Published in final edited form as:

Protein Expr Purif. 2007 May ; 53(1): 51–62.

Remedial Strategies in Structural Proteomics: Expression, Purification and Crystallization of the Vav1/Rac1 Complex

Alexei Brooun^{1,*}, Scott A. Foster¹, Jill E. Chrencik¹, Ellen Y.T. Chien¹, Anand R. Kolatkar¹, Markus Streiff², Paul Ramage², Hans Widmer², Gisbert Weckbecker², and Peter Kuhn¹

¹Department of Cellular Biology, The Scripps Research Institute, La Jolla, CA 92037 USA ²NIBR (Novartis Institutes for BioMedical Research), CH-4002 Basel, Switzerland

Abstract

The signal transduction pathway involving the Vav1 guanine nucleotide exchange factor (GEF) and the Rac1 GTPase plays several key roles in the immune response mediated by the T cell receptor. Vav1 is also a unique member of the GEF family in that it contains a cysteine-rich domain (CRD) that is critical for Rac1 binding and maximal guanine nucleotide exchange activity, and thus may provide a unique protein-protein interface compared to other GEF/GTPase pairs. Here we have applied a number of remedial structural proteomics strategies, such as construct and expression optimization, surface mutagenesis, limited proteolysis, and protein formulation to successfully express, purify, and crystallize the Vav1-DH-PH-CRD/Rac1 complex in an active conformation. We have also systematically characterized various Vav1 domains in a GEF assay, and Rac1 *in vitro* binding experiments. In the context of Vav1-DH-PH-CRD, the zinc finger motif of the CRD is required for the expression of stable Vav1, as well as for activity in both a GEF assay and *in vitro* formation of a Vav1/Rac1 complex suitable for biophysical and structural characterization. Our data also indicate that the isolated CRD maintains a low level of specific binding to Rac1, appears to be folded based on 1D-NMR analysis and coordinates two zinc ions based on ICP-MS analysis. The protein reagents generated here are essential tools for the determination of a three dimensional Vav1/Rac1 complex crystal structure and possibly for the identification of inhibitors of the Vav1/Rac1 protein-protein interaction with potential to inhibit lymphocyte activation.

Introduction

The Dbl family of guanine nucleotide exchange factors (GEFs) plays important roles in mediating vital cellular processes such as cytoskeletal rearrangements, mitogenesis, transcriptional changes and lymphocyte activation. In response to extracellular stimuli, GEF family members are responsible for the transition of Rho/Rac GTPases from the inactive GDP bound state to the active GTP bound state [1]. GTPase activating proteins (GAP) family members, on the other hand, bind to activated GTPases to promote GTP hydrolysis and antagonize downstream signaling. Rho proteins are overexpressed or hyperactivated in human cancers which is a result of dysregulation of the GDP-GTP cycle [2].

The Vav subgroup of Dbl GEFs contains three members (Vav1, Vav2, and Vav3) in mammals with distinctive patterns of cellular expression [3,4]. Vav1 is exclusively expressed in

*Corresponding Author and Requests for Reprints and Materials: Alexei Brooun, Department of Cellular Biology, The Scripps Research Institute 10550 N. Torrey Pines Rd MB-201, La Jolla, CA 92037 USA (858) 784-7332; (858) 784-7355 fax

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

hematopoietic cells in adult mouse, while Vav2 and Vav3 have more ubiquitous expression patterns [5]. Vav is a multidomain protein which consists of a calponin-homology domain (CH), an acidic domain (AC), a Dbl-homology domain (DH), a plekstrin-homology domain (PH), a cysteine-rich domain (CRD), and two SH3 domains flanking a SH2 domain [6](Figure 1). The DH domain has been identified as the catalytic domain, and is responsible for the release of GDP from the GTPase. The catalytic activity of the DH domain is regulated by several autoinhibitory interactions [7,8]. Recent electron microscopy (EM) work involving Vav3 has described the transition between the inactive and active conformations resulting in the displacement of the CH/CRD interaction and the dislocation of the CH and AC domains from the catalytic site [8]. Displacement of the CH and AC domains has been described *in vivo* to result in the primary relief of autoinhibition, and involves three tyrosine residues (Y142, Y160, and Y174) in the acidic domain [9]. These residues become phosphorylated by several receptor and nonreceptor tyrosine kinases [3,10]. The phosphorylation of the first two tyrosine residues results in an increase in affinity for the kinase to Y174. Displacement of Y174 exposes a hydrophobic surface on the DH domain that is required for binding to, and displacement of, GDP from the GTPase. Several oncogenic versions of Vav have been identified ($\Delta 1-66$, $\Delta 1-186$, $\Delta 1-186 + \Delta 608-845$, and Y142,160,174F) which relieve autoinhibition [11]. A number of reports have also indicated that PI3 kinase provides an additional control mechanism for Vav family members via binding of PIP3 to the PH domain, which leads to an increase in association with its cognate GTPase [12,13].

Structural studies with members of the Dbl GEF family have revealed the determinants of the DH domain-GTPase interaction. Those studies were performed with DH-PH constructs, and have resulted in variable orientations of the PH domain with respect to the GTPase [14-16]. Interestingly, Vav proteins are unique from other members of the Dbl GEF family in that they contain a cysteine-rich domain (CRD) C-terminal to the PH domain. The CRD has been shown to control the “on-rate” of the nucleotide exchange reaction and to associate with Rac1 by itself, therefore making the CRD/GTPase binding interface a novel feature of the Vav1/Rac1 protein-protein interaction [17]. The Raf-CRD is another example of a CRD/GTPase interaction that plays an important role in Ras oncogenic transformation [18]. The NMR structure for isolated Raf-CRD is available, but its relevance in terms of the mode of the CRD interaction with its cognate GTPase is unclear [19].

Evidence for the potential specificity of the Vav1-Rac1 interaction is supported by the identification of compounds capable of modulating two other GEF family members, Trio and Tiam1, which also interact with Rac1. Both Trio and Tiam1 contain the DH-PH module similar to Vav, but unlike Vav lack a CRD. NCS23766, an inhibitor of the Trio/Rac1 protein-protein interaction, was recently identified using a virtual ligand screening (VLS) approach against a surface groove of Rac1 known to be critical for GEF specification [20,21]. Interestingly, while NCS23766 was not capable of inhibiting the interaction between Rho-specific GEF Lbc, the CDC42-specific GEF intersectin, it also had no effect on the more promiscuous GEF Vav1, which interacts with Rac1 [20]. Additionally, the widely used immunosuppressant azathioprine was recently shown to invoke its action through modulation of the Vav (1 or 3)/Rac1 protein-protein interaction, while Vav1 interactions with other GTPases were characterized by normal activity [22]. This is the first study that points to the therapeutic potential of the Vav1/Rac1 interaction for development of immunosuppressants. In addition, wild type Vav1 was recently shown to be overexpressed in 53% of pancreatic adenocarcinomas and associated with decreased survival rates [23]. Vav RNAi abrogates neoplastic cellular proliferation *in vitro* and *in vivo* even in the presence of oncogenic *KRAS* [23]. Thus, intervention with Vav1-mediated signaling pathways through modulation of its interaction with Rac1 protein may serve as a potential therapeutic target in pancreatic cancer. The inherent specificity of the Vav1/Rac1 interaction is likely due to the unique CRD/GTPase interface making this protein-protein interaction a potential target for drug discovery initiatives.

Structure based drug design is a valid approach towards the identification of small molecule inhibitors of protein-protein interactions [24]. The identification of drugable protein-protein interactions represents an important avenue in the drug discovery paradigm, as the number of protein targets that can be directly modulated by small molecule compounds or antibodies is rapidly declining [25]. Recent advances in structural proteomics technologies, such as parallel protein expression/purification platforms and nanovolume crystallization, allow fairly routine determination of crystal structures of full length prokaryotic and eukaryotic proteins, as well as isolated domains [26]. Protein complexes, on the other hand, are more challenging to characterize structurally, because two well behaved proteins are required in a form capable of binding their cognate partner. Inherently, this becomes a more challenging problem when the protein-protein interaction involves a multidomain protein which binds to a single domain protein. The Vav1-Rac1 interaction represents such a paradigm which is further complicated by low levels of Vav1 expression in insect cells and reports of insoluble Vav1 post cleavage from an MBP fusion construct expressed in *E. coli* [17,27]. An additional obstacle is the intricate nature of the intermolecular inhibition of Vav1 binding to Rac1 controlled via phosphorylation by upstream tyrosine kinases.

While a number of structural proteomics strategies for the crystallization of challenging proteins are in process development, the systematic and sequential application of those strategies to a more difficult set of protein targets has not been extensively described. In this report we present an exhaustive structural proteomics approach applied to the expression, purification and crystallization of the Vav1/Rac1 complex. The resultant protein molecules also represent a toolbox for the identification of potential modulators of the Vav1/Rac1 protein-protein interaction for novel therapeutic strategies for immunological disorder and cancer.

Materials and Methods

Cloning and mutagenesis

All constructs were amplified from cDNA clones corresponding to GeneBank accession nos. AAC25011.1 (human Vav1), CAA45713.1 (mouse Vav1) and AAA36537.1 (Rac1). Human Vav1-AC-DH-PH-CRD, DH-PH-CRD, and DH (specific constructs listed in table 1), mouse Vav1-DH ("open" and "closed" conformation, Table 1) and Rac1 (full length and 1-177) constructs were amplified by PCR and ligated into a modified pET28b vector (Novagen, Madison, WI) containing a Tev (tobacco etch virus) protease site between the *NdeI* and *XhoI* sites. In general, primers were designed containing 20-25 bases specific to 5'- and 3'-ends of DNA fragments and preceded by a restriction site sequence at the 5'end of the primer which was protected by an additional 6 random bases to achieve efficient cleavage with the restriction endonuclease. Site-directed mutagenesis primers were designed according to the guidelines suggested in the manual for the Quick Change protocol (Stratagene, San Diego, CA). All primers were synthesized at a 10 μ M scale and purchased from IDT (Coraville, IA). Vav1-CRD constructs were amplified by PCR and ligated into modified pET28b vectors with an inserted GST cassette, followed by a PreScission protease cleavage site. The Rac1 (1-177) construct was PCR amplified and ligated into the pACYCDuet-1 vector with a GST cassette inserted under the control of a T7 promoter. hVav-DH to mVav-DH mutants were generated using the Quickchange Mutagenesis Kit (Stratagene, La Jolla, CA).

Microexpression and Micropurification

Microexpression analysis of all Vav1 constructs was performed on a 0.75-10 ml scale using either autoinduction or regular induction protocols in BL21-AI *E. coli* strain (Invitrogen, Carlsbad, CA). Overnight Express Autoinduction System 1 (Novagen) solutions were supplemented into LB/kanamycin (50 μ g/ml) with 0.05% L-arabinose. A single colony from a freshly transformed plate was grown in supplemented LB media supplemented with the

Overnight Express Autoinduction System 1 at 37 °C to an OD₆₀₀ of 1. hVav-DH-PH-CRD and CRD cultures were supplemented with ZnSO₄ (100 μM or 500 μM, respectively) and grown overnight at 22° C. Vav1 constructs expressed by a regular induction protocol were grown in LB media supplemented with kanamycin (50 μg/ml) at 37 °C until an OD₆₀₀ of 1. Protein expression was induced with 1mM IPTG and 0.2% arabinose for 3 hours at 22 °C. hVav1-DH-PH-CRD and hVav1-CRD cultures were also supplemented with ZnSO₄ (100 μM or 500 μM, respectively) at the point of induction. Cells were harvested by centrifugation at 2,000 × g for 20 min. Small scale immobilized metal affinity chromatography (IMAC) purification was performed using 50 μl of Ni-NTA resin (Qiagen, Valencia, CA) in a manner similar to that as previously described [29].

Large scale expression

Vav1-DH, Vav1-CRD, Rac1 (full length, and 1-177) and Rac1 (1-177) GST were freshly transformed into *E. coli* strain BL21-AI (Invitrogen, Carlsbad, CA), and grown in LB media containing the appropriate antibiotic at 37° C until an OD₆₀₀ of 1. Protein expression was induced with 1 mM IPTG and 0.2% arabinose for 3 hours at 22° C. Vav1-CRD expression media was supplemented with 500 μM ZnSO₄ at the point of induction. hVav1-DH-PH-CRD was freshly transformed into the *E. coli* strain BL21-AI, and fermented using the Overnight Express Autoinduction System 1 (Novagen, Madison, WI). Autoinduction solutions were diluted into 500 ml of LB/kanamycin (50 μg/ml) containing 0.05% L-arabinose, and grown at 37°C until an OD₆₀₀ of 1. The cultures were supplemented with 100 μM ZnSO₄, and expressed overnight at 22° C. All cells were harvested by centrifugation at 6,000 × g for 15 minutes.

Purification of Vav1-Rac1 complex

Cell pellets containing His-tagged Vav1 and Rac1 (1-177) proteins were lysed in 50 mM Tris, pH 7.8 (or 50 mM HEPES, pH 7.2), 50 mM NaCl, 1 mM MgCl₂, and 0.25 mM TCEP with 0.25 U/μl benzonase (Novagen, Madison, WI) and 24.4 U/μl lysozyme (Epicentre, Madison, WI) at room temperature for 30 minutes. Lysates were clarified by centrifugation at 30,000 × g for 55 minutes at 4° C. Clarified supernatants were purified by IMAC using Ni-NTA Superflow beads (Qiagen, Valencia, CA) at a ratio of 1 ml of settled resin per 10 mg of purified protein on a 20 ml gravity disposable column (Bio-Rad, Hercules, CA). The protein concentration was determined using Coomassie Plus Protein Assay Reagent (Pierce, Rockford, IL) as described by the manufacturer. The Vav1 (170-575) IMAC eluate was incubated with a 1.5-2-fold molar excess of Rac1 (1-177) and dialyzed against 25 mM HEPES, pH 7.2, 100 mM NaCl, and 5 mM DTT (Buffer A) overnight in the presence of Tev protease (tobacco etch virus). Approximately 20 μl of a 2 mg/ml Tev protease solution was used to cleave 1 mg of protein. Following dialysis, the protein was subjected to limited proteolysis using *Staphylococcus aureus* (V8) protease (Pierce, Rockland, IL) in Buffer A. The completion of V8 protease cleavage was confirmed by SDS-PAGE analysis. The protein was loaded onto a Superdex 75 size-exclusion column (GE Healthcare, Piscataway, NJ) equilibrated in Buffer A. Fractions corresponding to the primary Vav1/Rac1 peak were pooled together and supplemented with 5% glycerol and 5 mM DTT (an additional amount of reducing agent was added to account for the degradation of DTT over time; alternatively, 1 mM TCEP was used throughout purification). The protein complex was concentrated to 3-7 mg/ml for structural studies using a Amicon Ultra-15 concentrator with a 10,000 MW cut off membrane (Millipore, Billerica, MA).

Guanine Nucleotide Exchange (GEF) Assay

A GEF assay was established to monitor the release of a fluorescent GDP analog preloaded on the GTPase (Rac1) upon addition of the guanine nucleotide exchange factor (Vav1). Rac1 (1 μM) was preloaded with 60 nM Bodipy-GDP (Molecular Probes) and mixed with 200 μM GTP

and Vav1-DH or -DH-PH-CRD (specific molar ratios noted elsewhere). Fluorescence measurements were taken at 2-minute intervals for a total of 120 minutes on a Magellan Tecan (Tecan Instruments, Zurich, Switzerland) microplate reader at 535 nm (excitation wavelength of 488 nm). Data from three independent replicates were subsequently averaged and fitted using Graph Pad Prism 4.0.

Purification of Vav1-CRD

Isolated Vav1-CRD (501-567 or -575) was expressed in *E. coli* as a GST fusion protein in regular induction conditions overnight at 22° C. The pellet corresponding to 1 L of cell culture was lysed in 20 ml Phosphate Buffered Saline (137 mM Sodium Chloride, 10 mM sodium phosphate, 2.8 mM Potassium Chloride, pH 7.4) supplemented with 10 μ M ZnSO₄, and 5 mM β ME with 0.25 U/ μ l benzonase and 24.4 U/ μ l lysozyme (lysis buffer) at room temperature for 30 minutes. Lysates were clarified by centrifugation at 31,000 g for 55 minutes at 4° C. Clarified supernatants were purified on glutathione resin at a ratio of 1 ml of resin per 3 mg of purified protein (GE Healthcare, Piscataway, NJ). Briefly, gravity columns were equilibrated with lysis buffer (containing an additional 1mM ATP and 10 mM MgCl₂), washed with lysis buffer and eluted with 30 mM reduced glutathione in lysis buffer. The fusion protein was subsequently cleaved overnight with 20 units of PreScission protease per 1 mg of fusion protein (GE Healthcare, Piscataway, NJ) in dialysis against 1X PBS, pH 7.1, 20 μ M ZnSO₄, and 2 mM β ME. The Vav1-CRD was isolated from its fusion partner on a Superdex 75 size exclusion column equilibrated with 1X PBS, pH 7.1, 100 μ M ZnSO₄, and 10 mM β ME.

Inductively coupled plasma mass spectrometry (ICP-MS)

The zinc content of the isolated Vav1 CRD was determined by inductively coupled plasma mass spectrometry (ICP-MS) using a ThermoQuest/Finnigan Element2 mass spectrometer. Prior to ICP-MS analysis, the isolated CRD was purified as described above. The ICP-MS was externally calibrated with standard solutions prepared using zinc sulfate. A calibration curve was generated by linear regression. The zinc content was determined for three independent Vav-CRD samples in the 30-60 μ M protein concentration range from the ratio of total zinc in the sample and the protein concentration.

Purification of GST-Rac and Pull-down Experiments

Rac1 (1-177) was expressed in *E. coli* as a GST fusion using regular induction protocols overnight at 22 °C. The pellet corresponding to 1 L of cell culture was lysed in 20 ml of Phosphate Buffered Saline (137 mM sodium Chloride, 100 mM sodium phosphate, 2.8 mM Potassium Chloride, pH 7.4) supplemented with 0.25 mM TCEP, 1 mM MgCl₂, 0.25 U/ μ l benzonase and 24.4 U/ μ l lysozyme at 25° C. Lysates were clarified by centrifugation at 31,000 $\times$ g for 55 minutes at 4° C, and loaded onto glutathione resin (GE Healthcare) which had been pre-equilibrated in lysis buffer. GST-Rac1-bound glutathione beads were washed with 40 column volumes of 25 mM MOPS, pH 6.8, 100 mM NaCl, and 5 mM DTT (pull down buffer). Purified Vav1 constructs (mVav1-DH, hVav1-CRD, and hVav1-DH-PH-CRD) were dialyzed into the pull-down buffer for all pull-down experiments. A 2-fold molar excess of Vav1 constructs was preloaded onto GST-Rac1 bound glutathione beads, and incubated at 4° C for 3 hours. The column was washed extensively with pull down buffer and GST-Rac1/Vav1 complexes were subsequently eluted with 30 mM reduced glutathione.

Crystallization of Vav1/Rac1 complex

The hVav1-DH-PH-CRD-Rac1 complex was concentrated to 3-7 mg/mL in a buffer containing 25 mM Hepes, pH 7.2, 100 mM NaCl, and 10 mM DTT (or 1 mM TCEP; crystallization buffer), and crystallized in a sitting drop by vapor diffusion at 293 K. Data were collected at Beamlines 9-1 and 11-1 of the Stanford Synchrotron Radiation Lab (SSRL; Menlo Park, CA) or at

Beamline ID-29 of the European Synchrotron Radiation Facility (ESRF; Grenoble, France), and the diffraction data were processed and reduced using the HKL2000 program package [28].

Results and discussion

Cloning and expression of Vav1 domains

A single expression vector strategy was chosen over a multiple domain expression strategy with the goal of identifying fully active Vav1 constructs for structural, biochemical, and biophysical studies. A modified pET28b vector containing a Tev protease cleavage site was constructed and utilized for the generation of constructs summarized in Table 1. These constructs represent a select subset based on the micro-expression data of a sequential truncation series around hVav1-DH, -DH-PH, -DH-PH-CRD and -CRD domains (Figure 1). In general, sequential truncations of 3-5 amino acids from each domain termini were generated to identify constructs with soluble expression patterns in microexpression experiments. Small scale IMAC purification was performed as described above, and the resulting IMAC eluates were analyzed for purity, yield, and polydispersity using size exclusion chromatography (SEC).

hVav1-DH domain

Interestingly, the “closed” conformation of the hVav1-DH construct, consisting of an intact N-terminal helix and containing a regulatory Y174, provided a significant advantage for protein stability (Figure 1). The “open” conformation constructs (deletion of the N-terminal helix) yielded significantly lower amounts of soluble protein in IMAC purification (Fig.2, lane 8). This is notably different from microexpression studies with mouse Vav1 (mVav DH) residues 170-375 and 181-375, which were previously reported to express in the soluble fraction [7]. In addition, a NMR structure was determined for both the autoinhibited “closed” conformation and the phosphorylated, inhibition-relieved “open” conformation [7]. The open conformation of the Vav1-DH is of significant scientific value in the development of screening assays for the identification of modulators of DH domain interactions with Rac1. To fully understand the discrepancies in expression between mVav-DH and hVav1-DH, a stringent sequence comparison analysis was completed which identified seven non-conserved amino acid regions on the surface of the human and mouse DH domains containing a total of 12 amino acids. The selected amino acids showed significant amino acid variability between the mouse and human DH domain. Based on this information, we generated a panel of human Vav1-DH domain mutants containing the related point mutations. Each of the seven sites of the Vav1-DH domain was individually mutated to the corresponding residue of mVav1-DH. Although most of the mutants had no significant effects on hVav1-DH solubility, one mutant, M351T, restored expression levels and yields similar to the reported levels for mVav1-DH (Fig. 2, lanes 7, 9 respectively). Interestingly, hVav2 and Vav3 maintain a threonine and a proline, respectively, at the position corresponding to M351 of hVav1-DH. A change in the total expression level for hVav-DH M351T could be explained either by changes in mRNA copy number or by improved protein folding of the mutant version which could reduce its susceptibility to proteolytic degradation in the *E. coli* expression host. Interestingly, the M351T mutation in the context of the DH-PH-CRD protein has no effect on expression level compared to wild type, but increases the yield of mutant protein 2-3 fold when compared to wild type. This data would argue against the role of M351T in regulating mRNA levels for DH domain alone.

hVav1-DH-PH and PH-CRD domains

A set of constructs were generated for human hVav1-DH-PH domain with three N-terminal starting points at residues -170, -181, and -190. Those N-terminal starting points were complemented with a C-terminal five-amino acid truncation series. While overall expression levels were quite high for all hVav1-DH-PH constructs, the protein expressed in the insoluble

fraction, and attempts to refold hVav1-DH-PH proteins were unsuccessful (P. Ramage, unpublished data). Successful expression of the hVav1- DH-PH module has been reported for other members of the GEF family. We have also generated a series of constructs for the hVav1-PH-CRD domain with a five-amino acid N-terminal truncation series (starting at residue -370), all containing an optimized C-terminus at amino acid 575. Similar to the results obtained with the DH-PH module, all constructs over-expressed optimally in *E. coli*, but were inevitably insoluble.

However, we have successfully expressed soluble derivatives of hVav1 for the isolated DH and CRD domains respectively, along with a DH-PH-CRD module. It is possible that PH domain serves as rigid linker that allows CRD to interact with DH domain, as both domains are required for stable protein expression of the whole DH-PH-CRD module.

hVav1-CRD and hVav1-DH-PH-CRD domains

Previous studies indicated that the hVav1-CRD controls Vav1/Rac1 complex formation, however no direct binding experiments established whether such a complex could be isolated in quantities sufficient for structural studies [4,12,17]. While this work was in progress, work on *E. coli* expression of the hVav1-CRD, hVav1 PH-CRD, and hVav2 DH-PH-CRD MBP fusion modules and their detailed kinetic characterization of Rac1 binding was reported [17]. While MBP fusions provide a robust expression/purification platform for the generation of soluble eukaryotic proteins in a prokaryotic expression host, characterization of these fusion proteins often indicates the presence of soluble aggregates [30]. Upon cleavage of the MBP solubility partner, the target protein is often in an unsuitable form for structural characterization. The expression of various domains of hVav3 was reported in an insect cell expression system, and indicates insufficient yields for structural studies. In addition, purification of hVav3 constructs required purification in the presence of detergents, which are also unsuitable for structural studies [4]. Thus, we have pursued the identification of a hVav1-DH-PH-CRD construct with a small cleavable N-terminal tag to generate protein material for structural and functional studies.

Based on secondary structure prediction analysis and previously published reports, we designed a C-terminal truncation series between amino acids 567 and 608, with both open and closed conformations of the hVav1 N-terminus (amino acids 170 and 180 or 190 accordingly) with the aim to identify the smallest possible hVav1-DH-PH-CRD module suitable for structural work. Simultaneously, we explored the possibility of obtaining the isolated CRD domain using a GST expression cassette. Several Zn-finger binding motifs of various transcription factors have been successfully expressed as GST fusions [31]. Due to the presence of the Zn-finger motif, we have explored Zn^{2+} as an additive in our expression experiments using regular induction (IPTG + arabinose in the case BL21-AI strain) and autoinduction protocols to maximize our chances of obtaining soluble protein. Recently developed autoinduction expression protocols have been successful in the expression of toxic proteins, and often increase the yields for difficult proteins [32]. We have therefore utilized a commercially available autoinduction system to screen for hVav1-DH-PH-CRD and hVav1-CRD constructs. Almost all constructs in the closed conformation (with the N-terminal helix, starting at residue 170) yielded a sufficient quantity of soluble protein, but only under the conditions of the autoinduction media (Figure 1). The C-terminus of the CRD was mapped to amino acid 575 based on comparative analyses with the other members of truncation series using analytical SEC as a readout and length/homogeneity/stability as a criteria (data not shown). Further microexpression studies were conducted with increasing amounts of ZnSO_4 in both autoinduction and regular induction conditions. Upon supplementation with ZnSO_4 , small improvements in the yields of hVav1-DH-PH-CRD were achieved in both regular induction and autoinduction conditions, while overall protein yield was higher in the

autoinduction condition (Figure 3). Additional analysis of the soluble fraction at 500 and 1000 μM Zn^{2+} concentrations indicated that most of the Vav1-DH-PH-CRD was insoluble (data not shown). It is possible that at higher Zn^{2+} concentrations, interaction of the metal ions with additional cysteine residues in Vav1 results in a loss of solubility. The autoinduction protocol was chosen for preparative expression/purification work due to higher expression levels of recovered soluble protein, and the robustness of expression levels from batch to batch. In the case of the GST-hVav1-CRD fusion, higher yields were obtained using the regular induction protocol in the presence of 500 μM ZnSO_4 (data not shown).

Similar to the isolated hVav1-DH constructs, the open conformation construct of hVav1-DH-PH-CRD yielded minimal amounts of soluble protein. However, unlike the comparisons involving the isolated hVav1-DH (see above), the overall expression level of the open conformation constructs was comparable to Vav DH-PH-CRD in the closed conformation (data not shown). The expression analysis of Vav1-DH-PH-CRD indicated that the removal of the N-terminal inhibitory helix results in significant losses in protein solubility, which is likely related to the exposure of the hydrophobic surface of the DH domain involved in guanine nucleotide exchange with the solvent.

These results suggested that a co-expression approach would be a viable option for the generation of a hVav DH-PH-CRD/Rac1 complex. We have attempted a variety of formats for hVav1-DH-PH-CRD coexpression with GST-Rac1 and were unsuccessful in identifying appropriate conditions for the generation of complex Vav1-Rac1 with proper stoichiometry (S. Foster and A. Brooun, unpublished data). In general co-expression is a viable strategy in many cases for generation of protein complexes for structural studies, but the limitation of the co-expression strategy is the lack of stable individual components required for biochemical and biophysical characterization of a protein-protein interaction as well as screening for small molecule inhibitors of those interactions.

Analysis of hVav1 binding to Rac1 (GST pull down and SEC analysis)

Structural characterization of protein-protein interactions requires the isolation of stable and homogenous protein complexes. With the suite of expressed and purified Vav1 domains available, we have examined our constructs in terms of their ability to form a complex with its cognate GTPase Rac1. Unlike previous studies, we were compelled to use more direct methods to assess complex formation, such as gel visualization, and analytical SEC, to determine complex stoichiometry. Previous pull-down studies, where binding was assessed by Western blotting of Rac1 bound GEF factor, have yet to provide a quantitative assessment of Vav1/Rac1 complex stoichiometry [11].

We have purified mVav-DH (181-375), DH (170-375), DH-PH-CRD (170-575), and CRD (501-575) domains to homogeneity and used a GST-Rac1 fusion for pull down experiments. The results of the pull down experiments indicate that only Vav1-DH-PH-CRD effectively binds to GST-Rac1 in a quantitative fashion (Fig. 4A), while the isolated Vav1-CRD domain shows minimal detectable binding to Rac1 (Fig. 4B). These results confirm previous findings that the Vav1-CRD domain is critical for Vav1 binding to Rac1, but its role is pronounced in the context of the Vav1-DH-PH-CRD (Fig. 4B). In addition, the isolated Vav1-DH domain had no detectable binding to a GST-Rac1 fusion under our experimental conditions (Fig. 4C). The additional bands detected in the lanes containing GST eluate represent contaminants that co-purify with GST-Rac1 from *E. coli* cells or degradation products of the GST-Rac1 fusion.

While the *in vitro* pull down assay data with various Vav1 domains confirms previous mutational analysis and binding data that the CRD is indeed important for Rac1 complex formation, the notion that its functionality is more potent in association with the DH-PH module is intriguing. To further verify the results of the GST-Rac1 pull down experiments are not

associated with the properties of the isolated CRD, we determined the zinc content for the isolated Vav1-CRD sample using ICP-MS analysis. Three independent measurements of zinc content at various Vav1-CRD concentrations resulted in a zinc to Vav1-CRD ratio of 1.8, close to the predicted ratio of 2 based on similarity with the structurally characterized Raf1-CRD. We have also analyzed folding of the isolated Vav1-CRD using one dimensional (1D) ^1H -NMR analysis which has been successfully used for the prediction of protein quality for crystallization [33]. The primary characteristic indicating a folded globular domain is the high level of ^1H chemical shift dispersion in the regions of the spectrum corresponding to methyl protons (-1 to 1.5 ppm), α -protons (3.5–6 ppm) and amide protons (6–10 ppm). The significant chemical shift dispersion in these areas, as is seen in Figure 5, results from the wide variety of different structural environments within the tertiary structure of a folded, globular protein, and is not observed in unfolded or ‘random coil’ polypeptide chains (Figure 5) [33].

Previous studies of the hVav1-CRD interaction with Rac1 indicated a number of residues in Rac1 which undergo chemical shift upon addition of MBP-Vav1-CRD or MBP-Vav1-PH-CRD in NMR experiments [17]. However, we were unable to isolate a stable hVav1-CRD/Rac1 complex. Also MBP fusion of Vav3-CRD was shown to form a complex with RhoA that was detected from Western blotting analysis, however such analysis is not sufficient to determine complex stoichiometry. While additional factors may affect the interpretation of Vav1-CRD binding data to Rac1, it is possible that the hVav1-CRD fused to MBP has more structural constraints and similarities with binding properties of the hVav1-DH-PH-CRD. Similar observations were made with binding studies of the Raf1-CRD to h-Ras, where full N-terminal Raf1 construct (1-330) had higher binding affinity to h-Ras than isolated CRD [34].

Purification of Vav1/Rac1 complex for structural studies

In vitro binding experiments clearly indicate that the hVav1-DH-PH-CRD (170-575) construct is capable of forming a complex with Rac1 with proper stoichiometry. We have developed a purification procedure for the isolation of the hVav1-DH-PH-CRD/Rac1 complex using a combination of IMAC and SEC chromatography. The stability of the CRD domain was addressed by adding relatively high concentrations of the reducing agents DTT or TCEP (10 mM and 1 mM, respectively). A number of lysis and complex formation buffers were screened varying both the NaCl concentration and the pH. Our studies indicate the optimal formulation for complex formation to be 25 mM Hepes, pH 7.2, 100 mM NaCl, and 10 mM DTT. Following isolation of the hVav1-DH-PH-CRD/Rac1 complex by preparative SEC, we were unable to concentrate the protein complex to concentrations above 1 mg/ml without significant loss of material. The remaining small fraction of the soluble material that reached a concentration range of 4mg/mL was subjected to crystallization trials. A heavy precipitate was observed in almost every well of a sparse matrix screen. Supplementation of either detergents or ligands to the delivery buffers also failed to promote complex formation at higher concentrations. We believe that the autoinhibitory conformation of the DH domain with the N-terminal helix on (“closed conformation”) prevents formation of the second binding interface between the GEF and GTPase and thus impairs formation of the protein complex suitable for structural studies.

Limited proteolysis has previously been described as a valuable tool to remove disordered or flexible regions in purified proteins [35]. A panel of three proteases (trypsin, V8, and ArgC) was chosen for limited proteolysis studies. The degree of cleavage of the hVav1-DH-PH-CRD/Rac1 complex by limited proteolysis was assessed at various time points by gel electrophoresis and MS analysis. ArgC had no detectable cleavage in either of the proteins, while a number of smaller fragments were detected upon treatment with trypsin. However, V8 protease produced a distinct hVav1-DH-PH-CRD species with a molecular weight of 46,960, which corresponds to the MW of Vav1 with the N-terminal residue of 176 (Fig. 1B,). No additional fragments of

Vav1 or Rac1 were detected as a result of limited proteolysis with the V8 protease. Limited proteolysis analysis of isolated hVav1-DH and hVav1-DH-PH-CRD identified the same cleavage site, however the reaction did not proceed to completion, likely because Rac1 binding caused a conformation change in Vav1, rendering the N-terminal helix more accessible to the protease. Also, the overall stability of isolated hVav1-DH-PH-CRD is quite low, and resulted in a 50% protein loss compared to hVav1-DH-PH-CRD in complex with Rac1. The isolated cleaved fragment was recovered, and its N-terminal sequence was determined to correspond to amino acids 176-182 of Vav1. V8 protease cleaves following a glutamic acid residue at neutral pH, which corresponds to residue 175 in Vav1, located C-terminal to the regulatory Tyr174.

Based on the limited proteolysis data we modified our large scale purification to include a V8 cleavage step. The simultaneous cleavage with Tev protease and V8 protease has resulted in significant protein loss, but sequential cleavage with Tev and V8 of the Vav1-DH-PH-CRD/Rac1 complex had little impact on overall protein yields. Unlike the hVav1-DH-PH-CRD (170-575) fragment, following preparative SEC, V8 treated Vav1 concentrated well in complex with Rac1 (final concentrations 7-10 mg/ml). V8 protease cleavage clips in the middle of the N-terminal autoinhibitory helix, likely due to structural disorder in this region, and results in a homogeneous population of hVav1-DH-PH-CRD. The activity of V8 treated material in a GEF functional assay is also higher, which confirms our assumption that the hVav1-DH-PH-CRD exists in an open conformation (lacking a structured N-terminal auto inhibitory helix) following removal of residues 176-182. However, these residues are required during expression, as constructs starting at residue 176 yielded high levels of expressed material but minimal amounts of soluble protein in our hands.

GEF assay analysis

Vav1 expression is restricted to hematopoietic cells, making Vav1/Rac1 an attractive target for the development of novel immunosuppressants [5]. One of the goals of this study was to generate a minimal functional unit of hVav1 for the identification of small molecule modulators of the Vav1/Rac1 protein-protein interaction. Previous reports indicated that a 50:1 molar ratio of the DH domain to Rac1 was required for its GEF activity [7]. The activity of hVav1-DH-PH-CRD module was not previously examined in GEF experiments. We have conducted a systematic comparative analysis of various Vav1 domains in GEF assay and compared the relative activity of the fragments to their ability to form a complex with Rac1 *in vitro*.

First, we have examined the activity of the hVav1 DH domains in the “closed” and “open” conformation and compared their activities in the GEF assay. These experiments have confirmed that the DH domain in an “open” conformation has slightly higher activity in the GEF assay than in the “closed” conformation (Fig. 6A). Small changes in fluorescence were detected at a 1:1 molar ratio between “open” mVav and Rac1 (not shown), while a signal with good dynamic range was observed at 10:1 molar ratio (Fig. 6A). It is also apparent that the “closed” conformation exists in equilibrium with the “open” conformation, as it has activity in the GEF assay at a molar ratio 5-fold higher than the closed conformation. This is an important observation in the context of any structural characterization of hVav1, as such variants of the DH domain would lead to conformation heterogeneity. The “closed” conformation of the hVav1-DH domain was also subjected to limited proteolysis with V8 protease, leading to material that was comparably active to mVav-DH (data not shown). These data provides experimental evidence supporting the idea that treatment with V8 protease results in cleavage of the region containing Tyr174, and converts hVav1 into fully active protein.

The activity of “open and “closed” conformation constructs in the context of the hVav1-DH-PH-CRD module were also examined in a GEF assay. Similar to the DH domain alone, the hVav1-DH-PH-CRD constructs that were subjected to limited proteolysis with V8 protease

had higher activity than the “closed” conformation, and activity was detected at a substoichiometric molar ratio 1:5 of GEF to GTPase (Fig. 6B). hVav1-DH-PH-CRD (190-575) M351T was also found to be active in GEF assay up to a 1:20 molar ratio between GEF and GTPase (Fig. 7). To further evaluate the ability of Rac1 to bind to the isolated hVav1-CRD, we have performed titration experiments with isolated CRD and found that it had no effect on the GEF activity of hVav1-DH-PH-CRD. The activity data generated for various hVav1-DH-PH-CRD constructs supports the *in vitro* binding data since the molar ratios required for protein activity in GEF assay correspond to their ability to bind GST-Rac1 directly. hVav1-DH-PH-CRD (190-575) M351T maintains good stability at 4° C over a 5-7 day time period as indicated by analytical SEC and GEF assays. The stability data and activity data on Vav1-DH-PH-CRD (190-575) M351T indicates that this particular construct is a good tool for use in assay development protocols to identify antagonists of the Vav1/Rac1 protein-protein interactions.

Crystallization of Vav1/Rac1 complex

Crystallization of three different Vav1/Rac1 complexes was systematically explored using nanovolume crystallization technology. In addition to the Vav1/Rac1 complex generated by V8 limited proteolysis, we have attempted crystallization of Vav1-DH-PH-CRD (176-575) M351T and hVav1-DH-PH-CRD (190-575) M351T. Crystallization trials of Vav1/Rac1 complexes generated using hVav1-DH-PH-CRD (176-575) M351T and hVav1-DH-PH-CRD (190-575) M351T yielded no crystal hits. Rod-shaped crystals were obtained for both wild type and selenomethionine derivatized (Vav1) complex in a variety of precipitants from Qiagen PEG and PACT course screens, and optimized by fine screening around favorable precipitant conditions (0.1 M Hepes, pH 7.5, 17% PEG-3350, 0.2 M ammonium Chloride or 0.1 M HEPES, pH 7.5, 17% PEG-6000, 0.2 M LiCl₂). Vav1/Rac1 crystals belonged to the monoclinic P2₁ spacegroup with unit cell dimensions a=78.35, b= 74.58, c= 115.09 Å; $\alpha=\gamma=90^\circ$, $\beta=103.75^\circ$. The calculated Matthews coefficient (V_M) suggests that this protein-protein interaction exists as a heterotetramer (or a dimer of dimers) in the unit cell.

Conclusions

Structural determination of complex eukaryotic targets and their complexes with biologically relevant partners is one of the major challenges for modern structural proteomics. Even with the development of high-throughput structural proteomics strategies, one of the key questions during such efforts is the decision to discontinue efforts to pursue a three dimensional structure of a selected target. How does one determine that additional experiments will not generate desired structural information? As the complexity of the protein targets that are being pursued by structural biologists is gradually increasing, the number of variables that need to be tested at various stages of a structural proteomics pipeline becomes disproportionally larger. High-throughput strategies were developed to address the bottlenecks in protein expression, purification and crystallization, but their effectiveness could be substantially augmented by designing experiments using existing knowledge with regards to protein target function.

We have successfully applied a number of sequential structural proteomics strategies to the expression, purification and crystallization of the Vav1/Rac1 complex (Table 3). Such a sequential approach provides a blueprint for processing complex protein targets, and illustrates how an outcome at each stage of a decision making process defines the next step. In the case of Vav1 we have used iterative cycles of protein engineering, microexpression and protein characterization to arrive at a biologically relevant protein construct and protein conformation. At that stage both the N- and C-termini were mapped to within five amino acids each, which has resulted in a stable/minimal construct of hVav1-DH-PH-CRD, likely due to better folding of the CRD domain. To better understand how the Vav1 protein has been affected by protein engineering efforts, we have also systematically characterized the key variant in a functional

assay. This was followed by the identification of protein complex formation conditions and limited proteolysis studies in the presence of binding partner to remove a flexible region at the N-terminus to allow crystallization of Vav1/Rac1 complex. Detailed characterization of Vav1 domains in the functional assay also provides an entry point for screening of inhibitors of Vav1/Rac1 protein-protein interaction.

Acknowledgements

The authors would like to gratefully acknowledge Ray Stevens for guidance and attention to the project. The authors are grateful to Margaret Johnson for NMR analysis. We also would like to thank Mike Hanson and Michelle Kraus for insightful discussions. This study was supported by SFP 1543 and NIH Protein Structure Initiative Specialized Centers Grant GM074961. TSRI manuscript number: 18391.

References

1. Rossman KL, Der CJ, Sondek J. GEF means go: turning on RHO GTPases with guanine nucleotide-exchange factors. *Nat. Rev. Mol. Cell. Biol* 2005;6:167–80. [PubMed: 15688002]
2. Kusama T, Mukai M, Endo H, Ishikawa O, Tatsuta M, Nakamura H, Inoue M. Inactivation of Rho GTPases by p190 RhoGAP reduces human pancreatic cancer cell invasion and metastasis. *Cancer Sci* 2006;97:848–53. [PubMed: 16776779]
3. Katzav S, Martin-Zanca D, Barbacid M. Vav, a novel human oncogene derived from a locus ubiquitously expressed in hematopoietic cells. *EMBO J* 1989;8:2283–90. [PubMed: 2477241]
4. Movilla N, Bustelo XR. Biological and regulatory properties of Vav-3, a new member of the Vav family of oncoproteins. *Mol. Cell. Biol* 1999;19:7870–85. [PubMed: 10523675]
5. Tybulewicz VL. Vav-family proteins in T-cell signaling. *Curr. Opin. Immunol* 2005;17:267–74. [PubMed: 15886116]
6. Bustelo XR. Vav proteins, adaptors and cell signaling. *Oncogene* 2001;20:6372–81. [PubMed: 11607839]
7. Aghazadeh B, Lowry WE, Huang XY, Rosen MK. Structural basis for relief of autoinhibition of the Dbl homology domain of proto-oncogene Vav by tyrosine phosphorylation. *Cell* 2000;102:625–33. [PubMed: 11007481]
8. Llorca O, Arias-Palomo E, Zugaza JL, Bustelo XR. Global conformational rearrangements during the activation of the GDP/GTP exchange factor Vav3. *EMBO J* 2005;24:1330–40. [PubMed: 15775967]
9. Amarasinghe GK, Rosen MK. Acidic region tyrosines provide access points for allosteric activation of the autoinhibited Vav1 Dbl homology domain. *Biochemistry* 2005;44:15257–68. [PubMed: 16285729]
10. Han J, Das B, Wei W, Van Aelst L, Mosteller RD, Khosravi-Far R, Westwick JK, Der CJ, Broek D. Lck regulates Vav activation of members of the Rho family of GTPases. *Mol. Cell. Biol* 1997;17:1346–53. [PubMed: 9032261]
11. Zugaza JL, Lopez-Lago MA, Caloca MJ, Dosil M, Movilla N, Bustelo XR. Structural determinants for the biological activity of Vav proteins. *J. Biol. Chem* 2002;277:45377–92. [PubMed: 12228230]
12. Booden MA, Campbell SL, Der CJ. Critical but distinct roles for the pleckstrin homology and cysteine-rich domains as positive modulators of Vav2 signaling and transformation. *Mol. Cell. Biol* 2002;22:2487–97. [PubMed: 11909943]
13. Palmby TR, Abe K, Der CJ. Critical role of the pleckstrin homology and cysteine-rich domains in Vav signaling and transforming activity. *J Biol. Chem* 2002;277:39350–9. [PubMed: 12177050]
14. Derewenda U, Oleksy A, Stevenson AS, Korczynska J, Dauter Z, Somlyo AP, Otlewski J, Somlyo AV, Derewenda ZS. The crystal structure of RhoA in complex with the DH/PH fragment of PDZ-RhoGEF, an activator of the Ca(2+) sensitization pathway in smooth muscle. *Structure* 2004;12:1955–65. [PubMed: 15530360]
15. Snyder JT, Worthylake DK, Rossman KL, Betts L, Pruitt WM, Siderovski DP, Der CJ, Sondek J. Structural basis for the selective activation of Rho GTPases by Dbl exchange factors. *Nat. Struct. Biol* 2002;9:468–75. [PubMed: 12006984]
16. Worthylake DK, Rossman KL, Sondek J. Crystal structure of Rac1 in complex with the guanine nucleotide exchange region of Tiam1. *Nature* 2000;408:682–8. [PubMed: 11130063]

17. Heo J, Thapar R, Campbell SL. Recognition and activation of Rho GTPases by Vav1 and Vav2 guanine nucleotide exchange factors. *Biochemistry* 2005;44:6573–85. [PubMed: 15850391]
18. Drugan JK, Khosravi-Far R, White MA, Der CJ, Sung YJ, Hwang YW, Campbell SL. Ras interaction with two distinct binding domains in Raf-1 may be required for Ras transformation. *J Biol. Chem* 1996;271:233–7. [PubMed: 8550565]
19. Hommel U, Zurini M, Luyten M. Solution structure of a cysteine rich domain of rat protein kinase C. *Nat. Struct. Biol* 1994;1:383–7. [PubMed: 7664052]
20. Gao Y, Dickerson JB, Guo F, Zheng J, Zheng Y. Rational design and characterization of a Rac GTPase-specific small molecule inhibitor. *Proc. Natl. Acad. Sci. U S A* 2004;101:7618–23. [PubMed: 15128949]
21. Nassar N, Cancelas J, Zheng J, Williams DA, Zheng Y. Structure-function based design of small molecule inhibitors targeting rho family GTPases. *Curr. Top. Med. Chem* 2006;6:1109–16. [PubMed: 16842149]
22. Poppe D, Tiede I, Fritz G, Becker C, Bartsch B, Wirtz S, Strand D, Tanaka S, Galle PR, Bustelo XR, Neurath MF. Azathioprine suppresses ezrin-radixin-moesin-dependent T cell-APC conjugation through inhibition of Vav guanosine exchange activity on Rac proteins. *J Immunol* 2006;176:640–51. [PubMed: 16365460]
23. Fernandez-Zapico ME, Gonzalez-Paz NC, Weiss E, Savoy DN, Molina JR, Fonseca R, Smyrk TC, Chari ST, Urrutia R, Billadeau DD. Ectopic expression of VAV1 reveals an unexpected role in pancreatic cancer tumorigenesis. *Cancer Cell* 2005;7:39–49. [PubMed: 15652748]
24. Scapin G. Structural biology and drug discovery. *Curr. Pharm. Des* 2006;12:2087–97. [PubMed: 16796557]
25. Arkin M. Protein-protein interactions and cancer: small molecules going in for the kill. *Curr Opin Chem Biol* 2005;9:317–24. [PubMed: 15939335]
26. Kuhn P, Wilson K, Patch MG, Stevens RC. The genesis of high-throughput structure-based drug discovery using protein crystallography. *Curr. Opin. Chem. Biol* 2002;6:704–10. [PubMed: 12413557]
27. Movilla N, Dosil M, Zheng Y, Bustelo XR. How Vav proteins discriminate the GTPases Rac1 and RhoA from Cdc42. *Oncogene* 2001;20:8057–65. [PubMed: 11781818]
28. Otwinowski, HKL. 2000.
29. Page R, Moy K, Sims EC, Velasquez J, McManus B, Grittini C, Clayton TL, Stevens RC. Scalable high-throughput micro-expression device for recombinant proteins. *Biotechniques* 2004;37:364, 366, 368. [PubMed: 15470889]
30. Nomine Y, Ristriani T, Laurent C, Lefevre JF, Weiss E, Trave G. Formation of soluble inclusion bodies by HPV E6 oncoprotein fused to maltose-binding protein. *Protein. Expr. Purif* 2001;23:22–32. [PubMed: 11570842]
31. Matthews JM, Kowalski K, Liew CK, Sharpe BK, Fox AH, Crossley M, MacKay JP. A class of zinc fingers involved in protein-protein interactions biophysical characterization of CCHC fingers from fog and U-shaped. *Eur. J Biochem* 2000;267:1030–8. [PubMed: 10672011]
32. Studier FW. Protein production by auto-induction in high density shaking cultures. *Protein. Expr. Purif* 2005;41:207–34. [PubMed: 15915565]
33. Page R, Peti W, Wilson IA, Stevens RC, Wuthrich K. NMR screening and crystal quality of bacterially expressed prokaryotic and eukaryotic proteins in a structural genomics pipeline. *Proc. Natl. Acad. Sci. U S A* 2005;102:1901–5. [PubMed: 15677718]
34. Ghosh S, Bell RM. Identification of discrete segments of human Raf-1 kinase critical for high affinity binding to Ha-Ras. *J Biol. Chem* 1994;269:30785–8. [PubMed: 7983008]
35. Cheetham GM, Knegt RM, Coll JT, Renwick SB, Swenson L, Weber P, Lippke JA, Austen DA. Crystal structure of aurora-2, an oncogenic serine/threonine kinase. *J Biol. Chem* 2002;277:42419–22. [PubMed: 12237287]

Figure 1.

A) Vav1 domain boundaries. Vav1 consists of several domains, including a calmodulin homology domain (CH, 1-116), an acidic region (Ac, 132-176), a Dbl homology domain (DH; 185-375), a pleckstrin homology domain (PH, 398-508), a unique cysteine rich domain (CRD, 516-565), a short proline rich region (607-610), and an SH2-SH3-SH2 domain (612-844). B) Key elements of Vav1-DH-PH-CRD unit as related to hVav1 construct used for crystallization of Vav1-Rac1 complex. Y174 is regulatory tyrosine residue that controls opening/closure of DH domain. V8 protease cleaves after residue E175. Amino acids 170, 180 and 190 are starting points for the majority of Vav1 constructs in this study. Constructs starting from residue 170 are considered to be in the “closed conformation”, while constructs starting from residues 180 or 190 are considered to be in the “open conformation” and refers to the presence or absence of a N-terminal autoinhibitory helix.

Figure 2.

Microexpression analysis of hVav DH to mVav DH mutants (whole cell lysates and IMAC elutes). 1) hVav 181-375 P182S, S184P, M185T, 2) hVav 181-375 I236T, 3) hVav 181-375 L245F, R246S, 4) hVav 181-375 181-375 T260G, 5) hVav 181-375A264T, N265T, 6) hVav 181-375 R296Q, A299T, 7) hVav 181-375 M351T, 8) hVav 181-375 WT, and 9) mVav 181-375 WT.

Figure 3.

Zn²⁺ titrations supplemented to both autoinduction and regular induction conditions for the fermentation of hVav DH-PH-CRD (170-575). A) Autoinduction protocol: 1-4 - whole cell extract in the presence of 0, 100, 500, 1000 μM Zn²⁺; 4-8 – IMAC elutes in the presence of 0, 100, 500, 1000 μM Zn²⁺. In the case of autoinduction protocol ZnCl₂ was added prior to shifting the temperature to 25 °C. B) Regular induction protocol: 1-4 - whole cell extract in the presence of 0, 100, 500, 1000 μM Zn²⁺; 4-8 – IMAC elutes in the presence of 0, 100, 500, 1000 μM Zn²⁺.

Figure 4.

In vitro pull-down experiments using GST-Rac1 and (a) hVav-CRD, (b) hVav DH-PH-CRD, and (c) mVav DH. 1) resin control load, 2) resin control wash, 3) resin control GSH elute, 4) experimental load, 5) experimental final wash, and 6) experimental GSH elute.

Figure 5.

1D ¹H NMR spectrum of Vav1-CRD at 400 μ M in PBS (pH 7.4) with 11 mM TCEP. The spectrum was recorded at 12 $^{\circ}$ C, with 256 scans and with water suppression by presaturation, on a Bruker Avance 600 spectrometer equipped with a TXI HCN z-gradient probe

Figure 6.

(a) Guanine nucleotide exchange assay analysis of "open" and "closed" conformations of hVav DH and "open" conformation of mVav DH at various molar ratios between Vav and Rac1. (b) Guanine nucleotide exchange assay analysis of "open" and "closed" conformations of hVav DH-PH-CRD at various molar ratios between Vav1 and Rac1. Domain boundaries for mVav1 and hVav1 are as indicated in the figure.

Figure 7.
Guanine nucleotide exchange assay analysis of Vav1-DH-PH-CRD (190-575) M351T at various molar ratios between Vav1 and Rac1.

Table 1

Small scale characterization of human Vav1 constructs.

Domain	Boundary	Small scale purification score [*]	Analytical SEC
DH	170-375	4	monomer
DH	181-375	1	NA ^{**}
DH	181-375 M351T	4	Monomer
DH-PH	170-527	0	NA
DH-PH	181-527	0	NA
DH-PH-CRD	170-567	2	Aggregate
DH-PH-CRD	170-575	2	Monomer/aggregate
DH-PH-CRD	170-608	2	Monomer/aggregate
DH-PH-CRD	190-575	1	N/A
DH-PH-CRD	190-575 M351T	3	Monomer

* 0- not soluble; 4- highly soluble (based on the amount of protein in IMAC elute).

** NA- not applicable (sufficient amount of soluble protein was not recovered for analytical).

Table 2
Purification of hVav1 DH-PH-CRD in complex with Rac1

Purification Step	Total Protein (mg)	Relative Purity	Percent Loss
Binding Experiment	52	85	
Tev Protease Cleavage	48	90	8%
V8 Protease Cleavage	44	90	15%
Pooled SEC fractions for Crystallization [*]	6	>99	88.8%

^{*}
- 2 of total 6 peak fractions carried further and pooled at this step.

Table 3

Remedial strategies in structural proteomics applied to expression, purification and crystallization of Vav1/Rac1 complex.

Structural proteomics strategy	Number of variables explored	Outcome for Vav1/Rac1
Construct optimization	~200 clones generated for Vav-DH, Vav-AC-DH-PH-CRD, Vav-DH-PH-CRD, Vav-PH-CRD and Vav-CRD	Soluble protein generated for all domain combinations with exception of DH-PH
Expression optimization	3 microexpression protocols tested	Autoinduction protocol with Zn supplementation increased yield of hVav1-DH-PH-CRD
Expression optimization by homology based protein engineering	7 surface mutants generated for both DH domain and DH-PH-CRD	Human to mouse mutagenesis significantly improves yields of "open" hVav-DH
Formulation of protein for stability	Formulation screen with 12 buffer conditions at different ionic strength, pH and detergent concentrations to identify conditions for lysis/complex formation/co-concentration	Optimized conditions for hVav1 purification and hVav1/Rac1 complex formation
Limited proteolysis	Panel of 3 proteases applied to hVav1-DH-PH-CRD/Rac1 complex	Generated homogenous/active conformation of hVav1-DH domain suitable for structural studies
Protein crystallization	3 versions of Vav1/Rac1 complex	Single variant was crystallized