A Modular Assembly Strategy for Improving the Substrate Specificity of Small Catalytic Peptides

2 AUTHORS, INCLUDING:

Carlos F Barbas
The Scripps Research Institute
363 PUBLICATIONS **30,429** CITATIONS

A Modular Assembly Strategy for Improving the Substrate Specificity of Small Catalytic Peptides

Fujie Tanaka* and Carlos F. Barbas, III*

The Skaggs Institute for Chemical Biology and the Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Substrate specificity is one of the hallmarks of enzymes. Substrate specificity allows an enzyme to catalyze a reaction involving substrate molecules found within a complex mixture of compounds possessing the same functional groups. Generation of designer protein catalysts that possess substrate specificity has been demonstrated by modification of Nature's enzymes[1] and by the preparation of catalytic antibodies.[2] In contrast to large proteins, small peptide catalysts have demonstrated limited specificity for small-molecule substrates.[3] This is presumably a result of the limited opportunities small peptides have to fold in a manner that provides for the formation of an isolated reaction vessel that effectively binds and sequesters substrates from bulk solvent while at the same time catalyzing their transformation. To explore routes for the preparation of small peptide enzymes that possess improved substrate specificity, we have examined a modular assembly strategy. Here we demonstrate the potential of this strategy with the construction of a small 35-amino acid residue aldolase peptide with improved substrate specificity.

Our design strategy attempts to recruit a substrate-specific module for providing substrate specificity to an otherwise promiscuous catalyst. Covalent combination of binding- and catalytic-domain modules might improve the substrate specificity of the catalyst. When the binding site is in close proximity to the catalytic site, the catalytic site would receive the benefit of a higher local substrate concentration provided by sequestering the substrate in close proximity to the catalytic site. The potential advantages of this approach are that it reduces the demand on the functionalization of the catalytic site, which is limited in small peptides, and it is modular, therefore making its adaptation to a variety of specificities rapid.

Recently, we reported a 24-amino acid residue aldolase peptide FT-YLK3 that catalyzes retro-aldol reactions via an enamine mechanism.[4] This peptide was the product of covalent mechanism-based β-diketone selection of a phage-displayed peptide library where six randomized residues were appended to an 18-residue amphipathic α-helix[5] (YKLLKELLAKLKWLLRKL) containing five lysine residues. The ϵ-amino groups of the lysyl residues are essential for the catalytic reaction. Peptide FT-YLK3 is highly α-helical and catalyzes the retro-aldol reaction of **1** with a k_{cat}/k_{uncat} of 1900, albeit with a K_m of 1.8 mM indicative of relatively poor substrate recognition.

To endow this catalyst with improved substrate specificity, we sought to covalently connect this peptide to another peptide that binds a substrate molecule with high affinity, anticipating that the local concentration of substrate proximal to the catalytic peptide would be increased. Recently, Rozinov and Nolan reported small peptides that bind to small-molecule fluorophores.[6] One of these

peptides, YPNEFDWWDYYY (FluS303 in their report) binds to fluorescein. This peptide was selected for binding to fluorescein derivative **2** from a combinatorial library of 12-mer peptides displayed on phage.[7] We chose this peptide as our substrate-binding module and attached it to the N-terminus of FT-YLK3. Since FT-YLK3 has a tyrosine residue at the N-terminus and FluS303 has three tyrosine residues at the C-terminus, one tyrosine was removed at the junction. Peptide FluS303-FTYLK3 was chemically synthesized (C-terminal amide) and characterized.
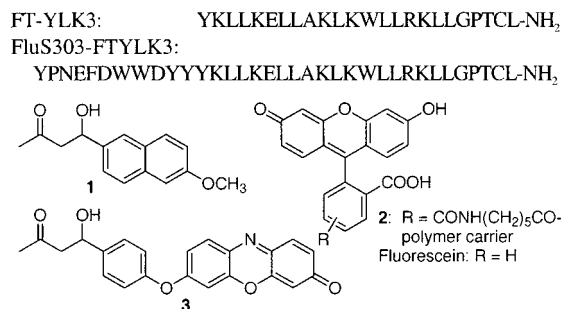
Retro-aldol substrate **3** contains a structural moiety from fluorescein and was designed and synthesized as a specific substrate for FluS303-FTYLK3.[8] Peptide-catalyzed retro-aldol reactions of **3** were examined using FluS303-FTYLK3 and FT-YLK3. Reactions were performed using 50 μM peptide in 10% CH$_3$CN−42.5 mM Na phosphate (pH 7.5), and the product aldehyde **4** was monitored by HPLC. Both peptides FluS303-FTYLK3 and FT-YLK3 displayed saturation kinetics. Peptide-catalyzed retro-aldol reactions of **1** were also studied.[9] Peptide FluS303-FTYLK3 catalyzed the retro-aldol reaction of **3** with a K_m of 8 μM and a k_{cat} of 2.3 × 10^{-4} min^{-1} (Table 1).[10] FluS303-FTYLK3 catalyzed the retro-aldol reaction of **1** with a 1.1 mM K_m.[11] The K_m value of substrate **3** was 140-fold lower than that of **1** in the FluS303-FTYLK3-catalyzed reaction and the k_{cat}/K_m value of **3** was 43-fold higher than that of **1**. Thus FluS303-FTYLK3 successfully discriminates between substrates **3** and **1**. As calculated from the K_m values, the differential binding energy ($\Delta\Delta G$) in favor of substrate **3** over **1** in the FluS303-FTYLK3-catalyzed reactions is 2.9 kcal/mol.[12] The catalytic module itself, FT-YLK3, catalyzed the reaction of **3** with a K_m of 130 μM. Fortuitously, the K_m value of substrate **3** was 14-fold lower than that of substrate **1** in the FT-YLK3-catalyzed reactions. Thus, for FT-YLK3-catalyzed reactions, differential binding of substrate **3** over **1** is favored by 1.6 kcal/mol.[12] Note that the K_m value of FluS303-FTYLK3 with substrate **3** was 16-fold lower than that of FT-YLK3 with **3**, while the K_m values of **1** were similar in both the FluS303-FTYLK3- and FT-YLK3-catalyzed reactions. These results indicate that the enhanced substrate specificity of FluS303-FTYLK3 for **3** is the result of two effects. First, binding of **3** to

* Corresponding authors. E-mail: carlos@scripps.edu and ftanaka@scripps.edu.

Table 1. Kinetic Parameters of Peptide-Catalyzed Retro-Aldol Reactions

peptide	substrate 3[a]			substrate 1[b]		
	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m (M^{-1} min^{-1})	K_m (mM)	k_{cat} (min^{-1})	k_{cat}/K_m (M^{-1}min^{-1})
FluS303-FTYLK3	8	2.3×10^{-4}	29	1.1	$7.4 \times 10^{-4\ c}$	0.67
FT-YLK3	130	2.0×10^{-4}	1.5	1.8[d]	$5.6 \times 10^{-4\ d}$	0.31[d]

[a] Reaction conditions: [Peptide] 50 μM in 10% CH_3CN−42.5 mM Na phosphate (pH 7.5) at 25 °C. See ref 11. [b] [Peptide] 50 μM in 5% CH_3CN−42.5 mM Na phosphate (pH 7.5) at 25 °C. The first-order kinetic constant of the background reaction (k_{uncat}) was 3.9×10^{-7} min^{-1}. [c] k_{cat}/k_{uncat} 1900. [d] Data were taken from ref 4. [Peptide] 100 μM.

the FT-YLK3 domain as compared to substrate **1** is enhanced by 1.6 kcal/mol. While the exact mechanism for this increase is not clear at present and was not part of the design strategy, heteroatom presentation within **3** may play a role (see Supporting Information for discussion and model). Second, the addition of the FluS303 sequence to FTYLK3 further increases the affinity of the peptide for substrate **3** ($\Delta\Delta G = 1.7$ kcal/mol)[13] while having a minimal effect on the binding of substrate **1** ($\Delta\Delta G = 0.3$ kcal/mol). We assign this increase in substrate specificity to the effect of the modular assembly strategy.[14]

To further probe the role of the substrate-binding domain FluS303 in the peptide enzyme FluS303-FTYLK3, binding to fluorescein and the effect of fluorescein on the catalyzed reaction was examined. The dissociation constant (K_d) of fluorescein for FluS303-FTYLK3 was determined to be 3.6 μM by fluorescence-quenching, indicating that the binding domain retains its original function in the fusion peptide.[15] Further, when the FluS303-FTYLK3-catalyzed reaction of **3** was performed in the presence of fluorescein (500 μM), the kinetic parameters were K_m 60 μM, k_{cat} 2.5 $\times$ 10^{-4} min^{-1}, and k_{cat}/K_m 4.2 M^{-1} min^{-1}. Thus, the K_m value substantially increased while the k_{cat} remained unchanged (within experimental deviation). This result indicates that when fluorescein occupies the substrate-binding domain, in competition with **3,** the specificity constant is greatly reduced. With respect to the catalytic domain, it is likely that the highly α-helical structure of FT-YLK3 is maintained in the fusion peptide since in all of our studies the k_{cat} of the catalytic domain FT-YLK3 is not altered as compared to that of the fusion peptide. Our earlier study of FT-YLK3 found a strong correlation between α-helical content and k_{cat}.[4]

Circular dichroism studies indicate that FluS303-FTYLK3 retains the α-helical structure of the parental peptide FTYLK3.[16] FluS303-FTYLK3 also formed a UV-observable enaminone upon addition of 2,4-pentanedione.[4] Denaturation of the α-helical structure (4.5 M guanidine hydrochloride) as observed in the CD spectra was accompanied by complete loss of catalytic activity.

In summary, we have shown that a modular assembly strategy can be used to create small peptide enzymes with good specificity for small molecule substrates. Since our approach accepts substrate-binding domains derived from phage display selections, a wide variety of substrate binding specificities might be readily selected and recruited to rapidly create peptide enzymes with a wide variety of substrate specificities. This is significant since peptide domains that bind small molecules are not currently accessible by rational design. We anticipate that this approach will be applicable to other small peptide catalytic domains as well, allowing for the combinatorial preparation of substrate specific peptide enzymes. The modular assembly strategy studied here may therefore present advantages over exhaustive searches of large random-sequence peptide libraries[17] for peptides with singular function.

Supporting Information Available: (1) Data for compounds **3** and **4**, (2) assays, (3) additional discussion concerning the binding of **3**, (4) circular dichroism studies, (5) multiturnover experiment, (6) catalysis of the aldol reaction, and (7) effect of peptide concentration on enaminone formation (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Petrounia, I. P.; Arnold, F. H. *Curr. Opin. Biotechnol.* **2000**, *11*, 325. Cedrone, F.; Menez, A.; Quemeneur, E. *Curr. Opin. Struct. Biol.* **2000**, *10*, 405.
(2) Reymond, J.-L. *Top. Curr. Chem.* **1999**, *200*, 59. Barbas, C. F., III; Rader, C.; Segal, D. J.; List, B.; Turner, J. M. *Adv. Protein Chem.* **2000**, *55*, 317. Schultz, P. G.; Lerner, R. A. *Science* **1995**, *269*, 1835. Tanaka, F.; Lerner, R. A.; Barbas, C. F., III. *J. Am. Chem. Soc.* **2000**, *122*, 4835.
(3) Broo, K. S.; Nilsson, H.; Nilsson, J.; Baltzer, L. *J. Am. Chem. Soc.* **1998**, *120*, 10287. Broo, K. S.; Brive, L.; Ahlberg, P.; Baltzer, L. *J. Am. Chem. Soc.* **1997**, *119*, 11362. Johnsson, K.; Allemann, R. K.; Widmer, H.; Benner, S. A. *Nature* **1993**, *365*, 530. Peptide ligases operate via template-assisted catalysis and demonstrate good substrate specificity: Kennan, A. K.; Haridas, V.; Severin, K.; Lee, D. H.; Ghadiri, M. R. *J. Am. Chem. Soc.* **2001**, *123*, 1797. Yao, S.; Ghosh, I.; Zutshi, R.; Chmielewski, J. *Nature* **1998**, *396*, 447.
(4) Tanaka, F.; Barbas, C. F., III. *Chem. Commun.* **2001**, 769.
(5) Perez-Paya, E.; Houghton, R. A.; Blondelle, S. E. *J. Biol. Chem.* **1996**, *271*, 4120.
(6) Rozinov, M. N.; Nolan, G. P. *Chem. Biol.* **1998**, *5*, 713.
(7) The affinity of FluS303 displayed pentavalently on phage for **2** was reported to be K_d 2.7 nM in ref 6. Lacking in avidity, the affinity of monovalent peptide FluS303 for **2** is lower.
(8) Aldehyde **4** was synthesized from resorufin, sodium salt, and 4-fluorobenzaldehyde according to the procedure reported by Yeager et al. Yeager, G. Y.; Schissel, D. N. *Synthesis* **1991**, 63. Subsequently, the DL-proline-catalyzed aldol reaction with acetone according to the procedure reported by Sakthivel et al. gave aldol **3**. Sakthivel, K.; Notz, W.; Bui, T.; Barbas, C. F., III. *J. Am. Chem. Soc.* **2001**, *123*, 5260.

(9) The reaction of **1** was followed by monitoring the increase in fluorescence (λ_{ex} 330 nm, λ_{em} 452 nm). See ref 4 and List, B.; Barbas, C. F., III; Lerner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 15351.
(10) Since near-equal peptide-substrate concentrations were employed in the assay, the k_{cat} and K_m were determined according to the procedure reported by Smith et al. Smith, G. D.; Eisenthal, R.; Harrison, R. *Anal. Biochem.* **1977**, *79*, 643.
(11) The velocity of the catalyzed reaction was identical in 10 and 5% CH_3-CN. To increase the solubility of **3**, 10% CH_3CN was used. The first-order kinetic constant of the background reaction (k_{uncat}) of **3** was significantly lower than that of **1**.
(12) $\Delta\Delta G = -RT \ln\{(K_m$ of **3**)/(K_m of **1**)$\}$ at 25 °C.
(13) $\Delta\Delta G = -RT \ln\{(K_m$ of FluS303-FTYLK3-catalyzed reaction of **3**)/(K_m of FT-YLK3-catalyzed reaction of **3**)$\}$ at 25 °C.
(14) At this point, two potential mechanisms for providing enhanced substrate discrimination are: (1) the binding domain provides a higher local substrate concentration to the catalytic domain with the two domains acting independently, and (2) the binding and catalytic domains work cooperatively to form a more specific active site.
(15) Quenching study was performed using fluorescein (5 μM) and FluS303-FTYLK3 (1.5−50 μM) in 10% CH_3CN−42.5 mM Na phosphate (pH 7.5) by monitoring the fluorescence of fluorescein (λ_{ex} 472 nm, λ_{em} 518 nm). See ref 6. The K_d was calculated according to the procedure described in ref 10.
(16) The CD spectra of FluS303-FTYLK3 (50 μM) in 45 mM Na phosphate (pH 7.5) at 25 °C indicates that the peptide contains α-helical and β-sheet structures. Studies using the program k2d indicate the α-helical content is 27% (max error 22.7%). Since the α-helical content of FT-YLK3 (100 μM) was estimated to be $\sim$70%,[4] the ideal content of FluS303-FTYLK3 is calculated to be 48% (24 residue x 0.7/35 residue). For the k2d program see, Andrate, M. A.; Chacon, P.; Merelo, J. J.; Moran, F. *Protein Eng.* **1993**, *6*, 383.
(17) Keefe, A. D.; Szostak, J. W. *Nature* **2001**, *410*, 715.

JA0171815