



# Analysing Point Mutations in Protein Cleavage Sites by Using Enzyme Specificity Matrices

Jakob Triebel<sup>1\*</sup>, Sandeep Silawal<sup>2</sup>, Maximilian Willauschus<sup>2</sup>, Gundula Schulze-Tanzil<sup>2</sup> and Thomas Bertsch<sup>1</sup>

<sup>1</sup> Institute for Clinical Chemistry, Laboratory Medicine and Transfusion Medicine, Nuremberg General Hospital and Paracelsus Medical University, Nuremberg, Germany, <sup>2</sup> Department of Anatomy, Paracelsus Medical University, Nuremberg, Germany

**Keywords:** MEROPS, Ensembl, UniProt, MEGA, cleavage sites, enzymes, point mutations

Public databases contain a rapidly expanding amount of information about protein sequences of diverse species, genetic mutations in humans, as well as enzymes and their substrates. This opens new possibilities of data analysis, hypothesis generation, and testing. It is the purpose of the present article to briefly highlight two exemplary analyses which are of broad applicability: the effect of point mutations on the generation of hormones, and the investigation of the molecular evolution of hormonal systems.

The pituitary hormone prolactin, for example, can undergo proteolytic cleavage by cathepsin D and other proteases. Such cleavage can result in the generation of vasoinhibin, a hormone with a range of effects and signaling mechanisms distinct from uncleaved prolactin. It was projected that missense point mutations in human prolactin, which lead to an amino acid change within a specific cleavage site, can inhibit or facilitate the enzymatic cleavage, thus altering the levels of vasoinhibin (1). The general principle underlying this projection, is that point mutations in cleavage sites can affect the levels of the effector hormone which is arising from the cleavage of its precursor. This alteration of hormone levels may be clinically relevant in terms of a protective or an aggravating factor, depending on the function of the hormone or protein in a disease. Indeed, if leucine on position P1 (the first position next to the cleavage site NH<sub>2</sub>-terminal) is affected by a missense mutation and replaced by tryptophan, it is very likely that the cleavage that usually occurs at this site is significantly inhibited. This is because leucine on position P1 is the most preferred amino acid by cathepsin D, as shown by its entry in the MEROPS data base<sup>1</sup> (2), which lists 416 observed cleavages in which leucine was present at this position in the substrate. For tryptophan, this number is merely 29, indicating a lower preference of cathepsin D for this amino acid at this position. If, in another scenario, serine in P1 (12 observed cleavages) is affected by a mutation and replaced by leucine, it is likely that the change significantly enhances the proteolytic cleavage, because, as mentioned, leucine is the most preferred amino acid at P1. Of note, there is experimental evidence supporting this type of projection as single amino acid substitutions in prolactin and the subsequent cleavage efficiency tested with cathepsin D correspond well with the expected changes according to the numbers of observed cleavages in the MEROPS enzyme specificity matrix (1, 3).

When analyzing point mutations in prolactin located in cleavage sites generating vasoinhibin, retrieved from the Ensembl-database<sup>2</sup> (4), significant changes of the projected cleavage efficiency were found. The cleavage efficiency was determined by addition of the numbers of observed cleavages for each amino acid at a particular position, generating an 8P-score as a surrogate parameter for the efficiency of a given sequence. For example, the mutation leading to the replacement of methionine by valine in position P3 of a cleavage site for the generation of a 15 kDa

## OPEN ACCESS

### Edited by:

Nils P. Krone,  
University of Sheffield,  
United Kingdom

### Reviewed by:

Consolato Sergi,  
University of Alberta Hospital, Canada

### \*Correspondence:

Jakob Triebel  
Jakob.Triebel@gmx.de

### Specialty section:

This article was submitted to  
Genomic Endocrinology,  
a section of the journal  
Frontiers in Endocrinology

**Received:** 17 August 2018

**Accepted:** 11 April 2019

**Published:** 03 May 2019

### Citation:

Triebel J, Silawal S, Willauschus M,  
Schulze-Tanzil G and Bertsch T (2019)  
Analysing Point Mutations in Protein  
Cleavage Sites by Using Enzyme  
Specificity Matrices.  
Front. Endocrinol. 10:267.  
doi: 10.3389/fendo.2019.00267

<sup>1</sup> Available online at: <https://www.ebi.ac.uk/merops/>

<sup>2</sup> Available online at: <http://www.ensembl.org>

**TABLE 1 |** Effect of mutation p.M158V in prolactin on cleavage efficiency.

Cleavage site for the 15 kDa vasoinhibin isoform in prolactin generated by Cathepsin D, ancestral sequence								
Amino acid position	157	158	159	160	161	162	163	164
Cleavage site position	P4	P3	P2	P1	P1'	P2'	P3'	P4'
Amino acid	Gly	Met	Glu	Leu	Ile	Val	Ser	Gln
Merops specificity CD	42	26	94	416	119	131	55	52
8P-score				935				
Prolactin sequence carrying point mutation p.M158V, ID: rs749978269								
Amino acid position	157	158	159	160	161	162	163	164
Cleavage site position	P4	P3	P2	P1	P1'	P2'	P3'	P4'
Amino acid	Gly	Val	Glu	Leu	Ile	Val	Ser	Gln
Merops specificity CD	42	94	94	416	119	131	55	52
8P-score				1003 (+68)				

The point mutation rs749978269 leading to a substitution of methionine by valine in position 158 is projected to increase the cleavage efficiency of cathepsin D (CD) at this site, according to the MEROPS specificity data for cathepsin D and the calculated 8P-score. Modified from Table 1 as published by Triebel et al. (1).

vasoinhibin-isoform, resulted in an increase of the 8P-score by 68 points, suggestive of a higher cleavage efficiency because of this point mutation (1) (Table 1). In the case of an alteration of vasoinhibin levels because of enhanced or inhibited cleavage, the clinical relevance of the point mutations can be assumed for diseases in which a dysregulation of vasoinhibin levels has been reported, such as diabetic retinopathy (5) and peripartum cardiomyopathy (6). Point mutations in cleavage sites may therefore constitute a “second-hit”-factor in the pathophysiology of diseases in which the cleavage product is an important detrimental player, or they could constitute a protective factor. Hence, we propose that this type of analyses of detected point mutations for which no phenotype is known, will lead to the identification of clinically relevant point mutations. As a matter of course, the clinical relevance of mutations can only be

demonstrated if these mutations are found in patients with the disease under investigation and if their effect on the cleavage is experimentally confirmed, but the theoretical projection of the cleavage efficiency will help to identify promising candidates. On the background of the many hormones and other proteins which arise from enzymatic cleavage of a precursor, for example adrenocorticotrophic hormone and insulin, this appears as a rewarding option.

Another application of the strategy using the enzyme substrate specificity data from MEROPS is its use in conjunction with protein sequences retrievable from UniProt<sup>3</sup> (7) or Ensembl. Multiple protein sequence alignments of sequences from various species can be build, using MEGA<sup>4</sup> (8) for example, and the conservation of cleavage sites can be evaluated, including an estimation of how the cleavage changes in response to sequence variations. This will likely produce valuable insights into the evolution of endocrine systems during speciation, but also into the evolution of any other protein which is generated by posttranslational proteolytic processing.

The use of MEROPS and its possible applications has been described in detail by its creators, Rawlings et al. (2). Here, we briefly point out methods which can be used to identify clinically relevant point mutations and to generate insights into the evolution of hormones and other proteins which emerge from proteolytic cleavage of a precursor. This shows once more the great value that emerges from public databases such as MEROPS, UniProt, and Ensembl. The existence of these databases, the efforts of their creators, curators, and funding bodies are much appreciated by the authors.

## AUTHOR CONTRIBUTIONS

JT wrote the manuscript. SS, MW, GS-T, and TB edited the manuscript. All authors approved the final version.

<sup>3</sup> Available online at: <https://www.uniprot.org/>

<sup>4</sup> Available online at: <https://megasoftware.net/>

## REFERENCES

1. Triebel J, Friedrich CJ, Leuchs A, Martinez de la Escalera G, Clapp C, Bertsch T. Human prolactin point mutations and their projected effect on vasoinhibin generation and vasoinhibin-related diseases. *Front. Endocrinol.* (2017) 8:294. doi: 10.3389/fendo.2017.00294
2. Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res.* (2018) 46:D624–32. doi: 10.1093/nar/gkx1134
3. Piwnica D, Touraine P, Struman I, Tabruyn S, Bolbach G, Clapp C, et al. Cathepsin D processes human prolactin into multiple 16K-like N-terminal fragments: study of their antiangiogenic properties and physiological relevance. *Mol Endocrinol.* (2004) 18:2522–42. doi: 10.1210/me.2004-0200
4. Zerbinio DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, et al. Ensembl 2018. *Nucleic Acids Res.* (2018) 46:D754–61. doi: 10.1093/nar/gkx1098
5. Triebel J, Robles-Osorio ML, Garcia-Franco R, Martinez de la Escalera G, Clapp C, Bertsch T. From bench to bedside: translating the prolactin/vasoinhibin axis. *Front Endocrinol.* (2017) 8:342. doi: 10.3389/fendo.2017.00342

6. Hilfiker-Kleiner D, Kaminski K, Podewski E, Bonda T, Schaefer A, Sliwa K, et al. A cathepsin D-cleaved 16 kDa form of prolactin mediates postpartum cardiomyopathy. *Cell.* (2007) 128:589–600. doi: 10.1016/j.cell.2006.12.036
7. UniProt Consortium T. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* (2018) 46:2699. doi: 10.1093/nar/gky092
8. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol.* (2018) 35:1547–9. doi: 10.1093/molbev/msy096

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Triebel, Silawal, Willauschus, Schulze-Tanzil and Bertsch. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.