Pepsin properties, structure, and its accurate measurement: a narrative review

Kyle J. Stanforth¹, Matthew D. Wilcox¹, Matthew D. Wilcox¹, Peter I. Chater¹, Iain A. Brownlee², Maria I. Zakhour¹, Katherine M. R. M. Banecki³, Jeffery P. Pearson¹

¹Newcastle University Biosciences Institute, Medical School, Framlington Place, Newcastle upon Tyne, Tyne and Wear, UK; ²Applied Sciences, Northumbria University, Sandyford Rd, Newcastle upon Tyne, UK; ³Department of Pharmacology, Medical Sciences Division, University of Oxford, Level 3, John Radcliffe Hospital, Oxford, UK

Contributions: (I) Conception and design: All authors; (II) Administrative support: All authors; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Matthew D. Wilcox. Newcastle University Biosciences Institute, Medical School, Framlington Place, Newcastle upon Tyne, Tyne and Wear. NE2 4HH, UK. Email: Matthew.wilcox@ncl.ac.uk.

Abstract: Pepsin is an aspartate protease that is generated from its proenzyme, pepsinogen by autocatalysis initiated by a fall in pH below 5. Human gastric juice contains eight isoenzymes of pepsin. The peptides released on conversion of pepsinogen to pepsin of which there are potentially five, have been shown to have antimicrobial activity against a wide range of bacteria including Escherichia coli, Pseudomonas and Staphylococcus which have also been shown to have biofilm formation inhibiting properties. The stability in response to changes in pH varies between pepsin and pepsinogen. Pepsinogen is stable up to pH 10, pepsin is only stable to pH just above 7.0 and is completely denatured at pH 8.0. Many diseases of the aerodigestive tract have been linked to reflux and the presence of pepsin. Therefore, the measurement of pepsin in tissue and lavages or in saliva or sputum, could be a good screening tool for the diagnosis of reflux related disease. However, there is no current consensus as to the best methods to measure it or the best time to sample it. For an effective pepsin ELISA, the following is required; a monoclonal/monospecific polyclonal antibody with a good lowest level of detection (LLOD) and sensitivity 1–25 ng/mL (depending on dilution) and an adequate supply of purified human pepsin as a standard for antibody-based assays. If possible, an activity assay for pepsin should also be used as the presence of pepsin protein does not indicate it is capable of damaging activity. Finally, if pepsin is associated with a disease large studies are required to confirm it with multiple samples. This review deals with several studies where pepsin quantitation is attempted, and their measurement techniques assessed.

Keywords: Pepsin; pepsinogen; pepsin measurement; pepsin detection

Received: 01 December 2020. Accepted: 22 February 2021.

do: 10.21037/aoe-20-95

View this article at: http://dx.doi.org/10.21037/aoe-20-95

Pepsin types and structure

Pepsins are aspartate proteases active in acidic conditions and are the major proteases in human gastric juice. Pepsins are therefore important in normal digestion of proteins and along with acid in the protection against ingested pathogenic organisms gaining a foothold in the gastrointestinal tract. The active enzyme is derived from its pro-enzyme pepsinogen which is secreted into the

^ ORCID of Matthew D. Wilcox: 0000-0003-0150-0106; Peter I. Chater: 0000-0001-8679-0004; Jeffery P. Pearson: 0000-0002-9816-7866.
pH 3.5. Bacterial viability was measured as CFU/mL and *Salmonella typhimurium* which has resistance to low pH had ~50% survival after 45 minutes exposure to pH 3.5. When the full-length peptide was added this was reduced to ~20% (12).

The two shorter peptides performed worse but were still effective around 50 µM (*Table 1*). To determine if this effect also occurred at acidic conditions pertaining to the stomach the full-length peptide was tested at pH 3.5. Bacterial viability was measured as CFU/mL and *Salmonella typhimurium* which has resistance to low pH had ~50% survival after 45 minutes exposure to pH 3.5. When the full-length peptide was added this was reduced to ~20% (12).

*Table 1* Antimicrobial activity of pepsinogen activation peptides

<table>
<thead>
<tr>
<th>Microbe</th>
<th>Minimum inhibitory concentration (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1–47</td>
</tr>
<tr>
<td><em>Escherichia coli</em></td>
<td>6.3</td>
</tr>
<tr>
<td><em>Salmonella typhimurium</em></td>
<td>6.3</td>
</tr>
<tr>
<td><em>Pseudomonas aeruginosa</em></td>
<td>6.3</td>
</tr>
<tr>
<td><em>Staphylococcus aureus</em></td>
<td>6.3</td>
</tr>
<tr>
<td><em>Listeria monocytogenes</em></td>
<td>&gt;50</td>
</tr>
</tbody>
</table>

*Data taken from Pane et al. (16).*

Pepsin activation and function of activation peptides

The mechanism of activation by just a change in pH is well understood in human and porcine pepsin (5,11). The active peptide, the N-terminal of the pepsinogen molecule which is cleaved off when pepsin is activated, has recently been shown to have antimicrobial properties (12). Human pepsinogen has a predicted amino acid sequence of 373 amino acids (13). In human pepsin the cleaved peptide is 47 amino acids long and is generally released as 2 peptides generating an active enzyme with a molecular weight 34–37 KD (5,14,15). The cleavage occurs in the region Leu23-Lys24-Asp25-Phe26 but the full-length peptide can also be detected, resulting in the appearance of 5 potential peptides in gastric juice 1–23, 1–25, 24–47, 26–47 and 1–47 (5). In order to test the antimicrobial properties of thesepane et al. expressed recombinant proteins in *E. coli* BL21 (DE3) producing the full length 47 amino acid peptide, a 1–25 and a 26–47 amino acid peptide. The only difference between these peptides and the original pepsinogen peptides is a proline at the N-terminal which was placed in the pepsinogen sequence to produce acid labile sequences and this proline is unlikely to affect any anti-microbial properties (16,17). Computational analysis was used to identify cryptic antimicrobial peptides (AMPs) based on amino acid composition, charge, and hydrophobicity of known AMPs. This analysis indicated that the pepsinogen peptides would have antimicrobial activity. The full-length peptide was shown to have a minimum inhibitory concentration (MIC) of ≤12.5 µM against a wide range of bacteria including *E. coli*, *Pseudomonas*, and *Staphylococcus*, but greater than 50 µM for *Listeria*.

The two shorter peptides performed worse but were still effective around 50 µM (*Table 1*). To determine if this effect also occurred at acidic conditions pertaining to the stomach the full-length peptide was tested at pH 3.5. Bacterial viability was measured as CFU/mL and *Salmonella typhimurium* which has resistance to low pH had ~50% survival after 45 minutes exposure to pH 3.5. When the full-length peptide was added this was reduced to ~20% (12).

The 1–25 peptide was also effective in preventing biofilm formation of *Pseudomonas aeruginosa* and *Staphylococcus aureus*. Finally, in a mouse model of *Pseudomonas aeruginosa* infection the full-length peptide showed 4 orders of magnitude reduction in CFU/mL and the two shorter peptides ~2 orders of magnitude (12).

These findings highlight that pepsin’s activation peptides generated from secreted pepsinogen may form an important element of the antimicrobial barrier function of gastric lumen of the gastric glands from chief cells (peptic cells). Pepsinogens consist of two immunological groups PGI containing pepsinogens 1, 2, 3, 4 and 5 and PGII containing pepsinogens 6 and 7 (1,2). Pepsinogen is a symmetrical molecule with N and C terminal lobes with a molecular weight between 40–42 K (3) and is stable up to pH values of 10 (4). However, when it is exposed to pH levels below 5, which occurs in the gastric glands, an auto-catalytic activation occurs. This involves the removal of a section of the pepsinogen's N-terminal (5-7). The generated pepsins consist of 1, 2, 3a, 3b, 3c, 4, 5, 6, 7 and can be separated by anion exchange high performance liquid chromatography (HPLC) and agar gel electrophoresis (8,9). Pepsin 7 isolated from gastric mucosa also referred to as a slow-moving protease is an aspartate protease. However, it is not a pepsin but instead a cathepsin E which is an intracellular enzyme (10).
juice. This may be particularly important in the case of microbial resistance to low pH.

**Pepsinogen and pepsin stability**

Unlike pepsinogen, pepsin is not stable at alkaline pH. When purified pepsin 3 was incubated over 24 hours at 37 °C with pH ranging from 2–8, when measuring at pH 2, 100% of activity could be recovered after the pre-incubation between pH 4–6.8 but no activity was recovered after pH 7.0. The activity after pre-incubation at pH 2.0 shows only 60% is recovered, due to direct denaturation as a result of pepsin auto-digestion. When this is repeated with human gastric juice pepsin activity is recoverable after incubation up to pH 7.5 but completely lost at pH 8.0 (18). Studies by Johnston et al. (19) carrying out similar experiments with pepsin 3 showed no activity could be recovered after 24-hour incubation at pH 8.0. The upshot of these studies suggests that pepsin is irreversibly denatured at pH between 7.0 and 8.0. The reason for the differences in stability between pepsinogen and pepsin is the missing activation peptides, meaning that the C-terminal can refold but not the N-terminal (20-22). In considering activity, pepsin shows proteolytic activity up to pH 6.5 so at pH 6.5 pepsin will have no activity but is still stable and can be re-activated with a drop in pH. Consequently, pepsin present in a reflux episode which leaves the oesophagus can bind to the mucosa of the aerodigestive tract, remain inactive but native after neutralization by saliva and bicarbonate. The bound pepsin can be re-activated by a new reflux event if the pH is below 6.0 with resulting tissue damage (19,23). Pepsin from refluxate is recognised as a biomarker of reflux as well as in tissue damage. In addition to pepsin, bile acids have been strongly implicated when present in the refluxate as biomarkers of reflux and damaging agents (19,23). However, many of the papers dealing with bile acids are using assays which do not have the required sensitivity (24-28).

### Diseases associated with reflux identified by the presence of pepsin

Over the past ten years or more the presence of pepsin has been associated with many diseases of the aerodigestive tract including gastro-oesophageal reflux disease (GORD), Laryngopharyngeal reflux (LPR), rhinitis and sinusitis, vocal fold leucoplakia (VFL), laryngomalacia and several lung diseases (28-43) (Table 2). Reflux of gastric contents associated diseases have become so prevalent that in the recent Incredibles 2 film a new superhero Reflux was born.

### What do we need to measure pepsin? Is pepsin being measured correctly? Pepsin or pepsinogen?

Pepsin is not secreted at proximal sites in the gastrointestinal tract. Therefore, it represents a rational and objective marker of recent reflux events when detected in biological samples from the aerodigestive tract, like saliva and sputum. Measurement of pepsin concentration would not provide information on whether pepsin detected was active or denatured. Due to the characteristic proteolytic action at low pH of pepsins, it is also possible to define their activity by kinetics which may provide further insights into their damaging potential in reflux.

A monoclonal/monospecific polyclonal antibody ELISA with a good lowest level of detection (LLOD) and sensitivity 1–25 ng/mL (depending on dilution) and an adequate supply of human pepsin as a standard are needed. If possible, an activity assay for pepsin should also be used as the presence of pepsin protein does not indicate it is capable of damaging activity. Finally, if pepsin is associated with a disease large studies are required to confirm it.

We have isolated pepsin using anion exchange HPLC, from human gastric juice obtained at endoscopy. Figure 1 shows pepsin eluting at 10–15 minutes from the HPLC column has a molecular weight just below 37,000. The fraction collected between 10–15 minutes also showed proteolytic activity at pH 2.2 using an N-terminal plate assay (44-46). This is characteristic of human pepsin.

There are several examples in the literature where these...
criteria are not followed. Iannella *et al.* (47) investigated pepsin in tears of children with LPR and suggested a route for pepsin in the refluxate to the pre-corneal film via the nasal fossa, inferior meatus, and nasolacrimal duct. The problems with this study involves the assay used to measure pepsin. It was an ELISA supplied by DRG Inc., Germany. However, this company has never had a pepsin ELISA only a pepsinogen one. The authors determined the LLOD as 2.5 ng/mL and considered the sensitivity as a percentage increase over the blank e.g., 2.5 ng/mL gave a 33.5% increase in absorbance over the blank. It is difficult to ascertain the absorbance changes and to know if they are in the spectrophotometer’s measurable range. They reported pepsin levels in tears between 3.5–5.4 ng/mL in 4/20 children with LPR, confirmed by 24-hour impedance and with non LPR controls having no pepsin. The problem with values this small is that serum levels of pepsinogen are 50–87 ng/mL (18) and the ELISA used will measure pepsinogen.

A study by O’Reilly *et al.* (36) investigating the role of gastric pepsin in the inflammatory cascade of paediatric otitis media used a large group of 129 subjects with 50% positive for pepsin in the middle ear effusions (glue). The authors used a monoclonal antibody supplied by Bio-Rad Serotec against human pepsin A in a standard ELISA with a claimed lower level of sensitivity of 0.1 ng/mL and with a cut off for pepsin detection starting at 0.25 ng/mL. Alongside this pepsin effusions were assayed for bacterial infection and several cytokines. The results demonstrated that the presence of pepsin correlated with age younger than three years. Pepsin presence was not associated with increased bacterial infection, but IL-8 levels were significantly associated with bacterial infection levels. The treatment for glue ear is to drain the middle ear cleft and place a tympanostomy tube in the ear drum. Interestingly this study showed that the levels of pepsin were significantly associated with IL-8 and the need for second and third tubes. This suggests that when reflux persists as identified by the presence of pepsin, second and third operations would be required. This data suggests a cascade as shown in Figure 2 with pepsin causing damage to the middle ear and inflammation resulting in increased synthesis and secretion of IL-8, a recognised mucin secretagogue (47-49); causing an increase in mucin secretion accumulating in the middle ear cleft. Again, this study has some problems relating to the measurement of pepsin. There is no report of the actual pepsin levels. The Bio-Rad Serotec company website confirms the antibody is monoclonal and it recognises both human pepsinogen 3 and pepsin 3 and that it recognises a 47,000 molecular weight protein from gastric juice. Pepsin has a molecular weight ~35,000 and pepsinogen ~42,000.

### Figure 1
Human gastric juice was obtained from volunteers undergoing routine gastroscopy. Pepsin was isolated through HPLC on an anion exchange column as per the methods of Coyle (in 2006) and Jones *et al.* (1993) with slight modification. Sample fractions were collected at 0–5, 5–10, 10–15, and 15–30 minutes, and run on an SDS-Page (45,46).

### Figure 2
Pepsin mediated cascade in otitis media. This figure was produced from an interpretation of the data from (36).
26 patients and 20 controls. It used immunohistochemistry with a polyclonal pepsin antibody and the results showed a significant increase in pepsin staining in VFL, compared to the controls suggesting LPR could be a risk factor for the development of VFL. It is difficult to assess the validity of this paper as there is little information from the supplier, Cloud-Clone Corp as to the specificity and cross reactivity of the antibody. Luebke et al. (38) carried out a small study of 10 laryngomalacia (LM) patients under 3 years old and 5 age matched control subjects. Pepsin was measured in supraglottic lavage samples in the controls and laryngomalacia patients and in biopsies from the laryngomalacia patients only. The results showed that 8/10 of the laryngomalacia subjects had lavage pepsin but none of the controls had. Four out of 10 of the LM subject biopsies showed the presence of pepsin. If the biopsy was pepsin positive so was the lavage. They concluded that refluxed pepsin was implicated as having a role in laryngomalacia. This paper used correct antibodies and detected pepsin. They measured pepsin using SDS PAGE and western blot in biopsies and lavage samples using a monoclonal antibody raised to human pepsin A which will recognise both pepsin and pepsinogen but as they have different molecular weights they could be differentiated on western blots. The antibody used in an ELISA to quantitate the pepsin levels in the lavages was raised against the N-terminus of human pepsin 3b, so it will not be expected to recognise human pepsinogen as it has a different N-terminus.

The authors do not give a LLLOD value for the ELISA so it is difficult to know if the pepsin values can be reported in ng/mL to one decimal point. In addition, they very diligently reported significantly higher median pepsin was observed in the LM patients compared to the controls in which no pepsin was detected. Calvo-Henríquez et al. (32) in 2017 carried out a systematic review to investigate if pepsin was a reliable marker of LPR. They initially included all studies up to December 2016 that contained pepsin measurements in LPR. 146 studies were found and any studies with no controls or a small sample size (less than 20) were excluded. This left 12 studies. The methods used to measure pepsin in the 12 studies were biopsy immunohistochemistry, western blot, and saliva ELISA and Peptest—a lateral flow device using monoclonal antibodies. Ten of the twelve studies found pepsin was significantly increased in LPR cases versus controls (32). The two studies that failed to show a significant link could be explained as follow, Komatsu et al. (30) measuring pharyngeal biopsies and western blots did not use non-symptomatic controls but instead used GORD patients as controls. Yadlapati et al. (51) using the Peptest measuring salivary pepsin also failed to show a significant increase in positive pepsin in LPR compared to controls. This may in part be due to the patient group being defined by reflux symptom index (RSI) (52), not the most reliable method of diagnosis of LPR. However, if pepsin (semi-quantitative) concentrations were compared rather than just pepsin positivity then significant differences were found between laryngeal and oesophageal symptoms and the control groups. This systematic review concluded that pepsin could be a reliable marker of LPR but optimal sampling time and threshold/cut off levels need to be defined. In addressing this Klimara et al. (in 2020) (28) investigated the optimum time for collection of samples of saliva and nasal lavage as markers of LPR. The study used 26 patients tested with 24-hour multi-channel intraluminal impedance (MII-pH) and reflux finding score (RFS) (53) and RSI to evaluate LPR. The control group was 13 subjects defined as reflux absent. Nasal lavage samples were collected at tube placement and saliva samples collected at different times. Firstly, in the clinic before MII-pH probe placement then an hour after each meal and on waking the next morning. Pepsin was measured with an ELISA using the monoclonal antibody previously described in Luebke et al. (38) specific for human pepsin. No pepsin was detected in any saliva samples or nasal lavage from the control subjects. The results showed that of the patients with suspected LPR, 19 out of 26 were confirmed by MII-pH and eight of the 19 had pepsin in any collected sample with 5 in the morning saliva sample. In addition, three of the seven patients negative with MII-pH had pepsin in any collected sample and all three of these had pepsin in the morning sample. The highest levels of pepsin were found in the nasal lavage samples with a mean of 7,662 ng/mL which could mean pepsin had accumulated in the nose from previous reflux events. The highest levels of salivary pepsin were in the morning after waking with a mean of 187 ng/mL with much lower values at the other time points. The pepsin levels on waking also significantly correlated with MII-pH parameters indicating proximal reflux and RSI scores but not with RFS. In conclusion if salivary pepsin is to be a useful diagnostic tool for LPR the levels should be measured on waking or additionally at a time when the patient identifies a reflux event.

Over recent years there has been accumulating evidence that measurement of pepsin in saliva could make a good screening tool in the diagnosis of reflux related diseases (28,43,54-60).

However, a paper published in 2019 (61) placed doubt on the usefulness of salivary pepsin in patients with GORD.
This study used a relatively small study group, 30 GORD patients and 20 asymptomatic controls and collected saliva samples at different time points when a dual channel pH catheter was in place. They compared a commercial lateral flow device (Peptest, RD biomed Hull UK) which uses pepsin specific monoclonal antibodies with a pepsin ELISA, using a monoclonal antibody supplied by Santa Cruz (sc365680).

They concluded that Peptest did not identify the patients from the controls based on pepsin levels. In addition, the ELISA did not show a difference in pepsin levels between the patients and the controls. Also, there was no correlation between the Peptest and the pepsin ELISA. This paper has several flaws. Firstly, the Peptest does measure the major human pepsin in gastric juice shown by its reactivity with human pepsin of the correct molecular in ELISA in Race et al. using an antibody specifically designed as a research tool and not for diagnostic purposes. Figure 3 taken from the Santa Cruz website shows a western blot of human stomach tissue extract using the antibody raised against a recombinant protein mapping to amino acids 281–324 near the C terminus of pepsin A (a member of the PGI group). Consequently, as pepsinogens and pepsins have the same C-terminal it will recognise both. Based on Figure 3 the antibody is recognising several components of stomach tissue, most if not all are larger than the molecular weight of pepsin or pepsinogen (3,5,13-15). This indicates a lack of specificity for pepsin. The antibody does recognise a recombinant protein containing amino acids 15–388 of pepsin A and this is not in dispute, but it is what else it recognises that is the pertinent question. This stresses the importance of using human pepsin isolated from gastric juice to test antibody specificity and to produce standard curves.

Secondly the 37% detection of pepsin in the control group samples needs to be addressed as other studies (28,62) have shown that healthy controls were negative for pepsin. In any study involving intubation a control without intubation should be carried out to determine if intubation causes reflux (63).

Thirdly the authors suggest that the presence of pepsin in saliva is not from reflux but from expression of pepsin by the tongue. This can be dismissed as the Human Expression Atlas lists nine studies on tissue expression of pepsinogen and only one reports any in the tongue at levels too low to account for the levels detected in saliva (64).

Conclusions

Pepsins in human gastric juice exist as several isoforms and pepsin is activated from pepsinogen by simply a fall in pH. In terms of proteolytic activity pepsin should not be regarded as simply a low pH protease as it has activity close to neutral. Pepsin has a key function in the gastric juice as an antimicrobial agent via its proteolytic activity, but we now know the activation peptides from pepsinogen can also act as antimicrobial peptides. Pepsin has the potential to be a biomarker of diseases associated with reflux of gastric contents. However, it is key to ensure that any assay system has the required specificity and sensitivity.

Acknowledgments

Funding: This work was funded by the Biotechnology and Biological Sciences Research Council Super Follow-on Fund (BB/R019657/1).

Footnote

Reporting Checklist: The authors have completed the Narrative Review reporting checklist. Available at http://dx.doi.org/10.21037/aoe-20-95

Provenance and Peer Review: This article was commissioned by the editorial office, Annals of Esophagus for the series “Epidemiology, Biomarkers and Modelling of
Gastroesophageal Reflux Disease”. The article has undergone external peer review.

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/aoe-20-95). The series “Epidemiology, Biomarkers and Modelling of Gastroesophageal Reflux Disease” was commissioned by the editorial office without any funding or sponsorship. JPP served as the unpaid Guest Editor of the series. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References
22. Privalov PL, Mateo PL, Khechinashvili NN, et al. Comparative thermodynamic study of pepsinogen and
49. Smirnova MG, Guo L, Birchall JP, et al. LPS up-regulates mucin and cytokine mRNA expression and stimulates...


doi: 10.21037/aoe-20-95