation (Figure 6). The most bright beads (56 beads) were manually isolated

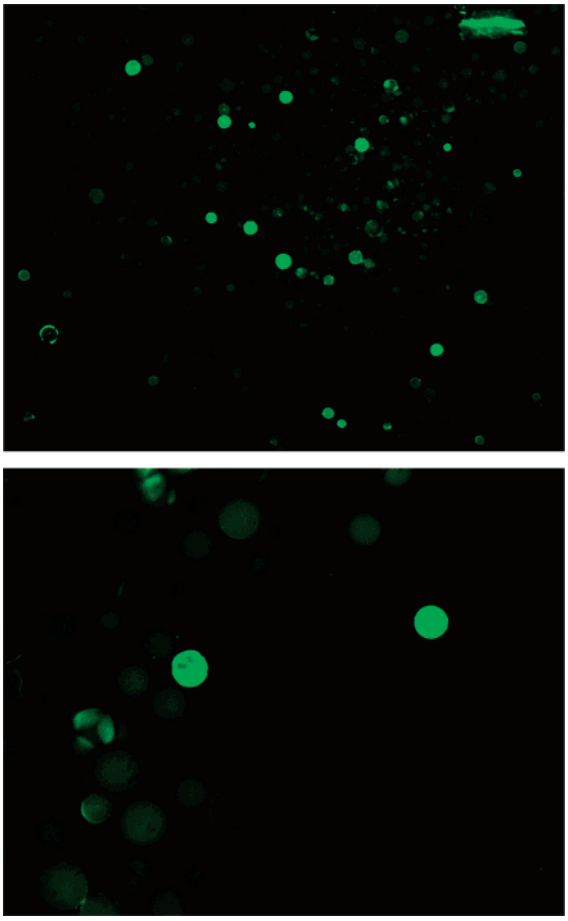


Figure 6. Fluorescence microscopy images of two screened samples of beads from the library L300. Highly fluorescent (positive) beads are easily distinguished.

Table 1. Positive Sequences Identified from the Screening of the L300 Library[a]

peptide	sequence	no. times identified
H3$_{P1}$	Ac−DPSQ***F***AN***K***LAE***A***KK***L***NDAQAPK−OH	11
H3$_{P2}$	Ac−DPSQ***F***AN***V***LAE***A***KK***L***NDAQAPK−OH	11
H3$_{P3}$	Ac−DPSQ***F***AN***V***LAE***A***KK***F***NDAQAPK−OH	7
H3$_{P4}$	Ac−DPSQ***T***AN***L***LAE***A***KK***L***NDAQAPK−OH	7
H3$_{P5}$	Ac−DPSQ***F***AN***L***LAE***A***KK***L***NDAQAPK−OH	6
H3$_{P6}$	Ac−DPSQ***S***AN***V***LAE***A***KK***L***NDAQAPK−OH	5
H3$_{P7}$	Ac−DPSQ***T***AN***V***LAE***A***KK***L***NDAQAPK−OH	4
H3$_{P8}$	Ac−DPSQ***F***AN***V***LAE***A***KK***S***NDAQAPK−OH	1
H3$_{P9}$ (H3$_{wt}$)	Ac−DPSQ***S***AN***L***LAEAKK***L***NDAQAPK−OH	1

[a] Amino acids corresponding to variable positions in the library are represented in italic and bold. The number of times each sequence was identified in the 53 sequenced beads is also specified.

from the mixture and washed. Active peptides were directly released on a MALDI-TOF sample plate and sequenced by PSD, as described above for **H3$_{wt}$**. High quality fragmentation spectra were recorded in most cases (see Supporting Information) allowing the identification of 53 full sequences out of the 56 isolated beads (Table 1).

The 53 identified sequences correspond to only nine different peptides from the library and include the natural **H3$_{wt}$**, which was used as a positive control, thus confirming the validity of the developed strategy. The frequency with which some of the sequences were found is further evidence of the robustness of the method: if there had been false positives (i.e., nonspecific responses), then random sequences would have been obtained, with poor or null probability of repetition.[47] Indeed, nearly all of the positive sequences appeared with a higher frequency than that for **H3$_{wt}$**, suggesting that they had stronger interactions in the noncovalent system (**H1**−**H2** + **H3$_{wt}$** or analogue + IgG) than **H3$_{wt}$** and thus merited further study. We therefore chose the two sequences identified a higher number of times in the library, **H3$_{P1}$** (S44F, L47K) and **H3$_{P2}$** (S44F, L47V), to verify the results from the library screening.

Test of Lead Peptides. To verify the results from the library, **H3$_{P1}$**, **H3$_{P2}$**, **H3$_{wt}$**, and **H3$_{s}$** were individually synthesized as free C-terminal amidopeptides via standard Fmoc/tBu procedures.

Evaluation of the interaction between the different versions of the C-terminal α-helix of the B domain and the natural N-terminal helices was first studied by circular dichroism. The results obtained are summarized in Figure 7. **H3$_{wt}$** and **H1**−**H2**, as well as **H3$_{P1}$** and **H3$_{P2}$**, tended to adopt helical structures in solution, whereas **H3$_{s}$** remained as a random coil. No self-interaction was detected for any of the peptides, as demonstrated by the lack of change in the molar ellipticity of the CD spectra recorded at higher concentrations (data not shown). However, introduction of **H1**−**H2** into the C-terminal helix samples in a 1:1 ratio increased the amount of α-helix detected in solution for all peptides except for the negative control (**H3$_{s}$**), for which no difference between the recorded spectrum and the addition spectra of the individual components of the mixture was detected (Figure 7A). The difference between the recorded and the arithmetically simulated spectra for the other mixtures revealed that **H3$_{wt}$** and its active analogues interact with **H1**−**H2**. Comparison of the spectra for each mixture indicated **H3$_{wt}$** and **H3$_{P1}$** had similar capacities to form noncovalent domains with **H1**−**H2**, whereas the strongest interaction was for **H3$_{P2}$** (Figure 7D). The domain formed by **H3$_{P2}$** and **H1**−**H2** was the most stable assembly as indicated by the magnitude and cooperativity of the unfolding melting curves recorded at 222 nm for all the mixtures (see Supporting Information).

CD data provided evidence that intradomain noncovalent interactions are established and importantly that these lead to the formation of α-helical structure, the type of structure adopted by the wild type B domain. This further validates the dissection strategy we have followed to redesign the hydrophobic core of the B domain.

Study of the Covalent Versions of the Modified B Domain. One of the aims of this work was to obtain new proteins with similar structure and activity to those of the natural B domain. To determine if we had achieved the aforementioned aim, we chemically synthesized the natural B domain as well as the modified B domains that include the mutations found in **H3$_{P1}$** and **H3$_{P2}$**. These covalent B domains (see Table 2) were prepared on a Rink amide MBHA resin via Fmoc/^{t}Bu standard procedures. The synthesized proteins were purified by semi-preparative reversed-phase HPLC, and their structures were studied by CD and NMR (^{1}H, TOCSY and NOESY).

CD spectra for the three versions of the covalent B domain evidenced the characteristic signature of an α-helix (Figure 8). As expected, the presence of a covalent bond between **H1**−**H2** and **H3$_{x}$** resulted in a clear increase in the α-helix content when compared to the noncovalent domains (see Figure 7 and Supporting Information Figure 15). In spite of this, the helicity ranking of the three analogues is different whether in nonco-

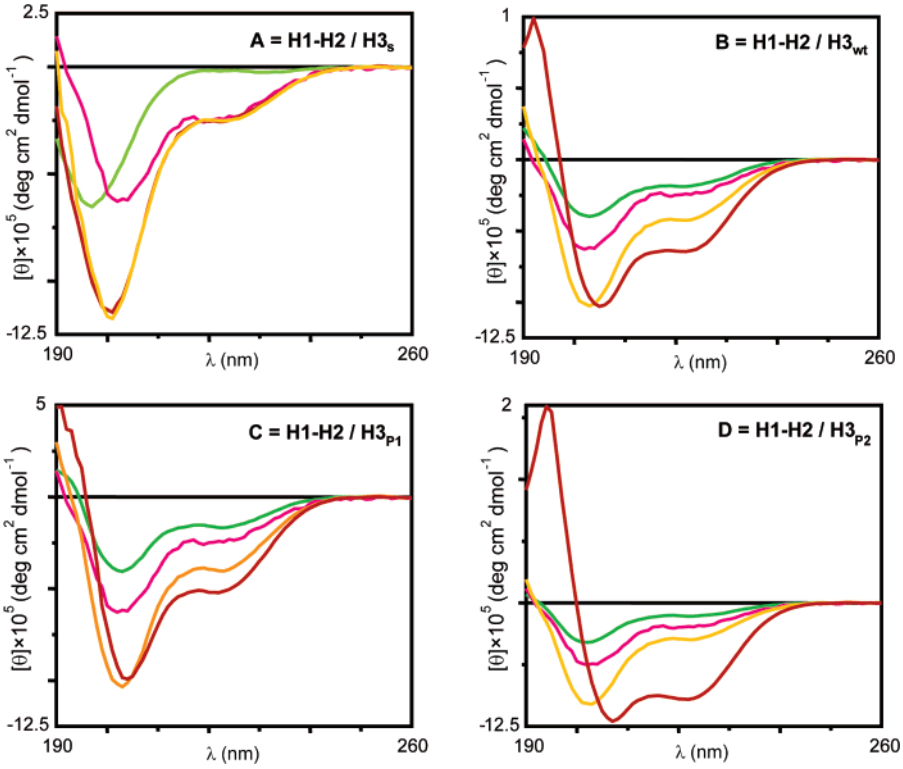


Figure 7. Circular dichroism study of the interaction between peptides **H3$_{wt}$**, **H3$_{P1}$**, **H3$_{P2}$**, or **H3$_s$** with the N-terminal α-helices of the B domain (**H1**−**H2**). (A) **H3$_s$** and **H1**−**H2** interaction; (B) **H3$_{wt}$** and **H1**−**H2** interaction; (C) **H3$_{P1}$** and **H1**−**H2** interaction; (D) **H3$_{P2}$** and **H1**−**H2** interaction. Color key: (i) green: C-terminal helix peptides (**H3$_s$**, **H3$_{wt}$**, **H3$_{P1}$**, and **H3$_{P2}$** for A, B, C, and D, respectively); (ii) magenta: **H1**−**H2**; (iii) orange: arithmetic addition spectra of **H3$_x$** analogues and **H1**−**H2**; and (iv) mars red: experimental spectra from equimolar mixtures.

Table 2. Sequences of the Three Versions of the Covalent B Domain Synthesized (Wild-Type Plus **H3$_{P1}$** and **H3$_{P2}$** Modified Domains)[a]

peptide	sequences
H3$_s$	Ac−AQKQASPDALEKNLPKAD NASL−NH_2
H1−**H2**	Ac−APKADNKFNKEQQNAFYEI LHLPNLNEEQRNGFIQSLKD−NH_2
H1−**H2**−**H3$_{wt}$**	Ac−APKADNKFNKEQQNAFYEILH LPNLNEEQRNGFIQSLKDDPSQS ANLLAEAKKLNDAQAPK−NH_2
H1−**H2**−**H3$_{p1}$**	Ac−APKADNKFNKEQQNAFYEILHLP NLNEEQRNGFIQSLKDDPSQFANKL AEAKKLNDAQAPK−NH_2
H1−**H2**−**H3$_{p2}$**	Ac−APKADNKFNKEQQNAFYEILHLPN LNEEQRNGFIQSLKDDPSQFANVLAEA KKLNDAQAPK−NH_2

[a] The sequences of the scrambled version of the natural C-terminal α-helix, **H3$_s$**, and of the peptide containing the two N-terminal helices of the B domain (**H1**−**H2**) are also specified.

valent or covalent forms. In the covalent form, the wild-type domain and the **H1**−**H2**−**H3$_{P2}$** protein showed a similar degree of α-helical structuration, whereas the **H1**−**H2**−**H3$_{P1}$** domain displayed a marked increase. This is in contrast to the results recorded for the noncovalent assembly domains where **H1**−**H2** + **H3$_{P2}$** presented a higher degree of α-helical structure than **H1**−**H2** + **H3$_{P1}$** and **H1**−**H2** + **H3$_{wt}$**. To explain this difference, it is important to consider the noncovalent nature of our approach. The higher conformational space available in the noncovalent systems may allow certain interactions that promote the structure to become partially/totally disrupted in the covalent form because of the conformational restrain that the covalent bond imposes. For example, while **H3$_{wt}$** interacts with **H1**−**H2** in an antiparallel manner in the wild-type domain, the sequence of **H3$_{wt}$** is highly palindromic. Thus, in the noncovalent system, both parallel and antiparallel interactions of **H3$_x$** with **H1**−**H2** could occur. If active parallel complexes **H1**−**H2** + **H3$_x$** were found during the screening test, they would become antiparallel in the covalent form which could result in a less structured molecule than expected. From this analysis we would like to point out that the choice of which covalent bond or bonds are removed from the protein is very important when applying this strategy to the modification of other globular proteins, since it can greatly influence the amount and quality of positive sequences found.

CD thermal denaturation, monitored at 222 nm, showed that the three domains unfolded in a cooperative manner, a characteristic behavior of native proteins. These data allowed T_m determination revealing that all the covalent domains were thermally stable, the most stable one being **H1**−**H2**−**H3$_{P2}$** with a T_m 9 °C higher than that for the natural protein. This result is not contradictory with the fact that **H1**−**H2**−**H3$_{P2}$** is not the most structured at 5 °C. The observed lack of correlation between thermal stability and amount of structure is similar to that reported for other *de novo* systems.[65]

Thus, CD analysis revealed that the two new proteins adopt an α-helical structure as the wild-type protein does. In order to learn about the overall fold of these new proteins, that is if they adopt a three helix bundle fold as the wild-type protein does, NMR analysis of the covalent domain was carried out (see Figure 9).[66]

(65) Walsh, S. T.; Sukharev, V. I.; Betz, S. F.; Vekshin, N. L.; DeGrado, W. F. *J. Mol. Biol.* **2000**, *305*, 361−373.
(66) Gibney, B. R.; Rabanal, F.; Skalicky, J. J.; Wand, A. J.; Dutton, P. L. *J. Am. Chem. Soc.* **1997**, *119*, 2323−2324.

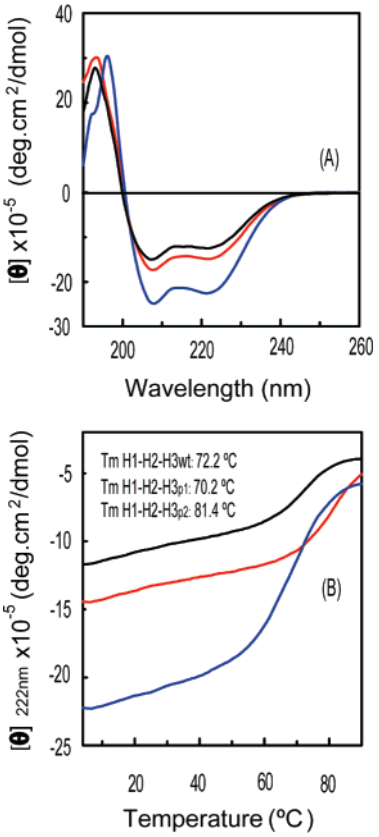


Figure 8. CD spectra (top) and melting curves recorded at 222 nm (bottom) of the three B domains synthesized. Color key: (i) black: **H1−H2−H3$_{wt}$**; (ii) blue: **H1−H2−H3$_{P1}$**; and (iii) red: **H1−H2−H3$_{P2}$**.

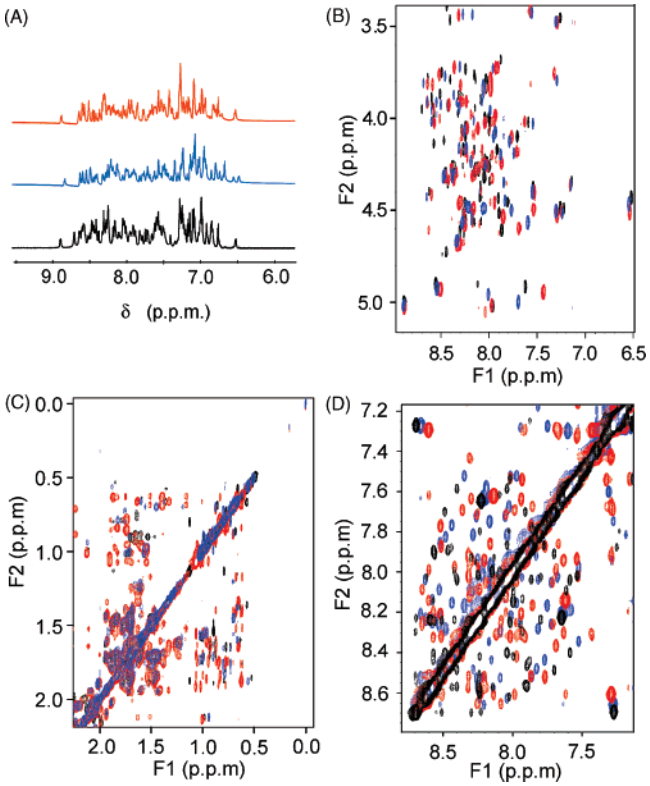


Figure 9. NMR spectra of the three covalent B domains studied. (A) monodimensional ^{1}H NMR of the amide/aromatic region; (B) fingerprint of the CαH-amide region of the TOCSY spectra; (C) fingerprint of the methyl region of the TOCSY spectra; and (D) amide−amide region of the NOESY spectra. Color key: (i) black: **H1−H2−H3$_{wt}$**; (ii) blue: **H1−H2−H3$_{P1}$**; and (iii) red: **H1−H2−H3$_{P2}$**.

Comparison of mono- and bidimensional proton experiments showed a large dispersion of both the NH backbone signals and the high field methyl signals indicating that the proteins adopt a single, well-folded structure. Chemical shifts of the synthesized wild-type domain are the same as those previously reported in the literature.[56,57] The fact that the redesigned proteins also show similar chemical shifts and peak widths at half-height to the wild-type protein suggests that, under the conditions at which the NMR data was measured, the proteins are monomeric. Preliminary evidence that the new redesigned proteins adopt a three helix bundle structure as the wild-type domain does come from the superposition of TOCSY spectra that revealed similar chemical shifts for the three domains and from the identification of intense sequential NH−NH NOE cross-peaks in similar areas of the three proteins, characteristic of an α-helical structure. Further characterization of the three proteins by NMR is currently being carried out. This will reveal any differences among the three proteins at the atomic level. In spite of this, the NMR data analyzed so far allow us to confirm that the analogues adopt a single three helix bundle globular structure similar to that of the natural B domain. The fact that the two new redesigned proteins present a similar fold to the native protein may imply that they also present a similar activity to the wild-type protein.

Analysis of the Hydrophobic Core of the B Domain. With the analysis presented so far, we have proven that the two new redesigned proteins are structurally similar to the natural B domain. However, another aim of the work was to study the type of interactions present in the hydrophobic core and the tolerance to its modification. Even in the absence of real three-dimensional structures for the modified B domains, some interesting structural information about the protein core can be deduced from analysis of the active sequences. Figure 10 shows a graphical representation of the percentage of each amino acid found in the randomized positions of the sequence.

Natural Ala 51 and Leu54 are highly conserved in all the active peptides found in the library. This result clearly suggests that, in this area, the protein core is specifically packed and has little tolerance for amino acid substitutions. Hydrophobic interactions seem to be predominant in both positions, as deduced from the absence of polar or charged functionalities on the positive sequences as well as their location in the central hydrophobic core of the protein (positions a and d, Figure 4). Position 44 of the sequence has a clear preference for Phe over the natural Ser of the domain (nonconservative change). Hydrophobic interactions for the amino acid in this position, located in the natural protein near the loop between the second and third α-helices (see Figure 4), seem to be more important for the stability of the bundle than the possible interactions involving the hydroxyl group of the natural serine. This hypothesis is corroborated by the presence of mainly hydrophobic amino acids in the area near position 44, such as Phe8 and Leu37. Finally, natural Leu47 is conserved or exchanged for Val, a conservative change as the latter has a similar hydrophobic side chain to that of the former. However, and more surprising, a significant substitution for Lys was also found, which at first sight represents the insertion of a positively charged amino acid into the hydrophobic protein core. Interestingly, the **H1−H2−H3$_{P1}$** protein carrying the Lys mutation is

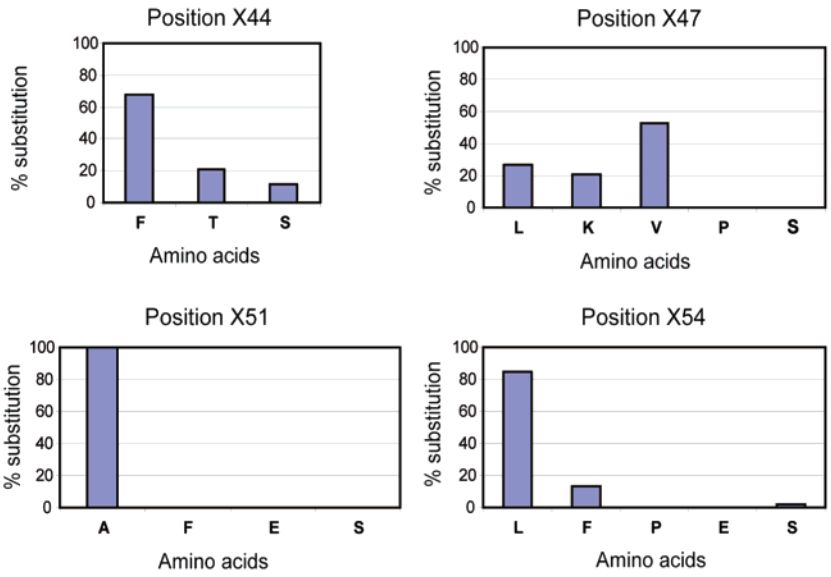


Figure 10. Graphical representation of the amino acids found at randomized positions in the active sequences obtained from the screening of library L300. Substitution is illustrated as a precentage of each amino acid at a given position, referenced to the 53 active peptide sequences.

the one with the highest α-helical structures by means of CD. This may indicate that some sort of ionic interaction involving the Lys side chain may stabilize the protein structure.[67,68] Detailed characterization of this structure is currently underway. From this analysis, we conclude that the strategy used has successfully allowed the redesign of the hydrophobic core of the B domain of protein A using some nonconservative mutations. Analysis of the two sequences most found in the library in a covalent form has allowed us to determine that these nonconservative mutations are accommodated in the hydrophobic core and that they lead to structurally very similar proteins to the wild-type protein.

Conclusions

We have developed a novel combinatorial strategy for the redesign of natural proteins. The strategy is based on the premise that noncovalent interactions among the different secondary elements of a protein are both sufficiently strong and specific to induce the correct folding of the protein and to retain activity even in the absence of a covalent link present in the native protein. The strategy thus thought has two important implications. First, the protein can be cut in two fragments allowing the use of one-bead-one-compound combinatorial chemistry to modify one of the fragments and thus overcome the drawbacks associated with the synthesis of large proteins. Also, by applying a synthetic methodology previously developed in our laboratory,[42] the modified protein fragment can be synthesized unprotected and attached to a biocompatible resin which permits the screening of the peptides while they are bound to the resin. Second, since one of the premises of the strategy is that the noncovalent version of the domain maintains the activity of the natural sequence, a screening test based on selecting active sequences can be used. We have successfully tested this strategy by redesigning the hydrophobic core of the B domain of protein A.

Using the aforementioned strategy, we have synthesized and evaluated a library of 300 $\mathbf{H3_{wt}}$ analogues with mutations at positions 44, 47, 51, and 54 of the B domain. Screening of the library has yielded eight different active sequences that contain some nonconservative mutations in addition to the natural $\mathbf{H3_{wt}}$, included as a positive control to validate our approach. The most frequently found sequences (peptides $\mathbf{H3_{P1}}$ and $\mathbf{H3_{P2}}$) were further studied. CD studies of these two peptides in the presence of **H1**−**H2** proved that the active peptides interacted with the remainder of the natural domain. The redesign strategy was further validated by the chemical synthesis and subsequent structural study of $\mathbf{H1-H2-H3_{wt}}$, $\mathbf{H1-H2-H3_{P1}}$, and $\mathbf{H1-H2-H3_{P2}}$ proteins in a covalent form. The CD and NMR analyses performed on the native and modified covalent B domains revealed that the amino acid substitutions lead to two new proteins which fold in unique and well-defined structures adopting a three helix bundle as the wild-type protein does.

We believe that our combinatorial strategy is amenable and very useful to redesign other proteins different from the B domain, including other secondary structures different from the α-helix. Main efforts in applying the strategy should be focused in the development of a proper screening test. This should involve the use of a suitable labeled molecule (antibody, a substrate, other protein...) able to interact with the redesigned proteins and recognize those with the desired structure and/or activity. With the right screening test developed, this methodology is rich in potential for the *chemical modification* of proteins overcoming the drawbacks associated with the total synthesis of large protein domains.

Experimental Section

Materials and General Methods. All solvents were used without further purification. Amino PEGA resin, *p*-MBHA resin, and Rink amide MBHA resin were obtained from NovaBiochem. Suitable Boc and Fmoc protected amino acids were from NovaBiochem, Neosystem, and Advanced ChemTech, and coupling agents, from Fluka and NovaBiochem. Human IgG and FITC-labeled rabbit IgG were purchased from Sigma. HMFS linker was synthesized as previously reported.[44]

(67) McLachlan, A. D.; Karn, J. *J. Mol. Biol.* **1983**, *164*, 605−626.
(68) O'Shea, E. K.; Rutkowski, R., III, W. F. S.; Kim, S. P. *Science* **1989**, *245*, 646−648.

Analytical and semipreparative HPLC separations were performed on Symmetry C18 columns (4.6 mm × 150 mm and 30 mm × 100 mm, respectively) at flow rates of 1 mL/min and 10 mL/min. Detection was at 220 nm with a double-wavelength UV detector. Solvent system **A**: 0.045% TFA in water for analytical HPLC and 0.1% TFA in water for semipreparative HPLC; solvent system **B**: 0.036% TFA in acetonitrile for analytical HPLC and 0.1% TFA in acetonitrile for semipreparative HPLC. HPLC-MS analyses were performed on a Symmetry300 C18 column (3.9 mm × 150 mm). Solvent system **A**: 0.1% formic acid in water; solvent system **B**: 0.07% formic acid in acetonitrile. Water was obtained as nanopure using a Millipore Academy A10 system. MALDI-TOF MS spectra and MALDI-TOF PSD analysis were recorded on a Voyager-DE SRT (PerSeptive Biosystems) system, using a mirror ratio range from 1 to 0.1, with 0.1 intervals, and α-cyanohydroxycinnamic acid as matrix. Fluorescence microscopy experiments were performed on Leica *DMBR* and *MZ FLIII* microscopes.

Synthesis of Unprotected Resin-Bound $H3_{wt}$ and $H3_s$. The wild type C-terminal α-helix of the B domain and its scrambled version were synthesized unprotected and reversibly attached to HMFS-PEGA resin, as previously reported.[42] The peptides were characterized after cleavage of small samples of peptide-resins by treatment with 20% morpholine in DMF. Volatiles were removed by vacuum evaporation, and the peptides were analyzed by HPLC, amino acid analysis, and MALDI-TOF PSD (see Supporting Information).

Synthesis of H1−H2. The N-terminal α-helices of the B domain were automatically synthesized on 430 mg of *p*-MBHA resin (0.52 mmol/g). Boc/Bn protected amino acids were incorporated following a standard *in situ* neutralization procedure using TBTU/DIEA as coupling agents. Side-chain deprotection and cleavage were carried out by treatment of the peptide-resin with HF at 0 °C. The peptides were purified by semipreparative HPLC (88 mg, yield = 8.5%). HPLC (R_t) = 7.0 min (linear gradient from 0 to 100% of B in 15 min). MALDI-TOF: $(M + H)^+ = 4642.1$ *m/z*. Amino acid analysis: Asp: 8.2(8), Ser: 0.9(1), Glu: 8.4(8), Gly: 1.1(1), Ala: 2.7(3), Ile: 1.8(2), Leu: 4.1(4), Tyr: 1.0(1), Phe: 2.8(3), His: 0.6(1), Lys: 3.7(4), Pro: 2.6-(2), Arg: 1.0(1).

Screening Tests. Resin-bound peptides were transferred to 2 mL reactors equipped with a porous polyethylene filter at the bottom. The desired amount of beads transferred to each reactor was controlled using homogeneous DMF suspensions of peptide-resin (10 mg/mL, whereby 400 μL = *ca.* 1000 beads). The solvent was first changed from DMF to water by washing with DMF/H_2O mixtures containing increasing amounts of water, and then peptide-resin was blocked by treatment with Buffer 1 (0.1% gelatine in PBS buffer: 10 mM phosphate, 100 mM NaCl, pH = 7.4−7.5) for 1 h and Buffer 2 (0.05% Tween 20 in PBS: 10 mM phosphate, 100 mM NaCl, pH = 7.4−7.5) 10 × 1 min. The solvents were removed by filtration. After resin conditioning, 250 μL of a solution of peptide **H1−H2** in Buffer 1 with desired concentration was added to each sample. Incubation was run for 16 h at room temperature with orbital shaking, at which point 250 μL of the FITC-IgG solution were then added (different dilutions for stock solution were tested) and incubation continued for 6 more hours in the dark. The peptide-resin was finally washed 3 × 1 min with 300 μL of Buffer 1 and transferred to a Petri dish suspended in water. Fluorescent microscopy observation was carried out with the resin in suspension. The fluorescent beads were manually isolated using a 2 μL micropipette.

Synthesis of the L300 Library. The L300 library was synthesized on 300 mg of PEGA resin in the same coupling conditions as those described for the synthesis of resin-bound $\mathbf{H3_{wt}}$. Boc-protected amino acids (5 equiv) were incorporated by activation with DIPCDI/HOBt, except for the C-terminal amino acid which was attached to the HMFS linker using a mixture of DIPCDI/DMAP. Randomized positions in the library were generated using the split and mix approach. The split was accomplished generating DMF suspensions of the resin and then dividing them into equi-volume aliquots as different amino acids to be incorporated at a given position. After coupling of individual amino acids, all beads were pooled into a single reactor and mixed before Boc removal. Coupling times ranged from 1 to 2 h, and reaction completion was checked by the Kaiser test. After final coupling, the peptides were acetylated with acetic anhydride (25 equiv acetic anhydride and 25 equiv of pyridine in DMF) and submitted to side-chain deprotection. HF treatment was carried out in a special Teflon/Kel-F system. The resin was taken up in DCM, introduced into the reactor, and then treated with a mixture 5:95 of *p*-cresol/HF at 0 °C for 90 min, in the absence of magnetical stirring. After HF evaporation, the peptide-resin was washed with DMF and DCM and then stored at −20 °C. Characterization of the library is detailed in the Supporting Information.

MALDI-TOF PSD Sequencing of Peptides from Individual Beads. Fluorescent beads isolated from the library were first washed with water and DMF and then transferred to single positions of a MALDI-TOF sample plate, provided with 1 μL of a solution of 20% morpholine in DMF. Beads were transferred using an end-capped glass capillary, taking advantage of the high affinity of the PEGA resin for glass. Cleavage was run until the morpholine solution had fully evaporated. Matrix solution was then added to each position (1 μL per position), and the samples were analyzed by MALDI-TOF PSD in positive mode, using the mirror ratios already detailed. Fragmentation was manually assigned due to the high complexity of the recorded spectra (see Supporting Information).

Synthesis of Lead Peptides $H3_{P1}$, $H3_{P2}$, $H3_{wt}$, and $H3_s$. The lead peptides were automatically synthesized in parallel on 137 mg of Rink amide resin each (0.7 mmol/g). Fmoc/*t*Bu protected amino acids were incorporated using standard procedures and TBTU/DIEA as coupling agents. After N-terminal acetylation under the same conditions as those described for library L300, concomitant side chain deprotection and cleavage were performed by adding 2 mL of a mixture of TFA/H_2O/triisopropylsilane/phenol (80:15:5:5 v/v/v/w) for 1 h. The TFA was removed by evaporation, and the peptides were purified by semi-preparative HPLC. The products were characterized by HPLC, amino acid analysis, and MALDI-TOF PSD (see Supporting Information).

Circular Dichroism of Noncovalent B Domains. Far-UV CD spectra of peptides were measured between 260 and 190 nm using a JASCO J-810 spectropolarimeter (Jasco, Tokyo, Japan). Peptide concentration was adjusted to 40 μM in 10 mM sodium phosphate buffer (100 mM NaCl, pH = 7.4). Spectra were recorded at 25 °C using 1 mm path-length cuvette. In a typical experiment, three spectra were collected, averaged, and corrected by subtraction of a buffer blank spectrum. The CD data were converted to molar ellipticity [θ], expressed in deg × cm^2 × $dmol^{-1}$, defined as

$$[\theta] = \theta/(10cl)$$

whereby θ is the ellipticity in millidegrees, c is the molar concentration of the peptide, and l is the path length in cm.

Synthesis of Covalent Natural and Modified B Domains. Covalent domains were automatically synthesized in parallel on 500 mg of Rink amide MBHA resin (0.18 mmol/g). Fmoc/*t*Bu protected amino acids were incorporated using standard procedures, and TBTU/DIEA, as coupling agents. After final deprotection, peptides were acetylated by anhydride acetic treatment (40 equiv, 9 equiv of DIEA, DMF). Concomitant side-chain deprotection and cleavage were performed by treatment with 5 mL of a mixture of TFA/H_2O/triisopropylsilane/ethanedithiol (92:4:2:2 v/v/v/v) for 1.5 h. TFA was removed by evaporation, and the crude was precipitated with *tert*-butyl methyl ether. The products were characterized by HPLC, amino acid analysis, and MALDI-TOF MS. $\mathbf{H1-H2-H3_{wt}}$ purification was carried out by semipreparative HPLC (100 mg, yield = 13%). HPLC (R_t) = 10.9 min (linear gradient from 0 to 50% of B in 15 min). MALDI-TOF: $(M + H)^+ = 6935.7$ *m/z*. Amino acid analysis: Asp: 12.1(12), Ser: 2.57-(3), Glu: 11.4(11), Gly: 0.9(1), Ala: 8.1(8), Ile: 1.7(2), Leu: 7.0(7), Tyr: 0.7(1), Phe: 2.6(3), His: 0.9(1), Lys: 6.9(7), Arg: 0.8(1). **H1−**

H2−H3$_{P1}$, peptide purification was carried out by semipreparative HPLC (100 mg, yield = 14%). HPLC (R_t) = 10.9 min (linear gradient from 0 to 50% of B in 15 min). MALDI-TOF: $(M + H)^+ = 7011.8$ *m*/*z*. Amino acid analysis: Asp: 12.3(12), Ser: 1.6(2), Glu: 10.7(11), Gly: 1.1(1), Ala: 7.5(8), Ile: 1.9(2), Leu: 6.1(6), Tyr: 0.7(1), Phe: 3.9(4), His: 0.9(1), Lys: 8.7(8), Arg: 0.9(1). **H1−H2−H3$_{P2}$**, peptide purification was carried out by semipreparative HPLC (100 mg, yield = 14%). HPLC (R_t) = 10.6 min (linear gradient from 0 to 50% of B in 15 min). MALDI-TOF: $(M + H)^+ = 6981.7$*m*/*z*. Amino acid analysis: Asp: 12.0(12), Ser: 1.5(2), Glu: 11.3(11), Pro: 4.4(4), Gly: 0.7(1), Ala: 7.8(8), Val: 1.1(1), Ile: 1.7(2), Leu: 6.1(6), Tyr: 0.6(1), Phe: 4.8(4), His: 0.9(1), Lys: 7.1(7), Arg: 0.7(1)

Circular Dichroism of Covalent B Domains. Far-UV CD spectra and thermal denaturation curves of covalent B domains were recorded in the same conditions as those described for noncovalent B domains. The melting temperature was determined by assuming a two-state mechanism and using a least-squares fitting routine derived from the van't Hoff equation.[69]

NMR of Covalent B Domains. TOCSY and NOESY experiments were recorded on a 800 MHz spectrometer (Bruker Avance 800) at 303 K. Proteins were dissolved in 10 mM phosphate buffer (100 mM NaCl, H_2O/D_2O 9:1, pH = 5.8) to a final concentration of 1.0 mM, with 0.1 mM DSS and 0.01% sodium azide as additives. ^{1}H chemical shifts were referenced to sodium 2,2-dimethyl-2-silapentane-5-sulfonate sodium salt (DSS) at 0.00 ppm. TOCSY and NOESY experiments were recorded with a matrix size of 4096 × 512. The mixing time was adjusted to 65 ms for TOCSY and 100 ms or 200 ms for NOESY.

(69) Greenfield, N. J. *Methods Enzymol.* **2004**, *383*, 282−317.

Acknowledgment. This work was supported by MCYT-FEDER (Bio2005-00295 and GEN2003-20642-C09-04) and Generalitat de Catalunya (CERBA and 2005SGR-00663).

Supporting Information Available: Full characterization data for unprotected resin-bound **H3$_{wt}$** and **H3$_s$**; characterization data from the synthetic efficiency analysis of library L300; MALDI-TOF PSD spectra and fragment assignment for isolated positive sequences from the screening of library L300; full characterization data for the synthesized lead peptides **H3$_{P1}$, H3$_{P2}$, H3$_{wt}$,** and **H3$_s$**; CD melting curves for noncovalent complexes **H1−H2 + H3$_x$**; table of molar ellipticities at 222 nm for noncovalent complexes (**H1−H2 + H3$_x$**) and covalent bundles (**H1−H2−H3$_x$**). This material is available free of charge via the Internet at http://pubs.acs.org.

JA073969X