Dan Piraner is a senior majoring in biology (with a specialization in biochemistry) and chemistry. He began his undergraduate research as a freshman in Dr. Christine Hrycyna’s lab under the tutelage of Dr. Amy Griggs, and joined Dr. Carol Post’s lab the following summer when he undertook the project of elucidating the Vav1-SH2 structure. Under the guidance of Dr. Chih-Hong Chen, Dan learned about the concepts and methods of NMR and contributed to the NMR data analysis of the project. After graduating from Purdue, Dan hopes to pursue a Ph.D. in biochemistry or biophysics, specifically to study protein design and engineering. In addition to academics and undergraduate research, Dan participates in the Purdue Aikiju jutsu Club and the Purdue Thai Boxing Club.

Dr. Carol Post is Professor of Medicinal Chemistry and Molecular Pharmacology at Purdue University. Her research focuses on understanding the molecular behavior of proteins, particularly their role in misregulation of the immune system and in cancer. Her research group uses nuclear magnetic resonance spectroscopy and computational methods to study certain protein-protein interactions. The purpose of this research is to design new inhibitors of signaling proteins that might be useful in the development of drug compounds. Additionally, Dr. Post engages in structural and computational studies of viral proteins that cause human diseases ranging from the common cold to AIDS. These studies aid in the discovery and exploration of anti-viral drugs. She received the Lion’s Club Award for Cancer Research in 2009.

Protein signaling is the key method by which cells recognize a stimulus from their environment and convert it into a response. Signaling occurs in many forms: hormones, growth factors, and even proteins may act as signals from the environment. The response to their detection must be carried from the cell surface, where the signal is detected, to the nucleus, where the cell alters its DNA expression. This study analyzes one component in the signaling pathway of Spleen Tyrosine Kinase (Syk). The Syk protein receives a signal from B-cell receptors and amplifies it, resulting in an adaptive immune response and the production of antibodies for targeting foreign molecules. Strikingly, constant amplification of the Syk signal can transform B-cells into a cancerous phenotype and has been associated with cancers of the lymphatic system. Syk has also been implicated in autoimmune diseases such as rheumatoid arthritis. This study uses nuclear magnetic resonance spectroscopy to determine the structure of a segment from the Syk protein bound to Vav1, a downstream signaling protein that has been implicated in various cancers. The preliminary result will be refined using a molecular dynamics simulation to obtain the most energetically stable conformation of the protein-ligand interaction. The terminal results of this experiment will reveal the structural details of the Syk-Vav1 interaction, which may describe a possible therapeutic target in the treatment of proliferative B-cell disorders, as well as some types of cancer. Using this information, one or more drugs may be synthesized to block the interaction in cancerous cells, providing a therapeutic avenue to treat the disease.


**Keywords**

NMR Spectroscopy  
Vav1  
Syk  
SH2  
Rheumatoid Arthritis  
B Lymphocyte  
Protein Structure
INTRODUCTION

Purpose of study

Many serious human disorders are caused by errors in cell signaling pathways. Just as a short circuit can induce dysfunction in an electrical system, broken cellular components can cause improper behavior in a cell, which may damage the host organism. The most well-known example of such a disorder is cancer. Caused by a plethora of mutations in the large and complex cell signaling network, cancer is the uncontrolled growth and reproduction of a cell with a broken signaling pathway. The results of this disease can devastate the patient, underscoring the urgent need for continuing cancer research. This study models the interaction of proteins in one signaling junction of B-cells. Both proteins modeled have been implicated in various cancers when mutated, suggesting that a physical structure representing their bound state could shed light on a possible method of cancer treatment.

This experiment models the interaction between Spleen Tyrosine Kinase (Syk) and Vav1. Syk is a non-receptor protein tyrosine kinase (Sada, Takano, Yanagi, & Yamamura, 2001); that is, it functions to add a phosphate group to a tyrosine amino acid of another protein. Syk is activated upon stimulation of the B-cell receptor. Two major signaling outcomes are Ca$^{2+}$ release within the cell and the activation of the MAP Kinase cascade, both leading to changes in DNA expression and cell morphology (Marx, 1993).

Vav1 is a multifunctional hub protein. It has several components, including a GEF domain for protein signal activation, SRC Homology domains for mediating interactions with other proteins, and several other domains (Katzav, 2005). Vav1 interacts with Syk via its SH2 (SRC Homology 2) domain. This domain has high affinity for phosphorylated tyrosine residues. Because tyrosine kinases often autophosphorylate (Murray, Padrique, Pinko, & McTigue, 2001), kinase-substrate binding via SH2-phosphotyrosine affinity is conserved across a spectrum of cellular tyrosine kinases. Despite the fact that structures of SH2 domains are known, each such domain is specific to one or a few target proteins, requiring a model to be constructed de novo (Birge & Hanafusa, 1993). This study may therefore provide useful data about the interaction of Syk and Vav1 through Vav1’s SH2 domain, which may serve as a basis for rational drug design targeting either of the two proteins studied.

The role of Syk kinase in cell signaling

Syk is a non-receptor protein kinase whose activity is essential to the clonal expansion of B-cells (Autoimmune Diseases, 2010). A crucial immune component, B-cells are responsible for producing antibodies. Activation of the B-cell receptor by an antigen initiates a signaling sequence in which Syk binds to the ITAM protein, which induces Syk autophosphorylation. Many downstream proteins in the Syk signal are directly implicated in cellular proliferation, such as Cbl, PLC-γ, PI 3-K, MAP Kinase, and Vav1. It is therefore unsurprising that Syk regulation is essential for normal cell development and transformation into cancer.

Although most effector proteins in the Syk pathway induce cellular proliferation, they differ in their modes of action. While inhibition of Syk greatly impedes B-cell...
development, disruption of a Syk-substrate interaction could have a pathway-specific result. Therefore, the structural determination of a Syk-substrate interface could provide the necessary information for the synthesis of a drug that would normalize cell function. The development of such drugs could aid in efforts to combat autoimmune diseases such as rheumatoid arthritis.

The polyfunctional protooncogene product Vav1

Vav1 is a hub protein containing multiple domains, each with a purpose in signaling. The N-terminal region contains a CH domain that is used to bind to the cellular scaffold in other proteins. In Vav1, however, it may regulate Ca$^{2+}$ release (Billadeau & Turner, 2002). The Vav1 DH domain is the active site; its GEF activity exchanges GDP for GTP in a class of proteins known as GTPases, thereby activating them. Neighboring regions in the Vav1 protein serve to bind GEF targets to facilitate their activation. The remaining region contains a set of conserved sequences called SH (SRC Homology) domains, which bind specific protein regions. Vav1 contains two SH3 domains, which bind to proline-rich motifs. Vav1 binds to Syk via the SH2 domain residing between the two SH3 regions.

Vav1 has been implicated in cancers via incorrect localization, overexpression, and overactivation (Billadeau, 2010). Research has shown that expression of Vav1 in pancreatic cells results in tumorigenesis in this organ (Fernandez-Zapico, et al., 2005). Similarly, expression of Vav1 in breast tissue may contribute to tumor formation (Lin & Van Golen, 2004). Vav1 is activated through phosphorylation by Syk (Deckert, 1996). Structural study of the Syk-Vav1 complex may lead to rationally designed inhibitors functioning to prevent Vav1 activation in cells overexpressing or constitutively activating this protein.

The role of NMR spectroscopy in structural studies

NMR, or Nuclear Magnetic Resonance spectroscopy, uses the magnetic properties of some atomic nuclei (mainly $^1$H, $^{19}$F, $^{13}$C, $^{15}$N) to provide information about their through-bond or through-space connectivity (NMR Spectroscopy, n.d.). The nuclei involved in these interactions are characterized by their nonzero spin quantum number and associated nuclear magnetic moment. In the presence of an external magnetic field, the ensemble of these nuclei within the sample generates a net magnetic moment, also called magnetization, which aligns parallel to the applied field. This magnetization can be pictured as a vector whose length is governed by the Boltzmann distribution law applied to nuclear magnetic moments. A short electromagnetic pulse at radiofrequency (RF pulse) perpendicular to the main magnetic field deflects the magnetization vector from its original direction, causing it to rotate. Information about the structure is gathered from the differing precession speeds of the magnetization vectors belonging to the different types of nuclei, or to the same nucleus but in different chemical environments. A Fourier transform of the signal induced in the receiver coil of the spectrometer by the magnetization allows the identification of a characteristic precession frequency for each nucleus in the protein.

The utility of the nuclear resonance for chemical and biological applications results from the microenvironment of the atom. As a magnetic field is imposed upon the electrons surrounding the nucleus, they redistribute and generate their own field in the opposite direction of the applied field. This electron field lowers the strength of the total magnetic field acting on the nucleus. This process, known as shielding, causes distinct resonant frequencies of most atoms in a molecule. Increased shielding decreases the necessary energy to flip the nuclear orientation in the magnetic field, which lowers the resonance frequency of the nucleus. A variety of NMR experiments have been devised to aid in assigning each resonance frequency to a specific nucleus in a protein. Furthermore, expected resonance frequencies for each atom across all amino acids that make up proteins are available in an online database.

Determining the three-dimensional structure of a large protein molecule (>1,000 atoms) is challenging. The crucial piece of information that NMR is able to resolve is the Nuclear Overhauser Effect (NOE), which is a change in signal intensity caused by the magnetic interaction of two closely spaced nuclei (Gemmecker, 1999). The NOE signal depends on distance between two protons ($r$), specifically on $r^{-6}$, and is a through-space—as opposed to a through-bond—interaction (Wüthrich, 2002). Because the effect occurs at the resonance frequency of both interacting nuclei, it can be correlated to known signals. These data enable the spectroscopist to obtain distance restraints between many pairs of protons within the protein.

While the distance restraints severely limit the number of possible conformations of the polypeptide chain, giving it a general shape, the exact conformation of the protein cannot be obtained from NMR data alone. Rather, simulations utilizing molecular mechanics calculate possible conformations consistent with the distance restraint data and energy considerations. Generally, the correct protein structure has the lowest free energy and is therefore in the most thermodynamically stable state. However, other methods have been devised that combine energy considerations with other relevant factors to predict the most likely structure (Güntert, 2004).
MATERIALS AND METHODS

Protein isolation

Recombinant *E. coli* expressing GST-tagged Vav1 SH2 were cultured in vitro. The bacteria were then lysed using a French pressure cell press (French press) and centrifuged to separate nuclear and cytoplasmic components. The resulting homogenate was filtered and purified using an ÄKTA (GE Healthcare) HPLC platform equipped with a GST-affinity column. GST was cleaved during a second elution through a His-affinity column after addition of TEV protease. Protein levels during elution were measured using the built-in optical spectrometry capability of the ÄKTA platform and were then confirmed using electrophoresis. Purified protein was transported on ice for experimentation.

NMR experimentation

A 13-residue peptide homologous in sequence to the SH2-interacting Syk loop was titrated with the purified SH2 domain. Following titration, several protein, carbon, and nitrogen NMR experiments were performed for chemical shift assignment. Three NOESY spectra were also recorded for NOE assignment. All experiments were performed on a Bruker Avance-III-800 spectrometer. A full list of experiments performed can be found in Table 1.

<table>
<thead>
<tr>
<th>Region</th>
<th>Experiment Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backbone</td>
<td>N-HSQC</td>
</tr>
<tr>
<td>Backbone</td>
<td>CBCA(CO)NH</td>
</tr>
<tr>
<td>Backbone</td>
<td>HNCACB</td>
</tr>
<tr>
<td>Backbone</td>
<td>HNCA</td>
</tr>
<tr>
<td>Backbone</td>
<td>HN(CO)CA</td>
</tr>
<tr>
<td>Aliphatic Side Chain</td>
<td>C(CO)NH-TOCSY</td>
</tr>
<tr>
<td>Aliphatic Side Chain</td>
<td>H(CCO)NH-TOCSY</td>
</tr>
<tr>
<td>Aliphatic Side Chain</td>
<td>TOCSY-HSQC</td>
</tr>
<tr>
<td>Aliphatic Side Chain</td>
<td>H(C)CH-TOCSY</td>
</tr>
<tr>
<td>Aliphatic Side Chain</td>
<td>(H)CCH-TOCSY</td>
</tr>
<tr>
<td>Aromatic Side Chain</td>
<td>HBCBCGCDHD</td>
</tr>
<tr>
<td>Aromatic Side Chain</td>
<td>HBCBCGCDCEHE</td>
</tr>
<tr>
<td>Aromatic Side Chain</td>
<td>$^{13}$C-HSQC</td>
</tr>
<tr>
<td>Peptide Backbone &amp; Side Chain</td>
<td>2D Double-Filtered TOCSY</td>
</tr>
<tr>
<td>Peptide Backbone &amp; Side Chain</td>
<td>2D Double-Filtered NOESY</td>
</tr>
<tr>
<td>Intramolecular NOE</td>
<td>$^{15}$N/$^{13}$C-NOESY-HSQC</td>
</tr>
<tr>
<td>Intramolecular NOE</td>
<td>$^{13}$C-NOESY-HSQC</td>
</tr>
<tr>
<td>Intermolecular NOE</td>
<td>3D ($^{13}$C/$^{15}$N)-filtered $^{13}$C-edited NOESY-HSQC</td>
</tr>
</tbody>
</table>

Table 1. Experiments performed on Bruker Avance-III-800 spectrometer. (Courtesy of Dr. Chih-Hong Chen.)
NOE and structure refinement

Intermolecular NOEs were picked manually out of the list provided by Piraner NMR Suite based on their conformity to the projected Syk binding site of the SH2 domain, as projected by a previous component of this study. These peaks were incorporated into CYANA calculations using the KEEP subroutine. The molecular dynamics simulator XPLOR-NIH (Schwieters, 2003) refined the CYANA structure to obtain the lowest energy conformation.

RESULTS AND DISCUSSION

Sparky data

Chemical shift assignment was initialized from the pre-assigned 2D $^{15}$N-HSQC (see Figure 1). This spectrum contained the chemical shifts of the amide nitrogen and protons for all residues except 1-3, 15, 38-45 (flexible loop), 57, 60-61, and 104. The chemical shifts of other atoms were assigned using the N-HSQC data. The chemical shift coverage of the protein as assigned in Sparky was nearly complete; according to the CYANA initialization report, 90.6% of the atoms in the protein were matched to a chemical shift.

Chemical shift determination of the peptide was far less successful. A combination of a leaky filter and poor sensitivity rendered chemical shift assignment difficult. In total, 76.8% of the peptide chemicals shifts were assigned. This poor coverage likely contributed to the lack of NOE restraints generated between the peptide and the protein.

CYANA calculations

To create a consensus protein structure, two CYANA calculations were performed. In the first, CYANA was allowed to select and drop NOE peaks at the discretion of its algorithm. In the second, the KEEP subroutine was used to preserve the 24 protein-peptide NOE assignments generated by Piraner NMR Suite. The CYANA results show a precise protein structure being determined, with a heavy atom RMSD near one Å. Table 2 displays the calculation statistics for both CYANA runs.
The first CYANA calculation resulted in a low RMSD and few distance or dihedral restraint violations. However, the resulting structure did not show any association between the protein and the peptide. Examination of the NOE peaks assigned by CYANA revealed that none of the 24 intermolecular peaks assigned by Piraner NMR Suite had been similarly labeled by the CYANA NOEassign function. This may be attributed to the small amount of such peaks. The CYANA NOEassign routine takes “network anchoring” into account. A lack of similar peaks would induce the program to treat the intermolecular peaks as noise and discard them. The second calculation has many restraint violations in every structure. However, the resulting structure may show the peptide associating with the protein. The structural violations are due to erroneous peak assignment. Because Piraner NMR Suite cannot check the validity of assignment for a structure without a template, it outputs results based solely on shift matching of the NOE peaks to the chemical shift library. Furthermore, because some peaks may correlate to the chemical shift of more than one atom, the program outputs all possible chemical shift assignments as duplicate peaks. The output peak list contains some impossible combinations, such as duplicate peaks in which one copy relates an interaction between a peptide atom and an N-terminal protein region, while the other contains the same peptide atom and a protein C-terminal atom. These discrepancies must be analyzed based on previous work with Vav1 SH2 to remove impossible distance restraints, thereby removing the cause for the multitude of restraint violations.

### Protein structure output

CYANA generates an ensemble of the 20 structures best conforming to the NMR data. Figure 2 shows the structure ensemble for the SH2 domain only. The location of the peptide is not visualized. Throughout the structure ensemble, some conserved components may be clearly seen, including the central β sheet (B) flanked by two α helices (A1 & A2) typical of SH2 domains. Other regions

---

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Manually Assigned Intermolecular NOEs</th>
<th>Automatically Assigned Intermolecular NOEs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extent of assignments</td>
<td>88.5% (of protein + peptide)</td>
<td>88.5% (of protein + peptide)</td>
</tr>
<tr>
<td>Number of distance constraints</td>
<td>1499</td>
<td>1681</td>
</tr>
<tr>
<td>Number of dihedral restraints</td>
<td>1161</td>
<td>1161</td>
</tr>
<tr>
<td>Distance constraint violations</td>
<td>125</td>
<td>3</td>
</tr>
<tr>
<td>Dihedral restraint violations</td>
<td>58</td>
<td>0</td>
</tr>
<tr>
<td>Average backbone RMSD to mean</td>
<td>1.35 +/- 0.26 Å</td>
<td>0.9 +/- 0.23 Å</td>
</tr>
<tr>
<td>Average heavy atom RMSD to mean</td>
<td>1.88 +/- 0.25 Å</td>
<td>1.41 +/- 1.21 Å</td>
</tr>
</tbody>
</table>

Table 2. Results of CYANA calculations.

---

**Figure 2.** Structure ensemble for the SH2 domain.  
**Figure 3.** Lack of binding between the protein and the peptide in the SH2 domain.  
**Figure 4.** The SH2 domain possibly bound to the Syk peptide, with restraints retained between the protein and peptide.
of interest include the small, transient flanking β sheet (C) and another α helix (D). This overall SH2 structure coincides well with that of SH2-pYpY obtained in an earlier part of this study by Dr. Chen.

Figures 3 and 4 display the same SH2 domain bound to the Syk peptide. Unlike in the previous study conducted by Dr. Chen, which modeled the interaction between Vav1 SH2 and a doubly phosphorylated pYpY peptide, this model studies the binding of a singly phosphorylated YpY chain. The phosphorylation sites are separated by three amino acids, with tyrosines 5 and 9 of the peptide phosphorylated in the pYpY structure, and only Y9 phosphorylated in the YpY-bound structure examined in the study. Due to the variability in peptide placement in the calculations, a full 20 structure ensemble is difficult to visualize. Therefore, only the first structures, which correspond to the lowest CYANA target functions, are visualized in each case.

Figure 3 clearly displays a lack of binding between the protein and the peptide. Note that due to the method of linking the peptide to the protein used in the CYANA calculation, the peptide resembles an extension of the SH2 domain. The lack of NOE restraints in this structure combined with a Van der Waals repulsion between the protein and peptide lead to a minimization of the CYANA target function when the peptide is placed far from the protein. Figure 4, in which restraints between the protein and peptide are retained, may show some interaction between the protein and peptide. However, the large number of distance and angle violations in this structure renders it only a starting point for determination of the actual structure. Manual resolution of the intermolecular distance restraints will enable a true structure to be determined. Figures 5 and 6 display the Van der Waals surface of main chain atoms of the protein and peptide with automatic assignment (Figure 5) and manual assignment (Figure 6).

Further experimentation

Clearly, correct determination of the protein-peptide structure requires manual assignment of the intermolecular NOE peaks. The current contradictory information leads to far more torsion and distance restraint violations than can be tolerated for a true protein structure. Resolution of the contradictory distance restraints will lead to the determination of the correct structure of the Vav1-Syk interaction. Refinement of the CYANA structure will be performed with XPLOR-NIH. This program will use molecular dynamics computations to minimize the energy of the structure created by CYANA. Though no drastic alterations are expected, the calculation may slightly modify some topological features of the protein, which are important for interaction with Syk and possibly the rest of the Syk-complexed “signalosome.”

Conclusion

The NMR structure determination of Vav1 SH2 bound to the YpY Syk peptide will provide information about SH2 binding behavior in this complex. This information could be used to design an inhibitor that would compete with Syk for Vav1 binding, thereby altering the signaling outcome in B-cells. The implication of constitutively
active Syk and Vav1 in cancer and autoimmune disorders suggests that such alteration of their activity could represent a possible avenue in the treatment of such diseases. Once the correct SH2-YpY structure is determined, the difference between it and the pYpY model may be used to assess the influence of phosphorylation on the structure, possibly providing information leading to the development of drugs that counteract excess signaling by these two proteins.

**References**


