Journal of Structural Biology 187 (2014) 58–65

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

In silico analysis and experimental verification of OSR1 kinase – Peptide interaction

Thomas M. Austin [a], David P. Nannemann [b], Samuel L. Deluca [b], Jens Meiler [b,c], Eric Delpire [a,*]

[a] Department of Anesthesiology, Vanderbilt University School of Medicine, Nashville, TN 37232, United States
[b] Department of Chemistry, Vanderbilt University, Nashville, TN, United States
[c] Departments of Pharmacology and Bioinformatics, Vanderbilt University, Nashville, TN, United States

ARTICLE INFO

Article history:
Received 10 December 2013
Received in revised form 25 April 2014
Accepted 4 May 2014
Available online 9 May 2014

Keywords:
Computer modeling
Protein design
Peptide conformation
Ste20 kinases
Rosetta Modeling Suite
Chloride transport

ABSTRACT

The oxidative-stress-responsive kinase 1 (OSR1) and the STE20/SPS1-related proline/alanine-rich kinase (SPAK) are key enzymes in a signaling cascade regulating the activity of Na^+–K^+–$2Cl^-$ cotransporters (NKCC1–2) and Na^+–Cl^- cotransporter (NCC). Both kinases have a conserved carboxyl-terminal (CCT) domain, which recognizes a unique peptide motif present in OSR1- and SPAK-activating kinases (with-no-lysine kinase 1 (WNK1) and WNK4) as well as their substrates (NKCC1, NKCC2, and NCC). Utilizing various modalities of the Rosetta Molecular Modeling Software Suite including flexible peptide docking and protein design, we comprehensively explored the sequence space recognized by the CCT domain. Specifically, we studied single residue mutations as well as complete unbiased designs of a hexapeptide substrate. The computational study started from a crystal structure of the CCT domain of OSR1 in complex with a hexapeptide derived from WNK4. Point mutations predicted to be favorable include Arg to His or Trp substitutions at position 2 and a Phe to Tyr substitution at position 3 of the hexapeptide. In addition, de novo design yielded two peptides predicted to bind to the CCT domain: FRFQVT and TRFDVT. These results, which indicate a little bit more freedom in the composition of the peptide, were confirmed through the use of yeast two-hybrid screening.

1. Introduction

SPAK (SPS1-related proline/alanine-rich kinase) and OSR1 (oxidative stress response 1 kinase) are members of the Ste20-related family of protein kinases (Dan et al., 2001; Delpire, 2009). They modulate the activity of cation-chloride cotransporters (Gagnon et al., 2006; Grimm et al., 2012; Lin et al., 2011; McCormick et al., 2011; Rafiqi et al., 2010), which are involved ion secretion (Kurihara et al., 2002; Matthews et al., 1993) and reabsorption (Gimenez and Forbush, 2005; Pacheco-Alvarez et al., 2006) across a variety of epithelia, Cl^- homeostasis in neurons (Austin and Delpire, 2011; Blaesse et al., 2009; Delpire, 2000; Delpire and Austin, 2010), and cell volume maintenance and regulation in many cells (Gagnon and Delpire, 2012). Molecular studies have demonstrated that kinase binding to the substrate is a prerequisite for the function of SPAK and OSR1 (Gagnon et al., 2006; Piechotta et al., 2003, 2002). The interaction between SPAK/OSR1 and substrates involves a ~90 residue domain located at the C-terminal tail of the kinase and a short conserved peptide located within the substrate (Piechotta et al., 2002). Yeast 2-hybrid analyses indicated that the peptide needed to be nine residues long, at least when located at the extreme C-terminus of the bait protein. Amino acid alignment between the cytosolic N-terminal tails of membrane transporters (related targets) revealed the preliminary conserved sequence: Arg-Phe-Xaa-Val (Piechotta et al., 2002). Following a large yeast 2-hybrid screen that used the conserved carboxyl-terminal domain of SPAK as bait (Piechotta et al., 2003), the motif was expanded to [Val/Ser/Gly]-Arg-Phe-Xaa-[Val/Iso]-Xaa-Xaa-[Thre/Ser/Val/Iso]. A whole mouse proteome search identified some 170 proteins containing this expanded motif (Delpire and Gagnon, 2007).

The crystal structure of the ~90 amino acid human OSR1 domain, which includes an embedded GRFQVT hexapeptide from human WNK4, was resolved at 1.95 Å (PDB ID: 2v3s, (Villa et al., 2007)). This domain, which was termed conserved carboxyl-terminal (CCT) (Villa et al., 2007) or protein fold 2 (PF2) (Lee et al., 2009), is represented in Fig. 1. The most salient feature is a hydrophobic groove that accommodates the Phe and Val residues of the peptide. While the protein fold was originally thought to be unique to SPAK

* Corresponding author. Address: Department of Anesthesiology, Vanderbilt University School of Medicine, T-4202 Medical Center North, 1161 21st Avenue South, Nashville, TN 37232, United States. Fax: +1 (615) 343 3916.
E-mail address: eric.delpire@vanderbilt.edu (E. Delpire).

http://dx.doi.org/10.1016/j.jsb.2014.05.001
1047-8477/

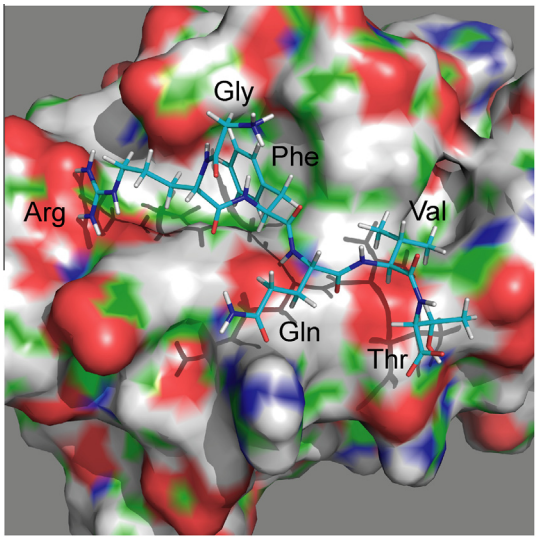

Fig.1. PyMol rendering of the hydrophobic pocket of OSR1 with the GRFQVT peptide of WNK4. The surface representation of the OSR1 domain highlights negative (red), positive (blue), and polar (green) moieties.

and OSR1 (Villa et al., 2007), we later showed that it is also present, at least partially, downstream of the catalytic domain of WNK4 (Delpire and Gagnon, 2008; Gagnon and Delpire, 2012). Availability of this domain – peptide structure allowed us to use the docking and design applications of Rosetta, a protein structure prediction and functional design software package, to assess the binding of various hexapeptides in the pocket and estimate binding energies. Thus, this analysis allows us to better understand the amino acid requirements for the interaction by ranking peptides with favorable to unfavorable energies. It also comprehensively determines the peptide sequences consistent with OSR1 and SPAK interaction thereby identifying potential sequences that have not yet been implied but can now be tested experimentally.

2. Material and methods

2.1. Computational modeling

We started with a three-dimensional representation of the crystal structure of the CCT/PF2 domain of OSR1 kinase in complex with a hexapeptide (GRFQVT) derived from WNK4 (Fig. 1). This file was obtained from the Protein Data Bank (http://www.rcsb.org/) as 2v3s, representing the work performed by Villa and collaborators (Villa et al., 2007). Crystallographic coordinates for extraneous molecules and fragments were removed, leaving a lone CCT-peptide complex. The resulting structure was then energy minimized using the Rosetta 3.4 relax application (Raman et al., 2009; Verma and Wenzel, 2007), according to the score12 energy function. This protocol adjusts the protein backbone and side chain torsion angles as a means of correcting local crystallographic bias, minimizing internal clashes, and moving the structure into an energy minimum on the Rosetta score12 energy function. As Rosetta employs a stochastic Monte Carlo Metropolis sampling strategy, multiple trajectories are needed to search the conformational (and sequence) space comprehensively. Here one hundred relaxed CCT-peptide complexes were produced and the top five models were chosen based on lowest Rosetta total energy score.

The native hexapeptides of the five relaxed complexes were then copied into separate files. Utilizing a simple python script, each residue of the hexapeptides was mutated into the other 19 canonical amino acids at all six positions, resulting in a total of 114 mutated peptides for each relaxed structure. The backbone atoms of all of these mutants had similar three-dimensional coordinates as their starting hexapeptides with the alterations only occurring at the side chains. These engineered peptides were then recombined with their corresponding unbound CCT domains (CCT-peptide complexes with peptides removed).

The mutated hexapeptides, along with the native forms, were then docked into the CCT domains by utilizing the FlexPepDock application of Rosetta 3.4 (London et al., 2011; Raveh et al., 2010) with the following flags: -use_input_sc, -ex1, -ex2, -pep_refine, and -unboundrot. Two hundred models were produced for each relaxed structure with a total of one thousand models created for each mutated hexapeptide. A low-resolution pre-optimization flag (-lowres_preoptimize) was employed in half of the docking runs in order to sample a larger peptide conformational space. A Rosetta binding energy (ddG) was calculated to assess the stability of the docked protein–peptide complex. The top ten models for each mutation, regardless of initial relaxed structure, were determined based on ddG and reweighted total energy score. The reweighted score is a combination of interface score, peptide score, and total score and is a better scoring function than score12 in the case of flexible peptides docking onto their receptors (Raveh et al., 2011). After averaging the scores of the ten models, each mutant was compared to the native CCT-peptide and other mutated peptide complexes.

Separately, the design module (Jha et al., 2010; Kuhlman and Baker, 2000) of Rosetta 3.4.1 was applied to the native hexapeptide in order to sample the sequence space consistent with CCT binding in a more unbiased fashion. In our procedure, only the hexapeptide was targeted for redesign while the CCT domain was left untouched. High resolution docking of these designed peptides into the binding pocket of the CCT domain was achieved through the use of FlexPepDock with similar options as previously described. After docking, another iterative round of design was performed on the hexapeptide to further refine the list of mutable residues. Again, two hundred decoys were produced for each relaxed structure with a total of one thousand models being created. The top one hundred models based on ddG were analyzed and compared to the native structure using PyMOL (PyMOL Molecular Graphics System, Version 1.5.0.4, Schrödinger, LLC). This entire process is outlined in Fig. 2.

2.2. Yeast-2-hybrid analysis

The regulatory domain of mouse SPAK (residues 353–556) fused to the GAL4 activating domain in pACT2 was originally isolated from a Clontech mouse brain library (Piechotta et al., 2002). The clone was re-transformed into PJ69-4A cells (James et al., 1996)) according to standard yeast handling procedures (Yeast Handbook, Clontech) and plated on -LEU plates. Sense and antisense oligonucleotides were purchased from Sigma Genosys. Upon annealing, the oligonucleotides create overhanging 5′ *Eco*RI and 3′ *Bam*HI sites that are directly used for ligation. The annealed adaptors encode 14 amino acid peptides that includes EF (*Eco*RI site), QLVG (linker), RFQVT or mutant (PF2 target peptides), and SSK followed by a stop codon. The QLVGRFQVTSSK sequence is original to the SPAK binding site in WNK4 ((Piechotta et al., 2003; Villa et al., 2007). The adaptors are ligated downstream of the Gal4 binding domain in the pGBDUc2 vector. Yeast cells containing the regulatory domain of SPAK in pACT2 were then transformed with individual peptide clones in pGBDUc2 and plated on double dropout -LEU, -URA plates. Yeast clones were then re-streaked on double-dropout plates as controls and triple dropout -LEU, -URA, -His, 2 mM 3-amino-1,2,4-triazole plates.