Proteolytic Activation of Human-specific Olduvai Domains by the Furin Protease

Ashley Pacheco[1,†], **Aaron Issaian**[1,†], **Jonathan Davis**[1], **Nathan Anderson**[1], **Travis Nemkov**[1], **Natasia Paukovich**[1], **Morkos A. Henen**[1], **Beat Vögeli**[1], **James M. Sikela**[1,*], **Kirk Hansen**[1,*]

[1]Department of Biochemistry and Molecular Genetics, University of Colorado, Aurora, CO, USA

Abstract

Olduvai protein domains (formerly DUF1220) show the greatest human-specific increase in copy number of any coding region in the genome and are highly correlated with human brain evolution and cognitive disease. The majority of human copies are found within four *NBPF* genes organized in a variable number of a tandemly arranged three-domain blocks called Olduvai triplets. Here we show that these human-specific Olduvai domains are posttranslationally processed by the furin protease, with a cleavage site occurring once at each triplet. These findings suggest that all expanded human-specific *NBPF* genes encode proproteins consisting of many independent Olduvai triplet proteins which are activated by furin processing. The exceptional correlation of Olduvai copy number and brain size taken together with our new furin data, indicates the ultimate target of selection was a rapid increase in dosage of autonomously functioning Olduvai triplet proteins, and that these proteins are the primary active agent underlying Olduvai's role in human brain expansion.

Keywords

furin; protease; gene expression; genetic disease; brain

[*]Author for Correspondence: Kirk Hansen, Department of Biochemistry and Molecular Genetics, University of Colorado, Aurora, CO, USA, kirk.hansen@cuanschutz.edu; James M. Sikela, Department of Biochemistry and Molecular Genetics, University of Colorado, Aurora, CO, USA, james.sikela@cuanschutz.edu.
[†]These authors contributed equally to this work.

Introduction

Olduvai protein domains (formerly called DUF1220) are approximately 65 amino acids in length and are found almost exclusively within *NBPF* genes (1). Studies have shown that Olduvai copy number is linearly associated with primate brain size, neuron number, and several other brain size-related phenotypes (2–6). In humans, Olduvai copy number is associated in a dosage-dependent manner to brain size, gray matter volume, and cognitive aptitude (3, 7). Olduvai domains have undergone strong positive selection among primates and Olduvai copy number increases with evolutionary proximity to humans (8). Non-primate mammals have 1–9 copies, monkeys have 25–40 copies, great apes have 90–125 copies, and humans have 250–350 copies (9). The hyper-amplification of Olduvai in humans represents the largest human-specific increase in copy number of any coding region in the genome (6, 8, 9). Interestingly, all four of the NBPF genes with the largest human-specific Olduvai expansions map to the 1q21.1–3 region of chromosome 1, and three of these are adjacent to, and tightly co-regulated with, three human specific *NOTCH2NL* genes (10–12). The *NOTCH2NL* genes are known to contribute to the expansion of cortical progenitors, neuronal output and ultimately the expansion of the human cortex (10, 13, 14). Taken together, these findings suggest that Olduvai and *NOTCH2NL* may function in a coordinated, complementary, dosage-related manner to promote human brain enlargement through expansion of radial glial cell populations (10).

The majority of Olduvai domains follow a unique genomic organization that allows copy number to be increased both rapidly and in an open-ended manner (5, 9). Olduvai domains are encoded by a small and large exon doublet (8, 15, 16) and can be phylogenetically subdivided into six primary clades: conserved (CON1-3) and human lineage-specific (HLS1-3) (9). The typical order of domain subtypes within human *NBPF* genes is: one to several CON1 domains; one CON2 domain; one-to-many instances of a triplet composed of HLS1, HLS2, and HLS3 domains; lastly one CON3 domain. The three-domain block of HLS sequences is known as the Olduvai triplet (9, 17).

The rapid and extreme human lineage-specific Olduvai increase is predominantly the result of tandem additions of Olduvai triplets within *NBPF* genes supplemented by gene duplications after Olduvai triplet expansion (9, 10). The proposed mechanism for the human expansion of *NBPF* genes is through a recombination event between HLS1 and CON3 which introduced a putative G-quadraplex (pG4) sequence pattern (18). G4 structures promote instability which in conjunction with selective pressure is thought to have contributed to the human-specific extreme expansion of *NBPF* genes through tandem additions of Olduvai triplets (18). In *NBPF* genes that contain one or more Olduvai triplets, the first triplet contains a canonical HLS1 small exon (HLS1-C), while every additional triplet contains a larger HLS1 small exon (HLS1-E) that is also found in the CON3 subtype (18). In addition to the pG4 sequence, the extended exon of HLS1 also contains additional protein processing elements that have importance for human specific Olduvai function and will be discussed in this work.

The Olduvai protein domain family has been proposed to represent a genomic trade-off where selective pressure for increased copies resulted in increased brain size and cognition

but, in certain situations, copies can also vary in deleterious ways (17). Indeed, Olduvai copy number variation has been linked with multiple brain disorders including autism spectrum disorder (ASD), macrocephaly, microcephaly and schizophrenia (2, 3, 19–22). Genomic data indicates that there are 23 human *NBPF* genes that carry from 5 to 60 tandemly arranged Olduvai domains. Such an arrangement suggests that the proteins encoded by human *NPBF* genes would span a wide range of sizes. Surprisingly, this is not the case. Rather Western analysis shows only a single primary band of ~36 kDa. This result has remained an enigma since it was first observed approximately 15 years ago (8). Here we present novel furin processing data confirming that expanded human-specific *NBPF* genes are posttranslationally cleaved, converting larger NBPF proproteins into smaller active protein units.

Results

Mass Spec Analysis and Validation of Olduvai HLS Peptide Size

As mentioned above, predicted protein sizes of human NBPF proteins span a wide range, but western blot analysis of human tissues shows only a single band for the HLS domains at ~36 kDa (Supporting Figure 1A depicts previously published data from Popesco et al. 2006)(8). This product could contain 3–4 Olduvai domains and suggests that the functional size of HLS protein is smaller (5).

To further investigate this observation, the 36 kDa band was enriched and analyzed by mass spectrometry-based proteomics. Chromatographic pre-enrichment of the 36 kDa band combined with data-dependent global proteomic analysis identified a characteristic peptide of the Olduvai HLS subtype (Supporting Figure 1B). Subsequent targeted proteomics analysis was able to identify additional peptides specific to NBPF including HLS domains (Supplemental Figure 2).

Predicted Furin Processing Sites

Analysis of potential protease cut sites determined that translated HLS1-E exons contain a predicted furin cleavage site (Figure 1A). The HLS1-E exon, including the predicted furin site, is found within every expanded Olduvai triplet and is absent from all unexpanded triplets (i.e. those containing the HLS-C exon) (Figure 1B).

Furin Protease Processing Cleaves Extended HLS Domains

To evaluate potential furin processing of Olduvai triplets, both HLS1-C and HLS1-E were co-expressed with furin. A single band is observed for HLS1-C while a multiple banding pattern is observed for HLS1-E with furin co-expression (Figure 2A). These results confirm that extended HLS1 domains are proteolytically processed by furin from longer 'proproteins'. The resulting predicted protein products from furin processing would include the longer N-terminal region of the protein including CON1 and CON2 and Olduvai triplets with HLS1-C, all Olduvai triplets containing the alternative HLS1-E, and finally the CON3 domain (Figure 2B). Furin processing of HLS1 would occur once at every Olduvai triplet within expanded *NBPF* genes generating numerous Olduvai triplet products from each gene. Confirmation of processing by furin resolves the long-standing question as to why only one

protein band size is observed for HLS domains, despite the wide range of predicted protein sizes.

The resulting predicted size for the expanded Olduvai triplet protein is approximately ~28 kDa, which is lower than the observed ~36 kDa band. This higher than expected observed molecular weight (MW) for this product could simply be due to the intrinsically disordered nature of the triplet protein that has been previously determined (23, 24). Alternatively, it may also indicate presence of posttranslational modifications (PTMs). As protein phosphorylation is the most common form of PTM, and intrinsically disordered proteins are disproportionately phosphorylated relative to other proteins (25), predictive analysis for protein phosphorylation was conducted. Prediction of protein phosphorylation sites shows the presence of multiple predicted sites within the cleaved protein product, which could account for the increased observed MW of the cleaved Olduvai triplet protein (Supporting Figure 3). This finding is preliminary and further investigation will be needed to determine whether other forms of PTM are involved.

Discussion

The evolutionary hyper-amplification of the Olduvai protein domain family in humans represents one of the most dramatic human lineage-specific genomic events. Due to several recent studies, we are beginning to understand the specific steps of how this occurred primarily through the following two processes. The first is that human Olduvai expansion has employed a very rapid and highly expandable mechanism for increasing copy number by addition of tandemly arrayed Olduvai triplets (18). The second being the addition of a furin cleavage site within each triplet which allows expanded copies to be readily converted via posttranslational processing to smaller independent active protein units, i.e., Olduvai triplets. It is likely that these events were driven by strong evolutionary selection pressures as they are seen within several human *NBPF* genes, and have added over 50 Olduvai triplets to the human genome over the past 2–3 million years (10). Taken together, these findings indicate that, rather than simply raising copy number, the primary target of selection in humans has been to rapidly increase the dosage of autonomously functioning Olduvai triplet proteins, and these are the primary active agents of Olduvai function in humans.

These results also have relevance to the findings that human-specific Olduvai domains and adjacent NOTCH2NL genes may work in a coordinated manner to increase brain size (10). If independent Olduvai triplets are the active subunits of NBPF proteins, they may drive coordinated function with NOTCH2NL proteins. Previous findings therefore suggest that NOTCH2NL and NBPF proteins have shared transcriptional co-regulation (10). This coordinated function may underly selective pressure for expanded Olduvai triplets in the three human NBPF genes that are in close proximity to, and co-evolved with, human NOTCH2NL genes (10). Interestingly, NBPF20, which has the highest copy number of human-specific Olduvai triplets and also maps to 1q21, is not adjacent to a NOTCH2NL gene. Phylogenetic analysis indicates that the NBPF20 Olduvai expansion arose independently from other highly expanded NBPF genes (10). This suggests that selective pressure favoring the addition of more human-specific Olduvai triplets is

independent of NOTCH2NL and that Olduvai has autonomous expression and functionality from NOTCH2NL genes.

Notably, processing by furin indicates that the N-terminal region including CON1 and CON2 as well as the C-terminal CON3 may be separated from extended HLS domains after translation. This finding suggests that CON and HLS domains may have independent functionality, which is supported by differing associations for various Olduvai domain subtypes. For example, one of the most striking and well-supported findings is the linear association between CON1 copy number and ASD severity (10, 20, 21). This link is not found with any other Olduvai subtypes, despite their being adjacent to CON1 in the NBPF proprotein. Independent function of the NBPF N-terminal protein product may have relevance to associations between CON1 and ASD severity and is beyond the scope of this paper.

Finally, the new insights presented here resolve a long-standing question regarding the function of Olduvai domains in humans, and show how selection has favored increasing levels of individual Olduvai triplet proteins. How this dosage phenomenon is tied to the evolutionary increase in human brain size is an interesting future area of investigation.

Experimental procedures

Mass Spec Analysis

Untargeted proteomic identification was performed on protein digests using liquid chromatography-tandem mass spectrometry as previously reported (26). Briefly, nanoflow reverse-phase LC-MS/MS was performed using an Eksigent nanoLC2D system coupled to a LTQ Orbitrap-Velos mass spectrometer (Thermo Fisher). Data acquisition was performed using XcaliburTM (Version 2.1, Thermo Fisher) software. Raw data files were converted into Mascot generic files using PAVA. These peak lists were searched against SwissProt human database using Mascot server (Version 2.2.06, Matrix Science). Mass tolerances were 15 ppm for MS peaks, and 0.5 Da for MS/MS fragment ions.

Enrichment of 36 kDa Band

To enrich the 36 kDa protein with antibody positivity for proteomics analysis, 100 mg of human brain tissue (4) was extracted and subjected to a series of fractionations during which the presence of the 36 kDa band was followed by western analysis using an antibody directed against the HLS subtype sequence. Antibody positive fractions were then dissolved in 0.1% formic acid and analyzed by targeted proteomics. Multiple reaction monitoring MRM was performed using an LC-MS/MS system interfaced with a an UPLC system (QTRAP 5500 and Ultimate 3000, respectively, Sciex and Thermo Fisher Scientific). The mass spectrometer was run in positive ionization mode with the following settings: a source temperature of 200°C, spray voltage of 5300 V, curtain gas of 20 psi, and a source gas of 35 psi (nitrogen gas). Multiple SRM transitions were monitored using unit resolution in both Q1 and Q3 quadrupoles to maximize specificity. SRM assay optimization was performed with the aid of computer software (Skyline, Version 3.1). Collision energies and declustering potential were optimized for each transition. Method building and acquisition

were performed using the instrument-supplied software (Analyst, Version 1.5.2, AB Sciex). Raw SRM data files were imported to Skyline Version 3.1 software for data processing.

Protein Expression and Purification

Codon optimized gene fragments of HLS-C and HLS-E were cloned into pET21b(+) with the NdeI and XhoI restriction sites. Each expression construct contained an N-terminal MBP tag, a TEV cleavage site, and C-terminal 6xHis-tag. The plasmids were transformed and expressed in *Escherichia coli* strain Rosetta™ 2(DE3)pLysS cells with appropriate antibiotic selection. All bacterial expressed recombinant proteins were grown in Luria–Bertani (LB) broth.

HLS-C and HLS-E were expressed as follows. Three colonies from an overnight culture plate were resuspended in 1.5 L of LB with 0.001% Antifoam 204 and shaken at 26 °C for 14 hours. The A600 was measured to ensure the culture was below 0.5. Fresh LB was inoculated with the overnight culture (3:1) and shaken at 37 °C. The incubator temperature was lowered to 25 °C once the A600 reached 0.4. The culture was induced with 0.5 mM Isopropyl β- d-1-thiogalactopyranoside (IPTG) once the A600 reached 0.6 and shaken for 6 hours at 25 °C. Cells were harvested by centrifugation at 4 °C for 10 min at 4500 rpm.

HLS-C and HLS-E were purified as follows. Cell pellets were resuspended in 50 mM Tris (pH 7.5), 300 mM NaCl, 8 M urea, 2% glycerol, 10 mM imidazole, and 10 mM βME. Cell disruption was performed by sonication (7×20 s) at 4 °C. The lysates were clarified by centrifugation at 13,000 rpm for 30 min at 4 °C and loaded onto a pre-equilibrated column packed with Ni Sepharose™ excel resin. The column was washed with 5 CV of resuspension buffer and His-tagged proteins were eluted with 50 mM Tris (pH 7.5), 50 mM NaCl, 8 M urea, 3 % glycerol, and 400 mM imidazole. The protein sample was diluted and loaded onto a pre-equilibrated column packed with Source™ 15Q. Proteins were separated using the following gradient: 0 – 50 % B (0 – 12 CV), 50 – 75 % B (12 – 15 CV), 75 – 100 % B (15 – 15.5 CV), 100 % B (15.5 – 16.5 CV) with buffer A (20 mM Tris (pH 8.0), 8 M urea) and buffer B (20 mM Tris (pH 8.0), 8 M urea, 1 M NaCl). The relevant fractions were pooled and dialyzed overnight at 4 °C against 20 mM Tris (pH 7.5), 50 mM NaCl, and 5 mM DTT. TEV protease (in-house) was added to the sample and incubated overnight at 4 °C with gentle rocking. The protein sample was loaded onto a pre-equilibrated column packed with Source™ 15Q. Proteins were separated using the following gradient: 0 – 50 % B (0 – 12 CV), 50 – 75 % B (12 – 15 CV), 75 – 100 % B (15 – 15.5 CV), 100 % B (15.5 – 16.5 CV) with buffer A (20 mM Tris (pH 8.0), 8 M urea) and buffer B (20 mM Tris (pH 8.0), 8 M urea, 1 M NaCl). Relevant fractions were pooled and dialyzed overnight at 4 °C against 20 mM HEPES (pH 7.5), 200 mM NaCl, and 5 mM DTT. The samples were concentrated to 1.5 mL using a 3000 MWCO concentrator (Sartorius) and diluted to 3 mL with urea buffer (20 mM Tris (pH 7.5), 300 mM NaCl, 8 M urea, 5 mM DTT). The samples were injected onto a Superose™ 6 Increase column (Cytiva) pre-equilibrated with native storage buffer (20 mM HEPES (pH 7.5), 200 mM NaCl, 5 mM DTT). Fractions containing pure target protein were pooled, concentrated, and stored at −80 °C.

In-Solution Enzymatic Digestion

Both HLS-C and HLS-E protein at 1 mg/mL were incubated with human Furin (EC 3.4.21.75, P09958, obtained from NEB P8077) in digestion buffer (10 mM HEPES (pH 7.5), 0.2 mM CaCl2, 0.2 mM βME, 0.1% Triton X-100) for 16 hours at RT (25°C). Additional samples lacking either Olduvai Triplet or Furin were assembled as experimental controls. Samples were mixed with 1X LDS Sample Buffer (GenScript) supplemented with 100 mM βME and incubated at 95 °C for 5 min. Equal volumes of each sample were loaded onto a 4–12% SDS-PAGE gel along with the BLUEstainTM 2 (GoldBio) protein ladder. The gel was stained with SimplyBlueTM Safe Stain (Thermo Fisher Scientific) and visualized with a CCD equipped imager (Bio-Rad).

In-situ Analyses

Furin site prediction was carried out by submitting full length NBPF protein sequences to the ProP – 1.0 server (27). Protein phosphorylation site prediction was conducted by submitting Olduvai triplet protein sequences containing cleaved HLS1-E exons to the NetPhos – 3.1 server (28).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Data availability

The data underlying this article are available in the article and in its online supplementary material.

References

1. Sikela JM, and van Roy F (2018) Changing the name of the NBPF/DUF1220 domain to the Olduvai domain. F1000Research. 10.12688/f1000research.13586.2

2. Dumas L, and Sikela JM (2009) DUF1220 domains, cognitive disease, and human brain evolution. in Cold Spring Harbor Symposia on Quantitative Biology, 10.1101/sqb.2009.74.025

3. Dumas LJ, O'bleness MS, Davis JM, Dickens CM, Anderson N, Keeney JG, Jackson J, Sikela M, Raznahan A, Giedd J, Rapoport J, Nagamani SSC, Erez A, Brunetti-Pierri N, Sugalski R, Lupski JR, Fingerlin T, Cheung SW, and Sikela JM (2012) DUF1220-domain copy number implicated in human brain-size pathology and evolution. Am. J. Hum. Genet. 10.1016/j.ajhg.2012.07.016

4. Keeney JG, Davis JM, Siegenthaler J, Post MD, Nielsen BS, Hopkins WD, and Sikela JM (2015) DUF1220 protein domains drive proliferation in human neural stem cells and are associated with increased cortical volume in anthropoid primates. Brain Struct. Funct. 10.1007/s00429-014-0814-9

5. Keeney JG, Dumas L, and Sikela JM (2014) The case for DUF1220 domain dosage as a primary contributor to anthropoid brain expansion. Front. Hum. Neurosci. 10.3389/fnhum.2014.00427

6. Zimmer F, and Montgomery SH (2015) Phylogenetic analysis supports a link between DUF1220 domain number and primate brain expansion. Genome Biol. Evol. 10.1093/gbe/evv122

7. Davis JM, Searles VB, Anderson N, Keeney J, Raznahan A, Horwood, · John L, Fergusson DM, Kennedy MA, Giedd J, and Sikela JM (2015) DUF1220 copy number is linearly associated with increased cognitive function as measured by total IQ and mathematical aptitude scores. Hum Genet. 134, 67–75 [PubMed: 25287832]

8. Popesco MC, MacLaren EJ, Hopkins J, Dumas L, Cox M, Meltesen L, McGavran L, Wyckoff GJ, and Sikela JM (2006) Human lineage-specific amplification, selection, and neuronal expression of DUF1220 domains. Science (80-.). 10.1126/science.1127980

9. O'Bleness MS, Michael Dickens C, Dumas LJ, Kehrer-Sawatzki H, Wyckoff GJ, and Sikela JM (2012) Evolutionary history and genome organization of duf1220 protein domains. G3 Genes, Genomes, Genet. 10.1534/g3.112.003061

10. Fiddes IT, Pollen AA, Davis JM, and Sikela JM (2019) Paired involvement of human-specific Olduvai domains and NOTCH2NL genes in human brain evolution. Hum. Genet. 10.1007/s00439-019-02018-4

11. O'Bleness M, Searles VB, Dickens CM, Astling D, Albracht D, Mak ACY, Lai YYY, Lin C, Chu C, Graves T, Kwok PY, Wilson RK, and Sikela JM (2014) Finished sequence and assembly of the DUF1220-rich 1q21 region using a haploid human genome. BMC Genomics. 10.1186/1471-2164-15-387

12. Fiddes IT, Lodewijk GA, Mooring M, Bosworth CM, Ewing AD, Mantalas GL, Novak AM, van den Bout A, Bishara A, Rosenkrantz JL, Lorig-Roach R, Field AR, Haeussler M, Russo L, Bhaduri A, Nowakowski TJ, Pollen AA, Dougherty ML, Nuttle X, Addor MC, Zwolinski S, Katzman S, Kriegstein A, Eichler EE, Salama SR, Jacobs FMJ, and Haussler D (2018) Human-Specific NOTCH2NL Genes Affect Notch Signaling and Cortical Neurogenesis. Cell. 10.1016/j.cell.2018.03.051

13. Florio M, Heide M, Pinson A, Brandl H, Albert M, Winkler S, Wimberger P, Huttner WB, and Hiller M (2018) Evolution and cell-type specificity of human-specific genes preferentially expressed in progenitors of fetal neocortex. Elife. 10.7554/eLife.32332

14. Suzuki IK, Gacquer D, Van Heurck R, Kumar D, Wojno M, Bilheu A, Herpoel A, Lambert N, Cheron J, Polleux F, Detours V, and Vanderhaeghen P (2018) Human-Specific NOTCH2NL Genes Expand Cortical Neurogenesis through Delta/Notch Regulation. Cell. 10.1016/j.cell.2018.03.067

15. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE, Paladin L, Raj S, Richardson LJ, Finn RD, and Bateman A (2021) Pfam: The protein families database in 2021. Nucleic Acids Res. 10.1093/nar/gkaa913

16. Vandepoele K, Van Roy N, Staes K, Speleman F, and Van Roy F (2005) A novel gene family NBPF: Intricate structure generated by gene duplications during primate evolution. Mol. Biol. Evol. 10.1093/molbev/msi222

17. Sikela JM, and Searles Quick VB (2018) Genomic trade-offs: are autism and schizophrenia the steep price of the human brain? Hum. Genet. 10.1007/s00439-017-1865-9

18. Heft IE, Mostovoy Y, Levy-Sakin M, Ma W, Stevens AJ, Pastor S, McCaffrey J, Boffelli D, Martin DI, Xiao M, Kennedy MA, Kwok PY, and Sikela JM (2020) The driver of extreme human-specific olduvai repeat expansion remains highly active in the human genome. Genetics. 10.1534/genetics.119.302782

19. Davis JM, Heft I, Scherer SW, and Sikela JM (2019) A third linear association between Olduvai (DUF1220) copy number and severity of the classic symptoms of inherited autism. Am. J. Psychiatry. 10.1176/appi.ajp.2018.18080993

20. Davis JM, Searles Quick VB, and Sikela JM (2015) Replicated linear association between DUF1220 copy number and severity of social impairment in autism. Hum. Genet. 10.1007/s00439-015-1537-6

21. Davis JM, Searles VB, Anderson N, Keeney J, Dumas L, and Sikela JM (2014) DUF1220 Dosage Is Linearly Associated with Increasing Severity of the Three Primary Symptoms of Autism. PLoS Genet. 10.1371/journal.pgen.1004241

22. Searles Quick VB, Davis JM, Olincy A, and Sikela JM (2015) DUF1220 copy number is associated with schizophrenia risk and severity: Implications for understanding autism and schizophrenia as related diseases. Transl. Psychiatry. 10.1038/tp.2015.192

23. Issaian A, Schmitt L, Born A, Nichols PJ, Sikela J, Hansen K, Vögeli B, and Henen MA (2019) Solution NMR backbone assignment reveals interaction-free tumbling of human lineage-specific Olduvai protein domains. Biomol. NMR Assign. 10.1007/s12104-019-09902-0

24. Wu H, Zhai LT, Guo XX, Rety S, and Xi XG (2020) The N-terminal of NBPF15 causes multiple types of aggregates and mediates phase transition. Biochem. J. 10.1042/BCJ20190566

25. Bah A, and Forman-Kay JD (2016) Modulation of intrinsically disordered protein function by post-translational modifications. J. Biol. Chem. 10.1074/jbc.R115.695056

26. Dzieciatkowska M, Hill R, and Hansen KC (2014) GeLC-MS/MS analysis of complex protein mixtures. Methods Mol. Biol. 10.1007/978-1-4939-0685-7_4

27. Duckert P, Brunak S, and Blom N (2004) Prediction of proprotein convertase cleavage sites. Protein Eng. Des. Sel. 10.1093/protein/gzh013

28. Blom N, Gammeltoft S, and Brunak S (1999) Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. J. Mol. Biol. 10.1006/jmbi.1999.3310
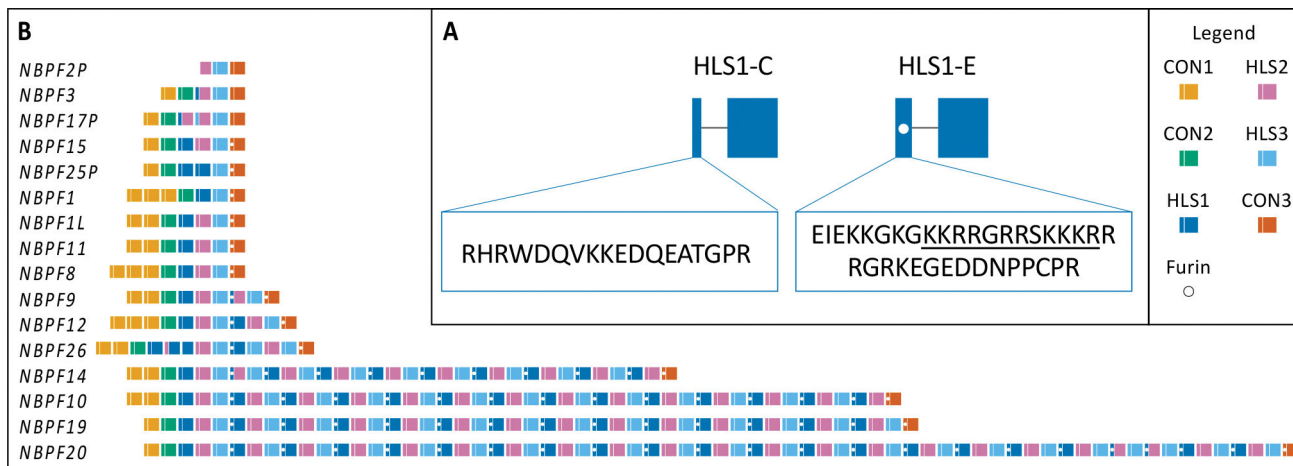
Fig. 1.
(A) Analysis of protease cleavage sites predicted a furin site located in the small exon of HLS1-E indicated by underlined amino acid sequence and the white dot in the figure. The small exons of canonical HLS1 (HLS1-C) and expanded HLS1 (HLS1-E) domains are enlarged to show the translated amino acid sequence. (B) All expanded NBPF Olduvai triplets contain a predicted furin cleavage site located within the small exons of HLS1-E and CON3 indicated by the white dots in the figure. Figure adapted from Heft et al., 2020.
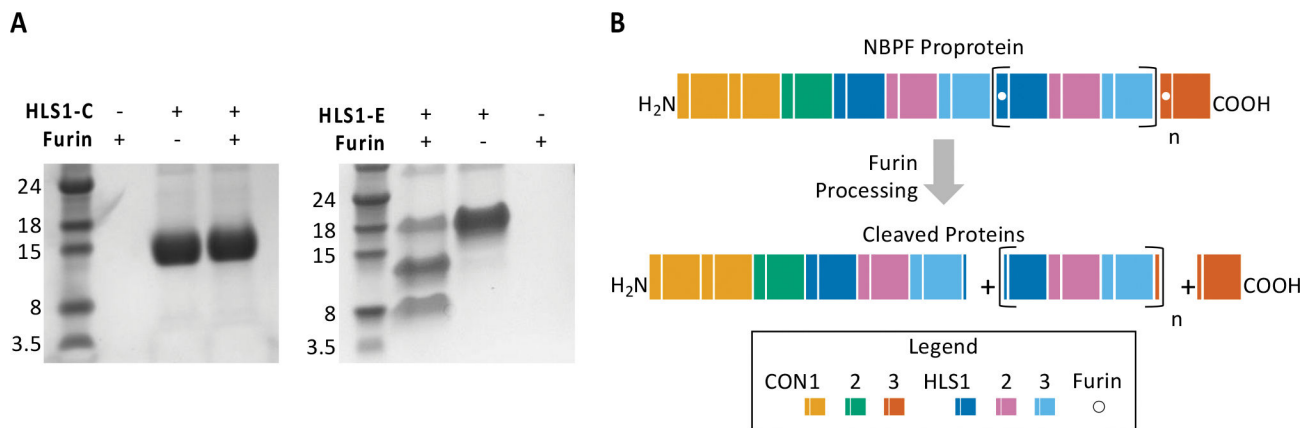
Fig. 2.
(A) Furin is shown to cleave expanded HLS1 (HLS1-E). To analyze furin cleavage of HLS1, both HLS1-C and HLS1-E were independently co-expressed with furin. A single banding pattern is shown when HLS1-C is co-expressed with furin (as seen on the left), whereas a multiple banding pattern is seen when HLS1-E is co-expressed with furin (as seen on the right). This finding confirms the previously predicted furin cleavage site that is located within HLS1-E. **(B)** Furin processing of the NBPF proprotein would result in three categories of peptides: the NBPF N-terminal peptide, "n" number of independent Olduvai triplet peptides, and the C-terminal CON3 peptide. A hypothetical NBPF protein with "n" number of Olduvai triplets is shown before and after furin processing. Furin cleavage sites are indicated by the white dots in HLS1-E and CON3.