Considerations for Small Peptide Analyses in SGF

Scott McClain, Ph.D.        Syngenta Crop Protection, LLC

Copenhagen, 18 October 2018
Interpreting the Putative Impact of Small Peptides

It seems it remains an open question of how to attempt an assessment on small peptides...particularly in the case of considering a wet chemistry methods approach

**Point of discussion** – it is assumed that in the guidance, the reference to 9 amino acids is specific to celiac-related peptides, but remains unclear whether Type I allergens are also being referred.

**Question** – for the purpose of novel protein risk assessment – when is it appropriate to apply small peptide fragment analysis, as part of pepsin (SGF) analysis?

- From Naegeli H., et al., 2017
  
  - if a protein digest is **composed of peptides < 9 amino acid** residues in length, the allergenic potential can be considered to be low;
  - if a protein digest is **composed of peptides ≥ 9 amino acid** residues in length but transient, the allergenic potential can be assumed to be low;
  - if digestion **fragments of ≥ 9 amino acids** or longer are identified and are persistent, further consideration is required.
  - In such a case, the **abundance** of the stable peptides ≥ 9 amino acids in length within the whole mixture of digestion products should be considered in the risk assessment process.
When is it appropriate to sample for, and identify small peptides in pepsin digestion assays (i.e., SGF) ?

- Any fragments from SGF have already been defined as not relevant for celiac if celiac bioinformatics is performed. The presence of fragments in SGF, in and of themselves, would not define a hazard...why?
- Pepsin, by defined action, creates variable fragments based on specific activity and a protein’s unique pepsin cleavage site and fragmentation pattern.

Risk is the consideration of both hazard (which protein, where is it from and its similarity with celiac peptides) AND Exposure – persistence plus exposure threshold (20 ppm for gluten)

Bioinformatics and Source organism/Gene HOSU identify potential hazard

No matches with celiac peptides

SGF – a focus on only the pepsin stability, as it relates to gut exposure

Putative matches with celiac peptides
Risk = Hazard X Exposure

Where can SGF be placed for best interpretation of results?

The question of stability is about exposure, as defined by pepsin digestion and downstream contact in the gut.

By definition, pepsin creates small peptides from most proteins.

Exposure itself is not a hazard, as pepsin digests most proteins the same...

It is only because we know about peptide size for celiac proteins that makes a very specific size (9-mer) a part of the well defined intracellular mode-of-action...

But, there is no mode of action for normal food proteins, because they are not a hazard.

\[ \text{Hazard} \times \text{Exposure} \]

\[ \sim 3 \text{kDA} = 33 \text{ mer} \]

How much does size matter here?

\[ \sim 1 \text{kDA} = 9 \text{ mer} \]
Broad Profiling - Pre-conceived SGF “outcomes” regarding fragment risk provides the lowest potential to identify important peptides.
Thinking about *Context*

- Analyses of peptides taken out of **exposure context** is not very useful for predicting a hazard...

- How do you characterize a peptide hazard by wet chemistry, unless you already know you have a gluten protein OR, you have a perfect 9mer match to known celiac peptides?
  - Peptide matches to celiac peptides can be known without SGF
  - Pepsin reduction of protein to putative peptides based on cleavage sites can be known without SGF
    - Do you need a Mass Spec assay to know this?
SGF – A complex matrix – the Novel Protein is not the only putative target

● Three main ingredients

- Active Ingredient
- Pepsin
- E. coli proteins

[Images of gel electrophoresis results with molecular weights and time points]
Focus on relevant fragments of New Protein provides best knowledge and least chance of false positives or inability to detect

Relevant fragments assessed just for exposure – degree of persistence assessed – how much?

FSQQQQSPF

Irrelevant fragments ignored

Fragments < 9AA

LPRRSEDKGRDEHGPG

Fragments ≥ 9AA

SGF as a digestion assay becomes informative when used in conjunction with good knowledge of any peptide hazards – fragment size itself predicts nothing

With focus on important fragments in only the New Protein, false positives are low
Technical Challenges – Mass Spec for very small peptides

● There are many technical challenges;
  - Not least of which – if the target peptide is not defined, then very difficult to measure in complex matrix – Novel protein not the only source in SGF

● Example:
  - Not all peptide that are as short as 9mer would even ionize properly for mass spec detection, much less for a quantitative measure
  - i.e., you would miss neutral peptides...in other words, new or different peptides that are different from the known celiac peptides would not be detected. (every purification method for NEP results in different E. coli profile).
  - Again, the size itself is not known to be determinant, other than for known celiac peptides.
    • If size was determinant, there would be far more hazardous celiac proteins than just those from wheat, barley, rye and oat.
Abundance

● Challenge question - Even if you had knowledge of abundance, what could it tell you about a protein risk with regard to celiac disease, if?
  - Protein is not from gluten containing organism
  - Is not otherwise similar to glutens
  - There is no specific “Dosing threshold” or measurable “mass balance” of peptides known for those that are DQ specific celiac peptides.

● Thought experiment – you can have knowledge of putative peptides from a protein prior to putting it in SGF?
  - Premise: we know exactly how much Novel Protein is loaded into each well of an SGF assay...so then, how much theoretical peptide (µ moles) of a peptide would be significant?
  - So then, would you measure small peptides after pepsin treatment, or just predict any matches with celiac peptides first, and then only if there were a match measure them after treatment with pepsin?
Exposure Scenario to Discuss in Panel Discussion

Consideration - 20 ppm threshold for gluten sensitivity – Food

- A threshold for a molecule that produces many peptides (i.e., a greater mass molar amount of celiac peptides than a single peptide)
- In contrast, The premise is that only one peptide match, not many as in a gluten, is triggering a concern in even a naturally occurring protein, like zein in maize.

- **Question** – if a GMO or other new-use food protein has the following condition...
  - Expressed in commodity food or processed food fraction @ less than 20 ppm

- **Then, would Celiac concerns would be negligible?**
  - What would be the practical, science-based position on this?
  - Is neither, celiac specific bioinformatics or SGF analyses of peptides required
    - One, but not the other?

- If a 9mer is a direct agent of celiac disease in the gut without processing, is there data to confirm this?
  - in other words, can a human consume a 9mer peptide and does it act in vivo the same as a processed gluten protein?
    - You would ask this question to determine the usefulness of SGF for celiac risk assessment
Thank You