Alphavirus nsP2 protease structure and cleavage prediction: Possible relevance to the pathogenesis of viral arthritis

Lucas Prescott[1]*

[1] Independent Researcher, Palo Alto, California, United States of America

* Corresponding author
E-mail: lskywalker2015@gmail.com

Keywords: Togaviridae, alphavirus, protease, proteomics, machine learning, neural networks

Abstract

Alphaviruses are a diverse genus of arboviruses capable of infecting many vertebrates including humans. Human infection is common in equatorial and subtropical regions and is often accompanied by arthralgia or encephalitis depending on viral lineage. No antivirals or vaccines have been approved, and many alphavirus lineages have only recently been discovered and classified. Alphavirus nsP2 protease is an important virulence factor yet is commonly thought to be a simple papain-like protease which only cleaves viral polyproteins. Here, I reveal novel molecular mechanisms of these proteases via sequence and predicted structure alignment and propose novel cellular mechanisms for the pathogenesis of viral arthritis by predicting which human proteins are likely cleaved by these proteases. In addition to the known primary cysteine mechanism in all alphaviruses and a secondary serine mechanism documented in chikungunya virus (CHIKV), I discovered secondary cysteine and threonine mechanisms exist in many other alphaviruses and that these secondary mechanisms coevolve with their viral polyprotein cleavages. As for cleavage prediction, neural networks trained on 93 different putative viral polyprotein cleavages achieved a Matthews correlation coefficient of 0.965, and, when applied to the human proteome, predicted that hundreds of proteins may be vulnerable. Notable pathways likely affected by cleavages include the cytoskeleton and extracellular matrix, antiproteases, protein translation/folding/glycosylation/ubiquitination, cellular differentiation, inflammation, and vesicle trafficking, hinting that this viral protease is a more important virulence factor than previously believed.

Introduction

Alphavirus genomes contain two open reading frames encoding non-structural and structural polyproteins. Although the structural polyprotein is proteolytically processed by the capsid protein and host furin and signal peptidases, the non-structural polyprotein is processed typically by a cysteine protease contained within nsP2 (nsP2pro). A cleft for substrate binding exists between nsP2's C-terminal protease and S-adenosyl-L-methionine-dependent methyltransferase (MTase)-like subdomains connected by a flexible linker, and long-range interactions with nsP2's N-terminal helicase[1, 2] or any nsP3 domains before its separation are important for polyprotein processing and virulence yet remain poorly characterized.[3, 4] CHIKV nsP2pro has been found to not only contain a papain-like cysteine mechanism, but also an adjacent serine with similar activity.[5] This mechanism has not yet been found in any other alphaviruses, but it likely dramatically affects the stability, activity, and selectivity of nsP2pro. Additionally, a single mutation (N475A) near the N-terminus of the protease subdomain was found to cause the flexible N-terminal residues to occupy the cleft and inhibit catalysis.[6] This mutation does not, however, exist in any natural variants, and this study was not performed on full-length nsP2. The nuclear localization signal and RNA-binding helicase determining nuclear and virion localization of nsP2 likely also drive nsP2pro's selective pressures and multiple activities.[7] In addition to these subdomain interactions, most alphaviruses contain a stop codon at the end of nsP3 which is read through in 5-20% of polyproteins.[7, 9] Depending on this termination suppression, nsP2pro cleaves either two or three sites within the non-structural polyprotein with kinetic rates varying up to 25 fold[10, 11] and, as with many viral proteases, is expected to cleave many host factors. To my knowledge, the antiviral TRIM14 is the only host protein experimentally verified to be cleaved by an alphavirus nsP2pro (Venezuelan equine encephalitis virus (VEEV) and somewhat by other New World alphaviruses).[12]

Due to the few cleavages per polyprotein and the continually expanding taxonomy of alphaviruses, few viral proteases or their cleavages have been characterized.[13] Following the successful application of machine learning methods to other viral proteases for cleavage prediction in human proteins,[14-16] computational analysis of alphavirus proteases will likely be an important step toward discovering additional therapeutic targets.

Methods

Data Set Preparation

A complete, manually reviewed human proteome containing 20,350 sequences (not including alternative isoforms) was retrieved from UniProt/Swiss-Prot (proteome:up000005640 AND reviewed:yes).[17] All polyprotein sequences within the family *Togaviridae* were collected from GenBank,[18] and 93 different cleavages were manually discovered using the Clustal Omega multiple sequence alignment server.[19-21] Similar cleavages are discoverable in divergent species within the order *Martellivirales* but were not included here because none infect animals. The next closest species that can infect humans are rubella (RUBV) and hepatitis E (HEV) viruses within the broader class *Alsuviricetes*, but their non-structural proteases have drastically different structures and activities than those within *Togaviridae* and were therefore also omitted. All unbalanced positive cleavages were used for subsequent classifier training in addition to all other 5,461 uncleaved alphavirus sequence windows with glycines in the P2 position, totaling 5,554 samples.

Protease Structure Prediction and Analysis

AlphaFold was used to predict the structures of alphavirus nsP2 sequences (only the C-terminal protease and MTase-like subdomains).[22] Predicted backbones were nearly identical to experimental data, and predicted catalytic dyad side chain distances ranged from 4 to 8 Å. Although AlphaFold does not have the ability to accurately predict the impact of single missense mutations on protein structures,[23, 24] this set of predicted structures and simulations, albeit on often nearly identical sequences, serves as a starting point to understanding the diversity of mechanisms within alphaviruses. CABS-flex was used to predict alternate conformations and flexibilities,[25] and zinc binding prediction was performed with ZincBind.[26] Molecular graphics and analyses were performed using UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases.[27]

Cleavage Prediction and Analysis

As in my previous work on 3CLpro and PLpro,[14, 15] sequence logo-based logistic regression and naïve Bayes classification and physiochemical and one-hot encoded neural networks (NNs) were used for cleavage prediction.[28] To reduce any potential false positives, only proteins expressed in relevant cell types with cleavages with agreement between all five NN replicates and with total solvent-accessible surface areas (SASAs) of more than 100 $Å^2$ between positions P5 and P5' were reported. SASAs were calculated from AlphaFold predicted human protein structures[22] with the FreeSASA package.[29] Multiple synovial fluid and associated cell type proteomes and transcriptomes were compiled and cross-referenced to remove cleavages irrelevant to arthritic pathogenesis.[30-35] Predictions of the effects of cleavage on subcellular localizations were performed using the DeepLoc server.[36] All training data, prediction methods, and results can be found on GitHub (https://github.com/Luke8472NN/NetProtease).

Results

Although alphavirus proteases are diverse and not necessarily only papain-like, their cleavages resemble those of coronavirus papain-like protease (PLpro) but not of papain itself (Figure 1).[37] The repeated glycines and alanines in alphavirus cleavage positions P2, P1, and P1' were, however, easier to align than coronavirus PLpro cleavages. Dimensionality reduction of putative cleavages clustered by order within the polyprotein much more than by lineage (except for previously discovered cleavages in tymoviruses[38]), indicating that all cleavages within *Togaviridae* but not *Alsuviricetes* can be combined into a single data set to train machine learning models to apply to human sequences (Figure 2). Although no cleavages from species outside *Togaviridae* were included for training here, it is noteworthy that

turnip yellow mosaic virus (TYMV) protease is but HEV and RUBV proteases are not structurally related to alphavirus nsP2pro. TYMV protease includes an equivalent catalytic cysteine helix and activity-tuning histidine flexible loop[39] yet does not contain an MTase-like domain to form a cleft as in alphaviruses. In addition to this more accessible active site, TYMV protease includes two hydrophobic patches required for interaction with ubiquitin for its deubiquitinating activity (Figure 3).[40]

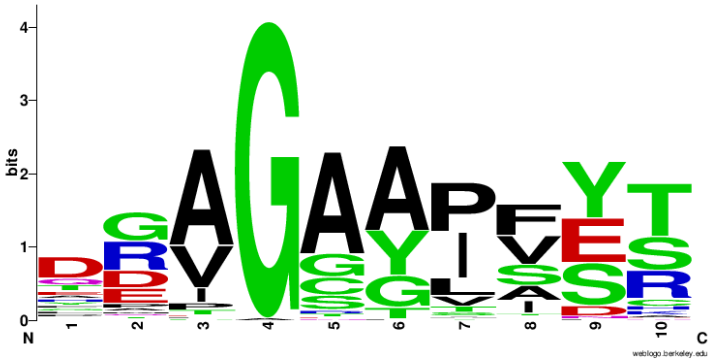

Figure 1: Sequence logo for all 93 putative cleavage sites.[41]

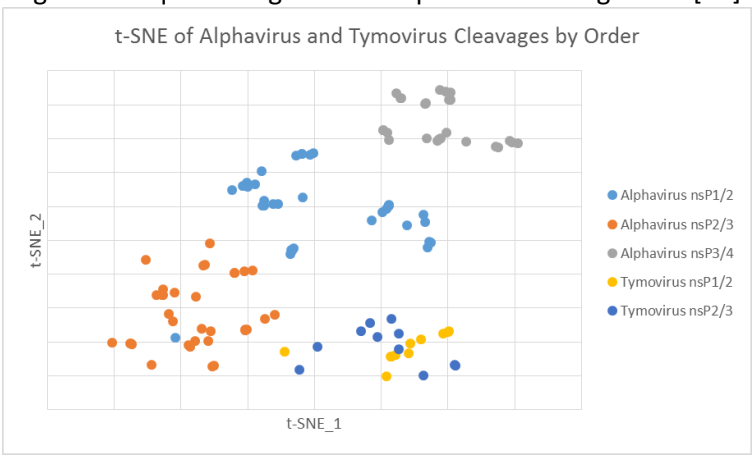

Figure 2: One-hot encoded cleavage t-SNE colored by order within the polyprotein and by genus.[42]

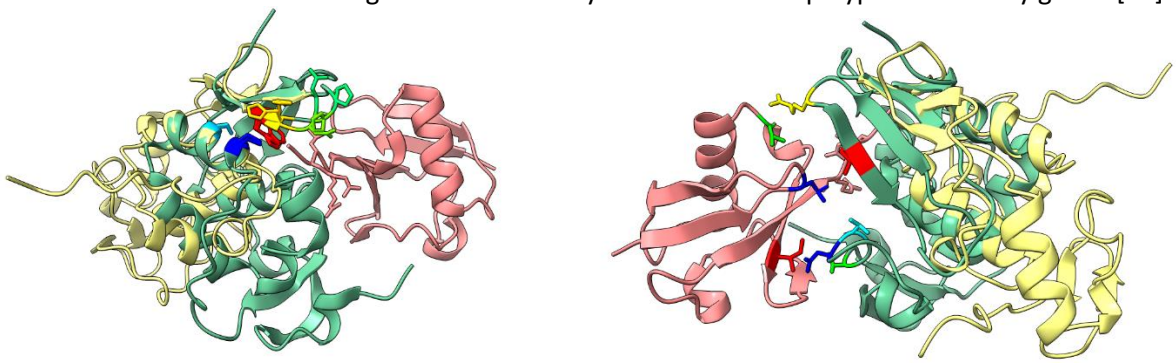

Figure 3: (A) Structural similarity between CHIKV and TYMV proteases near their active sites but (B) not in ubiquitin binding regions. Only TYMV protease binds ubiquitin's I36 and I44 hydrophobic patches and its L8 loop. Tan ribbon is CHIKV, green ribbon is TYMV, red ribbon is ubiquitin (PDB code 6YPT).[27, 40]

Alignment of all known alphavirus proteases (Figure 4) indicated that, in addition to a primary cysteine mechanism in all alphaviruses and a secondary serine mechanism found in at least CHIKV,[5] some proteases have secondary cysteine or threonine mechanisms. These secondary mechanisms may

restrict or extend the possible acidic residues aligning and polarizing the catalytic histidine,[43] and serine and threonine mechanisms may extend catalytic activity to higher pH.[44] By investigating how substrate sequences, particularly the P1 residue, coevolve with these different mechanisms, multiple functional hypotheses can be proposed: (1) cleavage after a P1 cysteine is most efficient when the secondary catalytic residue is another cysteine or serine, possible for an inert secondary alanine, and least efficient for a secondary threonine, (2) a secondary threonine is required for cleavage after a P1 serine, and (3) an inert secondary alanine is required for cleavage after a P1 arginine.

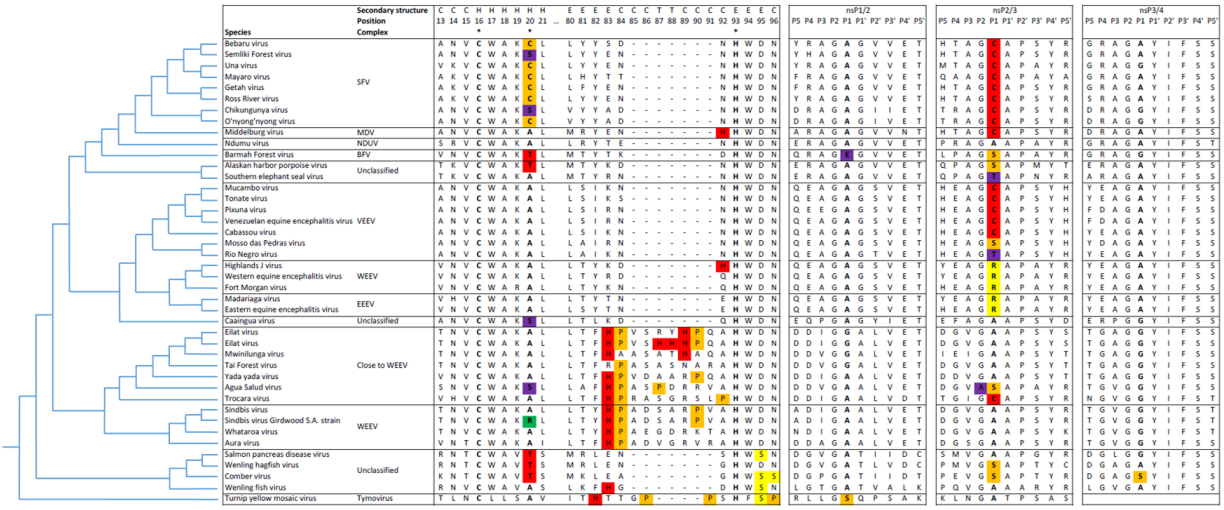

Figure 4: (A) Cladogram and multiple sequence alignment of the catalytic dyad and flexible loop of alphavirus and TYMV proteases and (B) their respective aligned cleavages.

Near the catalytic histidine, other histidines or related positively charged amino acids in the longer flexible loops of Western equine encephalitis complex and related viruses are close enough in proximity with each other that they may bind ordered water molecules as in other proteases[45] or metal ions in multiple conformations.[26] Unlike RUBV cysteine protease[46] and hepatitis C virus (HCV) NS3 serine protease[47] which require metal ions in either a structural or catalytic (as in metalloprotease) role for activity, metal ions are known to inhibit CHIKV protease.[48] In these alphavirus proteases, metal binding may disrupt normal histidine alignment with aspartic acid[49] and aim its protonated side toward the catalytic cysteine, serine, or threonine, preventing the charge relay mechanism required for proteolysis (Figure 5A). Additionally, metal binding to another histidine in the center of the loop may aid its flexing backward to allow substrate loading. In some divergent, unclassified alphavirus proteases, the adjacent aspartic acid is replaced with serine, but in these cases there is always another nearby potential metal-binding residue (glutamic acid, aspartic acid, or another histidine)(Figure 5B).

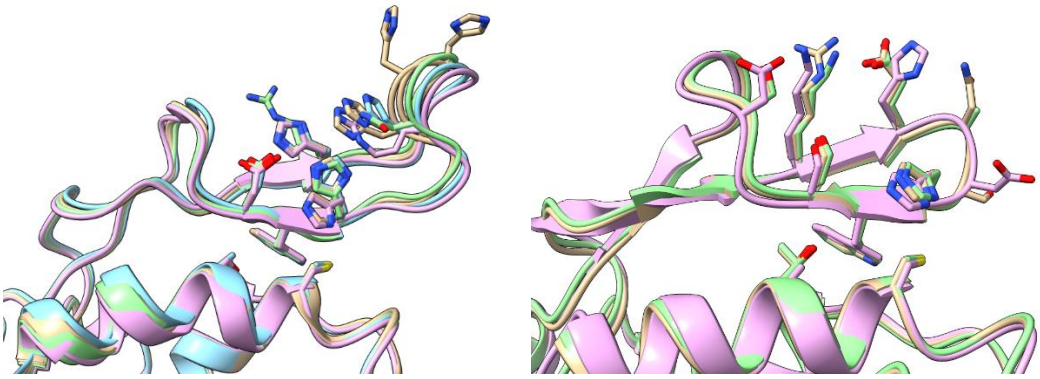

Figure 5: (A) Proposed metal binding histidines and adjacent aspartic acid guiding charge relaying in proteolysis. Tan ribbons are Eilat viruses, one of which includes five histidine residues,[50] purple ribbon is Agua Salud virus (ASALV), green ribbon is Tai Forest virus, and blue ribbon is Mwinilunga virus. (B) Proposed alternative metal binding residues in divergent proteases. Tan ribbon is Salmonid virus, purple ribbon is Wenling fish virus, and green ribbon is Comber virus).[27]

To my knowledge, no P2 glycine substitutions have been discovered in alphavirus cleavages, so it is noteworthy that ASALV[51] nsP2/3 cleavage (DGVAS^APAYR in MK959114.1 and MK959115.1) contains an alanine in this position (Figure 4). No sequence features obviously correlated with this substitution, and ASALV protease's predicted structure is extremely similar to those of related alphavirus proteases, indicating that other alphavirus proteases may also cleave alanine-containing substrates albeit possibly with suboptimal kinetics. Tryptophan is typically thought not to be directly involved in the active site, yet it appeared here to obstruct the secondary catalytic mechanism in some conformations (Figure 6A). In addition to the flexibility of the catalytic dyad, the size and flexibility of the variable loop between the protease domain β1 and β2 strands and its interaction with the MTase-like domain loop between β7 strand and α9 helix (Figures 6B and 6C) likely determine the rate of substrate loading into the cleft and therefore cleavage kinetics.[52] Deletion of the exposed and most proximal MTase-like domain residue (typically leucine, phenylalanine, or tryptophan) and sharper backbone twisting by subsequent prolines in divergent alphaviruses may also widen the gap between these two loops and affect substrate loading or may allow serine to better fit in the P1 pocket instead of the more common alanine (Figure 6D). Sodium[53, 54] or other salt binding within this cleft may affect structure and substrate binding, and these deep cleft residues are not conserved between alphaviruses. No matter the width of this cleft, the nsP2/3 site remained over 40 Å away from the active site, supporting this site's proposed *trans* cleavage.[3] Even Salmonid virus' large insertions at the C-terminus of its nsP2 MTase-like domain and between its nsP3 macro and zinc-binding domains did not affect this distance.

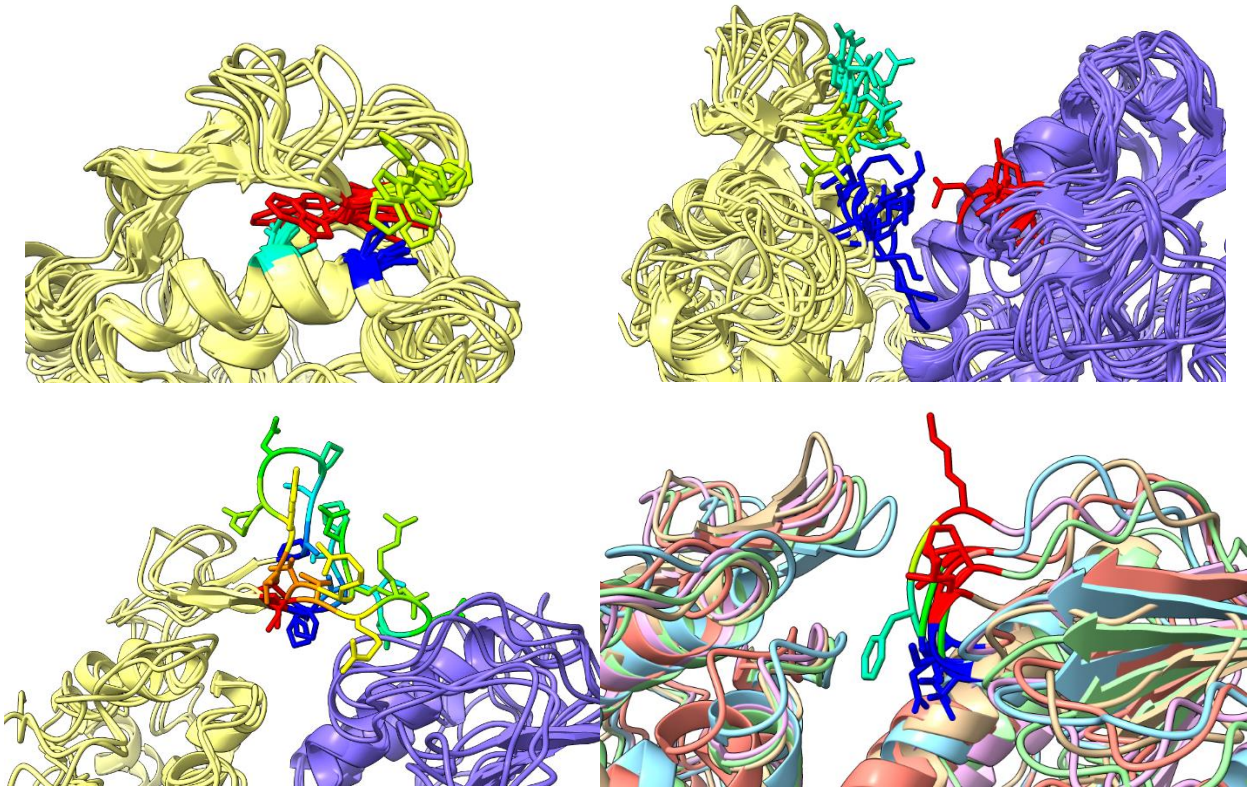

Figure 6: (A) CABS-flex ensemble of CHIKV nsP2pro flexible catalytic core and (B) flexible loop between protease subdomain (tan ribbon) β1 and β2 strands interacting with loop between MTase-like domain (purple ribbon) β7 strand and α9 helix. (C) ASALV insertion within flexible loop.[25] (D) Similarity between interacting MTase-like domain loops with noteworthy deletions. Tan ribbon is Alaskan harbor porpoise virus, blue ribbon is Salmonid virus, purple ribbon is Wenling fish virus, green ribbon is Wenling hagfish virus, and red ribbon is Comber virus.[27]

As with both coronavirus protease cleavage predictions, NNs outperformed all other classifiers (Figure 7). The optimized hyperparameters for NNs with one-hot encoding were Adam solver, rectifier (ReLU) activation, 1e-8 regularization, no oversampling, and 1 hidden layer with 10 neurons. Combining networks into ensembles again improved accuracy and stability, so the final results were generated with 5 replicates of 10-fold cross-validated (CV) networks with an average Matthews correlation coefficient (MCC) of 0.965. Very few false positives existed for any prediction method, but two putative sites were somewhat conserved within the Semliki Forest (SF) complex nsP1 MTase-GTase core.[55] These two sites are predicted to be ordered and not solvent exposed and so are likely not biologically important. When applied to the human proteome, 714 of 20,350 proteins were predicted to be cleaved at least once. Enrichment analysis did not return useful results as for coronavirus protease predictions, so this large list was instead reduced as described in Methods to discuss only the most likely meaningful cleavages.

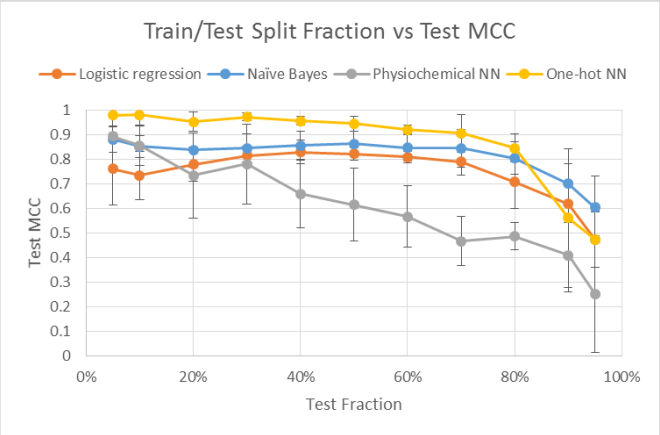

Figure 7: Train/test split fraction versus MCC demonstrating that the entire data set is not required for satisfactory accuracy.

Discussion

Similar biases and caveats exist for nsP2pro predictions as with previous viral protease predictions,[14, 15] yet many host pathways likely perturbed by nsP2pro are discoverable. Experimental validation is, however, required for all of the following hypotheses. Reviewing only predicted cleavages with the highest scores and sufficient SASAs and relevant tissue expression produced a more targeted list of cleavages for interpretation. This list contains many proteins involved in the cytoskeleton and extracellular matrix (ECM), protease inhibition, protein translation/folding/glycosylation/ubiquitination, cellular differentiation including the transforming growth factor beta (TGF-β) and tumor necrosis factor alpha (TNF-α) pathways, inflammation, and vesicle trafficking.

Cleavage of many cytoskeletal proteins likely contributes to altered virus and host component trafficking, yet cytoskeletal drugs have remained mostly ineffective against alphaviruses. These drugs may help prevent initial endocytosis,[56, 57] internalization of replication complexes in spherules into cytopathic vacuoles,[58] and virion release, yet alphaviruses are known to still replicate in their presence.[59, 60]

In the extracellular matrix, the most obviously affected region in arthritis, nsP2pro likely cleaves many structural proteins. In particular, cleavage of lubricin's hemopexin-like domain, similar to its normal cleavage by a human subtilisin-like proprotein convertase (SPC), may disrupt its ability to bind other proteins at the cartilage surface, reducing lubrication and promoting inflammation.[61] Cleavage of collagen alpha-1(XII) may disrupt normal shock-absorbing function.[62] In addition to the general matrix disruption likely resulting from cleavage of perlecan, elastin, von Willebrand factor A domain-containing protein 1 (WARP),[63] laminin, nidogen-1,[64] and thrombospondin-3 and -4,[65] degradation products of elastin and perlecan (laminin-like globular domain (LG3) of endorepellin) are documented to promote joint inflammation[66] and prevent angiogenesis in avascular cartilage,[67] respectively. Cleavage of specifically laminin subunit gamma-1 would not disrupt laminin binding membrane-bound integrins and dystrophins or extracellular collagen, but it may disrupt gamma subunit binding to nidogen and polymerization,[68, 69] freeing it up to be a more accessible alphavirus receptor.[70, 71] Similarly extracellularly secreted although not structural, nsP2pro has predicted cleavages near and within the bait regions of the serum and synovial fluid antiproteases alpha-2-macroglubulin (A2M) and pregnancy zone protein (PZP).[72, 73]

Noteworthy cleaved proteins involved in translation include signal recognition particle receptor subunit alpha (SRPRA), eukaryotic peptide chain release factor GTP-binding subunits (ERF3A/B), and La-related protein 1 (LARP1). Cleavage of SRPRA between its N-terminal SRX domain and its C-terminal targeting complex may reduce translocation of many proteins into the endoplasmic reticulum (ER). Only

the viral structural polyprotein contains a signal peptide, so this cleavage may also relate the kinetics of its N-terminal capsid protein autocleavage to its subsequent SRP association.[74, 75] Cleavage of release factors ERF3A/B expressed in osteoblasts and osteoclasts may promote stop codon readthrough critical to alphavirus infection.[76, 77] Cleavage of LARP1 may have effects similar to its documented inhibition by alphavirus capsid protein or by mTOR often activated early in infection.[78] This mechanism of inhibiting host translation differs between mosquito and vertebrate cells, is not required for viral production, but may be required for internalization of the replication complex.[79] Unlike viruses requiring eukaryotic initiation factor 2 (eIF2) for translation, alphavirus downstream hairpin loop (DLP)-mediated translation does not benefit from amino acid starvation, so the effects of alphaviruses on mTOR are more straightforward than those of picornaviruses, flaviviruses, etc. For example, cleavage of the E3 ubiquitin ligase tetratricopeptide repeat protein 3 (TTC3) may help activate AKT and downstream protein synthesis similar to its direct interaction with nsP3,[79, 80] and cleavage of TRIM63 may stabilize many proteins against amino acid starvation-associated degradation.

Also after translation, cleavage of the immunophilins peptidyl-prolyl cis-trans isomerases (PPI) H and FKBP10 may affect folding of many relevant proteins. These may (1) have immunosuppressive effects similar to inhibition by tacrolimus, (2) modulate calcineurin, ribonuclease A, and some interleukins with cis-prolines in their native states, (3) affect proline- and hydroxyproline-rich collagen structure (supported by nsP2pro's cleavage of prolyl 4-hydroxylase subunit alpha-1 and by documented excretion of proline and hydroxyproline in the urine of CHIKV-infected patients),[81] and (4) modulate alphavirus nsP3 proline-rich domain binding to amphiphysins involved in membrane bending of alphavirus-induced membrane organelles.[82, 83]

Enzymes involved in glycosylation also have predicted cleavages, although the differential effects this would have on viral versus host protein glycosylation remain unknown. Cleavage of mannosyl-oligosaccharide glucosidase (MOGS) may broadly disrupt viral protein glycosylation as in congenital disorders of this enzyme,[84] and cleavage of beta-1,4-galactosyltransferase 3 (B4GALT3) may disrupt complex N-linked glycans on immunoglobulins as in RA.[85] Cleavage of ER degradation-enhancing alpha-mannosidase-like protein 2 (EDEM2)[86] may disrupt host ERAD of viral glycoproteins[87] or redirect viral glycoproteins away from the cell membrane for internal budding as in SINV-infected mosquito cells.[88, 89] Cleavage of phosphoacetylglucosamine mutase (PAGM) may contribute to disruption of glycosaminoglycan polymers in cartilage and contribute to arthritic symptoms,[90] however N-acetylglucosamine (GlcNAc) supplementation sometimes used to treat osteoarthritis (OA) may be counterproductive given (1) it makes up some alphavirus glycans,[91] (2) it promotes replication of many other viruses *in vitro* and *in vivo*,[92] and (3) O-linked GlcNAc glycosylation of p65 aggravates TNF-α-mediated inflammation in rheumatoid arthritis (RA).[93]

Other predicted cleavages involved in protein degradation include the SUMO-specific E1 enzyme SAE1, the E2 enzyme UBE2Q1, the E3 enzyme UBR4, ubiquitin-1, and tripeptidyl peptidase 2 (TPP2). As with other viruses, modulation of SUMOylation is nontrivial and likely time-dependent; depletion of the only E2 for SUMO, UBC9, protects against CHIKV infection in mice,[94] yet depletion of SUMOylation enhances SFV replication in mosquito cells.[95] Cleavage of UBE2Q1 in muscle may disrupt B4GALT1-mediated cell adhesion to laminin and promote myoblast and satellite cell differentiation and syncytia formation[96, 97] to allow infection of myofibers without virion egress.[98-100] This is supported by the ability of alphaviruses to form filopodia-like protrusions mediating cell-to-cell transmission[101] and possibly to shield the virus from antibodies, making effective vaccination more difficult.[102] Cleavage of UBR4 may disrupt the N-end rule[103] to stabilize an inhibitor of apoptosis as with a picorna-like virus[104] or to stabilize cleaved viral functional proteins with less stable N-termini.[105] Cleavage of ubiquilin-1 between its ubiquitin-associated (UBA) and ubiquitin-like (UBL) domains may disrupt its trafficking ubiquitinated proteins to the proteasome[106] or its targeting of transmembrane proteins.[107, 108] Cleavage of the proteolytic TPP2 downstream of the 26S proteasome may promote

viral susceptibility as in TRIANGLE disease[109] and may affect the pool of short peptides available for MHC class I presentation.[110]

TGF-β is known to be elevated in RA and during alphavirus infection,[111] and its inhibition can reduce joint swelling yet does not reduce viral titer[112] and can even promote CHIKV-mediated cell death *in vitro*.[113] Cleavage of latent TGF-β binding protein 3 (LTBP3) near its ECM-binding C-terminus may be one mechanism alphaviruses employ to encourage release of TGF-β from latency-associated peptide (LAP) when combined with normal cleavage by host proteases in its N-terminal hinge and C-terminus regions.[114, 115] These pathways are counterintuitive due to the many differential effects of TGF-β and bone morphogenetic proteins (BMPs) on different cell types and their interactions throughout their differentiation.[116] In the mesenchymal (MSC) lineage, TGF-β stimulates proliferation and differentiation of MSCs into chondrocytes and osteoblast progenitors into osteoblasts with downregulated RANKL. The closely related BMPs, however, can oppose TGF-β and are required for late stage osteoblast differentiation through their different SMAD signal transducers.[117] In the hematopoietic (HSC) lineage, TGF-β keeps HSCs in hibernation and prevents osteoclast progenitor differentiation into mature osteoclasts at least partially due to disrupted RANKL/OPG ratio.[118] Alphavirus infection is, however, more complex than elevated TGF-β alone and is associated with increased RANKL/OPG ratio and therefore osteoclastogenesis[119, 120] likely via upregulated IL-6 positive feedback[121, 122] and disrupted osteoblastogenesis via reduced RUNX2.[123] This TGF-β disruption may direct stem cells toward differentiated lineages more susceptible to infection, yet disruption of other pathways may be able to prevent apoptosis in these differentiated cells. For example, cleavage of TNFR2 may prevent TNF-α transduction and even shed soluble TNFR2 which can, like its alternately spliced products, antagonize its full-length activity and downstream apoptosis of infected cells.[124]

Likely also in an attempt to prevent inflammation and death of infected cells, cleavage of PYCARD between its pyrin domain (PYD) and caspase recruitment domain (CARD) may disrupt normal inflammasome formation and act like host CARD only proteins (COPs) or pyrin only proteins (POPs), inhibiting caspase 1 and its target cytokines.[125] As for other inflammatory molecular classes likely affected, polyamines are generally required by RNA viruses,[126, 127] and exogenous polyamines can restore inflammation and immune dysregulation in RA and OA.[128-130] Cleavage of diamine acetyltransferase 1 (SAT1) would increase intracellular polyamine concentration by preventing export and may (1) promote DNA methylation and reduce host transcription via S-adenosyl methionine (SAM) metabolism,[131] (2) downregulate IL-2 in PBMCs and contribute to decreased T cell effector function as in RA,[132] (3) enhance translation of polyproline motifs such as collagen and alphavirus nsP3 proline-rich domain,[133] (4) speed up peptidyl-tRNA hydrolysis by termination factor eRF1 via its hypusine modification,[134] and (5) promote stop codon readthrough by altering tRNA conformation.[135] Cleavage of histidine decarboxylase (HDC), glutathione hydrolase 5 (GGT5), and phospholipase A and acyltransferase 3 (PLAAT3) would, however, oppose typical inflammation in RA (where histamine[136] and lysophospholipids[137] are elevated and glutathione[138] is depleted), serving as a reminder that viral arthritis is caused more by the immune response to infection than by the virus or nsP2pro.

Lastly, the effects of predicted cleavages in vesicle transporting proteins are difficult to interpret because both retrograde and anterograde pathways are affected. Major differences exist between alphavirus-induced mammalian and mosquito membrane rearrangements, so experimental characterization is required to understand the relevance of these cleavages in each host.[139, 140]

Conclusion

These predicted cleavages hint at many expected and novel mechanisms and indicate that nsP2pro is a much more important virulence factor than previously believed. Substrate docking and molecular dynamics may provide additional information about molecular mechanisms of these

proteases, and protease-specific kinetics and biological significance of these cleavages require experimental verification. Expansion of this data set to include all of *Martellivirales* or *Alsuviricetes* may also provide insight into how these molecular mechanisms evolved, but their inclusion into a cleavage prediction training data set would likely worsen the trained model's accuracy for the human viruses discussed here. Even though many caveats exist without experimentation, similar prediction and interpretation should be performed for all other viral proteases.

References

1. Law YS, Utt A, Tan YB, Zheng J, Wang S, Chen MW, et al. Structural insights into RNA recognition by the Chikungunya virus nsP2 helicase. PNAS. 2019 May;116(19):9558-67. doi: 10.1073/pnas.1900656116 PMID: 31000599
2. Law YS, Wang S, Tan YB, Shih O, Utt A, Goh WY, et al. Interdomain flexibility of chikungunya virus nsP2 helicase-protease differentially influences viral RNA replication and infectivity. J Virol. 2021 Feb;95(6):e01470-20. doi: 10.1128/JVI.01470-20 PMID: 33328310
3. Shin G, Yost SA, Miller MT, Elrod EJ, Grakoui A, Marcotrigiano J. Structural and functional insights into alphavirus polyprotein processing and pathogenesis. PNAS. 2012 Oct;109(41):16534-9. doi: 10.1073/pnas.1210418109 PMID: 23010928
4. Saul S, Ferguson M, Cordonin C, Fragkoudis R, Ool M, Tamberg N, et al. Differenced in processing determinants of nonstructural polyprotein and in the sequence of nonstructural protein 3 affect neurovirulence of Semliki Forest virus. J Virol. 2015 Nov;89(21):11030-45. doi: 10.1128/JVI.01186-15 PMID: 26311875
5. Saisawang C, Saitornuang S, Sillapee P, Ubol S, Smith DR, Ketterman AJ. Chikungunya nsP2 protease is not a papain-like cysteine protease and the catalytic dyad cysteine is interchangeable with a proximal serine. Sci Rep. 2015 Nov;5:17125. doi: 10.1038/srep17125 PMID: 26597768
6. Compton JR, Mickey MJ, Hu X, Marugan JJ, Legler PM. Mutation of Asn-475 in the Venezuelan equine encephalitis virus nsP2 cysteine protease leads to a self-inhibited state. Biochemistry. 2017 Nov;56(47):6221-30. doi: 10.1021/acs.biochem.7b00746 PMID: 29064679
7. Schuchman R, Kilianski A, Piper A, Vancini R, Ribeiro JMC, Sprague TR, et al. Comparative characterization of the Sindbis virus proteome from mammalian and invertebrate hosts identifies nsP2 as a component of the virion and sorting nexin 5 as a significant host factor for alphavirus replication. J Virol. 2018 Jul;92(14):e00694-18. doi: 10.1128/JVI.00694-18 PMID: 29743363
8. de Groot RJ, Hardy WR, Shirako Y, Strauss JH. Cleavage-site preferences of Sindbis virus polyproteins containing the non-structural proteinase. Evidence for temporal regulation of polyprotein processing *in vivo*. EMBO J. 1990 Aug;9(8):2631-8. PMID: 2142454
9. Shirako Y, Strauss JH. Regulation of Sindbis virus RNA replication: Uncleaved P123 and nsP4 function in minus-strand RNA synthesis, whereas cleaved products from P123 are required for efficient plus-strand RNA synthesis. J Virol. 1994 Mar;68(3):1874-85. doi: 10.1128/JVI.68.3.1874-1885.1994 PMID: 8107248
10. Hu X, Compton JR, Leary DH, Olson MA, Lee MS, Cheung J, et al. Kinetic, mutational, and structural studies of the Venezuelan equine encephalitis virus nonstructural protein 2 cysteine protease. Biochemistry. 2016 May;55(21):3007-19. doi: 10.1021/acs.biochem.5b00992 PMID: 27030368
11. Bozoki B, Motyan JA, Hoffka G, Waugh DS, Tozser J. Specificity studies of the Venezuelan equine encephalitis virus non-structural protein 2 protease using recombinant fluorescent substrates. Int J Mol Sci. 2020 Oct;21(20:7686. doi: 10.3390/ijms21207686 PMID: 33081394
12. Morazzani EM, Compton JR, Leary DH, Berry AV, Hu X, Marugan JJ, et al. Proteolytic cleavage of host proteins by the Group IV viral proteases of Venezuelan equine encephalitis virus and Zika virus. Antiviral Res. 2019 Apr;164:106-22. doi: 10.1016/j.antiviral.2019.02.001 PMID: 30742841

13. Abu Bakar F, Ng LFP. Nonstructural proteins of alphavirus—Potential targets for drug development. Viruses. 2018 Feb;10(2):71. doi: 10.3390/v10020071 PMID: 29425115
14. Prescott L. SARS-CoV-2 3CLpro whole human proteome cleavage prediction and enrichment/depletion analysis. bioRxiv [Preprint]. 2020 Aug. doi: 10.1101/2020.08.24.265645
15. Prescott L. SARS-CoV-2 PLpro whole human proteome cleavage prediction and enrichment/depletion analysis. bioRxiv [Preprint]. 2021 Oct. doi: 10.1101/2021.10.04.462902
16. Stanley JT, Gilchrist AR, Stabell AC, Allen MA, Sawyer SL, Dowell RD. Two-stage ML classifier for identifying host protein targets of the dengue protease. Pac Symp Biocomput. 2020 Jan;25:487-98. PMID: 31797621
17. The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 2019 Jan;47(D1):D505-15. doi: 10.1093/nar/gky1049 PMID: 30395287
18. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J. GenBank. Nucleic Acids Res. 2017 Jan;45(D1):D37-42. doi: 10.1093/nar/gkw1070 PMID: 27899564
19. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011 Oct;7:539. doi: 10.1038/msb.2011.75 PMID: 21988835
20. Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, et al. A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res. 2010 Jul;38(2):W695-9. doi: 10.1093/nar/gkq313 PMID: 20439314
21. McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, et al. Analysis tool web services from the EMBL-EBI. Nucleic Acids Res. 2013 Jul;41(W1):W597-600. doi: 10.1093/nar/gkt376 PMID: 23671338
22. Tunyasuvunakool K, Adler J, Wu Z, Green T, Zielinski M, Zídek A, et al. (2021) Highly accurate protein structure prediction for the human proteome. Nature. 2021 Jul;596:590-6. doi: 10.1038/s41586-021-03828-1 PMID: 34293799
23. Buel GR, Walters KJ. Can AlphaFold2 predict the impact of missense mutations on structure? Nat Struct Mol Biol. 2022 Jan;29(1):1-2. doi: 10.1038/s41594-021-00714-2 PMID: 35046575
24. Pak MA, Markhieva KA, Novikova MS, Petrov DS, Vorobyev IS, Maksimova ES, et al. Using AlphaFold to predict the impact of single mutations on protein stability and function. bioRxiv [Preprint]. 2021 Sep. doi: 10.1101/2021.09.19.460937
25. Kuriata A, Gierut AM, Oleniecki T, Ciemny MP, Kolinski A, Kurcinski M, et al. CABS-flex 2.0: a web server for fast simulations of flexibility of protein structures. Nucleic Acids Res. 2018 Jul;46(W1):W338-43. doi: 10.1093/nar/gky356 PMID: 29762700
26. Ireland SM, Martin ACR. ZincBind—the database of zinc binding sites. Database (Oxford). 2019 Jan;2019:baz006. doi: 10.1093/database/baz006 PMID: 30722040
27. Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci. 2021 Jan;30(1):70-82. doi: 10.1002/pro.3943 PMID: 32881101
28. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. J Mach Learn Res. 2011 Oct;12:2825-30. http://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html
29. Mitternacht S. FreeSASA: An open source C library for solvent accessible surface area calculations. F1000Research. 2016 Feb;5:189. doi: 10.12688/f1000research.7931.1 PMID: 26973785
30. Birkelund S, Bennike TB, Kastaniegaard K, Lausen M, Poulsen TBG, Kragstrup TW, et al. Proteomic analysis of synovial fluid from rheumatic arthritis and spondyloarthritis patients. Clin Proteomics. 2020 Aug;17:29. doi: 10.1186/s12014-020-09292-9 PMID: 32782445
31. Lambrecht S, Dhaenens M, Almqvist F, Verdonk P, Verbruggen G, Deforce D, et al. Proteome characterization of human articular chondrocytes leads to novel insights in the function of small

heat-shock proteins in chondrocyte homeostasis. Osteoarthritis Cartilage. 2010 Mar;18(3):440-6. doi: 10.1016/j.joca.2009.10.008 PMID: 19879248

32. Alves RDAM, Eijken M, Swagemakers S, Chiba H, Titulaer MK, Buruegrs PC, et al. Proteomic analysis of human osteoblastic cells: relevant proteins and functional categories for differentiation. J Proteome Res. 2010 Sep;9(9):4688-700. doi: 10.1021/pr100400d PMID: 20690663
33. An E, Narayanan M, Manes NP, Nita-Lazar A. Characterization of functional reprogramming during osteoclast development using quantitative proteomics and mRNA profiling. Mol Cell Proteomics. 2014 Oct;13(10):2687-704. doi: 10.1074/mcp.M113.034371 PMID: 25044017
34. Chou CH, Jain V, Gibson J, Attarian DE, Haraden CA, Yohn CB, et al. Synovial cell cross-talk with cartilage plays a major role in the pathogenesis of osteoarthritis. Sci Rep. 2020 Jul;10(1):10868. doi: 10.1038/s41598-020-67730-y PMID: 32616761
35. Sukkaew A, Suksatu A, Roytrakul S, Smith DR, Ubol S. Proteomic analysis of CHIKV-infected human fibroblast-like synoviocytes: Identification of host factors potentially associated with CHIKV replication and cellular pathogenesis. Microbiol Immunol. 2020 Jun;64(6):445-57. doi: 10.1111/1348-0421.12793 PMID: 32246487
36. Almagro Armenteros JJ, Sonderby CK, Sonderby SK, Nielsen H, Winther O. DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics. 2017 Jul;33(21):3387-95. doi: 10.1093/bioinformatics/btx431 PMID: 29036616
37. Rawlings ND, Barrett AJ, Bateman A. *MEROPS*: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res. 2012 Jan. 40(Database issue):D343-50. doi: 10.1093/nar/gkr987 PMID: 22086950
38. Jakubiec A, Drugeon G, Camborde L, Jupin I. Proteolytic processing of *turnip yellow mosaic virus* replication proteins and functional impact on infectivity. J Virol. 2007 Oct;81(20):11402-12. doi: 10.1128/JVI.01428-07 PMID: 17686855
39. Jupin I, Ayach M, Jomat L, Fieulaine S, Bressanelli S. A mobile loop near the active site acts as a switch between the dual activities of a viral protease/deubiquitinase. PLOS Pathog. 2017 Nov;13(11):e1006714. doi: 10.1371/journal.ppat.1006714 PMID: 29117247
40. Fieulaine S, Witte MD, Theile CS, Ayach M, Ploegh HL, Jupin I, et al. Turnip yellow mosaic virus protease binds ubiquitin suboptimally to fine-tune its deubiquitinase activity. J Biol Chem. 2020 Oct;295(40):13769-83. doi: 10.1074/jbc.RA120.014628 PMID: 32732284
41. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: A sequence logo generator. Genome Res. 2004 Jun;14:1188-90. doi: 10.1101/gr.849004 PMID: 15173120
42. van der Maaten L, Hinton G. Visualizing data using t-SNE. J Mach Learn Res. 2008 Nov;9:2579-605. doi: http://www.jmlr.org/papers/v9/vandermaaten08a.html
43. Ekici OD, Paetzel M, Dalbey RE. Unconventional serine proteases: Variations on the catalytic Ser/His/Asp triad configuration. Protein Sci. 2008 Dec;17(12):2023-37. doi: 10.1110/ps.035436.108 PMID: 18824507
44. Hofer F, Kraml J, Kahler U, Kamenik AS, Liedl KR. Catalytic site pKa values of aspartic, cysteine, and serine proteases: Constant pH MD simulations. J Chem Inf Model. 2020 Jun;60(6):3030-42. doi: 10.1021/acs.jcim.0c00190 PMID: 32348143
45. Zephyr J, Yilmaz NK, Schiffer CA. Viral proteases: Structure, mechanism and inhibition. Enzymes. 2021 Nov;50:301-33. doi: 10.1016/bs.enz.2021.09.004 PMID: 34861941
46. Liu X, Yang J, Chazi AM, Frey TK. Characterization of the zinc binding activity of the rubella virus nonstructural protease. J Virol. 2000 Jul;74(13):5949-56. doi: 10.1128/jvi.74.13.5949-5956.2000 PMID: 10846076
47. Stempniak M, Hostomska Z, Nodes BR, Hostomsky Z. The NS3 proteinase of hepatitis C virus is a zinc-containing enzyme. J Virol. 1997 Apr;71(4):2881-6. doi: 10.1128/JVI.71.4.2881-2886.1997 PMID: 9060645

48. Saisawang C, Sillapee P, Sinsirimongkol K, Ubol S, Smith DR, Ketterman AJ. Full length and protease domain activity of chikungunya virus nsP2 differ from other alphavirus nsP2 proteases in recognition of small peptide substrates. Biosci Rep. 2015 Jun;35(3):e00196. doi: 10.1042/BSR20150086 PMID: 26182358
49. Lin YL, Lim C. Factors governing the protonation state of Zn-bound histidine in proteins: a DFT/DCM study. J Am Chem Soc. 2004 Mar;126(8):2606-12. doi: 10.1021/ja038827r PMID: 14982470
50. Bennouna A, Gil P, El Rhaffouli H, Exbrayat A, Loire E, Balenghien T, et al. Identification of Eilat virus and prevalence of infection among *Culex pipiens L.* populations, Morocco, 2016. Virology. 2019 Apr;530:85-8. doi: 10.1016/j.virol.2019.02.007 PMID: 30782566
51. Hermanns K, Marklewitz M, Zirkel F, Overheul GJ, Page RA, Loaiza JR, et al. Agua Salud alphavirus defines a novel lineage of insect-specific alphaviruses discovered in the New World. J Gen Virol. 2020 Jan;101(1):96-104. doi: 10.1099/jgv.0.001344 PMID: 31674898
52. Narwal M, Singth H, Pratap S, Malik A, Kuhn RJ, Kumar P, et al. Crystal structure of chikungunya virus nsP2 cysteine protease reveals a putative flexible loop blocking its active site. Int J Biol Macromol. 2018 Sep;116:451-62. doi: 10.1016/j.ijbiomac.2018.05.007 PMID: 29730006
53. (a) Research Collaboratory for Structural Bioinformatics Protein Data Bank. https://www.rcsb.org/structure/3TRK (accessed 2022-01-13). (b) Cheung J, Franklin M, Mancia F, Rudolph M, Cassidy M, Gary E, et al. Structure of the chikungunya virus nsP2 protease. Unpublished work. 2011 Sep.
54. Lev B, Roux B, Noskov SY. Sodium-binding site types in proteins. In: Kretsinger RH, Uversky VN, Permyakov EA, editors. Encyclopedia of metalloproteins. New York: Springer; 2013. pp. 2112-8. doi: 10.1007/978-1-4614-1533-6_242
55. Ahola T, Karlin DG. Sequence analysis reveals a conserved extension in the capping enzyme of the alphavirus supergroup, and a homologous domain in nodaviruses. Biol Direct. 2015 Apr;10:16. doi: 10.1186/s13062-015-0050-0 PMID: 25886938
56. Bernard E, Solignat M, Gay B, Chazal N, Higgs S, Devaux C, et al. Endocytosis of chikungunya virus into mammalian cells: role of clathrin and early endosomal compartments. PLOS One. 2010 Jul;5(7):e11479. doi: 10.1371/journal.pone.0011479 PMID: 20628602
57. Hoornweg TE, Bouma EM, van de Pol DPI, Rodenhuis-Zybert IA, Smit JM. Chikungunya virus requires an intact microtubule network for efficient viral genome delivery. PLOS Negl Trop Dis. 2020 Aug;14(8):e0008469. doi: 10.1371/journal.pntd.0008469 PMID: 32764759
58. Spuul P, Balistreri G, Kaariainen L, Ahola T. Phosphatidylinositol 3-kinase-, actin-, and microtubule-dependent transport of Semliki Forest virus replication complexes from the plasma membrane to modified lysosomes. J Virol. 2010 Aug;84(15):7543-57. doi: 10.1128/JVI.00477-10 PMID: 20484502
59. Gil PI, Albrieu-Llinas G, Mlewski EC, Monetti M, Fozzatti L, Cuffini C, et al. Pixuna virus modifies host cell cytoskeleton to secure infection. Sci Rep. 2017 Jul;7(1);5757. doi: 10.1038/s41598-017-05983-w PMID: 28720756
60. Matthews JD, Morgan R, Sleigher C, Frey TK. Do viruses require the cytoskeleton? Virol J. 2013 Apr;10(1);121. doi: 10.1186/1743-422X-10-121 PMID: 23597412
61. Rhee DK, Marcelino J, Al-Mayouf S, Schelling DK, Bartels CF, Cui Y, et al. Consequences of disease-causing mutations on lubricin protein synthesis, secretion, and post-transcriptional processing. J Biol Chem. 2005 Sep;280(35):32325-32. doi: h10.1074/jbc.M505401200 PMID: 16000300
62. Luo Y, Sinkeviciute D, He Y, Karsdal M, Henrotin Y, Mobasheri A, et al. The minor collagens in articular cartilage. Protein Cell. 2017 Aug;8(8):560-72. doi: 10.1007/s13238-017-0377-7 PMID: 28213717
63. Allen JM, Bateman JF, Hansen U, Wilson R, Bruckner P, Owens RT, et al. WARP is a novel multimeric component of the chondrocyte pericellular matrix that interacts with perlecan. J Biol Chem. 2006 Mar;281(11):7341-9. doi: 10.1074/jbc.M513746200 PMID: 16407285

64. Kruegel J, Sadowski B, Miosge N. Nidogen-1 and nidogen-2 in healthy human cartilage and in late-stage osteoarthritis cartilage. Arthritis Rheum. 2008 May;58(5):1422-32. doi: 10.1002/art.23480 PMID: 18438862
65. Maly K, Schaible I, Riegger J, Brenner RE, Meurer A, Zaucke F. The expression of thrombospondin-4 correlates with disease severity in osteoarthritic knee cartilage. Int J Mol Sci. 2019 Jan;20(2):447. doi: 10.3390/ijms20020447 PMID: 30669608
66. Kobayashi K, Jokaji R, Miyazawa-Hira M, Takatsuka S, Tanaka A, Ooi K, et al. Elastin-derived peptides are involved in the process of human temporomandibular disorder by inducing inflammatory responses in synovial cells. Mol Med Rep. 2017 Sep;16(3):3147-54. doi: 10.3892/mmr.2017.7012 PMID: 28714016
67. Lord MS, Tang F, Rnjak-Kovacina J, Smith JGW, Melrose J, Whitelock JM. The multifaceted roles of perlecan in fibrosis. Matrix Biol. 2018 Aug;68-69:150-66. doi: 10.1016/j.matbio.2018.02.013 PMID: 29475023
68. Mayer U, Kohfeldt E, Timpl R. Structural and genetic analysis of laminin-nidogen interaction. Ann NY Acad Sci. 1998 Oct;857:130-42. doi: 10.1111/j.1749-6632.1998.tb10113.x PMID: 9917838
69. Hohenester E, Yurchenco PD. Laminins in basement membrane assembly. Cell Adh Migr. 2013 Jan:7(1):56-63. doi: 10.4161/cam.21831 PMID: 23076216
70. Funk SD, Miner JH. Muscular dystrophy meets protein biochemistry, the mother of invention. J Clin Invest. 2017 Mar;127(3):798-800. doi: 10.1172/JCI92847 PMID: 28218619
71. Holmes AC, Basore K, Fremont DH, Diamond MS. A molecular understanding of alphavirus entry. PLOS Pathog. 2020 Oct;16(10):e1008876. doi: 10.1371/journal.ppat.1008876 PMID: 33091085
72. Sottrup-Jensen L, Sand O, Kristensen L, Fey GH. The α-macroglobulin bait region: Sequence diversity and localization of cleavage sites for proteinases in five mammalian α-macroglobulins. J Biol Chem. 1989 Sep;264(27):15781-9. PMID: 2476433
73. Gettins PGW, Hahn KH, Crews BC. α2-macroglobulin bait region variants: A role for the bait region in tetramer formation. J Biol Chem. 1995 Jun;270(23):14160-7. doi: 10.1074/jbc.270.23.14160 PMID: 7539801
74. Carrasco L, Sanz MA, Gonzalez-Almela E. The regulation of translation in alphavirus-infected cells. Viruses. 2018 Feb;10(2):70. doi: 10.3390/v10020070 PMID: 29419763
75. Harrington HR, Zimmer MH, Chamness LM, Nash V, Penn WD, Miller TF, et al. Cotranslational folding stimulates programmed ribosomal frameshifting in the alphavirus structural polyprotein. J Biol Chem. 2020 May;295(20):6798-808. doi: 10.1074/jbc.RA120.012706 PMID: 32169904
76. Chauvin C, Salhi S, Le Goff C, Viranaicken W, Diop D, Jean-Jean O. Involvement of human release factors eRF3a and eRF3b in translation termination and regulation of the termination complex formation. Mol Cell Biol. 2005 Jul;25(14):5801-11. doi: 10.1128/MCB.25.14.5801-5811.2005 PMID: 15987998
77. Jaafar ZA, Kieft JS. Viral RNA structure-based strategies to manipulate translation. Nat Rev Microbiol. 2019 Jan;17(2):110-23. doi: 10.1038/s41579-018-0117-x PMID: 30514982
78. Landers VD, Wilkey DW, Merchant ML, Mitchell TC, Sokoloski KJ. The alphavirus capsid protein inhibits IRAK1-dependent TLR signaling. Viruses. 2021 Feb;13(3):377. doi: 10.3390/v13030377 PMID: 33673546
79. Lark T, Keck F, Narayanan A. Interactions of alphavirus nsP3 protein with host proteins. Front Microbiol. 2018 Jan;8:2652. doi: 10.3389/fmicb.2017.02652 PMID: 29375517
80. Mazzon M, Castro C, Thaa B, Liu L, Mutso M, Liu X, et al. Alphavirus-induced hyperactivation of PI3K/AKT directs pro-viral metabolic changes. PLOS Pathog. 2018 Jan;14(1):e1006835. doi: 10.1371/journal.ppat.1006835 PMID: 29377936
81. Lokireddy S, Vemula S, Vadde R. Connective tissue metabolism in chikungunya patients. Virol J. 2008 Feb;5:31. doi: 10.1186/1743-422X-5-31 PMID: 18302795

82. Neuvonen M, Kazlauskas A, Martikainen M, Hinkkanen A, Ahola T, Saksela K. SH3 domain-mediated recruitment of host cell amphiphysins by alphavirus nsP3 promotes viral RNA replication. PLOS Pathog. 2011 Nov;7(11):e1002383. doi: 10.1371/journal.ppat.1002383 PMID: 22114558
83. Tossavainen H, Aitio O, Hellman M, Saksela K, Permi P. Structural basis of the high affinity interaction between the *Alphavirus* nonstructural protein-3 (nsP3) and the SH3 domain of amphiphysin-2. J Biol Chem. 2016 Jul;291(31):16307-17. doi: 10.1074/jbc.M116.732412 PMID: 27268056
84. Chang J, Block TM, Guo JT. Viral resistance of MOGS-CDG patients implies a broad-spectrum strategy against acute virus infection. Antivir Ther. 2015 Jul;20(3):257-9. doi: 10.3851/IMP2907 PMID: 25318123
85. Alavi A, Axford J. Evaluation of β1,4-galactosyltransferase in rheumatoid arthritis and its role in the glycosylation network associated with this disease. Glycoconj J. 1995 Jun;12(3):206-10. doi: 10.1007/BF00731321 PMID: 7496133
86. Ninagawa S, Okada T, Sumitomo Y, Kamiya Y, Kato K, Horimoto S, et al. EDEM2 initiates mammalian glycoprotein ERAD by catalyzing the first mannose trimming step. J Cell Biol. 2014 Aug;206(3):347-56. doi: 10.1083/jcb.201404075 PMID: 25092655
87. Frabutt DA, Zheng YH. Arms race between enveloped viruses and the host ERAD machinery. Viruses. 2016 Sep;8(9):255. doi: 10.3390/v8090255 PMID: 27657106
88. McDowell W, Romero PA, Datema R, Schwarz RT. Glucose trimming and mannose trimming affect different phases of the maturation of Sindbis virus in infected BHK cells. Virology. 1987 Nov;161(1):37-44. doi: 10.1016/0042-6822(87)90168-1 PMID: 2960074
89. Gliedman JB, Smith JF, Brown DT. Morphogenesis of Sindbis virus in cultured *Aedes albopictus* cells. J Virol. 1975 Oct;16(4):913-26. doi: 10.1128/JVI.16.4.913-926.1975 PMID: 170422
90. Kubomura D, Ueno T, Yamada M, Tomonaga A, Nagaoka I. Effect of N-acetylglucosamine administration on cartilage metabolism and safety in healthy subjects without symptoms of arthritis: A case report. Exp Ther Med. 2017 Apr;13(4):1614-21. doi: 10.3892/etm.2017.4140 PMID: 28413518
91. Kaur R, Neetu, Mudgal R, Jose J, Kumar P, Tomar S. Glycan-dependent chikungunya viral infection divulged by antiviral activity of NAG specific chi-like lectin. Virology. 2019 Jan;526:91-8. doi: 10.1016/j.virol.2018.10.009 PMID: 30388630
92. Lin Y, Wu C, Wang X, Liu S, Zhao K, Kemper T, et al. Glucosamine promotes hepatitis B virus replication through its dual effects in suppressing autophagic degradation and inhibiting MTORC1 signaling. Autophagy. 2020 Mar;16(3):548-61. doi: 10.1080/15548627.2019.1632104 PMID: 31204557
93. Kim HB, Lee SW, Mun CH, Yoon JY, Pai J, Shin I, et al. O-linked N-acetylglucosamine glycosylation of p65 aggravated the inflammation in both fibroblast-like synoviocytes stimulated by tumor necrosis factor-α and mice with collagen induced arthritis. Arthritis Res Ther. 2015 Sep;17(1):248. doi: 10.1186/s13075-015-0762-7 PMID: 26370562
94. Decque A, Joffre O, Magalhaes JG, Cossec JC, Blecher-Gonen R, Lapaquette P, et al. Sumoylation coordinates the repression of inflammatory and anti-viral gene-expression programs during innate sensing. Nat Immunol. 2016 Feb;17(2):140-9. doi: 10.1038/ni.3342 PMID: 26657003
95. Stokes S, Almire F, Tatham MH, McFarlane S, Mertens P, Pondeville E, et al. The SUMOylation pathway suppresses arbovirus replication in *Aedes aegypti* cells. PLOS Pathog. 2020 Dec;16(12):e1009134. doi: 10.1371/journal.ppat.1009134 PMID: 33351855
96. Corso D, Chemello F, Alessio E, Urso I, Ferrarese G, Bazzega M, et al. MyoData: An expression knowledgebase at single cell/nucleus level for the discovery of coding-noncoding RNA functional interactions in skeletal muscle. Comput Struct Biotechnol J. 2021 Jul;19:4142-55. doi: 10.1016/j.csbj.2021.07.020 PMID: 34527188

97. Ozden S, Huerre M, Riviere JP, Coffey LL, Afonso PV, Mouly V, et al. Human muscle satellite cells as targets of chikungunya virus infection. PLOS One. 2007 Jun;2(6):e527. doi: 10.1371/journal.pone.0000527 PMID: 17565380
98. Rohatgi A, Corbo JC, Monte K, Higgs S, Vanlandingham DL, Kardon G, et al. Infection of myofibers contributes to increased pathogenicity during infection with an epidemic strain of chikungunya virus. J Virol. 2014 Mar;88(5):2414-25. doi: 10.1128/JVI.02716-13 PMID: 24335291
99. Metz SW, Geertsema C, Martina BE, Andrade P, Heldens JG, van Oers MM, et al. Functional processing and secretion of chikungunya virus E1 and E2 glycoproteins in insect cells. Virol J. 2011 Jul;8:353. doi: 10.1186/1743-422X-8-353 PMID: 21762510
100. Ruiz-Guillen M, Gabev E, Quetglas JI, Casales E, Ballesteros-Briones MC, Poutou J, et al. Capsid-deficient alphaviruses generate propagative infectious microvesicles at the plasma membrane. Cell Mol Life Sci. 2016 Apr;73(20):3897-916. doi: 10.1007/s00018-016-2230-1 PMID: 27117550
101. Martinez MG, Kielian M. Intercellular extensions are induced by the alphavirus structural proteins and mediate virus transmission. PLOS Pathog. 2016 Dec;12(12):e1006061. doi: 10.1371/journal.ppat.1006061 PMID: 27977778
102. Lee CY, Kam YW, Fric J, Malleret B, Koh EGL, Prakash C, et al. Chikungunya virus neutralization antigens and direct cell-to-cell transmission are revealed by human antibody-escape mutants. PLOS Pathog. 2011 Dec;7(12):e1002390. doi: 10.1371/journal.ppat.1002390 PMID: 22144891
103. Varshavsky A. N-degron and C-degron pathways of protein degradation. PNAS. 2019 Jan;116(2):358-66. doi: 10.1073/pnas.1816596116 PMID: 30622213
104. Wang Z, Xia X, Yang X, Zhang X, Liu Y, Wu D, et al. A picorna-like virus suppresses the N-end rule pathway to inhibit apoptosis. eLife. 2017 Dec;6:e30590. doi: 10.7554/eLife.30590 PMID: 29231806
105. Dissmeyer N, Rivas S, Graciet E. Life and death of proteins after protease cleavage: protein degradation by the N-end rule pathway. New Phytol. 2018 May;218(3):929-35. doi: 10.1111/nph.14619 PMID: 28581033
106. Madsen L, Schulze A, Seeger M, Hartmann-Petersen R. Ubiquitin domain proteins in disease. BMC Biochem. 2007 Nov;8(Suppl 1):S1. doi: 10.1186/1471-2091-8-S1-S1 PMID: 18047733
107. Fry MY, Saladi SM, Clemons WM. The STI1-domain is a flexible alpha-helical fold with a hydrophobic groove. Protein Sci. 2021 Apr;30(4):882-98. doi: 10.1002/pro.4049 PMID: 33620121
108. Itakura E, Zavodszky E, Shao S, Wohlever ML, Keenan RJ, Hegde RS. Ubiquilins chaperone and triage mitochondrial membrane proteins for degradation. Mol Cell. 2016 Jul;63(1):21-33. doi: 10.1016/j.molcel.2016.05.020 PMID: 27345149
109. Atallah I, Quinodoz M, Campos-Xavier B, Peter VG, Fouriki A, Bonvin C, et al. Immune deficiency, autoimmune disease and intellectual disability: A pleiotropic caused by biallelic variants in the TPP2 gene. Clin Genet. 2021 Jun;99(6):780-8. doi: 10.1111/cge.13942 PMID: 33586135
110. Reits E, Neijssen J, Herberts C, Benckhuijsen W, Janssen L, Drijfhout JW, et al. A major role for TPPII in trimming proteasomal degradation products for MHC class I antigen presentation. Immunity. 2004 Apr;20(4):495-506. doi: 10.1016/s1074-7613(04)00074-3 PMID: 15084277
111. Pohlers D, Beyer A, Koczan D, Wilhelm T, Thiesen HJ, Kinne RW. Constitutive upregulation of the transforming growth factor-β pathway in rheumatoid arthritis synovial fibroblasts. Arthritis Res Ther. 2007 Jun;9(3):R59. doi: 10.1186/ar2217 PMID: 17594488
112. Uhrlaub JL, Pulko V, DeFilippis VR, Broeckel R, Streblow DN, Coleman GD, et al. Dysregulated TGF-β production underlies the age-related vulnerability to chikungunya virus. PLOS Pathog. 2016 Oct:12(10):e1005891. doi: 10.1371/journal.ppat.1005891 PMID: 27736984
113. Saxena T, Tandon B, Sharma S, Chameettachal S, Ray P, Ray AR, et al. Combined miRNA and mRNA signature identifies key molecular players and pathways involved in chikungunya virus infection in human cells. PLOS One. 2013 Nov;8(11):e79886. doi: 10.1371/journal.pone.0079886 PMID: 24278205

114. Ge G, Greenspan DS. BMP1 controls TGFβ1 activation via cleavage of latent TGFβ-binding protein. J Cell Biol. 2006 Oct;175(1):111-20. doi: 10.1083/jcb.200606058 PMID: 17015622
115. Keski-Oja J. Transforming growth factor beta. In: Schwab M, editor. Encyclopedia of cancer. Berlin: Springer; 2008. pp. 3033-7. doi: 10.1007/978-3-540-47648-1_5916
116. Wu M, Chen G, Li YP. TGF-β and BMP signaling in osteoblast, skeletal development, and bone formation, homeostasis and disease. Bone Res. 2016 Apr;4:16009. doi: 10.1038/boneres.2016.9 PMID: 27563484
117. Spinella-Jaegle S, Roman-Roman S, Faucheu C, Dunn FW, Kawai S, Gallea S, et al. Opposite effects of bone morphogenic protein-2 and transforming growth factor-β1 on osteoblast differentiation. Bone. 2001 Oct;29(4):323-30. doi: 10.1016/s8756-3282(01)00580-4 PMID: 11595614
118. Kasagi S, Chen W. TGF-beta1 on osteoimmunology and the bone component cells. Cell Biosci. 2013 Jan;3(1):4. doi: 10.1186/2045-3701-3-4 PMID: 23321200
119. Chen W, Foo SS, Rulli NE, Taylor A, Sheng KC, Herrero LJ, et al. Arthritogenic alphaviral infection perturbs osteoblast function and triggers pathologic bone loss. PNAS. 2014 Apr;111(16):6040-5. doi: 10.1073/pnas.1318859111 PMID: 24733914
120. Chen W, Foo SS, Sims NA, Herrero LJ, Walsh NC, Mahalingam S. Arthritogenic alphaviruses: new insights into arthritis and bone pathology. Trends Microbiol. 2015 Jan;23(1):35-43. doi: 10.1016/j.tim.2014.09.005 PMID: 25449049
121. Noret M, Herrero L, Rulli N, Rolph M, Smith PN, Li RW, et al. Interleukin 6, RANKL, and osteoprotegerin expression by chikungunya virus–infected human osteoblasts. J Infect Dis. 2012 Aug;206(3):457-9. doi: 10.1093/infdis/jis368 PMID: 22634878
122. Phuklia W, Kasisith J, Modhiran N, Rodpai E, Thannagith M, Thongsakulprasert T, et al. Osteoclastogenesis induced by CHIKV-infected fibroblast-like synoviocytes: a possible interplay between synoviocytes and monocytes/macrophages in CHIKV-induced arthralgia/arthritis. Virus Res. 2013 Nov;177(2):179-88. doi: 10.1016/j.virusres.2013.08.011 PMID: 24012515
123. Roy E, Shi W, Duan B, Reid SP. Chikungunya virus infection impairs the function of osteogenic cells. mSphere. 2020 Jun;5(3):e00347-20. doi: 10.1128/mSphere.00347-20 PMID: 32493723
124. Lainez B, Fernandez-Real JM, Romero X, Esplugues E, Canete JD, Ricart W, et al. Identification and characterization of a novel spliced variant that encodes human soluble tumor necrosis factor receptor 2. Int Immunol. 2004 Jan;16(1):169-77. doi: 10.1093/intimm/dxh014 PMID: 14688072
125. Le HT, Harton JA. Pyrin- and CARD-only proteins as regulators of NLR functions. Front Immunol. 2013 Sep:4:275. doi: 10.3389/fimmu.2013.00275 PMID: 24062743
126. Mounce BC, Cesaro T, Moratorio G, Hooikaas PJ, Yakovleva A, Werneke SW, et al. Inhibition of polyamine biosynthesis is a broad-spectrum strategy against RNA viruses. J Virol. 2016 Oct;90(21):9683-92. doi: 10.1128/JVI.01347-16 PMID: 27535047
127. Mounce BC, Olsen ME, Vignuzzi M, Connor JH. Polyamines and their roles in virus infection. Microbiol Mol Biol Rev. 2017 Sep;81(4):e00029-17. doi: 10.1128/MMBR.00029-17 PMID: 28904024
128. Yuan H, Wu SX, Zhou YF, Peng F. Spermidine inhibits joints inflammation and macrophage activation in mice with collagen-induced arthritis. J Inflamm Res. 2021 Jun;14:2713-21. doi: 10.2147/JIR.S313179 PMID: 34194234
129. Sacitharan PK, Lwin S, Gharios GB, Edwards JR. Spermidine restores dysregulated autophagy and polyamine synthesis in aged and osteoarthritic chondrocytes via EP300. Exp Mol Med. 2018 Sep;50(9):1-10. doi: 10.1038/s12276-018-0149-3 PMID: 30232322
130. Chen Z, Lin CX, Song B, Li CC, Qiu JX, Li SX, et al. Spermidine activated RIP1 deubiquitination to inhibit TNF-α-induced NF-κB/p65 signaling pathway in osteoarthritis. Cell Death Dis. 2020 Jul;11(7):503. doi: 10.1038/s41419-020-2710-y PMID: 32632306

131. Karouzakis E, Gay RE, Gay S, Neidhart M. Increased recycling of polyamines is associated with global DNA hypomethylation in rheumatoid arthritis synovial fibroblasts. Arthritis Rheum. 2012 Jun;64(6):1809-17. doi: 10.1002/art.34340 PMID: 22170508

132. Flescher E, Bowlin TL, Ballester A, Houk R, Talal N. Increased polyamines may downregulate interleukin 2 production in rheumatoid arthritis. J Clin Invest. 1989 Apr;83(4):1356-62. doi: 10.1172/JCI114023 PMID: 2784801

133. Gutierrez E, Shin BS, Woolstenhulme CJ, Kim JR, Saini P, Buskirk AR, et al. eIF5A promotes translation of polyproline motifs. Mol Cell. 2013 Jul;51(1):35-45. doi: 10.1016/j.molcel.2013.04.021 PMID: 23727016

134. Schuller AP, Wu CCC, Dever TE, Buskirk AR, Green R. eIF5A functions globally in translation elongation and termination. Mol Cell. 2017 Apr;66(2):194-205.e5. doi: 10.1016/j.molcel.2017.03.003 PMID: 28392174

135. Hryniewicz MM, Vonder Haar RA. Polyamines enhance readthrough of the UGA termination codon in a mammalian messenger RNA. Mol Gen Genet. 1983 May;190(2):336-43. doi: 10.1007/BF00330661 PMID: 6576222

136. Kim KW, Kim BM, Lee KA, Lee SH, Firestein GS, Kim HR. Histamine and histamine H4 receptor promotes osteoclastogenesis in rheumatoid arthritis. Sci Rep. 2017 Apr;7(1):1197. doi: 10.1038/s41598-017-01101-y PMID: 28446753

137. Chiurchiu V, Leuti A, Maccarrone M. Bioactive lipids and chronic inflammation: managing the fire within. Front Immunol. 2018 Jan;9:38. doi: 10.3389/fimmu.2018.00038 PMID: 29434586

138. Hassan MQ, Hadi RA, Al-Rawi ZS, Padron VA, Stohs SJ. The glutathione defense system in the pathogenesis of rheumatoid arthritis. J Appl Toxicol. 2001 Feb;21(1):69-73. doi: 10.1002/jat.736 PMID: 11180282

139. Elmasri Z, Nasal BL, Jose J. Alphavirus-induced membrane rearrangements during replication, assembly, and budding. Pathogens. 2021 Aug;10(8):984. doi: 10.3390/pathogens10080984 PMID: 34451448

140. Jose J, Taylor AB, Kuhn RJ. Spatial and temporal analysis of alphavirus replication and assembly in mammalian and mosquito cells. mBio. 2017 Feb;8(1):e02294-16. doi: 10.1128/mBio.02294-16 PMID: 28196962