Analysis of SH2 domain phosphopeptide interactions

S. J. Campbell & R. M. Jackson

Src homology 2 (SH2) domains are highly homologous phosphotyrosine-binding motifs of approximately one hundred amino acids. They are found in a range of signal transduction proteins and are therefore important targets for therapeutic intervention. The domains are well-conserved, particularly at the two major binding regions known as the pTyr and pTyr+3 binding pockets.

We have applied a variety of computational & molecular modelling techniques to the system, including structural and sequence alignments, clustering of binding sites and residue conservation scoring. This has allowed us to investigate molecular recognition within and beyond the traditional pTyr to pTyr+3 linear motif. The clustering of binding sites is shown in Fig. 1, where the newly clustered domains are in qualitative agreement with the results of experimental screening, indicating that the method used to cluster is able to independently generate functional information about the domains.

Fig. 1: Clustering of Binding Sites. We have used the experimentally defined protein-ligand contacts to define the protein binding site, allowing the clustering of the domains according to amino acid conservation in the binding site.

Fig. 2 shows residue conservation within a large experimentally determined group of SH2 domains, versus a group of domains determined using the clustering method. All members of the theoretical group are present in the larger experimental ‘parent’ group. The conservation
maps reveal regions beyond the traditional pTyr to pTyr+3 binding motif where the theoretical group shows residue conservation but the experimental group does not. Such regions could be involved in the discrimination and recognition of ligands.

**Fig. 2: Diversity within the phosphotyrosyl binding site.** (i) ‘Parent’ group representing a group of SH2 domains based on experimental screening according to recognition within the pTyr to pTyr+3 motif. (ii) Group C SH2 domains from Fig. 1, determined computationally by comparison of binding site residues. (i) and (ii) are coloured by residue conservation score where blue indicates a high level of conservation and red indicates low conservation. (iii) Difference map showing the difference in residue conservation scores between the experimental group and the theoretical group mapped back to the same structure. Green regions represent a gain in conservation in the theoretical group relative to the experimental parent group. Grey regions represent a loss of conservation in the theoretical group relative to the experimental group. White regions represent no change in conservation score.

**Publications**

**Funding**
This work was supported by the BBSRC.