MD-based Energetic Heat Kernel Analysis: Insights to PDZ Allosterism

Ben Cowan[α], Kelly M. Thayer[•□]

College of Integrative Sciences Program[α], The Department of Chemistry [•], The Department of Computer Science[□]
Wesleyan University, Middletown, CT 06459

Introduction:

PDZ-domains compose an abundance of signaling and interaction centers found in many different protein families, themselves constituting the most abundant protein interaction module in mammals (1). PDZ domains display single domain allosteric modulation and are implicated in many protein networks involving allostery, stability, and folding interactions within single-domain proteins. Thus, investigating PDZ domain dynamics prove useful models for exploring new methodologies in allosteric pathway analysis while further enriching understandings of the signaling pathways characterizing specific PDZ-domain containing systems.

In this analysis, we investigate the energetic dynamics underlying the PDZ-containing system CRIB-Par6 through a novel heat-kernel based methodology. The 123-residue long CRIB-Par6 protein contains a canonical PDZ domain which are typically 80-100 amino acid residues in length and have a highly conserved single carboxylate binding site in a groove between their α2 and β2 structural elements. On its a1 helix, CRIB-Par6 directly interacts with Cdc4 while CRIB-Par6's semi-CRIB motif allosterically binds the Cdc42 effector protein in an extended antiparallel b-sheet conformation with the b2 strand of Cdc42.

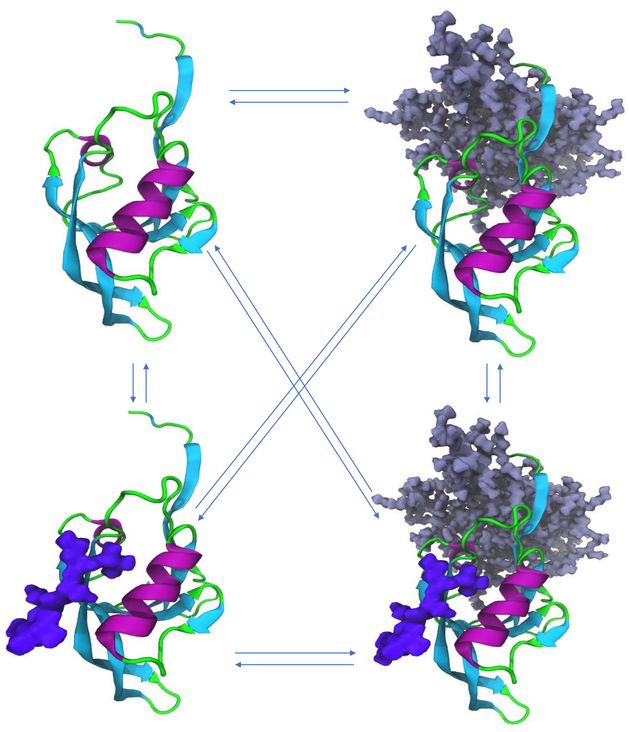


Figure 1: The four CRIB-Par6 systems investigated in our MD-Sector Analysis rendered in VMD. Cartoon model represents the CRIB-Par6 protein. The violet surf and grey surf models represent the KESLV peptide-ligand and the Cdc42 allosteric effector protein respectively.
Top left: CRIB-Par6, bottom left: KESLV Ligand bound CRIB-Par6, top right: Cdc42 Allosteric Effector bound CRIB-Par6, bottom right: Cdc42 and KESLV bound CRIB-Par6.

Methods of Heat-Kernel and Embedding Error Analysis:

The Heat Kernel Analysis methodology utilized in this study (2) captures interaction energy dynamics through the construction of temporally embedded energetic networks. These energetic networks are graph structures representing the locally-thresholded degree of pairwise interaction energy for residues of a protein for a frame of an MD-simulation on a pairwise atom-atom basis. Once generation of the heat-kernel is created for all residues for each simulation frame, the heat-kernel is temporally embedded into a principle-component graph representing the energetic associations for each residue over the course of an entire MD-simulation. Embedding error (EE) analysis is then performed on the embedded heat kernel PCA to reveal the degree to which a residue's energetic embeddings in a graph structure distribute differently between different constructs' graphs. We perform embedding error analysis on each bound construct in reference to unbound CRIB-Par6.

Heat-Kernel and Embedding Error Analysis of CRIB-Par6:

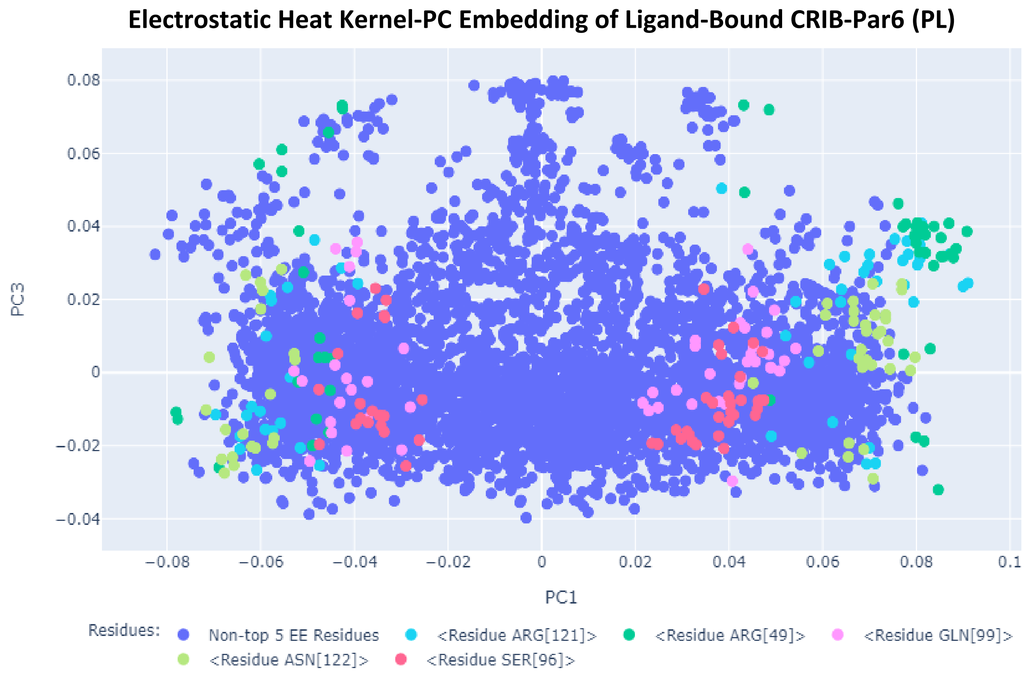
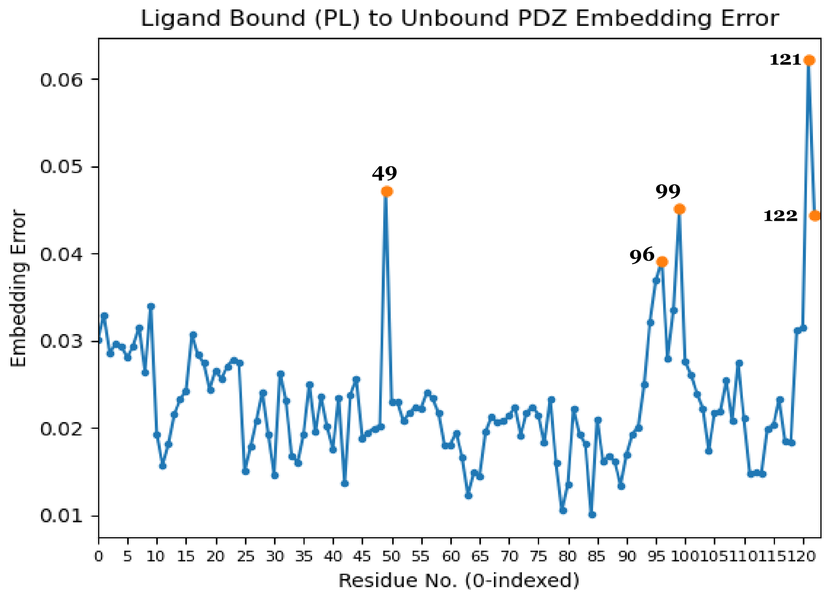


Fig 1a. Temporally-embedded electrostatic heat kernel PCA for all 123 residues of the ligand-bound construct of CRIB-Par6 over 51 sampled frames from a 100ns MD-simulation (6273 embedded nodes). The top 5 residues with highest embedding error from the unbound structure are highlighted in their respective colors as specified above.



- ***Fig 1b.*** *Line plot representation of the embedding error (EE) values for all 123 residues between the temporally embedded heat-kernels of the PL and UB states of CRIB-Par6.*
- *Top 5 residues (orange highlight) with the greatest embedding error (in descending order):*

  ARG[121], ARG[49], GLN[99], ASN[122], and

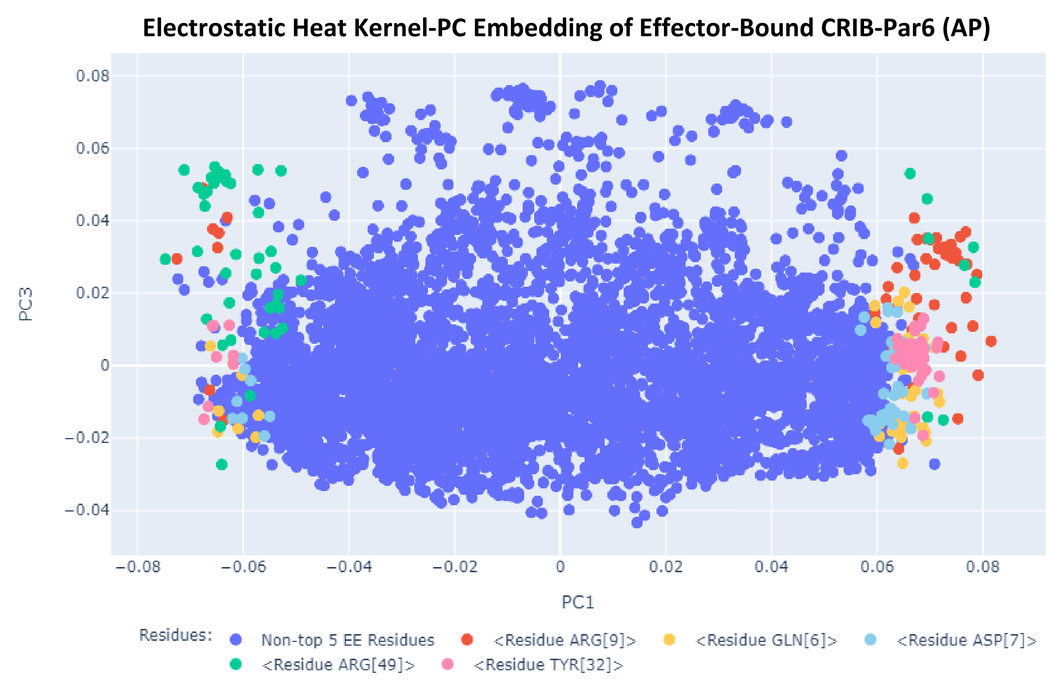
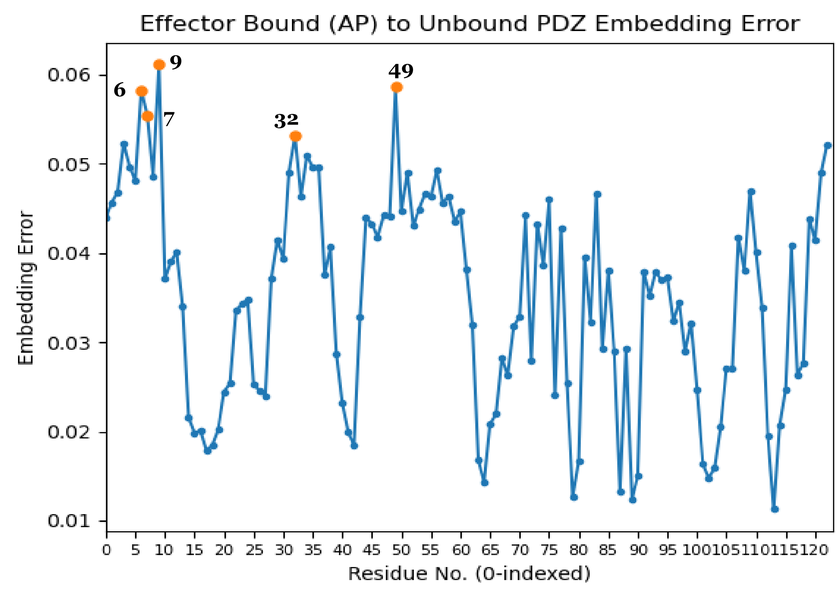


Fig 2a. Temporally-embedded electrostatic heat kernel PCA for all 123 residues of the effector-bound construct of CRIB-Par6 over 51 sampled frames from a 100ns MD-simulation (6273 embedded nodes). The top 5 residues with highest embedding error from the unbound structure are highlighted in their respective colors as specified above.



- **Fig 2b.** Line plot representation of the embedding error (EE) values for all 123 residues between the temporally embedded heat-kernels of the AP and UB states of CRIB-Par6.
- Top 5 residues (orange highlight) with the greatest embedding error (in descending order):

  ARG[9], GLN[6], ASP[7], ARG[49], and TYR[32]

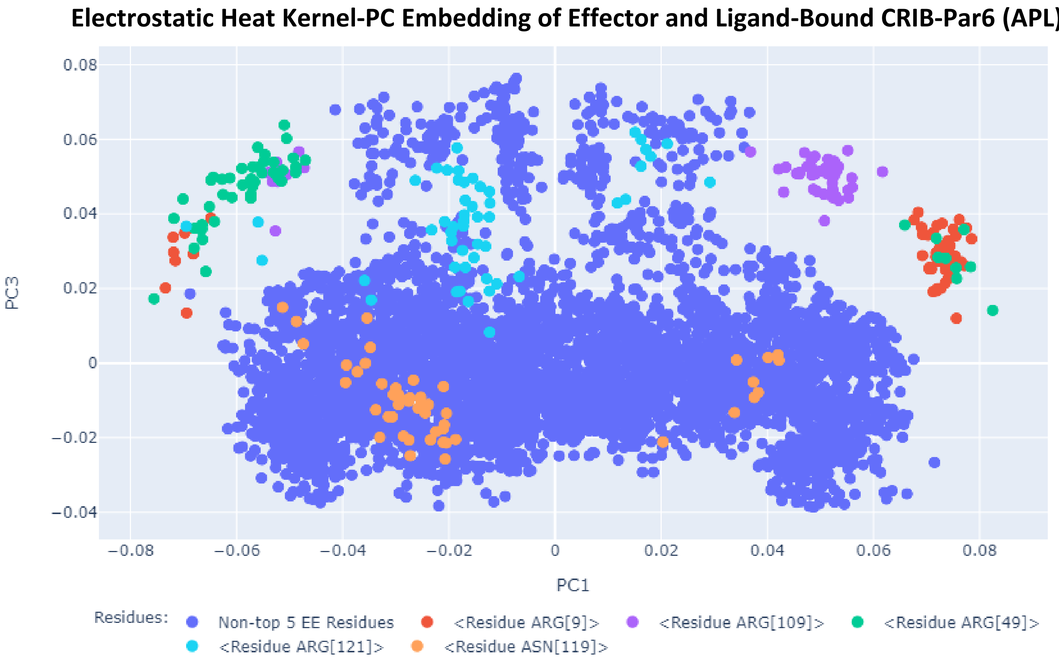
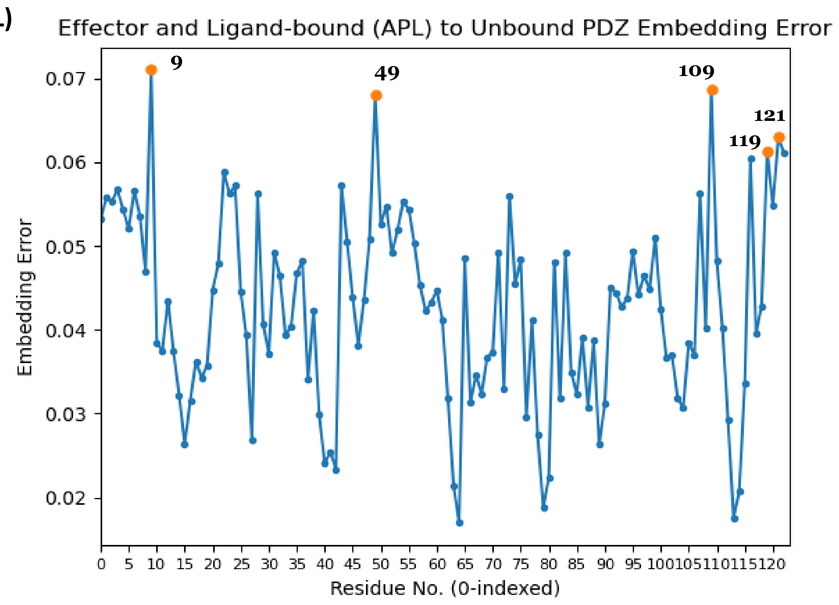


Fig 3a. Temporally-embedded electrostatic heat kernel PCA for all 123 residues of the effector and ligand-bound construct of CRIB-Par6 over 51 sampled frames from a 100ns MD-simulation (6273 embedded nodes). The top 5 residues with highest embedding error from the unbound structure are highlighted in their respective colors as specified above.



- **Fig 3b.** Line plot representation of the embedding error (EE) values for all 123 residues between the temporally embedded heat-kernels of the APL and UB states of CRIB-Par6.
- Top 5 residues (orange highlight) with the greatest embedding error (in descending order):

  ARG[9], ARG[109], ARG[49], ARG[121], and ASN[119]

Conclusions:

- Particularly evident are changes in both the clustering and embedding error for residues within allosteric bound constructs (AP and APL) and only ligand bound constructs (PL) in CRIB-Par6.
- Residues displaying the highest energetic variance in allosterically-bound states display greater energetic modulation from the unbound state compared to residues in purely ligand-bound conformations.
- Residues with the highest embedding error in allosterically modulated CRIB-Par6 constructs display greater clustering and energetic association as revealed through the heat kernel analysis.
- Changes in the energetic dynamics between allosterically-modulated states are reflected through heat-kernel analysis and enable for the identification of residues displaying high energetic variance and thus potential importance in the modulation of allosteric pathways.

References:

1) *Lee, H.-J., & Zheng, J. J. (2010). PDZ domains and their binding partners: structure, specificity, and modification. Cell Communication and Signaling, 8(1), 8. doi:10.1186/1478-811X-8-8*

2) *Abramson, Dylan (2021). Thesis. Wesleyan University.*

Acknowledgements:

- Dylan Abramson
- Thayer Lab Members
- The College of Integrative Sciences Summer Research Program
- Henk Meij
- NSF grant #CNS-0619508 and CNS-095985 for high performance computing facilities at Wesleyan
- NIH R15 GM128102-01 to KMT