

Bioinformatics Paper: Investigating the Function of ARAP2 PH 5

Gardner, Lauren

Biol 4010W

July 2020

### Abstract

The ARAP2 protein contains multiple PH (Pleckstrin homology) domains. The C-terminal PH domain in this protein is referred to as PH5, and the current function of this domain is unknown. Based on existing knowledge of the ARAP2 protein's involvement in the Akt signaling pathway, I predict that PH5 has a similar function to the PH domain located in Akt-related proteins. I used bioinformatics tools to assess this prediction. Results indicate it could be possible these two PH domains both undergo conformational changes for multiple instances of binding to occur.

## Introduction

The ARAP family of proteins are a family of Arf GAP proteins. Arf GAPs are essential when it comes to the regulation of the actin cytoskeleton. They are also known to be important components in membrane trafficking. They are GTPase activating proteins that control actin remodeling through the regulation of Arf and Rho cell signals. Within the Arf Gap domain in these proteins, a conserved zinc finger motif is responsible for the hydrolysis of GTP. However, this is just one function of the Arf GAPs, and there are several more diverse and often unknown functions as well (Randazzo 2004.) The ARAPs are named after their Rho GAP domains, and they consist of three different types: ARAP1, ARAP2, and ARAP3. They all contain a SAM (sterile alpha motif) domain at the N-terminus, an Arf Gap domain, a Rho Gap domain, ANK repeats, and five PH (pleckstrin homology) domains (Randazzo 2004.) Many of the functional roles of the ARAP proteins are still unknown.

Despite the three ARAPs having a similar domain structure, they differ however in their functional characteristics. ARAP1 and ARAP3 are central in the organization of the actin cytoskeleton, and overexpression of ARAP1 and ARAP3 lead to cell rounding and loss of stress fiber assembly. ARAP2, by contrast, is involved in focal adhesion assembly. Another major difference between ARAP2 and ARAP1/3 is that ARAP2 has not been found to contain activity in its Rho GAP domain while ARAP1/3 have Rho GAP activity (Yoon 2006) (Miura 2002.) There is also evidence that the substrate binding affinities between ARAP1, ARAP2, and ARAP3 can differ greatly as well. All the ARAPs bind to specific types of phospholipids, but the degree to which they bind vary. ARAP3 can bind to PtdIns (3,4,5)P<sub>3</sub> (or PIP<sub>3</sub>, a phospholipid found in the plasma membrane) strongly, while ARAP1 binds to the same phospholipid more weakly. However, in all three ARAP proteins, phospholipid binding is functionally associated with the PH domains (Craig 2010.)

The PH domains in the ARAPs are not well defined in a functional sense, despite the evidence of phospholipid binding. Each PH domain contains a conserved structural fold, and this is the case for all PH domains. In isolation, a PH domain may or may not bind specifically to membrane phospholipids (also known as phosphoinositides) at a high affinity. In fact, only a minority of isolated PH domains can bind

to these phospholipids at a high affinity, while most bind weakly or not at all. This would indicate that most high affinity binding between ARAP domains and membrane phospholipids most likely occur between multiple PH domains at the same time, in a coordinated fashion (Craig 2010.) In ARAP2,  $\text{PIP}_3$  binding has also been shown to be unrelated to the signaling activity responsible for focal adhesion assembly. Focal adhesion control mechanisms in ARAP2 can be traced to the Arf GAP domain binding with Arf6 protein (Casalou 2016). Yet, focal adhesion control is completely independent from  $\text{PIP}_3$  binding and associated cell signaling (Luo 2018). This indicates that the domains of ARAP2 could be working independently from one another in addition to coordinated efforts.

The differences between the ARAPs shows how ARAP2 could have a distinct, important function separate from the other ARAP proteins. Many Arf GAP proteins in general are thought to be related to cancer due to their roles in various cancer-associated cell signaling pathways (Ha 2008). Some scientists argue that ARAP2 may be integral in the Akt pathway, a signaling pathway studied extensively in relation to cancer. Because ARAP2 binds to  $\text{PIP}_3$ , and  $\text{PIP}_3$  has been shown to be important in Akt activation, scientists think that there is a relationship between ARAP2 and Akt inhibition, even though the exact mechanism is still being studied (Luo 2018).

ARAP2's association with Akt could illuminate the function of the PH 5 domain and the ARAP2 protein. In homo sapiens, Akt and its associated pathway P13K is involved in the regulation of cancer-related proteins. Akt deregulation is a common component of many types of cancer, and an increase in Akt kinase activity is reported in a large percentage of cancer tissues (Revathidevi 2019). Akt proteins also have a highly conserved PH domain which is known to mediate lipid-protein and protein-protein interactions. When this PH domain is phosphorylated, it binds with  $\text{PIP}_3$ , causing Akt to change shape so it is further available for additional binding (Revathidevi 2019.)

I hypothesize that the C-terminal PH domain (PH 5 domain), in ARAP2 has a similar function to that of the PH domain located within the Akt family of proteins. I will use comparisons of these two PH domain sequences to help illuminate the function of PH 5 domain in ARAP2. Hopefully, similarities between ARAP2 and Akt will also emerge from studying their PH domain sequence similarity. The PH 5

domain can also be studied in relation to other homologs, to understand how it can be compared with existing, known protein structures.

## Methods

Multiple types of bioinformatics tools were used to examine the PH5 domain. For identifying putative homologs of ARAP2, Blastp [Altschul et al., 2020] was used to search in the Refseq database [Pruitt et al., 2020] for specific organisms (*Danio rerio*, *Mus musculus*, *Xenopus tropicalis* and *Homo sapien*) using the Blosum62 matrix. Each protein was compared to ARAP2 in *Homo sapiens* using Emboss pairwise alignment tools [Madeira et al. 2020]. The FASTA sequences from the homolog proteins were compared using a multiple sequence alignment tool, Clustal Omega [Sievers, et al., 2020]. The alignment was then entered into JalView [Waterhouse, et al., 2020] for a more dynamic visualization.

For analyzing the secondary structure of PH 5, the FASTA sequence was retrieved from Genbank [Clark, et al., 2020]. Using SSP [SSP, 2020], the PH5 domain sequence was analyzed with PSI [Jones, 2020], PSS [Yan et. al., 2020] JPred [Drozdetskiy, et al., 2020], and Sable [Adamczak, et al., 2020] to find a consensus.

For the domain architecture, the ARAP2 FASTA sequence was analyzed by searching against five different Domain Architecture databases: SMART [Letunic, et al., 2020], Pfam [Finn, et al., 2020], CDD [Marchler-Bauer, et al., 2020], Prosite [Sigrist, et al., 2020], and InterPro [Mitchell, et al., 2020]. A graphic representation was produced of the consensus domain architecture after carefully deciding which residue ranges corresponded to each protein domain using the domain architecture analysis. The graphic was produced using the MyDomains tool on ProSite [Hulo, et. al., 2020].

For the tertiary structure, the FASTA sequence of the PH 5 domain from ARAP2 was again retrieved from GenBank [Clark, et al., 2020]. It was then used as input data for two tertiary structure prediction programs: HHPred [Zimmermann, et al., 2020], a homology modeling program, and Quark [Xu, et al., 2020], an ab initio program. PyMOL [Rigsby, et al., 2020] was used to produce a 3D graphic of each .pdb tertiary structural model post-output. A variety of evaluation tools were then used on each model: Verify3D [Bowie, et el., 2020], ProqQ3 [Uziela, et al., 2020], ProSA [Wiederstein, et al., 2020], and VoroMQA [Olechnovič, et al, 2020].

Structural analysis and visualization was performed on the 3D homology model as well as the ab initio model using PyMOL [Rigsby, et al., 2020]. Structural superposition was assessed on the homology model using Superpose [Maiti, et al., 2020].

The FASTA sequence for the PH domain within RAC-alpha serine/threonine-protein kinase was retrieved from Genbank [Clark, et al., 2020] and its PH domain assessed using CDD [Marchler-Bauer, et al., 2020]. The alignment between the PH domain in within RAC-alpha serine/threonine-protein kinase and PH 5 domain in ARAP was found using Emboss Needle [Madeira et al. 2020]. A secondary structure analysis was also performed on the PH domain within RAC-alpha serine/threonine-protein kinase using SSP [SSP, 2020] as well as PSI [Jones, 2020], PSS [Yan et. al., 2020] JPred [Drozdetskiy, et al., 2020], and Sable [Adamczak, et al., 2020]. Known PDB model 4EJN (a RAC-alpha serine/threonine-protein kinase), also found using HHPred, was aligned with the homology model based on the 4K2P template created previously from HHPred [Zimmermann, et al., 2020].

## Results

When analyzing homologs of the ARAP2 protein, ARAP2 proteins found in the organisms *Danio rerio*, *Mus musculus*, and *Xenopus tropicalis* showed enough similarity to ARAP2 to be considered putative homologs to the ARAP2 protein in homo sapiens. ARAP2 in *Danio rerio* had a 42.1% Identity and a 56.7% similarity compared with ARAP2 in *Homo sapiens*, ARAP2 in *Mus musculus* had a 82.8% Identity and a 88.7% similarity compared with ARAP 2 in *Homo sapiens*, and *Xenopus tropicalis* had a 46.3% Identity and a 60.6% similarity compared with ARAP 2 in *Homo sapiens*. ARAP1 in *Homo sapiens* was also found to be similar enough to be considered a homolog to ARAP2 with a 28.0% Identity and a 40.8% similarity. The multiple sequence alignment (Figure 1) of these homologs highlight the similarity between each amino acid sequence. In Figure 2, the image is enlarged, and the location of the PH 5 domain is identified. Conserved regions are highlighted in blue. According this multiple sequence alignment, the region encompassing the PH 5 domain appears to be highly conserved.

The domain architecture analysis for ARAP2 resulted in the following consensus: a SAM domain located near the N-terminus with residue range 4-70, PH 1 domain between residues 483-576, PH 2 domain between 588-675, an Arf GAP domain between residues 676-811, a PH 3 domain between residues 892-1005, PH 4 domain between residues 1015-1114, a RhoGap domain between residues 1114-1302, an RA (Ras associating) domain between residues 1325-1422, and finally, the C-terminal PH 5 domain between residues 1423-1543, or the domain location consistent with the CDD [Marchler-Bauer, et al., 2020]. Figure 3 exhibits this consensus in graphic form.

Secondary structural analysis on the PH 5 domain in ARAP2 (Figure 4) indicates there are seven beta strands and one alpha helix within the domain. While supersecondary structure elements are not accounted for, this structural analysis is suitable for understanding that the PH 5 domain within ARAP2 has a consistent structural makeup for a typical PH domain.

Tertiary structure was predicted using template based modeling tools. The PDB template 4K2P was found to be suitable enough for a somewhat reliable prediction of the tertiary structure of PH 5, based on a variety of evaluations. The 3D model (Figure 5) shows evidence of having the same seven beta



sheets and one alpha helix predicted by the secondary structure analysis. Electrostatic analysis reveals the model to also have both positive and negative regions, with many of the positive regions likely on the interior of the protein. The evaluation (Figure 6) of this homology model is mostly favorable: The evaluation from Verify3D (Figure 25) [Bowie, et al., 2020] indicated that 63.30% of residues have an average score above 0.2. The evaluation from ProQ3 (Figure 26) [Uziela, et al., 2020] indicated a score between .521 and .584. The evaluation from ProSA [Wiederstein, et al., 2020] (Figure 27) provided a Z-score of -3.94, which is consistent with existing PDB structures; there is a high energy region however in the beginning of the alignment. VoromQA (Figure 28) [Olechnovič, et al, 2020] indicated a score of 0.430, which shows that this model is reliable for assessing the PH 5 domain. The structure based sequence alignment (Figure 7) was somewhat strong, indicating that there are enough structural similarities in the model to draw a comparison.

The ab initio model (Figure 5) however had slightly more favorable evaluations; from Verify3D, the model received a “Pass” grade and had a 87.60% of residues above the 0.2 cutoff. The outcome of the model however was similar to the homology model, with seven beta strands, one alpha helix, and varying regions of positive and negative charges.

Using the alignment of the PH domain within RAC-alpha serine/threonine-protein kinase and the PH 5 domain in ARAP2, there is evidence of similar functioning residues, as many in the alignment are similar; there is a 34% similarity and a 52% identity (Figure 10). The identical residues were superimposed onto the 3D ab initio model, potentially mapping possible binding sites (Figure 11). The secondary structure analysis, also performed on the PH domain within RAC-alpha serine/threonine-protein kinase, indicates that the PH domain in this protein also contains a typical structure for a PH domain, except in contrast with the ARAP2 PH 5, this domain only has 5-6 beta sheets while ARAP2 PH 5 has one additional beta sheet. The identical residues found in the alignment between ARAP2 PH 5 and the PH domain from RAC-alpha serine/threonine-protein kinase, when mapped onto the ab initio model (Figure 12), indicate that potential binding sites congregate within the region surrounding the first three beta sheets, approximately between residues 2-51.

In addition, when the approximate residue range of the PH domain within the known PDB structure 4EJN, a RAC-alpha serine/threonine-protein kinase, was aligned with the homology model (Figure 5), the result indicated a RMSD equal to 0.680, showing the alignments are compatible in 3D space (Figure 13).

Figure 1: Multiple Sequence Alignment between ARAP2 in homo sapiens and homologs showing region of PH 5 Domain:

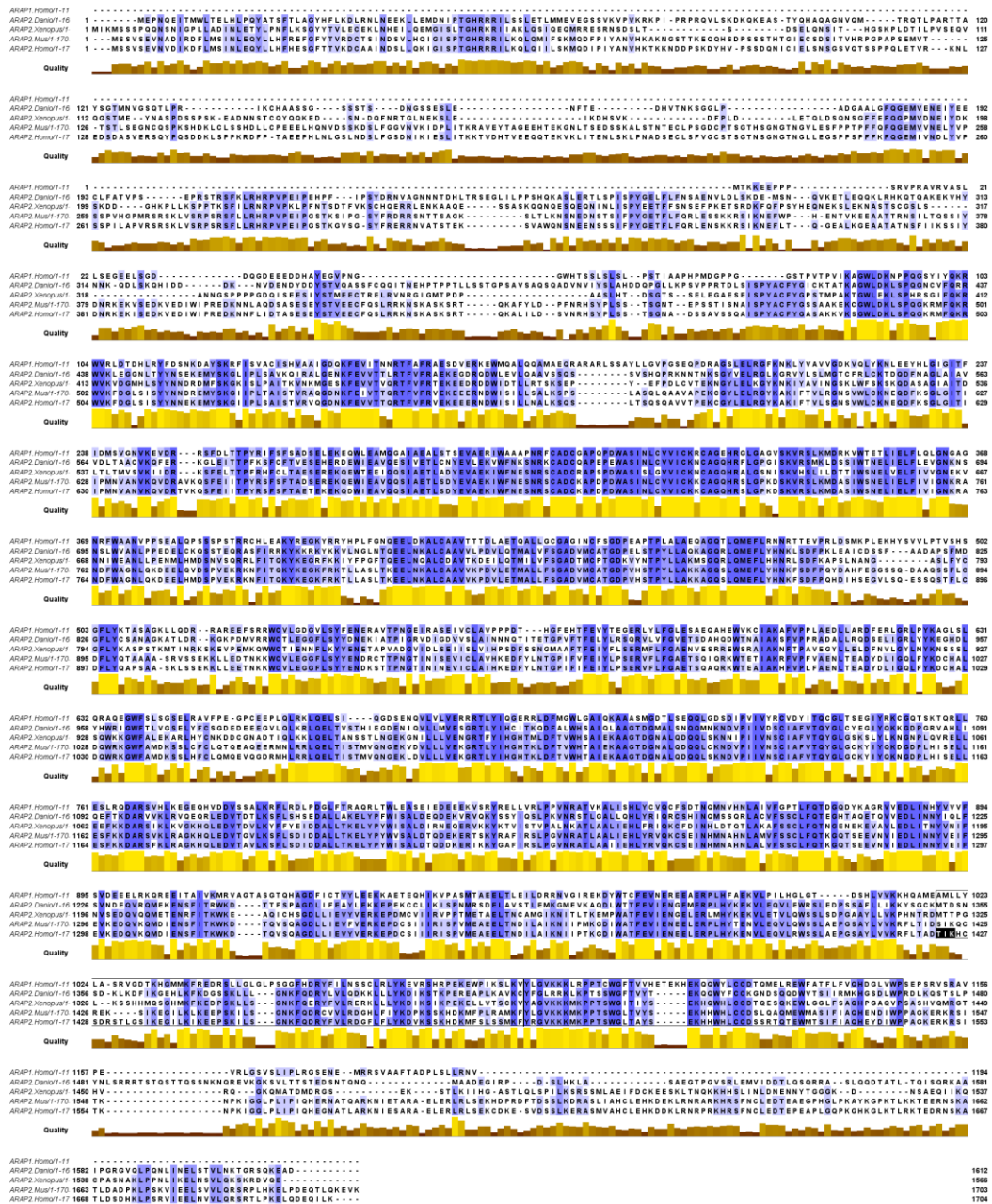


Figure 2: Enlargement of multiple Sequence Alignment between ARAP2 in homo sapiens and homologs showing region of PH 5 Domain:

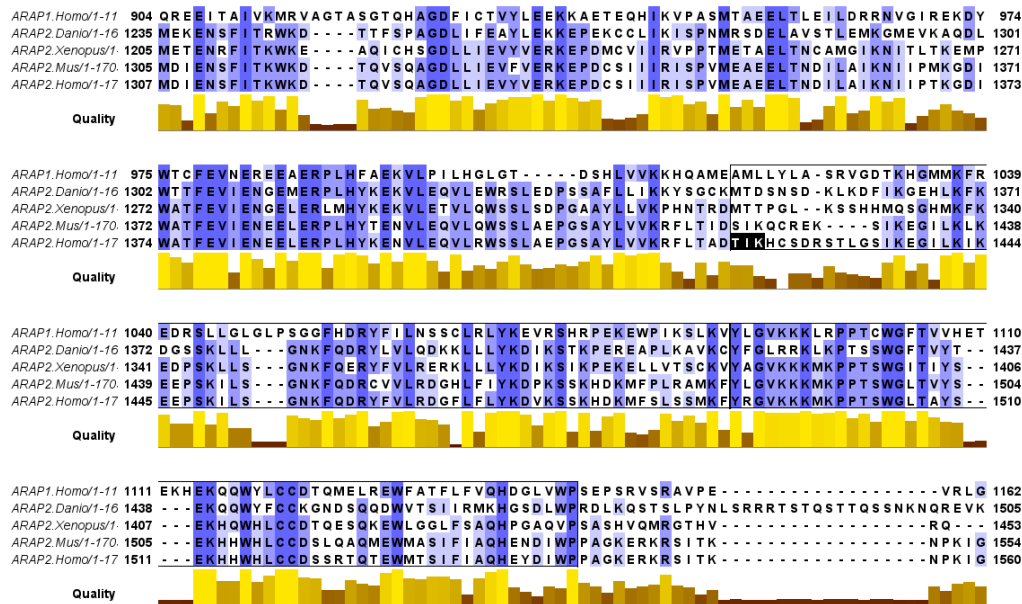


Figure 3: Domain Architecture consensus of the ARAP2 protein :

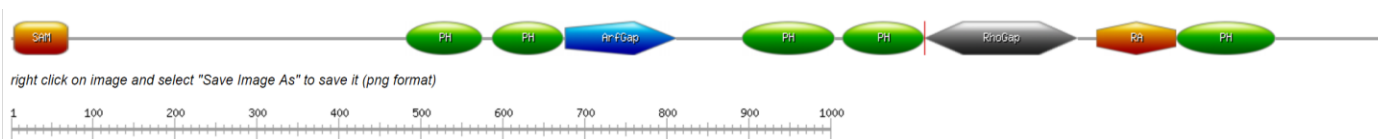


Figure 4: Secondary Structure Analysis Secondary Structure Prediction of

PH5\_ARAP:

	1	11	21	31	41	51
Sequence:	TIKHCSRSTLGSIKEGILKIKEEPSKILSGNKFQDRYFVL RDGFLFLYKDVKSSKHDKM					
PSI Pred:	CCCCCCCCCCCCCEEEEEEECCCCCCCCCCCCCEEEEEEECEEEEEEECCCCCCCCCE					
PSI Conf:	975567878899708999999379998778888706999997998999958888778806					
PSS Pred:	CCCCCCCCCCCCCCCCCEEEEEEECCCCCCCCCCCCCEEEEEEECEEEEEEECCCCCCCCCE					
PSS Conf:	975454445666642125898438875457876047799996996999825667885321					
JPRED Pred:	CCCCCCCCCCCCCCCCCEEEEEEECCCCCCCCCCCCCEEEEEEECEEEEEEECCCCCCCCCE					
JPRED Conf:	89887777777771336888731567777775318999855855898516777765113					
SABLE Pred:	CHHHHCCCCCCCCCCCCCEEEEEEECCCCCCCCCCCCCEEEEEEECEEEEEEECCCCCCCCCE					
SABLE Conf:	546544778888887766888668876668876445788875885899857888887665					
Majority Vote:	CCCCCCCCCCCCCXEEEEEECCCCCCCCCCCCCEEEEEEXCEEEEEEECCCCCCCCCX					

	61	71	81	91	101	111
Sequence:	FSLSSMKFYRGVKKMKPPTSWGLTAYSEKHHWHLCDDSSRTQTEWMTSIFIAQHEYDIW					
PSI Pred:	EECCCCCEEEEECCCCCCCCCEEEEEEECCCEEEEEEECCCHHHHHHHHHHHHHCCCCCCCC					
PSI Conf:	773554799940368899986089995799699999299999999999999861323589					
PSS Pred:	EECEEEEEEEEEEECCCCCCCCCEEEEEEECEEEEEEECCCHHHHHHHHHHHHHHHCCCCCCCC					
PSS Conf:	341115999712312589987369998167058974599889999999999973278668					
JPRED Pred:	EECCCCCEEEEECCCCCCCCCCCCCEEEEEEECCCEEEEEEECCCHHHHHHHHHHHHHHHCCCCCCCC					
JPRED Conf:	216761688601577777774589985586589861765899999999887622688889					
SABLE Pred:	EEEEEEEEEEEECCCCCCCCCCCCCEEEEEEECCCEEEEEEECCCHHHHHHHHHHHHHHHCCCCCCCC					
SABLE Conf:	566567899874456788887568988578757887588688999999988754688877					
Majority Vote:	EEXCXEEEEEECCCCCCCCCCCCCEEEEEEECCXEEEEEECCCHHHHHHHHHHHHHHHXCCCCC					

Figure 5: Tertiary Structure Prediction of PH5\_ARAP using template model 4K2P (left) with electrostatic analysis on surface of model (right):

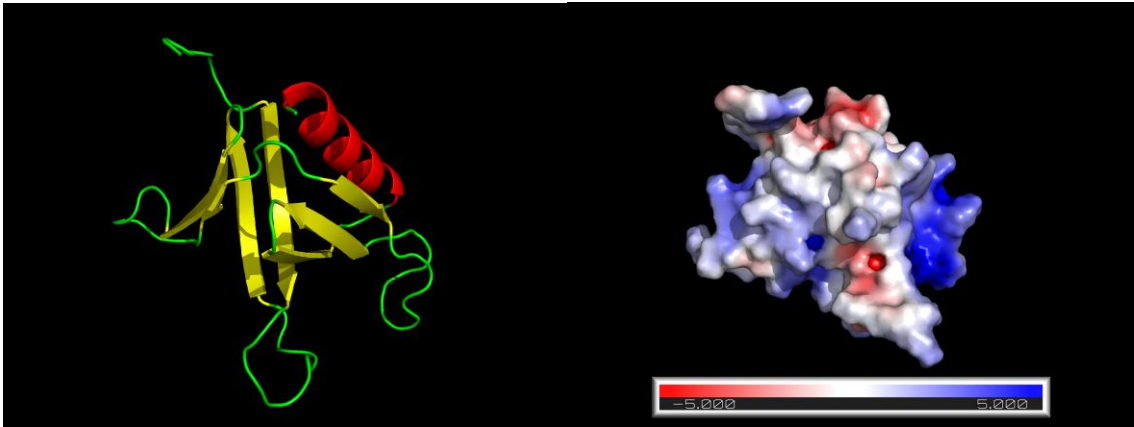
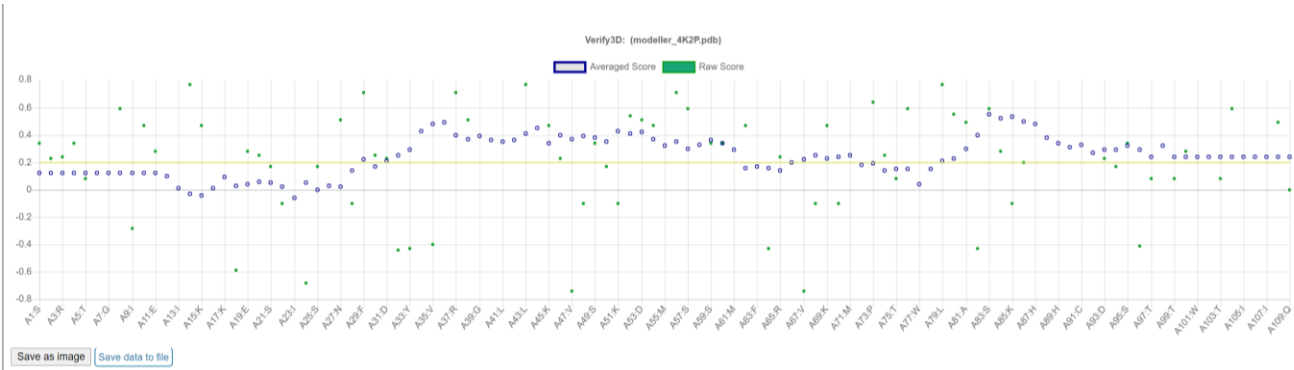


Figure 6: Evaluation of Tertiary Structure Prediction of PH5\_ARAP using template model 4K2P: ProQ3 (A), Verify3D (B), ProSA (C), and VoroMQA (D):

A:

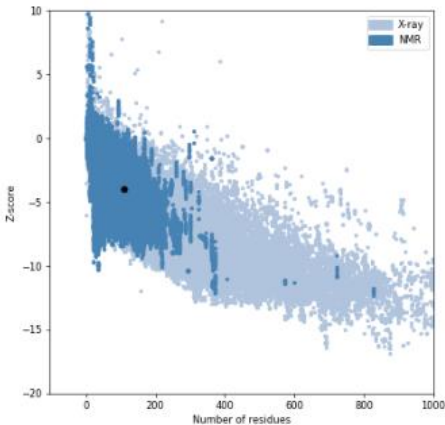
Model	Length	RunTime(s)	ProQ2D	ProQRosCenD	ProQRosFAD	ProQ3D
Model 0	104	36.5	0.438	0.482	0.521	0.514

B:



C:

Overall model quality  
Z-Score: -3.94



D:

Results:

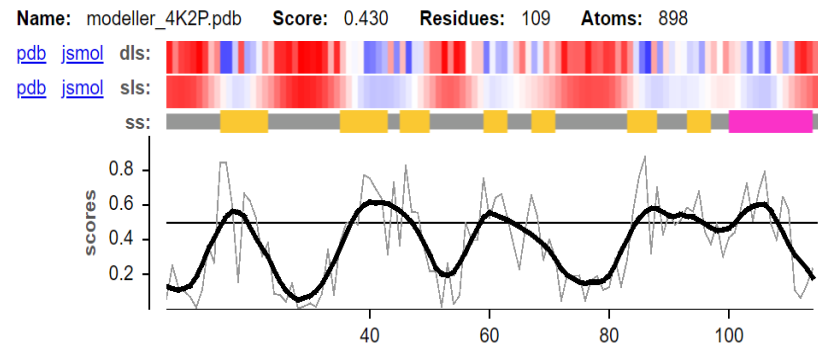


Figure 7: Structure Based Sequence Alignment between known PDB structure and newly created homology model (Figure 5) from 4K2P template using Superpose

4K2P_model_de	1	-----TVRKAGALAVKNFLVHKKNKKVESATRRKWKHYWVSLKGCTLF	43
		::  . . .:  :... : . . : : : : : : :	
PDBB_model_de	1	SDRSTLGSI-KEGILKIK-----EEPSKILSG--NKFQDRYFVLRDGLF	42
4K2P_model_de	44	FYESDGRSGIDHNSIPKHAVWVENSIVQAVPEHPKKDFVFCLSN-----	87
		. :..  ::   . .: . :	
PDBB_model_de	43	LYKD-----VKSS-----KHDKMFSLSMKFYRG	66
4K2P_model_de	88	-----SLG-DAF-----LFQTTSQTELENWITAIHSACATAVAR	120
		.  . :  ... : : .  :: . .	
PDBB_model_de	67	VKKKMKPPTSWGLTAYSEKHHWHLCCDSSRTQTE-WMTSIFIAQ-----	109
4K2P_model_de	121	HHHKEDTLRLKSEIKKLEQKIDMDEKLKKMGELQLSSVTDSKAAATILD	170
PDBB_model_de	109	-----	109
4K2P_model_de	171	QIFVWEQNLEQFQMDLFRFRCYLASLQGGELPNPKRLLAFASRPTKVAMG	220
PDBB_model_de	109	-----	109
4K2P_model_de	221	RLGIFSVSSFHALVAART	238
PDBB_model_de	109	-----	109

Figure 8: Ab initio Tertiary Structure Prediction of PH5\_ARAP (left) with electrostatic analysis on surface of model (right):

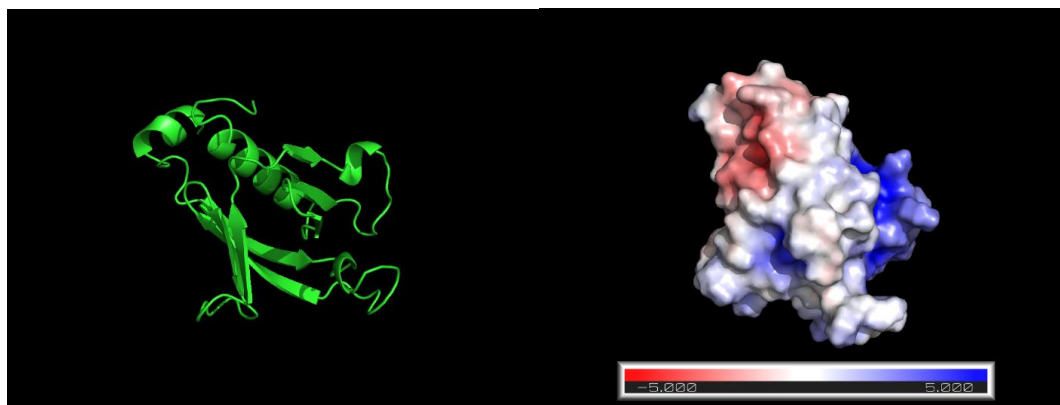
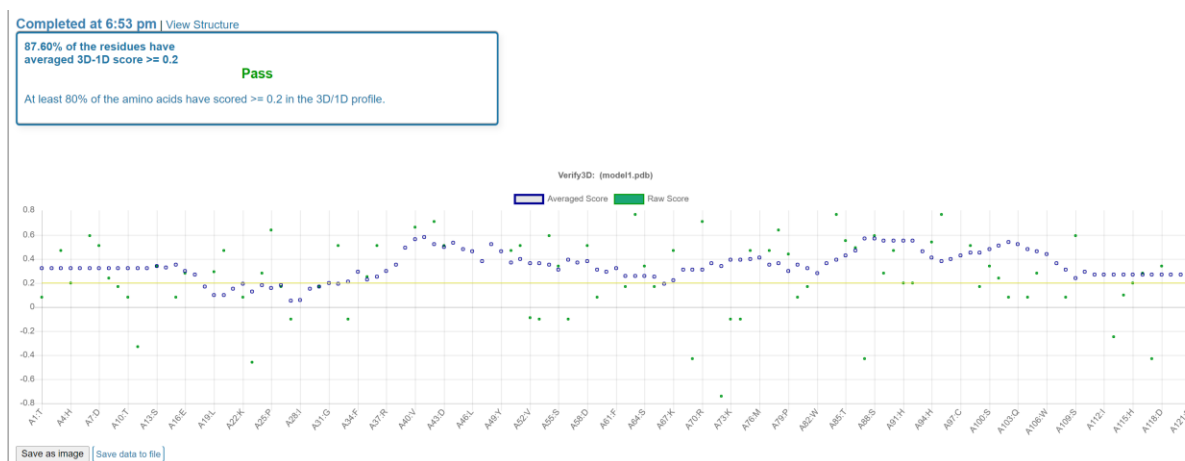


Figure 9: Evaluation of Tertiary Structure Prediction of PH5\_ARAP using ab initio model and Verify3D:





serine/threonine-protein kinase:

RAC-alpha	2	SDVAIVKEGWLHKGREYIK-----TWRPRYFLKNDGTFIGYKE	40
		.:.: : : .:.: .:.        .:.: : : : : : : : : :	
PH	9	STLGSIKEGILKIKEEPSKILSGNKFDQRYFVLR-DGFLFLYKD	51

alpha serine/threonine-protein kinase:

```

1          11          21          31          41          51
Sequence: MSDVAIVKEGWLHKRGEYIKTWPRPYFLLNKDGTFIGYKERPQVDQREAPLNNFSVAQC
PSI Pred: CCHHHEEEEEEEECCCCCCCCEEEEEEEECCCCCCCCCCCCCCCCCCCCCCCC
PSI Conf: 983320488999963887777324599993499899983795567888985326991455
PSS Pred: CCCCCEEEEEEEECCCCCCCCCCCCEEEEEEEECCCCEEEECCCCCCCCCCCCCCCCCCCC
PSS Conf: 987642555554214544455432799817970787504777666666655211332
JPRED Pred: CCCCCEEEEEEEECCCCCCCCEEEEEEEECCCCEEEECCCCCCCCCCCCCCCCCCCC
JPRED Conf: 9997413444566743664577701788521862674306776666777777777777777
Majority Vote: CCCCCEEEEEEEECCCCCCCCCCCCEEEEEEEECCCCEEEECCCCCCCCCCCCCCCCCCCC

61          71          81          91          101          111
Sequence: QLMKTERPRPNTFIIRCLQWTTVIERTFHVETPEEREWTTAIQTVDGLKKQEEEMDMF
PSI Pred: EEEEECCCCCCCCEEEEEECCCEEEEEEEEEECHHHHHHHHHHHHHHHHHHHHHHHHHHHCC
PSI Conf: 799659999966999929963899899981999999999999999998989984014
PSS Pred: EEEEECCCCCCCCEEEEEECCCCCCCCCCCCCCCCHHHHHHHHHHHHHHCCCCCCCCCCCC
PSS Conf: 133214678633898622554333443226896776766654321010234332111234
JPRED Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
JPRED Conf: 7776667777631113677776777777777777777777777777777777777777777
Majority Vote: EEEEECCCCCCCCEEEEEECCCCCCCCCCCCCCCCHHHHHHHHHHHHHHCCCCCCCCCCCC

```

Figure 12: Ab initio model of PH5\_ARAP2 highlighting identical residues between PH 5 domain of ARAP2 and PH domain of RAC-alpha serine/threonine-protein kinase:

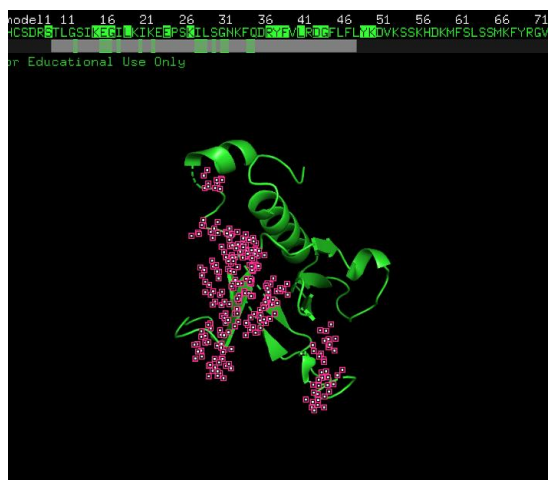
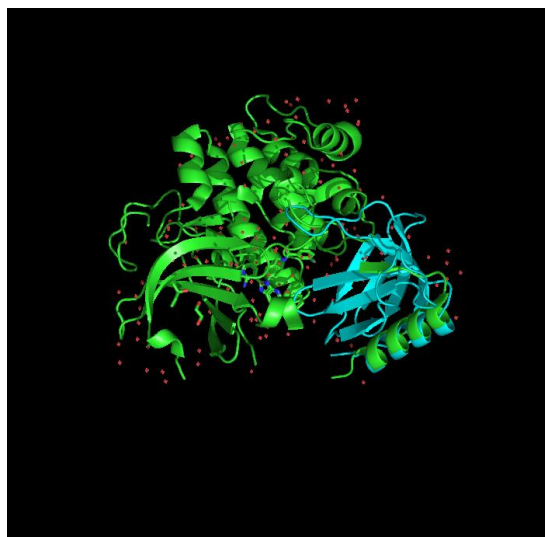


Figure 13: Known PDB model 4EJN (a RAC-alpha serine/threonine-protein kinase) aligned with PH 5 homology model from Figure 5 at location of the PH domain:



## Discussion

The discovery of putative homologs in relation to ARAP2 unlock the potential to understand ARAP2 as a part of a conserved system. With the identification of multiple conserved functional domains within the protein, there is evidence of PH 5 being conserved as well. According to the multiple sequence alignment (Figures 1 and 2), conserved areas include residues accounting for most of the identifiable domains according to the domain architecture consensus (Figure 3). Residues specifically for the PH 5 domain (1423-1543) are no exception. This establishes a basis for further studies regarding this domain.

The secondary structure analysis provides additional evidence that PH 5 is structurally consistent with other existing PH domains in terms of folding. The seven beta sheets and single alpha helix found in the secondary structure analysis show that this domain could operate in a similar manner to other known PH domains. The analyses could not however provide a clear consensus on where certain secondary structures begin and end within the chain, including where the alpha helix ends on the C-terminal end. The secondary structural analysis is reliable enough though to use for further investigation into tertiary structure.

The tertiary structure analysis revealed a bit more about the potential function of the ARAP2 PH 5 domain. The homology model based on the existing PDB template was satisfactory in predicting a reliable PH 5 domain structure. Therefore a 3D visualization was useful in understanding the overall tertiary dynamic of the domain. However, because the homology model could have scored higher, there are still ambiguous regions within the PH 5 domain. For example, from the Verify3D model (Figure 6B), residues at the beginning of the amino acid sequence scored lower than regions in the middle of the sequence. These residues, approximately TIKHCSDRSTLGSIKEGILKIKEEPSK (containing no hydrophobic residues), could indicate that there are functional properties here that are not the same as other PH 5 domains within the ARAP family.

The evaluation tools indicate that the ab initio model is a better fit for the PH 5 domain. According to the Verify3D for this model, there are no regions where the match is extremely poor, except for residues 16-33. However, when assessing the ARAP2 PH 5 alignment with the PH domain of RAC-

alpha serine/threonine-protein kinase, there are several identical residue matches between 16-33. In this region in the PH domain of RAC-alpha serine/threonine-protein kinase, there are known phospholipid binding sites according to the CDD [Marchler-Bauer, et al., 2020]. However, it is not clear if these phospholipid binding sites would be the same in PH 5, as the PyMOL electrostatic analysis indicated that the binding region could be at a different location.

When aligning the PH 5 domain with the PH domain sequence of known PDB structure containing RAC-alpha serine/threonine-protein kinase, the RMSD is 0.680, indicating this is a reliable alignment and can be used to further discover similar properties of these two domains (Figure 13). Knowing that PIP<sub>3</sub> binds to the more hydrophobic regions of most PH domains, it seems unlikely that the interior-located residues 16-33 (approximately within the first two beta sheets) would be responsible for the majority of this PIP<sub>3</sub> binding. It could potentially mean that the 16-33 residues instead bind additional types of phosphoinositides once the PH 5 domain undergoes a conformational change. This change would allow these residue regions to be more available to polar groups, similar to how Akt proteins also undergo conformational changes after initial binding has taken place.

## References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990 Oct 5;215(3):403-10.
- Bowie JU, Lüthy R, Eisenberg D. A method to identify protein sequences that fold into a known three-dimensional structure. *Science* 1991 Jul 12;253(5016):164-70.
- Casalou C, Faustino A, Barral DC. Arf proteins in cancer cell migration. *Small GTPases* 2016;7(4):270-82.
- Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic Acids Res.* 2016;44(D1):D67-D72. doi:10.1093/nar/gkv1276
- Craig HE, Coadwell J, Guillou H, Vermeren S. ARAP3 binding to phosphatidylinositol-(3,4,5)-trisphosphate depends on N-terminal tandem PH domains and adjacent sequences. *Cell Signal* 2010 Feb;22(2):257-64.
- Drozdetskiy A, Cole C, Procter J, Barton GJ. JPred4: A protein secondary structure prediction server. *Nucleic Acids Res* 2015;43:W389-94.
- Gille C, Föhling M, Weyand B, Wieland T, Gille A. Alignment-Annotator web server: rendering and annotating sequence alignments. *Nucleic Acids Res.* 2014;42(Web Server issue):W3-W6. doi:10.1093/nar/gku400
- Ha VL, Luo R, Nie Z, Randazzo PA. Contribution of AZAP-type arf GAPs to cancer cell migration and invasion. *Adv Cancer Res* 2008;101:1-28.
- Hoque, Huang, and Singh. Secondary structure visualization app. [Internet][Cited 2020 July 2] Available from:<https://sspred.herokuapp.com/>
- Luo R, Chen PW, Kuo JC, Jenkins L, Jian X, Waterman CM, Randazzo PA. ARAP2 inhibits akt independently of its effects on focal adhesions. *Biol Cell* 2018 Dec;110(12):257-70.
- Madeira F, Park YM, Lee J, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research.* 2019 Jul;47(W1):W636-W641. DOI: 10.1093/nar/gkz268.
- Marchler-Bauer A et al. (2017), "CDD/SPARCLE: functional classification of proteins via subfamily domain architectures.", *Nucleic Acids Res.*45(D)200-3.
- Maiti R, Van Domselaar G,H., Zhang H, Wishart DS. SuperPose: A simple server for sophisticated structural superposition. *Nucleic Acids Res* 2004;32:W590-4.

- Miura K, Jacques KM, Stauffer S, Kubosaki A, Zhu K, Hirsch DS, Resau J, Zheng Y, Randazzo PA. ARAP1: A point of convergence for arf and rho signaling. *Mol Cell* 2002;9(1):109-19.
- Olechnovič K, Venclovas Č. VoroMQA web server for assessing three-dimensional structures of proteins and protein complexes. *Nucleic Acids Res* 2019;47:W437-42
- Parasuraman S. Protein data bank. *Journal of Pharmacology & Pharmacotherapeutics* 2012;3(4):351-2.
- Randazzo PA, Hirsch DS. Arf GAPs: Multifunctional proteins that regulate membrane traffic and actin remodelling. *Cell Signal* 2004;16(4):401-13.
- Revathidevi S, Munirajan AK. Akt in cancer: Mediator and more. *Semin Cancer Biol* 2019;59:80-91.
- Rigsby RE, Parker AB. Using the PyMOL application to reinforce visual understanding of protein structure. *Biochem Mol Biol Educ* 2016 Sep 10;44(5):433-7.
- Sievers F, Higgins DG. Clustal omega, accurate alignment of very large numbers of sequences. *Methods Mol Biol* 2014;1079:105-16.
- SSP Visualization App [Internet]. Available from: <https://sspred.herokuapp.com/>.
- Uziela K, Shu N, Wallner B, Elofsson A. ProQ3: Improved model quality assessments using rosetta energy terms. *Sci Rep* 2016 Oct 4;6:33509.
- Wallner B, Fang H, Elofsson A. Automatic consensus-based fold recognition using pcons, ProQ, and pmodeller. *Proteins* 2003;53 Suppl 6:534-41.
- Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer T, A.P., Rempfer C, Bordoli L, et al. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res* 2018;46:W296-303.
- Wiederstein M, Sippl MJ. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 2007;35:W407-10.
- Xu D, Zhang Y. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins* 2012 Jul;80(7):1715-35.
- Yan R, Xu D, Yang J, Walker S, Zhang Y. A comparative assessment and analysis of 20 representative sequence alignment methods for protein structure prediction. *Scientific Reports* 2013;3(1):2619.
- Yoon H, Miura K, Cuthbert EJ, Davis KK, Ahvazi B, Casanova JE, Randazzo PA. ARAP2 effects on the actin cytoskeleton are dependent on Arf6-specific GTPase-activating-protein activity and binding to RhoA-GTP. *J Cell Sci* 2006;119(22):4650-66.

Zimmermann L, Stephens A, Nam S, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *Journal of Molecular Biology* 2018;430(15):2237-43.