Specificity, Structure and Inhibitor Screen of Tiam1 PDZ Domain/Syndecan1 Complex

1. INTRODUCTION

Cells sense and respond to the extracellular environment using signal transduction systems. Upon activation of cell-surface receptors, such as cell adhesion receptors, multi-protein complexes assemble at the cell membrane and induce downstream signaling. One important signal transduction system is Rho-family of GTPases, consisting Rac1, RhoA and Cdc42, which regulates cell polarity, migration, differentiation and division\(^1\). Rho GTPases are molecular switches that cycle between the GTP-bound active state and the GDP-bound inactive state. Spatiotemporal regulation of Rho GTPases comes from three classes of regulatory proteins: guanine nucleotide exchange factors (GEFs), GTPase-activating proteins (GAPs), and guanine nucleotide dissociation inhibitors (GDIs)\(^2\). GDIs sequester the GTPase in an inactive GDP-bound form. GAPs increase the intrinsic GTPase activity leading to deactivation of the GTPase. GEF proteins catalyze the release of bound GDP to be exchanged with GTP, thus activating the GTPase (\textbf{Figure1}). GEF proteins have catalytic Dbl-homology (DH)-Pleckstrin homology (PH) domain as well as other signaling domains. The interactions between these signaling domains and upstream signaling proteins dictate the localization, specificity and activity of Rho GEFs\(^3\).

The long-term goal of my research is to define how GEFs organize upstream signals to control GTPase activity. In this report, I will focus on the Rac1-specific GEF named T-cell lymphoma invasion and metastasis 1 (Tiam1) and determine how it regulates Rac1 signaling through a specific interaction with a cell adhesion receptor, syndecan1.

The TIAM1 gene was originally identified in a screen for genes inducing an invasive phenotype\(^4\). Tiam1 is involved in regulating cell-cell adhesion and cell migration in both epithelial and endothelial cells\(^5\). Tiam1 is a multi-domain GEF protein, which is composed of a
DH-PH catalytic domain, a Pleckstrin homology-coiled-coil-extension (PH-CC-Ex) domain, a Ras binding domain (RBD) and a post-synaptic density-95/discs large/zonula occludens-1 (PDZ) domain. The TIAM2 gene was identified from mouse brain tissue and determined to encode a GEF protein like Tiam1\(^6\). Tiam1 and Tiam2 share the same domain architecture, though the sequence conservation of individual domain varies largely (Figure 2). Several studies have attempted to understand the molecular details of Tiam1 and Tiam2 regulation and activation, however, little is known about the role of the PDZ domains in Tiam proteins. Given that PDZ domains share only 28\% identity in Tiam1 and Tiam2, the two proteins can be differentially regulated in a PDZ-dependent manner and may ultimately have different functions.

PDZ domains are small protein-protein interaction domains and share a similar structure including two α helices and six β strands. PDZ domains typically bind the 5-10 extreme C-terminal residues of their interaction partners by making strand-strand interactions between the peptide ligand and its β2 strand. The nomenclature within the PDZ-peptide interaction is as follows: the peptide residue at C-terminus position is referred to the P\(_0\) residue; subsequent residues towards the N-terminus are termed P\(_{-1}\), P\(_{-2}\), etc; the specificity pocket on the surface of PDZ domain occupied by P\(_0\), P\(_{-1}\) peptide residue is referred to S\(_0\), S\(_{-1}\), etc. The specificity of the interaction generally comes from the side chain interactions between the PDZ domain and the ligand, majorly from that the P\(_0\) and P\(_{-2}\) residues.

PDZ domain-containing proteins function \textit{in vivo} to organize multi-protein complexes. In epithelial and endothelial cells, these signaling complexes assemble at the plasma membrane through interaction with cell adhesion receptors. Among these receptors, syndecans1-4 is a four-member transmembrane proteoglycan family. Their extracellular domain recognizes the extracellular matrix (ECM) through the polysaccharide chains and activates syndecan signaling\(^7\).
Syndecans have a PDZ binding motif at the cytoplasmic tail, which can be recognized by synectin, syntenin and Tiam1 PDZ domains\textsuperscript{8-10}. Additionally, it has been shown that syndecan1 can be phosphorylated on the penultimate tyrosine (phos-syndecan1)\textsuperscript{11}. Furthermore, this (de)phosphorylation is indicated to be a signaling event in syndecan1/ syntenin pathway and determines the syntenin binding\textsuperscript{12}.

Numerous studies strongly correlate Tiam1/Rac1 deregulation to enhanced cell migration and invasion, resulting in metastatic potential in cancers\textsuperscript{13, 14}. Likewise, overexpression of syndecan1 has also been well documented in enabling the promotion of growth and invasion of cancer\textsuperscript{15}. The current model is that syndecan1 binds to the ECM which leads to the activation of Rac1 signaling to trigger the assembly of the actin cytoskeleton and intercellular junctional complexes at the cytoplasmic membrane\textsuperscript{16}. Dysfunction of these processes leads to disassembly of proper cell-cell junctions, increase of cell migration and finally causes tumor progression. In these pathways, PDZ-mediated protein interaction with syndecan1 is required for signaling complex assembly and Rac1 activation. Therefore, this interaction is an attractive target for development of small-molecule inhibitors that can assist in dissecting molecular signaling mechanisms and formulating pharmaceutical agents. The availability of high-resolution structures of the PDZ domain/syndecan1 complex will facilitate rational drug targeting, screen and design.

2. METHODOLOGY

**Protein expression and purification.** The polymerase chain reaction was used to amplify the human Tiam1 PDZ domain (residues 841–930) from the full-length DNA sequence. The amplified DNA was ligated into a modified pET21a vector (Novagen) that contains an N-terminal His6 tag and a recombinant tobacco etch virus (rTEV) protease cleavage site. PDZ mutations were produced using oligonucleotide-directed mutagenesis (QuikChange, Stratagene) with Tiam1 PDZ domain wild-type as a template. All mutants were verified by DNA sequencing.
(University of Iowa, DNA Facility). All protein expression was conducted in BL21 (DE3) (Invitrogen) *Escherichia coli* cells. Typically, *E. coli* cells were grown at 37°C in Luria-Bertani medium supplemented with ampicillin (100 mg/mL) under vigorous agitation until an $A_{600}$ of 0.6 was reached. Cultures were subsequently cooled to 25°C and protein expression was induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to 1 mM final concentration. Induced cells were incubated for an additional 6-8 h at 25°C and harvested by centrifugation.

The Tiam1 PDZ domain was purified by nickel-chelate and size-exclusion chromatography. The N-terminal His6 affinity tag was removed by proteolysis with rTEV protease for 12 h at 25°C. Undigested protein, cleaved His6 tag and His-tagged rTEV were separated from PDZ domain by nickel-chelate chromatography. The final yield of PDZ domain was ∼20 mg (>95% pure as verified by SDS- PAGE) from 1 L of culture.

**Synthetic peptides.** All peptides were synthesized by GenScript Inc. (Piscataway, NJ) and were >95% pure as verified by analytical HPLC and mass spectrometry. Peptides for fluorescence anisotropy-based binding assays and crystallography were dansylated at their N-terminus. Peptide concentrations were determined by absorbance measurements (A280) using extinction coefficient calculated from SEDNTERP (v1.09).

**Fluorescence anisotropy.** Fluorescence anisotropy was used to monitor the binding of dansylated peptides to the Tiam1 PDZ domain. PDZ protein (1.5 mM stock) was titrated into 1.3 mL of 0.2 μM dansylated peptide contained in a stirred quartz cuvette until saturation. Fluorescence anisotropy measurements were recorded at 25°C on a Fluorolog 3 (Jobin Yvon, Horiba) spectrofluorimeter with excitation of 340 nm and emission of 550 nm. Data were baseline corrected with a buffer blank. Bmax and Kd for each titration were determined by fitting baseline-corrected data with non-linear regression using following equation:
\[
(A - A_{\text{min}}) = \frac{B_{\text{max}}[\text{PDZ}]}{K_d + [\text{PDZ}]}
\]

For presentation, the data for each curve in Figures 3 were normalized to the fitted Bmax. Reported dissociation constants are the average of at least three independent experiments.

**Crystallization and data collection.** Crystallization conditions for the syndecan1 and phos-syndecan1 peptide-bound forms of the Tiam1 PDZ domain were determined by the hanging-drop, vapor-diffusion method using sparse-matrix screens automated by a Mosquito drop setter (TTP LabTech). Equal volumes (0.75 μL) of precipitant and protein (20 mg /mL) in 20 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$ (pH=6.9), 50 mM NaCl with 5 molar equivalents of each peptide were used for screens. Crystals of the PDZ/syndecan1 complex were obtained in condition of 0.1 M MES, pH 6.5, 20% PEG 8000, and the Tiam1 PDZ/phos-syndecan1 formed crystals in condition of 0.1 M sodium acetate, 25% PEG 4000, 8% isopropanol. Before data collection, crystals were soaked for 10 s in mother liquor supplemented with 10% glycerol and then flash-frozen in liquid nitrogen. Full X-ray diffraction datasets of the crystals were collected at the Advanced Light Source (Lawrence Berkeley National Laboratories, Berkeley, CA) beam line 4.2.2 (\(\gamma = 1.0 \text{ Å}\)) with a NOIR-1 CCD detector at 0.5° oscillation over 180°. The PDZ/syndecan1 complex crystallized in space group P2$_1$ with two molecules in the asymmetric unit and the PDZ/phos-syndecan1 complex crystallized in space group P2$_1$2$_1$2$_1$ with one molecule per asymmetric unit. Proper space group handedness was verified by analysis of the electron density.

**Structure determination and refinement.** Indexing, integration, and scaling were performed using d*TREK\textsuperscript{17}. The structure was solved by molecular replacement using MOLREP\textsuperscript{18}. The initial phases were determined using the apo-form structure of the Tiam1 PDZ domain (PDB codes 3KZD) as a model. Further refinement was performed in the REFMAC, and electron density and manual model building was done using Coot. Water molecules were modeled into
electron density using Coot. R$_{free}$ values were calculated using 10% of the reflections selected randomly. The structures were refined to 1.85 Å in the syndecan1-bound form and 1.54 Å in the phos-syndecan1-bound form. MolProbity was used to validate the structures$^{19}$. All structure figures were created by PyMOL (v1.4).

**In silico docking screen.** The structures of Tiam1 PDZ/syndecan1 and PDZ/phos-syndecan1 were used for *in silico* docking studies to identify small-molecule inhibitors. The S$_0$, S$_2$ specificity pockets and S$_0$, S$_1$, S$_2$ pockets of the two complexes (**Figure 5 A & B**) were targeted, respectively. The DOCK program$^{20}$ was used to screen 14 million compounds from the ZINC database$^{21}$, a virtual library of commercially available drug-like molecules. We selected 10 top-ranked compounds with best docking scores for the PDZ/syndecan1 complex and 12 for the PDZ/phos-syndecan1 complex. These primary hits were further validated by solution NMR technique. A series of 2D $^{15}$N-$^1$H Heteronuclear Single-Quantum Correlation (HSQC) spectra of $^{15}$N labeled PDZ domain with and without these compounds were collected on a Varian 600 MHz NMR spectrometer. Compounds were added into 0.5mM PDZ domain in several titration steps to a final molar ratio of compound to PDZ domain of 5:1.

**AlphaScreen-based high throughput screen (HTS).** Appropriate HTS with high sensitivity and low cost is necessary for large scale screen of potential inhibitors. We used Alpha (Amplified Luminescent Proximity Homogeneous Assay) screen. In this assay, two molecules are tagged with specific beads. The interaction between these two molecules would bring the two beads into close proximity ($\leq$200 nm). After a singlet oxygen transfer within this distance, the excitation of the donor beads results in chemiluminescent reaction, which is detected by Alpha signal (**Figure 5A**). For this assay, recombinant GST-PDZ was expressed and purified. Biotin-labeled Caspr4 peptide was used as a substitute of syndecan1 because it showed similar binding pattern with
syndecan1 but higher binding affinity. A cross-titration assay was first performed to determine the optimal reagent concentrations. GST-PDZ and biotin-Caspr4 at various concentrations (0-3μM, at half log increment) were incubated with 20μg/ml anti-GST acceptor beads and streptavidin donor beads (PerkinElmer), respectively. Then 15μL of two beads conjugated to various concentrations of PDZ and Caspr4 were added into each well of the plate and signal was detected at Synergy 2 plate reader (BioTek). In the competition assay, increasing amount of untagged Caspr4 peptide was titrated into preincubated GST-PDZ/biotin-Caspr4/beads complex and the inhibition effect was determined by IC50. To eliminate the interference of DMSO (a solvent to dissolve all compounds) in the subsequent screen, we performed DMSO tolerance assay. 5% DMSO was added into the positive GST-PDZ/biotin-Caspr4 binding and negative binding fully disrupted by untagged Caspr4 with Alpha signal detected afterwards.

3. RESULTS

**Syndecan family specificity.** We have previously shown that syndecan1 can bind to the Tiam1 PDZ domain. Different syndecans are expressed in various tissues and developmental stages. Likewise, expression of Tiam1 is also under spatiotemporal regulation. This suggests that specificity in binding to the Tiam1 PDZ domain may exist in the syndecan family. As syndecan1 and syndecan3 have most similar sequence in their PDZ binding motif (TKQEEFYA vs DKQEEFYA), it is intriguing to study the binding of syndecan2 and syndecan4 to PDZ domain. Fluorescence anisotropy was used to determine the dissociation constant (Kd) of the binding process. Both syndecan2 and syndecan4 have a Kd greater than 200 μM, at least 8-fold weaker than the 26 μM Kd found for syndecan1 (Figure 3).

**Tiam1 PDZ domain/syndecan1 complex structure.** Since all syndecan members have the same C-terminal residues: EFYA (Figure 3), we expect the molecular determinants of PDZ binding specificity to come from other residues. To understand the structural feature of this specificity,
we determined the crystal structure of Tiam1 PDZ domain with syndecan1 peptide. The PDZ domain had six β strands and two α helices. The entire syndecan1 peptide made β sheet interactions along the second β strand of the PDZ domain (858-866). Syndecan1 used two specificity pockets: pocket $S_0$ was formed by the side chains of Y858, F860, L862 and L915, which accommodated the buried methyl side chain of the P₀ alanine of syndecan1; pocket $S_2$ was created between the aliphatic side chains of L911 and K912 and accommodated the phenylalanine at P₂ of syndecan1 (Figure 5 A).

In addition to these typical PDZ/ligand interactions, hydrogen bond network was formed by the side chains of the P⁻³ glutamic acid, the P⁻⁶ lysine and N876 of the PDZ domain. The positively charged side chain of P⁻⁶ lysine was important. Both syndecan2 and syndecan4 have a proline at P⁻₆, which disrupts this hydrogen bond network and thus their binding to PDZ domain. Also, a favorable electrostatic interaction was observed between P⁻⁴ glutamic acid of syndecan1 and K912 of PDZ domain (Figure 4 A). Mutation analysis suggested that disruption of this ion pair dramatically impeded the binding between syndecan1 and the PDZ domain (Figure 4 B). In contrast, the P⁻⁴ residue of syndecan2 is an electrostatically unfavorable lysine. Therefore, we conclude that the residues at P⁻⁴~P⁻⁷ positions of syndecan1 provide a combined contribution to PDZ binding through hydrogen bond and electrostatic interaction and determine the binding specificity between different syndecans.

**Phos-syndecan1 uses a new binding pocket.** It is known that the Tiam1 PDZ domain interacts with both syndecan1 and phos-syndecan1¹⁰. To determine if there is a molecular difference in binding, the structure of Tiam1 PDZ/phos-syndecan1 complex was also solved. This structure showed significantly different phos-peptide interactions from the unmodified peptide. The small backbone RMSD (0.35 Å) indicated that there was no significant rearrangement of the PDZ
domain upon binding (Figure 5 A & B). In contrast, the peptide conformation was different with a new specificity pocket, S$_1$, being filled by the phosphorylated P$_{-1}$ tyrosine. The peptide interaction was stabilized by a salt bridge between the negatively charged phosphate adduct of the P$_{-1}$ tyrosine and the positively charged amine on the side chain of K879, which was distal from the syndecan1 binding pocket (Figure 5B). This interaction required the P$_{-1}$ tyrosine side chain to rotate by 107° compared to the syndecan1 structure. The phosphate group also made hydrogen bonds with the hydroxyl group on the side chain of T857.

To test the importance of K879 in pohos-syndecan1 binding, we performed a double-mutant cycle thermodynamic analysis$^{22}$. The Kd of four possible binding reactions made from syndecan1, phos-syndecan1, wild type PDZ and PDZ (K879E) were measured using fluorescence anisotropy. Following the Gibbs free energy calculation ($\Delta G$) and double-mutant cycle analysis ($\Delta \Delta G$), the $\Delta \Delta G_{\text{INT}}$ was found to be -0.83 kcal/mol (Figure 5C). This indicated that the phosphate group of syndean1 and the side chain of K897 were energetically coupled. The cooperativity between these two was required for phos-syndecan1 binding.

**In silico docking screen and HTS.** The validation of the primary hits discovered by in silico screen was performed by NMR titration experiments. Three compounds showed interaction with the PDZ domain (Figure 6, inset table). The $^{15}$N-HSQC spectrum of one compound (ZINC ID 18276277) was shown in Figure 6A. This spectrum clearly showed the chemical shifts perturbation of several residues, indicating the interaction between this compound and PDZ domain. By mapping these residues onto the PDZ domain surface, we showed that the distribution of these residues was consistent with the predicated binding site (Figure 6 B).

Regarding the unbiased HTS of inhibitors, we developed AlphaScreen assay through several pilot experiments. The cross titration assay showed the ‘hooking effect’: the signal increased first
and began to decrease when the biotin-Caspr4 concentration was above 1µM (Figure 7B). In order to keep the reagent concentration low while still maintain a high assay signal, we chose GST-PDZ at 300nM and biotin-Caspr4 at 100nM for the competition study. In competition assay, the maximal binding was measured in absence of untagged Caspr4. Alpha signal decreased with increasing concentration of untagged Caspr4. The determined IC$_{50}$ was 21±2.3µM (Figure 7C), consistent with $K_d$ measured by fluorescence anisotropy (19µM±0.4µM). In DMSO tolerance assay, high concentration of DMSO resulted in a non-significant loss of signal compared to DMSO-free in both positive and negative bindings (Figure 7D).

4. SUMMARY AND FUTURE DIRECTIONS

In this work we show that the Tiam1 PDZ domain specifically interacts with syndecan1 and has significantly less affinity for syndecan2 and syndecan4. Using X-ray crystallography, we elucidate the structure basis of the Tiam1 PDZ/syndecan1 binding specificity. The Tiam1 PDZ/phos-syndecan1 complex structure shows that the phosphate group induces a change in ligand conformation, which is accommodated by a new specificity pocket. Targeting these two structures, we identify several small-molecule inhibitors by in silico docking screen and further validate them by NMR. We have successfully initiated AlphaScreen-based HTS, a complementary way to discover inhibitors. In the future, 2320 compounds from the Spectrum Library (MicroSource) will be screened using AlphaScreen techniques. Further optimization on the screened compounds from both in silico and HTS will be performed using quantitative structure–activity relationship analysis. The optimized inhibitors will ultimately help investigate Tiam1/syndecan1 signaling in cells and further develop lead compounds that are effective against cancer.
5. FIGURES

Figure 1. The Regulatory cycle for Rho GTPase.
Regulation of RhoGTPases occurs through their interaction with GDI, GAP and GEF proteins. After activation by GEF proteins, GTP-bound GTPases is able to interact with downstream effectors and lead to various biological functions.

Figure 2. Conservation and domain architecture of the Tiam1 and Tiam2 GEF proteins. The percent identity (%ID) between each homologous domain is indicated (see texts for details).

Figure 3. The binding specificity of Tiam1 PDZ domain with syndecan family. Upper panel shows the binding curves of syndecan1, 2 and 4; lower panel shows the sequence alignment of syndecan1, 2 and 4 their affinities of binding to PDZ domain.

Figure 4. The crystal structure of Tiam1 PDZ domain bound to syndecan1 peptide. (A) A stick model showing the interactions in the PDZ/syndecan1 complex. PDZ residues involved in syndecan1 interactions are labeled and colored in yellow. Dotted lines indicate hydrogen bond. (B). Binding affinity between Tiam1 PDZ wt (or K912E mutant) and syndecan1 peptide.
Figure 5. Comparison between syndecan1 and phosphorylated syndecan1 bound to Tiam1 PDZ. (A)-(B). Space-filling model of these two complexes. Peptide is colored in green, residues for peptide binding are colored in red. (C) Double mutant cycle shows the importance of PDZ K879 in stabilizing the phospho-adduct.

Figure 6. Validation of in silico docking identified compounds by NMR. (A) HSQC spectra of Tiam1 PDZ domain in presence of compound 18276277. (B) The perturbed residues are mapped on PDZ surface to show the binding site. Inset table is the summary of five compounds tested by NMR.

Figure 7. AlphaScreen-based HTS. (A) Alpha screen-based platform for discovery of inhibitors targeting PDZ-Caspr4 interaction. (B) Cross-titration assay used to determine the best reagent concentration. (C) Competition assay using untagged Caspr4 to compete with biotin-Caspr4 for binding to GST-PDZ. (D) DMSO tolerance assay shows alpha signal is tolerant to 5% DMSO.
REFERENCES