Ph.D. Dissertation of Pharmacy

Structural Studies on Human FAM129B/MINERVA Protein Mediating Cell Invasion in Cancer

암세포 침윤에 영향을 끼치는 인간 유래 FAM129B/MINERVA 단백질의 구조 연구

February 2020

Graduate School of Pharmacy
Seoul National University
Pharmaceutical Biosciences Major

Hyunggu Hahn

Abstract

MINERVA (Melanoma invasion by Erk) also known as FAM129B is a member of FAM129 protein family only present in vertebrates. MINERVA is involved in key signaling pathways including EGFR/Erk and Wnt/β-catenin and found upregulated in many types of cancer to promote invasion. However exact function of the protein remain elusive. X-ray crystallographic methods were implemented in order to determine structure of MINERVA$^{9-553}$, lacking C-terminal flexible region. Trypsin-digestion was required before crystallization to obtain diffraction-quality crystals. While the N-terminal PH domain exhibits typical fold of a PH domains, amino acid composition at its ligand binding site directs affinity for specific lipids. The following helix-rich domain demonstrates a novel fold. The overall structure of MINERVA$^{9-553}$ hints possible plasma membrane association, which seem to be tightly regulated by its C-terminal flexible region. Elucidation of MINERVA$^{9-553}$ structure presents a new fold for an α-helix bundle domain and a mechanism for autoregulation of its PH domain to be pursued at the same time.

Keywords：MINERVA; FAM129B; NIBAN2; PH domain; Novel fold; Autoregulation

Student Number：2013－30516

Table of Contents

List of Figures

List of Tables

Abbreviations

ASU	Asymmetric unit
DTT	Dithiothreitol
ECL	Electrochemiluminescence
EGFR	Epidermal growth factor receptor
FAK	Focal adhesion kinase
IP_5	Inositol (2,3,4,5,6)-pentakisphosphate
Keap1	Kelch-like ECH-associated protein 1
LPA	Lysophosphatidic acid
LPC	Lysophosphocholine
MINERVA	Melanoma invasion by Erk
Nrf2	Nuclear factor-erythroid 2-related factor 2
PA	Phosphatidic acid
PC	Phosphatidylcholine
PCR	Polymerase chain reaction
PE	Phosphatidylethanolamine
PEG	Polyethylene glycol
PH domain	Pleckstrin homology domain
PISA	Proteins interfaces, surfaces, and assemblies
PS	Phosphatidylserine
PtdIns	Phosphatidylinositol
RMSD	Root-mean-square deviation
S1P	Sphingosine 1-phosphate
SAD	Single-wavelength anomalous diffraction
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SeMet	Selenomethionine
SPR	Surface plasmon resonance
TCEP	Tris(2-carboxyethyl)phosphine

Chapter 1. Introduction

1.1. Study Background

In human proteome, pleckstrin homology (PH) domain is present in nearly 300 proteins involved in diverse physiological processes. A typical PH domain comprises about 100 amino acid residues structured into a seven-stranded antiparallel β-sheet and a C-terminal α-helix (Yoon, et al., 1994, Nature, 1). The most well-known physiological property of PH domains are recognition of phosphoinositides with high affinity and specificity, which regulate localization of PH domain-containing proteins to and from intracellular membranes (Lemmon, 2007, Biochem Soc Symp, 2). PH domains recognizing phosphoinositide display polar surface charge distribution around its β1–β2 region for accommodation of negatively-charged head groups of phosphoinositides. On the other hand, PH domains are also capable of binding protein interaction partners, including G-protein coupled receptors (GPCRs), GTPase exchange factors (GEFs), and focal adhesion kinase (FAK) (Lemmon, 2004, Biochem Soc Trans, 3; Bar-Shavit, et al., 2016, Cell Cycle, 4). Phosphotyrosine-containing peptides and polyproline helixes may bind on pockets created by the β-sheet and the C-terminal α-helix of the PH domain (Scheffzek and Welti, 2012, FEBS Lett, 5). Considering its pivotal role in regulating protein subcellular localization, mutations within the PH domains can lead to aberrant cell signaling (Bar-Shavit, et al., 2016, Cell Cycle, 4).

FAM129B (Family with sequence similarity 129, member B), also known as MINERVA (Melanoma invasion by Erk), is a member of a small family of proteins that includes FAM129A and FAM129C that are present only in vertebrates.

Sequence analysis of the FAM129 family members reveal that they share a common PH domain on their N-terminal region with high sequence similarity (Old, et al., 2009, Mol Cell, 6), even though functions of their PH domain are not well understood. While FAM129A is an endoplasmic reticulum stress-induced protein that is upregulated in renal and thyroid cancer (Sun, et al., 2007, Biochem Biophys Res Commun, 7), FAM129C is a B-cell membrane protein that is overexpressed in chronic lymphocytic leukemia (Boyd, et al., 2003, Leukemia, 8).

Following the first identification of MINERVA as a target of ERK phosphorylation (Old, et al., 2009, Mol Cell, 6), several studies have identified that MINERVA is upregulated in many types of cancer, including breast, kidney, large intestine, lung and endometrial cancers as well as hematopoietic and central nervous system tumors (Bamford, et al., 2004, Br J Cancer, 9; Forbes, et al., 2015, Nucleic Acids Res, 10). Cancer patients overexpressing MINERVA exhibited bad prognosis and low survival rate (Zhou, et al., 2018, Onco Targets Ther, 11; Cheng, et al., 2019, EBioMedicine, 12). Overexpression of MINERVA was associated with activation of focal adhesion kinase (FAK) by facilitating phosphorylation of Tyr397 and Tyr925 residues on FAK (Zhou, et al., 2018, Onco Targets Ther, 11), which in turn promotes cell migration (Deramaudt, et al., 2011, Mol Biol Cell, 13). Similar observations were made from MINERVA-deficient mice that exhibited delayed rate of wound healing, suggesting critical role of MINERVA in promoting cell motility (Oishi, et al., 2012, J Biochem, 14).

MINERVA is a target of phosphorylation by both epidermal growth factor receptor (EGFR) and ERK, the central kinases in the ERK growth signaling pathway (Old, et al., 2009, Mol Cell, 6; Ji, et al., 2016, Proc Natl Acad Sci U S A, 15). EGFR is a receptor tyrosine kinase that phosphorylates MINERVA on its

Tyr593 residue in response to extracellular growth signaling molecule epidermal growth factor (EGF). MINERVA phosphorylated on Tyr593 is capable of direct binding to H-Ras and K-Ras for their constitutive activation by reducing association of Ras GTPase-activating proteins (GAPs) to inhibit Ras-bound GTP hydrolysis (Figure 1) (Ji, et al., 2016, Proc Natl Acad Sci U S A, 15). Constitutive activation of Ras proteins in turn activates downstream ERK for β-catenin-TCF/LEF transcriptional complex-mediated upregulation of cyclin D1, c-Myc, and PKM2 for promotion of Warburg effect, cell proliferation and tumorigenesis (Ji, et al., 2016, Proc Natl Acad Sci U S A, 15). ERK also phosphorylates the serine residues on the C-terminal flexible region of MINERVA as mentioned above, enhancing invasion capabilities of cells via an unknown mechanism (Old, et al., 2009, Mol Cell, 6). In sparse, exponentially growing HeLa cells, MINERVA is found in cytoplasm whereas the protein is recruited to plasma membrane and colocalizes with β-catenin at adherens junctions in growth-inhibited confluent HeLa cells (Figure 2) (Chen, et al., 2011, J Biol Chem, 16). At the adherens junctions, MINERVA seem to protect β-catenin from degradation by caspases, delaying onset of apoptosis elicited by the confluency in melanoma cells (Figure 3) (Chen, et al., 2011, J Biol Chem, 16). It was discovered that MINERVA interacts with LATS1 kinase that suppresses nuclear transportation of YAP, suggesting its involvement in Hippo signaling pathway activated in confluency (Figure 4) (Conrad, et al., 2013, F1000Res, 17). Overall, MINERVA governs crosstalk between essential cellular signaling pathways for cell proliferation, tumorigenesis, invasion, and apoptosis (Figure 5).
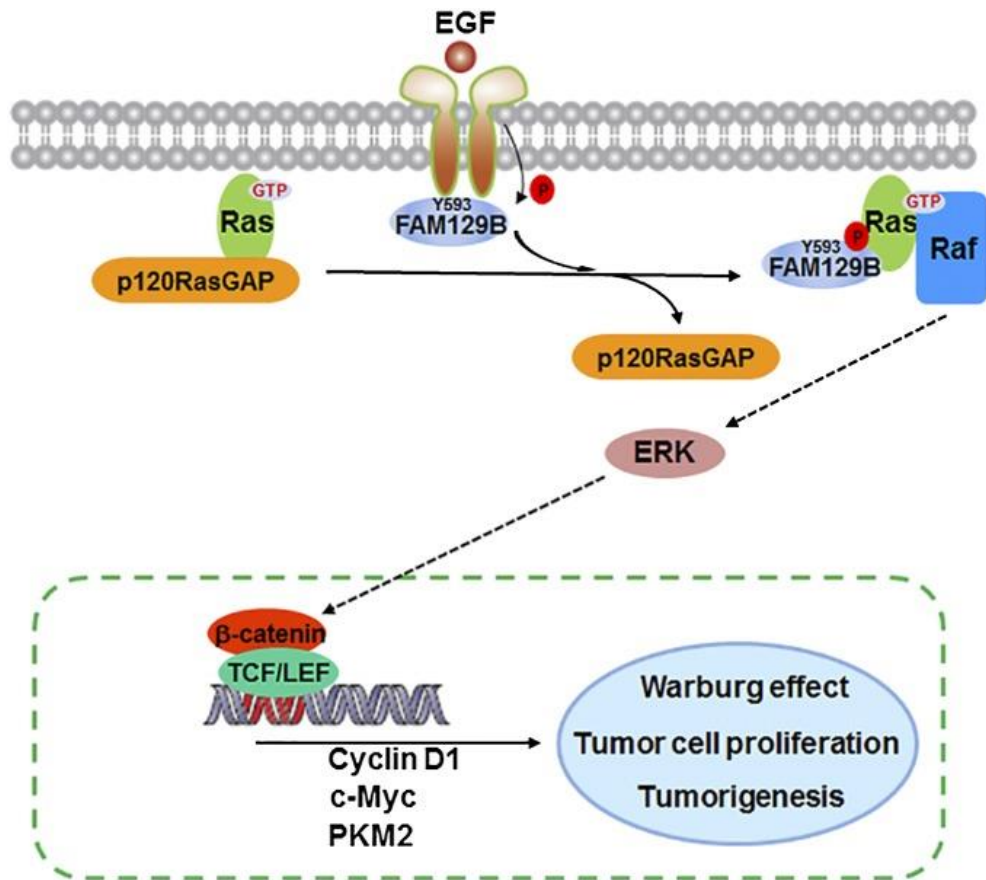
Figure 1. Proposed role of MINERVA (FAM129B) phosphorylated on Tyr593 residue by EGFR.

(Ji, et al., 2016, Proc Natl Acad Sci U S A, 15)
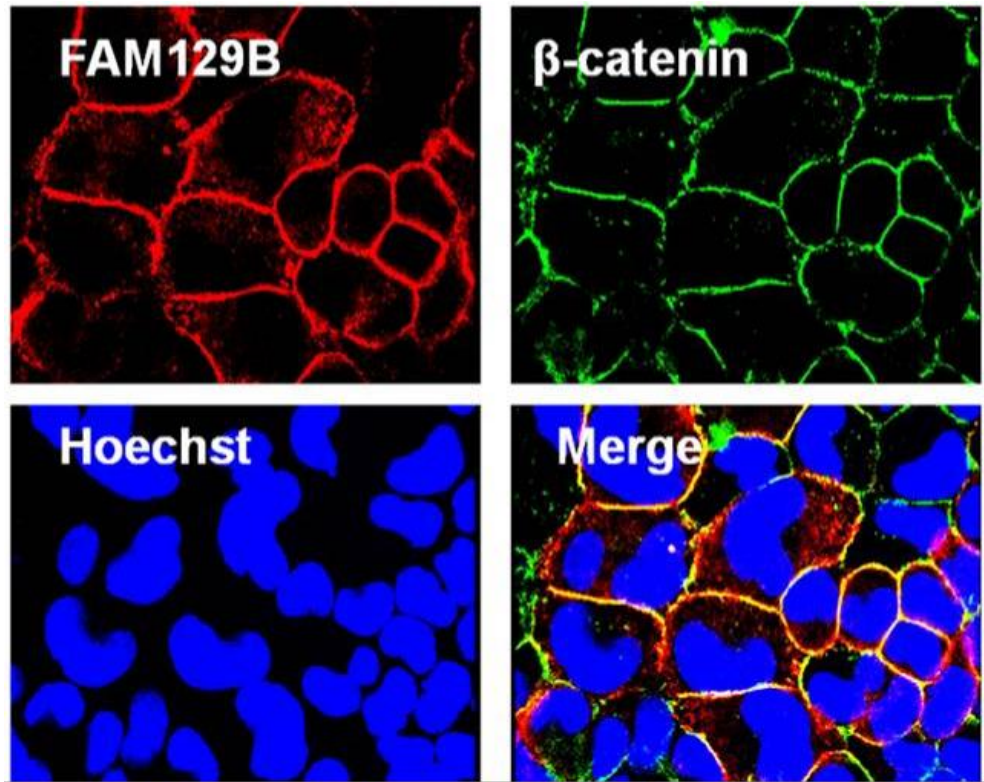
Figure 2. Colocalization of MINERVA (FAM129B) with β-catenin in confluent HeLa cells.

(Chen, et al., 2011, J Biol Chem, 16)
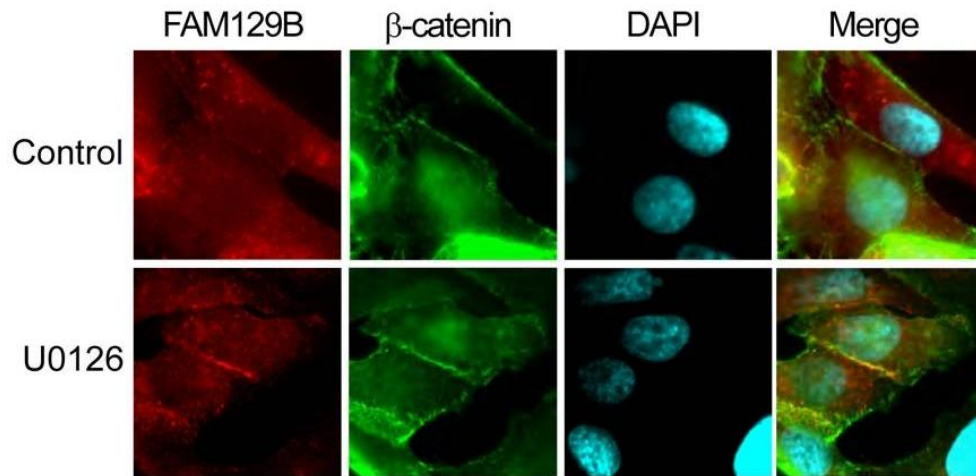
Figure 3. Colocalization of MINERVA (FAM129B) with β-catenin in U0126-treated WM115 melanoma cells.

(Old, et al., 2009, Mol Cell, 6)
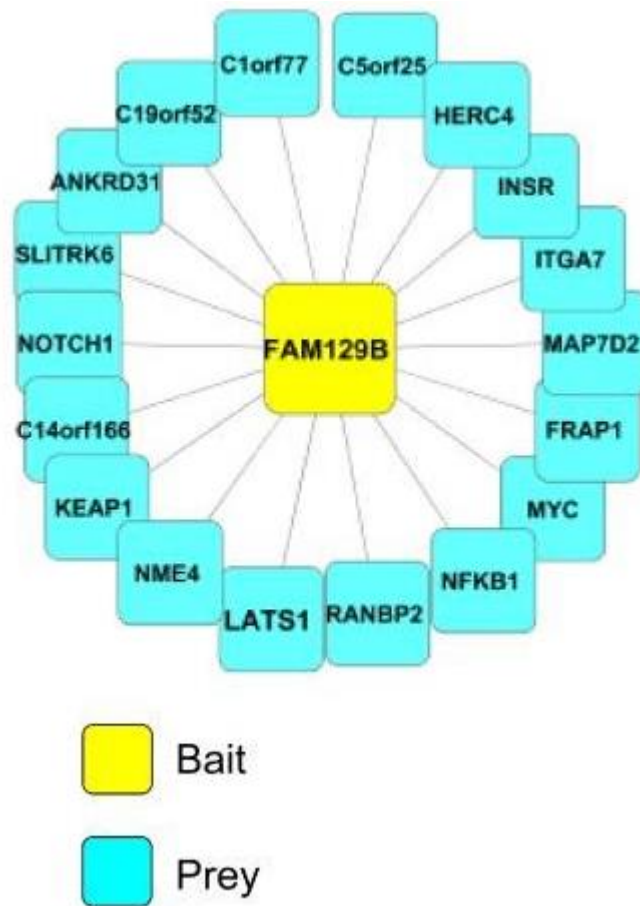
Figure 4. Interactors of MINERVA co-purified from affinity chromatography and analyzed by mass spectrometry.

(Conrad, et al., 2013, F1000Res, 17)
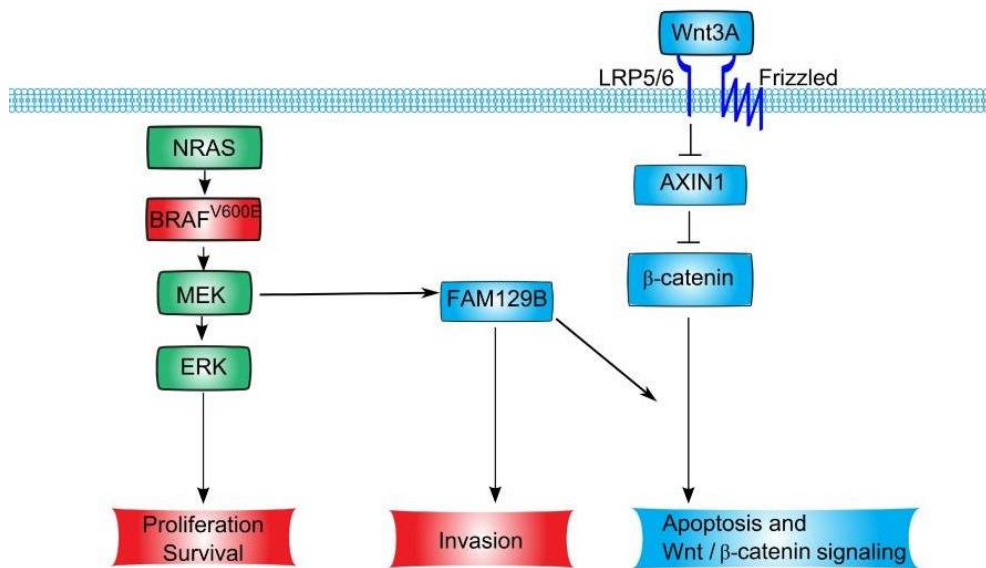
Figure 5. Proposed role of MINERVA (FAM129B) in Wnt/β-catenin signaling.

(Conrad, et al., 2013, F1000Res, 17)

Sequence and secondary structure analyses of 746 amino acid-long MINERVA protein indicates a PH domain on N-terminus (residues Asp66–Gly197) and a proline-rich region on C-terminus (residues Asp575–Phe746) which contains the phosphorylatable serine residues (Figure 6) (Old, et al., 2009, Mol Cell, 6). PH domains are well-known for binding phosphatidylinositols phosphates (PIPs) for regulating protein localization (Lemmon, 2007, Biochem Soc Symp, 2). Additionally, MINERVA contains an N-terminal myristoylation motif 1-MGX3S-6 (Thinon, et al., 2014, Nat Commun, 18), which supports intracellular membrane localization of MINERVA (Chen, et al., 2011, J Biol Chem, 16). Within the C-terminal flexible region, there is a 708-DLGX7ETGE-721 motif that interacts with Kelch domain on Kelch-like ECH-associated protein 1 (Keap1), which abrogates Keap1-mediated ubiquitination of nuclear factor-erythroid 2-related factor 2 (Nrf2) for proteasomal degradation (Figure 7) (Cheng, et al., 2019, EBioMedicine, 12). Nrf2 is a key regulator in cellular response against oxidative stress by regulating transcription of antioxidant gene expression that is targeted for degradation under quiescent conditions (Figure 8) (Furukawa and Xiong, 2005, Mol Cell Biol, 19; Kensler, et al., 2007, Annu Rev Pharmacol Toxicol, 20). Hence, MINERVA acts as an upstream regulator of Nrf2-mediated response to oxidative stress in cancer by hindering interaction of Keap1 and Nrf2 by competing for the Kelch domain binding through the the DLGXnETGE motif.
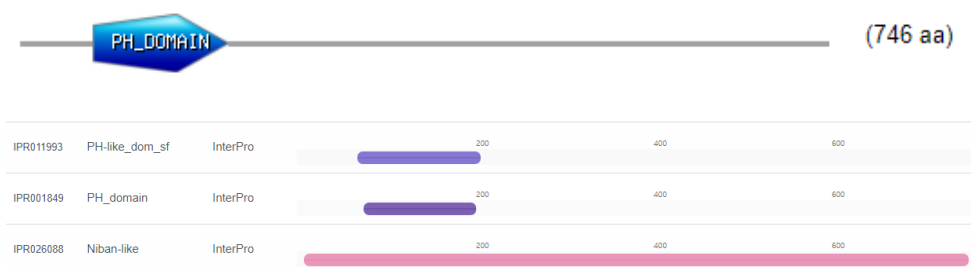
Figure 6. Protein domain analysis of MINERVA sequence indicates presence of a PH domain on N-terminus.

Top, sequence analysis result from *Prosite* (https://prosite.expasy.org/index.html). Bottom, sequence analysis result from *InterPro* (https://www.ebi.ac.uk/interpro/).
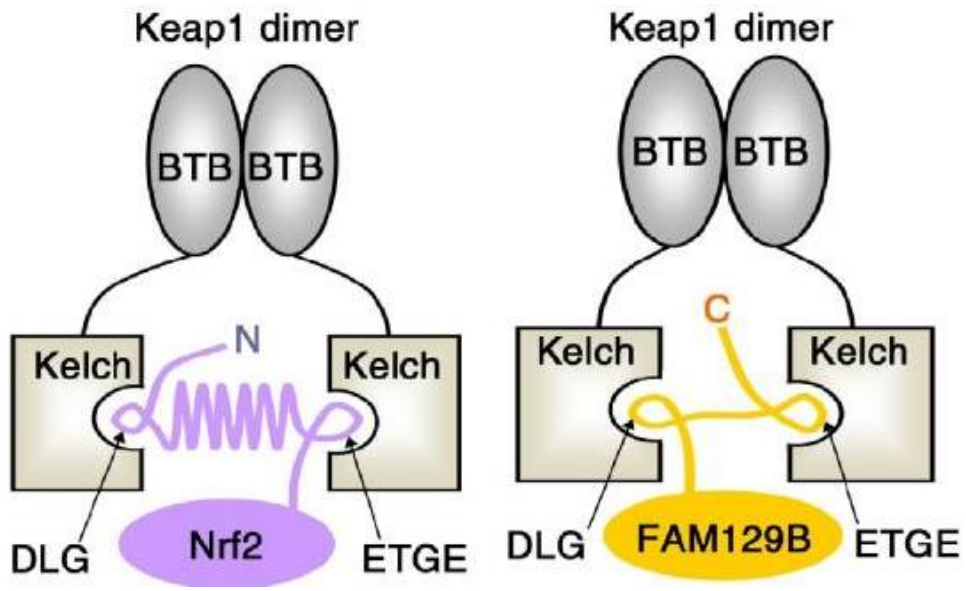
Figure 7. MINERVA (FAM129B) competes with Nrf2 for Keap1 binding through "DLGX$_n$ETGE" motif.

(Cheng, et al., 2019, EBioMedicine, 12)
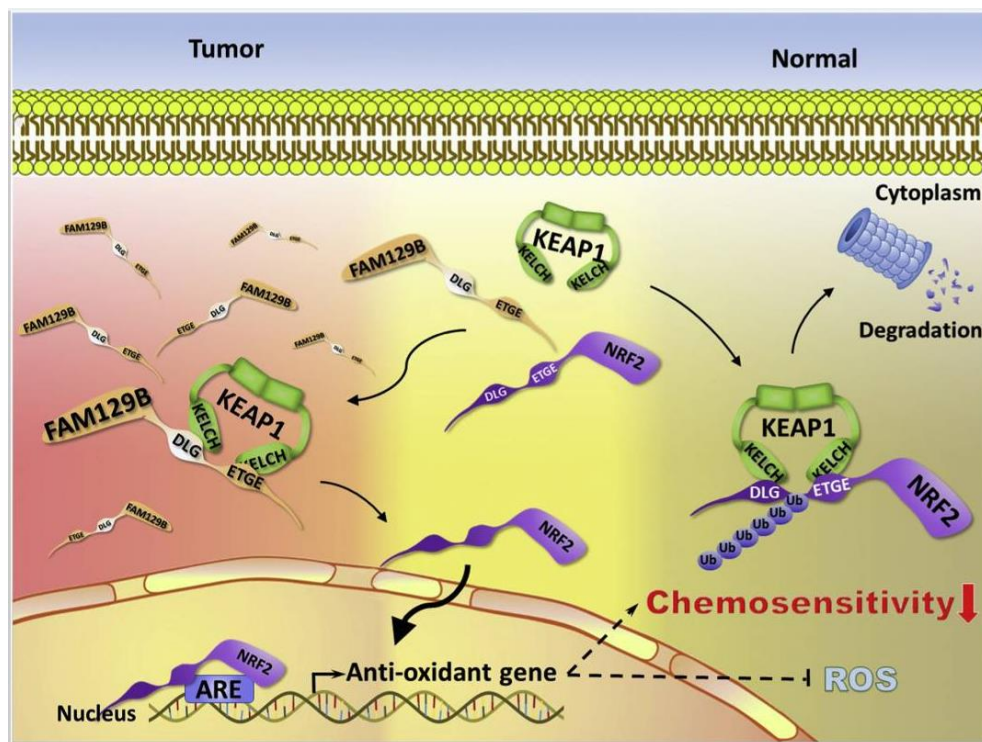
Figure 8. Proposed model of MINERVA (FAM129B) function in regulation of oxidative stress.

(Cheng, et al., 2019, EBioMedicine, 12)

1.2. Purpose of Research

Despite the importance of MINERVA in regulating cell growth, migration, apoptosis, and oxidative stress, no structural information of the protein was available. Here, I report the first crystal structure of MINERVA9–553, which contains the PH domain and a helix bundle domain with a novel fold. In addition, lipid binding ability of the PH domain is assessed for structural and functional analyses of MINERVA. This study shows structural basis for membrane association of MINERVA and its importance in governing multiple signaling pathways in cancer progression.

Chapter 2. Materials and methods

2.1. Cloning, expression, and purification of MINERVA constructs

Sequence analysis result of MINERVA using various web servers such as *PSIPRED*, *DISOPRED*, and *XtalPred* (Ward, et al., 2004, Bioinformatics, 21; McGuffin, et al., 2000, Bioinformatics, 22; Slabinski, et al., 2007, Bioinformatics, 23) predicted that the C-terminal region of MINERVA (Asp575–Phe746) is disordered, while the N-terminal PH domain and the following helix-rich region is well-conserved among the members of the FAM129 family (Figure 9, Figure 10, and Figure 11). The full length MINERVA^{FL} (residues Met1–Phe746) and N- and C-terminally truncated human MINERVA$^{9–553}$ (residues Leu9–Val553) were amplified using PCR and cloned into pET-28a(+) vectors (Novagen) to each contain an N-terminal His$_6$-tags. The plasmids containing MINERVA constructs were each transformed into Rosetta 2(DE3)pLysS *Escherichia coli* strain. C-terminally truncated human MINERVA$^{1–574}$ (residues Met1–Ser574) was amplified using PCR and was cloned into pET-21a(+) vector (Novagen) to contain an C-terminal His$_6$-tag. The MINERVA$^{1–574}$ containing plasmid was transformed into BL21-CodonPlus (DE3)-RIPL *E. coli* strain. The C-terminal flexible region MINERVA$^{575–746}$ (residues Asp575–Phe746) were amplified using PCR and cloned into pET-28a(+) vector to contain an N-terminal maltose binding protein (MBP)-tag. The MINERVA$^{575–746}$ containing plasmid was transformed into BL21-CodonPlus (DE3)-RIPL *E. coli* strain.

```
Conf: ███████████████████████████████████████████
Pred: ─────────────⬭━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━
Pred: CCCCCCCCCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
  AA: MGDVLSTHLDDARRQHIAEKTGKILTEFLQFYEDQYGVAL
                10        20        30        40

Conf: ███████████████████████████████████████████
Pred: ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━⟹━━━━━━━━━
Pred: HHHHHHHHCCCCCCCCCCCCCCCCCCCCCCEEEECCCCCCCCC
  AA: FNSMRHEIEGTGLPQAQLLWRKVPLDERIVFSGNLFQHQE
                50        60        70        80

Conf: ███████████████████████████████████████████
Pred: ─────────⟹━━━━━⟹━━━⬭━━━━━━━━━⟹━━━
Pred: CCCCCCCEEEEECCCCEEEECCCHHHHCCCCCCCEEEECC
  AA: DSKKWRNRFSLVPHNYGLVLYENKAAYERQVPPRAVINSA
                90        100       110       120

Conf: ███████████████████████████████████████████
Pred: ──⟹━⬭━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━■
Pred: CCEEECCHHHHHHHHHCCCCCCCCCCCCCCCCCCCCCCCCCE
  AA: GYKILTSVDQYLELIGNSLPGTTAKSGSAPILKCPTQFPL
                130       140       150       160

Conf: ███████████████████████████████████████████
Pred: ⟹━━━━━━━━━━━━━⬭━━━━━━━━━━━━━━━━━━━━━━
Pred: EEECCCCCCCCCCCCCCHHHHHHHHHHHHHHHHHHCCCCCCC
  AA: ILWHPYARHYYFCMMTEAEQDKWQAVLQDCIRHCNNGIPE
                170       180       190       200

Conf: ███████████████████████████████████████████
Pred: ──────⬭━━━━━━━━━━━━━━━━━━━━━━⬭━━━━━━
Pred: CCCCCCHHHHHHHHHHHHHHCCCCCCCCCCCCCCHHHHHHHH
  AA: DSKVEGPAFTDAIRMYRQSKELYGTWEMLCGNEVQILSNL
                210       220       230       240

Conf: ███████████████████████████████████████████
Pred: ━━━━━━━━━━━━━⬭━━━━━━━━━━━━━━━━━━━━━━━
Pred: HHHHHHHHHHHHHCCCCCCCHHHHHHHHHHHHHHHHHHHHHH
  AA: VMEELGPELKAELGPRLKGKPQERQRQWIQISDAVYHMVY
                250       260       270       280
```

15

```
Conf: ]███████████████████▄▄▄▄██▄█▄▄▄▄▄▄▄▄▄▄▄▄▄▄▄███▄[
Pred: ⬤────────────────⬤───⬤────────────
Pred: HHHHHHHHHHHHHHHHHCHHHHHCCCCHHHHHHHHHHHH
  AA: EQAKARFEEVLSKVQQVQPAMQAVIRTDMDQIITSKEHLA
              290       300       310       320

Conf: ]██▄▄██▄▄▄▄▄▄▄▄▄▄▄▄▄▄▄███▄█▄████▄▄▄▄██▄▄▄██[
Pred: ──────────⬤──────⬤──────⬤────
Pred: HHHHHHCCCHHHHHHHCCCCCCCHHHHHHHHHCCCCCHHHHH
  AA: SKIRAFILPKAEVCVRNHVQPYIPSILEALMVPTSQGFTE
              330       340       350       360

Conf: ]████████████▄▄▄▄▄▄▄▄▄▄██▄▄▄▄██████▄▄▄████▄[
Pred: ──────────────⬤────────────
Pred: HHHHHHHHHHHHHCCCCCCCCCCCHHHHHHHHHHHCCCCCCC
  AA: VRDVFFKEVTDMNLNVINEGGIDKLGEYMEKLSRLAYHPL
              370       380       390       400

Conf: ]▄▄██████████▄▄▄▄▄▄▄▄▄▄▄▄▄▄▄▄▄▄▄▄███████████▄[
Pred: ⬤────────⬤─────────────
Pred: HHHHHHHHHHHHCCCCCCCCCCCCCCCCCHHHHHHHHHHHHHHH
  AA: KMQSCYEKMESLRLDGLQQRFDVSSTSVFKQRAQIHMREQ
              410       420       430       440

Conf: ]▄▄▄▄█████████▄▄▄▄▄▄████▄█████████████▄▄▄▄██[
Pred: ──────────⬤────────────
Pred: HHHHHHHHHHHHHHHHHCCCCCHHHHHHHHHHHHHHHHHHHHC
  AA: MDNAVYTFETLLHQELGKGPTKEELCKSIQRVLERVLKKY
              450       460       470       480

Conf: ]███▄████████████▄▄▄▄▄▄▄██████▄▄▄█████████▄▄▄█[
Pred: ──⬤────────────────⬤─
Pred: CCCCHHHHHHHHHHHHHHHHHHHHHHHHHCCCCCCCCCCHHH
  AA: DYDSSSVRKRFFREALLQISIPFLLKKLAPTCKSELPRFQ
              490       500       510       520

Conf: ]██████▄▄▄▄▄▄▄▄▄▄███████████████████████████▄▄▄█[
Pred: ────────⬤──────────────
Pred: HHHHHHHCCCCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHC
  AA: ELIFEDFARFILVENTYEEVVLQTVMKDILQAVKEAAVQR
              530       540       550       560
```

16

```
Conf: }███████████████████████████████████████████E
Pred: _____
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
  AA: KHNLYRDSMVMHNSDPNLHLLAEGAPIDWGEEYSNSGGGG
            570       580       590       600

Conf: }███████████████████████████████████████E
Pred: _____      ⬭▬▬▬▬▬⬭      _____
Pred: CCCCCCCCCCCHHHHHHCCCCCCCCCCCCCCCCCCCCCCCC
  AA: SPSPSTPESATLSEKRRRAKQVVSVVQDEEVGLPFEASPE
            610       620       630       640

Conf: }███████████████████████████████████████E
Pred: _____
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
  AA: SPPPASPDGVTEIRGLLAQGLRPESPPPAGPLLNGAPAGE
            650       660       670       680

Conf: }███████████████████████████████████████E
Pred: _____
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
  AA: SPQPKAAPEASSPPASPLQHLLPGKAVDLGPPKPSDQETG
            690       700       710       720

Conf: }██████████████████████████████E
Pred: 
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
  AA: EQVSSPSSHPALHTTTEDSAGVQTEF
            730       740
```

Legend:
⬭▬▬▬ = helix   Conf: }▪▫▪▪▪E = confidence of prediction
                       − +
➡ = strand   Pred: predicted secondary structure
─── = coil   AA: target sequence

Figure 9. *PSIPRED* analysis of MINERVA sequence showing predicted secondary structures.

MINERVA contain structured regions and C-terminal disordered region.

Figure 10. *DISOPRED* **analysis of MINERVA sequence showing C-terminal disordered region.**

The C-terminal region approximately from residue number 550 is predicted to be intrinsically disordered.

```
1...*...10....*...20....*...30....*...40....*...50....*...60....*...70....*...80....*...90....*...100
MGDVLSTHLDDARRQHIAEKTGKILTEFLQFVEDQVGVALFNSMRHEIEGTGLPQAQLLWRKVPLDERIVFSGNLFQHQEDSKKWRNRFSLVPHNYGLVL

....*...110....*...120....*...130....*...140....*...150....*...160....*...170....*...180....*...190....*...200
VENKAAVERQVPPRAVINSAGYKILTSVDQYLELIGNSLPGTTAKSGSAPILKCPTQFPLILWHPYARHVYFCMMTEAEQDKWQAVLQDCIRHCNNGIPE

....*...210....*...220....*...230....*...240....*...250....*...260....*...270....*...280....*...290....*...300
DSKVEGPAFTDAIRMVRQSKELYGTWEMLCGNEVQILSNLVMEELGPELKAELGPRLKGKPQERQRQWIQISDAVVHMVYEQAKARFEEVLSKVQQVQPA

....*...310....*...320....*...330....*...340....*...350....*...360....*...370....*...380....*...390....*...400
MQAVIRTDMDQIITSKEHLASKIRAFILPKAEVCVRNHVQPYIPSILEALMVPTSQGFTEVRDVFFKEVTDMNLNVINEGGIDKLGEYMEKLSRLAVHPL

....*...410....*...420....*...430....*...440....*...450....*...460....*...470....*...480....*...490....*...500
KMQSCYEKMESLRLDGLQQRFDVSSTSVFKQRAQIHMREQMDNAVYTFETLLHQELGKGPTKEELCKSIQRVLERVLKKYDYDSSSVRKRFFREALLQIS

....*...510....*...520....*...530....*...540....*...550....*...560....*...570....*...580....*...590....*...600
IPFLLKKLAPTCKSELPRFQELIFEDFARFILVENTYEEVVLQTVMKDILQAVKEAAVQRKHNLYRDSMVMHNSDPNLHLLAEGAPIDWGEEYSNSGGGG

....*...610....*...620....*...630....*...640....*...650....*...660....*...670....*...680....*...690....*...700
SPSPSTPESATLSEKRRRAKQVVSVVQDEEVGLPFEASPESPPPASADGVTEIRGLLAQGLRPESPPPAGPLLNGAPAGESPQPKAAPEASSPPASALQH

....*...710....*...720....*...730....*...740....*.
LLPGKAVDLGPPKPSDQETGEQVSSPSSHPALHTTTEDSAGVQTEF
```

Legend

LOOP	loop secondary structure predicted by PSIPRED
HELIX	helix secondary structure predicted by PSIPRED
STRAND	strand secondary structure predicted by PSIPRED
DISORDER	disordered region predicted by DISOPRED2
LOW COMPLEXITY	low complexity region predicted by SEG
COILS	coiled coils region predicted by COILS
TRANSMEMBRANE HELICES	transmembrane helices predicted by TMHMM
SIGNAL PEPTIDES	signal peptites predicted by SignalP

Figure 11. *XtalPred* **analysis of MINERVA sequence showing helix-rich composition and the C-terminal disordered region.**

The C-terminal region approximately from residue number 550 is predicted to be intrinsically disordered.

The transformed cells were grown for 16 h in Luria Broth media without induction to be used as seeds for the scale-up culture. The uninduced *E. coli* seeds containing MINERVA plasmids were inoculated into a scale-up culture media at a ratio of 1:50. The cells in scale-up culture were grown in Luria Broth media and induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside, followed by further incubation at 20 °C for 16 h. Harvested cells were ly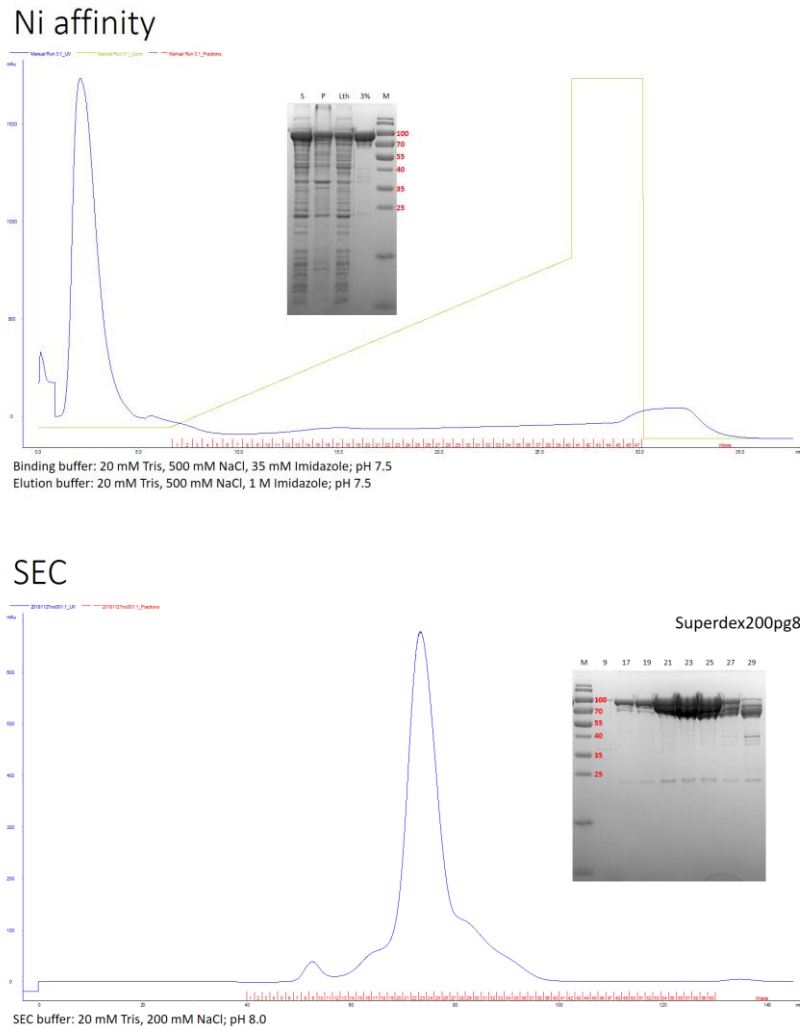sed by a cell sonicator (SONICS) in a lysis buffer containing 20 mM Tris-HCl (pH 7.5), 500 mM NaCl, 35 mM imidazole, and 1 mM phenylmethylsulfonyl fluoride. Cell debris was removed by centrifugation at 35,000 × g for 50 min at 4 °C, then the supernatant was filtered with 0.22 μm filter to remove cell debris and any aggregated proteins. The filtered supernatant was applied onto a HiTrap Chelating HP column (GE Healthcare) equilibrated with the lysis buffer for affinity chromatography. Unbound or weakly bound proteins were removed from the HiTrap Chelating HP column with six column volume elutions of the washing buffer containing 20 mM Tris-HCl (pH 7.5), 500 mM NaCl, and 50 mM imidazole. The retained proteins were eluted with a gradually increasing addition of a buffer containing 20 mM Tris-HCl (pH 7.5), 500 mM NaCl, and 500 mM imidazole. The eluted protein samples were further purified by size-exclusion chromatography with HiLoad 16/600 Superdex200 pg column (GE Healthcare) equilibrated with a buffer containing 20 mM Tris-HCl (pH 8.0) and 200 mM NaCl (Figure 12, Figure 13, and Figure 14).

The purification steps for SeMet-substituted MINERVA[1–574] were as above except that the buffer for the final size exclusion chromatography additionally contained 1 mM tris(2-carboxyethyl)phosphine (Figure 15).

Ni affinity

Binding buffer: 20 mM Tris, 500 mM NaCl, 35 mM Imidazole; pH 7.5
Elution buffer: 20 mM Tris, 500 mM NaCl, 1 M Imidazole; pH 7.5

SEC

SEC buffer: 20 mM Tris, 200 mM NaCl; pH 8.0

Figure 12. Purification profiles of the full length MINERVA.

The full length MINERVA was purified through nickel affinity chromatography
(top) followed by size exclusion chromatography (bottom) to high purity, as
analyzed by SDS-PAGE results (insets).

Ni affinity

Binding buffer: 20 mM Tris, 500 mM NaCl, 35 mM Imidazole; pH 7.5
Elution buffer: 20 mM Tris, 500 mM NaCl, 1 M Imidazole; pH 7.5

SEC

SEC buffer: 10 mM HEPES, 200 mM NaCl; pH 8.0
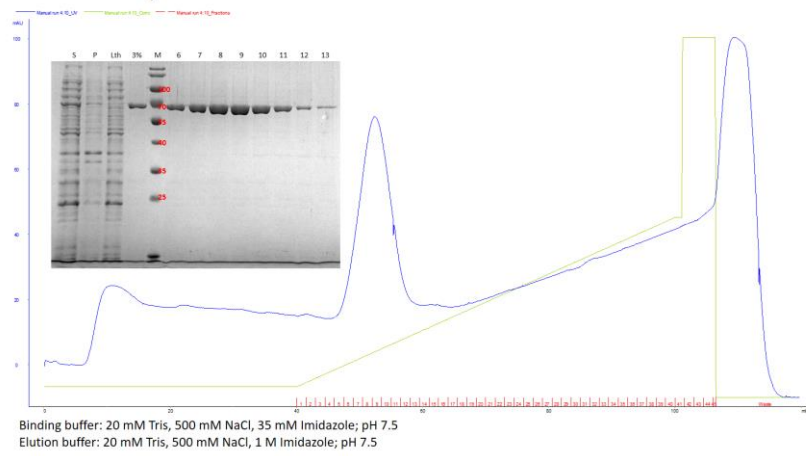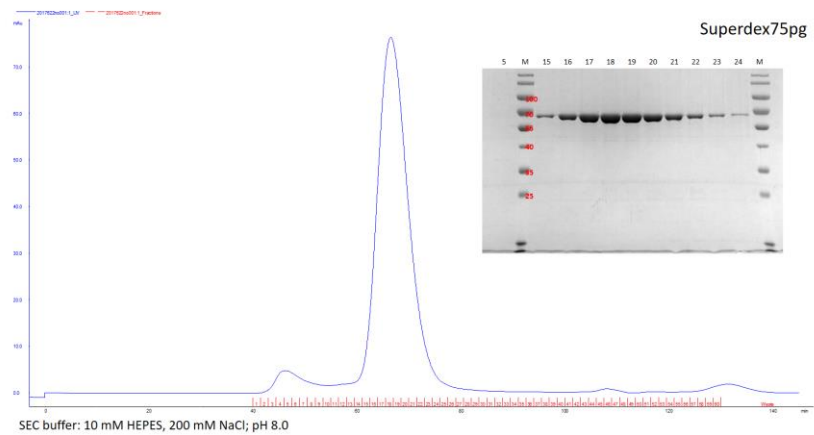
Figure 13. Purification profiles of MINERVA[1–574].

MINERVA[1-574] was purified through nickel affinity chromatography (top) followed by size exclusion chromatography (bottom) to high purity, as analyzed by SDS-PAGE results (insets).

Ni affinity

Binding buffer: 20 mM Tris, 500 mM NaCl, 35 mM Imidazole; pH 7.5
Elution buffer: 20 mM Tris, 500 mM NaCl, 1 M Imidazole; pH 7.5

SEC

Superdex200pg8

SEC buffer: 20 mM Tris, 200 mM NaCl; pH 7.9

Figure 14. Purification profiles of MINERVA[9–553].

MINERVA[9–553] was purified through nickel affinity chromatography (top) followed by size exclusion chromatography (bottom) to high purity, as analyzed by SDS-PAGE results (insets).

23

Ni affinity

Binding buffer: 20 mM Tris, 500 mM NaCl, 35 mM Imidazole, 3 mM β-ME; pH 7.5
Elution buffer: 20 mM Tris, 500 mM NaCl, 1 M Imidazole, 3 mM β-ME; pH 7.5

SEC

SEC buffer: 20 mM Tris, 200 mM NaCl, 1 mM TCEP; pH 8.0
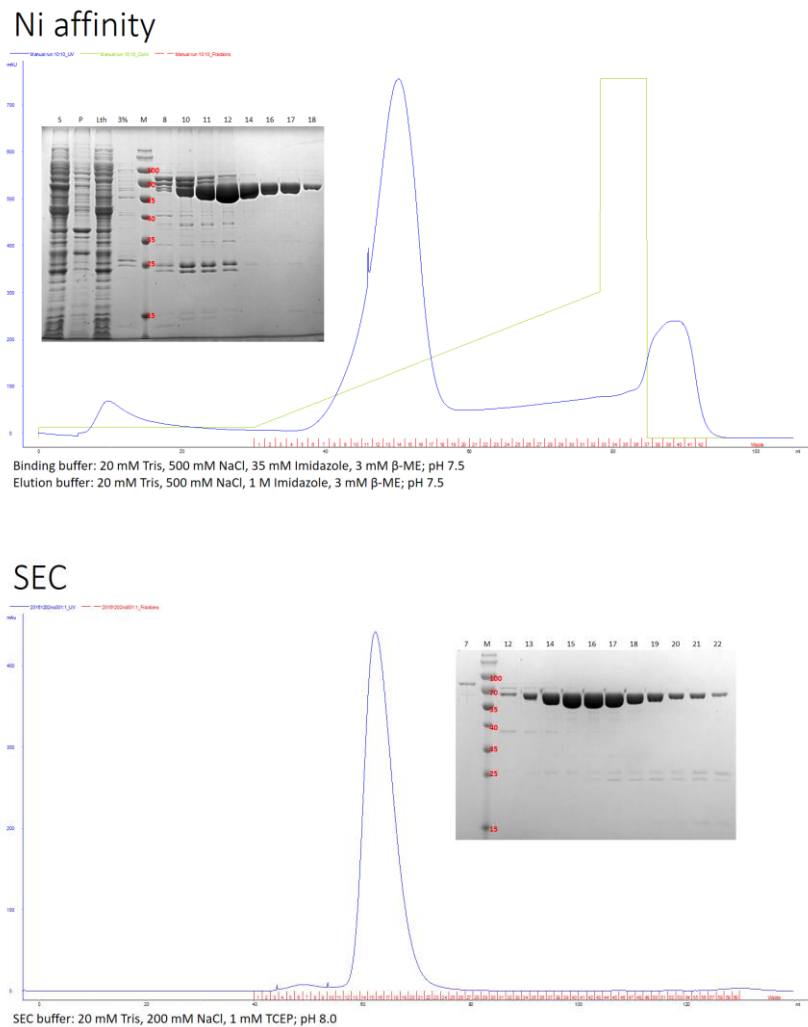
Figure 15. Purification profiles of the SeMet-substituted MINERVA[1–574].

The SeMet-substituted MINERVA[1-574] was purified through nickel affinity chromatography (top) followed by size exclusion chromatography (bottom) to high purity, as analyzed by SDS-PAGE results (insets).

2.2. Trypsin digestion of MINERVA constructs

Purified MINERVA proteins were subjected to digestion by trypsin (Sigma-Aldrich) by incubating MINERVA at a 5,000:1 trypsin ratio at 4°C for 20 h prior to crystallization in order to improve diffraction quality (Figure 16). Digested proteins were further purified by size-exclusion chromatography with HiLoad 16/600 Superdex200 pg column (GE Healthcare)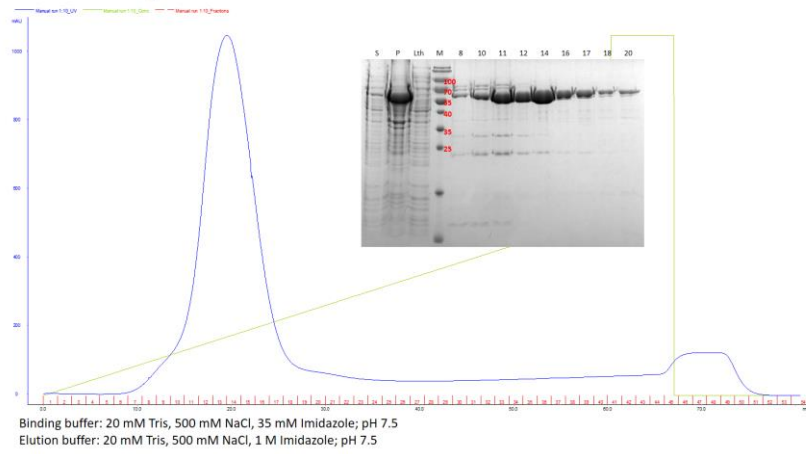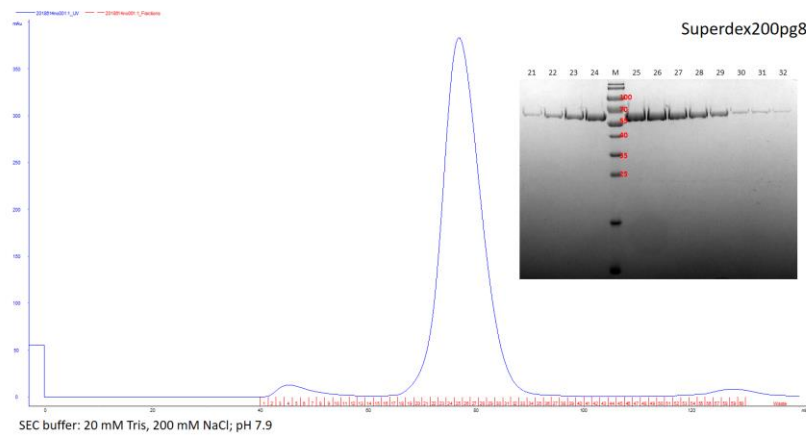 equilibrated with a buffer containing 20 mM Tris-HCl (pH 8.0) and 200 mM NaCl (Figure 17). The purified MINERVA proteins were concentrated to 4.8 mg mL^{-1} for crystallization.
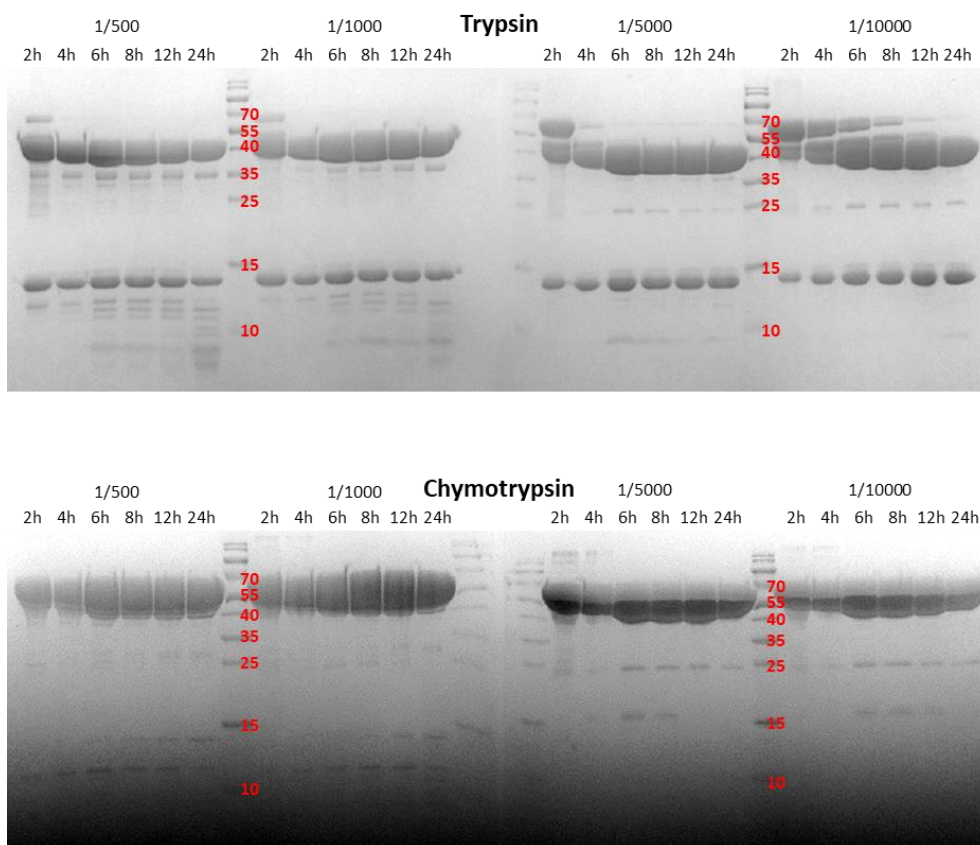
Figure 16. Preliminary trypsin-digestion of MINERVA[1–574].

MINERVA[1–574] was incubated with trypsin (top) or chymotrypsin (bottom) for determination of best molar ratio and incubation time of trypsin-digestion. The digests were analyzed by SDS-PAGE.
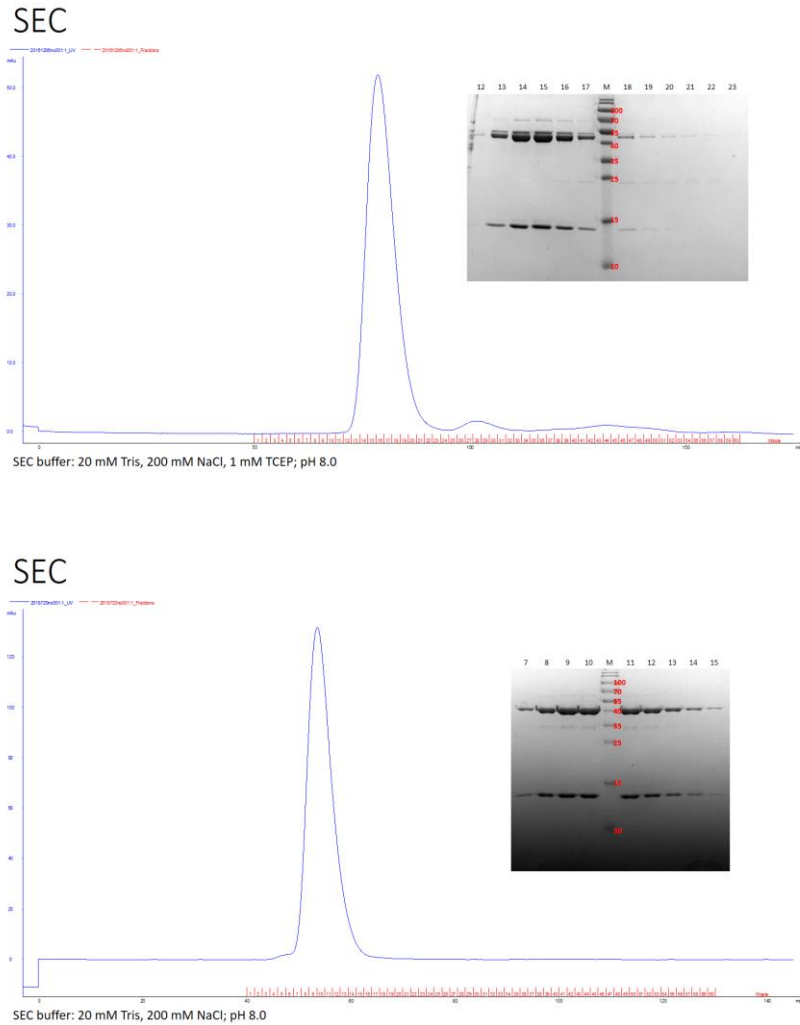
Figure 17. Purification profiles of trypsin-digested MINERVA proteins.

Trypsin-digested MINERVA[1–574] (top) and MINERVA[9–553] (bottom) were purified by size exclusion chromatography to high purity, as analyzed by SDS-PAGE (insets).

2.3. Crystallization, data collection and structure determination of MINERVA constructs
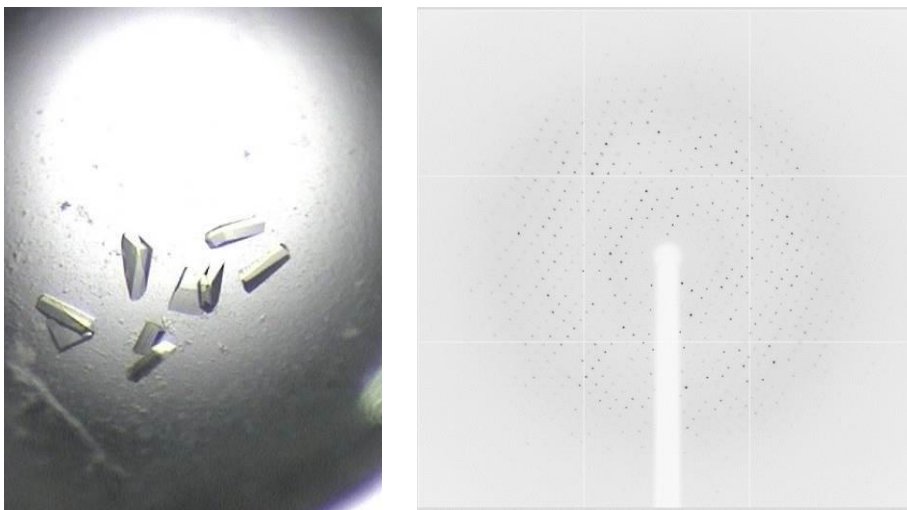
Crystals of the SeMet-substituted MINERVA[1–574] and the native MINERVA[9–553] were both grown at 22 °C by the hanging drop vapor diffusion method by mixing equal volumes of the protein at 4.8 mg mL^{-1} and a crystallization solution containing 200 mM sodium tartrate and 20% (w/v) PEG 3,350. Crystals of diffraction quality were cryoprotected with the reservoir solution containing additional 25% (v/v) glycerol and flash-cooled in a nitrogen gas stream at 100 K. Diffraction data from the SeMet-substituted and the native crystals were collected up to 2.385 Å and 2.600 Å, respectively (Figure 18 and Figure 19). Raw X-ray diffraction data were indexed and scaled using the *HKL-2000* program suite (Otwinowski and Minor, 1997, Methods Enzymol, 24). Single-wavelength anomalous diffraction (SAD) phases of the SeMet-substituted MINERVA[1–574] were initially calculated with *AUTOSOL* in the *PHENIX* software suite (Adams, et al., 2010, Acta crystallographica Section D, Biological crystallography, 25) and further improved by the automatic model building program *RESOLVE* (Terwilliger, 2003, Acta crystallographica Section D, Biological crystallography, 26), resulting in an initial model. The initial model was further refined to the final model using iterative cycles of model building with *Coot* (Emsley, et al., 2010, Acta crystallographica Section D, Biological crystallography, 27) and subsequent refinement with *Refmac5* in the *CCP4* suite (Murshudov, et al., 1997, Acta crystallographica Section D, Biological crystallography, 28) and *phenix.refine* (Afonine, et al., 2012, Acta crystallographica Section D, Biological crystallography, 29). The crystal structure of the native MINERVA[9–553] was determined by

molecular replacement with the *MolRep* program (Vagin and Teplyakov, 2010, Acta crystallographica Section D, Biological crystallography, 30), using the refined structures of SeMet-substituted MINERVA[1–574] as a phasing model. Validation of crystal structures was implemented with *MolProbity* (Chen, et al., 2010, Acta crystallographica Section D, Biological crystallography, 31) and the Research Collaboratory for Structural Bioinformatics Protein Data Bank (PDB) Validation server. Statistics for the data collection and refinement are summarized in Table 1. The coordinates and structure factors of the native MINERVA[9–553] has been deposited in the Protein Data Bank (http://www.rcsb.org) under ID codes 6L9A.

Figure 18. Representative crystals and X-ray diffraction pattern of MINERVA[1–574].

Left, Crystals of SeMet-substituted, trypsin-digested MINERVA[1–574] were grown under crystallization buffer containing 200 mM sodium tartrate and 20% (w/v) PEG 3,350. Right, a diffraction pattern of the SeMet-substituted, trypsin-digested MINERVA[1–574] crystal diffraction up to 2.385 Å resolution.
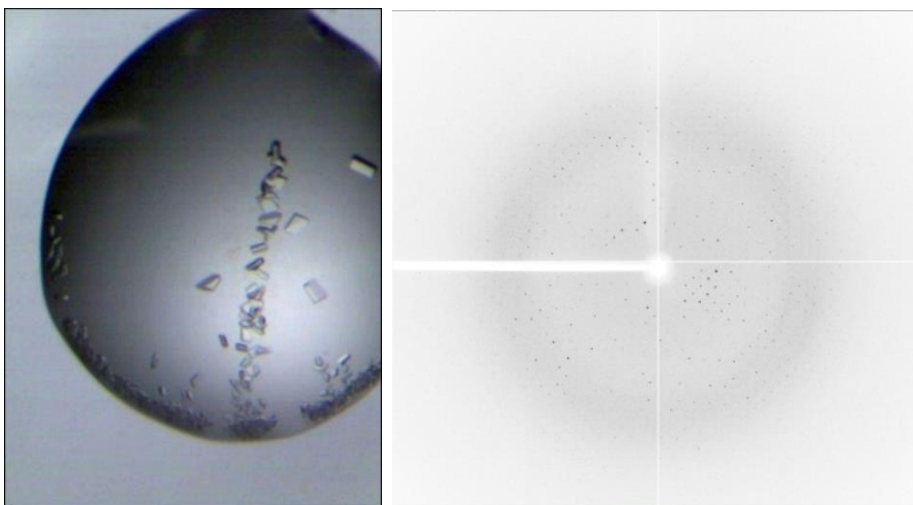
Figure 19. Representative crystals and X-ray diffraction pattern of MINERVA[9–553].

Left, Crystals of trypsin-digested MINERVA[9–553] were grown under crystallization buffer containing 200 mM sodium tartrate and 20% (w/v) PEG 3,350. Right, a diffraction pattern of the trypsin-digested MINERVA[9–553] crystal diffraction up to 2.600 Å resolution.

Table 1. Diffraction data collection and refinement statistics of the MINERVA[9–553] crystal.

PDB entry	6L9A
Diffraction source	PLS-5C
Wavelength (Å)	0.9793
Temperature (K)	100
Space group	$P2_1$
a, b, c (Å)	55.57, 55.46, 201.39
α, β, γ (°)	90.00, 89.99, 90.00
Resolution range (Å)	48.66–2.60 (2.69–2.60)
Total No. of reflections	
No. of unique reflections	36,957
Completeness (%)	96.4 (71.5)
Redundancy	3.8 (3.8)
$\langle I/\sigma(I) \rangle$	12.0 (1.8)
R_{meas}	0.125 (0.845)
R_{pim}	0.064 (0.432)
Overall B factor from Wilson plot (Å^2)	42.0
Resolution range (Å)	48.66–2.60
No. of reflections, working set	36,932 (2,716)
No. of reflections, test set	1,812 (124)
Final R_{work}	0.226 (0.298)
Final R_{free}	0.262 (0.333)
No. of non-H atoms	
Protein	8,790
Water	114
R.m.s. deviations	
Bonds (Å)	0.005
Angles (°)	1.050
Average B factors (Å^2)	
Protein	49.42
Water	39.59
Ramachandran plot	
Most favored (%)	98.95
Allowed (%)	1.05

2.4. Lipid dot blot assay

The PIP strip (Echelon Biosciences) was blocked with phosphate-buffered saline containing 0.5% (v/v) Tween® 20 (Amersham Biosciences) and 3% (w/v) bovine serum albumin (BSA; Bovogen) for 1 h and incubated with MINERVA[1–574] or the full length MINERVA for 1 h. The strip was washed with the blocking buffer for 5 min three times, before treating horseradish peroxidase (HRP)-conjugated α-His$_6$-tag antibody (Santa Cruz Biotechnology) for 1 h. After extensive wash with the blocking buffer, the strip was visualized with ECL kit (Amersham Biosciences).

2.5. Cloning, expression, purification, and crystallization of H-Ras[1−166]

H-Ras[1−166] was cloned, expressed and purified to assess interaction of H-Ras with phosphorylated Tyr593 residue (Lee, et al., 2016, Cell Cycle, 32) through ligand binding assays and co-crystallization. A MINERVA[588−596] peptide containing the phosphorylated Tyr593 ([588]-DWGEEpYSNS-[596]) was purchased (Peptron).

C-terminally truncated human H-Ras[1−166] (residues Met1–His166) was amplified using PCR and cloned into pET-28a(+) vectors (Novagen) to contain an N-terminal His6-tags. The plasmids containing H-Ras[1−166] construct was transformed into Rosetta 2(DE3)pLysS *Escherichia coli* strains. The transformed cells were grown for 16 h in Luria Broth media without induction to be used as seeds for the scale-up culture. The uninduced *E. coli* seeds containing H-Ras[1−166] plasmids were inoculated into a scale-up culture media at a ratio of 1:50. The cells in scale-up culture were grown in Luria Broth media and induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside at $O.D._{600nm}$ 0.5, followed by further incubation at 20 °C for 16 h. Harvested cells were lysed by a cell sonicator (SONICS) in a lysis buffer containing 20 mM Tris-HCl (pH 7.5), 500 mM NaCl, 35 mM imidazole, and 1 mM phenylmethylsulfonyl fluoride. Cell debris was removed by centrifugation at 35,000 × g for 50 min at 4 °C, then the supernatant was filtered with 0.22 μm filter to remove cell debris and any aggregated proteins. The filtered supernatant was applied onto a HiTrap Chelating HP column (GE Healthcare) equilibrated with the lysis buffer for affinity chromatography. Unbound or weakly bound proteins were removed from the HiTrap Chelating HP

column with six column volume elutions of the washing buffer containing 20 mM Tris-HCl (pH 7.5), 500 mM NaCl, and 50 mM imidazole. The retained proteins were eluted with a gradually increasing addition of a buffer containing 20 mM Tris-HCl (pH 7.5), 500 mM NaCl, and 500 mM imidazole. The eluted protein samples were further purified by size-exclusion chromatography with HiLoad 16/600 Superdex75 pg column (GE Healthcare) equilibrated with a buffer containing 20 mM Tris-HCl (pH 8.0) 50 mM NaCl, 5 mM $MgCl_2$, and 1 mM dithiothreitol (DTT) (Figure 20).

Ni affinity

Binding buffer: 20 mM Tris, 500 mM NaCl, 35 mM Imidazole; pH 7.5
Elution buffer: 20 mM Tris, 500 mM NaCl, 1 M Imidazole; pH 7.5

SEC

Superdex75pg

SEC buffer: 20 mM Tris, 50 mM NaCl; 5 mM MgCl$_2$, 1 mM DTT; pH 8.0
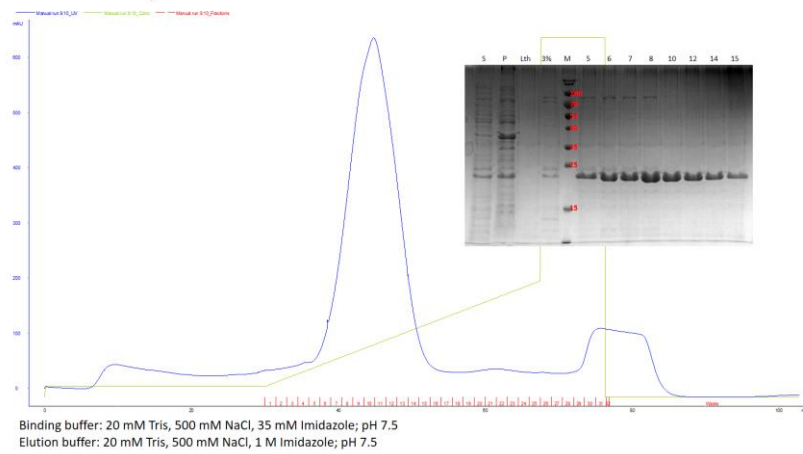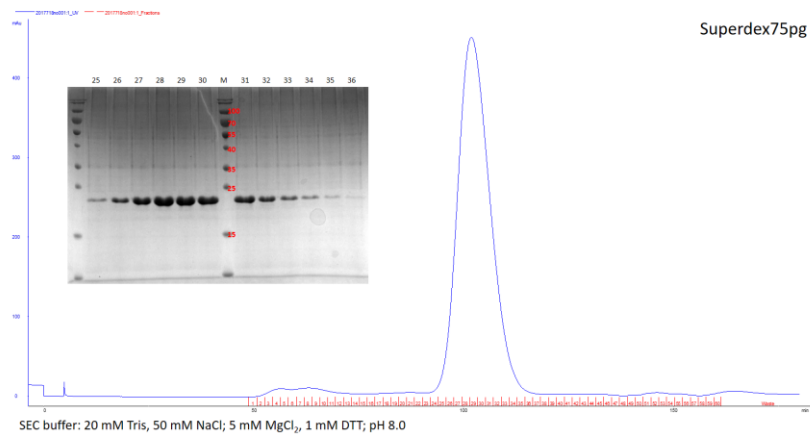
Figure 20. Purification profiles of H-Ras$^{1-166}$.

H-Ras$^{1-166}$ was purified through nickel affinity chromatography (top) followed by size exclusion chromatography (bottom) to high purity, as analyzed by SDS-PAGE results (insets).

Crystals of H-Ras[1–166] was grown at 22 °C by the hanging drop vapor diffusion method by mixing equal volumes of the protein at 14.8 mg mL^{-1} containing the MINERVA[588–596] peptide at 1:20 molar ratio and a crystallization solution containing 0.2 M lithium sulfate, 0.1 M Tris-HCl (pH 8.5) and 30% (w/v) polyethylene glycol (PEG) 4,000. Crystals of diffraction quality were cryoprotected with the reservoir solution containing additional 25% (v/v) glycerol and flash-cooled in a nitrogen gas stream at 100 K. Diffraction data from H-Ras[1–166] was collected up to 1.500 Å resolution (Figure 21). Raw X-ray diffraction data were indexed and scaled using the *HKL-2000* program suite (Otwinowski and Minor, 1997, Methods Enzymol, 24). The crystal structure of H-Ras[1–166] was determined by molecular replacement with the *MolRep* program (Vagin and Teplyakov, 2010, Acta crystallographica Section D, Biological crystallography, 30), using the structures of previous determined structure of H-Ras[1–166] as a phasing model (PDB ID: 4L9S) (Ostrem, et al., 2013, Nature, 33). The initial model was further refined to the final model using iterative cycles of model building with *Coot* (Emsley, et al., 2010, Acta crystallographica Section D, Biological crystallography, 27) and subsequent refinement with *Refmac5* in the *CCP4* suite (Murshudov, et al., 1997, Acta crystallographica Section D, Biological crystallography, 28). The determined crystal structure of H-Ras[1–166] did not contain bound MINERVA[588–596] peptide.
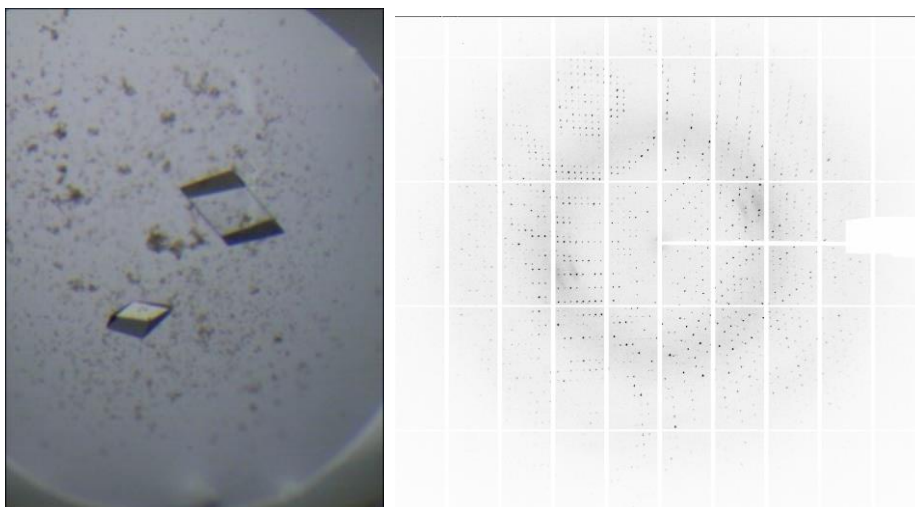
Figure 21. Representative crystals and X-ray diffraction pattern of H-Ras[1–166].

Left, Crystals of H-Ras[1–166] were grown under crystallization buffer containing 200 mM lithium sulfate, 100 mM Tris-HCl (pH 8.5) and 30% (w/v) PEG 4,000. Right, a diffraction pattern of H-Ras[1–166] crystal diffraction up to 1.500 Å resolution.

2.6. Cloning, expression, purification, and crystallization of Keap1[321–609]

Keap1[321–609] was cloned, expressed and purified to assess interaction of Keap1 with the DLGX[7]ETGE motif within the C-terminal region of MINERVA (Cheng, et al., 2019, EBioMedicine, 12) through SPR experiment.

N- and C-terminally truncated Keap1 protein (residues Ala321–Thr609) containing the Kelch domain was amplified using PCR and cloned into the pET-28a(+) vector toe contain an N-terminal His[6]-tag. The plasmid containing Keap1[321–609] was transformed into Rosetta 2(DE3)pLysS *Escherichia coli* strain. The transformed cells were grown for 16 h in Luria Broth media without induction to be used as seeds for the scale-up culture. The uninduced *E. coli* seeds containing H-Ras[1–166] plasmids were inoculated into a scale-up culture media at a ratio of 1:50. The cells in scale-up culture were grown in Luria Broth media and induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside at O.D.[600nm] 0.5, followed by further incubation at 20 °C for 16 h. Harvested cells were lysed by a cell sonicator (SONICS) in a lysis buffer containing 20 mM Tris-HCl (pH 7.5), 500 mM NaCl, 35 mM imidazole, and 1 mM phenylmethylsulfonyl fluoride. Cell debris was removed by centrifugation at 35,000 × g for 50 min at 4 °C, then the supernatant was filtered with 0.22 μm filter to remove cell debris and any aggregated proteins. The filtered supernatant was applied onto a HiTrap Chelating HP column (GE Healthcare) equilibrated with the lysis buffer for affinity chromatography. Unbound or weakly bound proteins were removed from the HiTrap Chelating HP column with six column volume elutions of the washing buffer containing 20 mM Tris-HCl (pH 7.5), 500 mM NaCl, and 50 mM imidazole. The retained proteins were eluted with a gradually increasing addition of a buffer containing 20 mM

Tris-HCl (pH 7.5), 500 mM NaCl, and 500 mM imidazole. The eluted protein samples were further purified by size-exclusion chromatography with HiLoad 16/600 Superdex75 pg column (GE Healthcare) equilibrated with a buffer containing 20 mM Tris-HCl (pH 7.5) and 5 mM Dithiothreitol (DTT) (Figure 22).

Ni affinity

Keap1 321-609 Ntag R2p 20191203 Ni HiTrap HG

Binding buffer: 20 mM Tris, 500 mM NaCl, 35 mM Imidazole; pH 7.5
Elution buffer: 20 mM Tris, 500 mM NaCl, 1 M Imidazole; pH 7.5

SEC

SEC buffer: 20 mM Tris (pH 7.5), 5 mM DTT
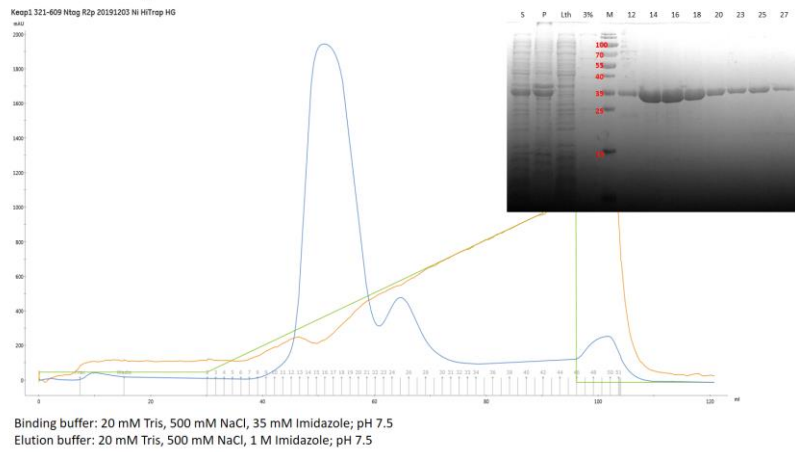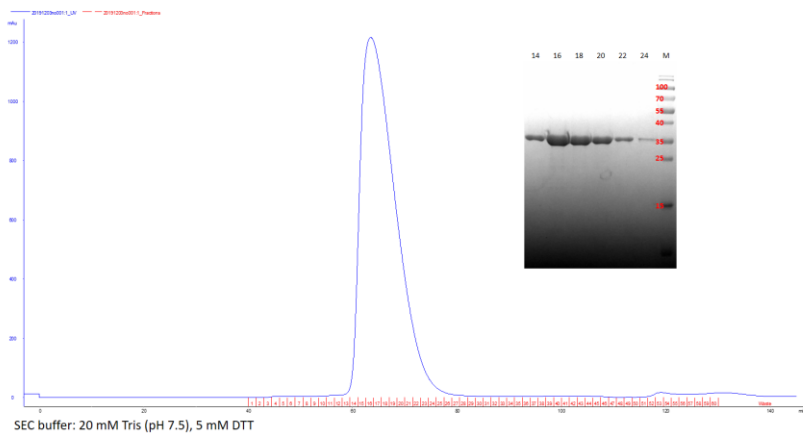
Figure 22. Purification profiles of Keap1$^{321-609}$.

Keap1$^{321-609}$ was purified through nickel affinity chromatography (top) followed by

size exclusion chromatography (bottom) to high purity, as analyzed by SDS-PAGE

results (insets).

41

Chapter 3. Results

3.1. Literature search and data mining

Sequence analysis result of MINERVA using various web servers such as *XtalPred*, *PSIPRED*, and *DISOPRED* (Ward, et al., 2004, Bioinformatics, 21; McGuffin, et al., 2000, Bioinformatics, 22; Slabinski, et al., 2007, Bioinformatics, 23) predicted that the C-terminal region of MINERVA (Asp575–Phe746) is disordered, while the N-terminal PH domain and the following helix-rich region is well-conserved among the members of the FAM129 family (Figure 9, Figure 10, and Figure 11). The flexible C-terminal region contains multiple phosphorylatable site, such as Tyr593 by EGFR and Ser641, Ser646, Ser692, and Ser696 by ERK that are guided by prolines (Pro642, Pro647, Pro693, and Pro697). The "[708]-DLGX$_7$ETGE-[721]" motif for binding Keap1 is also present within this C-terminal flexible region.

Although MINERVA is N-myristoylated during translation for regulating its cellular localization, it was revealed that the N-terminal PH domains was essential for the protein function (Oishi, et al., 2012, J Biochem, 14). PH domains are known to normally recognize phosphatidylinositols with multiple phosphates on their head group. However, an interaction study has showed that phosphatidylinositol-3-phosphate (PI3P) is the innate binding partner of MINERVA (Catimel, et al., 2013, J Proteomics, 34). PI3P is predominantly found on early endosomes and acts as a key regulator of intracellular membrane trafficking (Marat and Haucke, 2016, EMBO J, 35).

Adherens junctions comprise transmembrane protein E-cadherin that form contact between adjacent cells. E-cadherin mediate Ca^{2+}-dependent cell-cell adhesion and its cytoplasmic domains contain binding sites for various catenins for association with cytoskeleton and transducing intracellular signals (Harris and Tepass, 2010, Nat Rev Mol Cell Biol, 36). In cells receiving with growth signals, adherens junctions are destabilized and the bound β-catenin is released from E-cadherin. Accumulation of β-catenin in cytoplasm allows nuclear translocation of the protein and consequent expression of β-catenin-regulated downstream genes for proliferation (Daniels and Weis, 2005, Nat Struct Mol Biol, 37). Cadherin form disrupted adherens junctions can be internalized by endocytosis for either degradation or recycling (Kowalczyk and Nanes, 2012, Subcell Biochem, 38). While endosomes containing cadherins sorted for degradations merge with lysosomes, some accumulate in perinuclear compartment for recycling back to the plasma membrane (Kam and Quaranta, 2009, PLoS One, 39).

As mentioned above, the membrane of early endosomes are enriched with PI3P, the innate target of the PH domain on MINERVA. MINERVA was found on the plasma membrane whenever two cell were in contact in confluent HeLa cells, but was found to be dispersed in cytoplasm in cells with exponential growth (Chen, et al., 2011, J Biol Chem, 16). I hypothesize that the cytoplasmic MINERVA are not just dispersed freely but exist in complex with the internalized E-cadherins on the endocytic recycling compartments (ERC). In confluent cells, MINERVA would be translocated along with the E-cadherins on the ERC to the plasma membrane, displaying plasma membrane association of MINERVA (Figure 23) (Chen, et al., 2011, J Biol Chem, 16). β-catenins are dissociated from the E-cadherins upon adherens junction disruption to act as a transcription factor together with TCF/LEF

in the nucleus in response to Wnt signaling (van Noort, et al., 2002, J Biol Chem, 40).

Figure 23. Schematic of hypothesis on subcellular localization of MINERVA in exponentially growing and confluent cells.

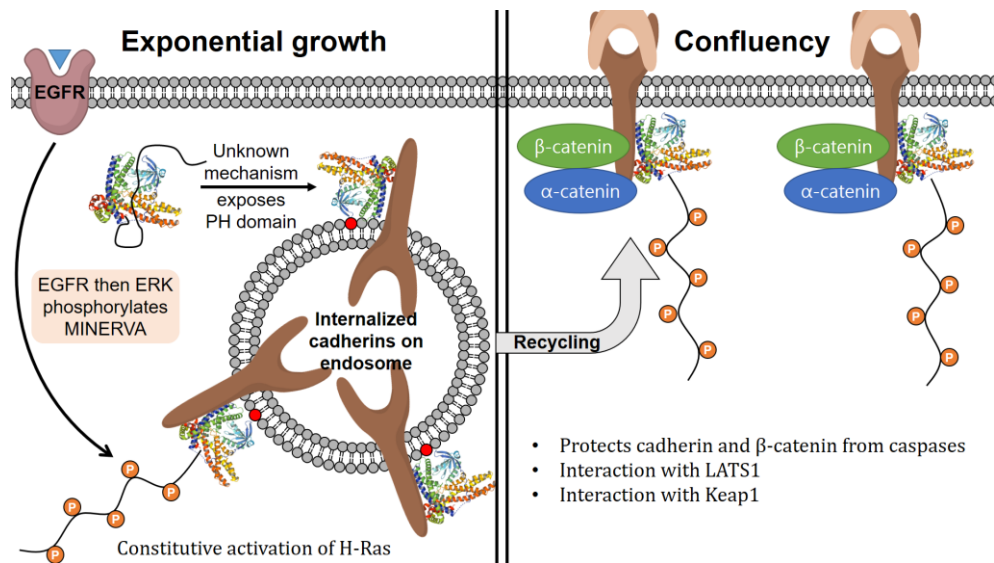## 3.2. Trypsin digestion is essential for producing diffraction quality crystals

Sequence analysis result of MINERVA using various web servers such as *XtalPred*, *PSIPRED*, and *DISOPRED* (Ward, et al., 2004, Bioinformatics, 21; McGuffin, et al., 2000, Bioinformatics, 22; Slabinski, et al., 2007, Bioinformatics, 23) predicted that the C-terminal region of MINERVA (Asp575–Phe746) is disordered, while the N-terminal PH domain and the following helix-rich region is well-conserved among the members of the FAM129 family (Figure 24). I attempted purification and crystallization of the full length MINERVA (MINERVA[FL]) and a construct lacking the C-terminal disordered region (MINERVA[1–574]). Unfortunately, not only the full length construct but also the C-terminally truncated MINERVA[1–574] produced poorly-diffracting rosette-shaped crystals, which were not fit for data collection. I speculated that internal disordered regions were hindering crystal formation and implemented limited proteolysis by trypsin. Trypsin effectively digested MINERVA[1–574] and yielded two separate fragments with molecular weights of approximately 14 kDa and 50 kDa, as confirmed by SDS-PAGE (Figure 25). N-terminal sequencing confirmed the first five residues of the fragments were Gly[2]-Asp[3]-Val[4]-Leu[5]-Ser[6] and Ser[146]-Gly[147]-Ser[148]-Ala[149]-Pro[150], respectively (Table 2). The sequencing result corresponded with the prediction from *PeptideCutter* (https://web.expasy.org/peptide_cutter), that trypsin would cleave the protein after Lys145 residue. Other predicted trypsin cleavage sites were likely protected by protein folding, which prevented trypsin approaching the sites. The trypsin-digested MINERVA[1–574] fragments were stable under the purification conditions and eluted together from the size-exclusion

chromatography column (Figure 26), meaning that trypsin-digestion did not disrupt the overall conformation of MINERVA[1–574]. Trypsin-digestion of the full length MINERVA also gave fragments of the same sizes, justifying my construct design of removing the C-terminal flexible region spanning residues from Asp575 to Phe746. The trypsin-digestion greatly aided production of large single crystals of MINERVA suitable for data collection, and hence was implemented for crystallization of other constructs.

I collected X-ray diffraction data from crystals of the trypsin-treated, SeMet-substituted MINERVA[1–574] and native MINERVA[9–553] at 2.4 Å and 2.6 Å resolutions, respectively. The phase was obtained from the SeMet-substituted MINERVA[1–574] crystal by the SAD method. Although the structure model obtained from the SeMet-substituted MINERVA[1–574] crystal was at higher resolution, I shall describe the structure based on the native MINERVA[9–553] model as loop regions connecting secondary structures are better visualized. I could successfully resolve MINERVA[9–553] structure model for Leu9–Gly50, Ala56–Pro140, and Cys154–Val553. N-terminal His$_6$-tag as well as Thr51–Gln55 and Gly141–Lys153 could not be modeled due to high flexibility in these regions. In an asymmetric unit (ASU) of $P2_1$ space group, there were two MINERVA monomers that are highly similar to each other, with root-mean-square deviation (RMSD) of 0.20 Å over 528 equivalent C_α positions for modeled residues. According to Proteins interfaces, surfaces, and assemblies (PISA) analysis (Krissinel and Henrick, 2007, J Mol Biol, 41) the two monomers within ASU each have surface area of 2,071.6 Å^2 on average, while the interface between them only has area of 284.3 Å^2, strongly suggesting that dimeric conformation within the ASU is due to crystallographic artifact. Oligomeric states of MINERVA[FL] and MINERVA[1–574] in solution were

verified by measuring molecular mass using SEC-MALS method, and confirmed that both constructs of MINERVA exist as monomers (Figure 27).

```
                                                                      1         10        20
                                                         0000000000000000000000000000

FAM129B   ...........................................MGDVLSTHLDDARRQHIAEKTGKILTEFL
FAM129A   ...........................................MGGSASSQLDEGKCAYIRGKTEAAIKNFS
FAM129C   MGPDRKEVPLSRGTQAVVVGKGRGAPGDDSSMGGRPSSPLDKQQRQHLRGQVDTLLRNFL

          α1                                                        β1              β2
          0000000000000000000000000                           ------------>    ----------->
          30        40        50        60        70        80

FAM129B   QFYEDQYGVALFNSMRHEIEGTGLPQACLLWRKVPLD.ERIVFSGNLFQHQEDSKKWRNR
FAM129A   PYYSRQYSVAFCNHVRTEVEQQRDLTSQFLKTKPPLAPGTILYEAELSQFSEDIKKWKER
FAM129C   PCYRGQLAASVLRQISRELGPQEPTGSQLLRSKKL..PRVREHRGPLTQLRGHPPRWQPI

                       β3        α2          β4         β5       α3
          ---------->    ---------->  000000      --------->  --------->   0000000000
          90        100       110       120       130       140

FAM129B   FSLVPHNYGLVLYENKAAYERQVPPRAVINSAGYKILTSVDQYLELIGNSLPGTTA...K
FAM129A   YVVVKNDYAVESYENKEAYQRGAAPKCRILPAGGKVLTSEDEYNLLSDRHFPDPLASSEK
FAM129C   FCVLRGDGRLEWFSHKEEYENGGHCLGSTALTGYTLLTSQREYLRLLDALCPESLGDHTQ

                          β6          β7          α4                       l
          --------->    --------->   000000000000000000       0
          150       160       170       180       190       200

FAM129B   SGSAPILKCPTQFPLILWHPYARHYYFCMMTEAEQDKWQAVLQDCIRHCNNGIPEDSKVE
FAM129A   ENTQPFVVLPKEFPVYLWQPFFRHGYFCFHEAADQKRFSALLSDCVRHLNHDYMKQMTFE
FAM129C   EEPDSLLEVPVSFPLFLQHPFRRHLCFSAATREAQHAWRLALQGGIRLQGIVLQRSQAPA

          α5                             α6                                 00000
          0000000000000000000       000000000000000000000000000000
          210       220       230       240       250       260

FAM129B   GPAFTDAIRMYRQSKELYGTWEMLCGNEVQILSNLVMEELGPELKAELGPRLKGKPQERQ
FAM129A   AQAFLEAVQFFRQEKGHYGSWEMITGDEIQILSNLVMEELLPTLQTDLLPKMKGKKNDRK
FAM129C   ARAFLDAVRLYRQHQGHFGDDDVTLGSDAEVLTAVLMREQLPALRAQTLPGLRGAGRARA

                                                                 α7
          000000000000000000000000000000000000000000000000000000000
          270       280       290       300       310       320

FAM129B   RQWIQISDAVYHMVYEQAKARFEEVLSKVQQVQPAMQAVIRTDMDQIITSKEHLASKIRA
FAM129A   RTWLGLLEEAYTLVQHQVSEGLSALKEECRALTKGLEGTIRSDMDQIVNSKNYLIGKIKA
FAM129C   WAWTELLDAVHAAVLAGASAGLCAFQPEKDELLASLEKTIRPDVDQLLRQRARVAGRLRT

          0000000000000000000000000000000000000000000000000000       0000
          330       340       350       360       370       380

FAM129B   FILPKAEVCVRNHVQPYIPSILEALMVPTSQGFTEVRDVFFKEVTDMNLNVINEGGIDKL
FAM129A   MVAQPAEKSCLESVQPFLASILEELMGPVSSGFSEVRVLFEKEVNEVSQNFQTTKDSVQL
FAM129C   DIRGPLESCLRREVDPQLPRVVQTLLRTVEASLEAVRTLLAQGMDRLSHRLRQSPSGTRL

          α8                α9                                    α10
          0000000000000   000000000000000                  000000000000000000
          390       400       410       420       430       440

FAM129B   GEYMEKLSRLAYHPLKMQSCYEKMESL..RLDGLQQRFDVSSTSVFKQRACIHMREQMDN
FAM129A   KEHLDRLMNLPLHSVKMEPCYTKVNLLHERLQDLKSRFRFPHIDLVVQRTQNYMQELMEN
FAM129C   RREVYSFGEMPWDLALMQTCYREAERSRGRLGQLAAPFGFLGMQSLVFGAQDLAQQLMAD

                                                             α11
          00000000000000000000000000000000000000000000   0000000000000000000
          450       460       470       480       490       500

FAM129B   AVYTFETLLHQELGKGPTKEELCKSIQRVLERVLKKYDYDSSSVRKRFFREALLQISIPF
FAM129A   AVFTFEQLLSPHLQGEASKT..AVAIEKVKLRVLKQYDYDSSTIRKKIFQEALVQITLPT
FAM129C   AVATFLQIADQCITTALNCDQAAQRLERVRGRVLKKFKSDSGLAQRRFIRGWGLCIFLPF

          0000000000000000000000000000       η1        α12
          000000000000000000000000000   00000000000  000000000000000000
          510       520       530       540       550       560

FAM129B   LLKKLAPTCKSELPRFQELIFEDFARFILVENTYEEVVLQTVMKDILQAVKEAAVQRKHN
FAM129A   VQKALASTCKPELQKYEQFIFADHTNMIHVENVYEEILHQILLDETLKVIKEAAILKKHN
FAM129C   VLSQLEPGCKKELPEFEGDVLAVGSQALTTEGIYEDVIRGCLLQRIDQELKKTLGANDVS

                    570       580       590       600             610

FAM129B   LYRDSMVMHNSDPNLHLLAEGAPIDWGEEYSNSGGGGSPSP............STPESATL
FAM129A   LFEDNMALPSESVSS..LTDLKP.........PTGSNQASP..........ARRASAIL
FAM129C   ..............C.TLDGCLEVPWEQEGAAPNLNLVSSFLAGRQAFTDFLCLPAKSSA
```

```
                                                                    650
                                                                     .
FAM129B    .................................................VTEIRGLLA
FAM129A    SRVDDPVVNPVATEDTAGLPGTCSSELEFGGTLEDEEPAQEEPEPITASGSLKALRKLLT
FAM129C    ............................................................


           660       670       680
            .         .         .
FAM129B    QGLRPESPPPAGPLLNGAPAGESPQPKAAPE.............................
FAM129A    ASV..EVPVDSAPVMEEDTNGESHVPQENEEEEEKEPSQAAAIHPDNCEESEVSEREAQP
FAM129C    ............................................................


                                   690       700       710       720
                                    .         .         .         .
FAM129B    ..........................ASSPPASPLQHLLPGKAVDLGPPKPSDQETGEQ
FAM129A    PCPEAHGEELGGFPEVGSPASPPASGGLTEEPLGPMEGELPGEACTLTAHEGRGGKCTEE
FAM129C    ............................................................


                      730       740
                       .         .
FAM129B    VS.......SPSSHPALH......TTTEDSAGVQTEF
FAM129A    GDASQQEGCTLGSDPICLSESQVSEEQEEMGGQSSAA
FAM129C    ....................................
```

Figure 24. Multiple sequence alignment of MINERVA (FAM129B) and its homologs FAM129A and FAM129C.

Multiple sequence alignment of MINERVA (FAM129B), FAM129A, and FAM129C using *Clustal Omega* (Madeira, et al., 2019, Nucleic Acids Res, 42), displayed with *ESPript 3.0* (Robert and Gouet, 2014, Nucleic Acids Res, 43). Sequences are well-conserved among the family members for the structured parts, while C-terminal flexible region show little conservation. Conserved residues and similar residues are indicated by brown highlights and brown texts, respectively. Secondary structure elements from MINERVA model are shown above the alignment in red and blue for PH domain and helix bundle domain, respectively.
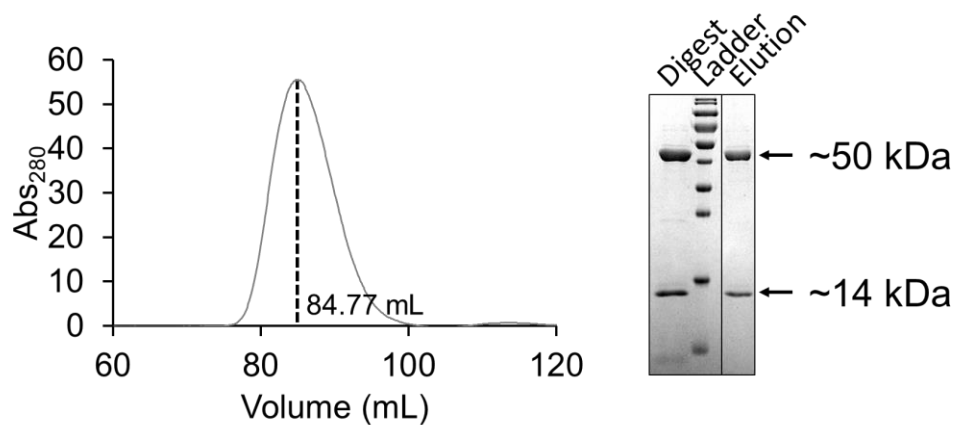
Figure 25. Fragments of trypsin-digested MINERVA[1–574] remain intact.

Left, trypsin-digested MINERVA[1–574] are eluted in a single peak from size exclusion chromatography column. Right, SDS-PAGE analysis of the eluted fraction reveal that the protein is separated into two fragments, each with molecular weight of approximately 14 kDa and 50 kDa.

Table 2. Possible trypsin cleavage sites and resulting peptides

Position	Resulting peptide sequence
13	MGDVLSTHLDDAR
14	R
20	QHIAEK
23	TGK
45	ILTEFLQFYEDQYGVALFNSMR
61	HEIEGTGLPQAQLLWR
86	KVPLDERIVFSGNLFQHQEDSKKWR
88	NR
104	FSLVPHNYGLVLYENK
109	AAYER
114	QVPPR
123	AVINSAGYK
145	ILTSVDQYLELIGNSLPGTTAK
153	SGSAPILK
168	CPTQFPLILWHPYAR
182	HYYFCMMTEAEQDK
192	WQAVLQDCIR
203	HCNNGIPEDSK
214	VEGPAFTDAIR
217	MYR
250	QSKELYGTWEMLCGNEVQILSNLVMEELGPELK
256	AELGPR
258	LK
264	GKPQER
266	QR
284	QWIQISDAVYHMVYEQAK
286	AR
293	FEEVLSK
306	VQQVQPAMQAVIR
322	TDMDQIITSKEHLASK
324	IR
330	AFILPK
336	AEVCVR
362	NHVQPYIPSILEALMVPTSQGFTEVR
367	DVFFK

391	EVTDMNLNVINEGGIDKLGEYMEK
394	LSR
401	LAYHPLK
408	MQSCYEK
413	MESLR
420	LDGLQQR
430	FDVSSTSVFK
432	QR
438	AQIHMR
458	EQMDNAVYTFETLLHQELGK
462	GPTK
467	EELCK
471	SIQR
475	VLER
479	VLKK
488	YDYDSSSVR
490	KR
513	FFREALLQISIPFLLKKLAPTCK
518	SELPR
529	FQELIFEDFAR
554	FILVENTYEEVVLQTVMKDILQAVK
560	EAAVQR
561	K
566	HNLYR
615	DSMVMHNSDPNLHLLAEGAPIDWGEEYSNSGGGGSPSPSTPESATLSEK
620	RRRAK
654	QVVSVVQDEEVGLPFEASPESPPPASPDGVTEIR
685	GLLAQGLRPESPPPAGPLLNGAPAGESPQPK
705	AAPEASSPPASPLQHLLPGK
746	AVDLGPPKPSDQETGEQVSSPSSHPALHTTTEDSAGVQTEF
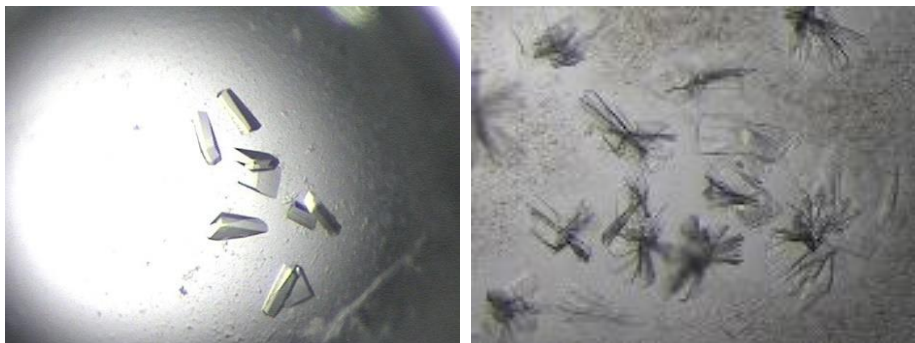
Figure 26. Comparison of the undigested and trypsin-digested MINERVA[1–574] crystals.

Left, single rigid crystals of trypsin-digested MINERVA[1–574] are more suitable for X-ray diffraction data collection than crystals of undigested MINERVA[1–574] on right.
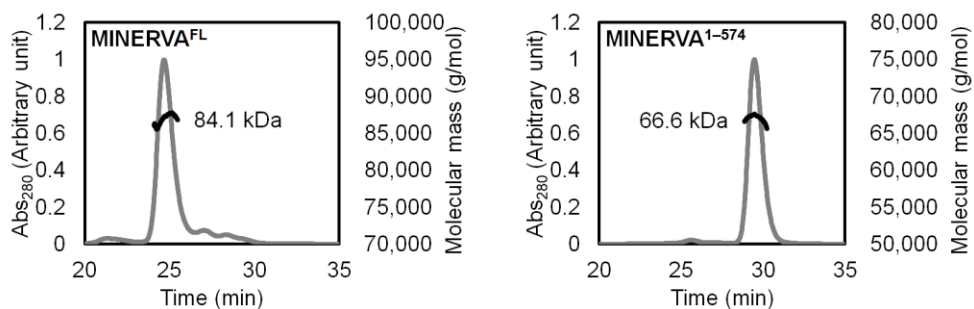
Figure 27. SEC-MALS results of MINERVA^{FL} and MINERVA$^{1-574}$ show monomeric states of the proteins.

Left, SEC-MALS result of MINERVA^{FL}. Right, SEC-MALS result of MINERVA$^{1-574}$.

3.3. Overall structure of MINERVA

The overall structure model of a MINERVA[9–553] monomer assumes a novel globular fold that spans approximately 70 Å by 70 Å by 60 Å in size. MINERVA[9–553] could be divided into two domains, the N-terminal PH domain (residues Asp66–Asn196) and an α-helix bundle domain comprising the rest of the molecule, which includes the N-terminal helix α1 (residues Asp11–Glu49) (Figure 28 and Figure 29). The N-terminus of the protein is located near the PH-domain, presumably for cooperative action of myristic acid membrane insertion and PI3P recognition for recruitment of MINERVA to membrane (Figure 29). Surface representation of the protein reveals that MINERVA possesses a positively-charged patch on putative membrane-interacting surface (Figure 30, middle). Additionally, the surface representation indicates that the protein possess two distinct crevices on the front and the back, which might serves as binding platform for interaction partners (Figure 30, left and right).

The helix bundle domain is composed of eight α-helices (α5–α12) and one independent 3_{10} helix (η6) that are held together mostly by hydrophobic contacts. The overall architecture of the helix bundle domain resembled no other known structure, as analyzed by the Dali structural similarity search algorithm (Holm, 2019, Bioinformatics, 44), suggesting that it is of a novel fold. The longest helix α7 spanning 119 residues (Pro261–Glu379) is severely bent to encircle the entire protein (Figure 29). Other helices in the bundle interact with the helix α7 via hydrophobic contacts with an exception of helix α5, which is in proximity to helix α1. Helices α1 and α6 are associated with the N-terminal half of the helix α7, while helices α8, α9, α10, α11, and α12 interacts with the C-terminal half of the helix α7

(Figure 29). The N-terminal PH domain is held by the helix bundle domain that acts as a scaffold. Residues on helices α6, 7, 9, and 10 comprise extensive network of hydrogen bond and salt bridges with residues on the PH domain.

The N-terminal PH domain adopts a typical fold of a PH domain containing a central curved antiparallel β-sheet (β5↑-β6↓-β7↑-β1↓-β2↑-β3↓-β4↑) followed by a C-terminal α-helix (α4) (Figure 29). Interestingly, there are two additional α-helices inserted between the β strands when compared to a typical PH domain structure from Pleckstrin (PDB ID: 2I5F) (Jackson, et al., 2007, BMC Struct Biol, 45), which the name PH domain derived from; α2 between β3 and β4, and α3 between β5 and β6 (Figure 31). Sequence alignment and structural comparison of the PH domains from MINERVA against Pleckstrin revealed that the core secondary structure including the characteristic antiparallel β-sheet is well conserved, with RMSD of 1.30 Å over 91 equivalent C_α positions. A major difference arises from a loop region connecting the helix α3 and the strand β6 (residues Gly136–Phe158) of MINERVA, which is unusually long and disordered (Figure 31). The corresponding region in Pleckstrin (residues Asn303–Glu311) was shorter but also disordered. Incidentally, the trypsin cleavage site (Lys145ˇSer146) is located within this α3–β6 loop, meaning that this loop is intrinsically disordered and exposed to solvation. Notably, intermolecular interaction analysis by the PISA program suggests that the cleavage allowed better stacking of the MINERVA molecules in the protein crystal; a portion of the helix α7' (residues Ser292'–Arg306') from an adjacent MINERVA approaches the region exposed by the cleavage to form several intermolecular hydrogen bonds with residues on helix α2 and helix α7 (residues Asp129–Asn137 and Lys260–Gln265; Figure 32). The residues forming hydrogen bonds would have been concealed by the α3–β6 loop

without trypsin-digestion, explaining poor diffraction quality of crystals of undigested MINERVA I had obtained.
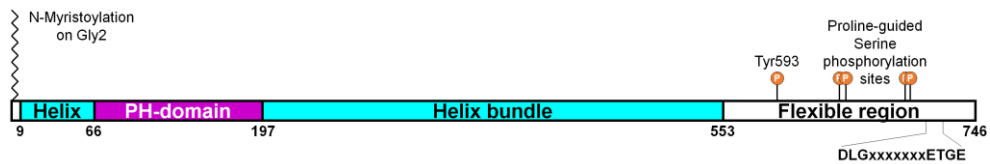
Figure 28. Schematic of MINERVA sequence showing domain composition.

Schematic of MINERVA sequence showing domain composition and posttranslational modification sites, as well as the Keap1-binding motif [708]-DLGX$_7$ETGE-[721]. The sites of phosphorylations by EGFR (Tyr593) and ERK (Ser628, Ser633, Ser679, and Ser683) are marked with circles.
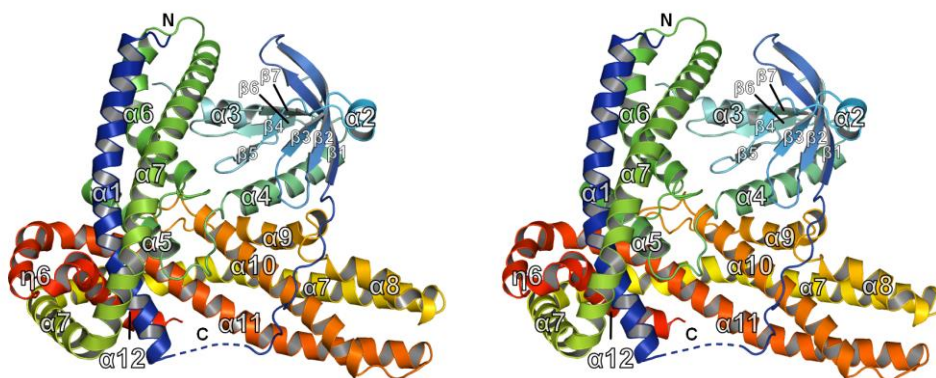
Figure 29. Stereoscopic view of MINERVA[9–553]

MINERVA[9–553] is represented in cartoon representation colored in rainbow scheme showing the PH domain and complexity of the helix bundle domain.
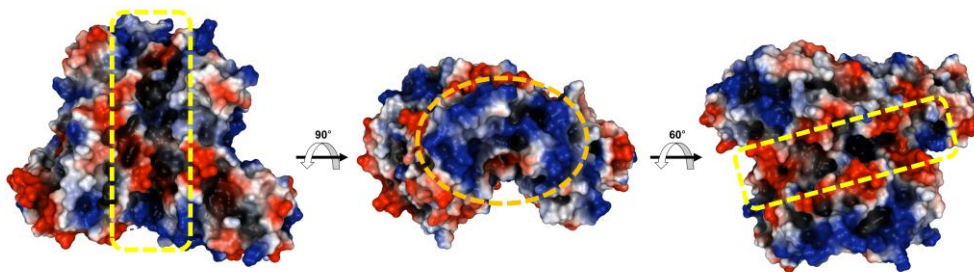
Figure 30. Overall structure of MINERVA$^{9–553}$ monomer.

Surface representation of the MINERVA$^{9–553}$ structure. The structure is rotated horizontally to show membrane-associating surface (orange dashed circle) and two crevices for presumed protein-protein interactions (yellow dashed circles)
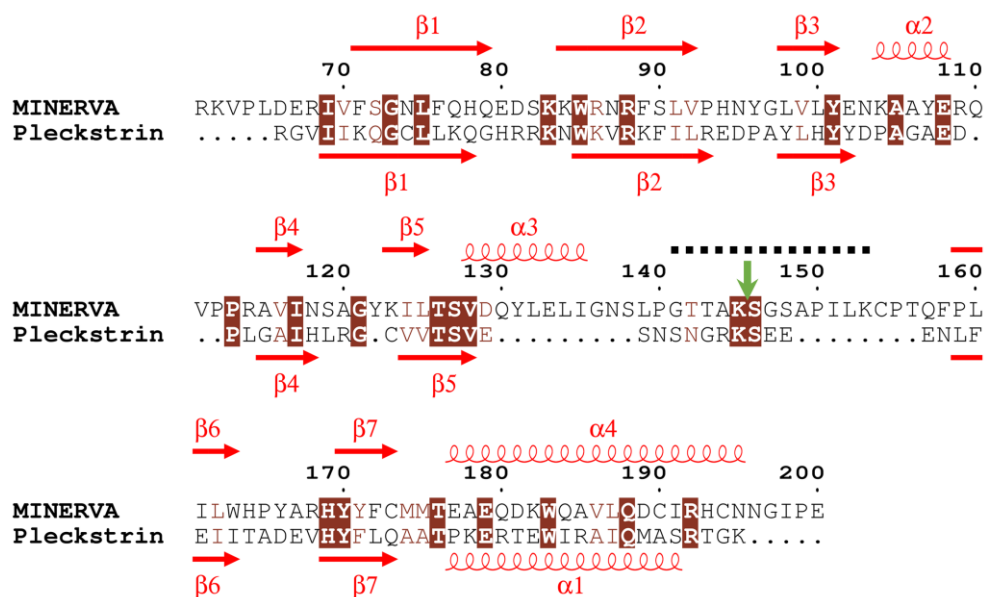
Figure 31. Structure-based sequence alignment of PH domains from MINERVA and Pleckstrin (PDB ID: 2I5F).

Residues within the PH domains of MINERVA and Pleckstrin are well-conserved. MINERVA contains extra helices α2 and α3 inserted in loops connecting β-strands. Conserved residues and similar residues are indicated by brown highlights and brown texts, respectively. Secondary structure elements from MINERVA and Pleckstrin model are shown in red, above and below the alignment, respectively. The trypsin cleavage site is indicated with a green arrow. Unmodeled residues (Gly141–Lys153) due to lack of electron density around the cleavage site are indicated with a black dashed line.
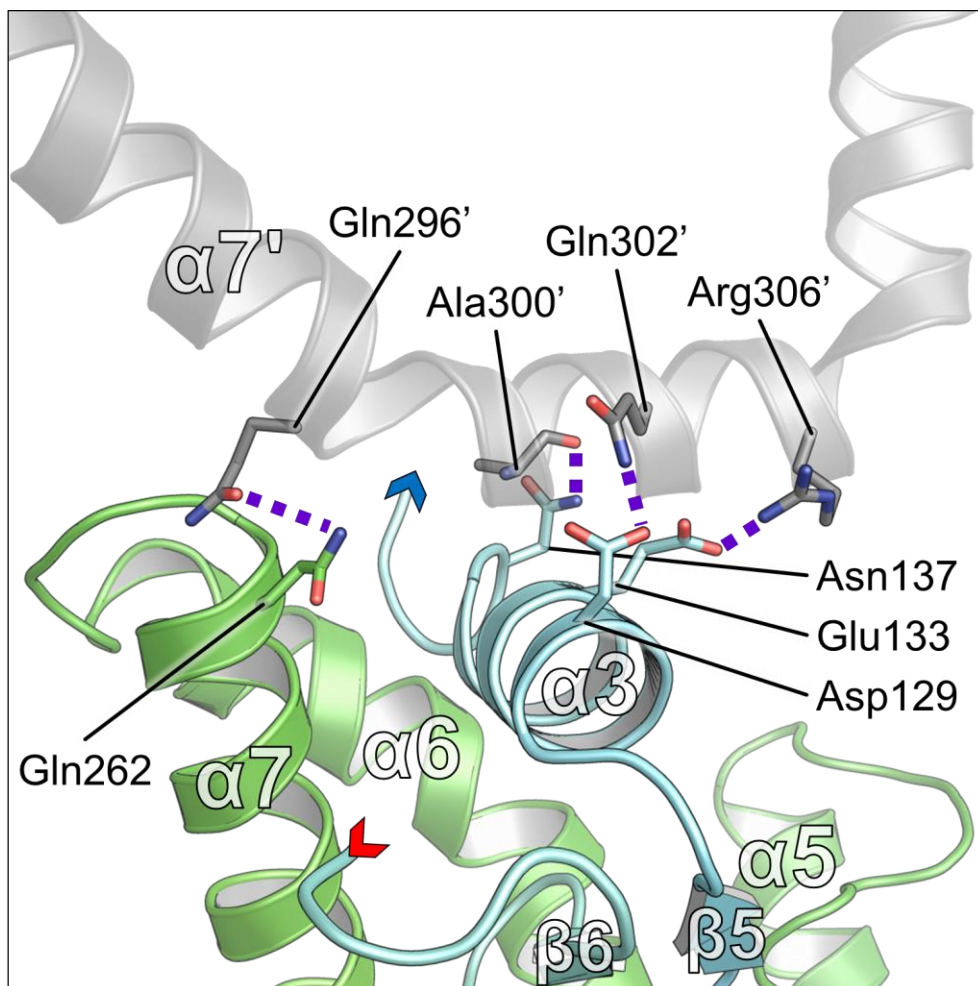
Figure 32. Trypsin cleavage allow better stacking of MINERVA molecules in crystal.

Intermolecular interaction between two adjacent molecules of MINERVA comprises multiple hydrogen bonds, indicated with purple dashed lines. The PH domain of a MINERVA molecule is shown as a cartoon representation in cyan-to-green color while helix α7' of the adjacent MINERVA is in gray. The terminal residue before and after the unmodeled gap (residues Gly141–Lys153) are indicated with a blue and a red arrows, respectively.

3.4. PH domain of MINERVA recognizes phosphatidic acid and inositol phosphates

PH domains are well-known for high affinity and specificity towards phosphatidylinositol phosphates (Lemmon, 2007, Biochem Soc Symp, 2). MINERVA are recruited to plasma membrane under certain growth-inhibited conditions (Old, et al., 2009, Mol Cell, 6), affinity towards the phosphatidylinositol phosphates of the PH domains were expected. In fact, PI3P was recognized as the innate ligand of the MINERVA PH domain (Catimel, et al., 2013, J Proteomics, 34). In order to elucidate the binding affinity of the phosphatidylinositol phosphates, we attempted to solve the structure of MINERVA in complex with inositol phosphates, the head groups of phosphatidylinositols phosphates. However, we could only obtain structures of apo-forms of MINERVA. Structure-based sequence alignment of the PH domains from MINERVA and Pleckstrin in complex with inositol (2,3,4,5,6)-pentakisphosphate (IP_5), revealed that putative binding site for phosphatidylinositols on MINERVA is composed of residues Gln77, Gln79, Lys84, Arg86, Arg88, Arg114, and Arg168, which are mainly located on strands β1 and β2, forming a ligand binding pocket (Figure 31 and Figure 33). On Pleckstrin, residues Lys253, His256, Arg257, Arg258, Lys262, Arg264, and Tyr277 establish extensive hydrogen bond network with the bound IP_5. Among these residues, His256–Arg258 seem to be the major contributors in stabilizing the bound IP_5 via encapsulating the ligand with their long and positively-charged sidechains (Figure 34). MINERVA has negatively-charged amino acid residues (Glu80–Asp81–Ser82) instead, on the positions corresponding to His256–Arg258 of Pleckstrin, which form a negatively-charged face within the ligand binding site (Figure 33). The oppositely-charged faces within the ligand binding site would

restrict types of ligands that can bind in the PH domain of MINERVA to those with fewer phosphates at their head groups. Furthermore, efficient IP_5 binding on Pleckstrin is facilitated by absence of side chain on Gly255 residue, while MINERVA has Gln79 in the corresponding position (Figure 33), which would cause steric hindrance to incoming ligand.

In order to assess lipid binding specificity, we implemented lipid dot blot assay using the PIP Strips™. The strip is a cellulose membraned dotted with 100 pmol of the following lipids; lysophosphatidic acid (LPA), lysophosphocholine (LPC), phosphatidylinositol (PtdIns), phosphatidylinositol (3)-phosphate (PtdIns(3)P), phosphatidylinositol (4)-phosphate (PtdIns(4)P), phosphatidylinositol (5)-phosphate (PtdIns(5)P), phosphatidylethanolamine (PE), phosphatidylcholine (PC), sphingosine 1-phosphate (S1P), phosphatidylinositol (3,4)-bisphosphate (PtdIns(3,4)P_2), phosphatidylinositol (3,5)-bisphosphate (PtdIns(3,5)P_2), phosphatidylinositol (4,5)-bisphosphate (PtdIns(4,5)P_2), phosphatidylinositol (3,4,5)-trisphosphate (PtdIns(3,4,5)P_3), phosphatidic acid (PA), and phosphatidylserine (PS). The dot blot assay revealed that MINERVA[1–574] shows affinity towards PA, PtdIns(3)P, PtdIns(4)P, PtdIns(5)P, PtdIns(3,4)P_2, PtdIns(3,5)P_2, and PtdIns(4,5)P_2, but not PtdIns(3,4,5)P_3 or any other lipids Figure 35). A common feature shared among the lipid bound to MINERVA[1–574] are that they have one or two unmasked phosphates at their head groups and two fatty acid chains, that may be the determinants of the specificity of the MINERVA PH domain. Surprisingly, PA showed higher affinity than PI3P, the innate ligand of the MINERVA PH domain. According to structure-based sequence alignment of PH domains of MINERVA and Pleckstrin, the lipid binding pocket of MINERVA is expected to possess specific affinity towards lipids with fewer phosphates at their

head groups due to oppositely-charged faces within the ligand binding site and the

steric hindrance elicited by side chain of the Gln79 residue (Figure 33).
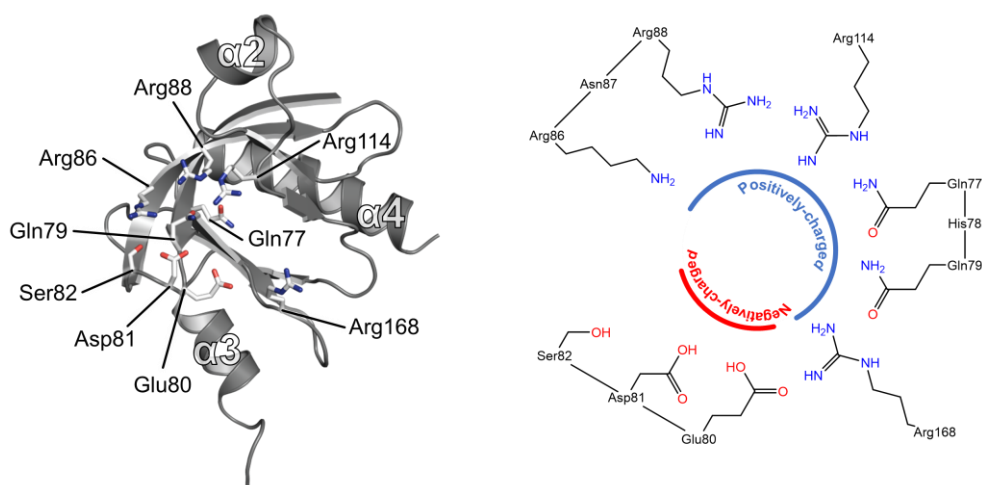
Figure 33. Lipid binding pocket on PH domain of MINERVA.

Left, lipid binding pocket of the PH domain of MINERVA is composed of multiple charged residues. Right, schematic of the PH domain of MINERVA lipid binding pocket illustrating positively-charged residues on one face and negatively-charged residues on the other of the ligand binding pocket.
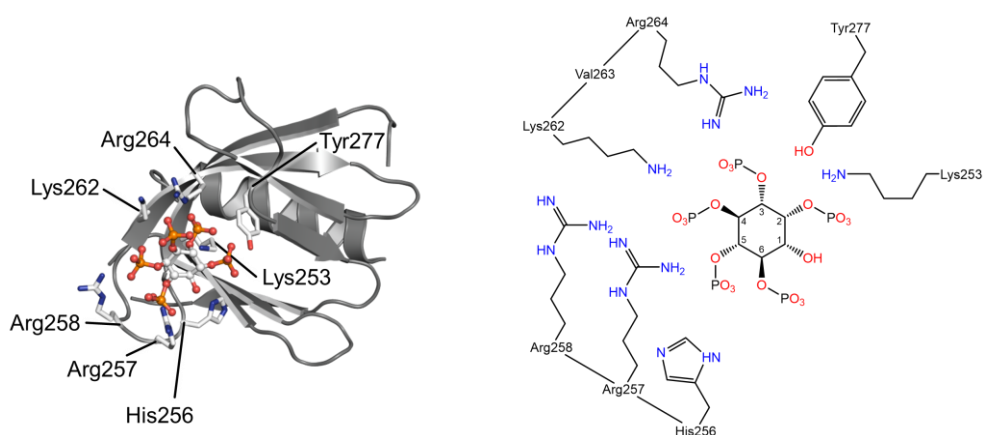
Figure 34. Lipid binding pocket on PH domain of Pleckstrin.

Left, lipid binding pocket of the Pleckstrin accommodates inositol (2,3,4,5,6)-pentakisphosphate (IP_5) via multiple positively-charged residues. Right, schematic of the PH domain of the Pleckstrin illustrating interaction with IP_5 via multiple positively-charged residues at the lipid binding pocket.
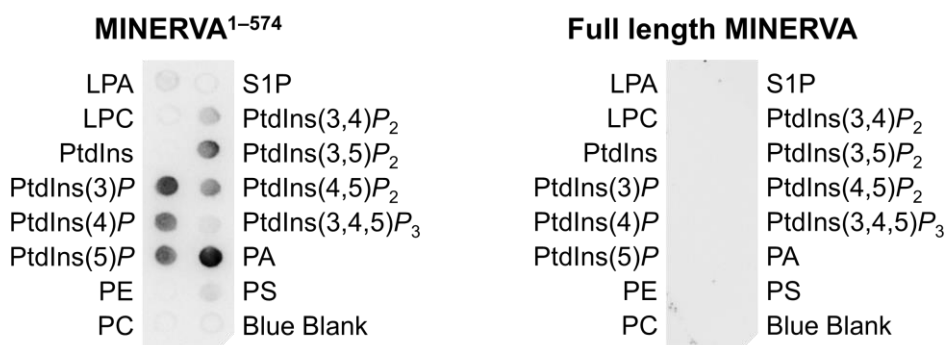
Figure 35. Lipid dot-blot assay results of MINERVA$^{1-574}$ and MINERVA^{FL}.

Lipid dot-blot assay of MINERVA$^{1-574}$ (left) and the full length MINERVA (right) showing specific interaction of MINERVA$^{1-574}$ with certain lipids, while the full length MINERVA fail to bind any of the lipids. Dotted lipids are as following; lysophosphatidic acid (**LPA**), lysophosphocholine (**LPC**), phosphatidylinositol (**PtdIns**), phosphatidylinositol (3)-phosphate (**PtdIns(3)P**), phosphatidylinositol (4)-phosphate (**PtdIns(4)P**), phosphatidylinositol (5)-phosphate (**PtdIns(5)P**), phosphatidylethanolamine (**PE**), phosphatidylcholine (**PC**), sphingosine 1-phosphate (**S1P**), phosphatidylinositol (3,4)-bisphosphate (**PtdIns(3,4)P_2**), phosphatidylinositol (3,5)-bisphosphate (**PtdIns(3,5)P_2**), phosphatidylinositol (4,5)-bisphosphate (**PtdIns(4,5)P_2**), phosphatidylinositol (3,4,5)-trisphosphate (**PtdIns(3,4,5)P_3**), phosphatidic acid (**PA**), and phosphatidylserine (**PS**).

3.5. The C-terminal region of MINERVA interacts with Keap1

Keap1$^{321-609}$ at concentrations of 78.1, 156, 313, 625, 1,250, 2,500, 5,000, and 10,000 nM were injected over sensor chips coated with either MINERVA^{FL} or MINERVA$^{1-574}$ to assess the interaction between the C-terminal flexible region of MINERVA and Keap1$^{321-609}$ by SPR experiment. Keap1$^{321-609}$ showed high affinity towards MINERVA^{FL} with K_D value of 1.72 μM (Figure 36), but not MINERVA$^{1-574}$, indicating that the DLGX$_7$ETGE motif within C-terminal flexible region of MINERVA is indeed responsible for the interaction with Keap1 (Cheng, et al., 2019, EBioMedicine, 12).
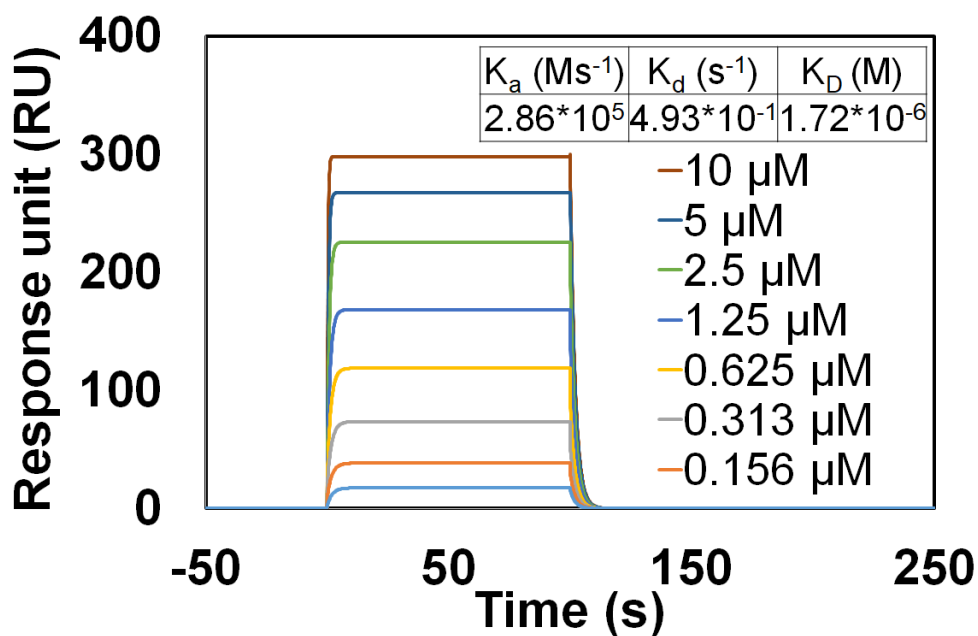
Figure 36. SPR analysis of interaction between MINERVA[FL] and Keap1[321−609].

Keap1[321−609] at concentrations of 78.1, 156, 313, 625, 1,250, 2,500, 5,000, and 10,000 nM were injected over a sensor chip coated with MINERVA[FL].

Chapter 4. Discussion

MINERVA is a protein of a yet unknown function, which is upregulated in many types of cancer. Sequence analysis have predicted presence of a PH domain at N-terminus, followed by an α-helix rich region and a C-terminal flexible region. Here, I report the structure of the MINERVA[9–553], which includes the PH domain and the α-helix rich scaffold. Since preliminary attempts to determine the structure of the protein was unsuccessful, I employed trypsin-digestion of the purified protein in order to improve diffraction quality of the crystals. Though MINERVA[9–553] was cleaved after the Lys145 residue by trypsin-digestion, the protein maintained its integrity under purification and crystallization conditions. It was revealed that the cleavage permitted additional intermolecular interactions between adjacent MINERVA molecules in crystals and hence the successful structure determination.

The α-helix bundle domain (residues Met8–Gly51 and Gly197–Val553) contains eight α-helices and six 3_{10} helices that caps four of the α-helices, bundled together mainly by hydrophobic effects. The domain structure is of a novel fold and no protein matching the entire helix bundle domain was found from the DALI search. Helix α7 was remarkably long and have several kinks that allow the helix to encircle and hold the whole protein together as a backbone of the scaffold.

Since PH domains generally have high affinity towards phosphatidylinositols, the PH domain in MINERVA was deemed responsible for the plasma membrane recruitment together with the N-terminal myristoylation on Gly2. I discovered that PH domain of MINERVA show affinity towards specific types of lipids, which

may be important in regulating localization of the protein. Structure-based sequence alignment of MINERVA with Pleckstrin revealed that MINERVA possess several negatively-charged amino acids on the lipid binding site, which might explain the different preference for lipids (Figure 31, Figure 33, and Figure 34). In MINERVA, Gln77 is present in place of Gly255 of Pleckstrin, which seem to cause steric hindrance with incoming lipid molecule. The high affinity of MINERVA PH domain towards phosphatidic acid can be explained by difference in amino acid at this site, as phosphatidylinositol phosphates have bulky cyclohexane moiety at its head group while the head group of the phosphatidic acid is only composed of a single phosphate. To our surprise, the C-terminal flexible region of MINERVA effectively blocks the lipid binding site on the PH domain in a self-inhibitory manner, according to the lipid-dot blot assay results (Figure 35). However, trypsin-digestion of the full length MINERVA produced similar, if not the same, fragments from trypsin-digestion of MINERVA[9–553] or MINERVA[1–574], the full length MINERVA. If the C-terminal flexible region was to bind the structured domain determined in this study, it would have been protected from trypsin as was the case for most other trypsin cleavage sites present in the MINERVA sequence. No additional peptide fragment was observed from the SDS-PAGE analysis of the trypsin-digested full length MINERVA, meaning that the C-terminal flexible region was completely obliterated. Our result suggests that the regulation mechanism of the PH domain would be more subtle than a direct binding of the C-terminal flexible region to the lipid binding site on the PH domain.

MINERVA is heavily involved in growth signaling pathway when dispersed in cytoplasm. The C-terminal flexible region harbors many target sites for phosphorylation by Erk (Old, et al., 2009, Mol Cell, 6) and EGFR (Ji, et al., 2016,

Proc Natl Acad Sci U S A, 15), and also contains $DLGX_nETGE$ motif to compete with Nrf2 for interaction with Keap1 (Cheng, et al., 2019, EBioMedicine, 12). Effect of cytosolic MINERVA includes increased proliferation, differentiation and invasion, as well as resistance against oxidative stress, all of which provide survival advantage to cancer cells. On the other hand, MINERVA migrates to plasma membrane and colocalizes with β-catenin at adherens junctions in confluent cells regardless of phosphorylation state of the C-terminal flexible region (Chen, et al., 2011, J Biol Chem, 16), suggesting that the auto-inhibition of the PH domain is not regulated by the phosphorylation of the C-terminal flexible region. Therefore, I speculated that there is a yet unknown regulation mechanism for inhibition of the lipid binding site on PH domain by the C-terminal flexible region, which is lifted in response to certain events such as contact inhibition.

In this study, I report the first structure on MINERVA, a member of Niban-like protein 1 family, which is also known as FAM129B. Phase data obtained from SeMet-substituted MINERVA[1–574] crystals were utilized to model the structure of MINERVA[9–553]. MINERVA is composed of an N-terminal PH domain and an α-helix bundle of novel fold. Lipid dot-blot assay showed that MINERVA display characteristic preference for specific lipids, and that C-terminal flexible region on MINERVA facilitate competent auto-inhibition of the PH domain. In conclusion, my findings on MINERVA provides valuable insight into the structural arrangement and supports previously reported functional analyses.

Bibliography

1.   Yoon HS, Hajduk PJ, Petros AM, Olejniczak ET, Meadows RP, Fesik SW. Solution structure of a pleckstrin-homology domain. Nature 1994;369(6482):672-675.
2.   Lemmon MA. Pleckstrin homology (PH) domains and phosphoinositides. Biochem Soc Symp 2007(74):81-93.
3.   Lemmon MA. Pleckstrin homology domains: not just for phosphoinositides. Biochem Soc Trans 2004;32(Pt 5):707-711.
4.   Bar-Shavit R, Grisaru-Granovsky S, Uziely B. PH-domains as central modulators driving tumor growth. Cell Cycle 2016;15(5):615-616.
5.   Scheffzek K, Welti S. Pleckstrin homology (PH) like domains - versatile modules in protein-protein interaction platforms. FEBS Lett 2012;586(17):2662-2673.
6.   Old WM, Shabb JB, Houel S, Wang H, Couts KL, Yen CY, Litman ES, Croy CH, Meyer-Arendt K, Miranda JG, Brown RA, Witze ES, Schweppe RE, Resing KA, Ahn NG. Functional proteomics identifies targets of phosphorylation by B-Raf signaling in melanoma. Mol Cell 2009;34(1):115-131.
7.   Sun GD, Kobayashi T, Abe M, Tada N, Adachi H, Shiota A, Totsuka Y, Hino O. The endoplasmic reticulum stress-inducible protein Niban regulates eIF2alpha and S6K1/4E-BP1 phosphorylation. Biochem Biophys Res Commun 2007;360(1):181-187.
8.   Boyd RS, Adam PJ, Patel S, Loader JA, Berry J, Redpath NT, Poyser HR, Fletcher GC, Burgess NA, Stamps AC, Hudson L, Smith P, Griffiths M, Willis TG, Karran EL, Oscier DG, Catovsky D, Terrett JA, Dyer MJ. Proteomic analysis of the cell-surface membrane in chronic lymphocytic leukemia: identification of two novel proteins, BCNP1 and MIG2B. Leukemia 2003;17(8):1605-1612.
9.   Bamford S, Dawson E, Forbes S, Clements J, Pettett R, Dogan A, Flanagan A, Teague J, Futreal PA, Stratton MR, Wooster R. The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. Br J Cancer 2004;91(2):355-358.
10.   Forbes SA, Beare D, Gunasekaran P, Leung K, Bindal N, Boutselakis H, Ding M, Bamford S, Cole C, Ward S, Kok CY, Jia M, De T, Teague JW, Stratton MR, McDermott U, Campbell PJ. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic Acids Res 2015;43(Database issue):D805-811.
11.   Zhou X, Yang F, Zhang Q, Miao Y, Hu X, Li A, Hou G, Wang Q, Kang J. FAM129B promoted tumor invasion and proliferation via

facilitating the phosphorylation of FAK signaling and associated with adverse clinical outcome of non-small cell lung cancer patients. Onco Targets Ther 2018;11:7493-7501.

12. Cheng KC, Lin RJ, Cheng JY, Wang SH, Yu JC, Wu JC, Liang YJ, Hsu HM, Yu J, Yu AL. FAM129B, an antioxidative protein, reduces chemosensitivity by competing with Nrf2 for Keap1 binding. EBioMedicine 2019;45:25-38.

13. Deramaudt TB, Dujardin D, Hamadi A, Noulet F, Kolli K, De Mey J, Takeda K, Ronde P. FAK phosphorylation at Tyr-925 regulates cross-talk between focal adhesion turnover and cell protrusion. Mol Biol Cell 2011;22(7):964-975.

14. Oishi H, Itoh S, Matsumoto K, Ishitobi H, Suzuki R, Ema M, Kojima T, Uchida K, Kato M, Miyata T, Takahashi S. Delayed cutaneous wound healing in Fam129b/Minerva-deficient mice. J Biochem 2012;152(6):549-555.

15. Ji H, Lee JH, Wang Y, Pang Y, Zhang T, Xia Y, Zhong L, Lyu J, Lu Z. EGFR phosphorylates FAM129B to promote Ras activation. Proc Natl Acad Sci U S A 2016;113(3):644-649.

16. Chen S, Evans HG, Evans DR. FAM129B/MINERVA, a novel adherens junction-associated protein, suppresses apoptosis in HeLa cells. J Biol Chem 2011;286(12):10201-10209.

17. Conrad W, Major MB, Cleary MA, Ferrer M, Roberts B, Marine S, Chung N, Arthur WT, Moon RT, Berndt JD, Chien AJ. FAM129B is a novel regulator of Wnt/beta-catenin signal transduction in melanoma cells. F1000Res 2013;2:134.

18. Thinon E, Serwa RA, Broncel M, Brannigan JA, Brassat U, Wright MH, Heal WP, Wilkinson AJ, Mann DJ, Tate EW. Global profiling of co- and post-translationally N-myristoylated proteomes in human cells. Nat Commun 2014;5:4919.

19. Furukawa M, Xiong Y. BTB protein Keap1 targets antioxidant transcription factor Nrf2 for ubiquitination by the Cullin 3-Roc1 ligase. Mol Cell Biol 2005;25(1):162-171.

20. Kensler TW, Wakabayashi N, Biswal S. Cell survival responses to environmental stresses via the Keap1-Nrf2-ARE pathway. Annu Rev Pharmacol Toxicol 2007;47:89-116.

21. Ward JJ, McGuffin LJ, Bryson K, Buxton BF, Jones DT. The DISOPRED server for the prediction of protein disorder. Bioinformatics 2004;20(13):2138-2139.

22. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. Bioinformatics 2000;16(4):404-405.

23. Slabinski L, Jaroszewski L, Rychlewski L, Wilson IA, Lesley SA, Godzik A. XtalPred: a web server for prediction of protein crystallizability. Bioinformatics 2007;23(24):3403-3405.

24. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307-326.

25. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta crystallographica Section D, Biological crystallography 2010;66(Pt 2):213-221.

26. Terwilliger TC. Automated main-chain model building by template matching and iterative fragment extension. Acta crystallographica Section D, Biological crystallography 2003;59(Pt 1):38-44.

27. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta crystallographica Section D, Biological crystallography 2010;66(Pt 4):486-501.

28. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta crystallographica Section D, Biological crystallography 1997;53(Pt 3):240-255.

29. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. Towards automated crystallographic structure refinement with phenix.refine. Acta crystallographica Section D, Biological crystallography 2012;68(Pt 4):352-367.

30. Vagin A, Teplyakov A. Molecular replacement with MOLREP. Acta crystallographica Section D, Biological crystallography 2010;66(Pt 1):22-25.

31. Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. Acta crystallographica Section D, Biological crystallography 2010;66(Pt 1):12-21.

32. Lee JH, Ji H, Lu Z. FAM129B activates Ras and promotes aerobic glycolysis. Cell Cycle 2016;15(11):1391-1392.

33. Ostrem JM, Peters U, Sos ML, Wells JA, Shokat KM. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. Nature 2013;503(7477):548-551.

34. Catimel B, Kapp E, Yin MX, Gregory M, Wong LS, Condron M, Church N, Kershaw N, Holmes AB, Burgess AW. The PI(3)P interactome from a colon cancer cell. J Proteomics 2013;82:35-51.

35. Marat AL, Haucke V. Phosphatidylinositol 3-phosphates-at the interface between cell signalling and membrane traffic. EMBO J 2016;35(6):561-579.

36. Harris TJ, Tepass U. Adherens junctions: from molecules to morphogenesis. Nat Rev Mol Cell Biol 2010;11(7):502-514.

37. Daniels DL, Weis WI. Beta-catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. Nat Struct Mol Biol 2005;12(4):364-371.

38. Kowalczyk AP, Nanes BA. Adherens junction turnover: regulating adhesion through cadherin endocytosis, degradation, and recycling. Subcell Biochem 2012;60:197-222.

39. Kam Y, Quaranta V. Cadherin-bound beta-catenin feeds into the Wnt pathway upon adherens junctions dissociation: evidence for an intersection between beta-catenin pools. PLoS One 2009;4(2):e4580.

40. van Noort M, Meeldijk J, van der Zee R, Destree O, Clevers H. Wnt signaling controls the phosphorylation status of beta-catenin. J Biol Chem 2002;277(20):17901-17905.

41. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J Mol Biol 2007;372(3):774-797.

42. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, Lopez R. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res 2019;47(W1):W636-W641.

43. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res 2014;42(Web Server issue):W320-324.

44. Holm L. Benchmarking Fold Detection by DaliLite v.5. Bioinformatics 2019.

45. Jackson SG, Zhang Y, Haslam RJ, Junop MS. Structural analysis of the carboxy terminal PH domain of pleckstrin bound to D-myo-inositol 1,2,3,5,6-pentakisphosphate. BMC Struct Biol 2007;7:80.

국문 초록

FAM129B 라는 이름으로도 알려진 MINERVA (Melanoma invasion by Erk) 단백질은 척추동물에만 존재하는 FAM129 단백질 군의 일원이다. MINERVA 는 EGFR/Erk 와 Wnt/β-catenin 를 비롯한 다양한 신호전달체계에 관련되어 있으며 많은 종류의 암에서 발현이 증가하여 암세포 침윤을 유발하는 것이 보고되었다. 그러나, MINERVA 의 정확한 작용기전은 알려져 있지 않다. 본 논문에서는 C-말단의 유연 부위가 제거된 MINERVA$^{9-553}$ 의 구조를 X-선 결정학 방법을 통하여 규명하였다. 구조를 규명하기에 충분한 회절을 하는 MINERVA$^{9-553}$ 결정을 얻는데 있어 Trypsin 소화과정이 필수적이었다. N-말단의 PH domain 은 전형적인 PH domain 의 구조를 갖고 있으나, 지질결합부위 주위의 아미노산 구성이 결합 가능한 지질을 제한 할 것으로 보인다. α-나선이 풍부한 이후 domain 은 지금껏 알려지지 않은 새로운 접힘 구조를 갖고 있는 것으로 나타났다. 전체적인 MINERVA$^{9-553}$ 의 구조로 원형질 막과의 상호작용을 추정할 수 있으며, C-말단의 유연 부위이 이 상호작용을 조절할 것으로 보인다. MINERVA$^{9-553}$ 의 구조를 밝힘으로써 새로운 단백질 접힘 구조를 발견하였으나, 동시에 C-말단 유연 부위의 PH domain 에 대한 자기조절작용 기제의 규명 과제 또한 제시한다.

주 요 어 : MINERVA; FAM129B; NIBAN2; PH domain; 새로운 접힘 구조; 자기조절작용

학   번 : 2013-30516

Acknowledgements