

Immunodominant proteins and fragments in multiple sclerosis

FIELD OF INVENTION

The disclosure relates to the treatment, diagnosis and/or prevention of multiple sclerosis by using an immunodominant protein or peptide. More particular the invention relates to the field of antigen specific immunotherapies, such as the induction of tolerance.

5

BACKGROUND OF THE INVENTION

Multiple sclerosis (MS) is a devastating autoimmune inflammatory disease mainly affecting young adults. MS is a prototypic example of an organ-specific autoimmune disease (AID), as the autoimmune response only targets the central nervous system (CNS) consisting of brain and spinal cord. Organ-specific AID means that the immune system of the patient damages a specific tissue or cell type by autoreactive T cells and/or antibodies.

MS preferentially affects young adults between 20 and 40 years, but children and older individuals can also develop MS. The disease is about 2-3 times more frequent in women than in men. MS usually becomes clinically manifest by temporary problems with vision (acute optic neuritis), sensation, or motor and autonomous function, but can lead to a broad range of neurological symptoms.

At the time of first manifestation, when differential diagnoses have been excluded, the disease is referred to as clinically isolated syndrome (CIS) provided that the cerebrospinal fluid (CSF) and magnetic resonance imaging (MRI) findings are consistent with the diagnosis. MRI discloses lesions in locations typical for MS, i.e. juxtacortical, periventricular, in the brain stem or spinal cord. If certain criteria are fulfilled that can be summarized as dissemination in space (more than one lesion or clinical symptom/sign) and time (more than one event) then the

diagnosis of relapsing-remitting multiple sclerosis (RRMS) can be made. A special scenario is the accidental discovery of MRI lesions compatible with MS without clinical symptoms. This is referred to as radiologically isolated syndrome (RIS) and can be considered a pre-stage of CIS and RRMS. More than 80% of patients suffer from one of these, and the majority of patients
5 develops later what is called secondary progressive MS (SPMS). At this time, relapses/exacerbations become less frequent or stop altogether and neurological disability increases steadily either between relapses or without these.

A special form of MS is primary progressive MS (PPMS), which never shows relapses, but
10 rather begins with steady worsening of neurological symptoms, e.g. of the ability to walk. PPMS affects approximately 10% of MS patients and males and females with equal frequency. Its onset is usually later than CIS or RRMS. With respect to causes and disease mechanisms PPMS is considered similar to the above RIS-CIS-RRMS-SPMS.

15 Typically, MS is diagnosed according to the revised McDonald or recently Lublin criteria. These criteria also allow distinguishing between the different forms and disease activity of MS (Thompson et al., 2018, Lancet Neurol, 17(2):162-173).

MS is a disease with a complex genetic background. More than 200 MS risk alleles or
20 quantitative traits (common variants of genes detected as single nucleotide polymorphisms, SNPs) have been identified in the last decade, however, by far the most important is the human leukocyte antigen (HLA)-DR15 haplotype. In addition, several environmental/lifestyle risk factors have been found. These include infection with Epstein Barr virus (EBV), smoking, low vitamin D3 levels and obesity as the most important ones.

25 All the genetic and environmental risk factors are common and shared by many individuals in the healthy population. The exact reasons, why the disease starts in individuals with certain genetic and environmental risk factors, are not clear, but one assumes that viral and bacterial infections, for instance by changes in the gut microbiota, can be triggers. The concordance rate
30 of monozygotic twins of 10-30% and the risk of first-degree relatives of an MS patient of approximately 2-4%, compared to a risk of 1/1000 in the general population, provide an estimate of the genetic versus the environmental risk, although the interplay between the two is also complex.

35 In order to identify the components of the CNS, against which the autoimmune response in MS is directed, researchers oriented their efforts towards the cells and structures that are affected in

MS, particularly myelin and axons/neurons and the proteins that are specific for these cells/structures. During the last thirty years, several myelin proteins such as myelin basic protein (MBP), proteolipid protein (PLP) and myelin oligodendroglia glycoprotein (MOG) have been identified as encephalitogenic in animal models (experimental autoimmune encephalomyelitis; EAE), i.e. their injection into susceptible rodent strains leads to a disease with similarities with MS, but also by examining immune cells from MS patients (Sospedra and Martin, 2005, *Annu Rev Immunol*, 23:683-747). The above autoantigens are CNS-specific and exclusively (PLP and MOG) or almost exclusively (MBP) expressed in the brain. In MS, a few autoantigens that are not CNS-specific such as alpha-B crystallin and transaldolase-H have also been described as potential targets.

Current evidence suggests CD4+ autoreactive T cells as a central factor for the autoimmune pathogenesis of MS probably relevant not only for the induction and maintenance of the autoimmune response, but also during tissue damage (Sospedra and Martin, 2005). The frequency of high avidity CD4+ T cells reactive to main constituents of the myelin sheath, such as MBP, PLP and MOG is increased in MS patients (Bielekova et al., 2004, *J Immunol*, 172:3893-3904). Due to their involvement in disease pathogenesis CD4+ T cells are a target for therapeutic interventions.

Detailed investigation of the immune response against the CNS-specific proteins showed that certain peptides thereof are recognized by a large fraction of patients and in the context of the disease-associated HLA-DR molecules. Such peptides are referred to as immunodominant (Bielekova et al., 2004).

The following characteristics indicate that a certain peptide of a protein is immunodominant in the context of MS:

a) frequent recognition of this peptide by T cells, i. e. by approximately 10% or more of MS patients, often in the context of a disease-associated HLA allele or haplotype (Sospedra and Martin, 2005), and

b) recognition of this peptide by disease-relevant T cells such as those that respond to peptides at low concentrations (high avidity T cells) (Bielekova et al., 2004) and are therefore considered particularly dangerous, and/or have a proinflammatory phenotype, and/or are isolated from the target organ or compartment (CNS), in the case of MS, brain-, spinal cord- or CSF-infiltrating T cells.

However, high avidity recognition is not a prerequisite, since low-avidity myelin-specific T cells have also been shown to be pathogenic in humanized transgenic mouse models (Quandt et al. 2012, J Immunol, 189(6): 2897-2908).

5 It has recently been demonstrated that T cells of MS patients show increased *in vitro* proliferation in the absence of an exogenous antigen (Mohme et al., 2013, Brain, 136:1783-1798). These "autoproliferating" T cells are enriched for cells that home to the CNS compartment of MS patients and can thus be considered as a peripheral blood source of brain-/CSF-infiltrating T cells (Jelcic et al., 2018, Cell, 175(1):85-100.e23).

10

In the case that data from testing T cells *in vitro* is not available or in addition to such testings, immune recognition of peptides can also be predicted/inferred from those peptides that will bind well to the HLA-class I or -class II alleles of the individual and for CD8+ and CD4+ T cells respectively. Peptide binding predictions are well known to the skilled person. They can be
15 performed by well-established prediction algorithms (NetMHCII - www.cbs.dtu.dk/services/NetMHCII/; IEDB - www.iedb.org/) and analysis of the HLA-binding motifs (SYFPEITHI - www.syfpeithi.de/).

Immunodominant peptides can be used in antigen-specific immunotherapies such as tolerance
20 induction. One example is EP 2 205 273 B1, which discloses immunodominant peptides of MBP, PLP and MOG and their application for MS treatment. In the approach disclosed therein, the peptides are coupled to white or red blood cells.

Tolerance induction is antigen-specific and renders autoreactive T cells non-functional or
25 anergic or induces regulatory T (Treg) cells that specifically suppress untoward autoimmunity to said target antigens. The induction of tolerance to target autoantigens is a highly important therapeutic goal in autoimmune diseases. It offers the opportunity to attenuate specifically the pathogenic autoimmune response in an effective way with few side effects. Tolerance induction can also be achieved by applying a whole protein instead of or in addition to an
30 immunodominant peptide being a fragment of the protein (Kennedy MK et al., 1990, J Immunol, 144(3):909-15).

Some pathological characteristics of MS are reflected in the EAE model, a paradigmatic animal
35 model of Th1/Th17 cell-driven autoimmune disease. Studies in relapsing EAE (R-EAE) in the SJL mouse have clearly shown that chronic demyelination involves the activation of T cell responses to an immunodominant myelin peptide, i.e. PLP 139-154, to which the first disease

exacerbation is directed. Subsequently, the immune response broadens to other myelin peptides of PLP, MBP and MOG, a process, which is referred to as epitope spreading. Unresponsiveness of T cells, i.e. tolerance, can for example be induced when antigen presenting cells (APC) pulsed with antigenic peptide are for example treated with the cross linker 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (ECDI; also abbreviated EDC).

Preclinical experiments have proven that a single i.v. injection of naïve murine splenocytes pulsed with a mixture of encephalitogenic myelin peptides and fixed with the cross-linker EDC is highly efficient in inducing peptide-specific tolerance *in vivo*. In EAE, this protocol not only prevented animals from disease, but even effectively reduced the onset and severity of all subsequent relapses when given after disease induction, indicating that specific tolerance can downregulate an ongoing autoimmune response (Miller et al., 1991, Acad Sci, 636:79-94). More relevant to the treatment of MS, studies in EAE have shown that tolerance can be simultaneously induced to multiple epitopes using a cocktail of encephalitogenic myelin peptides, thus providing the capacity to target autoreactive T cells with multiple specificities.

Tolerization of human T cells by autologous antigen-coupled cells, e. g. APCs (Vandenbark et al., 2000, Int Immunol, 12:57-66) or non-nucleated cells, i.e. red blood cells (RBCs), treated with EDC is effective *in vitro* as shown by failure of tolerized T cells to proliferate or to produce Th1 cytokines and a decreased expression of costimulatory molecules on these cells.

There is evidence that at least two distinct mechanisms are involved in the induction of antigen-specific tolerance by this regime:

1) Direct tolerance where Th1 clones encountering nominal antigen/MHC complexes on antigen-coupled APCs were anergized as a result of failure to receive adequate CD28-mediated costimulation, and

2) indirect mechanisms such as cross tolerance, where tolerance is induced by reprocessing and re-presentation of antigens by tolerogenic host APCs and/or expansion of Treg cells.

The latter cross tolerance likely involves the induction and/or expansion of antigen-specific Treg cells which assumption is also supported by data obtained in a phase Ib trial as disclosed herein. Further, treatment of cells with EDC induces apoptosis in a substantial percentage of treated cells. Thus, an indirect mechanism that involves fixed APC undergoing apoptosis, which are then processed and represented by host APC, is likely. This is further supported by effective

induction of tolerance in MHC-deficient and allogeneic mice. In-vitro bone marrow-derived dendritic cells effectively phagocytose and process antigen-pulsed, fixed APC.

5 Currently approved therapies for MS involve various antigen-nonspecific immunomodulating or immunosuppressive strategies, which are only partially effective. All current therapeutics need to be taken orally daily or injected/infused at various time intervals and for long periods of time. Further, they are associated with numerous and sometimes severe side effects.

10 A therapy that addresses the pathogenesis of MS at its roots should aim to specifically delete or functionally inhibit pathogenic autoreactive cells without altering the "normal" immune system. This is of importance because global immunomodulation and/or immunosuppression come at the cost of inhibiting beneficial regulatory cells and immune cells that serve protective functions against pathogens. Ideally, peptide-specific immune tolerance, that is the specific correction of the misdirected autoimmune response against brain/spinal cord tissue, should be achieved
15 early in the inflammatory phase of the disease, when blockade of the autoreactive immune response can inhibit dissemination and propagation of the disease and irreversible disability can be prevented. Therefore, the preferred targeted patient group are relapsing-remitting MS patients early in the disease course or even patients presenting with a first clinical event suggestive of MS, i.e. CIS, or patients, in whom the disease is discovered even earlier at the
20 stage of RIS. At this time point, MS patients generally have a low grade of neurologic disability, which allows them to participate in all activities of daily life and work without significant compromise.

SUMMARY OF THE INVENTION

25 It is an object of the present invention to identify MS-relevant antigens suitable for use in the treatment, diagnosis and/or prevention of MS, in particular in a tolerization approach. A further aspect of the present invention is the identification of a human subject, who is suitable for tolerization.

30 The present invention is based on a novel approach for identifying MS-relevant antigens: T cells clonally expanded in the brain of an MS patient (homozygous for the HLA DR15 haplotype, which is known to be the major genetic risk factor in MS), who died of a very aggressive form of MS were examined. Thereby, for the first time T cells originating from and clonally expanded in
35 the target organ were analyzed for the purpose of antigen identification. In all previous approaches, peripheral blood lymphocytes were analyzed to identify immunodominant antigens.

The clonal expansion of a T cell clone (TCC) in MS brain lesions suggests that the cells are MS-relevant. Herein, the target antigens of a specific TCC that has previously been described in Planas et al. (Planas et al., 2015, *Ann Clin Transl Neurol*, 2(9):875-893), TCC21.1, have been identified. Furthermore, the target antigens of another TCC (TCC14) of the same patient have
5 been identified. In the case of TCC14, the clone was isolated from a peripheral blood T cell population that was identified as disease-relevant by demonstrating increased spontaneous proliferation (autoproliferation) and enrichment for brain-homing autoreactive T cells. Specifically, TCC14 was also found to be clonally expanded in MS brain lesions, even though it had been isolated from the peripheral blood. The isolation and identification of disease-relevant
10 T cells like TCC21.1 and TCC14 is schematically depicted in Fig. 1.

Target epitopes that are recognized by biologically relevant (e. g. tissue-infiltrating) T cells can then be identified (Fig. 2).

15 This novel approach revealed the proteins GDP-L-fucose synthase (gene abbreviation: TSTA3; also known as: GDP-L-fucose:NADP⁺ 4-oxidoreductase (3,5-epimerizing) or GDP-4-keto-6-deoxy-D-mannose-3,5-epimerase-4-reductase) and proteins of the RASGRP (RAS guanyl releasing protein) family, including RASGRP1, RASGRP2, RASGRP3 and RASGRP4 as well as splice variants and isoforms thereof as particularly relevant in MS. RASGRP2 is especially
20 preferred. Thus, these proteins have been found to be immunodominant in MS and are autoantigens.

The relevance has also been tested in CSF-infiltrating CD4⁺ T cells from CIS/MS patients for GDP-L-fucose synthase and in peripheral blood-derived mononuclear cells for RASGRP2. In a
25 further analysis, both GDP-L-fucose synthase and RASGRP2 have been tested in CSF-infiltrating CD4⁺ T cells from CIS/MS patients.

Thus, the antigens as described herein in the examples are the first immunodominant antigens in MS, which have been discovered by examining the specificity of T cells that are clonally
30 expanded in MS brain lesions and therefore can be assumed to be involved in the damaging autoimmune reaction in the brain. The methodology that led to their identification, combinatorial peptide libraries, does not involve the above-mentioned focus on myelin/brain proteins, but is completely bias-free. The antigens have not been described as implicated in MS before. Further, RNA sequencing and proteomics have shown that both autoantigens, GDP-L-fucose
35 synthase and RASGRP2 and the related proteins RASGRP1 and -3, are expressed in MS brain tissue.

The identified antigens can be used in the treatment, diagnosis and/or prevention of MS, in particular in a tolerization approach, and for identifying a human subject, who is suitable for tolerization. By identifying a human subject, who is suitable for tolerization, the human subject can be *in vitro* diagnosed with MS. In other words, the identified autoantigens can be used in the *in vitro* diagnosis of MS. This *in vitro* testing can be complemented with clinical and imaging findings, i. e. the MS diagnosis according to the state of the art, in particular according to the revised McDonald criteria. Thus, the identified autoantigens may serve diagnosing MS in a human subject either with or without additional diagnostic tests.

10

BRIEF DESCRIPTION OF THE FIGURES

- Figure 1 Isolation of disease-relevant T cells
- Figure 2 Target epitope discovery
- Figure 3 Single and dual-defined decapeptide positional scanning mixtures in combination with biometric analysis to identify the specificity of TCC21.1
- Figure 4 Summary of human decapeptides predicted with the biometrical approach, synthesized and tested for stimulatory capacity
- Figure 5 GDP- L-fucose synthase transcripts and peptides identified in brain tissue
- Figure 6 GDP-L-fucose synthase, myelin and CEF Peptides (control peptides from cytomegalovirus, Epstein Barr virus and influenza virus)
- Figure 7 Recognition of GDP-L-fucose synthase peptides by CSF-infiltrating CD4⁺ T cells from CIS/MS patients
- Figure 8 Isolation of autoproliiferating T cells and autoproliiferating T cells found in MS brain lesions
- Figure 9 Screening procedure for peptide ligand identification of the autoproliiferating, peripheral blood isolated brain-homing TCC14 using a positional scanning library
- Figure 10 RASGRP2 reactivity of peripheral blood-derived memory T cells
- Figure 11 Expression of RASGRPs in peripheral blood B cells and brain
- Figure 12 Peptide identification of RASGRP1, 2 and 3 protein in brain tissue
- Figure 13 Staining of biotin-PLP1-coupled cells
- Figure 14 Recognition of GDP-L-fucose synthase and myelin peptides by CSF-infiltrating CD4⁺ T cells from 105 MS patients
- Figure 15 Recognition of RASGRP2 and myelin peptides by CSF-infiltrating CD4⁺ T cells from 57 MS patients
- Figure 16 Tolerance induction *in vivo* with red blood cells coupled to myelin peptides

DETAILED DESCRIPTION OF THE INVENTION

T cells used for identifying immunodominant peptides and the corresponding protein are ideally those T cells that are pathogenetically relevant for the disease. Regarding the latter
5 characteristic, those T cells that are clonally expanded in the target tissue of MS, the brain, spinal cord and CSF, are of greatest interest for identification of disease-relevant target antigens. According to a method as described by Planas et al. (2015) next-generation sequencing of T cell receptor (TCR) beta chain complementary- or genomic-DNA sequences has been used to identify clonally expanded T cells in brain autopsy lesions of MS patients and
10 to isolate these T cells as TCC from autologous CSF and/or tissue (comprising living cells and obtained by, e. g. biopsy or early autopsy) and characterize them with respect to functional phenotype and antigen specificity. Thereby disease-relevant TCC have been isolated. Specifically, TCC21.1 has been identified and characterized: TCC21.1 displayed a Th2 phenotype releasing mainly Th2 cytokines and was able to provide B cell help for antibody
15 production. This strategy has led to the identification of the relevance of the GDP-L-fucose synthase protein.

The above strategy, i.e. deep TCR sequencing of brain/spinal cord/CSF-infiltrating T cells, has also been used to isolate disease-relevant T cells from peripheral blood and clone identified T
20 cells from autoproliferating peripheral blood mononuclear cells (PBMC). This strategy has led to the identification of the relevance of the RASGRP protein family.

The present studies with CSF-infiltrating T cells (for GDP-L-fucose synthase and in a further analysis for both GDP-L-fucose synthase and RASGRP2) and with peripheral blood T cells (for
25 RASGRP2) thus demonstrate that these are immunodominant targets of the autoimmune response in MS, and both GDP-L-fucose synthase and RASGRP2 (and other members of the RASGRP family) are recognized by brain-infiltrating T cells in MS, including CIS, RRMS and SPMS. The NetMHCII and the IEDB in silico peptide binding prediction algorithms have also been used to identify immunodominant regions within the respective protein (as described in
30 "Examples").

The cytosolic enzyme GDP-L-fucose synthase converts GDP-4-keto-6-deoxy-D-mannose into GDP-L-fucose, which is then used by fucosyltransferases to fucosylate all oligosaccharides. In mammals, fucosylated glycans play important roles in many biological processes including
35 blood transfusion reactions, host-microbe interactions, cancer pathogenesis and maintenance of a non-inflammatory environment in the brain.

The RASGR proteins exist in at least the four variants RASGRP1, RASGRP2, RASGRP3 and RASGRP4. RASGRP1, RASGRP2, RASGRP3 in particular have been implicated in the present invention. The protein family is characterized by the presence of a Ras superfamily guanine nucleotide exchange factor (GEF) domain which functions as a diacylglycerol (DAG)-regulated nucleotide exchange factor specifically activating Ras through the exchange of bound GDP for GTP. The proteins of the protein family activate the Erk/MAP kinase cascade. They are involved in reduced apoptosis and tumorigenesis of EBV-infected B cells, B and T cell signaling, - adhesion, -motility and are crucial for maintaining B-T cell homeostasis. At least four isoforms, i. e. splice variants, of RASGRP2 exist.

The identified antigens GDP-L-fucose synthase and RASGRP2 (as well as RASGRP1, RASGRP3 and RASGRP4) have not been implicated in MS etiology or pathogenesis or its animal model, EAE, before. The present finding that both proteins or fragments, derivatives or splice variants thereof are an immunodominant target of the autoimmune response in MS allows using them in the treatment, diagnosis and/or prevention of MS.

A protein is intended to mean oligopeptides, polypeptides as well as proteins as such. A protein sequence may be defined by a GenBank entry. A protein sequence may also be defined by a UniProtKB/Swiss-Prot entry and/or by a GenPept entry. An entry may be defined by a number, e. g. an accession number. Where applicable, the database entries include the respective accession number (i. e. an entry number) and version number. A protein may also be defined by any other database known to the skilled person. Different isoforms, derivatives and/or splice variants may exist which are also encompassed by the present invention. Thereby, the sequence may vary from the known sequence from, for example, the GenBank or UniProtKB/Swiss-Prot entry.

“A” protein or “the” protein according to the present invention refers to either a GDP-L-fucose synthase protein or a protein of the RASGRP protein family, preferably RASGRP2, or refers to both a GDP-L-fucose synthase protein or a protein of the RASGRP protein family, preferably RASGRP2, unless otherwise explicitly mentioned.

A splice variant arises from alternative splicing during gene expression. The splice variant according to the invention is preferably immunodominant.

A fragment is preferably any part of the protein, which is shorter, i.e. has less amino acids, than the parent protein. A fragment may be a peptide. In one embodiment, the fragment comprises 5

to 50, preferably 5 to 20, more preferably 10 to 15 amino acids, even more preferably 15 amino acids. The fragment according to the invention is preferably immunodominant.

It is also possible to use more than one fragment according to the invention. Preferably, more
5 than three, more than five, more than 10, more than 15, or even more than 20 different fragments are used. In a preferred embodiment, between five and 20, preferably between five and 15 different fragments are used. A fragment is different from another fragment if it does not consist of the same amino acid sequence.

10 In another embodiment, at least one fragment of each protein according to the present invention (a GDP-L-fucose synthase protein or a protein of the RASGRP protein family, preferably RASGRP2) are used in combination. It is particularly advantageous to combine at least one fragment of each protein according to the present invention with at least one known peptide of the state of the art, in particular with at least one myelin peptide, in particular with at least one or
15 all of the myelin peptides as defined by SEQ ID NOs: 261 to 267.

A derivative of a sequence is preferably defined as an amino acid sequence which shares a homology or identity over its entire length with a corresponding part of the reference amino acid sequence of at least 75%, more preferably at least 80%, at least 85%, at least 90%, at least
20 93%, at least 95%, at least 97%, at least 98% or at least 99%. The "corresponding part" in the sense of the present invention preferably refers to the same stretch of amino acids of the same parent sequence. For example, if a derivative with a length of 100 amino acids differs from a stretch of amino acids of SEQ ID NO: 1 (amino acids 1 to 100 of SEQ ID NO: 1) by 20 amino acids, this particular derivative shares an identity of 80 % over its entire length with the
25 corresponding part, i. e. amino acids 1 to 100, of the reference amino acid sequence, i. e. SEQ ID NO: 1. The derivative according to the invention is preferably immunodominant.

A "homology" or "identity" of an amino acid sequence is preferably determined according to the invention over the entire length of the reference amino acid sequence or over the entire length
30 of the corresponding part of the reference amino acid sequence which corresponds to the sequence which homology or identity is defined.

An "identity" is defined as identical amino acids, a "homology" comprises identical amino acids as well as conservative substitutions. The person skilled in the art is aware of conservative
35 substitutions, such as

- aromatic and aromatic F and W/Y

- positively charged and positively charged R and K/H
- negatively charged E and D or
- aliphatic V and L/M/I, or A and S/T.

5 The nucleotide sequence encoding any of the proteins of the present invention or fragment, derivative or splice variant thereof refers to any coding nucleotide sequence, for example RNA or DNA, in particular mRNA or cDNA. In one embodiment, the nucleotide sequence is a plasmid or any type of vector known to the person skilled in the art. In a preferred embodiment, the nucleotide sequences do not comprise introns and the gene sequences comprise exons and
10 introns.

In one aspect of the invention, the protein for use in the treatment, diagnosis and/or prevention of multiple sclerosis (MS) is GDP-L-fucose synthase or a fragment, derivative or splice variant thereof. In another aspect of the invention, the protein is a member of the RASGRP family or a
15 fragment, derivative or splice variant thereof. The invention also refers to a nucleotide sequence encoding any of the proteins or fragment, derivative or splice variant thereof, for use in the treatment, diagnosis and/or prevention of multiple sclerosis (MS).

In a preferred embodiment, the proteins are human proteins and/or the nucleotide sequences and/or gene sequences are human sequences.
20

In one embodiment, the GDP-L-fucose synthase shows enzyme activity converting GDP-4-keto-6-deoxy-D-mannose into GDP-L-fucose.

25 In another embodiment, the member of the RASGRP family activates Ras through the exchange of bound GDP for GTP. Additionally or alternatively, the protein activates the Erk/MAP kinase cascade.

In a preferred embodiment, the GDP-L-fucose synthase protein
30 a) has the amino acid sequence as set forth in SEQ ID NO: 1 or
b) has an amino acid sequence which is at least 85 %, preferably at least 90 %, more preferably at least 95 % identical to the amino acid sequence as set forth in SEQ ID NO: 1 or
c) has an amino acid sequence which is at least 70 %, preferably at least 80 %, more
35 preferably at least 90 % homologous to the amino acid sequence as set forth in SEQ ID NO: 1 or

d) has an amino acid sequence which is at least 60 %, preferably at least 70 %, more preferably at least 80 %, even more preferably at least 90 % homologous to the amino acid sequence as set forth in SEQ ID NO: 1 and the protein or fragment or splice variant thereof binds to an autologous HLA allele, is recognized by a T cell and/or is recognized by an antibody which binds to or recognizes the amino acid sequence as set forth in SEQ ID NO: 1 or a fragment thereof or

e) is encoded by a TSTA3 gene, in particular by a gene sequence of nucleotides 143612618 to 143618048 of NC_000008.11, or is encoded by a gene which is at least 80 %, preferably at least 90 %, even more preferably at least 95 % identical to the gene sequence of nucleotides 143612618 to 143618048 of NC_000008.11.

In another preferred embodiment, the member of the RASGRP protein family

f) has the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 or SEQ ID NO: 9 or

g) has an amino acid sequence which is at least 85 %, preferably at least 90 %, more preferably at least 95 % identical to the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 or SEQ ID NO: 9 or

h) has an amino acid sequence which is at least 70 %, preferably at least 80 %, more preferably at least 90 % homologous to the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 or SEQ ID NO: 9 or

i) has an amino acid sequence which is at least 60 %, preferably at least 70 %, more preferably at least 80 %, even more preferably at least 90 % homologous to the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 or SEQ ID NO: 9 and the protein or fragment or splice variant thereof binds to an autologous HLA allele, is recognized by a T cell and/or is recognized by an antibody which binds to or recognizes the respective amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 or SEQ ID NO: 9 or a fragment thereof or

j) is encoded by a RASGRP gene, in particular by a gene sequence of nucleotides

- 38488101 to 38565575 of NC_000015.10,
- 64726911 to 64745456 of NC_000011.10,
- 33436324 to 33564750 of NC_000002.12, or

- 38409051 to 38426305 of NC_000019.10,
or is encoded by a gene which is at least 80 %, preferably at least 90 %, even more preferably at least 95 % identical to the gene sequence of nucleotides
- 38488101 to 38565575 of NC_000015.10,
- 5 – 64726911 to 64745456 of NC_000011.10,
- 33436324 to 33564750 of NC_000002.12, or
- 38409051 to 38426305 of NC_000019.10.

10 The binding to an autologous HLA allele, recognition by a T cell and/o recognition by an antibody may indicate immunodominance of the protein or fragment or splice variant thereof. Immunodominance can also be tested as disclosed below.

15 In a particularly preferred embodiment, a GDP-L-fucose synthase or RASGRP2 protein or splice variant thereof, preferably the GDP-L-fucose synthase or RASGRP2 protein, is used in the present invention. The GDP-L-fucose synthase protein has, for example, the sequence as set forth in SEQ ID NO: 1, the RASGRP2 protein has, for example, the sequence as set forth in SEQ ID NO: 2, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 or SEQ ID NO: 9.

20 In one embodiment, the fragment comprises 5 to 50, preferably 5 to 20, more preferably 10 to 15, even more preferably 15 amino acids.

In another embodiment, the fragment is

- a) at least 85 %, preferably at least 90 %, more preferably at least 95 % identical to a respective corresponding amino acid sequence or
- 25 b) at least 70 %, preferably at least 80 %, more preferably at least 90 % homologous to a respective corresponding amino acid sequence or
- c) at least 60 %, preferably at least 70 %, more preferably at least 80 %, even more preferably at least 90 % homologous to a respective corresponding amino acid sequence and binds to an autologous HLA allele, is recognized by a T cell and/or is
- 30 recognized by an antibody which binds to or recognizes the respective amino acid sequence.

35 The “respective corresponding amino acid sequence” refers to the respective fragment of the corresponding amino acid sequence (i. e. SEQ ID NO: 1) with the same length as the homologous fragment (see also definition of “corresponding part” *supra*). A fragment with these identities and/or homologies may comprise 5 to 50, preferably 5 to 20, more preferably 10 to 15,

even more preferably 15 amino acids. The identity and/or homology is determined over the entire length of the respective fragment. In other words, the “corresponding amino acid sequence” refers to the unaltered sequence, i. e. when a fragment of a sequence as set forth in SEQ ID NO: 1 is 85 % identical to a respective corresponding amino acid, the fragment is 85 % identical (over the entire length of the fragment) to the unaltered fragment as “excised”, i. e. directly taken or copied from SEQ ID NO: 1.

Therefore, in one embodiment, the protein (GDP-L-fucose synthase or a member of the RASGRP protein family) has an amino acid sequence with a certain homology (at least 60 %, preferably at least 70 %, more preferably at least 80 %, even more preferably at least 90 %) to the respective depicted sequence as set forth in the SEQ ID NOs with the additional requirement that the protein or fragment or splice variant thereof binds to an autologous HLA allele, is recognized by a T cell and/or is recognized by an antibody which binds to or recognizes the amino acid sequence as set forth in the respective SEQ ID NO or a fragment thereof. In another embodiment, the protein or fragment or splice variant thereof binds to an autologous HLA allele and is recognized by a T cell which binds to or recognizes the amino acid sequence as set forth in the respective SEQ ID NO or a fragment thereof.

Assays for measuring and/or predicting binding to an autologous HLA allele, recognition by a T cell or recognition by an antibody are well known to the person skilled in the art. Binding of a peptide to an HLA allele can for example be predicted by the using the well-accepted NetMHCII (<http://www.cbs.dtu.dk/services/NetMHCII/>) or IEDB (<http://www.iedb.org/>) in silico peptide binding prediction algorithms. T cell recognition can for example be measured by a T cell proliferation assay, for example by measuring incorporated radioactivity. Binding of a peptide and/or a protein to an antibody can be measured by standard assays known to the person skilled in the art, for example by ELISA. The binding to an autologous HLA allele, recognition by a T cell or recognition by an antibody may indicate immunodominance of a peptide or protein. Immunodominance can also be tested as disclosed below.

In one embodiment, the fragment represents a stretch of consecutive amino acids (for example, 20 to 30 amino acids) which are at least 90 % identical or homologous between proteins of the RASGRP family.

In a preferred embodiment the peptides used for the treatment according to the invention are fragments of GDP-L-fucose synthase or of a protein of the RASGRP protein family and comprise a sequence selected from the group consisting of SEQ ID NOs: 10 to 98. In especially

preferred embodiments the peptides consist of amino acid sequences as depicted in one of the SEQ ID NOs: 10 to 98. The sequences of SEQ ID NOs: 10 to 35 are preferred.

The sequences according to SEQ ID NOs: 10 to 35 have been identified as immunodominant peptides due to being recognized by disease-relevant T cells, and subsequent validation of recognition by CSF-infiltrating bulk T cells for GDP-L-fucose synthase and PBMC for RASGRP2, (cf. "Examples"). The amino acid sequences according to SEQ ID NOs: 36 and 37 represent a stretch of identical sequences between RASGRP2 und RASGRP3 (alignment of UniProtKB/Swiss-Prot Q7LDG7-1 and UniProtKB/Swiss-Prot Q8IV61). The amino acids according to SEQ ID NOs 38 to 98 are comprised in peptide pools which gave rise to reactions of memory T cells from MS patients (cf. 9. of Examples and Figure 10). Sequences with SEQ ID NOs: 12, 21, 23, 28 and 32 (GDP-L-fucose synthase) and with SEQ ID NO: 46 (RASGRP2) have been validated in a further analysis with CSF-infiltrating CD4⁺ T cells (cf. 13. of "Examples" and Figs. 14 and 15).

15

The sequences of SEQ ID NOs: 1-98 are also listed in the following Table 1:

Table 1: Protein sequences, position and database entry. The sequence Q7LDG7-1 is the canonical sequence of the RASGRP2 protein and also refers to RASGRP2 Isoform 1. Proteins and peptides depicted hereunder are particularly preferred embodiments of the invention.

20

SEQ ID NO:	Sequence (amino acids)	Position	Database
1	MGEPQGSMRILVTGGSLVGKAIQKVADGAGL PGEDWVFVSSKDADLTDTAQTRALFEKVQPTHVI HLAAMVGGLFRNIKYNLDFWRKNVHMNDNVLHS AFEVGARKVVSCLSTCIFPDKTTYPIDETMIHNGP PHNSNFGYSYAKRMIDVQNRAYFQQYGCTFTAVI PTNVFGPHDNFNIEDGHVLPGLIHKVHLAKSSGS ALTVWGTGNPRRQFIYSLDLAQLFIWVLREYNEV EPIILSVGEEDEVSIKEAAEAVVEAMDFHGEVTFD TTKSDGQFKKTASNSKLRTYLPDFRFTPFKQAVK ETCAWFTDNYEQARK	1-321 (whole protein)	GenBank AAH93061.1 or UniProtKB/Swiss -Prot: Q13630 (GDP L-fucose synthase)
2	AAAAARPAGGSARRWGRPGRCGLLAAGPKRVR SEPGGRLPERSLGPAPAPAAAMAGTLDLDKGCT	1-662 (whole protein)	GenBank AAI10307.1

	VEELLRGCI EAFDDSGKVRDPQLVRMFLMMHPW YIPSSQLAAKLLHIYQQSRKDNSNSLQVKTCHLVR YWISAFPAEFDLNPELAEQIKELKALLDQEGNRR HSSLIDIDSVPTYKWKRQVTQRNPVGQKKRKMSL LFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFVT HGCTVDNPVLERFISLFSVSQWVQLMILSKPTA PQRALVITHFVHVAEKLLQLQNFNTLMAVVGGLS HSSISRLKETHSHVSPETIKLWEGLTELVTATGNY GNYRRRLAACVGFRFPILGVHLKDLVALQLALPD WLDPARTRLNGAKMKQLFSILEELAMVTSLRPPV QANPDLLSLLTVSLDQYQTEDELYQLSLQREPRS KSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQAL VVEHIEKMVESVFRNFDVDGDGHISQEEFQIIRGN FPYLSAFGDLQDQDGCISREEMVSYFLRSSSVL GGRMGFVHNFQESNSLRPVACRHCKALILGIYKQ GLKCRACGVNCHKQCKDRLSVECRRAQSVSLE GSAPSPSPMHSHHHRAFSFSLPRPGRRGSRPPE IREEEVQTVEDGVFDIHL		(RASGRP2)
3	MGTLGKAREAPRKPSHGCRAASKARLEAKPANS PFPSHPSLAHITQFRMMVSLGHLAKGASLDDLID SCIQSFADAGNLCSNQLLQVMLTMHRIVISSAEL LQKVITLYKDALAKNSPGLCLKICYFVRYWITEFW VMFKMDASLTDTMEEFQELVKAKGEELHCRLIDT TQINARDWSRKL TQRIKSNTSKKRKVSL LFDHLE PEELSEHLTYLEFKSFRRISFSDYQNYLVNSCVKE NPTMERSIALCNGISQWVQLMVLSRPTPQLRAEV FIKFIQVAQKLHQLQNFNTLMAVIGGLCHSSISRLK ETSSHVPHEINKVLGEMTELLSSSRNYDNYRRAY GECTDFKIPILGVHLKDLISLYEAMPDYLEDGKVN VHKLLALYNHISELVQLQEVAAPPLEANKDLVHLLT LSLDLYYTEDEIYELSYAREPRNHRAPPLTPSKPP VVVDWASGVSPKPDPKTISKHVQRMVDSVFKNY DHDQDGYISQEEFEKIAASFPSFCVMDKDREGL ISRDEITAYFMRASSIYSKLGLGFPHNFQETTYLK PTFCDNCAGFLWGVIKQGYRCKDCGMNCHKQC KDLVVFECKKRAKNPVAPTENNTSVGPVSNLCSL	1-797 (whole protein)	GenBank AAC97349.1 or UniProtKB/Swiss -Prot: O95267 (RASGRP1)

	GAKDLLHAPEEGPFTFPNGEAVEHGEESKDRTIM LMGVSSQKISLRLKRAVAHKATQTESQPWIGSEG PSGPFVLSSPRKTAQDTLYVLPSTSPCPSPLV RKRAFVKWENKDSLKSKEELRHLRLPTYQELEQ EINTLKADNDALKIQLKYAQKKIESLQLEKSNHVLA QMEQGDCS		
4	MGSSGLGKAATLDELLCTCIEMFDDNGELDNSYL PRIVLLMHRWYLSSTELAELLCMYRNATGESCN EFRLKICYFMRYWILKFPAEFNLDGLIRMTEEFR EVASQLGYEKHVSLIDISSIPSYDWMRRVTQRKK VSKKGKACLLFDHLEPIELAEHLTFLEHKSFRISF TDYQSYVIHGCLENNPTLERSIALFNGISKWVQLM VLSKPTPQQRAEVITKFINVAKKLLQLKNFNTLMA VVGGLSHSSISRLKETHSHLSSEVTKNWNEMTEL VSSNGNYCNYRKAFADCDGFKIPILGVHLKDIAV HVIFPDWTEENKVNIVKMHQLSVTLSELVSLQNA SHHLEPNMDLINLLTSLDLYHTEDDIYKLSLVLEP RNSKSQPTSPTTPNKPVPVPLEWALGVMPKPDPT VINKHIRKLVESVFRNYDHDHGDYISQEDFESIAA NFPFLDSFCVLDKDQDGLISKDEMMAYFLRAKSQ LHCKMGPFGFIHNFQEMTYLKPTFCEHCAGFLWGI IKQGYKCKDCGANCHKQCKDLLVLACRRFARAP SLSSGHGSLPGSPSLPPAQDEVFEFPGVTAGHR DLDSRAITLVTGSSRKISVRLQRATTSQATQTEPV WSEAGWGDSGSHTFPKMKS FHDKAAKDKGFA KWENEKPRVHAGVDVDRGTEFELDQDEGEET RQDGEDG	1-690 (whole protein)	GenBank AAY15037.1 or UniProtKB/Swiss -Prot Q8IV61 (RASGRP3)
5	MNRKDSKRKSHQECTGKIGGRGRPRQVRRHKT CPSPREISKVMASMLGLLSEGGCSEDELLEKCI QSFDSAGSLCHEDHMLNMVLAMHSWVLPADLA ARLLTSYQKATGDTQELRRLQICHLVRYWLMRHP EVMHQDPQLEEVIGRFWATVAREGNSAQRR LGD SSDLLSPGGPGPPLPMSSPGLGKKRKVSLLFDHL ETGELAQHLYLEFRSFQAITPQDLRSYVLQGSV RGCPALEGSVGLSNSVSRWVQVMVLSRPGPLQ RAQVLDKFIHVAQRLHQLQNFNLTMAVTGGLCHS	1-673 (whole protein)	UniProtKB/Swiss -Prot Q8TDF6 or GenPept NP_733749.1 (RASGRP4)

	AISRLKDSHAHLSPDSTKALLELTELASHNNYAR YRRTWAGCAGFRLPVLGVHLKDLVSLHEAQPDR LPDGRHLPLKLNLYLRLQELVALQGQHPPCSAN EDLLHLLTSLDLFYTEDEIYELSYAREPRCPKSLP PSPFNAPLVVEWAPGVTPKPDRVTLGRHVEQLV ESVFKNYDPEGRGTISQEDFERLSGNFPFACHGL HPPPRQGRGSFSREELTGyllRASAICSKLGLAF LHTFHEVTFRKPTFCDCSCGFLWGVTKQGYRCR ECGLCCHKHCRDQVKVECKKRPgAKGDAGPPG APVPSTPAPHASCGSEENHSYTLSEPETGCQLR HAWTQTESPHPSWETDTVPCPVMdPPSTASSKL DS		
6	MAGTLDLDKGCTVEELLRGCIeAFDDSGKVRDP QLVRMFLMMHPWYIPSSQLAAKLLHIYQQSRKDN SNSLQVKtCHLVRYWISAFPAEFDLNPELAEQIKE LKALLDQEGNRRHSSLIDIDSVPTYKWKRQVTQR NPVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSF CKILFQDYHSFVTHGCTVDNPVLERFISLFNSVSQ WVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNF NTLMAVVGGLSHSSISRLKETHSHVSPETIKLWE GLTELVTATGNYGNYRRRLAACVGFRFPILGVHL KDLVALQLALPDWLDPARTRLNGAKMKQLFSILE ELAMVTSLRPPVQANPDLLSLLTVSLDQYQTEDE LYQLSLQREPRSKSSPTSPTCTPPPRPPVLEEW TSAAKPKLDQALVVEHIEKMVESVFRNFDVDGDG HISQEEFQIIRGNFPYLSAFGDLQNDGDCISREE MVSyFLRSSSVLGGRMGFVHNFQESNSLRPVAC RHCKALILGIYKQGLKCRACGVNCHKQCKDRLSV ECRRAQSVSLEGSAPSPSPMHSHHHRAFSFSL PRPGRRGSRPPEIREEEVQTVEDGVFDIHL	1-609 (whole protein)	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
7	MGTQRLCGRGTQGWPGSSEQHVQEATSSAGLH SGVDELGVRSEPGGRLPERSLGPAPAPAAMAG TLDLDKGCTVEELLRGCIeAFDDSGKVRDPQLVR MFLMMHPWYIPSSQLAAKLLHIYQQSRKDNSNSL QVKtCHLVRYWISAFPAEFDLNPELAEQIKELKAL LDQEGNRRHSSLIDIDSVPTYKWKRQVTQRNPV	1-671 (whole protein)	UniProtKB/Swiss -Prot Q7LDG7-2 or GenPept AAF07219.1 (RASGRRP2

	GQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKI LFQDYHSFVTHGCTVDNPVLERFISLFNSVSQWV QLMILSKPTAPQRALVITHFVHVAEKLLQLQNFNT LMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLT ELVTATGNYGNYRRRLAACVGRFPILGVHLKDL VALQLALPDWLDPARTRLNGAKMKQLFSILELA MVTSLRPPVQANPDLLSLLTVSLDQYQTEDELYQ LSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTS AKPKLDQALVVEHIEKMVESVFRNFDVDGDGHIS QEEFQIIRGNFPYLSAFGDLDQNQDGCISREEMV SYFLRSSSVLGGRMGFVHNFQESNSLRPVACRH CKALILGIYKQGLKCRACGVNCHKQCKDRLSVEC RRRAQSVSLEGSAPSPSPMHSHHHRAFSFSLPR PGRGRSRPPEIREEEVQTVEDGVFDIHL		Isoform 2)
8	MAGTLDLDKGCTVEELLRGCIEAFDDSGKVRDP QLVRMFLMMHPWYIPSSQLAAKLLHIYQQSRKDN SNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKE LKALLDQEGNRRHSSLIDIDSVCVGAEHRGLGGH SVSYTICA	1-144 (whole protein)	UniProtKB/Swiss -Prot Q7LDG7-3 (RASGRRP2 Isoform 3)
9	MAGTLDLDKGCTVEELLRGCIEAFDDSGKVRDP QLVRMFLMMHPWYIPSSQLAAKLLHIYQQSRKDN SNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKE LKALLDQEGNRRHSSLIDIDSVPYKWKQVTR NPVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSF CKILFQDYHSFVTHGCTVDNPVLERFISLFNSVSQ WVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNF NTLMAVVGGLSHSSISRLKETHSHVSPETIKLWE GLTELVTATGNYGNYRRRLAACVGRFPILGVHL KDLVALQLALPDWLDPARTRLNGAKMKQLFSILE ELAMVTSLRPPVQANPDLLSLLTVSLDQYQTEDE LYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEW TSAAPKLDQALVVEHIEKMVESVFRNFDVDGDG HISQEEFQIIRGNFPYLSAFGDLDQNQDGCISREE MVSYFLRSSSVLGGRMGFVHNFQESNSLRPVAC RHCKALILGIYKQGLKCRACGVNCHKQCKDRLSV ECRRRAQSVSLEGSAPSPSPMHSHHHRAFSFSL	1-610 (whole protein)	UniProtKB/Swiss -Prot Q7LDG7-4 or GenPept XP_011543022.1 (RASGRRP2 Isoform 4)

	PRPGRRGSRPPAEIREEEVQTVEDGVFDIHL		
10	MGEPQGSMRILVTGG	1-15	GenBank AAH93061.1 (GDP L-fucose synthase)
11	VVADGAGLPGEDWVF	26-40	GenBank AAH93061.1 (GDP L-fucose synthase)
12	TAQTRALFEKVQPTH	51-65	GenBank AAH93061.1 (GDP L-fucose synthase)
13	LFRNIKYNLDFWRKN	76-90	GenBank AAH93061.1 (GDP L-fucose synthase)
14	VHMNDNVLHSAFEVG	91-105	GenBank AAH93061.1 (GDP L-fucose synthase)
15	DNVLHSAFEV	95-104	GenBank AAH93061.1 (GDP L-fucose synthase)
16	NVLHSAFEVG	96-105	GenBank AAH93061.1 (GDP L-fucose synthase)
17	NVLHSAFEVGARKVV	96-110	GenBank AAH93061.1 (GDP L-fucose synthase)
18	VLHSAFEVGA	97-106	GenBank AAH93061.1 (GDP L-fucose

			synthase)
19	KTTYPIDETMIHNGP	121-135	GenBank AAH93061.1 (GDP L-fucose synthase)
20	IHNGPPHNSNFGYSY	131-145	GenBank AAH93061.1 (GDP L-fucose synthase)
21	PHNSNFGYSYAKRMI	136-150	GenBank AAH93061.1 (GDP L-fucose synthase)
22	AYFQQYGCTFTAVIP	156-170	GenBank AAH93061.1 (GDP L-fucose synthase)
23	YGCTFTAVIPTNVFG	161-175	GenBank AAH93061.1 (GDP L-fucose synthase)
24	LFIWVLREYNEVEPI	226-240	GenBank AAH93061.1 (GDP L-fucose synthase)
25	LREYNEVEPIILSVG	231-245	GenBank AAH93061.1 (GDP L-fucose synthase)
26	EVEPIILSVGEEDEV	236-250	GenBank AAH93061.1 (GDP L-fucose synthase)
27	ILSVGEEDEVSIKEA	241-255	GenBank AAH93061.1 (GDP L-fucose

			synthase)
28	EEDEVSIKEAAEAVV	246-260	GenBank AAH93061.1 (GDP L-fucose synthase)
29	SIKEAAEAVVEAMDF	251-265	GenBank AAH93061.1 (GDP L-fucose synthase)
30	AEAVVEAMDFHGEVT	256-270	GenBank AAH93061.1 (GDP L-fucose synthase)
31	FDTTKSDGQFKKTAS	271-285	GenBank AAH93061.1 (GDP L-fucose synthase)
32	FRFTPFKQAVKETCA	296-310	GenBank AAH93061.1 (GDP L-fucose synthase)
33	LVRWISAFP	131-140 or 78-87	GenBank AAI10307.1 (RASGRP2) or UniProtKB/Swiss -Prot Q7LDG7-1 (RASGRP2 Isoform 1)
34	FVRYWITEFW	128-137	GenBank AAC97349.1 (RASGRP1)
35	FMRYWILKFP	77-86	GenBank BAA74869.3 (RASGRP3)
36	LLFDHLEPMELAEHLTYLEYRSF	148-170	UniProtKB/Swiss

			-Prot Q7LDG7-1 (RASGRP2 Isoform 1)
37	NFNTLMAVVGGLSHSSISRLKETHSHVS	239-266	UniProtKB/Swiss -Prot Q7LDG7-1 (RASGRP2 Isoform 1)
38	PAAMAGTLDLDKGCT	1-12 and the additional amino acids PAA at the N-terminal end	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
39	DKGCTVEELLRGCIE	8-22	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
40	RGCIEAFDDSGKVRD	18-32	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
41	GKVRDPQLVRMFLMM	28-42	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
42	MFLMMHPWYIPSSQL	38-52	UniProtKB/Swiss

			-Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
43	PSSQLAAKLLHIYQQ	48-62	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
44	HIYQQSRKDNSNSLQ	58-72	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
45	SNSLQVKTCHLVRYW	68-82	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
46	LVRYWISAFPAEFDL	78-92	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
47	AEFDLNPELAEQIKE	88-102	UniProtKB/Swiss -Prot Q7LDG7-1

			or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
48	EQIKELKALLDQEGN	98-112	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
49	DQEGNRRHSSLIDID	108-122	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
50	LIDIDSVPTYKWKRQ	118-132	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
51	KWKQVQTQRNPVGQK	128-142	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
52	PVGQKKRKMSLLFDH	138-152	UniProtKB/Swiss -Prot Q7LDG7-1 or

			GenPept NP_722541.1 (RASGRRP2 Isoform 1)
53	LLFDHLEPMELAEHL	148-162	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
54	LAEHLTYLEYRSFCK	158-172	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
55	RSFCKILFQDYHSFV	168-182	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
56	YHSFVTHGCTVDNPV	178-192	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
57	VDNPVLERFISLFNS	188-202	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept

			NP_722541.1 (RASGRRP2 Isoform 1)
58	SLFNSVSQWVQLMIL	198-212	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
59	QLMILSKPTAPQRAL	208-222	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
60	PQRALVITHFVHVAE	218-232	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
61	VHVAEKLLQLQNFNT	228-242	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
62	QNFNTLMAVVGGLSH	238-252	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1

			(RASGRRP2 Isoform 1)
63	GGLSHSSISRLKETH	248-262	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
64	LKETHSHVSPETIKL	258-272	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
65	ETIKLWEGLTELVTA	268-282	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
66	ELVTATGNYGNYRRR	278-292	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
67	NYRRRLAACVGFRFP	288-302	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2

			Isoform 1)
68	GFRFPILGVHLKDLV	298-312	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
69	LKDLVALQLALPDWL	308-322	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
70	LPDWLDPARTRLNGA	318-332	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
71	RLNGAKMKQLFSILE	328-342	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
72	FSILEELAMVTSLRP	338-352	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)

73	TSLRPPVQANPDLLS	348-362	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
74	PDLLSLLTVSLDQYQ	358-372	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
75	LDQYQTEDELYQLSL	368-382	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
76	YQLSLQREPRSKSSP	378-392	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
77	SKSSPTSPTSCTPPP	388-402	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
78	CTPPPRPPVLEEWS	398-412	UniProtKB/Swiss

			-Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
79	EEWTSAAKPKLDQAL	408-422	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
80	LDQALVVEHIEKMVE	418-432	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
81	EKMVESVFRNFDVDG	428-442	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
82	FDVDGDGHISQEEFQ	438-452	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
83	QEEFQIIRGNFPYLS	448-462	UniProtKB/Swiss -Prot Q7LDG7-1

			or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
84	FPYLSAFGDLQDQD	458-472	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
85	DQNQDGCISREEMVS	468-482	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
86	EEMVSYFLRSSVLG	478-492	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
87	SSVLGGRMGFVHNFQ	488-502	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
88	VHNFQESNSLRPVAC	498-512	UniProtKB/Swiss -Prot Q7LDG7-1 or

			GenPept NP_722541.1 (RASGRRP2 Isoform 1)
89	RPVACRHCKALILGI	508-522	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
90	LILGIYKQGLKCRAC	518-532	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
91	KCRACGVNCHKQCKD	528-542	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
92	KQCKDRLSVECRRA	538-552	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
93	CRRRAQSVSLEGSAP	548-562	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept

			NP_722541.1 (RASGRRP2 Isoform 1)
94	EGSAPSPSPMHSHHH	558-572	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
95	HSHHHRAFSFSLPRP	568-582	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
96	SLPRPGRRGSRPPEI	578-592	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
97	RPPEIREEEVQTVED	588-602	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1 (RASGRRP2 Isoform 1)
98	EEVQTVEDGVFDIHL	595-609	UniProtKB/Swiss -Prot Q7LDG7-1 or GenPept NP_722541.1

			(RASGRP2 Isoform 1)
--	--	--	------------------------

The following gene sequences (Table 2) represent the gene sequences of the proteins GDP-L fucose synthase (gene: TSTA3), RASGRP1, RASGRP2, RASGRP3 and RASGRP4 (Table 2). The proteins or derivatives or splice variants thereof are preferably encoded by the respective

5

Table 2: Gene sequences and database entry

Gene name	Database entry
TSTA3	NCBI Reference Sequence: NC_000008.11 REGION: 143612618..143618048
RASGRP1	NCBI Reference Sequence: NC_000015.10 REGION: 38488101..38565575
RASGRP2	NCBI Reference Sequence: NC_000011.10 REGION: 64726911..64745456
RASGRP3	NCBI Reference Sequence: NC_000002.12 REGION: 33436324..33564750
RASGRP4	NCBI Reference Sequence: NC_000019.10 REGION: 38409051..38426305

The following nucleotide sequences (Table 3) represent preferred nucleotide sequences encoding any of the proteins or fragment, derivative or splice variant thereof of the present invention. The table also comprises coding sequences (CDS), i. e. proteins or peptides, which can also be used in the treatment, diagnosis and/or prevention of multiple sclerosis (MS).

10

Table 3: Nucleotide sequences and protein sequences and the respective database entries (Genbank Accession Nos)

15

Gene name	Database entry
TSTA3	RNA: XM_011517269.1 NM_003313.3 NM_001317783.1 XM_005251051.3 protein: XP_011515571.1 NP_003304.1

	NP_001304712.1 XP_005251108.2
RASGRP1	RNA: NM_001306086.1 NM_001128602.1 NM_005739.3 XM_017021860 XM_005254114.3 XM_011521151.3 XR_001751485.2 XR_001751486.2 protein: NP_001293015.1 NP_001122074.1 NP_005730.2 XP_016877349.1 XP_005254171.1 XP_011519453.1
RASGRP2	RNA: XM_011544722.2 NM_153819 XM_011544723.3 XM_011544721.1 NM_001098670.1 XM_017017084.2 XM_011544720.2 XM_011544718.2 XM_017017082.2 XM_005273707.4 XM_017017083.2 NM_001318398.1 NM_001098671.1 XM_011544725.2 XM_017017085.2 XM_017017086.1 XR_001747720.2

	XR_001747719.2 protein: XP_011543024.1 NP_722541.1 XP_011543025.1 XP_011543023.1 NP_001092140.1 XP_016872573.1 XP_011543022.1 XP_011543020.1 XP_016872571.1 XP_005273764.3 XP_016872572.1 NP_001305327.1 NP_001092141.1 XP_011543027.1 XP_016872574.1 XP_016872575.1
RASGRP3	RNA: NM_001349975.1 NM_170672.2 NM_001349978.1 XM_017003759.2 XM_011532746.3 XM_011532748.3 NM_001349979.1 NM_001349977.1 NM_001349980.1 NM_001139488.1 NM_001349981.1 NM_001349976.1 XM_017003761.2 NM_015376.2 XR_001738693.2 XR_001738692.2 XR_001738691.2

	NR_146505.1 protein: NP_001336904.1 NP_733772.1 NP_001336907.1 XP_016859248.1 XP_011531048.1 XP_011531050.2 NP_001336908.1 NP_001336906.1 NP_001336909.1 NP_001132960 NP_001336910.1 NP_001336905.1 XP_016859250.1 NP_056191.1
RASGRP4	RNA: NM_001146202.1 NM_001146206.1 NM_001146207.1 NM_001146205.1 NM_001146203.1 NM_001146204.1 NM_170604.2 XR_935732.2 protein: NP_001139674.1 NP_001139678.1 NP_001139679.1 NP_001139677.1 NP_001139675.1 NP_001139676.1 NP_733749.1

All sequences were retrieved from the respective online databases on June 22, 2018.

In one embodiment, the protein, fragment, derivative or splice variant according to the present invention may be used for identifying a human subject who is suitable for tolerization to autoantigens in MS, preferably early MS. Identifying a human subject, who is suitable for tolerization to autoantigens in MS, preferably early MS, preferably comprises measuring positive reactivity of the T cells and/or antibodies to the autoantigens in the human subject. Thereby, the human subject can also be *in vitro* diagnosed with MS. In other words, the identified autoantigens can also be used in the *in vitro* diagnosis of MS.

In vitro diagnosis of MS preferably comprises the following steps: isolating T cells, preferably CD4+ T cells, and/or antibodies from blood, CSF or other body fluid of the subject and measuring reactivity of the T cells and/or antibodies against a protein or fragment, derivative or splice variant thereof according to the present invention. The person skilled in the art is aware of methods for isolating T cells and/or antibodies from blood, CSF or other body fluid of the subject and measuring reactivity of the T cells and/or antibodies against the protein, fragment, derivative or splice variant. A reactivity of the T cells, preferably CD4+ T cells, and/or antibodies to the tested protein or fragment, derivative or splice variant thereof, may indicate that the subject suffers from MS. The diagnosis can also be combined with clinical and imaging findings, i. e. the MS diagnosis according to the state of the art, in particular according to the revised McDonald criteria.

In another embodiment, a protein, fragment, derivative or splice variant according to the present invention may be used for distinguishing between MS subgroups. In particular, the protein, fragment, derivative or splice variant according to the present invention may be used for diagnosing pattern II MS in a human subject.

The following characteristics indicate that a certain peptide of a protein is immunodominant in the context of MS:

a) frequent recognition of this peptide by T cells, i. e. by approximately 10% or more of MS patients, often in the context of a disease-associated HLA allele or haplotype (Sospedra and Martin, 2005), and

b) recognition of this peptide by disease-relevant T cells such as those that respond to peptides at low concentrations (high avidity T cells) (Bielekova et al., 2004) and are therefore considered particularly dangerous, and/or have a proinflammatory phenotype, and/or are isolated from the

target organ or compartment (CNS), in the case of MS, brain-, spinal cord- or CSF-infiltrating T cells.

5 However, high avidity recognition is not a prerequisite, since low-avidity myelin-specific T cells have also been shown to be pathogenic in humanized transgenic mouse models (Quandt et al. 2012).

10 Thus, it can be tested whether a protein or fragment, derivative or splice variant thereof is immunodominant in the context of MS. Such a test is preferably an *in vitro* test. Particularly suitable is an *in vitro* test that allows measuring the reactivity of T cells and/or antibodies obtained from the blood, CSF or other body fluid of a human subject that had been diagnosed with MS, preferably CSF-infiltrating CD4⁺ T cells, to the tested protein or fragment, derivative or splice variant. The person skilled in the art is aware of methods testing the reactivity of T cells, preferably CD4⁺ T cells, and/or antibodies. For example, the proliferation of CD4⁺ T cells and/or
15 their secretion of IFN- γ or reactivity in a ELISPOT/FLUOROSPOT assay or reactivity against HLA-peptide tetramers can be tested. If the tested protein or fragment, derivative or splice variant thereof induces reactivity in a human subject that had been diagnosed with MS, in case of T cell reactivity in particular a stimulatory index (SI) above 2 and/or an IFN- γ secretion above 20 pg/ml, the tested protein or fragment, derivative or splice variant may be termed
20 immunodominant. It is also possible to select 10 patients who had been diagnosed with MS for such a test. If reactivity is induced in at least 2 patients, the tested protein or fragment, derivative or splice variant may be termed immunodominant. Preferably, the 10 patients have been diagnosed with RRMS according to the established revised McDonald criteria.

25 It has recently been demonstrated that T cells of MS patients show increased *in vitro* proliferation in the absence of an exogenous antigen (Mohme et al., 2013; Jelcic et al., 2018). These "autoproliferating" T cells are enriched for cells that home to the CNS compartment of MS patients and can thus be considered as a peripheral blood source of brain-/CSF-infiltrating T cells.

30 In the case that data from testing T cells *in vitro* is not available or in addition to such testings, immune recognition of peptides can also be predicted/inferred from those peptides that will bind well to the HLA-class I or -class II alleles of the individual and for CD8⁺ and CD4⁺ T cells respectively. Peptide binding predictions are well known to the skilled person. They can be
35 performed by well-established prediction algorithms (NetMHCII -

www.cbs.dtu.dk/services/NetMHCII/; IEDB - www.iedb.org/) and analysis of the HLA-binding motifs (SYFPEITHI - www.syfpeithi.de/), cf. "Examples", under 11.

According to the present invention the proteins GDP-L-fucose synthase and a member of the
5 RASGRP protein family, in particular RASGRP2, have been identified as immunodominant in MS, thus they have been identified as autoantigens.

It is not necessary that the binding to the HLA allele is particularly strong. In fact, immunodominance can also occur for peptides that bind poorly to the HLA allele (Muraro et al.,
10 1997, J Clin Invest; 100(2):339-349).

Tolerance induction is antigen-specific and renders autoreactive T cells non-functional or anergic or induces Treg cells that specifically suppress untoward autoimmunity to said target antigens. The induction of tolerance to target autoantigens is a highly important therapeutic goal
15 in autoimmune diseases. It offers the opportunity to attenuate specifically the pathogenic autoimmune response in an effective way with few side effects. Tolerance induction can also be achieved by applying a whole protein instead of or in addition to an immunodominant peptide being a fragment of the protein (Kennedy MK et al., 1990). It has herein been shown that immunologic changes consistent with tolerance can be induced in human patients upon i.v.
20 injection of immunodominant peptides coupled to red blood cells (14. of Examples). Hence, immunodominant peptides can be used for tolerance induction.

The immunodominance of the proteins and/or fragments thus allows using the protein and/or a fragment, derivative or splice variant thereof for antigen-specific immunotherapies such as
25 tolerance induction.

According to the present invention, antigen-specific tolerization can be used in all forms of MS: At the time of first manifestation, when differential diagnoses have been excluded, the disease is referred to as CIS provided that the CSF and MRI findings are consistent with the diagnosis.
30 MRI discloses lesions in locations typical for MS, i.e. juxtacortical, periventricular, in the brain stem or spinal cord. If certain criteria are fulfilled that can be summarized as dissemination in space (more than one lesion or clinical symptom/sign) and time (more than one event) then the diagnosis of RRMS can be made. A special scenario is the accidental discovery of MRI lesions compatible with MS without clinical symptoms. This is referred to as RIS and can be considered
35 a pre-stage of CIS and RRMS. More than 80% of patients suffer from one of these, and the majority of patients develops later what is called SPMS. At this time, relapses/exacerbations

become less frequent or stop altogether and neurological disability increases steadily either between relapses or without these.

5 A special form of MS is PPMS, which never shows relapses, but rather begins with steady worsening of neurological symptoms, e.g. of the ability to walk. PPMS affects approximately 10% of MS patients and males and females with equal frequency. Its onset is usually later than CIS or RRMS. With respect to causes and disease mechanisms PPMS is considered similar to the above RIS-CIS-RRMS-SPMS.

10 Typically, MS is diagnosed according to the revised McDonald criteria. These criteria also allow distinguishing between the different forms and disease activity of MS (Thompson et al., 2018, Lancet Neurol, 17(2):162-173).

15 Preferably, the tolerization approach is applied at an early stage, i.e. RIS, CIS and early RRMS, since it is assumed that the immune processes at this stage are primarily mediated by autoreactive T lymphocytes, while tissue damage, so-called degenerative changes, become gradually more important when the disease advances. However, tolerization is meaningful as long as there is an autoreactive T cell response against the antigens used for tolerization, which could also be during SPMS and PPMS.

20 In a particularly preferred embodiment, a GDP-L-fucose synthase or RASGRP2 protein or splice variant thereof, preferably the GDP-L-fucose synthase or RASGRP2 protein, is used in a tolerization approach at an early stage, i.e. RIS, CIS and early RRMS. The GDP-L-fucose synthase protein has, for example, the sequence as set forth in SEQ ID NO: 1, the RASGRP2 protein has, for example the sequence as set forth in SEQ ID NO: 2, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 or SEQ ID NO: 9.

30 In one aspect of the invention, a method for inducing antigen-specific tolerance to autoantigens in a human subject suffering from or at risk of developing MS is provided. The method comprises the step of applying to a patient in the need thereof, i. e. to the human subject, at least one protein selected from the group consisting of GDP-L-fucose synthase and members of the RASGRP protein family, fragment (peptide), derivative and/or splice variant thereof, a nucleotide sequence encoding any of the proteins or fragment, derivative or splice variant thereof and/or a gene sequence as described herein or applying at least one carrier comprising
35 at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence as described herein.

In one embodiment, the nucleotide sequence or gene sequence is applied to the patient via a carrier, e.g. a cell. The antigen may then be expressed by a carrier, e.g. a cell. Transfer of autoantigen-encoding RNA/DNA into a carrier, e.g. a cell, and thus encoding the above autoantigens is also conceivable similar to tumor vaccination approaches that employ antigen-
5 encoding RNAs.

It is particularly preferred to use the whole protein of GDP-L-fucose synthase (SEQ ID NO: 1) or the whole protein of RASGRP2 (SEQ ID NO: 2 or SEQ ID NO: 6, 7, 8 or 9) for inducing antigen-specific tolerance. In another preferred embodiment, a fragment of any of these proteins is
10 used. It is particularly preferred to use a fragment as set forth in any of SEQ ID NOs: 10 to 98. Even more preferred is a peptide as set forth in any of SEQ ID NOs: 10 to 35.

The at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence may be applied by nasal, inhaled, oral, subcutaneous (s.c.), intracoelomic (i.c),
15 intramuscular (i.m.), intradermal (i.d.), transdermal (t.d.) or intravenous (i.v.) administration, preferably by routes of administration that are considered tolerogenic, for example by i.v., s.c., i.d., t.d., oral, inhaled, nasal or coupled to a tolerogenic carrier, preferably an RBC.

In particular, the method may be used for inducing antigen-specific tolerance to autoantigens in
20 early MS or even pre-clinical stages of the disease.

The antigen-specific tolerance protocol provided herein may selectively target both activated and naïve autoreactive T cells specific for multiple potential encephalitogenic epitopes that perpetuate the disease.
25

The tolerization approach can also be used to prevent MS. This approach may include identifying those individuals (e.g. in a family with a MS patient), who are at a high risk of developing MS. For example, it is possible to tolerize e.g. the children of a mother with MS or the identical twin of a patient with MS, in whom the risk of developing MS would be particularly
30 high.

Diagnosis of MS or one of its forms is made by demonstrating neurological deficits and/or MRI lesions compatible with MS that are disseminated in space and time. Positive laboratory testing for an autoimmune response against the novel target proteins GDP-L-fucose synthase and the
35 RASGRP family and/or fragments, derivatives and/or splice variants thereof can be used to identify patients particularly likely to profit from antigen-specific tolerance induction, i.e. allow

personalizing antigen-specific tolerance approaches. By identifying patients who particularly likely profit from antigen-specific tolerance induction, a patient can be *in vitro* diagnosed with MS. In other words, the identified autoantigens GDP-L-fucose synthase and the RASGRP family (in particular RASGRP2) and/or fragments, and/or derivatives and/or splice variants thereof can
5 be used in the *in vitro* diagnosis of MS. Thus, this *in vitro* testing can be complemented with clinical and imaging findings, i. e. the MS diagnosis according to the state of the art, in particular according to the revised McDonald criteria.

In another aspect of the invention, a method for identifying a human subject suitable for
10 tolerization to autoantigens in MS, preferably early MS, is provided. Thereby, an individual patient, who would profit from or is eligible for the treatment, is identified. This method can also be used for *in vitro* diagnosing MS. Patients with CIS, probably also RIS, and RRMS are best suited for tolerization, although tolerization appears meaningful in any form of MS including SPMS and PPMS as long as the patient responds at least to one of the antigenic peptides that
15 are included in the tolerization treatment. The steps of this method comprise isolating T cells, preferably CD4+ T cells, and/or antibodies from blood, CSF or other body fluid of the subject and measuring reactivity of the T cells and/or antibodies against the protein or fragment, derivative or splice variant thereof according to the present invention. The person skilled in the art is aware of methods for isolating T cells and/or antibodies from blood, CSF or other body
20 fluid of the subject and measuring reactivity of the T cells and/or antibodies against the protein, fragment, derivative or splice variant.

If an individual has a response against one of the protein/s or fragment/s he/she could thereby be *in vitro* diagnosed with MS and would be eligible and a good candidate for treatment. A
25 further selection step could be HLA-class II typing and presence of the MS-associated HLA-DR alleles.

Proteins, in particular the GDP-L-fucoase synthase protein or a protein of the RASGRP family, in particular RASGRP2, or fragments, derivatives or splice variants thereof derived from
30 disease-relevant antigens can thus be used to identify those patients or patient subgroups with an existing and/or particularly strong proinflammatory (potentially harmful) T cell or antibody response against the respective autoantigen. By identifying those patients or patient subgroups with an existing and/or particularly strong proinflammatory (potentially harmful) T cell or antibody response against the respective autoantigen, the patient can thus be *in vitro* diagnosed with
35 MS. In this situation, pretesting the patient with a suitable test to assess reactivity against the autoantigen would also allow tailoring the tolerizing treatment (e.g. composition of the

peptides/proteins used for tolerization) to the individual patient or patient subgroup with the aim to render the tolerization as specific as possible and also to avoid potential adverse effects. Antigen-specific tolerization, however, can also be performed in patients in whom a T cell response to the tolerizing antigen has not been shown. Thus, in one embodiment, the protein,
5 fragment, derivative and/or splice variant according to the invention can be used in the *in vitro* pre-testing of a human subject diagnosed with MS or a human subject at risk to develop MS.

In one aspect of the invention, a carrier is provided, which comprises at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence as described
10 herein. The carrier may be coupled to the at least one protein, fragment, derivative and/or splice variant and/or the carrier may contain the at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence. In one embodiment, the term "contains" means that the protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence is inside the carrier and not on its surface. A protein, fragment, derivative and/or
15 splice variant may also be coupled to and contained within the carrier, which means that one part of the protein, fragment, derivative and/or splice variant is coupled to the carrier and another part of the protein, fragment, derivative and/or splice variant is contained within the carrier.

20 The person skilled in the art is familiar with possible carriers. For example, the carrier may be any cell, protein, lipid, glycolipid, bead, nanoparticle, virus-like-particle (VLP), or molecule, such as a sugar molecule, or any combination thereof that is suitable for application in humans and to which protein/s and/or fragment/s can be coupled by a coupling process, e. g. by a chemical coupling process, preferably by EDC. The carrier can be derived from one existing in nature or
25 be synthetic. Preferably, the cell, molecule, bead, nanoparticle, or VLP is biodegradable *in vivo* or is at least applicable to living persons and broken down *in vivo* or is eliminated from the body to which the carrier is applied. The term cell also includes cell precursors, e.g. RBC precursors. Preferably, the carrier is a blood cell, even more preferably a red or white blood cell. The white blood cell may be a splenocyte or a PBMC or generally an APC.

30 In one embodiment, the protein, fragment, derivative and/or splice variant is expressed by the cell, preferably the blood cell. Thereby, the genetic information encoding the protein, fragment, derivative and/or splice variant is introduced into the cell before the protein, fragment, derivative and/or splice variant is expressed by the cell.

35

Any coupling agent or method for coupling a protein and/or a fragment thereof to a carrier may be used. For example, a synthetic or natural linker may be employed for coupling. One example of such a linker is glycophorin A, present on the surface of RBC. In one embodiment, chemical crosslinking is performed. In a preferred embodiment, the chemical crosslinker EDC catalyzing the formation of peptide bonds between free amino and carboxyl groups is used. Particularly in the presence of EDC, multiple peptides can be coupled to the surface of the carrier thereby allowing for the simultaneous targeting of multiple T cell specificities. Preferably, more than three, more than five, more than 10, more than 15, or even more than 20 different peptides are coupled to the surface of the carrier. In a preferred embodiment, between five and 20, preferably between five and 15 different peptides are used. A peptide is different from another peptide if it does not consist of the same amino acid sequence. The carrier is preferably, but not necessarily a cell. EDC can be used for coupling to any carrier as long as a free amino group is present.

In another embodiment, at least one peptide of each protein according to the present invention (a GDP-L-fucose synthase protein or a protein of the RASGRP protein family, preferably RASGRP2) are used in combination and coupled to a carrier. It is particularly advantageous to combine at least one peptide of each protein according to the present invention with at least one known peptide of the state of the art, in particular with at least one myelin peptide, in particular with at least one or all of the myelin peptides as defined by SEQ ID NOs: 261 to 267.

In one preferred embodiment, the carrier is a blood cell, and the blood cell is chemically coupled by a coupling agent, preferably by EDC, to the at least one protein, fragment, derivative and/or splice variant. A method of manufacturing such a chemically coupled, i. e. antigen-coupled blood cell is also provided, comprising isolating the blood cell from a human subject, adding the at least one protein, fragment, derivative and/or splice variant, i. e. the antigen, and subsequently adding the coupling agent, preferably EDC.

The mechanism of action of cells coupled with peptide by EDC is not fully understood, but involves covalently linking amino- and carboxy groups of peptide and cell surface molecules, subsequent programmed cell death (apoptosis in the case of nucleated cells; eryptosis in the case of RBC) of the peptide-coupled, i. e. antigen-coupled, cells and then tolerogenic presentation of the dying cells in vivo.

The dose of EDC used for the coupling reaction may be titrated to obtain a maximum of safety with best efficacy. In high concentrations, EDC may lead to lysis of cells, in particular of RBC.

For optimal stability of the RBC, a final concentration of EDC of less than 15 mg/ml, preferably less than 10 mg/ml, even more preferably less than 5 mg/ml, even more preferably of about 3 mg/ml may be used. The optimal dose may also vary. The person skilled in the art knows how to determine the optimal stability of the RBC and the optimal dose of EDC.

5

The protein, fragment, derivative and/or splice variant to be coupled may be added in an amount to be readily determined by the person skilled in the art. The person skilled is aware of measures to determine the optimal amount in the interplay with an optimal amount of EDC.

10 The incubation time can be varied and tailored for each specific coupling reaction (for example, 15 min, 30 min, 45 min, 60 min, 120 min). Coupling efficiency may be better with longer time of incubation. In one embodiment, a maximum is reached after 60 min.

The incubation temperature may also vary. For example, 15-25°C or 2-8°C may be used. In one
15 embodiment, coupling efficiency was higher when the coupling reaction was performed at 15-25°C.

Any excipient that allows the coupling reaction may be used. In one embodiment, the excipient is sterile and endotoxin free. In a preferred embodiment, the excipient is sterile, endotoxin free
20 saline (NaCl 0.9%). Saline is approved for use in humans and provides a maximum of safety.

The person skilled in the art knows how to determine an optimal incubation time and temperature and how to determine possible excipients.

25 In one aspect of the invention, a pharmaceutical composition is provided which comprises at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence as described herein and a pharmaceutically acceptable carrier.

In another aspect of the invention, a method for the therapeutic treatment, the prevention or
30 diagnosis of MS in a human subject is provided comprising administering into said subject a therapeutically effective amount of the protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence as described herein and/or the carrier as described herein. Thus, the carrier comprising the at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence according to the present invention, in particular the
35 carrier coupled to the at least one protein, fragment, derivative and/or splice variant according to the invention, and/or the carrier containing the at least one protein, fragment, derivative, splice

variant, nucleotide sequence and/or gene sequence according to the invention can be used in the treatment, diagnosis and/or prevention of MS.

5 In a further aspect of the invention, a peptide according to the present invention, preferably a peptide, which comprises 5 to 50, preferably 5 to 20, more preferably 10 to 15, even more preferably 15 amino acids, is used as a medicament.

A preferred peptide for use as a medicament is

- 10 a) at least 85 %, preferably at least 90 %, more preferably at least 95 % identical to a respective corresponding amino acid sequence or
- b) at least 70 %, preferably at least 80 %, more preferably at least 90 % homologous to a respective corresponding amino acid sequence or
- 15 c) at least 60 %, preferably at least 70 %, more preferably at least 80 %, even more preferably at least 90 % homologous to a respective corresponding amino acid sequence and binds to an autologous HLA allele, is recognized by a T cell and/or is recognized by an antibody which binds to or recognizes the respective amino acid sequence.

Even more preferred for use as a medicament is a peptide with a sequence selected from the group comprising SEQ ID NOs: 10 to 98, preferably SEQ ID NOs: 10 to 35, even more

20 preferably a peptide which consists of a sequence selected from the group comprising SEQ ID NO: 10 to 98, preferably SEQ ID NOs: 10 to 35.

As a further aspect, the peptide motives of Figure 3 (ii) E (SEQ ID NO: 99), the GDP-L-fucose synthase peptides as depicted in Figure 5 (SEQ ID NOs: 199-215), the GDP-L-fucose synthase

25 peptides of the table in Figure 6 (I and II) (SEQ ID NOs: 10-14, 19-20, 22-23, 25-32 and 216-290) and the RASGRP1, RASGRP2 and RASGRP3 peptides of the table in Figure 12 (SEQ ID NOs: 291-309) are all subject-matter of the present invention and can be used according to the present invention.

30 **EXAMPLES**

GDP-L-Fucose Synthase Specificity

1. Characterization of TCC21.1

The unbiased target antigen identification approach using the positional scanning synthetic

35 combinatorial libraries (ps-SCL) is schematically shown in Fig. 2. The clone TCC21.1 has been identified as described in Planas et al. (2015). TCC21.1 is a CD4⁺ TCC isolated from CSF-

infiltrating cells and clonally expanded in two active white matter demyelinating MS lesions (LI and LIII) from an SPMS patient (1154SA) with pattern II demyelination. TCC21.1 releases Th2 cytokines and helps proliferation and antibody production by autologous B cells upon unspecific activation. In order to study the peptides recognized by this TCC of unknown specificity using a

5 decapeptide positional scanning library, at first the HLA class II molecule used by TCC21.1 to recognize the peptide mixtures was identified. This is a prerequisite for using the peptide library/biometrical analysis approach for unbiased identification of target antigens (Zhao et al., 2001, J Immunol, 167:2130-2141; Sospedra et al., 2010, J Immunol Methods, 353(1-2):93-101). Since patient 1154SA was homozygous for the DR15 haplotype, TCC21.1 was initially tested

10 with the mixtures with amino acids (AAs) defined at position 5, presented by EBV immortalized bare lymphocyte syndrome (BLS) B cell line (BCL) cells transfected with different single autologous HLA DR/DQ molecules (DRA*01:01/DRB5*01:01 = DR2a, DRA*01:01/DRB1*15:01 = DR2b, and DQA1*01:02/DQB1*06:02). As readout of T cell activation GM-CSF production was used, since TCC21.1 releases this cytokine after unspecific stimulation with anti-CD3 and

15 PMA (Planas et al., 2015). Only mixtures presented by BCL expressing DRB1*15:01 (DR2b) class II molecules were stimulatory (data not shown).

2. Positional scanning synthetic combinatorial library as a method for identifying antigens / identification of GDP-L-fucose synthase

20 TCC21.1 was then tested with the complete decapeptide positional scanning library (200 mixtures; i.e. mixtures representing the 10 positions of a 10-mer and in each of the fixed positions one of the 20 L-amino acids), presented by BCL expressing DRB1*15:01 class II molecules. GM-CSF release in response to the complete library is shown in Fig. 3A. A similar response pattern was obtained with IL-10 release. Next, a biometrical analysis scoring matrix

25 was generated by assigning numerical values to the stimulatory potential of each of the 20 defined AAs in each of the ten positions of the decapeptide library as previously described (Zhao et al., 2001). Here, the values were calculated as the log base 10 of the median GM-CSF (pg/ml) secretion of three independent experiments in the presence of mixtures, minus the secretion in the absence of mixtures (Fig. 3B). Based on a model of independent contribution of

30 individual AAs to peptide antigen recognition, the predicted stimulatory score of a given peptide is the sum of the matrix values of each AA contained in the peptide at each position (Zhao et al., 2001). This scoring approach was applied to rank, according to their stimulatory score that is predictive of their stimulatory potency, all natural overlapping 10-mers peptides of the protein sequences within the UniProt human protein database. Based upon these predicted values, the

35 50 predicted human natural peptides with highest scores were synthesized and tested at 5 µg/ml (Fig. 4). Unexpectedly, none of the peptides was clearly stimulatory (Fig. 3C). The

predictive capacity of the above approach combining positional scanning library and biometrical analysis was lower for this TCC than for others in previous studies. The most stimulatory mixtures as shown in Fig. 3B have identical defined AAs in consecutive positions, suggesting a possible recognition of one AA motif in multiple frames. To be able to analyze the data using the approach outlined in Fig. 2, a set of 22 dual-defined mixtures (i.e. positional scanning mixtures with two defined positions) was designed, synthesized, and tested (Fig. 3D); the results confirmed the presence of a unique recognition motif (LHSXFEV, SEQ ID NO: 99) with different flanking residues (Fig. 3E). In order to apply this recognition motif in the two frames to the original GM-CSF-derived matrix and perform biometrical analysis, the harmonic mean model (HM) was used to integrate the stimulatory responses of some of the dual defined mixtures (Fig. 3D, Frame 1/2-HM mixtures) into the original matrix. Using the new "harmonic boost" frame 1 and frame 2 matrices, all natural overlapping 10-mers peptides derived from the UniProt human protein database were scored and ranked as schematically shown in Fig. 2. The 50 predicted natural peptides with highest scores for both matrices were synthesized and tested for GM-CSF release (Fig. 4). These two matrices allowed the identification of three clearly stimulatory peptides (Fig. 3F). The two most stimulatory peptides, NVLHSAFEVG (SEQ ID NO: 16) predicted with harmonic boost frame 1 and DNVLHSAFEV (SEQ ID NO: 15) predicted with harmonic boost frame 2, belong to GDP-L-fucose synthase encoded by the TSTA3 gene, and overlap by 9 AAs.

In detail, Figure 3 shows: A. GM-CSF production by TCC21.1 in response to a complete decapeptide positional scanning library (200 mixtures) presented by BLS cells expressing only DRB1*15:01. B. Score matrix designed with the log10 median of GM-CSF production of three independent experiments. Bold borders show mixtures selected for dual defined mixtures. C. GM-CSF production by TCC21.1 in response to the 50 peptides with highest scores predicted using the GM-CSF-based scoring matrix. D. GM-CSF production by TCC21.1 in response to 22 dual-defined mixtures. In grey, mixtures with defined AAs of the TCR motif in frame 1 and in black in frame 2. Stimulatory responses of mixtures in bold (Frame 1/2-HM) were integrated into the original matrix using the harmonic mean model. E. TCR motif and dual-defined mixture activity values selected based on the harmonic mean model and incorporated into the original matrix, Frame 1-HM in grey and Frame 2-HM in black. F. GM-CSF production by TCC21.1 in response to the 50 peptides with higher scores predicted using harmonic boost frame 1 and 2 score matrices. Complete decapeptide library, dual defined mixtures and individual decapeptides were presented by BLS cells expressing DRB1*15:01. Mixtures were tested at 200 µg/ml and individual decapeptides at 5 µg/ml. Histograms show mean ± standard error mean (SEM) and dot plots mean of three independent experiments. Cytokine released is always

expressed as pg/ml released by TCC21.1 in response to stimuli minus pg/ml released in absence of stimuli (negative control).

Figure 4 shows a summary of human decapeptides predicted with the biometrical approach, which have been synthesized and then tested for stimulatory capacity at 5 mg/ml based on GM-CSF release. The peptides are ranked from 1 to 50 for the HM Boost Frame 1 and 2, respectively.

3. RNASeq/transcriptome and proteome data demonstrating expression of GDP-L-fucose-synthase in brain tissue

Transcript level, expressed as “reads per kilo base of exon model per million mapped reads” (RPKM), of GDP-L-fucose synthase in the autologous brain LI and LIII is shown in Fig. 5. Transcript values for other brain-specific genes are also shown as quality control of the samples and as reference for genes expressed at high (MBP, PLP1), and medium (myelin-associated glycoprotein (MAG), myelin-associated oligodendrocyte basic protein (MOBP) and oligodendrocyte myelin glycoprotein (OMG)) level.

Seventeen GDP-L-fucose synthase peptides were identified in white and grey matter brain tissue from other MS patients and non-MS controls by proteomic analysis. The peptide sequences are listed in Fig. 5 as well as the peptide spectrum matches (PSM) in the different samples. The percentage of the GDP-L-fucose synthase AA sequence covered by identified peptides was 56%. The position refers to the sequence of UniProtKB/Swiss-Prot: Q13630.1. Analysis of other brain-specific proteins as reference is also included (Fig. 5).

In sum, GDP-L-fucose-synthase has been shown to be expressed in the brain (RNA and protein).

4. GDP-L-fucose synthase as the main autoantigen for a brain-infiltrating TCC

In order to identify autoantigens recognized by TCC21.1 in the two autologous brain lesions, within which TCC21.1 was known to be clonally expanded (LI and III (Planas et al., 2015)), the harmonic boost frame 1 and frame 2 matrices were then used to score and rank, according to their stimulatory score, all natural overlapping 10-mer peptides in the protein sequences within a brain-protein sub-database created with RNASeq-based transcriptome data from these two lesions (GSE60943). Of the 40 predicted natural brain peptides with highest scores for the frame 1 matrix, 38 peptides were already predicted from the unbiased UniProt human database. For frame 2 matrix, the top 40 peptides were previously predicted (Fig. 4). The two new

peptides for frame 1 were synthesized and tested. NVLHSAFEVG (GDP-L-fucose synthase 96-105, SEQ ID NO: 16) and DNVLHSAFEV (95-104, SEQ ID NO: 15) were found to stimulate TCC21.1 (Fig. 4).

5. Characterization of GDP-L-fucose synthase recognition / characterization of response

As mentioned above, the two GDP-L-fucose synthase peptides recognized by TCC21.1 overlap by 9 AAs. On DRB1*15:01-expressing BCL cells nine additional 10-mer peptides overlapping by 9 AAs were synthesized and tested and an additional peptide was identified, VLHSAFEVGA (97-106, SEQ ID NO: 18), which induced GM-CSF release. Peptide NVLHSAFEVG (96-105, SEQ ID NO: 16) gave the optimal response with an EC50 of 0.2 µg/ml and was presented by DRB1*15:01 and DQB1*06:02 molecules. The common AAs of the three stimulatory peptides are VLHSAFEV (97-104) (data not shown).

Next, the response of TCC21.1 to GDP-L-fucose synthase peptides presented by autologous irradiated PBMCs and BCL was characterized. GDP-L-fucose synthase peptides presented by the two types of APCs were able to induce proliferation. The functional phenotype of this response was also analyzed. TCC21.1 displayed a Th2 phenotype, releasing mainly Th2 cytokines and lower levels of IL-22 and IL-10. Unexpectedly, when peptides were presented by autologous BCL, they induced higher levels of IFNγ. In addition, TCC21.1 also released GM-CSF and IL-3 in response to GDP-L-fucose synthase peptides. The release of these two cytokines seems to be a specific feature of this TCC compared with other Th1, Th1* or Th1/2 TCCs that were generated from patients with different conditions. Intracellular cytokine staining confirmed the Th2 functional phenotype of TCC21.1 in response to GDP-L-fucose synthase (96-105, SEQ ID NO: 16). More than 60% of TCC21.1 cells were IL-4+ after stimulation with GDP-L-fucose synthase (96-105, SEQ ID NO: 16), while only around 12% were IFNγ+. Around 60% of the cells were GM-CSF+ and 47.7% of these also IL-4+. Further characterization of TCC21.1 demonstrated expression of CD28 and the chemokine receptor CRTh2 (data not shown).

6. Recognition of GDP-L-fucose synthase by CSF-infiltrating CD4+ T cells from patient 1154SA

Sixty-two 15-mer peptides overlapping by 10 AAs and covering the entire GDP-L-fucose synthase protein (Figure 6) were synthesized and tested for their ability to induce TCC21.1 proliferation when presented by autologous PBMCs. Seven immunodominant/encephalitogenic myelin peptides (Bielekova et al., 2004), CEF (cytomegalovirus, EBV, influenza virus and tetanus toxoid) peptide pool and control beads were tested in parallel (Figure 6). TCC21.1 recognized two overlapping GDP-L-fucose synthase peptides, 91-105 (SEQ ID NO: 14) and 96-

110 (SEQ ID NO: 17), containing the three stimulatory decapeptides (95-104 (SEQ ID NO: 15), 96-105 (SEQ ID NO: 16) and 97-106 (SEQ ID NO: 18)) that were identified previously.

7. Recognition of GDP-L-fucose synthase by CSF-infiltrating CD4+ T cells from CIS/MS patients

In order to find out if specific recognition of GDP-L-fucose synthase occurs in CSF-infiltrating CD4+ T cells in patients with different forms of MS (the majority CIS and RRMS), a new protocol was developed to expand fresh CSF-infiltrating CD4+ T cells to high numbers in a single round to minimize variations in the original T cell repertoire. Resting PHA-expanded CSF-infiltrating CD4+ T cells, T cells that are also derived from the CNS compartment, from 31 CIS/MS patients were tested in quadruplicate with the 62 overlapping GDP-L-fucose synthase peptides presented by autologous irradiated PBMCs, as well as with the seven myelin peptides, CEF peptide pool and control beads. All stimulatory indices (SIs; except control beads) were pooled and SI values less than one were treated as having unit value. Cluster k-means analysis was performed to determine the optimal cut-off to differentiate responsive SI values from nonresponsive in this patient population. K-means clustering resulted in a cut-off value of 1.455 to differentiate positive responses from negative. Subsequently, for each patient, all peptides having a median SI (of the quadruplicate wells) greater than 1.455 were identified as positive responses.

Next, patient scores were constructed by calculating the sum of each responsive peptide median SI, weighted by the number of total patients that responded to that peptide. In this way, both the SI values themselves for each peptide, as well as the relative immunogenicity of each peptide, were factored in to assess the frequency and strength of the immune response against the specific antigens. Three-cluster k-means analysis was performed on base 10 patient scores and clearly grouped patients into three categories: "nonresponders", "moderate responders", and "high responders". Nineteen patients (61.3%) were therefore characterized as nonresponders to GDP-L-fucose synthase, six (19.35%) as moderate, and six (19.35%) as high responders (data not shown). No significant differences between the three groups were found for CEF responses or for the positive or negative controls. Immunodominant peptides were defined as peptides able to induce positive responses in at least 10 % of the patients. The 14 GDP-L-fucose peptides achieving this criterion are shown in Figure 7. Functional analysis of the strongest responses to some immunodominant peptides revealed a Th1 phenotype with mainly production of IFN- γ (data not shown).

In detail, Figure 7 shows: Number of CIS/MS patients with CSF-infiltrating CD4+ T cells that responded to GDP-L-fucose synthase peptides. Immunodominant peptides that were positive in at least three patients are shown in black. Black squares are high responder and white squares moderate responder patients.

5

In conclusion, about 20-25% of the MS patients (CIS, RRMS, SPMS) show an immune response to different immunodominant GDP-L-fucose synthase peptides; most of these GDP-L-fucose synthase-specific T cells have a Th1 phenotype (most frequent phenotype in MS patients). A comparison with reaction to 7 myelin peptides (ETIMS peptides) shows that significantly fewer patients react to these peptides as compared to GDP-L-fucose synthase.

10

RASGRP2 Specificity

8. Identification of TCC14 and testing ps-SCL with TCC14 to identify RASGRP2

15

TCC14 was identified in an analogous approach from MS patient 1 (homozygous for HLA-DR15), however, in this case from the fraction of autoproliferating (proliferation without stimulation) peripheral blood T cells, which is enriched for brain-homing T cells. TCC14 was also shown by deep TCR sequencing of the cells infiltrating the patient's brain lesions that it is clonally expanded in the brain. TCC14 was generated as schematically shown in Fig. 1 (right part). The isolation of autoproliferating T cells is shown in more detail in Fig. 8A. In detail, peripheral blood mononuclear cells are seeded in replicate wells after labeling with the dye carboxyfluorescein diacetate N-succinimidyl ester (CFSE) without stimulus. After 7 days of culture, proliferating (CFSE^{dim}) and non-proliferating (CFSE^{hi}) cells are identified by flow cytometry and the autoproliferating T cells (CFSE^{dim}) isolated by cell sorting. The TCRβV sequences are compared between MS brain lesions and the CFSE^{dim} (autoproliferating) population. Over 20% of TCRβV sequences of the CFSE^{dim} population are also found in MS brain lesions (Fig. 8B).

20

25

30

35

TCC14 was then expanded and tested as shown in Fig. 2, using an unbiased approach to identify the target antigen/s of TCC14 using positional scanning combinatorial peptides libraries and biometrical analysis as described above. TCC14 expanded sufficiently well to be tested with the full set of 200 positional scanning library mixtures. The restriction of TCC14 was tested with BLS cells transfected with HLA-DR15 haplotype expressing HLA class II alleles (DR2a, DR2b or DQw6), since the TCCs are derived from a HLA-DR15 homozygous MS patient. After determining its HLA-class II restriction (DRB1*15:01), reactivity against all 200 samples was tested using BLS cells transfected with DRB1*15:01 and showed positive responses against single or multiple amino acids (aa) in each of the 10 positions. Proliferation (stimulatory index =

SI; dotted line SI = 2) based on thymidine incorporation assay was used after 72 h as readout for the TCC response (Figure 9 panel I). Using scoring matrices to summarize the reactivity of TCC14 against all 20 L-aa in each of the 10 positions of the decapeptide library after testing multiple doses, peptides to be recognized by TCC14 (Figure 9 panel II) were predicted using the biometrical analysis process (Zhao et al., 2001). Mean responses from three repetitive experiments were used to generate a matrix for optimal amino acid combinations of a potential peptide ligand. The brain transcriptome data and the respective proteins of patient 1, from whom the TCC had been isolated, were used as search database. 92 sequences were synthesized and tested for recognition by TCC14 based upon their appearance in the top 50 predicted peptides for at least one of the matrices used (stimulatory response with SI > 3) (Figure 9 panel III). As shown previously for other TCCs, there is a good relationship between predicted high ranking and T cell response (Sospedra et al., 2010; Zhao et al., 2001) since TCC14 recognized many of the high scoring peptides. To assess the functional avidity for these, dose titration experiments were performed with the 33 peptides that gave positive responses. Among these were the peptides with SEQ ID NOs: 33 to 35. The stimulatory peptides were tested in decreasing concentrations for the proliferative response of TCC14 using BLS DR2b after 72 h. (Figure 9 panel IV). A peptide from RASGRP2 was recognized with high antigen avidity ($EC_{50} = 0.012 \mu M$) (SEQ ID NO: 33), but peptides from several other RASGRP isoforms (RASGRP1, -3, 4) and other peptides gave positive responses as well (Figure 9 panel IV). The recognition of the RASGRP2 peptide by TCC14 resulted in secretion of Th2 cytokines and also of IFN- γ (data not shown).

In conclusion, multiple RASGRP versions are recognized by the clone TCC14 with RASGRP2 with by far the highest avidity (i.e. at lower antigen concentrations), underscoring its biological relevance.

9. RASGRP2 reactivity of peripheral blood cells

In order to test reactivity to RASGRP2 in peripheral blood-derived memory T cells, cryopreserved PBMCs (1×10^8 cells) of natalizumab-treated MS patients (NAT; n = 8) were thawed and afterwards depleted for CD45RA-expressing cells using magnetic cell sorting (Miltenyi). 2×10^5 CD45RA-depleted PBMCs were seeded per well (10-15 replicate wells per condition) in X-Vivo medium and were either treated with vehicle (DMSO), with CD2/CD3/CD28 beads or pulsed with RASGRP2 peptide pools (final concentration of pool $10 \mu M$) or whole purified RASGRP2 protein (Origene; $0.3 \mu g/ml$). The 15mer-overlapping peptides covering the whole RASGRP2 protein (SEQ ID NO:s 38-98) were organized in 9 peptide pools with 7 peptides per pool covering the sequence of RASGRP2 from N- (pool 1) to C-terminus (pool 9).

Thymidine incorporation assay was used to measure proliferation responses to RASGRP2. At day 7 the cells were pulsed with 1 μ Ci of methyl-3H-thymidine per well (Hartmann Analytic) and harvested after 15 h on a membrane (Tomtec). Incorporation was measured by β -scintillation counting (Wallac 1450, PerkinElmer). The results are shown as dots (mean $\pm$ SEM). Stimulatory index (SI) was calculated as ratio of peptide or protein stimulation vs. vehicle control. SI values > 2 were considered as positive (Figure 10).

In sum, Figure 10 shows that memory T cells from MS patients with high autoprolieration react to RASGRP2. All 8 donors responded to either individual or several peptide pools of RASGRP2 peptides (including pool 2, which contains the peptide with SEQ ID NO: 46 which overlaps with a target peptide of TCC 14 (SEQ ID NO: 33)) and/or entire protein demonstrating that RASGRP2 is an autoantigen that is broadly recognized by MS patients with a high degree of autoprolieration.

10. RNASeq/transcriptome and proteome data demonstrating expression of RASGRP2 in B cells and in the brain

Expression of RASGRP1-4 was tested on RNA and protein level in peripheral B cells and brain (Fig. 11). RASGRP1-3 are expressed both in the brain of patient 1 and in the transcriptome of autoprolierating memory B cells (Fig. 11).

In detail, Figure 11 shows: (A) Expression level of stimulatory peptide-originating transcripts in the active brain lesion III of MS patient 1 (RPKM) and autoprolierating (CFSE^{dim}), peripheral blood B cells (RPKM) from 6 RRMS (REM) patients. Expression levels under 0.1 or absent transcript expression were set as 0.1. Expression of control transcripts for brain (MOBP) and B cells (CD19) are also shown. (B) Mass spectrometry analysis of peripheral blood B cells of RRMS (REM, nihil; n = 4) and brain tissue (gray matter, pooled, n = 6) of MS patients. The protein coverage (columns) and spectral counts (numbers) of RASGRP1-4 are depicted as measure for protein abundance.

Brain tissue (white and gray matter) of controls and MS patients were also tested with proteomics using mass spectrometry. Peptides of RASGRP1, 2 and 3 proteins were identified in brain tissue, in particular with a high abundance of RASGRP2 (Figure 12). The position refers to the sequence of: GenBank AAC97349.1 (RASGRP1), GenBank AAI10307.1 (RASGRP2), and GenBank AAY15037.1 (RASGRP3).

Further, immunohistochemistry (IHC) studies were performed and showed expression of RASGRP2 protein in brain gray matter, specifically in cortical neurons, and spleen (data not shown).

5 **11. Identification of further immunodominant peptides within the GDP-L-fucose synthase and RASGRP2 sequences**

Furthermore, peptides have been identified based on commonly used search algorithms and assumptions to be potentially immunogenic. The approach has been adopted from the context of tumor vaccination, where the search for immunogenic peptides in self proteins (from a tumor)
10 is a standard procedure. In this scenario, a protein sequence is screened for peptides that are predicted to bind to disease-relevant HLA alleles (or the HLA alleles of a given tumor patient). Therefore, the GDP-L-fucose synthase and RASGRP2 sequences were taken and strong (SB) or weak (WB) binding peptides were predicted using the well-accepted NetMHCII (<http://www.cbs.dtu.dk/services/NetMHCII/>) and IEDB (<http://www.iedb.org/>) in silico peptide
15 binding prediction algorithms. Details how to perform the searches are readily available on the websites of the two search algorithms. In brief, the search entails copying the sequence of the protein of interest into the webtool and choosing a HLA-class II (or class I, if of interest) allele of interest. The algorithm will then yield the peptide sequences and their respective predicted binding to the HLA-class II allele. For alleles, for which NetMHCII searches were not possible,
20 IEDB was used. HLA alleles that are known to be associated with MS have been used. If one now considers all regions of the two proteins that are predicted to be WB (these are of highest interest in the context of autoantigens) or SB (NetMHCII searches) or have a predicted binding rank of 25% or lower (IEDB), these stretches of amino acids will cover almost the whole GDP-L-fucose synthase and RASGRP2 proteins:

25 For the GDP-L-fucose synthase protein, peptides spanning the area of amino acids 1 to 315 were predicted to bind.

30 For the RASGRP2 protein, peptides spanning the areas of amino acids 9 to 449 and 457 to 659 were predicted to bind.

The following alleles were used:

- 35 – HLA-DRB1*15:01 (NetMHCII)
- HLA-DRB5*01:01 (NetMHCII)
- HLA-DRB1*03:01 (NetMHCII)

- HLA-DRB1*13:03 (IEDB)
- HLA-DRB1*08:01 (IEDB)
- HLA-DRB3*02:02 (IEDB)
- HLA-DRB1*04:01 (NetMHCII)
- 5 – HLA-DRB1*04:04 (NetMHCII)
- HLA-DQw6 (DQA1*01:01; DQB1*06:02) (IEDB)

The following protein sequences were used:

- 10 GDP-L-fucose synthase: GenBank: AAH93061.1
 RASGRP2: GenBank: AAI10307.1

In conclusion, whole proteins and immunodominant peptides are suitable for use in the treatment, diagnosis and/or prevention of MS.

15

12. Manufacturing of a chemically coupled red blood cell

EDC as chemical cross linker

- A peptide that had been synthesized with a biotin residue (biotin-PLP1; PLP1 = PLP 139-154) has been used to allow highly specific detection of the peptide using fluorophore-conjugated streptavidin. Briefly, peripheral blood mononuclear cells were pulsed with either biotin-PLP1
20 (final concentration 0.05 mg/ml), EDC (final concentration 10 mg/ml), both in PBS for 1h at 4°C. After two washing steps cells were stained with fluorophore-conjugated streptavidin and analysed by flow cytometry (Streptavidin-APC).

- 25 As shown in Figure 13, efficient peptide binding was only observed in the presence of both (biotin-PLP peptide and EDC).

13. Further analysis with 105 MS patients (for GDP-L-fucose synthase) and 57 MS patients (for RASGRP2)

30

- A further analysis with 105 MS patients (for GDP-L-fucose synthase) and 57 MS patients (for RASGRP2) was performed to validate the immunodominance of GDP-L-fucose synthase peptides 51-65, 136-150, 161-175, 246-260 and 296-310 (SEQ ID NOs: 12, 21, 23, 28 and 32) and RASGRP2 peptide 78-92 (SEQ ID NO: 46). The reactivity of CSF-infiltrating CD4+ T cells,
35 which are highly relevant for the pathogenicity of MS, to the tested peptides was compared with

the reactivity to the known immunodominant reference peptides MBP 13-32, MBP 83-99, MBP 111-129, MBP 146-170, MOG 1-20, MOG 35-55, and PLP 139-154.

5 The results are shown in Figures 14 and 15. The data show that the tested peptides, which are fragments of the herein identified proteins GDP-L-fucose synthase and RASGRP2, are as reactive or even more reactive than the known immunodominant reference peptides. Thus, the tested peptides have again been confirmed as immunodominant. Studying CSF-infiltrating T cells and their response to a putative autoantigen is particularly meaningful because autoreactive T cells that have infiltrated the target organ, i.e. brain, spinal cord or CSF, are
10 considered likely biologically relevant

Proliferative responses and IFN- γ secretion were tested as described above under 7. and below under 15. Materials and Methods, unless otherwise specified.

15 In detail, Figure 14 shows: A. Proliferative responses expressed as stimulatory indices (SI) and IFN- γ secretion (pg/ml) of CSF-infiltrating CD4⁺ T cells (single round of PHA expansion) to five GDP-L-fucose synthase peptides (51-65, 136-150, 161-175, 246-260 and 296-310), four MBP peptides (13-32, 83-99, 111-129 and 146-170), two MOG peptides (1-20 and 35-55), one PLP peptide (139-154) and CEF peptides, presented by autologous PBMCs. All peptides have been
20 tested in four wells per patient. Each dot represents one well and each peptide has been tested in 420 wells (4 wells x 105 patients). Positive wells are wells with SIs above 2 (dotted line) or with IFN- γ above 20 pg/ml (dotted line). Percentage of positive wells are also shown as well as the ratio between the percentage of positive wells for IFN- γ and SI. B. Percentage of positive patients with CSF-infiltrating CD4⁺ T cells specific for the different peptides or without identified
25 specificity. Positive patients are defined as patients with more than 2 of the 4 wells positive for SI (left histogram), IFN- γ (middle histogram) and SI or IFN- γ (right histogram).

In detail, Figure 15 shows: A. Proliferative responses expressed as stimulatory indices (SI) and IFN- γ secretion (pg/ml) of CSF-infiltrating CD4⁺ T cells (single round of PHA expansion) to one
30 RASGRP2 peptide (78-92), four MBP peptides (13-32, 83-99, 111-129 and 146-170), two MOG peptides (1-20 and 35-55), one PLP peptide (139-154) and CEF peptides, presented by autologous PBMCs. All peptides have been tested in four wells per patient. Each dot represents one well and each peptide has been tested in 228 wells (4 wells x 57 patients). Positive wells are wells with SIs above 2 (dotted line) or with IFN- γ above 20 pg/ml (dotted line). Percentage of
35 positive wells are also shown as well as the ratio between the percentage of positive wells for IFN- γ and SI. B. Percentage of positive patients with CSF-infiltrating CD4⁺ T cells specific for

the different peptides or without identified specificity. Positive patients are defined as patients with more than 2 of the 4 wells positive for SI (left histogram), IFN- γ (middle histogram) and SI or IFN- γ (right histogram).

5 **14. Tolerance induction *in vivo* with myelin peptides in a Phase Ib trial**

10 patients who were diagnosed with MS were tested for tolerance induction. In particular, they were tested for safety and tolerability of peptide-coupled, EDC-fixed red blood cells and for indicators of tolerance induction *in vivo* by i.v. injection of autologous red blood cells chemically (by EDC) coupled to a set of myelin peptides (MBP 13-32, MBP 83-99, MBP 111-129, MBP
10 146-170, MOG 1-20, MOG 35-55, and PLP 139-154) (SEQ ID Nos: 261-267) in context of a Phase Ib trial. In short, blood is taken from each patient and red blood cells are separated. The red blood cells are then chemically coupled to the myelin peptides *ex vivo* under sterile conditions and i.v. injected into the patient. 2 patients received 1×10^{10} cells, 3 patients received 1×10^{11} cells and 5 patients received 3×10^{11} cells. The day of injection was defined as Day 0.

15 Blood was also taken 6 weeks before injection (= Pre Tolerization) and 12 weeks after injection (= Post Tolerization). On these dates, the peripheral blood lymphocytes were obtained and examined by flow cytometry with fluorescently labeled antibodies for the presence of a wide range of different cell types including several subpopulations of T cells, B cells, monocytes and
20 dendritic cells, natural killer cells including those expressing markers of induced T regulatory cells (Tr1) or natural T regulatory cells (nTregs). The latter two cell types can be characterized by the following markers:

- T regulatory 1 (Tr1) cells (CD3+ CD4+ CD45RA- CD49b+ LAG3+)
- FoxP3+ natural T regulatory (nTreg) cells (CD4+ CD25hi FOXP3+)

25 In addition, T cell reactivity to all seven peptides individually was measured for peripheral blood T cells and for CSF-infiltrating T cells as described below under “Testing of CSF-derived bulk T cells from CIS- and RRMS patients against GDP-L-fucose synthase and myelin peptides” and “T cell stimulation” on the same dates, i. e. 6 weeks before injection and 12 weeks after
30 injection.

Figure 16 shows that signs of antigen-specific tolerance induction upon injection of the coupled red blood cells were detectable as measured by an increase in Tr1 and FoxP3+ nTreg cells and a decrease of peptide-specific T cell reactivity. These data indicate that by administering
35 immunodominant peptides tolerance to the respective immunodominant peptides can be induced.

In detail, Figure 16 shows:

A. Percentage of Tr1 cells of CD4⁺ memory cells 6 weeks before injection (= Pre Tolerization) and 12 weeks after injection (= Post Tolerization). B. Percentage of FoxP3⁺ nTreg cells of CD4⁺ memory cells 6 weeks before injection (= Pre Tolerization) and 12 weeks after injection (= Post Tolerization). C. T cell reactivity in five patients that received 3×10^{11} cells 6 weeks before injection and 12 weeks after injection (grey dots: before injection; black dots: after injection). Top graph shows the percentage of cells responding to all peptides. Bottom graph shows the proliferation of each individual microwell (60 in total).

15. Materials and Methods

Patient Material

GDP-L-fucose synthase

Patient 1154SA: CSF-derived mononuclear cells and PBMCs were obtained from a SPMS patient with pattern II demyelinating lesions as previously described (Planas et al., 2015). HLA-class I and II types of this patient were: A*32:01, A*33:01, B*14:02, B*51:01, DRB1*15:01, DRB5* 01:01, DQB1*06:02 and DQA1*01:02.

CSF from diagnostic lumbar puncture and paired peripheral blood were collected from 31 untreated MS patients: 8 patients with CIS, 20 patients with RRMS and 3 patients with SPMS. Patients were recruited from the inims outpatient clinic and day hospital at the University Medical Center Hamburg-Eppendorf and the nims section, Neurology Clinic, University Hospital Zurich. MS diagnosis was based on the revised McDonald criteria. All CIS patients had CSF-specific oligoclonal bands detected by isoelectric focusing (IEF). Patients, who had not received steroids at least 4 weeks prior to enrolment or any immunomodulatory or immunosuppressive agent during the last 3 months, were considered untreated and included in the study. Fresh CSF cells from these patients were expanded *in vitro* (see below). PBMCs were freshly isolated from EDTA-containing blood tubes by Ficoll density gradient centrifugation (PAA, Pasching, Austria) and cryopreserved. The Ethik Kommission der Ärztekammer Hamburg, protocol No. 2758 and the Cantonal Ethical Committee of Zurich, EC-No. of the research project 2013-0001, approved the study procedures. Informed consent was obtained from all patients or relatives.

Brain autopsy tissue from 13 MS patients (7 SPMS, 5 PPMS and 1 primary relapsing (PR) MS) and 7 non-MS controls were obtained from the UK Multiple Sclerosis Tissue Bank (UK Multicentre Research Ethics Committee, MREC/02/2/39).

For the further analysis of GDP-L-fucose synthase (13. of "Examples"), CSF from 105 patients were collected. Among the 105 patients, the ratio females to males was 1.9, the mean age was 35.68 (range 17-58), 4 patients were diagnosed with RIS, 10 patients were diagnosed with CIS, 82 patients were diagnosed with RRMS, 4 patients were diagnosed with SPMS and 5 patients were diagnosed with PPMS.

RASGRP2

PBMC were isolated from patient 1154SA by Ficoll density centrifugation.

For the further analysis of RASGRP2 (13. of "Examples"), CSF from 57 patients were collected. The 57 patients were among the 105 patients tested for GDP-L-fucose synthase. Among the 57 patients, the ratio females to males was 2.5, the mean age was 35.33 (range 17-55), 2 patients were diagnosed with RIS, 8 patients were diagnosed with CIS, 43 patients were diagnosed with RRMS, 1 patient was diagnosed with SPMS and 3 patients were diagnosed with PPMS.

Autoproliferation assays

PBMCs were thawed with complete IMDM media (GE Healthcare) containing 100 U/mL penicillin/streptomycin (Corning), 50 µg/mL gentamicin (Sigma-Aldrich), 2 mmol/L L-glutamine (PAA) and 5% heat-decomplemented human serum (HS, PAA) and afterwards washed once with serum-free AIM-V medium (GIBCO, Thermo Fisher Scientific), containing human albumin. Cells were incubated for 15 min in AIM-V medium containing 50 U/ml DNase (Roche) at 37°C to avoid cell clump formation. Following two wash steps with PBS containing 0.1% HS, cells were resuspended at a concentration of 10×10^6 cells/ml in PBS/0.1% HS and were then labeled at a final concentration of 0.5 µM CFSE(Sigma-Aldrich) for 3 min at room temperature. The labeling was stopped by quenching with 5x excess volume of cold complete RPMI (PAN Biotech) medium containing 10% HS. After one further wash step with AIM-V, CFSE-labeled cells were seeded at 2×10^5 PBMCs/200 µl per well in AIM-V (10-12 replicate wells per donor and condition) in 96-well U-bottom microtiter plates (Greiner Bio-One) at 37°C, 5% CO₂, in the absence of exogenous stimuli (= autoproliferation). For conventional T cell reactions, for the same donors PHA (0.5 µg/ml) as TCR-independent stimulus, tetanus toxoid (TTx, 5 µg/ml, Novartis Behring) as foreign antigen stimulus and mixed lymphocyte reaction (MLR) as allogeneic antigen stimulus were used. After 7 days CFSE-labeled cells were collected and pooled from replicate wells, washed with PBS, Fc-blocked with human IgG (Sigma-Aldrich) and labeled with Live/Dead® Aqua (Invitrogen, Thermo Fisher Scientific) at 4°C. After washing with cold PBS containing 2mM EDTA and 2% FCS, cells were directly stained for surface markers using the fluorochrome-conjugated antibodies (Key Resource Table). Measurements were

performed on an LSR Fortessa Flow Cytometer (BD Biosciences), and data were analyzed with FlowJo (Tree Star). This assay was further used to test competition of autoproliiferation by incubating CFSE-labeled PBMCs in the presence of anti-HLA-DR, anti-CD4, anti-IFN- γ and anti-GM-CSF antibodies (10 μ g/ml) or appropriate isotype controls for 7 days. For thymidine incorporation assay, 2×10^5 PBMCs/well (10-12 replicate wells per donor and condition) were cultured with serum-free AIM-V medium in 96-well U-bottom microtiter plates at 37°C, 5% CO₂ and at day 7 pulsed with 1 μ Ci of methyl-³H-thymidine per well (Hartmann Analytic) and harvested cells after 15 h (Tomtec). Incorporation was measured by β -scintillation counting (Wallac 1450, PerkinElmer).

T cell cloning

In order to generate TCCs from the autoproliiferating compartment, 500 CFSE^{dim} cells from the sorted cell pool of CFSE^{dim} cells (20.000 cells) of MS patient 1 (for the TCRV β sequencing, described above) were split and limiting dilution was performed as previously described (Aly et al., 2011). TCCs were enriched using an expansion protocol with PHA (Sigma) and human IL-2 (Aly et al., 2011). Sequencing of TCR rearrangements of the generated TCCs was analyzed, as previously described (Yousef et al., 2012). To assess the cytokine response of the TCCs, anti-CD2/CD3/CD28 antibody-loaded MACSibead particles (Miltenyi) were used to stimulate each TCC in 7 replicate wells with each 200.000 cells in X-Vivo medium (Lonza). After 48 h supernatants were collected and cytokine responses measured, as described above. For chemokine receptor expression, TCCs were stained with Live/dead Aqua and antibodies against CXCR3 and CCR6 after thawing and resting cells overnight.

Transcriptomic Analysis

Transcriptomic analysis of brain lesions was performed as previously described (Planas et al., 2015). The data discussed was deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE60943.

Proteomic Analysis

For proteomic analysis, pressure-assisted protein extraction and digestion was performed with a barocycler (2320EXT, BioSciences, Inc, South Easton, MA). Reduction and alkylation was applied on the homogenate before samples were digested with Lys-C and trypsin. Peptides were desalted on solid phase extraction columns (C18 Finisterre, Wicom Germany), vacuum dried, re-dissolved and measured (Nanodrop 1000, spectrophotometer (Thermo Scientific, Wilmington, DE, USA). Resulting peptides were purified and separated by hydrophilic interaction chromatography (HILIC, Agilent LC1200 equipped with a column Polyamin II 250 x

3.0 mm 120 Å, 5µm) before they were injected on a nano liquid chromatography system Easy-nLC linked to an Orbitrap Fusion instrument (Thermo Fisher). Data analysis was performed with MASCOT software using a human UniProtKB/Swiss-Prot protein database (March 22, 2016 with 40912 entries). Search parameters were 0.05 Da fragment mass tolerance and 10ppm precursor mass, minimal number of peptides 2, and FDR (false discovery rate) of 0.1%, allowing 2 mis-cleavages on trypsin fragments. Carbamidomethylation at cysteine was set as a fixed modification, and oxidation of methionine, n-terminal acetylation as variable modifications.

Positional Scanning Peptide Libraries and Individual Peptides

10 *GDP-L-fucose synthase*

A synthetic N-acetylated, C-amide L-amino acid (AA) decapeptide combinatorial library in a positional scanning format (200 mixtures) and twenty-two dual defined mixtures were prepared. Individual peptides (Fig. 4 and Fig. 6) were synthesized by Peptides and Elephants GmbH (Potsdam, Germany).

15

RASGRP2

An L-aa decapeptide positional scanning library (N-acetylated and C-amide TPI 2040) was prepared by standard methods. Each of the 200 mixtures of the library was tested for their proliferative activity by TCC14 at 40, 120, and 200 µg/ml using thymidine incorporation assay.

20 The restriction of TCC14 was tested with BLS cells transfected with HLA-DR15 haplotype expressing HLA class II alleles (DR2a (= DRB5*01:01), DR2b (= DRB1*15:01) or DQw6 (= DQB1*0602)), since the TCC is derived from a HLA-DR15 homozygous MS patient. HLA class II expression of BLS cell lines was verified with specific antibodies against DR2a, DR2b and DQ and cells were tested negative for mycoplasma. The results were organized into four matrices (data not shown): three matrices each representing the activity at one of the above doses, and a matrix using the concentration to achieve 3-fold proliferation to combine all three doses into a single activity. Using the biometrical analysis process (Zhao et al., 2001) against the transcriptome protein database from the brain of MS patient 1154SA, predicted peptide lists were generated for each of the four matrices. Due to a large amount of agreement between the predicted lists, a total of 92 distinct decamer peptides occurred within the top 50 predicted peptides in at least one matrix's prediction list. These peptides were chosen to be synthesized (by Peptides and Elephants GmbH, Potsdam, Germany) and tested.

Cells and culture conditions

35 Bulk CSF-derived mononuclear cells from patient 1154SA were expanded as previously reported (Planas et al., 2015). Briefly, 2000 cells per well were seeded in 96-well U-bottom

microtiter plates together with 2×10^5 allogeneic irradiated PBMC (45 Gy), 1 $\mu\text{g/ml}$ of PHA-L (Sigma, St Louis, MO) and IL-2 supernatant (500 U/ml). Medium consisted of IMDM (PAA) containing 100 U/ml penicillin/streptomycin (PAA), 50 $\mu\text{g/ml}$ gentamicin (BioWhittaker, Cambrex), 2 mM L-glutamine (Gibco, Invitrogen, Carlsbad, CA) and 5% heat-decomplemented human serum (PAA). Additional IL-2 was added every 3-4 days. CSF-infiltrating CD4+ T cells were positively selected using anti-CD4 magnetic beads (CD4 Micro Beads human MACS, Miltenyi Biotec Inc, CA, USA) and restimulated once again with PHA-L, IL-2 and allogeneic irradiated PBMC.

TCC21.1 was established from CSF-infiltrating cells and TCC14 from PBMC-derived autoproliiferating T cells as previously described (Planas et al., 2015).

Testing of CSF-derived bulk T cells from CIS- and RRMS patients against GDP-L-fucose synthase and myelin peptides

Fresh bulk CSF-derived mononuclear cells from the 31 CIS/MS patients were mixed with 5×10^6 allogeneic irradiated PBMCs and CD4+ T cells were positively selected with anti-CD4 magnetic beads. CD4+ fractions were then seeded at 1500 cells per well in 96-well U-bottom microtiter plates together with 1.5×10^5 allogeneic irradiated PBMC, 1 $\mu\text{g/ml}$ of PHA-L and IL-2 supernatant. Medium consisted of RPMI 1640 without Hepes (Pan-Biotech, Aidenbach, Germany) supplemented with 2 mM glutamine (Pan-Biotech), 1% (vol/vol) nonessential amino acids (Gibco), 1% (vol/vol) sodium pyruvate (Gibco), 50 $\mu\text{g/ml}$ penicillin-streptomycin (Corning, NY, USA), 0.00001% β -mercaptoethanol (Gibco) and 5% human serum (Blood Bank Basel). Additional IL-2 was added every 4 days. Growing wells were transferred to 48 well plates and finally to 75 cm³ flask until cells were fully rested (20-25 days). Cells were highly expanded in a single round of stimulation.

For the further analysis of GDP-L-fucose synthase peptides and a RASGRP2 peptide (13. of "Examples"), the same method was performed.

An autologous BCL from patient 1154SA was generated by EBV-transformation. BLS cells were transfected with single HLA class II molecules, DR2a (DRA1*01:01, DRB5*01:01), DR2b (DRA1*01:01, DRB1*15:01) and DQw6 (DQA1*01:02, DQB1*06:02).

T cell stimulation

TCC responses to single/dual defined peptide mixtures or individual decapeptides were tested by seeding in duplicate 2×10^4 T cells and 5×10^4 irradiated BLS cell lines or autologous BCL or

1 x 10⁵ irradiated PBMC (as indicated) with or without combinatorial peptide mixtures or individual decapeptides. 2.5 µg/ml PHA and 10⁻⁷ M PMA (Sigma), 1 µg/ml of surface-coated anti-CD3 (OKT3, Ortho Biotech Products, Raritan, NJ) and 0.5 µg/ml of soluble anti-CD28 (Biolegend, San Diego, CA), and a T Cell Activation Kit (anti-CD3, anti-CD28, anti-CD2 beads) (Miltenyi Biotec) served as positive controls as indicated.

The response of PHA-expanded CSF-infiltrating CD4⁺ T cells to GDP-L-fucose synthase, myelin and CEF peptides (Figure 6) was tested by seeding in quadruplicate 6 x 10⁴ T cells and 2 x 10⁵ irradiated autologous PBMCs with or without peptides. For EdU experiments BLS DRB1*15:01 were used as APCs. T Cell Activation Kit was used as positive control.

For the further analysis of GDP-L-fucose synthase peptides and a RASGRP2 peptide (13. of "Examples"), the same method was performed as for the response of PHA-expanded CSF-infiltrating CD4⁺ T cells to GDP-L-fucose synthase.

15

Cytokine measurement

GDP-L-fucose synthase

Cytokines in the supernatant of stimulated TCC21.1 and expanded CSF cells were measured 48 h after stimulation using the Human T Helper Cytokine Panel LEGENDplex bead-based immunoassay (Biolegend), GM-CSF ELISA (BD Biosciences, Franklin Lakes, NJ) and IL-3 ELISA (Biolegend) according to the manufacturer's instructions.

For intracellular cytokine staining, TCC21.1 was analyzed 48 h after stimulation. After 5 h in presence of GolgiStop protein transport inhibitor (BD Biosciences), T cells were labeled with Live/Dead® Aqua (Invitrogen). Following fixation and permeabilization with Cytofix/Cytoperm (BD Biosciences), cells were stained with antibodies against CD4 (APC-Cy7, Biolegend), IFN-γ (FITC, Biolegend), IL-4 (PE, BD Bioscience), GM-CSF (APC, Biolegend) and IL-3 (PE, Biolegend) in PBS containing saponin and BSA, and analyzed by flow cytometry.

For the further analysis of GDP-L-fucose synthase peptides and a RASGRP2 peptide (13. of "Examples"), the secretion of IFN-γ in the supernatant of stimulated and unstimulated expanded CSF-infiltrating T cells was measured after 48 h of culture using IFN-γ ELISA (Biolegend) per duplicate in all single wells according to the manufacturer's instructions.

35

Proliferative Responses*GDP-L-fucose synthase*

Proliferation was measured 72 h after stimulation by 3H-thymidine (Hartmann Analytic, Braunschweig, Germany) incorporation in a scintillation counter (Wallac 1450, PerkinElmer, Rodgau-Jürgesheim, Germany). The stimulatory index (SI) was calculated as follows: SI = Median (replicates cpm peptide) / Median (replicates cpm without peptide). Proliferation was also measured using a Click-iTTM EdU Flow Cytometry Assay Kit (APC, Molecular Probes, Invitrogen) following manufacturer's instructions. Cells were stained with the following antibodies anti-CD3 (PE-Cy-7, e-Bioscience, San Diego, CA) and anti-TRBV-21 (FITC, Beckman Coulter, Brea, CA) and analyzed by flow cytometry.

RASGRP2

The proliferative responses were tested as described under 9.

For the further analysis of GDP-L-fucose synthase peptides and a RASGRP2 peptide (13. of "Examples"), proliferative responses were measured as described for GDP-L-fucose synthase above.

Surface receptor expression

Resting TCC21.1 was stained with antibodies against CD4 (PE-Texas Red, Thermo Fischer, Waltham, MA), TRBV21 (FITC, Beckman Coulter), CD28 (PE-Cy7, BioLegend), CCR4 (APC, BioLegend), CCR6 (BV785, BioLegend) and CRTh2 (PE, BioLegend) and analyzed by flow cytometry.

Flow Cytometric Analysis

Sample acquisition was conducted using a LSR Fortessa Flow Cytometer (BD Biosciences) with Diva software, and data were analyzed with FlowJo (Tree Star, Ashland, OR).

RT-PCR and Sequencing of TCR Rearrangements

RNA extraction, reverse transcription and TCRα/β-chain (TRA/BV) sequencing of TCC21.1 was assessed as previously reported (Planas et al., 2015). TCR gene designations are in accord with IMGT nomenclature (ImMunoGeneTics, www.IMGT.org).

HLA

Individuals were typed for HLA-class I and II molecules at Histogenetics LLC, NY, USA. Isolation of DNA from whole blood with a final concentration of 15 ng/μl was performed with a

standard DNA isolation protocol using a Triton lysis buffer and proteinase K treatment. The samples were typed for HLA class I (A* and B*) and HLA class II (DRB1*, DRB3*, DRB4*, DRB5*, DQA1* and DQB1*) using high-resolution HLA sequence-based typing (SBT). The HLA class II binding predictions were made using the IEDB analysis resource consensus tool.

5

Statistical analysis

Three-cluster k-means analysis was performed on patient scores to group patients into three categories. Associations between response levels of peptides, patients, and HLA status were all performed using Fisher's Exact Test with Bonferroni-Holm correction applied as appropriate, with 5% significance.

10

EMBODIMENTS

1. A GDP-L-fucose synthase protein or a protein of the RASGRP protein family, or a fragment, derivative or splice variant thereof, or a nucleotide sequence encoding any of the proteins or fragment, derivative or splice variant thereof, for use in the treatment, diagnosis and/or prevention of multiple sclerosis (MS).
A GDP-L-fucose synthase protein or a protein of the RASGRP protein family, in particular RASGRP2, or a fragment thereof is especially preferred.
2. The protein, fragment, derivative or splice variant according to embodiment 1, wherein the GDP-L-fucose synthase protein
 - a) has the amino acid sequence as set forth in SEQ ID NO: 1 or
 - b) has an amino acid sequence which is at least 85 %, preferably at least 90 %, more preferably at least 95 % identical to the amino acid sequence as set forth in SEQ ID NO: 1 or
 - c) has an amino acid sequence which is at least 70 %, preferably at least 80 %, more preferably at least 90 % homologous to the amino acid sequence as set forth in SEQ ID NO: 1 or
 - d) has an amino acid sequence which is at least 60 %, preferably at least 70 %, more preferably at least 80 %, even more preferably at least 90 % homologous to the amino acid sequence as set forth in SEQ ID NO: 1 and the protein or fragment or splice variant thereof binds to an autologous HLA allele, is recognized by a T cell and/or is recognized by an antibody which binds to or recognizes the amino acid sequence as set forth in SEQ ID NO: 1 or a fragment thereof or
 - e) is encoded by a TSTA3 gene, in particular by a gene sequence of nucleotides 143612618 to 143618048 of NC_000008.11, or is encoded by a gene which is at least 80 %, preferably at least 90 %, even more preferably at least 95 % identical to the gene sequence of nucleotides 143612618 to 143618048 of NC_000008.11
and/or
wherein the member of the RASGRP protein family
 - f) has the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 or SEQ ID NO: 9 or
 - g) has an amino acid sequence which is at least 85 %, preferably at least 90 %, more preferably at least 95 % identical to the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 or SEQ ID NO: 9 or
 - h) has an amino acid sequence which is at least 70 %, preferably at least 80 %, more preferably at least 90 % homologous to the amino acid sequence as set forth in any of SEQ ID NO: 2,

SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 or SEQ ID NO: 9 or

5 i) has an amino acid sequence which is at least 60 %, preferably at least 70 %, more preferably at least 80 %, even more preferably at least 90 % homologous to the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 or SEQ ID NO: 9 and the protein or fragment or splice variant thereof binds to an autologous HLA allele, is recognized by a T cell and/or is recognized by an antibody which binds to or recognizes the respective amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 or SEQ ID NO: 9 or a fragment thereof or

10 j) is encoded by a RASGRP gene, in particular by a gene sequence of nucleotides

- 38488101 to 38565575 of NC_000015.10,
- 64726911 to 64745456 of NC_000011.10,
- 33436324 to 33564750 of NC_000002.12, or
- 15 – 38409051 to 38426305 of NC_000019.10,

or is encoded by a gene which is at least 80 %, preferably at least 90 %, even more preferably at least 95 % identical to the gene sequence of nucleotides

- 38488101 to 38565575 of NC_000015.10,
- 64726911 to 64745456 of NC_000011.10,
- 20 – 33436324 to 33564750 of NC_000002.12, or
- 38409051 to 38426305 of NC_000019.10.

The GDP-L-fucose synthase protein with the amino acid sequence as set forth in SEQ ID NO: 1 or a protein with at least 90 % identity or a fragment thereof is especially preferred. A RASGRP2 protein with the amino acid sequence as set forth in SEQ ID NO: 2, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 or SEQ ID NO: 9 or a protein with at least 90 % identity or a fragment thereof is also especially preferred.

3. The protein, fragment, derivative or splice variant according to embodiment 1 or 2, wherein the fragment comprises 5 to 50, preferably 5 to 20, more preferably 10 to 15 amino acids, even more preferably 15 amino acids.

30 A fragment of a length of 10 to 15 amino acids is particularly preferred.

4. The protein, fragment, derivative or splice variant according to embodiment 2 or 3, wherein the fragment is

35 a) at least 85 %, preferably at least 90 %, more preferably at least 95 % identical to a respective corresponding amino acid sequence or

b) at least 70 %, preferably at least 80 %, more preferably at least 90 % homologous to a respective corresponding amino acid sequence or

c) at least 60 %, preferably at least 70 %, more preferably at least 80 %, even more preferably at least 90 % homologous to a respective corresponding amino acid sequence and binds to an autologous HLA allele, is recognized by a T cell and/or is recognized by an antibody which binds to or recognizes the respective amino acid sequence.

It is especially preferred that the fragment is at least 90 % identical to the respective corresponding amino acid sequence.

5. The protein, fragment, derivative or splice variant according to any of the preceding embodiments, wherein the fragment comprises a sequence selected from the group comprising SEQ ID NOs: 10 to 98, preferably SEQ ID NOs: 10 to 35, preferably consists of a sequence selected from the group comprising SEQ ID NO: 10 to 98, preferably SEQ ID NOs: 10 to 35.

The fragment preferably comprises a sequence selected from the group comprising SEQ ID NOs: 10 to 35.

6. The protein, fragment, derivative or splice variant according to any of the preceding embodiments for identifying a human subject who is suitable for tolerization to autoantigens in MS, preferably early MS. It is particularly preferred to use the protein or fragment thereof.

7. The protein, fragment, derivative or splice variant according to any of embodiments 1 to 5 for diagnosing pattern II MS in a human subject. It is particularly preferred to use the protein or fragment thereof.

8. A carrier comprising at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence according to any of embodiments 1 to 5. It is particularly preferred that the carrier comprises at least one protein, fragment or nucleotide sequence.

9. The carrier according to embodiment 8, wherein the carrier is coupled to the at least one protein, fragment, derivative and/or splice variant, and/or the carrier contains the at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence.

It is particularly preferred that the carrier is coupled to the at least one protein or fragment and/or that the carrier contains the nucleotide sequence.

10. The carrier according to embodiment 8 or 9, wherein the carrier is selected from the group comprising a cell, preferably a blood cell, a protein, a lipid, a glycolipid, a bead, a nanoparticle, a virus-like-particle (VLP) and a molecule, such as a sugar molecule, and any combination thereof.
- 5 The carrier is preferably a blood cell.
11. The carrier according to embodiment 10, wherein the protein, fragment, derivative and/or splice variant is expressed by the cell, preferably the blood cell.
It is especially preferred that the protein is expressed by a blood cell.
- 10 12. The carrier according to embodiment 10 or 11, wherein the blood cell is a red or white blood cell.
13. The carrier according to any of embodiments 10 to 12, wherein the carrier is a blood cell
15 and the blood cell is chemically coupled by a coupling agent, preferably by 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (ECDI/EDC), to the at least one protein, fragment, derivative and/or splice variant. Preferably, the at least one protein or fragment is coupled by EDC to a blood cell.
- 20 14. A method of manufacturing the chemically coupled blood cell of embodiment 13, comprising isolating the blood cell from a human subject, adding the at least one protein, fragment, derivative and/or splice variant and subsequently adding the coupling agent, preferably EDC. Preferably, at least one protein or fragment is added.
- 25 15. A pharmaceutical composition comprising at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence according to any of embodiments 1 to 5 and a pharmaceutically acceptable carrier. The pharmaceutical composition preferably comprises at least one protein or fragment thereof or a nucleotide sequence and a pharmaceutically acceptable carrier.
- 30 16. A method for inducing antigen-specific tolerance to autoantigens in a human subject suffering from or at risk of developing MS comprising the step of applying to the human subject
- a) at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene
35 sequence according to any of embodiments 1 to 5, and/or
- b) at least one carrier according to any of embodiments 8 to 13.

A method for inducing antigen-specific tolerance is preferred, wherein at least one protein or fragment thereof or a carrier which coupled to the at least one protein or fragment and/or a carrier which contains the nucleotide sequence is applied.

- 5 17. The method according to embodiment 16, wherein the at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence is applied by nasal, inhaled, oral, subcutaneous (s.c.), intracoelomic (i.c), intramuscular (i.m.), intradermal (i.d.), transdermal (t.d.) or intravenous (i.v.) administration, preferably by i.v., s.c., i.d., t.d., oral, inhaled, nasal or coupled to a carrier, preferably a red blood cell.
- 10 18. The method according to embodiment 16 or 17 for inducing antigen-specific tolerance to autoantigens in early MS.
- 15 19. A method for identifying a human subject suitable for tolerization to autoantigens in MS, preferably early MS, comprising isolating T cells and/or antibodies from blood, CSF or other body fluid of the subject and measuring reactivity of the T cells and/or antibodies against the protein, fragment, derivative and/or splice variant according to any of embodiments 1 to 5. It is particularly preferred to measure the reactivity of the T cells and/or antibodies against a fragment.
- 20 20. The fragment according to any of embodiments 3 to 5, preferably the fragment according to embodiment 5, for use as a medicament. The fragment particularly comprises a sequence selected from the group comprising SEQ ID NOs: 10 to 35.
- 25 21. Use of the protein, fragment, derivative and/or splice variant according to any of embodiments 1 to 5 in the *in vitro* diagnosis of MS. It is particularly preferred to use the protein or a fragment thereof.
- 30 22. A method for *in vitro* diagnosing MS using the protein, fragment, derivative and/or splice variant according to any of embodiments 1 to 5. It is particularly preferred to use the protein or a fragment thereof.
- 35 23. Use of the protein, fragment, derivative and/or splice variant according to any of embodiments 1 to 5 in the *in vitro* pretesting of a human subject diagnosed with MS or a human subject at risk to develop MS. It is particularly preferred to use the protein or a fragment thereof.

24. Use of the protein, fragment, derivative and/or splice variant according to any of embodiments 1 to 5 for the manufacture of a medicament for the treatment, diagnosis and/or prevention of MS. It is particularly preferred to use the protein or a fragment thereof.
- 5
25. The protein, fragment, derivative and/or splice variant according to any of embodiments 1 to 5, wherein the derivative is an amino acid sequence which shares a homology or identity over its entire length with a corresponding part of the reference amino acid sequence of at least 75%, more preferably at least 80%, at least 85%, at least 90%, at least 93%, at least 95%, at least 97%, at least 98% or at least 99%.
- 10
26. An *in vitro* method for identifying a human subject suitable for tolerization to autoantigens in MS, preferably early MS, comprising measuring reactivity of T cells and/or antibodies against the protein, fragment, derivative and/or splice variant according to any of embodiments 1 to 5 with previously obtained T cells and/or antibodies from blood, CSF or other body fluid of the subject. It is particularly preferred to use the protein or a fragment thereof.
- 15
27. The at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence according to any of embodiments 1 to 5, and/or the at least one carrier according to any of embodiments 8 to 13 for use in a method for inducing antigen-specific tolerance to autoantigens in a human subject suffering from or at risk of developing MS comprising the step of applying to the human subject the at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence according to any of embodiments 1 to 5, and/or the at least one carrier according to any of embodiments 8 to 13. It is particularly preferred to use the protein or a fragment thereof and/or the at least one carrier coupled to the protein or the fragment thereof.
- 20
28. The at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence according to embodiment 27 and/or the at least one carrier according to embodiment 27, wherein the at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence is applied by nasal, inhaled, oral, subcutaneous (s.c.), intracoelomic (i.c), intramuscular (i.m.), intradermal (i.d.), transdermal (t.d.) or intravenous (i.v.) administration, preferably by i.v., s.c., i.d., t.d., oral, inhaled, nasal or coupled to a carrier, preferably a red blood cell. It is particularly preferred to use the
- 25
- 30
- 35

protein or a fragment thereof and/or the at least one carrier coupled to the protein or the fragment thereof.

- 5 29. The at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence according to any of embodiments 1 to 5, and/or the at least one carrier according to any of embodiments 8 to 13 for inducing antigen-specific tolerance to autoantigens in early MS. It is particularly preferred to use the protein or a fragment thereof and/or the at least one carrier coupled to the protein or the fragment thereof.

CLAIMS

1. A GDP-L-fucose synthase protein or a protein of the RASGRP protein family, or a fragment, derivative or splice variant thereof, or a nucleotide sequence encoding any of the proteins or fragment, derivative or splice variant thereof, for use in the treatment, diagnosis and/or prevention of multiple sclerosis (MS).
2. The protein, fragment, derivative or splice variant for use according to claim 1, wherein the GDP-L-fucose synthase protein
- a) has the amino acid sequence as set forth in SEQ ID NO: 1 or
- b) has an amino acid sequence which is at least 85 %, preferably at least 90 %, more preferably at least 95 % identical to the amino acid sequence as set forth in SEQ ID NO: 1 or
- c) has an amino acid sequence which is at least 70 %, preferably at least 80 %, more preferably at least 90 % homologous to the amino acid sequence as set forth in SEQ ID NO: 1 or
- d) has an amino acid sequence which is at least 60 %, preferably at least 70 %, more preferably at least 80 %, even more preferably at least 90 % homologous to the amino acid sequence as set forth in SEQ ID NO: 1 and the protein or fragment or splice variant thereof binds to an autologous HLA allele, is recognized by a T cell and/or is recognized by an antibody which binds to or recognizes the amino acid sequence as set forth in SEQ ID NO: 1 or a fragment thereof or
- e) is encoded by a TSTA3 gene, in particular by a gene sequence of nucleotides 143612618 to 143618048 of NC_000008.11, or is encoded by a gene which is at least 80 %, preferably at least 90 %, even more preferably at least 95 % identical to the gene sequence of nucleotides 143612618 to 143618048 of NC_000008.11 and/or
- wherein the member of the RASGRP protein family
- f) has the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 or SEQ ID NO: 9 or
- g) has an amino acid sequence which is at least 85 %, preferably at least 90 %, more preferably at least 95 % identical to the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 or SEQ ID NO: 9 or
- h) has an amino acid sequence which is at least 70 %, preferably at least 80 %, more preferably at least 90 % homologous to the amino acid sequence as set forth in any of

SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 or SEQ ID NO: 9 or

i) has an amino acid sequence which is at least 60 %, preferably at least 70 %, more preferably at least 80 %, even more preferably at least 90 % homologous to the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 or SEQ ID NO: 9 and the protein or fragment or splice variant thereof binds to an autologous HLA allele, is recognized by a T cell and/or is recognized by an antibody which binds to or recognizes the respective amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 or SEQ ID NO: 9 or a fragment thereof or

j) is encoded by a RASGRP gene, in particular by a gene sequence of nucleotides

- 38488101 to 38565575 of NC_000015.10,
- 64726911 to 64745456 of NC_000011.10,
- 33436324 to 33564750 of NC_000002.12, or
- 38409051 to 38426305 of NC_000019.10,

or is encoded by a gene which is at least 80 %, preferably at least 90 %, even more preferably at least 95 % identical to the gene sequence of nucleotides

- 38488101 to 38565575 of NC_000015.10,
- 64726911 to 64745456 of NC_000011.10,
- 33436324 to 33564750 of NC_000002.12, or
- 38409051 to 38426305 of NC_000019.10.

3. The protein, fragment, derivative or splice variant for use according to claim 1 or 2, wherein the fragment comprises 5 to 50, preferably 5 to 20, more preferably 10 to 15 amino acids, even more preferably 15 amino acids.

4. The protein, fragment, derivative or splice variant for use according to claim 2 or 3, wherein the fragment is

a) at least 85 %, preferably at least 90 %, more preferably at least 95 % identical to a respective corresponding amino acid sequence or

b) at least 70 %, preferably at least 80 %, more preferably at least 90 % homologous to a respective corresponding amino acid sequence or

c) at least 60 %, preferably at least 70 %, more preferably at least 80 %, even more preferably at least 90 % homologous to a respective corresponding amino acid sequence and binds to an autologous HLA allele, is recognized by a T cell and/or is

recognized by an antibody which binds to or recognizes the respective amino acid sequence.

- 5 5. The protein, fragment, derivative or splice variant for use according to any of the preceding claims, wherein the fragment comprises a sequence selected from the group comprising SEQ ID NOs: 10 to 98, preferably SEQ ID NOs: 10 to 35, preferably consists of a sequence selected from the group comprising SEQ ID NO: 10 to 98, preferably SEQ ID NOs: 10 to 35.
- 10 6. The protein, fragment, derivative or splice variant for use according to any of the preceding claims for identifying a human subject who is suitable for tolerization to autoantigens in MS, preferably early MS.
- 15 7. The protein, fragment, derivative or splice variant for use according to any of claims 1 to 5 for diagnosing pattern II MS in a human subject.
- 20 8. A carrier comprising at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence according to any of claims 1 to 5 for use in the treatment, diagnosis and/or prevention of MS.
- 25 9. The carrier according to claim 8, wherein the carrier is coupled to the at least one protein, fragment, derivative and/or splice variant, and/or the carrier contains the at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence.
- 30 10. The carrier according to claim 8 or 9, wherein the carrier is selected from the group comprising a cell, preferably a blood cell, a protein, a lipid, a glycolipid, a bead, a nanoparticle, a virus-like-particle (VLP) and a molecule, such as a sugar molecule, and any combination thereof.
- 35 11. The carrier according to claim 10, wherein the protein, fragment, derivative and/or splice variant is expressed by the cell, preferably the blood cell.
12. The carrier according to claim 10 or 11, wherein the blood cell is a red or white blood cell.

13. The carrier according to any of claims 10 to 12, wherein the carrier is a blood cell and the blood cell is chemically coupled by a coupling agent, preferably by 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (ECDI/EDC), to the at least one protein, fragment, derivative and/or splice variant.

5

14. A method of manufacturing the chemically coupled blood cell of claim 13, comprising isolating the blood cell from a human subject, adding the at least one protein, fragment, derivative and/or splice variant and subsequently adding the coupling agent, preferably EDC.

10

15. A pharmaceutical composition comprising at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence according to any of claims 1 to 5 and a pharmaceutically acceptable carrier.

15

16. A method for inducing antigen-specific tolerance to autoantigens in a human subject suffering from or at risk of developing MS comprising the step of applying to the human subject

a) at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence according to any of claims 1 to 5, and/or

20

b) at least one carrier according to any of claims 8 to 13.

17. The method according to claim 16, wherein the at least one protein, fragment, derivative, splice variant, nucleotide sequence and/or gene sequence is applied by nasal, inhaled, oral, subcutaneous (s.c.), intracoelomic (i.c), intramuscular (i.m.), intradermal (i.d.), transdermal (t.d.) or intravenous (i.v.) administration, preferably by i.v., s.c., i.d., t.d., oral, inhaled, nasal or coupled to a carrier, preferably a red blood cell.

25

18. The method according to claim 16 or 17 for inducing antigen-specific tolerance to autoantigens in early MS.

30

19. A method for identifying a human subject suitable for tolerization to autoantigens in MS, preferably early MS, comprising isolating T cells and/or antibodies from blood, CSF or other body fluid of the subject and measuring reactivity of the T cells and/or antibodies against the protein, fragment, derivative and/or splice variant according to any of claims 1 to 5.

35

20. The fragment according to any of claims 3 to 5, preferably the fragment according to claim 5, for use as a medicament.