Communications

Engineering Peptide Inhibitors To Overcome PDZ Binding Promiscuity**

Lars Vouilleme, Patrick R. Cushing, Rudolf Volkmer, Dean R. Madden, and Prisca Boisguerin**

Protein–protein interaction domains (PPIDs) are key elements in assembling functional protein complexes and controlling cellular activities. A major class of PPIDs is mediated by PDZ (for PSD-95, Dlg, ZO-1) domains,[1–3] widespread scaffolding modules essential for the localization and activity of numerous cellular effector proteins. Among the diverse protein interaction domains, PDZ domains are highly conserved in organisms from bacteria to humans.[4] They usually bind the C-terminus of their ligands.

Consistent with their structural homology, PDZ domains exhibit overlapping recognition sequences, meaning that a given partner typically can interact with multiple domains. Some years ago, we proposed a general and efficient procedure for profiling PDZ–peptide interactions that provides a picture of specificity and selectivity covering the complete PDZ–ligand sequence space by combining SPOT synthesis and K_d prediction.[5] As expected, among the three PDZ domains that were analyzed (AF6, ERBIN, and SNA1), the overlap of ligand sequences recognized at K_d values between 50–100 μM was substantial. Recent studies have suggested that there is more diversity among PDZ sequence preferences than originally thought.[6,7] Nevertheless, the set of PDZ domains interacting with a given protein necessarily shares overlapping binding motifs, and it remains challenging to develop a canonical peptide that will inhibit only a single PDZ domain out of this set.

[*] L. Vouilleme,[+] Dr. R. Volkmer, Dr. P. Boisguerin
Institute of Medical Immunology
Charité – Universitätsmedizin Berlin
Hessische Strasse 3–4, 10115 Berlin (Germany)
Fax: (+ 49) 30-450-524942
E-mail: prisca.boisguerin@charite.de

P. R. Cushing,[+] Dr. D. R. Madden
Department of Biochemistry, Dartmouth Medical School
7200 Vail Building, Hanover, NH 03755 (USA)
Fax: (+ 1) 603-650-1128
E-mail: drm0001@dartmouth.edu

[+] These authors contributed equally to the work.

Supporting information for this article, including Experimental Methods, is available on the WWW under http://dx.doi.org/10.1002/anie.201005575.

To address this issue, we focused on a set of five PDZ domains known to interact with the cystic fibrosis (CF) transmembrane conductance regulator (CFTR). The PDZ-containing proteins CAL (CFTR-associated ligand)[8,9] and its antagonists NHERF1 and NHERF2 (Na+/H+ exchanger regulatory factors 1 and 2),[8,10] compete for binding to CFTR. CAL contains one (CALP) and each NHERF protein two PDZ domains (N1P1, N1P2, N2P1 and N2P2) which control both the activity and the cell surface abundance of CFTR. NHERF family members increase CFTR activity at the apical membrane, whereas CAL promotes its lysosomal degradation. Thus, to explore novel therapeutic strategies for increasing the cell-surface abundance of CFTR, our goal was to design a selective inhibitor of the CFTR–CAL interaction that does not affect the biologically relevant PDZ competitors NHERF1 and NHERF2.[8]

Here, we present a strategy for the parallel evolution of inhibitor affinity and selectivity, optimizing binding determinants distributed along the length of a decameric sequence. Because the individual contributions can be modest, our challenge is to survey them efficiently and with high accuracy. Peptide libraries provide the necessary throughput, while fluorescence polarization (FP) measurements provide precise estimates of affinity for all five domains.

The general approach (Figure 1) involves the synthesis of a variety of different cellulose-bound peptide libraries with the method of inverted peptides[11] based on SPOT technology[12]—a simple but robust technique for the parallel synthesis of up to 6000 peptides with free C-termini.[13] As promising extensions or sequence modifications are identified, FP[14] assays are used to determine binding constants for all five PDZ domains of the CFTR trafficking system. Each modification is then evaluated for its contribution to the affinity for CAL and to the loss of affinity for the individual NHERF domains. (For details see Experimental Section in Supporting Information).

To initiate the project, arrays encoding a human C-terminal peptide library (6223HumLib)[11] were synthesized and probed with each of the five PDZ domains (Figure 1). The binding sequences that were detected included published interactions such as the platelet-derived growth factor,[15] the β2-adrenergic receptor,[16] and CFTR[17] for N1P1, as well as the N2P2 partner β-catenin[18] (see Supporting Information, Table S2).

The 80 best binding sequences of the HumLib incubations were analyzed using the WebLogo algorithm,[19] revealing clear C-terminal binding motifs (Figure 2A and Tables S1–S5). Consistent with their shared affinity for the CFTR C-terminus, the resulting consensus motifs are very similar, especially for ligand positions 0 (P^0: ligand positions are

© 2010 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

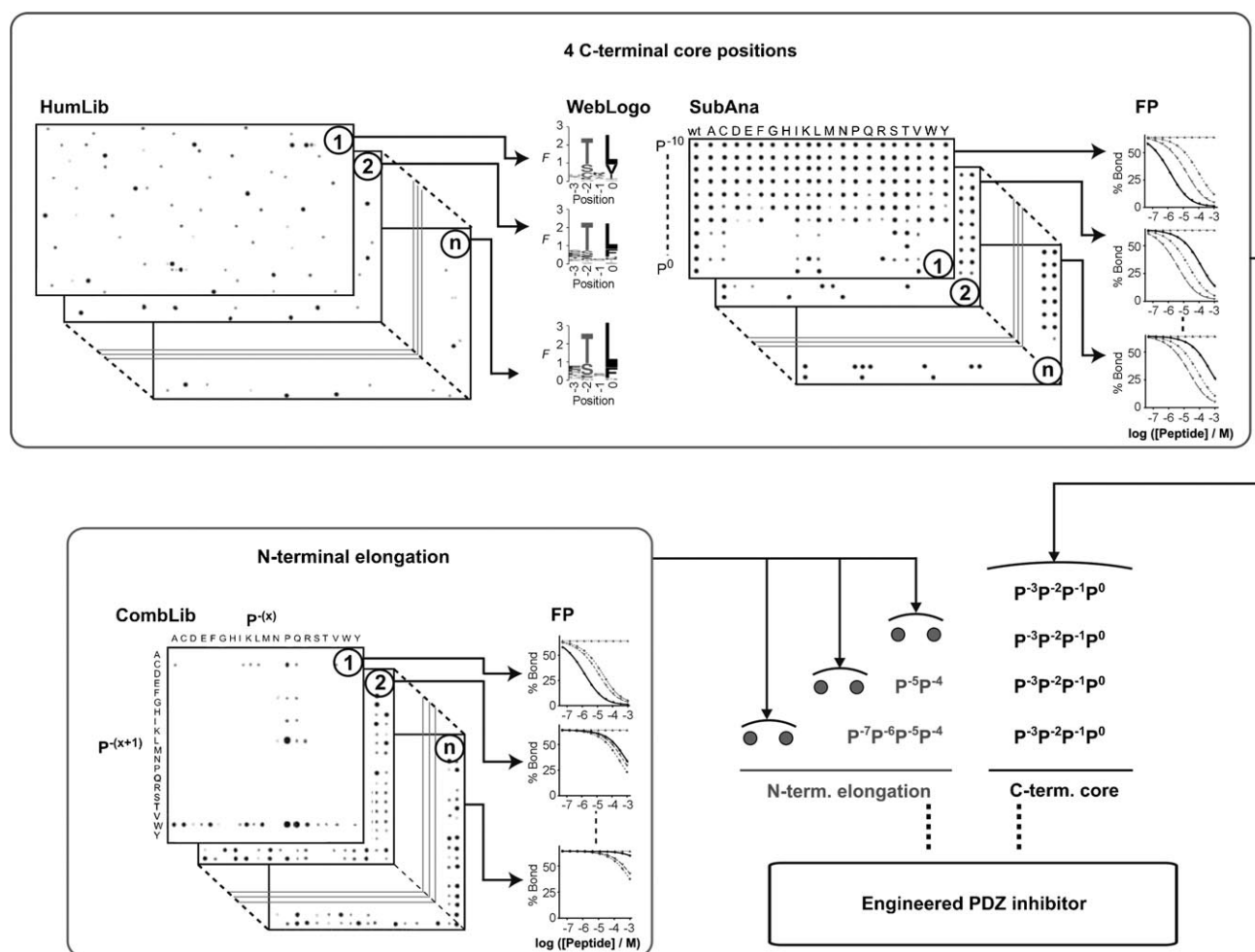*Angew. Chem. Int. Ed.* 2010, 49, 9912–9916

Figure 1. Engineering selective PDZ inhibitors. To find a C-terminal "core", HumLib arrays were incubated with individual PDZ domains and immunoblotted with antibody to detect bound PDZ protein as a function of peptide sequence. Top binding sequences were aligned (WebLogo) (F = frequency), and arrays of individual side-chain substitutions were synthesized (SubAna) to establish C-terminal sequence motifs with affinity for CALP, but not for the NHERF PDZ domains. Candidate peptides were synthesized and tested by FP inhibition assay for binding to all five PDZ domains. Once a selective C-terminal sequence was established for positions P^{-3}–P^0, CombLibs were synthesized. Pairs of amino acids were selected based on strong binding to the CALP and weak or reduced binding to the NHERF domains. FP inhibition was used for validation. Once a residue pair was added to the core sequence, the pairwise N-terminal extension process was iterated.

numbered in reverse from the C-terminal ligand residue, which is denoted as 0) and P^{-2} requiring L and S/T, respectively. Importantly, we also found distinctions in the alternative side-chain preferences at P^0, with CAL sequences containing I or V, and NHERF sequences containing F. A similar N1P1 binding motif corresponding to x-S/T-R-F was reported by Joo and Pei.[20]

To investigate the context dependence of the amino acid preferences, we performed substitutional analyses (SubAna) on a series of peptides from the 80 best HumLib peptides, including the C-terminus of the somatostatin receptor type 5 ($SSR5_{10}$) which has the highest affinity for CALP among a series of known binding sequences[14] (Figure 1, 2B, and S1–S2). The SubAnas confirm the difference in the amino acid preferences at P^0 between the five PDZ domains, and show a clear tolerance of CALP for I at P^0. In contrast, NHERF PDZ domains show weak or absent binding for I at these positions (Figure S1, S2). As determined by FP measurements, the single P^0 L/I substitution generates a ca. 7-fold increase in the selectivity index for CALP compared to the NHERF PDZ domains (0.4 for $SSR5_6$ vs. 2.7 for $iCAL05_6$, Table 1).

SubAna data also show a modest preference of CALP for I at P^{-1}, (Figures S1, S2). With the double substitution at P^0 and P^{-1} we increase the selectivity index ca. 25-fold ($SSR5_6$ vs. $iCAL06_6$, Table 1). Although the motif and SubAna data on $P^{-2/-3}$ were ambiguous, previous FP measurements had shown that the T/S sequence provided ca. 5-fold more selectivity versus N1P2 than the S/T sequence (data not shown).

Beyond the P^{-3} position, classical motif analysis reveals no meaningful preferences for any of the five PDZ domains (Figure 2B and S1, S2). However, other studies have shown that upstream residues can contribute to PDZ binding affinity.[14] Our hypothesis was that different C-terminal anchor sequences might have distinct upstream preferences, washing out signals in a global motif analysis. To detect context-dependent preferences, we generated combinatorial
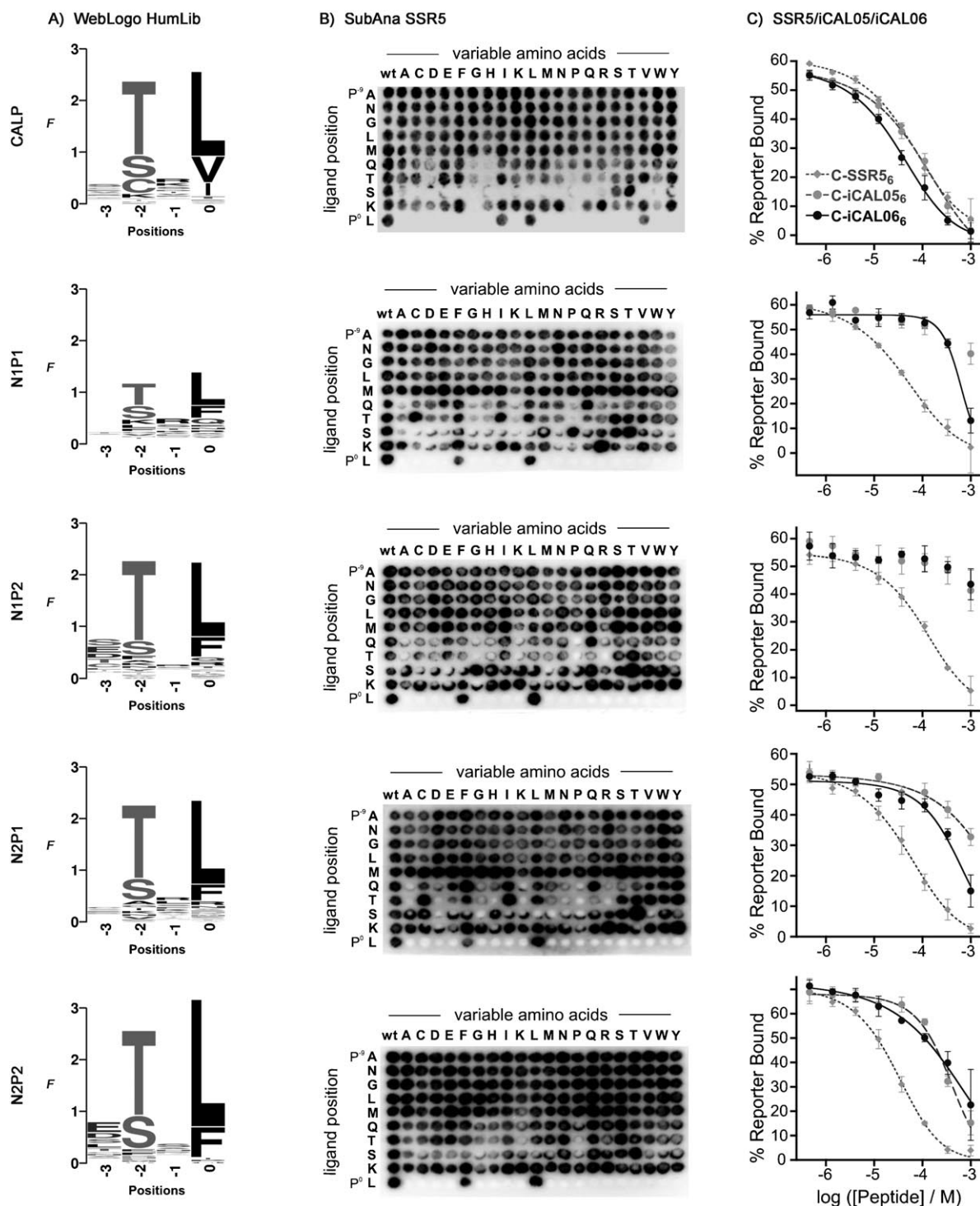
Figure 2. Consensus motifs of the five PDZ domains involved in CFTR binding. A) Five 6223-HumLibs were synthesized, each containing 6223 human C-terminal peptides, which were incubated with the CALP and the four NHERF PDZ domains, respectively. The most frequent amino acids were plotted for the four C-terminal residues using WebLogo analysis of the 80 best binding sequences. Sequences are listed in the Supporting Information. B) Substitutional analyses (SubAna) of the C-terminal SSR5 sequence demonstrate the interaction determinants of the CALP and the NHERF PDZ domains. C) FP measurements comparing SSR5$_6$ with iCAL06$_6$ and with iCAL05$_6$ clearly revealed that I at P^0 does not disturb the interaction with CALP, but increases the K_i values of the NHERF PDZ domains. Values shown are mean $\pm$ SD, $N=3$.

libraries (CombLib). Due to chemical restriction, the core sequence at least has to be a 4-mer to allow cyclization for the inversion during peptide synthesis. The double permutation provides information beyond that available through SubAnas,

since it can potentially identify cross-talk between adjacent positions. As the starting point for our investigations, we selected P$^{-(x+1)}$–P$^{-(x)}$–TSII (P$^{-(x+1)}$, P$^{-(x)}$: simultaneous permutation with the 20 L-amino acids, other positions were held

Table 1: Evaluation of K_i values (in µM) for peptide engineering.

Peptide	Sequence	CALP	N1P1	N1P2	N2P1	N2P2	Selectivity[a]
CFTR$_{10}$	TEEEVQDTRL	390 ± 20	0.45 ± 0.02	1.9 ± 0.1	1.1 ± 0.1	0.10 ± 0.02	0.00026
SSR5$_{10}$	ANGLMQTSKL	21.4 ± 1.7	24.5 ± 1.9	130 ± 48	38.6 ± 7.7	19.8 ± 2.7	0.93
SSR5$_6$[b]	MQTSKL	25.3 ± 3.4	19.3 ± 5.6	50.2 ± 9.1	31.7 ± 8.9	10.0 ± 0.6	0.40
iCAL05$_6$[b]	MQTSKI	40.8 ± 10.3	914 ± 186	1700 ± 1300	806 ± 84	109 ± 10	2.7
iCAL06$_6$[b]	MQTSII	16.9 ± 1.6	245 ± 71	2400 ± 1400	245 ± 26	166 ± 104	9.8
iCAL35$_6$[b]	WQTSII	16.3 ± 2.1	430 ± 109	> 5000	496 ± 151	246 ± 186	15
iCAL36$_6$[b]	WPTSII	32.8 ± 0.3	> 5000	> 5000	> 5000	> 5000	150
iCAL36$_8$	SRWPTSII	20.3 ± 2.6	> 5000	> 5000	516 ± 183	433 ± 154	21
iCAL36$_{10}$	ANSRWPTSII	17.3 ± 4.3	> 5000	> 5000	> 5000	> 3000	170

[a] Selectivity index = min(K_iN1P1, K_iN1P2, K_iN2P1, K_iN2P2)/K_iCAL. [b] Peptides include an N-terminal cysteine to permit labeling. $N = 3$ for all experiments.

constant, Figure 1) to elucidate $P^{-4/-5}$ preferences. With this peptide library, we were able to determine that Q/P at P^{-4} as well as W at P^{-5} gave the highest signal intensities for CALP (Figure 3 A). To establish the affinity contribution of the single substitutions rigorously, we analyzed the sequences WQTSII (iCAL35$_6$) and WPTSII (iCAL36$_6$) compared to MQTSII (iCAL06$_6$) by FP measurements (Figure 3 B–F; Table 1).
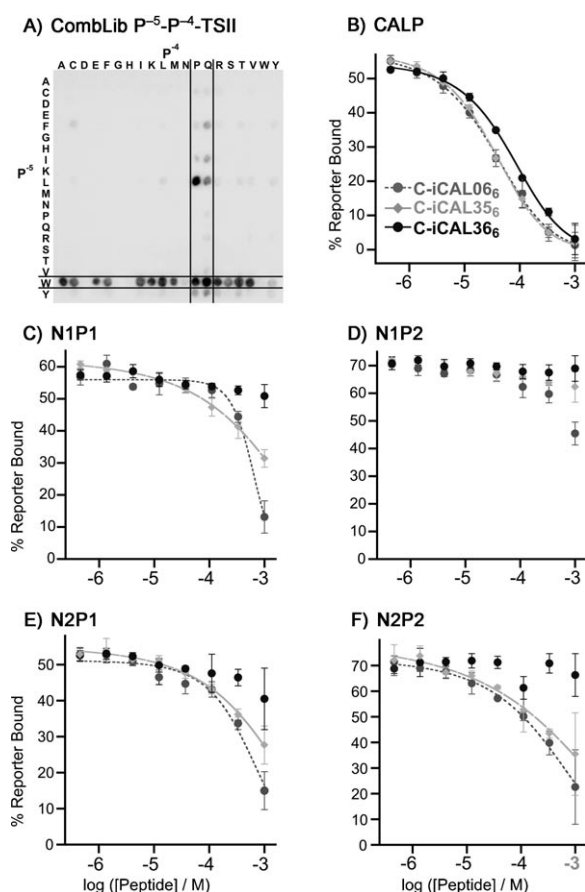

Figure 3. Enhancing CAL PDZ selectivity by amino acid substitution. A) The recognition pattern of the CombLib incubation shows a clear preference for the combination of Q/P and W at P^{-4} and P^{-5}. B–F) FP binding isotherms for the three peptides iCAL06$_6$, iCAL35$_6$ and iCAL36$_6$ were measured with the five different PDZ domains.

The substitution of W for M at P^{-5} (iCAL35$_6$) has only a minimal effect on CALP affinity, but further weakens peptide interactions with all of the NHERF domains, increasing the selectivity index from 9.8 to 15. The additional substitution of P for Q at P^{-4} (iCAL36$_6$) abolished interactions with the NHERF PDZ domains (all > 5000 µM), consistent with CombLib data showing that NHERF domains bind poorly to peptides containing proline at P^{-4} in multiple sequence contexts (Figure S4). The ability to visualize negative contributions to peptide affinity represents another important advantage of the CombLib approach over WebLogo-based motif analysis. Even though affinity for CALP is also reduced by nearly two-fold, the resulting peptide has a selectivity index of 150 (Table 1).

Based on our knowledge about the importance of the C-terminal peptide length on CALP affinity,[14] we decided to elongate the iCAL36$_6$ sequence N-terminally, in an effort to further optimize both elements. CombLibs of the type P^{-7}–P^{-6}–WQTSII and P^{-7}–P^{-6}–WPTSII were incubated with CALP determine the amino acid preferences at P^{-6} and P^{-7}, but reflected little specificity at either position (Figure S3). This was confirmed by SubAna incubations, which revealed no specificity upstream of P^{-6} (Figure S1). Nevertheless, P^{-7}–P^{-6}–WPSTRV CombLib incubated with CALP demonstrates a slight preference for R at P^{-6} (Figure S3). Addition of an SR pair at P^{-7}/P^{-6} enhanced CALP affinity, but also reduced selectivity vs. NHERF2 PDZ domains. Selectivity was restored without loss of CAL affinity by a further addition of an AN pair at P^{-9}/P^{-8}. As the CombLibs also did not reveal any clear side-chain preference at these positions (Figure S3), these affinity effects are presumably mediated primarily by peptide main-chain interactions. These extensions of the peptide sequences resulted in the peptide sequence iCAL36$_{10}$ (ANSRWPTSII). Retrospective sequence analysis using a SubAna library shows a similar binding pattern for CALP with P^0, P^{-2}, and P^{-5} as key residues (Figure 4).

Our final modification involved the N-terminal attachment of fluorescein to a decamer sequence, which previous studies had shown to enhance CALP binding more than NHERF1 binding.[14] As expected, the resulting F*-iCAL36 sequence (F*-ANRSWPTSII) exhibits substantially improved affinity for the CAL PDZ domain ($K_d = 1.3 \pm 0.1$ µM). This is comparable to the affinity of other reported PDZ inhibitors.[21,22]
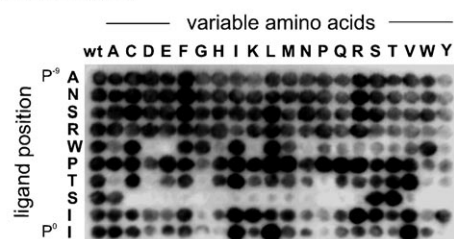
Communications

CALP:iCAL36

Figure 4. Substitutional analyses of iCAL36 incubated with the CAL PDZ domain.

Titrations with NHERF PDZ domains did not reveal significant binding at concentrations as high as 100 μm, indicative of K_d values of more than 1 mm. As a result, the selectivity index for $F*$-iCAL36 exceeds 750. Compared to the non-selective SSR5 sequence, $F*$-iCAL36$_{10}$ has 16-fold higher affinity for CALP and a more than 800-fold higher selectivity index. Compared to the CFTR C-terminus, which competes for binding, $F*$-iCAL36$_{10}$ has 300-fold higher affinity for CAL and a circa 3×10^6-fold increase in selectivity. A key aspect of our approach was the combination of multiple affinity determinants along the length of the peptide. By themselves, C-terminal motif-driven changes yielded a selectivity index of only 14 (iCAL05). The rest was contributed by optimization of upstream elements. As has been seen previously in other PPID systems, optimization also required alternating tradeoffs between affinity and selectivity.[23] Sequence–activity studies and/or structure determination followed by rational peptidomimetic approaches may be required to enhance affinity, selectivity and stability.

Overall, we have clearly achieved both our positive and negative design goals: $F*$-iCAL36 has robust affinity for the CAL PDZ domain, but strong selectivity against the four NHERF1 and NHERF2 domains, despite shared sequence specificity of the domains. This provides proof-of-principle for selective PDZ inhibition. Since our strategy (Figure 1) can be easily adapted to other PDZ domain networks (e.g., neuronal PDZ scaffolds), it represents a milestone in the development of peptidic inhibitors of this common class of PPIDs. In principle, this approach could be also adapted to other PPIDs having a pronounced peptide-binding pocket.

Selective CAL inhibitors also represent a potentially novel class of CFTR modulators. As described in the accompanying report,[24] these peptides enable us to analyze the effects of CAL inhibition on the cell-surface abundance of CFTR in bronchial epithelial cells. They also allow us to explore the possibility of cooperative rescue in conjunction with correctors of the primary folding defect of the most common disease-associated allele, ΔF508-CFTR.

[1] D. A. Doyle, A. Lee, J. Lewis, E. Kim, M. Sheng, R. MacKinnon, *Cell* **1996**, *85*, 1067.
[2] Z. Songyang, A. S. Fanning, C. Fu, J. Xu, S. M. Marfatia, A. H. Chishti, A. Crompton, A. C. Chan, J. M. Anderson, L. C. Cantley, *Science* **1997**, *275*, 73.
[3] M. van Ham, W. Hendriks, *Mol. Biol. Rep.* **2003**, *30*, 69.
[4] C. Nourry, S. G. Grant, J. P. Borg, *Sci. STKE* **2003**, *2003*, 7re.
[5] U. Wiedemann, P. Boisguerin, R. Leben, D. Leitner, G. Krause, K. Moelling, R. Volkmer-Engert, H. Oschkinat, *J. Mol. Biol.* **2004**, *343*, 703.
[6] M. A. Stiffler, J. R. Chen, V. P. Grantcharova, Y. Lei, D. Fuchs, J. E. Allen, L. A. Zaslavskaia, G. MacBeath, *Science* **2007**, *317*, 364.
[7] R. Tonikian, Y. Zhang, S. L. Sazinsky, B. Currell, J. H. Yeh, B. Reva, H. A. Held, B. A. Appleton, M. Evangelista, Y. Wu, X. Xin, A. C. Chan, S. Seshagiri, L. A. Lasky, C. Sander, C. Boone, G. D. Bader, S. S. Sidhu, *PLoS Biol.* **2008**, *6*, e239.
[8] W. B. Guggino, B. A. Stanton, *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 426.
[9] J. Li, Z. Dai, D. Jana, D. J. Callaway, Z. Bu, *J. Biol. Chem.* **2005**, *280*, 37634.
[10] M. Favia, T. Fanelli, A. Bagorda, F. Di Sole, S. J. Reshkin, P. G. Suh, L. Guerra, V. Casavola, *Biochem. Biophys. Res. Commun.* **2006**, *347*, 452.
[11] P. Boisguerin, B. Ay, G. Radziwill, R. D. Fritz, K. Moelling, R. Volkmer, *ChemBioChem* **2007**, *8*, 2302.
[12] R. Frank, *J. Immunol. Methods* **2002**, *267*, 13.
[13] R. Volkmer, *ChemBioChem* **2009**, *10*, 1431.
[14] P. R. Cushing, A. Fellows, D. Villone, P. Boisguerin, D. R. Madden, *Biochemistry* **2008**, *47*, 10084.
[15] S. Maudsley, A. M. Zamah, N. Rahman, J. T. Blitzer, L. M. Luttrell, R. J. Lefkowitz, R. A. Hall, *Mol. Cell. Biol.* **2000**, *20*, 8352.
[16] R. A. Hall, R. T. Premont, C. W. Chow, J. T. Blitzer, J. A. Pitcher, A. Claing, R. H. Stoffel, L. S. Barak, S. Shenolikar, E. J. Weinman, S. Grinstein, R. J. Lefkowitz, *Nature* **1998**, *392*, 626.
[17] R. A. Hall, L. S. Ostedgaard, R. T. Premont, J. T. Blitzer, N. Rahman, M. J. Welsh, R. J. Lefkowitz, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 8496.
[18] T. Shibata, M. Chuma, A. Kokubu, M. Sakamoto, S. Hirohashi, *Hepatology* **2003**, *38*, 178.
[19] G. E. Crooks, G. Hon, J. M. Chandonia, S. E. Brenner, *Genome Res.* **2004**, *14*, 1188.
[20] S. H. Joo, D. Pei, *Biochemistry* **2008**, *47*, 3061.
[21] M. R. Arkin, J. A. Wells, *Nat. Rev. Drug Discovery* **2004**, *3*, 301.
[22] A. Mayasundari, A. M. Ferreira, L. He, N. Mahindroo, D. Bashford, N. Fujii, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 942.
[23] G. Grigoryan, A. W. Reinke, A. E. Keating, *Nature* **2009**, *458*, 859.
[24] P. R. Cushing, L. Vouilleme, M. Pellegrini, P. Boisguerin, D. R. Madden, *Angew. Chem.* **2010**, DOI: 10.1002/ange.201005585; *Angew. Chem. Int. Ed.* **2010**, DOI: 10.1002/anie.201005585.