

art, a peptide, or a fragment of a larger polypeptide, is identified as immunogenic for a specific human subject if it is presented by a HLA molecule that is expressed by the subject. In other words, the state of the art describes immunogenic peptides as HLA-restricted epitopes. However, HLA restricted epitopes induce T cell responses in only a fraction of individuals who express the HLA molecule. Peptides that activate a T cell response in one individual are inactive in others despite HLA allele matching. Therefore, it was unknown how an individual's HLA molecules present the antigen-derived epitopes that positively activate T cell responses.

As provided herein multiple HLA expressed by an individual need to present the same peptide in order to trigger a T cell response. Therefore the fragments of a polypeptide antigen that are immunogenic for a specific individual are those that can bind to multiple class I (activate cytotoxic T cells) or class II (activate helper T cells) HLAs expressed by that individual.

In one aspect, the disclosure refers to a method of identifying a fragment of a polypeptide as immunogenic for a specific human subject or predicting whether a polypeptide or a fragment of a polypeptide is immunogenic for a specific human subject, the method comprising the steps of

(i) determining whether the polypeptide comprises:

(a) an amino acid sequence that is a T cell epitope capable of binding to at least two HLA class I molecules of the subject; or

(b) an amino acid sequence that is a T cell epitope capable of binding to at least two HLA class II molecules of the subject; and

(ii) (I) identifying said sequence as a fragment of the polypeptide that is immunogenic for the subject; or

(II) predicting

A. that the polypeptide is immunogenic for the subject if the polypeptide comprises at least one sequence that meets the requirements of step (i); or

B. that the polypeptide is not immunogenic for the subject if the polypeptide does not comprise at least one sequence that meets the requirements of step (i).

In another aspect, the disclosure refers to the use of a checkpoint inhibitor in the preparation of a medicament for treating cancer in a human subject in need thereof, wherein the

subject has been predicted to respond, or to be likely to respond, to administration of the checkpoint inhibitor by a method as described herein.

In another aspect, the disclosure refers to an *in vitro* or *in silico* method of designing or preparing a human subject-specific pharmaceutical composition or kit or panel of polypeptides for treatment of a specific human subject, the method comprising:

(i) selecting a fragment of a polypeptide, which fragment has been identified as immunogenic for the subject by the method of any one of claims 1 to 3;

(ii) if the fragment selected in step (i) is an HLA class I-binding epitope, optionally selecting a longer fragment of the polypeptide, which longer fragment

a. comprises the fragment selected in step (i); and

b. is a T cell epitope capable of binding at least three or to the most possible HLA class II molecules of the subject;

(iii) selecting a first sequence of up to 50 consecutive amino acids of the polypeptide, which consecutive amino acids comprise the amino acid sequence of the fragment

selected in step (i) or the longer fragment selected in step (ii);

(iv) repeating steps (i) to (iii) to select a second amino acid sequence of up to 50 consecutive amino acids of the same or a different polypeptide to the first amino acid sequence;

(v) optionally further repeating steps (i) to (iii) to select one or more additional amino

acid sequences of up to 50 consecutive amino acids of the same or different polypeptides to the first and second amino acid sequences; and

(vi) designing or preparing a subject-specific pharmaceutical composition, kit or panel of polypeptides having as active ingredients one or more polypeptides that together have all of the amino acid sequences selected in the preceding steps, optionally wherein one or more or each sequence is flanked at the N and/or C terminus by additional amino acids that are not part of the sequence of the polypeptides.

In yet another aspect, the disclosure refers to a human subject-specific pharmaceutical composition, kit or panel of polypeptides for use in therapy for inducing an immune response in a specific human subject, and designed or prepared for the subject according to the method as

described herein, wherein the composition or kit optionally comprises at least one pharmaceutically acceptable diluent, carrier, or preservative, optionally wherein the therapy is prophylactic.

In yet another aspect, the disclosure refers to a human subject-specific composition, kit or panel of polypeptides for use in therapy of a specific human subject in need thereof, the composition, kit or panel comprising as active ingredients

(a) a first and a second peptide and optionally one of more additional peptides, wherein each peptide comprises an amino acid sequence that is a T cell epitope capable of binding to at least two HLA class I molecules and/or at least two HLA class II molecules of the subject, wherein the amino acid sequence of the T cell epitope of the first, second and optionally any additional peptides are different from each other, and wherein the pharmaceutical composition or kit optionally comprises at least one pharmaceutically acceptable diluent, carrier, or preservative;

(b) a first region and a second region and optionally one of more additional regions, wherein each region comprises an amino acid sequence that is a T cell epitope capable of binding to at least two HLA class I molecules and/or at least two HLA class II molecules of the subject, wherein the amino acid sequence of the T cell epitope of the first, second and optionally any additional regions are different from each other, and wherein the pharmaceutical composition or kit optionally comprises at least one pharmaceutically acceptable diluent, carrier, or preservative; optionally wherein the therapy is prophylactic.

In yet another aspect, the disclosure refers to a method of designing or preparing a polypeptide for inducing an immune response in a specific human subject the method comprising selecting an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules or at least three HLA class II molecules of the subject, and designing or preparing a polypeptide comprising the selected amino acid sequence; optionally wherein the method is (a) a method of designing or preparing a polypeptide for inducing a cytotoxic T cell response in a specific human subject, the method comprising selecting an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject, and designing or preparing a polypeptide comprising the selected amino acid sequence; or

(b) a method of designing or preparing a polypeptide for inducing a helper T cell response, the method comprising selecting an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class II molecules of the subject, and designing or preparing a polypeptide comprising the selected amino acid sequence.

5 In yet another aspect, the disclosure refers to a system comprising

(a) a storage module configured to store data comprising the class I and/or class II HLA genotype of a subject and the amino acid sequence of one or more test polypeptides; and

(b) a computation module configured to identify and/or quantify amino acid sequences in the one or more test polypeptides that are capable of binding to multiple HLA class I molecules of the subject and/or that are are capable of binding to multiple HLA class II molecules of the
10 subject; and optionally further comprising

(c) an output module configured to display

(i) a prediction of whether the one or more polypeptides is immunogenic for the subject; or the sequence of one or more fragments of the one or more polypeptides that are
15 predicted to be immunogenic for the subject;

(ii) a prediction of whether the individual will have an immune response to administration of the one or more polypeptides or one or more pharmaceutical compositions comprising the one or more polypeptides as active ingredients;

(iii) a prediction of whether the subject will have a clinical response to a
20 method of treatment comprising administering to the subject one or more pharmaceutical compositions comprising the one or more polypeptides as active ingredients;

(iv) the likelihood that the subject will have a clinical response to administration of one or more pharmaceutical compositions comprising the one or more polypeptides as active ingredients;

25 (v) a prediction of whether administration of the one or more polypeptides or one or more pharmaceutical compositions comprising the one or more polypeptides will induce a toxic immune response in the subject;

(vi) a prediction that the one or more polypeptides is associated with an autoimmune disorder in the subject;

(vii) a prediction of whether the subject will have a clinical response to administration of a checkpoint inhibitor; or

(viii) a recommendation of whether or not the subject should be treated by administration of the one or more polypeptides and/or one or more pharmaceutical compositions.

5 Accordingly, in a first example the disclosure provides methods of predicting whether a polypeptide or a fragment of a polypeptide is immunogenic for a specific human subject, the methods comprising the steps of

(i) determining whether the polypeptide comprises:

(a) an amino acid sequence that is a T cell epitope capable of binding to at least two
10 HLA class I molecules of the subject; or

(b) an amino acid sequence that is a T cell epitope capable of binding to at least two HLA class II molecules of the subject; and

(ii) predicting

A. that the polypeptide is immunogenic for the subject if the polypeptide comprises at
15 least one sequence that meets the requirements of step (i); or

B. that the polypeptide is not immunogenic for the subject if the polypeptide does not comprise at least one sequence that meets the requirements of step (i).

The disclosure also provides methods of identifying a fragment of a polypeptide as immunogenic for a specific human subject, the methods comprising the steps of

20 (i) determining that the polypeptide comprises:

(a) an amino acid sequence that is a T cell epitope capable of binding to at least two HLA class I molecules of the subject; or

(b) an amino acid sequence that is a T cell epitope capable of binding to at least two HLA class II molecules of the subject; and

25 (ii) identifying said sequence as a fragment of the polypeptide that is immunogenic for the subject.

In some embodiments the methods of the disclosure comprise the step of determining or obtaining the HLA class I genotype and/or the HLA class II genotype of the specific human subject.

A specific polypeptide antigen may comprise more than one fragment that is a T cell epitope capable of binding to multiple HLA of a specific individual. The combined group of all such fragments characterize the individual's antigen specific T cell response set, wherein the amino acid sequence of each fragment characterizes the specificity of each activated T cell clone.

5 Accordingly in some cases the method is repeated until all of the fragments of the polypeptide that are a T cell epitope capable of binding to at least two HLA class I and/or at least two HLA class II of the subject have been identified. This method characterises the immune response of the subject to the polypeptide.

10 The disclosure further provides methods of treatment of a human subject in need thereof, the method comprising administering to the subject a polypeptide, pharmaceutical composition or kit of the polypeptides of a panel of polypeptides that has been identified or selected by any of the methods above or comprising a fragment of a polypeptide that has been identified or selected by any of the methods above; their use in a method of treatment of a relevant human subject; and their use in the manufacture of a medicament for treating a relevant subject.

15 The fragments of polypeptide that are determined to be immunogenic for a specific human subject in accordance with the methods above can be used to prepare human subject-specific immunogenic compositions.

20 Accordingly in a further example, the disclosure provides methods of designing or preparing a human subject-specific pharmaceutical composition or kit or panel of polypeptides for use in a method of treatment of a specific human subject, the methods comprising:

- (i) selecting a fragment of a polypeptide, which fragment has been identified as immunogenic for the subject by the method above;
- (ii) if the fragment selected in step (i) is an HLA class I-binding epitope, optionally selecting a longer fragment of the polypeptide, which longer fragment
 - 25 a. comprises the fragment selected in step (i); and

- b. is a T cell epitope capable of binding at least three or to the most possible HLA class II molecules of the subject;
- (iii) selecting a first sequence of up to 50 consecutive amino acids of the polypeptide, which consecutive amino acids comprise the amino acid sequence of the fragment selected in step (i) or the longer fragment selected in step (ii);
- (iv) repeating steps (i) to (iii) to select a second amino acid sequence of up to 50 consecutive amino acids of the same or a different polypeptide to the first amino acid sequence;
- (v) optionally further repeating steps (i) to (iii) to select one or more additional amino acid sequences of up to 50 consecutive amino acids of the same or different polypeptides to the first and second amino acid sequences; and
- (vi) designing or preparing a subject-specific pharmaceutical composition, kit or panel of polypeptides having as active ingredients one or more polypeptides that together have all of the amino acid sequences selected in the preceding steps, optionally wherein one or more or each sequence is flanked at the N and/or C terminus by additional amino acids that are not part of the sequence of the polypeptides.

In some cases each peptide either consists of one of the selected amino acid sequences, or consists of two or more of the amino acid sequences arranged end to end or overlapping in a single peptide.

The disclosure further provides a human subject-specific pharmaceutical composition, kit or panel of polypeptides for use in a method of treatment of a specific human subject in need thereof, the composition, kit or panel comprising as active ingredients a first and a second peptide and optionally one of more additional peptides, wherein each peptide comprises an amino acid sequence that is a T cell epitope capable of binding to at least two HLA class I molecules and/or at least two HLA class II molecules of the subject, wherein the amino acid sequence of the T cell epitope of the first, second and optionally any additional peptides are different from each other, and wherein the pharmaceutical composition or kit optionally comprises at least one pharmaceutically acceptable diluent, carrier, or preservative.

The disclosure further provides a human subject-specific pharmaceutical composition, kit or panel of polypeptides for use in a method of treatment of a specific human subject in need thereof, the composition or kit comprising as an active ingredient a polypeptide comprising a first region and a second region and optionally one of more additional regions, wherein each region comprises an amino acid sequence that is a T cell epitope capable of binding to at least two HLA class I molecules and/or at least two HLA class II molecules of the subject, wherein the amino acid sequence of the T cell epitope of the first, second and optionally any additional regions are different from each other, and wherein the pharmaceutical composition or kit optionally comprises at least one pharmaceutically acceptable diluent, carrier, or preservative.

The disclosure further provides a method of designing or preparing a polypeptide for inducing an immune response in a specific human subject the method comprising selecting an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules or at least three HLA class II molecules of the subject, and designing or preparing a polypeptide comprising the selected amino acid sequence.

In further examples, the disclosure provides

- a method of inducing an immune response or a method of treatment comprising administering to a human subject in need thereof a human subject-specific pharmaceutical composition, or the polypeptides a kit or panel as described above, wherein the composition, kit or panel of polypeptides is specific for the subject;
- a human subject-specific immunogenic composition, kit or panel as described above for use in a method of inducing an immune response or a method of treatment of the specific human subject; and
- use of a human subject-specific pharmaceutical composition or the polypeptides of a kit or panel as described above in the manufacture of a medicament, wherein the medicament is for inducing an immune response in or treating the specific subject.

In a further example the disclosure provides a system comprising
(a) a storage module configured to store data comprising the class I and/or class II HLA genotype of a subject and the amino acid sequence of one or more test polypeptides; and

(b) a computation module configured to identify and/or quantify amino acid sequences in the one or more test polypeptides that are capable of binding to multiple HLA class I molecules of the subject and/or that are are capable of binding to multiple HLA class II molecules of the subject.

The disclosure provides a method of treatment of a human subject in need thereof, the method comprising administering to the subject a polypeptide, a panel of polypeptides, a pharmaceutical composition or the active ingredient polypeptides of a kit described above, wherein the subject has been determined to express at least three HLA class I molecules and/or at least three HLA class II molecules capable of binding to the polypeptide or to one or more of the active ingredient polypeptides of the pharmaceutical composition or kit.

The disclosure will now be described in more detail, by way of example and not limitation, and by reference to the accompanying drawings. Many equivalent modifications and variations will be apparent, to those skilled in the art when given this disclosure. Accordingly, the exemplary embodiments of the disclosure set forth are considered to be illustrative and not limiting. Various changes to the described embodiments may be made without departing from the scope of the disclosure. All documents cited herein, whether supra or infra, are expressly incorporated by reference in their entirety.

The present disclosure includes the combination of the examples and preferred features described except where such a combination is clearly impermissible or is stated to be expressly avoided. As used in this specification and the appended claims, the singular forms “a”, “an”, and “the” include plural referents unless the content clearly dictates otherwise. Thus, for example, reference to “a peptide” includes two or more such peptides.

Section headings are used herein for convenience only and are not to be construed as limiting in any way.

Description of the Figures

Fig. 1

ROC curve of HLA restricted PEPI biomarkers.

Fig. 2

ROC curve of ≥ 1 PEPI3+ Test for the determination of the diagnostic accuracy.

Fig. 3

Distribution of HLA class I PEPI3+ compared to CD8+ T cell responses measured by a state of art assay among peptide pools used in the CD8+ T cell response assays. A: HLA class I restricted PEPI3+s. The 90% Overall Percent of Agreement (OPA) among the T cell responses and PEPI3+ peptides demonstrate the utility of the disclosed peptides for prediction of vaccine induced T cell response set of individuals. B: Class I HLA restricted epitopes (PEPI1+). The OPA between predicted epitopes and CD8+ T cell responses was 28% (not statistically significant). Darkest grey: True positive (TP), both peptide and T cell responses were detected; Light grey: False negative (FN), only T cell responses were detected; Lightest grey: False positive (FP), only peptide were detected; Dark grey: True negative (TN): neither peptides nor T cell responses were detected.

Fig. 4

Distribution of HLA class II PEPIs compared to CD4+ T cell responses measured by a state of art assay among peptide pools used in the assays. A: HLA class II restricted PEPI4+s. 67% OPA between PEPI4+ and CD4+ T-cell responses ($p=0.002$). B: The class II HLA restricted epitopes. OPA between class II HLA restricted epitopes and CD4+ T cell responses was 66% (not statistically significant). Darkest grey: True positive (TP), both peptide and T cell responses were detected; Light grey: False negative (FN), only T cell responses were detected; Lightest grey:

CLAIMS

1. A method of identifying a fragment of a polypeptide as immunogenic for a specific human subject or predicting whether a polypeptide or a fragment of a polypeptide is immunogenic for a specific human subject, the method comprising the steps of
 - (i) determining whether the polypeptide comprises:
 - (a) an amino acid sequence that is a T cell epitope capable of binding to at least two HLA class I molecules of the subject; or
 - (b) an amino acid sequence that is a T cell epitope capable of binding to at least two HLA class II molecules of the subject; and
 - (ii) (I) identifying said sequence as a fragment of the polypeptide that is immunogenic for the subject; or
 - (II) predicting
 - A. that the polypeptide is immunogenic for the subject if the polypeptide comprises at least one sequence that meets the requirements of step (i); or
 - B. that the polypeptide is not immunogenic for the subject if the polypeptide does not comprise at least one sequence that meets the requirements of step (i)
2. The method of claim 1, wherein:
 - (a) the T cell epitope is capable of binding to at least two HLA class I molecules of the subject and consists of 9 consecutive amino acids of the polypeptide, or wherein the T cell epitope is capable of binding to at least two HLA class II molecules of the subject and consists of 15 consecutive amino acids of the polypeptide;
 - (b) step (i) comprises determining that the polypeptide comprises an amino acid sequence that is a T cell epitope capable of binding to at least two HLA class I molecules of the subject;
 - (c) step (i) comprises determining that the polypeptide comprises an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject;
 - (d) step (i) comprises determining that the polypeptide comprises an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class II molecules of the subject; and/or
 - (e) step (i) comprises determining that the polypeptide comprises an amino acid sequence that is a T cell epitope capable of binding to at least two HLA class I molecules of the subject

and identifying a fragment of the polypeptide that is a T cell epitope capable of binding to at least one HLA class II molecule of the subject, wherein the HLA class II-binding epitope comprises the amino acid sequence of the HLA class I-binding T cell epitope.

3. The method of claim 1 or claim 2, wherein the polypeptide

(a) is expressed by a pathogenic organism, a virus or a cancer cell, is associated with an autoimmune disorder, or is an allergen or an ingredient of a pharmaceutical composition;

(b) is selected from the antigens listed in Tables 2 to 6;

(c) is an antigen or neoantigen expressed by a cancer cell, optionally wherein the cancer cell, the antigen or the neoantigen is in a sample taken from the subject;

(d) is a mutational neoantigen, optionally wherein

A. the neoantigen is present in a sample obtained from the subject; and/or

B. the immunogenic fragment comprises a neoantigen specific mutation; or

(e) is associated with or suspected of being associated with an autoimmune disorder or an autoimmune response in the subject, wherein determining that the polypeptide comprises an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject identifies the polypeptide and/or the fragment as immunogenic or associated with the autoimmune disorder or autoimmune response in the subject.

4. The method of any one of claims 1 to 3, wherein all of the fragments of the polypeptide that are a T cell epitope capable of binding to at least two HLA class I molecules and/or all of the fragments of the polypeptide that are a T cell epitope capable of binding to at least two HLA class II molecules of the subject are identified, optionally wherein the method is repeated for each polypeptide that is an active ingredient of a specific pharmaceutical composition.

5. The method of any one of the preceding claims, further comprising predicting whether the subject will have a cytotoxic T cell response or a helper T cell response to administration of one or more polypeptide or a pharmaceutical composition or kit comprising one or more polypeptides as active ingredients, wherein

A. a cytotoxic T cell response is predicted if the polypeptide(s) comprises at least one amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject;

B. a helper T cell response is predicted if the polypeptide(s) comprises at least one amino acid sequence that is a T cell epitope capable of binding to at least three HLA class II molecules of the subject;

C. no cytotoxic T cell response is predicted if the polypeptide(s) does not comprise any amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject; or

D. no helper T cell response is predicted if the polypeptide(s) does not comprise any amino acid sequence that is a T cell epitope capable of binding to at least three HLA class II molecules of the subject,

optionally wherein

E. the method further comprises determine the likelihood that the subject will have a cytotoxic T cell response and/or a helper T cell response that targets a polypeptide antigen that is expressed in the subject, the method comprising

(i) identifying one or more polypeptide antigens that comprises an amino acid sequence that is

(a) a T cell epitope capable of binding to at least three HLA class I or at least three HLA class II molecules of the subject; and

(b) comprised in the amino acid sequence of the polypeptide(s)

(ii) using population expression frequency data for the one or more polypeptide antigens identified in step (i) to determine the likelihood that the subject will have a cytotoxic T cell response and/or a helper T cell response that targets a polypeptide antigen that is expressed in the subject; or

F. the polypeptide is a component of a pharmaceutical composition and the method comprises determining the likelihood that the subject will develop anti-drug antibodies (ADA) following administration of the polypeptide, wherein a predicted T helper cell response corresponds to a higher likelihood of ADA and no predicted T helper cell response corresponds to a lower likelihood of ADA, further optionally wherein the polypeptide is a checkpoint inhibitor.

6. The method of any one of claims 1 to 4 further comprising:

(I) predicting whether the subject will have a clinical response to administration of a pharmaceutical composition, kit or panel of polypeptides comprising one or more polypeptides as active ingredients, the method comprising determining whether the one or more active

ingredient polypeptides together comprise at least two different amino acid sequences each of which is a T cell epitope capable of binding to at least three HLA class I molecules of the subject; and predicting

A. that the subject will have a clinical response to administration of the pharmaceutical composition, kit or panel of polypeptides if the one or more active ingredient polypeptides together comprise at least two different sequences each of which is a T cell epitope capable of binding to at least three HLA class I molecules of the subject; or

B. that the subject will not have a clinical response to administration of the pharmaceutical composition, kit or panel of polypeptides if the one or more active ingredient polypeptides together comprise no more than one sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject, optionally wherein the at least two different amino acid sequences are comprised in the amino acid sequence of two different polypeptide antigens targeted by the active ingredient polypeptide(s); and/or

(II) determining the likelihood that the specific human subject will have a clinical response to administration of a pharmaceutical composition, kit or panel of polypeptides comprising one or more polypeptides as active ingredients, wherein

(1) one or more of the following factors corresponds to a higher likelihood of a clinical response:

(a) presence in the active ingredient polypeptide(s) of a higher number of amino acid sequences and/or different amino acid sequences that are each a T cell epitope capable of binding to at least three HLA class I of the subject;

(b) a higher number of target polypeptide antigens, comprising at least one amino acid sequence that is both

A. comprised in an active ingredient polypeptide; and

B. a T cell epitope capable of binding to at least three HLA class I of the subject; optionally wherein the target polypeptide antigens are expressed in the subject, further optionally wherein the target polypeptides antigens are in one or more samples obtained from the subject;

(c) a higher probability that the subject expresses target polypeptide antigens, optionally a threshold number of the target polypeptide antigens and/or optionally target polypeptide antigens that have been determined to comprise at least one amino acid sequence that is both

A. comprised in in an active ingredient polypeptide; and

B. a T cell epitope capable of binding to at least three HLA class I of the subject; and/or

(d) a higher number of target polypeptide antigens that the subject is predicted to express, optionally a higher number of target polypeptide antigens that the subject expresses with a threshold probability, and/or optionally the target polypeptide antigens that have been determined to comprise at least one amino acid sequence that is both

A. comprised in an active ingredient polypeptide; and

B. a T cell epitope capable of binding to at least three HLA class I of the subject; and/or

(2) the method comprises

(i) identifying which polypeptide antigens targeted by the active ingredient polypeptide(s) comprise an amino acid sequence that is both

A. comprised in an active ingredient polypeptide; and

B. a T cell epitope capable of binding to at least three HLA class I of the subject;

(ii) using population expression data for each antigen identified in step (i) to determine the probability that the subject expresses one or more of the antigens identified in step (i) that together comprise at least two different amino acid sequences of step (i), optionally using population expression data for each antigen identified in step (i) to determine the probability that the subject expresses two or more of the antigens identified in step (i) that together comprise at least two different amino acid sequences of step (i), further optionally wherein the at least two different amino acid sequences are comprised in the amino acid sequence of two different polypeptide antigens targeted by the active ingredient polypeptide(s); and

(iii) determining the likelihood that the subject will have a clinical response to administration of the pharmaceutical composition, kit or panel of polypeptides, wherein a higher probability determined in step (ii) corresponds to a more likely clinical response; and/or

(3) one or more of the following factors correspond to a higher likelihood of a clinical response:

(a) presence in the active ingredient polypeptide(s) of a higher number of amino acid sequences and/or different amino acid sequences that are each a T cell epitope capable of binding to at least three HLA class II of the subject;

(b) a higher number of target polypeptide antigens comprising at least one amino acid sequence that is both

- A. comprised in an active ingredient polypeptide; and
 - B. a T cell epitope-capable of binding to at least three HLA class II of the subject, optionally wherein the target polypeptide antigens are expressed in the subject, optionally wherein the target polypeptides antigens are in one or more samples obtained from the subject;
- (c) a higher number of target polypeptide antigens comprising
 - i. at least one amino acid sequence that is both
 - A. comprised in an active ingredient polypeptide; and
 - B. a T cell epitope capable of binding to at least three HLA class I of the subject; and
 - ii. at least one amino acid sequence that is both
 - A. comprised in an active ingredient polypeptide; and
 - B. a T cell epitope capable of binding to at least three HLA class II of the subject;
- (d) a higher probability that the subject expresses target polypeptide antigens, optionally a threshold number of the target polypeptide antigens, that have been determined to comprise at least one amino acid sequence that is both
 - A. comprised in in an active ingredient polypeptide; and
 - B. a T cell epitope capable of binding to at least three HLA class II of the subject
- (e) a higher probability that the subject expresses target polypeptide antigens, optionally a threshold number of the target polypeptide antigens, that have been determined to comprise
 - i. at least one amino acid sequence that is both
 - A. comprised in an active ingredient polypeptide; and
 - B. a T cell epitope capable of binding to at least three HLA class I of the subject; and
 - ii. at least one amino acid sequence that is both
 - A. comprised in an active ingredient polypeptide; and
 - B. a T cell epitope capable of binding to at least three HLA class II of the subject;
- (f) a higher number of target polypeptide antigens that the subject is predicted to express, optionally a higher number of target polypeptide antigens that the subject expresses with

a threshold probability, and that have been determined to comprise at least one amino acid sequence that is both

A. comprised in an active ingredient polypeptide; and

B. a T cell epitope capable of binding to at least three HLA class II of the subject; and/or

(g) a higher number of target polypeptide antigens that the subject is predicted to express, optionally a higher number of target polypeptide antigens that the subject expresses with a threshold probability, and that have been determined to comprise

i. at least one amino acid sequence that is both

A. comprised in an active ingredient polypeptide; and

B. a T cell epitope capable of binding to at least three HLA class I of the subject; and

ii. at least one amino acid sequence that is both

A. comprised in an active ingredient polypeptide; and

B. a T cell epitope capable of binding to at least three HLA class II of the subject.

7. The method of claim 6, further comprising

(I) repeating the method for one or more further pharmaceutical compositions, kits or panels of polypeptides and ranking the compositions, kits or panels of polypeptides by their likelihood to induce a clinical response in the subject; and/or

(II) predicting whether administration of the polypeptide, pharmaceutical composition, kit or panel of polypeptides will induce a toxic immune response in the subject, wherein

(a) the polypeptide(s) comprises at least one amino acid sequence that

i. is capable of binding to at least three HLA class I of the subject; and

ii. corresponds to a fragment of a human polypeptide expressed in healthy cells;

and a toxic immune response is predicted; or

(b) the polypeptide(s) do not comprise any amino acid sequence that

A. is capable of binding to at least three HLA class I of the subject; and

B. corresponds to a fragment of a human polypeptide expressed in healthy cells;

and no toxic immune response is predicted.

8. The method of any one of the preceding claims further comprising selecting or recommending for treatment of the specific human subject administration to the subject of a polypeptide that comprises a polypeptide fragment that is identified as immunogenic for the subject, or of a polypeptide that is predicted to be immunogenic, or to induce a cytotoxic T cell or helper T cell response, or of a pharmaceutical composition, kit or panel of polypeptides that is predicted to induce a clinical response, or of a polypeptide or pharmaceutical composition that is predicted not to induce a toxic immune response or not to induce ADA in the subject.

9. The method of any one of claims 1 to 4 further comprising predicting whether the subject will have a clinical response to administration of a checkpoint inhibitor to treat cancer, the method comprising determining whether one or more cancer associated antigens together comprise at least two different amino acid sequences each of which is a T cell epitope capable of binding to at least three HLA class I of the subject and predicting

A. that the subject will have a clinical response to administration of a checkpoint inhibitor if the one or more cancer associated antigens together comprise at least two different sequences each of which is a T cell epitope capable of binding to at least three HLA class I molecules of the subject; or

B. that the subject will not have a clinical response to administration of a checkpoint inhibitor if the one or more cancer associated antigens together comprise no more than one sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject;

optionally wherein the method further comprises determining the likelihood that the subject will have a clinical response to administration of a checkpoint inhibitor to treat cancer, the method comprising

(i) selecting a plurality of polypeptide antigens that are associated with the cancer type of the subject;

(ii) identifying which of said cancer associated antigens comprise an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject; and

(iii) using population expression data for each cancer associated antigen identified in step to determine the likelihood that the subject will have a clinical response to administration of

a checkpoint inhibitor to treat cancer, wherein a higher probability that the subject expresses one or more of the cancer associated antigens identified in step (ii) that together comprise at least two amino acid sequences each of which is a T cell epitope capable of binding to at least three HLA class I molecules of the subject corresponds to a more likely clinical response.

10. The method of claim 9 further comprising selecting or recommending administration of a checkpoint inhibitor for treatment of the subject.

11. Use of a checkpoint inhibitor in the preparation of a medicament for treating cancer in a human subject in need thereof, wherein the subject has been predicted to respond, or to be likely to respond, to administration of the checkpoint inhibitor by a method according to claim 9.

12. The method of any one of claims 5, 6 and 9, wherein the subject has been predicted to have a toxic immune response or ADA development, or not to have a cytotoxic T cell or helper T cell or clinical response, or not to respond to treatment with a checkpoint inhibitor and the method further comprises selecting or recommending a different treatment for the subject.

13. An *in vitro* or *in silico* method of designing or preparing a human subject-specific pharmaceutical composition or kit or panel of polypeptides for treatment of a specific human subject, the method comprising:

(i) selecting a fragment of a polypeptide, which fragment has been identified as immunogenic for the subject by the method of any one of claims 1 to 3;

(ii) if the fragment selected in step (i) is an HLA class I-binding epitope, optionally selecting a longer fragment of the polypeptide, which longer fragment

a. comprises the fragment selected in step (i); and

b. is a T cell epitope capable of binding at least three or to the most possible HLA class II molecules of the subject;

(iii) selecting a first sequence of up to 50 consecutive amino acids of the polypeptide, which consecutive amino acids comprise the amino acid sequence of the fragment selected in step (i) or the longer fragment selected in step (ii);

(iv) repeating steps (i) to (iii) to select a second amino acid sequence of up to 50 consecutive amino acids of the same or a different polypeptide to the first amino acid sequence;

(v) optionally further repeating steps (i) to (iii) to select one or more additional amino acid sequences of up to 50 consecutive amino acids of the same or different polypeptides to the first and second amino acid sequences; and

(vi) designing or preparing a subject-specific pharmaceutical composition, kit or panel of polypeptides having as active ingredients one or more polypeptides that together have all of the amino acid sequences selected in the preceding steps, optionally wherein one or more or each sequence is flanked at the N and/or C terminus by additional amino acids that are not part of the sequence of the polypeptides.

14. The method of claim 13, wherein

(a) each polypeptide either consists of one of the selected amino acid sequences, or comprises or consists of two or more of the selected amino acid sequences arranged end to end or overlapping in a single peptide, optionally wherein all of the neoepitopes formed at the join between any two of the selected amino acid sequences arranged end to end in a single polypeptide have been screened to eliminate polypeptides comprising a neoepitope amino acid sequence that

(i) corresponds to a fragment of a human polypeptide expressed in healthy cells;

(ii) is a T cell epitope capable of binding to at least two HLA class I molecules of the subject; or

(iii) meets both requirements (i) and (ii); and/or

(b) the one or more polypeptides have been screened to eliminate polypeptides comprising an amino acid sequence that

(i) corresponds to a fragment of a human polypeptide expressed in healthy cells; or

(ii) corresponds to a fragment of a human polypeptide expressed in healthy cells and is a T cell epitope capable of binding to at least two HLA class I molecules of the subject.

15. A human subject-specific pharmaceutical composition, kit or panel of polypeptides for use in therapy for inducing an immune response in a specific human subject, and designed or prepared for the subject according to the method of claim 13 or claim 14, wherein the composition or kit optionally comprises at least one pharmaceutically acceptable diluent, carrier, or preservative, optionally wherein the therapy is prophylactic.

16. A human subject-specific composition, kit or panel of polypeptides for use in therapy of a specific human subject in need thereof, the composition, kit or panel comprising as active ingredients

(a) a first and a second peptide and optionally one of more additional peptides, wherein each peptide comprises an amino acid sequence that is a T cell epitope capable of binding to at least two HLA class I molecules and/or at least two HLA class II molecules of the subject, wherein the amino acid sequence of the T cell epitope of the first, second and optionally any additional peptides are different from each other, and wherein the pharmaceutical composition or kit optionally comprises at least one pharmaceutically acceptable diluent, carrier, or preservative;

(b) a first region and a second region and optionally one of more additional regions, wherein each region comprises an amino acid sequence that is a T cell epitope capable of binding to at least two HLA class I molecules and/or at least two HLA class II molecules of the subject, wherein the amino acid sequence of the T cell epitope of the first, second and optionally any additional regions are different from each other, and wherein the pharmaceutical composition or kit optionally comprises at least one pharmaceutically acceptable diluent, carrier, or preservative; optionally wherein the therapy is prophylactic.

17. The human subject-specific pharmaceutical composition, kit or panel according to claim 16, wherein

(a) one or more or each of the peptides or regions comprises an amino acid sequence that is a T cell epitope capable of binding to at least two HLA class I molecules of the subject;

(b) one or more or each of the peptides or regions comprises an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject;

(c) one or more or each of the peptides or regions comprises an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class II molecules of the subject;

(d) one or more or each of the peptides or regions comprises an amino acid sequence that is a T cell epitope capable of binding at least one HLA class II molecule of the subject, wherein the HLA class II-binding T cell epitope comprises an amino acid sequence that is a T cell epitope capable of binding to at least two HLA class I molecules of the subject;

(e) one or more polypeptides do not comprise any amino acid sequences that

(i) corresponds to a fragment of a human polypeptide expressed in healthy cells;

or

(ii) corresponds to a fragment of a human polypeptide expressed in healthy cells and is a T cell epitope capable of binding to at least two HLA class I molecules of the subject;

(f) one or more or each of the peptides or regions comprises a sequence of up to 50 consecutive amino acids of a polypeptide that is

A. expressed by a pathogenic organism, a virus or a cancer cell, is associated with an autoimmune disorder, or is an allergen, wherein the sequence comprises the T cell epitope of the peptide or region that is capable of binding to at least two HLA class I and/or class II molecules of the subject;

B. selected from the antigens listed in Tables 2 to 6;

C. an antigen or neoantigen expressed by a cancer cell, optionally wherein the cancer cell is in a sample taken from the subject;

D. a mutational neoantigen, optionally wherein the neoantigen is present in a sample obtained from the subject; and/or the T cell epitope(s) each comprise a neoantigen specific mutation;

optionally wherein

E. one or more or each of the polypeptide sequences is flanked at the N and/or C terminus by additional amino acids that are not part of the amino acid sequence of the polypeptide(s);

F. two or more or each of the polypeptide sequences of up to 50 consecutive amino acids are from different polypeptides;

G. one or more or each of the sequences of up to 50 consecutive amino acids comprises an amino acid sequence that

(i) comprises an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject; and

(ii) is a T cell epitope capable of binding to at least three HLA class II molecules of the subject or to the most possible HLA class II molecules of the subject for a sequence comprising the HLA class I-binding epitope of (i);

H. one or more or each active ingredient peptide or polypeptide of the pharmaceutical composition, kit or panel of polypeptides either

(i) consists of one of said sequence of up to 50 consecutive amino acids from a polypeptide that is expressed by a pathogenic organism, a virus or a cancer cell, is associated with an autoimmune disorder or is an allergen; or

(ii) comprises or consist of two or more of said sequences of up to 50 consecutive amino acids arranged end to end or overlapping in a single peptide; further optionally wherein the one or more peptides or polypeptides do not comprise any neoepitopes that span a join between any two of said amino acid sequences that are arranged end to end in a single peptide and that

(1) corresponds to a fragment of a human polypeptide expressed in healthy cells;

(2) is a T cell epitope capable of binding to at least two HLA class I molecules of the subject; or

(3) meets both requirements (1) and (2).

18. The human subject-specific composition, kit or panel of polypeptides according to any one of claims 15, 16 and 17, wherein the therapy is for the treatment of cancer, and/or wherein the treatment is to be administered in combination with chemotherapy, targeted therapy or a checkpoint inhibitor.

19. A method of designing or preparing a polypeptide for inducing an immune response in a specific human subject the method comprising selecting an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules or at least three HLA class II molecules of the subject, and designing or preparing a polypeptide comprising the selected amino acid sequence; optionally wherein the method is (a) a method of designing or preparing a polypeptide for inducing a cytotoxic T cell response in a specific human subject, the method comprising selecting an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class I molecules of the subject, and designing or preparing a polypeptide comprising the selected amino acid sequence; or

(b) a method of designing or preparing a polypeptide for inducing a helper T cell response, the method comprising selecting an amino acid sequence that is a T cell epitope capable of binding to at least three HLA class II molecules of the subject, and designing or preparing a polypeptide comprising the selected amino acid sequence.

20. A system comprising

(a) a storage module configured to store data comprising the class I and/or class II HLA genotype of a subject and the amino acid sequence of one or more test polypeptides; and

(b) a computation module configured to identify and/or quantify amino acid sequences in the one or more test polypeptides that are capable of binding to multiple HLA class I molecules of the subject and/or that are capable of binding to multiple HLA class II molecules of the subject; and optionally further comprising

(c) an output module configured to display

(i) a prediction of whether the one or more polypeptides is immunogenic for the subject; or the sequence of one or more fragments of the one or more polypeptides that are predicted to be immunogenic for the subject;

(ii) a prediction of whether the individual will have an immune response to administration of the one or more polypeptides or one or more pharmaceutical compositions comprising the one or more polypeptides as active ingredients;

(iii) a prediction of whether the subject will have a clinical response to a method of treatment comprising administering to the subject one or more pharmaceutical compositions comprising the one or more polypeptides as active ingredients;

(iv) the likelihood that the subject will have a clinical response to administration of one or more pharmaceutical compositions comprising the one or more polypeptides as active ingredients;

(v) a prediction of whether administration of the one or more polypeptides or one or more pharmaceutical compositions comprising the one or more polypeptides will induce a toxic immune response in the subject;

(vi) a prediction that the one or more polypeptides is associated with an autoimmune disorder in the subject;

(vii) a prediction of whether the subject will have a clinical response to administration of a checkpoint inhibitor; or

(viii) a recommendation of whether or not the subject should be treated by administration of the one or more polypeptides and/or one or more pharmaceutical compositions.