ABSTRACT

Nck-1 and Nck-2 are ubiquitously expressed a Src homology 2 (SH2)- and three SH3-domain-containing adaptor proteins that interact with numerous signal transduction proteins. To elucidate the structure of SH3 domains from Ncks, we turned to use Bioinformatics method to understand the properties of each individual SH3 domains. The results showed that whole Nck1 and Nck2 share about 68% sequence identity, however, homologous domains have very strong identity: 82% for SH3-1, 80% for SH3-2, and 73% for SH3-3. Theoretical structures of SH3 domains are obtained from homology modeling. These structures showed the difference of residues distributing in the binding sites and also difference in charge distribution in the surface of each SH3 domains. Our data could be implied that each SH3 domains bind to different ligands which could be useful for further studying of ligand interaction with this domains.

Keyword: Nck, proteins, structure modeling

K. Choowongkomon : fsciktc@ku.ac.th
INTRODUCTION

SH3 domains are small proteins containing approximately 50 amino acids (Mayer, 2001, Pawson and Schlessingert, 1993). These domains play critical roles in a wide variety of biological processes including regulation of enzymes, increasing the local concentration, and mediating the assembly of large multiprotein complexes (Mayer, 2001). They have a characteristic fold of five or six $\beta$-strands arranged as two tightly packed antiparallel $\beta$-sheets. The linker regions may contain short helices (Kuriyan and Cowburn, 1997). The surface of the SH3 domain bears a relatively flat, hydrophobic ligand-binding surface, which consists of three shallow pockets defined by conserved aromatic residues. The ligand adopts a left-handed poly-proline type II helix that lies along the binding site of the SH3 domain, with its prolines interacting with the aromatic residues on the hydrophobic face of the SH3 domain.

The non-catalytic region of tyrosine kinase adaptor proteins or Ncks are a family of adaptor proteins which has no enzymatic activity (Lehmann et al., 1990). The Nck family contains two known members in humans, Nck-1/Nck-$\alpha$ and Nck-2/Nck-$\beta$, both of which are widely expressed in many tissues, and share about 68% sequence identity. Both Nck-1 and Nck-2 are cytosolic 47-kDa molecules and comprised of three Src homology 3 (SH3) domains and one Src homology 2 (SH2) domain. Ncks play important roles in cellular signaling as a physical bridge between activated cell surface receptors and various intracellular signal transduction proteins (Reviewed in McCarty, 1998). For example, The SH2 domain of Nck-2 has been shown to bind with epidermal growth factor (EGF) receptor depending on the activation state whereas all SH3 domains are shown to interact with EGF receptor regardless of EGF stimulation (Tu et al., 1998), placing the adaptor in position to interact with numerous downstream effector proteins via their SH3 domains. Many known SH3-interacting effectors interact with the actin cytoskeleton while a smaller percentage are associated with gene expression, translation, and DNA synthesis (Li et al., 2001). With the many different intracellular Nck binding targets, several functional Nck pools in the cell may allow for the diversity of interactions.

In this study, we use Bioinformatics techniques in order to understand the properties of each SH3 domains. Since several of structures of SH3 domains were solved by X-ray crystallography and NMR techniques, we can use protein modeling techniques to construct the 3D structure models of SH3 domains from Nck1 and Nck2. The sequence comparison and 3D theoretical structures were revealed differences of residues in binding sites and charge distributions in peptide binding site pockets.

METHODS
The Nck sequences were obtained from the NCBI protein database for Nck-1 and Nck-2. These proteins were identified as entries P16333 and O43639, respectively. All SH3 domains were identified by variety of methods including NCBI (http://www.ncbi.nih.gov/), ScanProsite (http://us.expasy.org/tools/scanprosite/), Motif Scan (http://myhits.isb-sib.ch/cgi-bin/motif_scan), SMART (http://smart.embl-heidelberg.de/), and CD Search (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). For each of the 6 domains, the longest sequences from the aforementioned step were submitted to Swiss-Model server (http://swissmodel.expasy.org/SWISS-MODEL.html). The sequence comparisons between SH3 domains were done by both ClustalW (http://www.ebi.ac.uk/clustalw/) and Blast 2 (http://www.ebi.ac.uk/blast2/) to report identity and similarity among 6 SH3 domains. Nck-11, Nck-12, Nck-13 are the first, second and third SH3 domains of Nck-1, respectively while Nck-21, Nck-22, Nck-23 are the first, second and third SH3 domains of Nck-2, respectively. The PDB templates for Nck SH3 modelling were generated from 18 known SH3 structures as described by (Larson and Davidson, 2000). The 18 sequences of SH3 domains were downloaded and amino acids sequences of all 18 plus 6 SH3 domans from Ncks optimally aligned using the ClustalW and manual fitting. All 18 structures were shown to have adequate identity (> 30%) among Nck SH3 domains. These optimized sequences were sent to the Swiss-model server (http://swissmodel.expasy.org/SWISS-MODEL.html) for optimized automated modeling. The structure coordinate outputs from server were further energy minimization by program GROMACS (E. Lindahl et al., 2001). To evaluate the quality of the models, the six minimized models were analyzed by program PROCHECK (Laskowski et al., 1993) and VADAR (Willard et al., 2003) and visualized by either SPDBV (Kaplan and Littlejohn, 2001) or PyMol (DeLano, 2002).

RESULTS

The sequences of whole Nck-1 and Nck-2 revealed that the two proteins are 68% identical and 79% similar in respect to individual residue comparisons (Table 1). The intra-Nck and inter-Nck domain comparisons revealed that homologous domains had very strong identity: 82% for SH3-1, 80% for SH3-2, and 73% for SH3-3. Similarities between the same domain types gave values of 95%, 92%, and 82%, respectively, showing that the first SH3 domains form both Ncks are the most conserved among the three domains, closely followed by the second SH3 domains. Identities between different types of domains ran between 22% and 31%, while similarities fell between 53% and 64%. Further comparisons were also made between the twenty-five residues identified as directly involved in ligand binding, either by directly contacting the ligand, or by supporting and shaping the binding pockets.
Figure 1: Schematic representation of the SH3 domains from Nck1 and Nck2 and alignment of the SH3 domains from Nck1 and Nck2. Positions of the RT-loop, N-Src, and Distal loop are indicated.

Table 1 comparison of all six SH3 domains from Nck-1 and Nck-2.

<table>
<thead>
<tr>
<th>ClustalW Identities:</th>
<th>Blast2 Similarities:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>11 12 13 21 22 23</td>
<td>11 12 13 21 22 23</td>
</tr>
<tr>
<td>11 26 25 82 29 27</td>
<td>11 64 54 95 62 58</td>
</tr>
<tr>
<td>12 26 29 22 80 31</td>
<td>12 64 56 62 92 53</td>
</tr>
<tr>
<td>13 25 29 25 25 73</td>
<td>13 54 56 54 54 82</td>
</tr>
<tr>
<td>21 82 22 25 24 25</td>
<td>21 95 62 54 60 58</td>
</tr>
<tr>
<td>22 29 80 25 24 27</td>
<td>22 62 92 54 60 58</td>
</tr>
<tr>
<td>23 27 31 73 25 27</td>
<td>23 58 53 82 58 58</td>
</tr>
</tbody>
</table>

Whole Nck-1 vs Nck-2: 68% Identity
Whole Nck-1 vs Nck-2: 79% Similarity

The six minimized models were obtained from Swiss-Model server were run on PROCHECK which was found that the models did not have any backbone or amino acid clashes and that the Ramachandaran plot was acceptable. Furthermore, the models were shown to be visually similar to the known SH3 structures (Figure 2). Next, the VADAR was used to evaluate the quality of the six Nck models which added important data such as surface accessibility of individual residues and overall packing quality. The overall VADAR data further supported the data obtained from PROCHECK that indicated that the final Nck models were of a very high stereochemical quality.
Surfaces were then generated by program SPDBV. There were very great differences among the domains, particularly when the electrostatic potentials were calculated (Figure 3). Surfaces were seen to vary the most at the regions surrounding Site 1, the specificity determining site, as the ‘top’ of the pocket is formed by the variable N-Src loop and the bottom by the RT-Src loop (Figure 4). Variations in loop lengths, particularly in the N-Src loop, led to apparent steric interferences among the domains, particularly the SH3-3s and their very long N-Src loop inserts (two residues longer than SH3-2, and three residues longer than SH3-1). This data could be indicated that each SH3 domains from Nck1 and Nck2 binds to different ligands due to not only different residues in the binding pockets but also the electrostatic charges which were from neighboring pocket site area.

Figure 2: Theoretical models of SH3 domains from all
DISCUSSION

Our results indicated that SH3 domains from Ncks have different properties in binding pockets which could imply that each SH3 domains binds with different ligand which were strengthen the concept that Ncks can act as a scaffold protein to bring two different proteins close together such as substrate and enzyme rather than recruiting the same type of proteins. Furthermore since there is no structural difference between solving structures of SH3 with and without bound ligand (Gosser et al., 1995, Pisabarro et al., 1998), our theoretical structures from Ncks can also further use as templates for computational identifying ligands.
REFERENCES


