

- (ii) identifying an antigenic epitope by identifying a surface-exposed epitope among the at least one surface-exposed peptide that is present in a region of the protein that results in a lack of, or significantly altered, biological function of the protein when the peptide is cleaved off or removed from the protein during the limited or restricted proteolysis; or
- selecting at least one target region within the protein based on bioinformatics and/or known data of biological function of the protein and identifying an antigenic epitope by identifying a surface-exposed epitope among the at least one surface-exposed peptide present in said at least one target region.

10 In one aspect, the present disclosure refers to a method of identifying an epitope on a native protein that can be bound by an antibody, said method comprising:

- (i) identifying sites at which one or more protease has cut said native protein after said native protein has been exposed to limited or restricted proteolysis by said one or more protease; and
- 15 (ii) probing a plurality of epitopes on said native protein that overlap with a cut site or that are in a region that flanks a cut site with antibodies directed to said epitopes, thereby identifying one or more epitopes that can be bound by an antibody, wherein said plurality of epitopes probed in step (ii) are, or at least a portion of any given epitope in said plurality of epitopes probed in step (ii) is, within 50 amino acids of said cut site, and wherein said
- 20 method further comprises a step prior to step (ii) of generating a plurality of isolated epitopes having sequences that correspond to epitopes on said native protein that overlap with a cut site or that are in a region that flanks a cut site, and generating antibodies that bind to said isolated epitopes, and using said antibodies in step (ii) for probing said plurality of epitopes on said native protein.

25 In another aspect, the present disclosure refers to a method of identifying an epitope on a native protein that can be bound by an antibody, said method comprising:

- (i) performing *in silico* protease digestion of a protein with one or more proteases to identify sites on the protein that are predicted to be cut by said one or more proteases and
- 30 a. performing protein homology modelling to predict which *in silico* predicted protease cut sites are likely to be exposed in the native protein, and/or
- b. performing *in silico* docking of antibody fragments or said one or more proteases to predict which *in silico* predicted cut sites are likely to be cut *in vitro* in the native protein;
- (ii) performing *in vitro* limited or restricted proteolysis of said native protein with one or
- 35 more proteases;
- (iii) identifying peptides released from said native protein by the *in vitro* protease digestion of step (ii), thereby identifying cut sites;

- (iv) comparing the *in silico* predicted cut sites identified that are predicted to be exposed or predicted to be cut *in vitro* in step (i) with cut sites identified in step (iii);
- (v) probing one or more epitopes in a region of the native protein containing or flanking a cut site that is an *in silico* predicted protease cut site identified that is predicted to be exposed or predicted to be cut *in vitro* in step (i) but that is not a cut site identified in step (iii) with one or more antibodies; and
- (vi) identifying whether or not said one or more antibodies bind to said one or more epitopes thereby identifying an epitope on a native protein that can be bound by an antibody, wherein said one or more epitopes probed in step (v) are, or at least a portion of any given epitope probed in step (v) is, within 50 amino acids of said cut site, and wherein said method further comprises a step prior to step (v) of generating one or more isolated epitopes having sequences that correspond to one or more epitopes on said native protein that are in a region of the protein containing or flanking a cut site that is an *in silico* predicted protease cut site identified that is predicted to be exposed or predicted to be cut *in vitro* in step (i) but that is not a cut site identified in step (iii), and generating antibodies that bind to said isolated epitopes, and using said antibodies in step (v) for probing said one or more epitopes of said native protein.

The present disclosure relates to methods of detecting and identifying amino acid sequences in proteins where said amino acid sequences are well-exposed, and functionally relevant, at least they are well-exposed. These amino acid sequences which we refer to as "hot spots", thus, may be utilized as antigenic epitopes that guides antibody targeting, discovery, and development. Furthermore, these amino acid sequences can be ranked based on their appearance after proteolytic digestion, and based on functional relevance from already known bioinformatic data or from functional/pharmacological testing. Thus, from a list of several amino acid sequences resulting from a proteolytic digestion, the best suited amino acid sequences (based on functional and structural arguments) can be picked for antigenic epitope discovery and development. The proteolytic digestion is performed under limiting conditions, i.e. the activity of the protease or several proteases is very low such that just one or a few surface-exposed peptides are cleaved off from the target protein at a time. The proteases are thus used as druggability probes for antibody binding to a target protein. In an example, the antibodies are pharmacologically active. In another example, the antibodies are pharmacologically active and developed for therapeutic usage. More specifically, such methods include proteomic tools to reveal hot spot epitopes of target proteins. In an example of the disclosure, a protein is digested, deconstructed and/or truncated through protease action and all well-exposed amino acid sequences are used for antigenic epitope generation, and antibodies developed based on said antigenic epitopes are tried for potency, efficacy, pharmacological profiling, and other testing as customary in antibody discovery used in the pharmaceutical industry. In an example of the disclosure, a protein is digested, deconstructed and/or truncated through protease action and in parallel probed by a functional assay on the digested, deconstructed

CLAIMS

1. A method of identifying an epitope on a native protein that can be bound by an antibody, said method comprising:
 - 5 (i) identifying sites at which one or more protease has cut said native protein after said native protein has been exposed to limited or restricted proteolysis by said one or more protease; and
 - (ii) probing a plurality of epitopes on said native protein that overlap with a cut site or that are in a region that flanks a cut site with antibodies directed to said epitopes, thereby identifying one or more epitopes that can be bound by an antibody,

10 wherein said plurality of epitopes probed in step (ii) are, or at least a portion of any given epitope in said plurality of epitopes probed in step (ii) is, within 50 amino acids of said cut site, and

15 wherein said method further comprises a step prior to step (ii) of generating a plurality of isolated epitopes having sequences that correspond to epitopes on said native protein that overlap with a cut site or that are in a region that flanks a cut site, and generating antibodies that bind to said isolated epitopes, and using said antibodies in step (ii) for probing said plurality of epitopes on said native protein.

20
2. The method of claim 1, wherein a single protease is used.
3. The method of claim 1, wherein multiple proteases are used.
- 25 4. The method of any one of claims 1 to 3, wherein said protease is selected from the group consisting of: trypsin, Arg-C proteinase, Asp-N endopeptidase, Clostripain, Glutamyl endopeptidase, Lys-C, Lys-N, Chymotrypsin, Proteinase K, Thermolysin, Pepsin, Caspase 1, Caspase 2, Caspase 3, Caspase 4, Caspase 5, Caspase 6, Caspase 7, Caspase 8, Caspase 9, Caspase 10, Enterokinase, Factor Xa,

30 GranzymeB, Neutrophil elastase, Proline-endopeptidase, Staphylococcal peptidase I, and Thrombin.
5. The method of any one of claims 1 to 4, wherein said protease is selected from the group consisting of: trypsin, Asp-N endopeptidase, Chymotrypsin, pepsin and

35 Proteinase K.

6. The method of any one of claims 1 to 5, wherein said protein is a membrane protein that is present in a proteoliposome derived from cells.
7. The method of claim 6, wherein said proteoliposome is immobilized in a flow cell to create a stationary phase of membrane proteins.
8. The method of any one of claims 1 to 7, wherein said sites at which one or more protease has cut said protein are identified using mass spectrometry.
9. The method of any one of claims 1 to 8, wherein said plurality of epitopes probed in step (ii) are, or at least a portion of any given epitope in said plurality of epitopes probed in step (ii) is, within 20 amino acids of said cut site.
10. The method of any one of claims 1 to 9, wherein said plurality of epitopes is a set of epitopes wherein the sequence of each epitope in the set is offset from another epitope in the set by 1, 2 or 3 amino acids.
11. The method of any one of claims 1 to 10, wherein said method further comprises a step of producing an antibody against an epitope identified by any one of claims 1 to 10.
12. A method of identifying an epitope on a native protein that can be bound by an antibody, said method comprising:
 - (i) performing *in silico* protease digestion of a protein with one or more proteases to identify sites on the protein that are predicted to be cut by said one or more proteases and
 - a. performing protein homology modelling to predict which *in silico* predicted protease cut sites are likely to be exposed in the native protein, and/or
 - b. performing *in silico* docking of antibody fragments or said one or more proteases to predict which *in silico* predicted cut sites are likely to be cut *in vitro* in the native protein;
 - (ii) performing *in vitro* limited or restricted proteolysis of said native protein with one or more proteases;

- (iii) identifying peptides released from said native protein by the *in vitro* protease digestion of step (ii), thereby identifying cut sites;
 - (iv) comparing the *in silico* predicted cut sites identified that are predicted to be exposed or predicted to be cut *in vitro* in step (i) with cut sites identified in step (iii);
 - 5 (v) probing one or more epitopes in a region of the native protein containing or flanking a cut site that is an *in silico* predicted protease cut site identified that is predicted to be exposed or predicted to be cut *in vitro* in step (i) but that is not a cut site identified in step (iii) with one or more antibodies; and
 - (vi) identifying whether or not said one or more antibodies bind to said one or more
10 epitopes thereby identifying an epitope on a native protein that can be bound by an antibody,
- wherein said one or more epitopes probed in step (v) are, or at least a portion of any given epitope probed in step (v) is, within 50 amino acids of said cut site, and
- wherein said method further comprises a step prior to step (v) of generating
15 one or more isolated epitopes having sequences that correspond to one or more epitopes on said native protein that are in a region of the protein containing or flanking a cut site that is an *in silico* predicted protease cut site identified that is predicted to be exposed or predicted to be cut *in vitro* in step (i) but that is not a cut site identified in step (iii), and generating antibodies that bind to said isolated epitopes, and using said
20 antibodies in step (v) for probing said one or more epitopes of said native protein.
13. The method of claim 12, wherein a single protease is used.
14. The method of claim 12, wherein multiple proteases are used.
- 25 15. The method of any one of claims 12 to 14, wherein said protease is selected from the group consisting of: trypsin, Arg-C proteinase, Asp-N endopeptidase, Clostripain, Glutamyl endopeptidase, Lys-C, Lys-N, Chymotrypsin, Proteinase K, Thermolysin, Pepsin, Caspase 1, Caspase 2, Caspase 3, Caspase 4, Caspase 5, Caspase 6,
30 Caspase 7, Caspase 8, Caspase 9, Caspase 10, Enterokinase, Factor Xa, GranzymeB, Neutrophil elastase, Proline-endopeptidase, Staphylococcal peptidase I, and Thrombin.
16. The method of any one of claims 12 to 15, wherein said protease is selected from the
35 group consisting of: trypsin, Asp-N endopeptidase, Chymotrypsin, pepsin and Proteinase K.

17. The method of any one of claims 12 to 16, wherein said protein is a membrane protein that is present in a proteoliposome derived from cells.
18. The method of claim 17, wherein said proteoliposome is immobilized in a flow cell to
5 create a stationary phase of membrane proteins.
19. The method of any one of claims 12 to 18, wherein identifying peptides released from said protein by the *in vitro* protease digestion of step (ii), thereby identifying cut sites, is done by mass spectrometry.
10
20. The method of any one of claims 12 to 19, wherein a plurality of epitopes is probed.
21. The method of any one of claims 12 to 20, wherein said epitope is within 20 amino acids of said cut site that is an *in silico* predicted protease cut site identified that is
15 predicted to be exposed or predicted to be cut *in vitro* in step (i) but that is not a cut site identified in step (iii).
22. The method of any one of claims 12 to 21, wherein a plurality of epitopes is probed and said plurality of epitopes is a set of epitopes wherein the sequence of each
20 epitope in the set is offset from another epitope in the set by 1, 2 or 3 amino acids.
23. The method of any one of claims 12 to 22, wherein said method further comprises a step of producing an antibody against an epitope identified by any one of claims 12 to 22.