

6 Summary

A notable fraction of protein-protein interactions are mediated by a small number of adaptor domain families (e.g. PDZ, SH2, SH3, WW), which recognize linear sequence motifs. The analysis of the resulting interaction networks requires a quantification of domain specificity and selectivity towards all possible ligands with physiologically relevant affinity.

PDZ domains are ubiquitous protein interaction modules that play a key role in cellular signaling. They bind mainly to the carboxyl-termini (C-termini) of their interaction partners, which often belong to receptor and ion channel families. As representative examples for the analysis of protein-protein interactions, we determined the specificity of the AF6, the ERBIN and the α 1-syntrophin (SNA1) PDZ domains.

The aim of this study was to provide a PDZ domain/ligand interaction dataset that, in conjunction with those by others, substantially expands our knowledge on putative PDZ domain/ligand recognition modes. One driven force in this work was to understand how PDZ domains bind *specifically* to a variety of other proteins. It was a particular goal to unravel the mechanism how PDZ domains are ‘tuned’ towards a certain ligand spectrum. The knowledge of this mechanism would allow the rational drug design or the engineering of novel functions.

A prerequisite for reaching these goals was the development of a new strategy to generate cellulose membrane-bound peptides with free C-termini via standard SPOT synthesis for large-scale screening of PDZ domains. Most solid support-bound peptide libraries lack a free C-terminus due to C-terminal fixation using standard SPOT synthesis. To overcome this restriction, we developed a robust methodology based on a chemoselective cyclization/cleavage step to create peptides with free C-termini (Boisguerin *et al.*, *Chem. Biol.*, (2004), **11**, 459-559). A peptide library including all (6223) non redundant human C-termini (6223-Humlib) known at this juncture was screened to obtain new ligands for the AF6, the ERBIN and the SNA1 PDZ domains. The results of the SNA1 screen were successfully compared with earlier publications as a validation of the improved method. As an example, the newfound interaction between the ERBIN PDZ domain and the breakpoint cluster region (BCR) protein kinase was substantiated by *in vivo* studies using the full-length and/or the endogenous proteins.

Multiple substitutional analyses, dissociation constant determination by surface plasmon resonance (SPR) measurements and nuclear magnetic resonance (NMR) titrations were performed to obtain a description of the individual PDZ domain/ligand specificity-relationships. Furthermore, a focused peptide library ('profile library') was designed to derive the individual contribution of the last four C-terminal ligand residues to the total binding affinity. The amino acid type-specific contributions were quantified using an Analysis of Variance (ANOVA) model, which was visualized by means of 'term schemes'. This approach, which quantifies for the first time the specificity of the three investigated PDZ domains, was validated by a successful prediction of super-binding peptides for each PDZ domain. The predicted dissociation constants (K_d) for the interaction with peptides showing all potential combinations of the four C-terminal ligand residues are used to compare the degree of selectivity and promiscuity of these PDZ domains (<http://www.fmp-berlin.de/nmr/pdz> or Boisguerin, Wiedemann *et al.*, accepted at *J. Mol. Biol.*).

Moreover, a putative and novel binding mode, which we denoted 'shifted motif', was derived from the 6223-Humlib screen of the AF6 and the ERBIN PDZ domains. This binding motif represents the conventional binding mode (E(S/T)xV_{COOH}) of PDZ domains with an additional amino acid at the ligand C-terminus (E(S/T)xVX_{COOH}). The specificity of the PDZ domain/shifted motif interaction was proven through *in vitro* experiments such as binding studies of substitution analogues, determination of K_d values, and further through *in vivo* co-localization experiments, exemplified on the ERBIN/angiotensin type II receptor.

In this work, we present a generally applicable and novel analysis of the interactions of adaptor domains with peptide ligands. Peptide libraries, which are synthesized with the improved method of inverted peptides, are applied to 'quantify' the specificity of the PDZ domains towards all possible ligands. The resulting specificity profiles are a display of amino acid type-specific relative affinity contributions for residues in each of the relevant C-terminal ligand positions of the conserved PDZ domain binding mechanism. Our approach includes the novel abilities to rationally design peptides with maximum affinity (super-binders) and to calculate dissociation constants towards the complete potential peptide ligand sequence space. This analysis and the term schemes make it now possible to rationalize to which extent PDZ domains bind to a diverse set of peptide sequences.

Zusammenfassung

Nur eine kleine Auswahl von Protein-Domänen (z.B. PDZ, SH2, SH3, WW), die lineare Sequenzmotive in ihren Liganden erkennen, vermittelt eine beachtliche Anzahl von Protein-Proteininteraktionen. Dabei spielen PDZ Domänen, ubiquitäre Proteinmodule, eine Schlüsselrolle in einer Vielzahl von zellulären Prozessen. Sie wechselwirken meistens mit den Carboxyltermini (C-Termini) bestimmter Proteine, insbesondere mit den Rezeptoren und den Ionenkanälen. Um dieses Interaktionsnetzwerk zu untersuchen, musste die Spezifität und die Selektivität aller möglichen Domänen/Liganden Wechselwirkungen innerhalb einer physiologisch relevanten Bindungsaffinität quantifiziert werden. Dazu wurden repräsentativ die AF6, die ERBIN und die *α 1-syntrophin* (SNA1) PDZ Domänen ausgewählt.

Ziel der vorliegenden Arbeit war es also, einen PDZ Domänen/Liganden Interaktionsdatensatz zu erhalten, der unser Verständnis über den PDZ Domänen/Liganden Bindungsmechanismus wesentlich erweitert. Dabei war eines der Hauptanliegen zu verstehen, wie PDZ Domänen *spezifisch* eine Vielzahl von Proteinen binden können und den Mechanismus zu erläutern, wie sie ein bestimmtes Ligandenspektrum selektieren. Die Kenntnis über diesen Mechanismus, die neuen Funktionen und Wechselwirkungen, erlauben letztlich die rationale Entwicklung von Arzneimitteln.

Unabdingbar musste dazu eine neue Strategie entwickelt werden, um zellulosemembran-gebundene Peptide mit freien C-Termini mittels SPOT Synthese herzustellen. Bedingt durch ihre Synthese haben aber die, mit konventionellen Methoden hergestellten, träger-gebundenen Peptide keinen freien C-Termini. Um dieses zu umgehen, wurde eine Methode entwickelt, die mittels eines chemoselektiven Zyklisierungs- und Spaltungsschrittes die Synthese von Peptiden mit freien C-Termini ermöglicht (,invertierte Peptide', Boisguerin, et al., *Chem. Biol.*, (2004), **11**, 459-559). Eine Peptidbibliothek, die alle (6223) zu diesem Zeitpunkt bekannten und nicht-redundanten humanen C-Termini umfasst (6223-Humlib), wurde mit den PDZ Domänen von AF6, ERBIN und SNA1 *gescreent*, um bisher unbekannte Liganden zu finden. Innerhalb der besten detektierten Wechselwirkungen konnten die in der Literatur bekannten PDZ Domänen/Peptid Interaktionspaare wieder gefunden werden. So wurde die Methode der ,invertierten Peptide' validiert. Die neu gefundene Interaktion zwischen der ERBIN PDZ Domäne und dem C-Terminus des BCR (*breakpoint cluster region*) Proteins konnte mittels *in vivo* Experimenten mit dem jeweiligen *full-length* und endogenen Proteinen bestätigt werden.

Multiple Substitutionsanalysen, Messungen von Dissoziationskonstanten (K_d) und NMR Titrationsen wurden verwendet, um eine detaillierte Beschreibung der spezifischen PDZ Domäne/Liganden Wechselwirkung zu erhalten. Darüber hinaus wurde eine fokussierte Peptidbibliothek (*profile library*) entworfen, um die Einzelbeiträge der letzten vier C-terminalen Aminosäuren zur gesamten Bindungsaffinität zu bestimmen. Der spezifische Beitrag der einzelnen Aminosäuren wurde mittels des ANOVA Modells (*Analysis of Variance*) quantifiziert und mit Hilfe der ‚Termschemen‘ visualisiert.

Diese Herangehensweise, die erstmalig die Spezifität einer PDZ Domäne mittels K_d s quantifiziert, erlaubt es erfolgreich super-bindende Peptide für alle drei untersuchten PDZ Domänen vorherzusagen und damit zu validieren. Abschließend konnte die Vorhersage der K_d s für alle potentiellen 130321 (19^4) Ligandensequenzen verwendet werden, um die Selektivität sowie der Überlappungsbereiche der drei PDZ Domänen zu vergleichen (<http://www.fmp-berlin.de/nmr/pdz> oder Boisguerin, Wiedemann, et al., angenommen bei *J. Mol. Biol.*).

Außerdem konnte aus dem 6223-Humlib *screen* ein potentieller und neuer Bindungsmodus für die AF6 und die ERBIN PDZ Domänen abgeleitet werden, der als *shifted motif* bezeichnet wird. Dieser Bindungsmodus beinhaltet das konventionelle Bindungsmotiv (ETxV_{COOH}) von PDZ Domänen, ist aber durch eine zusätzliche Aminosäure am äußersten C-Terminus erweitert (ETxVX_{COOH}). Die Spezifität der PDZ Domäne/*shifted motif* Interaktion wurde exemplarisch am ERBIN/Angiotensin Type II Rezeptor mit Hilfe von Versuchen wie Substitutionsanalysen, K_d Bestimmungen (beides *in vitro*) und durch *in vivo* Co-Lokalisationsexperimente bewiesen.

Somit stellt diese Arbeit eine generell anwendbare und neuartige Möglichkeit zur Untersuchung der Interaktionen zwischen Domänen und ihren peptidischen Liganden vor. Peptidbibliotheken, mittels der Methode der ‚invertierten Peptide‘ hergestellt, wurden zur ‚Quantifizierung‘ der Spezifität von PDZ Domänen, bezogen auf alle möglichen Liganden, verwendet. Das resultierende Spezifitätsprofil stellt den aminosäure-spezifischen relativen Affinitätsbeitrag in jeder der relevanten C-terminalen Ligandenpositionen des konservierten PDZ Domänen-Bindungsmechanismus dar. Unsere Herangehensweise beschreibt eine neue Möglichkeit, Peptide mit maximaler Affinität rational zu entwerfen (Superbinder) und ebenso Bindungskonstanten für den gesamten möglichen Ligandensequenzraum vorherzusagen. Zusammen mit den Termschemen erklärt diese Analyse also rational, warum PDZ Domänen an diversen Peptidsequenzen binden können.