

CLAIMS

1. A method for predicting immunogenic modified peptides, the method comprising the steps:
 - 5 a) ascertaining a score (M_{wt}) for binding of a non-modified peptide to one or more MHC molecules,
 - b) ascertaining a score (M_{mut}) for binding of a modified peptide to one or more MHC molecules, wherein the modified peptide comprises the amino acid sequence of the non-modified peptide with an amino acid substituted at a position corresponding to the same
10 relative position in the non-modified peptide, and
 - c) ascertaining a T score for binding of the modified peptide when present in a MHC-peptide complex to one or more T cell receptors, wherein the T score is based on the chemical and physical similarities between the substituted amino acid in the modified peptide and amino acid at the corresponding position in the non-modified peptide,
 - 15 wherein
the modified peptide is predicted to be immunogenic if (i) the M_{wt} satisfies a threshold indicating binding to one or more MHC molecules, (ii) the M_{mut} satisfies a threshold indicating binding to one or more MHC molecules, and (iii) the T score satisfies a threshold indicating binding of the modified peptide when present in a MHC-peptide complex to one or
20 more T cell receptors.
2. The method of claim 1, wherein the modified peptide comprises a fragment of a modified protein, said fragment comprising the modification(s) present in the protein.
- 25 3. The method of claim 1 or 2, wherein the non-modified peptide has the germline amino acid at the position(s) corresponding to the position(s) of the modification(s) in the modified peptide.
4. The method of any one of claims 1 to 3, wherein the non-modified peptide and
30 modified peptide are identical but for the modification(s).
5. The method of any one of claims 1 to 4, wherein the non-modified peptide and modified peptide are 8 to 15, preferably 8 to 12 amino acids in length.

6. The method of any one of claims 1 to 5, wherein the one or more MHC molecules comprise different MHC molecule types, in particular different MHC alleles.
7. The method of any one of claims 1 to 6, wherein the one or more MHC molecules are
5 MHC class I molecules and/or MHC class II molecules.
8. The method of any one of claims 1 to 7, wherein the score for binding to one or more MHC molecules is ascertained by a process comprising a sequence comparison with a database of MHC-binding motifs.
10
9. The method of any one of claims 1 to 8, wherein the threshold applied for M_{wt} is different to the threshold applied for M_{mut} .
10. The method of any one of claims 1 to 9, wherein the threshold for binding to one or
15 more MHC molecules reflects a probability for binding to one or more MHC molecules.
11. The method of any one of claims 1 to 10, wherein the score for the chemical and physical similarities is ascertained on the basis of the probability of amino acids being interchanged in nature.
20
12. The method of claim 11, wherein the more frequently amino acids are interchanged in nature the more similar the amino acids are considered.
13. The method of any one of claims 1 to 12, wherein the chemical and physical
25 similarities are determined using evolutionary based log-odds matrices.
14. The method of any one of claims 1 to 13, wherein the modification is not in an anchor position for binding to one or more MHC molecules.
- 30 15. The method of any one of claims 1 to 14, wherein the modification is in an anchor position for binding to one or more MHC molecules.

16. The method of any one of claims 1 to 15, which comprises performing step b) on two or more different modified peptides, said two or more different modified peptides comprising the same modification(s).
- 5 17. The method of claim 16, wherein the two or more different modified peptides comprising the same modification(s) comprise different fragments of a modified protein, said different fragments comprising the same modification(s) present in the protein.
- 10 18. The method of claim 16 or 17, wherein the two or more different modified peptides comprising the same modification(s) comprise all potential MHC binding fragments of a modified protein, said fragments comprising the same modification(s) present in the protein.
- 15 19. The method of any one of claims 16 to 18, further comprising selecting (the) modified peptide(s) from the two or more different modified peptides comprising the same modification(s) having a probability or having the highest probability for binding to one or more MHC molecules.
- 20 20. The method of any one of claims 16 to 19, wherein the two or more different modified peptides comprising the same modification(s) differ in length and/or position of the modification(s).
- 25 21. The method of any one of claims 1 to 15, which comprises performing step b) on two or more different modified peptides, said two or more different modified peptides comprising different modifications.
22. The method of claim 21, wherein the different modifications are present in the same and/or in different proteins.
- 30 23. The method of any one of claims 16 to 22, which comprises comparing the scores of two or more of said different modified peptides.
24. The method of any one of claims 1 to 23, further comprising identifying non-synonymous mutations in one or more protein-coding regions.

25. The method of any one of claims 1 to 24, wherein modifications are identified by partially or completely sequencing the genome or transcriptome of one or more cells such as one or more cancer cells and optionally one or more non-cancerous cells and identifying mutations in one or more protein-coding regions.

5

26. The method of claim 24 or 25, wherein said mutations are somatic mutations.

27. The method of any one of claims 24 to 26, wherein said mutations are cancer mutations.

10