Title: Human Betacoronavirus Lineage C and identification of N-terminal dipeptidyl peptidase as its virus receptor

The invention provides a new previously undescribed Coronavirus isolated from cases of unexplained disease in September 2012 and identified herein as belonging to a newly recognized and previously undescribed species of human Corona Virus (HCoV), herein identified as HCoV-SA1 or HCoV EMC or Middle East Respiratory Syndrome-Coronavirus (MERS-CoV). In particular the nucleic acid and/or amino acid sequences of the MERS-CoV genome and sequences specifically encoding (parts of) viral proteins and antigenic polypeptides are provided. Further, the invention relates to diagnostic means and methods, prophylactic means and methods and therapeutic means and methods to be employed in the diagnosis, prevention and/or treatment of disease, in particular of respiratory disease, in particular of mammals, more in particular in humans. It particularly also relates to an isolated virus and its receptor.

A fundamental yet unresolved puzzle in virology is how viruses evolve to recognize their receptor proteins on the cells they need to enter in order to replicate. Specifically, how do different viruses recognize the same receptor protein and how do similar viruses recognize different receptor proteins? Do viruses select their receptor proteins by chance or do they target specific virus binding hotspots on these receptor proteins? Structural information of virus–receptor interfaces can potentially answer these questions. To date, although a few studies have obtained structural information for a single virus–receptor interface, even less studies have provided structural information for the interfaces between different viruses and their common receptor protein.

The invention in particular relates to coronaviruses that are the second leading cause of adult colds. Of the more than 30 kinds, three or four infect humans. The 2003 SARS virus is a coronavirus. Coronaviruses are rather difficult to grow in the laboratory, so they have not been studied to the same extent as other viruses. NL63 coronavirus (NL63 CoV), a prevalent human respiratory virus, is a group I coronavirus known to use angiotensin converting enzyme 2 (ACE2, a cell membrane bound carboxy terminal dipeptidyl peptidase) as its receptor. Incidentally, ACE2 is also used by group II SARS coronavirus (SARS CoV).

The distribution of coronavirus receptors is critical to the pathogenic outcome of the disease they cause. In this regard, it is notable that coronavirus spikes exhibit a wide range of receptor specificities; human aminopeptidase N (a metalloprotease) is a receptor for human coronavirus 229E, mouse hepatitis virus enters after binding members of a pleiotropic family of carcinoembryonic antigen cell adhesion molecules (CEACAMs); feline and porcine coronaviruses also bind various metalloproteases; and bovine coronaviruses recognize 9 O acetylated sialic acids.

Coronaviruses enter cells through a large spike protein on their envelopes. The coronavirus spike protein is a membrane anchored trimer and contains two subunits, receptor binding subunit S1 and membrane fusion subunit S2. The S2 subunits from group I and group II coronaviruses share both sequence and structural homology; they contain homologous heptad repeat segments that fold into a conserved trimers of hairpin structure, which is essential for membrane fusion. Surprisingly, the S1 subunits from group I and group II coronaviruses have no obvious sequence homology. Nevertheless, they can be divided approximately into N terminal region, central region, and C terminal region. Coronaviruses are believed to have common ancestors because they share similar replication mechanisms, genomic structures, and overall gene sequences.

Among all of the coronavirus genes, the one encoding the spike protein is the most variable. Between the spike protein subunits, S1 is more variable than S2. The current structural divergences of the S1

subunits reveal the tremendous evolutionary pressure that coronaviruses face to adapt to different host receptors, and they also reflect on the evolutionary history of coronaviruses and their receptor selections.

In general, coronaviruses are well known and most of those who are diagnosed with it recover completely with no complications after receiving the needed supportive therapy. However, in some of the patients who are infected, serious complications can develop affecting the respiratory system and the kidneys and can cause death, especially among the elderly and in patients with chronic respiratory and cardiac conditions and among immune compromised patients.

Coronaviruses (CoVs), a genus of the Coronaviridae family, are positive strand RNA viruses with the largest viral genome of all RNA viruses (27 – 32 Kb). The genomic RNA is capped, polyadenylated and covered with nucleocapsid proteins. The virus is enveloped and carries large spike glycoproteins. All CoVs employ a common genome organization where the replicase gene encompasses the 5′ two thirds of the genome and is comprised of two overlapping open reading frames (ORFs), ORF1a and ORF1b.

The structural gene region, which covers the 3′ third of the genome, encodes the canonical set of structural protein genes in the order 5′ spike (S) envelope (E) membrane (M) and nucleocapsid (N) – 3′. Some beta CoVs carry an additional structural protein encoding a hemagglutinin esterase (HE). The gene is located between the ORF1b and S gene. Expression of the nonstructural replicase proteins is mediated by translation of the genomic RNA that gives rise to the biosynthesis of two large polyproteins, pp1a (encoded by ORF1a) and pp1ab (encoded by ORF1a and ORF1b) facilitated by a ribosomal frame shift at the ORF1a/1b junction.

In contrast, the structural proteins are translated from sub genomic (sg) mRNAs. These sg mRNAs are the result of discontinuous transcription, a hallmark of CoV gene expression. The structural gene region also harbors several ORFs that are interspersed along the structural protein coding genes. The number and location of these accessory ORFs varies between the CoV species.

Although coronaviruses were first identified nearly 60 years ago, they only received notoriety in 2003 when one of their members was identified as the aetiological agent of severe acute respiratory syndrome (SARS). Previously these viruses were known to be important agents of respiratory and enteric infections of domestic and companion animals and to cause approximately 15% of all cases of the common cold. Coronaviruses (CoVs), a genus of the Coronaviridae family, are positive strand RNA viruses with the largest viral genome of all RNA viruses (27 – 32 Kb). The genomic RNA is capped, polyadenylated and covered with nucleocapsid proteins. The virus is enveloped and carries large spike glycoproteins. All CoVs employ a common genome organization where the replicase gene encompasses the 5'-two thirds of the genome and is comprised of two overlapping open reading frames (ORFs), ORF1a and ORF1b. The structural gene region, which covers the 3'-third of the genome, encodes the canonical set of structural protein genes in the order 5' - spike (S) - envelope (E) - membrane (M) and nucleocapsid (N) – 3'. Some beta-CoVs carry an additional structural protein encoding a heamagglutinin-esterase (HE). The gene is located between the ORF1b and S gene. Expression of the nonstructural replicase proteins is mediated by translation of the genomic RNA that gives rise to the biosynthesis of two large polyproteins, pp1a (encoded by ORF1a) and pp1ab (encoded by ORF1a and ORF1b) facilitated by a ribosomal frame shift at the ORF1a/1b junction. In contrast, the structural proteins are translated from sub genomic (sg) mRNAs. These sg mRNAs are the result of discontinuous transcription, a hallmark of CoV gene expression. The structural gene region also harbors several ORFs that are interspersed along the structural protein coding genes. The number and location of these accessory ORFs varies between the CoV species. In animals CoV infections can lead to a variety of syndromes, e.g. bronchitis, gastroenteritis, progressive demyelinating encephalitis, diarrhea, peritonitis and

respiratory tract disease. The first reports on human CoVs (HCoV) appeared in the mid-1960s. The human viruses were isolated from persons with common cold, and two species were detected: HCoV-229E and HCoV-OC43. Almost 40 years later, SARS-CoV was identified as the causative agent of the Severe Acute Respiratory Syndrome (SARS). A highly effective global public health response prevented further spread of this virus, and as a result SARS-CoV was eradicated from the human population. Soon thereafter it became clear that there are more HCoVs. HCoV-NL63 was identified in 2004 and HCoV-HKU1 in 2005. Both viruses are not emerging viruses like SARS-CoV but were previously unidentified. In fact, infections by these viruses are as common and wide spread as HCoV-229E and HCoV-OC43 infections. The SARS outbreak intensified the research on the unknown animal CoVs. As much as 16 new animal CoV species were identified till 2008. There are currently at around 29 complete reference genome sequences available in Genbank of the various viruses. Recently, the Coronavirus Study Group of the International Committee for Taxonomy of Viruses has proposed renaming the traditional group 1, 2, and 3 coronaviruses into the genus Alphacoronavirus, Betacoronavirus, and Gammacoronavirus, respectively (http://talk.ictvonline.org/media/p/1230.aspx). Each genus is subdivided into different species on the basis of sequence identity in the replicase domains of the polyprotein pp1ab.

The classification of the family Coronaviridae and the organization of the established subfamily Coronavirinae is based upon rooted phylogeny and pair-wise comparisons using Coronaviridae-wide conserved domains in replicase polyprotein pp1ab as well as the structural proteins S, E, M and N. In rooted trees, the proposed genera Alpha-, Beta- and Gammacoronavirus consistently form three distinct monophyletic groups and in pair-wise comparisons, they form three robust non-overlapping clusters. The inter-group pair-wise scores for coronaviruses are comparable to those calculated for structural and non-structural proteins of different genera in other RNA virus families (e.g. Potyviridae, Picornaviridae). Based on this defacto criterion phylogroups 1 through 3 are named into genera designated Alpha-, Beta and Gammacoronavirus, respectively. The 90% aa sequence identity threshold now proposed as a species demarcation criterion within each genus has been determined from the analysis of pair-wise aa distances in seven conserved replicase domains (nsp3 ADRP, nsp5 (3CLpro), nsp12 (RdRp), nsp13 (He11), nsp14 (ExoN), nsp15 (NendoU) and nsp16 (O-MT)) of 156 viruses in the Coronaviridae. In this analysis, 20 distinct groups (17 coronaviruses, 2 toroviruses, 1 bafinivirus) are unambiguously recognized as non-overlapping clusters (with the largest intra-cluster distance being smaller than the smallest inter-cluster distance). Of these clusters, at least 7 fall into the genus Betacoronavirus, each of which represents a distinct betacoronavirus species (Betacoronavirus 1, Murine coronavirus, Human coronavirus HKU1 , Rousettus bat coronavirus HKU9, Tylonycteris bat coronavirus HKU4, Pipistrellus bat coronavirus HKU5, Severe acute respiratory syndrome-related coronavirus (SARS-CoV). The Betacoronavirus genus is additionally considered to contain 4 lineages (A, B, C and D). Human coronaviruses HCoV-HKU1 and HCoV-OC43 belong to lineage A while human coronavirus SARS-CoV belongs to lineage B. Lineage C and D are not known to contain any human representatives. Other human coronaviruses, such as HCoV-NL63 and HCoV-229E, are even more distinct since these two human pathogens belong to a different genus, the Alphacoronavirus genus.

The invention also relates to so called “pull down” experiments, which are methods for the identification of protein protein interactions based on affinity purification of interacting proteins from complex proteinaceous substances such as cellular extracts. Pull down experiments with, for example, fusion proteins attached to inert beads are a screening technique for isolating proteinaceous substances having specific protein components that bind to each other and thus lead to identification of protein protein interactions.

Typically, pull down experiments are used to identify interactions between a probe protein and unknown targets and to confirm suspected interactions between a probe protein and a known protein. When

coupled with peptide digests of pulled down proteins and with mass spectrometry to sequence those peptides and identify targets, pull downs can be considered as the protein based equivalent of a yeast two hybrid screen.

To improve the isolation of specific binding partners, pull down methods have been developed involving the use of cross linking, of large scale tissue lysates, and of spin columns. Appropriate methods of sample preparation for mass spectrometry based identification of interacting proteins have been developed as well, including specialized gel staining techniques, band excision, and in gel tryptic digestion. Data interpretation and most commonly encountered problems are, for example, discussed in Current Protocols in Cell Biology, "UNIT 17.5 Protein Protein Interactions Identified by Pull Down Experiments and Mass Spectrometry," Adam Brymora, Valentina A. Valova, and Phillip J. Robinson, Published Online: 1 May 2004 DOI: 10.1002/0471143030.cb1705s22, and included herein by reference.

The invention

Summary

According to a first aspect, there is provided an essentially mammalian positive-sense single stranded RNA virus comprising one or more of the nucleic acid or comprising one or more of the amino acid sequences selected from figure 3 depicting the partial open reading frame of HCoV-SA1, figure 5 file N.rtf depicting the nucleocapsid (N) protein, figure 6 file M.rtf depicting the matrix (M) protein, figure 7 file E.rtf depicting the small envelope € protein, figure 8 file NS3d.5 rtf depicting the non-structural gene NS3d, figure 9 file NS3c.rtf depicting the non-structural gene NS3c, figure 10 file NS3b.rtf depicting the non-structural gene NS3b, figure 11 file NS3a.rtf depicting the non-structural gene NS3a, figure 12, file S.rtf depicting the spike surface glycoprotein (S), figure 13 file Orf1ab.rtf, encoding many enzymatic products among which the replicase, figure 14 file HCoV-SA1.rtf depicting isolate HCoV-SA1 or figure 15 HCoVSA1. rtf depicting its 3 translation frames, or comprising amino acid sequences having at least 90% identical with a nucleic acid or amino acid sequence depicted in said figure 3 or figure 5 to 15.

According to a second aspect, there is provided a nucleic acid, preferably a cDNA, obtainable from a virus as defined herein, said nucleic acid having at least an identity of 75% with a nucleic acid encoding a protein selected from any of the figure 3 or 5 - 15.

According to a third aspect, there is provided a vector comprising a nucleic acid as defined herein.

According to a fourth aspect, there is provided a cell comprising a virus as defined herein, a nucleic acid as defined herein or a vector as defined herein.

According to a fifth aspect, there is provided a protein encoded by a nucleic acid as defined herein.

According to a sixth aspect, there is provided an antigen comprising a protein as defined herein.

According to a seventh aspect, there is provided an antibody specifically directed against an antigen as defined herein.

According to an eighth aspect, there is provided a method for identifying a viral isolate as a MERS-CoV comprising reacting said viral isolate with a nucleic acid as defined herein and/or with an antibody as defined herein.

According to a ninth aspect, there is provided a method for virologically diagnosing a MERS-CoV infection of a mammal comprising determining in a sample of said mammal the presence of a viral isolate by reacting said sample with a nucleic acid as defined herein or an antibody as defined herein.

According to a tenth aspect, there is provided a method for serologically diagnosing a MERS-CoV infection of a mammal comprising determining in a sample of said mammal the presence of an antibody specifically directed against a Betacoronavirus, preferably Lineage C virus or component thereof by reacting said sample with a protein or fragment thereof as defined herein or an antigen as defined herein.

According to an eleventh aspect, there is provided a diagnostic kit for diagnosing a MERS-CoV infection comprising a virus as defined herein, and/or a nucleic acid as defined herein, and/or a protein as defined herein, and/or an antigen as defined herein and/or an antibody as defined herein.

According to a twelfth aspect, there is provided use of a virus as defined herein, and/or a nucleic acid as defined herein, and/or a vector as defined herein, and/or a cell as defined herein, and/or a protein or fragment thereof as defined herein, and/or an antigen as defined herein, and/or or an antibody as defined herein for use in the treatment or prevention of a Betacoronavirus, infection with a MERS-Cov.

According to a thirteenth aspect, there is provided a pharmaceutical composition comprising a virus as defined herein, and/or a nucleic acid as defined herein , and/or a vector as defined herein , and/or a cell as defined herein , and/or a protein or fragment thereof as defined herein , and/or an antigen as defined herein , and/or or an antibody as defined herein .

According to a fourteenth aspect, there is provided a pharmaceutical composition as defined herein for use in the treatment or prevention of a MERS-CoV, infection in a mammal, and/or in the treatment or prevention of atypical pneumonia and/or renal failure caused by such a MERS-CoV infection in a mammal.

According to a fifteenth aspect, there is provided a proteinaceous substance having been provided with a protein as defined herein and additionally comprising an N-terminal dipeptidyl peptidase protein.

According to a sixteenth aspect, there is provided a container provided with a virus as defined herein, and/or a nucleic acid as defined herein, and/or a vector as defined herein, and/or a cell as defined herein, and/or a protein or fragment thereof as defined herein, and/or an antigen as defined herein, and/or an antibody as defined herein and/or a pharmaceutical composition as defined herein and/or a substance as defined herein.

According to a seventeenth aspect, there is provided a method of identifying a candidate modulator as an agent that modulates the function or the binding of a virus to a dipeptidyl peptidase, said method comprising: a. providing a substance as defined herein in the presence and absence of said candidate modulator or said candidate antiviral agent under conditions permitting binding of a virus with a peptidase protein; b. measuring binding of said virus to said protein, wherein a decrease or increase in binding in the presence of said candidate modulator or said antiviral agent, relative to binding in the absence of said candidate modulator, identifies said candidate modulator as an agent that modulates the function of a dipeptidyl peptidaseor identifies

said antiviral agent as an agent that modulates that function of a dipeptidyl peptidase, preferably wherein said virus and/or said protein is detectably labeled, preferably with a moiety selected from the group consisting of a radioisotope, a fluorophore, a quencher of flurescence, an enzyme, and an affinity tag.

Description

The invention provides an isolated essentially mammalian positive-sense single stranded RNA virus classifiable as belonging to the Order: Nidovirales; Family: Coronaviridae; Subfamily: Coronavirinae; Genus: Betacoronavirus; and non-Lineage A, non-Lineage B or non-Lineage D, human betacoronavirus. The invention also provides a human virus (the prefix human stands for the virus classified as capable of infecting humans where a bat virus is capable of infecting bats), preferably isolated or isolatable from humans, having a receptor binding domain (RBD) capable of binding to a dipeptidyl peptidase 4 (DPP4), no such isolates have been deposited or in any other way made available to the art until now. In particular, the invention provides an isolated essentially mammalian positive-sense single stranded RNA virus classifiable as belonging to the Order: Nidovirales; Family: Coronaviridae; Subfamily: Coronavirinae; Genus: Betacoronavirus; Lineage C human betacoronavirus. The invention also provides an isolated essentially mammalian positive-sense single stranded RNA virus classifiable as belonging to the Order: Nidovirales; Family: Coronaviridae; Subfamily: Coronavirinae; Genus: Betacoronavirus; Lineage: C and isolatable from humans, and components thereof. Until now, no Betacoronavirus isolates have been isolated from humans that were then classified as belonging to Lineage: C of Betacoronavirus. In particular, the invention provides a Lineage: C human Betacoronavirus having a receptor binding domain (RBD) capable of binding to a dipeptidyl peptidase 4 (DPP4). In particular, no such isolates have been deposited or in any other way made available to the art until now. In a preferred embodiment, a virus according to the invention is isolated or isolatable from a human. In particular, the invention provides a new previously undescribed Coronavirus isolated from cases of unexplained disease in September 2012 and identified herein as belonging to a newly recognized and previously undescribed species of human Corona Virus (HCoV), herein identified as HCoV-SA1 or HCoV EMC or Middle East Respiratory Syndrome-Coronavirus (MERS-CoV). In particular the specific nucleic acid and/or amino acid sequences of the MERS-CoV genome and sequences encoding (parts of) viral proteins and antigenic polypeptides are provided, as demonstrated by phylogenetic analyses. Further, the invention relates to diagnostic means and methods, prophylactic means and methods and therapeutic means and methods to be employed in the diagnosis, prevention and/or treatment of disease, in particular of respiratory disease, in particular of mammals, more in particular in humans, most in particular specific for MERS-CoV. It particularly also relates to an isolated virus and its receptor. The invention also provides identification of N-terminal dipeptidyl peptidase as virus receptor and uses thereof, identification of the receptor binding domain of MERS-CoV mapping to a 231-residue region 2 in the spike protein that efficiently elicits neutralizing antibodies identification and uses thereof and dipeptidyl peptidase 4 receptor determinants of respiratory MERS-coronavirus infection, and uses thereof. The invention in particular provides specific diagnostics of MERS-CoV, sub-unit compositions of S1-MERS CoV protein

for vaccine purposes, screening tests for detecting compounds capable of interfering with MER-CoV-DPP4 binding, and animal models for determining activity of compounds capable of interfering with MERS-CoV-DPP4 binding.

The invention also provides a virus according to the invention comprising one or more of a nucleic acid or fragment thereof selected from any of figures 3 or 5 to 15, preferably wherein said virus provided herein is having an amino acid sequence of its conserved replicase domain that is at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with the amino acid sequence of the conserved replicase domain of an isolated essentially mammalian positive-sense single stranded RNA virus classifiable as belonging to the Order: Nidovirales; Family: Coronaviridae; Subfamily: Coronavirinae; Genus: Betacoronavirus; Lineage C isolatable from humans and comprising one or more of the sequences selected from any of figures 3 or 5 to 15, preferably wherein said conserved replicase domain comprises ORF1ab.

The invention also provides a virus according to the invention having an amino acid sequence of its receptor binding domain that is at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with the amino acid sequence of the receptor binding domain of an isolated essentially mammalian positive-sense single stranded RNA virus classifiable as belonging to the Order: Nidovirales; Family: Coronaviridae; Subfamily: Coronavirinae; Genus: Betacoronavirus; Lineage C isolatable from humans and comprising one or more of the sequences selected from any of figures 3 or 5 to 15, preferably wherein said receptor domain comprises residues 1 747 of the S1 spike protein, preferably residues 358 588 of the S1 spike protein.

In a one embodiment, a virus is provided that belongs to the Coronaviruses, genus Betacoronavirus and is identifiable as phylogenetically corresponding or specific to the MERS-CoV thereto by determining a nucleic acid or amino acid sequence of said virus or fragments thereof and testing it in phylogenetic tree analyses wherein maximum likelihood trees are generated and finding it, the virus or fragment, to be more closely phylogenetically corresponding to a virus isolate or fragment thereof having the sequences as depicted in any of figures 3 or 5 to 15 than it is corresponding to a bat coronavirus HKU4 or HKU5, or fragments thereof, in another embodiment, a virus is provided that belongs to the Coronaviruses and is identifiable as phylogenetically corresponding or specific to the MERS-CoV thereto by determining a nucleic acid sequence or amino acid sequence of said virus and testing it in phylogenetic tree analyses wherein maximum likelihood trees are generated and finding it to be more closely phylogenetically corresponding to a virus isolate isolatable from humans having the sequences as depicted in any of figures 3 or 5 to 15 than it is corresponding to a human coronavirus virus isolate HCoV-HKU1 or HCoV-OC43 or SARS-CoV, or fragments thereof.

In a preferred embodiment, a virus is provided herein that belongs to the Coronaviruses, genus Betacoronavirus and is identifiable as phylogenetically corresponding thereto by determining a nucleic acid or amino acid sequence of said virus and testing it in phylogenetic tree analyses wherein maximum likelihood trees are generated and finding it to be more closely phylogenetically corresponding to a virus isolate having the sequences as depicted in any of figures 3 or 5 to 15 than it is corresponding to a bat coronavirus HKU4 or HKU5 or to a human coronavirus virus isolate HCoV-HKU1 or HCoV-OC43 or SARS-CoV.

The invention also provides a cell, preferably a host cell, and a culture of such a cell or host cell, i.e. a cultured cell, comprising a virus according to the invention. Preferred examples of such cells and cell cultures comprise a Vero cell or LLC-MK2 cell and cultures thereof; other preferred examples comprise a Huh-7 cell, a primary nonciliated human airway epithelial cell, a primary human fibroblast, a primary human kidney cell, a primary human alveolar type 2 cell, or a primary kidney cell of Pipistrellus pipistrellu, and cultures of said cells.

The invention also provides a nucleic acid, preferably a cDNA, or MERS-CoV-specific fragment thereof obtainable, derived or obtained from an isolated essentially mammalian positive-sense single stranded

RNA virus classifiable as belonging to the Order: Nidovirales; Family: Coronaviridae; Subfamily: Coronavirinae; Genus: Betacoronavirus; and non-Lineage A, non-Lineage B or non-Lineage D, human betacoronavirus. In a preferred embodiment, the invention provides a nucleic acid isolatable from a human virus, preferably isolatable from humans, having a receptor binding domain (RBD) capable of binding to a dipeptidyl peptidase 4 (DPP4), In particular, a nucleic acid is provided by the invention obtainable, derived or obtained from an isolated essentially mammalian positive-sense single stranded RNA virus classifiable as belonging to the Order: Nidovirales; Family: Coronaviridae; Subfamily: Coronavirinae; Genus: Betacoronavirus; Lineage C human betacoronavirus. a nucleic acid is provided by the invention obtainable, derived or obtained from an isolated essentially mammalian positive-sense single stranded RNA virus classifiable as belonging to the Order: Nidovirales; Family: Coronaviridae; Subfamily: Coronavirinae; Genus: Betacoronavirus; Lineage: C and isolatable from humans, and components thereof. Until now, no Betacoronavirus isolates have been isolated from humans that were then classified as belonging to Lineage: C of Betacoronavirus. In particular, a nucleic acid is provided by the invention obtainable, derived or obtained from a Lineage: C Betacoronavirus having a receptor binding domain (RBD) capable of binding to a dipeptidyl peptidase 4 (DPP4), preferably from a virus having a nucleic acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with a nucleic acid sequence provided in figure 3 or figures 5 to 15. In particular, a MERS-CoV specific fragment of a nucleic acid, RNA or DNA or cDNA is provided by the invention which comprises one or more of the sequences of MERS-CoV as depicted in figures 3, or 5 to 15 or a nucleic acid sequence which can hybridize with any of these sequences under stringent conditions. The invention also provides a vector comprising a nucleic acid according to the invention, and a host cell comprising a nucleic acid according to the invention or a vector according to the invention.

The invention also provides an isolated or recombinant proteinaceous molecule or MERS-CoV-specific fragment thereof encoded by a nucleic acid according to the invention. In a preferred embodiment, the invention provides a proteinaceous molecule or MERS-CoV-specific viral protein or fragment thereof encoded by a nucleic acid according to the invention. Useful proteinaceous molecules are for example derived from any of the genes or genomic fragments or open reading frames (ORFs) derivable from a virus according to the invention. Such molecules, or antigenic fragments thereof, as provided herein, are for example useful in diagnostic methods or kits and in pharmaceutical compositions such as sub-unit vaccines and inhibitory peptides. Particularly useful is the viral polymerase protein, the spike protein, the nucleocapsid or antigenic fragments thereof for inclusion as antigen or subunit immunogen in a vaccine, but inactivated whole virus can also be used. Particularly useful are also those proteinaceous substances that are encoded by recombinant nucleic acid fragments that are identified by phylogenetic analyses as being MERS-CoV specific fragments, of course preferred are those that are within the preferred bounds and metes of ORFs useful in phylogenetic analyses, in particular for eliciting MERS-CoV specific antibodies, whether in vivo (e.g. for protective purposes or for providing diagnostic antibodies) or in vitro (e.g. by phage display technology or another technique useful for generating synthetic antibodies).

In one embodiment, the invention provides a viral replicase or MERS-CoV-specific fragment thereof having an amino acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with an amino acid sequence provided in figure 13, said viral replicase or MERS-CoV-specific fragment thereof preferably encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof as provided herein, having a nucleic acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with a nucleic acid sequence provided in figure 13.

In another embodiment, the invention provides a viral spike protein or MERS-CoV-specific fragment thereof having an amino acid sequence at least 75%, more preferably at least 80%, more preferably at

least 85%, more preferably at least 90%, most preferably at least 95% identical with an amino acid sequence provided in figure 12, said viral spike protein or MERS-CoV-specific fragment thereof preferably encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof as provided herein, having a nucleic acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with a nucleic acid sequence provided in figure 12.

In another embodiment, the invention provides an S1 spike protein or MERS-CoV-specific fragment thereof having an amino acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with an amino acid sequence provided for residues 1 588 in figure 17.

In another embodiment, the invention provides an S1 spike protein or MERS-CoV-specific fragment thereof having an amino acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with an amino acid sequence provided for residues 1 357 in figure 17.

In another embodiment, the invention provides an S1 spike protein or MERS-CoV-specific fragment thereof having an amino acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with an amino acid sequence provided for residues 358 747 in figure 17.

In another embodiment, the invention provides an S1 spike protein or MERS-CoV-specific fragment thereof having an amino acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with an amino acid sequence provided for residues 358-588 in figure 17.

In another embodiment, the invention provides an S1 spike protein or MERS-CoV-specific fragment thereof having an amino acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with an amino acid sequence provided for residues 589 747 in figure 17.

In another embodiment, the invention provides an S1 spike protein or MERS-CoV-specific fragment thereof having an amino acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with an amino acid sequence provided for residues 1 747 in figure 17.

In another embodiment, the invention provides an S1 spike protein or MERS-CoV-specific fragment thereof having an amino acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with an amino acid sequence provided for residues 1 747 in figure 17.

In another embodiment, the invention provides a viral non-structural gene protein or MERS-CoV-specific fragment thereof having an amino acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with an amino acid sequence provided in figure 8, said viral non-structural gene protein or MERS-CoV-specific fragment preferably encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof as provided herein, having a nucleic acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with a nucleic acid sequence provided in figure 8.

In another embodiment, the invention provides a viral non-structural gene protein or MERS-CoV-specific fragment thereof having an amino acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with an amino acid sequence provided in figure 9, said viral non-structural gene protein or MERS-CoV-specific fragment preferably encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof as provided herein, having a

nucleic acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with a nucleic acid sequence provided in figure 9.

In another embodiment, the invention provides a viral non-structural gene protein or MERS-CoV-specific fragment thereof having an amino acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with an amino acid sequence provided in figure 10, said viral non-structural gene protein or MERS-CoV-specific fragment preferably encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof as provided herein, having a nucleic acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with a nucleic acid sequence provided in figure 10.

In another embodiment, the invention provides a viral non-structural gene protein or MERS-CoV-specific fragment thereof having an amino acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with an amino acid sequence provided in figure 11, said viral non-structural gene protein or MERS-CoV-specific fragment preferably encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof as provided herein, having a nucleic acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with a nucleic acid sequence provided in figure 11.

In another embodiment, the invention provides a viral small envelope (E) protein or MERS-CoV-specific fragment thereof having an amino acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with an amino acid sequence provided in figure 7, said viral small envelope (E) protein or MERS-CoV-specific fragment preferably encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof as provided herein, having a nucleic acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with a nucleic acid sequence provided in figure 7.

In another embodiment, the invention provides a viral matrix (M) protein or MERS-CoV-specific fragment thereof having an amino acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with an amino acid sequence provided in figure 6, said viral matrix (M) protein or MERS-CoV-specific fragment preferably encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof as provided herein, having a nucleic acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with a nucleic acid sequence provided in figure 6.

In another embodiment, the invention provides a nucleocapsid (N) protein or MERS-CoV-specific fragment thereof having an amino acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with an amino acid sequence provided in figure 5, said nucleocapsid (N) protein or MERS-CoV-specific fragment preferably encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof as provided herein, having a nucleic acid sequence at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with a nucleic acid sequence provided in figure 5.

The invention also provides an antigen comprising a proteinaceous molecule or MERS-CoV-specific fragment thereof as provided herein. In a preferred embodiment, said proteinaceous molecule comprises or consists of a nucleocapsid (N) protein or MERS-CoV-specific fragment thereof as provided herein, or a viral matrix (M) protein or MERS-CoV-specific fragment thereof as provided herein, or a viral small envelope (E) protein or MERS-CoV-specific fragment thereof as provided herein, or a viral non-structural gene protein or MERS-CoV-specific fragment as provided herein, or an S1 spike protein or MERS-CoV-specific fragment thereof as provided

herein, or a viral spike protein or MERS-CoV-specific fragment thereof as provided herein or of a viral replicase or MERS-CoV-specific fragment thereof as provided herein.

Also provided herein are antibodies, be it natural polyclonal or monoclonal, or synthetic (e.g. (phage) library-derived binding molecules) antibodies that specifically react with an antigen comprising a proteinaceous molecule or MERS-CoV-specific fragment thereof according to the invention. A person skilled in the art will be able to develop (monoclonal) antibodies using isolated virus material and/or recombinantly expressed viral proteins. In particular the invention provides a rabbit antibody specifically directed against an antigen according to the invention, rabbits being particularly well suited to raise antibodies against an antigen according to the invention. Such antibodies are also useful in a method for identifying a viral isolate as a MERS-CoV comprising reacting said viral isolate or a component thereof with an antibody as provided herein. This can for example be achieved by using purified or non-purified MERS-CoV or parts thereof (proteins, peptides) using ELISA, RIA, FACS or similar formats of antigen detection assays (Current Protocols in Immunology). Alternatively, infected cells or cell cultures may be used to identify viral antigens using classical immunofluorescence or immunohistochemical techniques. Specifically useful in this respect are antibodies raised against MERS-CoV proteins or peptides which are encoded by a nucleotide sequence comprising one or more of the fragments disclosed in figures 3 and 5 to 15. Antibodies, both monoclonal and polyclonal, or fragments thereof, can also be used for detection purpose in the present invention, for example, in immunoassays in which they can be utilized in liquid phase or bound to a solid phase carrier. In addition, the monoclonal antibodies in these immunoassays can be detectably labeled in various ways. A variety of immunoassay formats may be used to select antibodies specifically reactive with a particular protein (or other analyte). For example, solid-phase ELISA immunoassays are routinely used to select monoclonal antibodies specifically immunoreactive with a protein. See Harlow and Lane, Antibodies, A Laboratory Manual, Cold Spring Harbor Publications, New York (1988), for a description of immunoassay formats and conditions that can be used to determine selective binding. Examples of types of immunoassays that can utilize antibodies of the invention are competitive and non-competitive immunoassays in either a direct or indirect format. Examples of such immunoassays are the radioimmunoassay (RIA) and the sandwich (immunometric) assay. Detection of the antigens using the antibodies of the invention can be done utilizing immunoassays that are run in either the forward, reverse, or simultaneous modes, including immunohistochemical assays on physiological samples. Those of skill in the art will know, or can readily discern, other immunoassay formats without undue experimentation.

Antibodies can be bound to many different carriers and used to detect the presence of the target molecules. Examples of well-known carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, agaroses and magnetite. The nature of the carrier can be either soluble or insoluble for purposes of the invention. Those skilled in the art will know of other suitable carriers for binding monoclonal antibodies, or will be able to ascertain such using routine experimentation.

The invention also provides method for identifying a viral isolate as a Betacoronavirus, Lineage C comprising reacting said viral isolate or a component thereof with a nucleic acid according to the invention and/or with an antibody according to the invention. The invention for example provides a method for virologically diagnosing a MERS-CoV infection of an animal, in particular of a mammal, more in particular of a human being, comprising determining in a sample of said animal the presence of a viral isolate or component thereof by reacting said sample with a MERS-CoV specific nucleic acid or antibody according to the invention, and a method for serologically diagnosing a MERS-CoV infection of a mammal comprising determining in a sample of said mammal the presence of an antibody specifically directed against a MERS-CoV or component thereof by reacting said

sample with a MERS-CoV-specific proteinaceous molecule or fragment thereof or an antigen according to the invention.

The invention also provides a diagnostic kit for diagnosing a MERS-CoV infection comprising a MERS-CoV, a MERS-CoV-specific nucleic acid, proteinaceous molecule or fragment thereof, antigen and/or an antibody according to the invention, and preferably a means for detecting said MERS-CoV, MERS-CoV-specific nucleic acid, proteinaceous molecule or fragment thereof, antigen and/or an antibody, said means for example comprising an excitable group such as a fluorophore or enzymatic detection system used in the art (examples of suitable diagnostic kit format comprise IF, ELISA, neutralization assay, RT-PCR assay). To determine whether an as yet unidentified virus component or synthetic analogue thereof such as nucleic acid, proteinaceous molecule or fragment thereof can be identified as MERS-CoV-specific, it suffices to analyze the nucleic acid or amino acid sequence of said component, for example for a stretch of said nucleic acid or amino acid, preferably of at least 10, more preferably at least 25, more preferably at least 40 nucleotides or amino acids (respectively), by sequence homology comparison with the provided MERS-CoV nucleic acid or amino acid sequences and with known non-MERS-CoV nucleic acid or amino acid sequences using for example phylogenetic analyses as provided herein. Depending on the degree of relationship with said MERS-CoV or non-MERS-CoV viral sequences, the component or synthetic analogue can be identified.

The invention also provides use of a virus according to the invention, and/or a nucleic acid according to the invention , a vector according to the invention, a host cell according the invention, a proteinaceous molecule or fragment thereof according to the invention, an antigen according to the invention, or an antibody according to the invention for the production of a pharmaceutical composition, preferably for the production of a pharmaceutical composition for therapeutical use, preferably for use in antiviral therapy, preferably for the treatment or prevention of a Betacoronavirus, Lineage C virus infection, preferably a human infection, preferably an infection with a MERS-Cov, or for the production of a pharmaceutical composition for the treatment or prevention of atypical pneumonia and/or renal failure, preferably wherein said atypical pneumonia and/or renal failure is a human disease. Preferably a peptide comprising part of the amino acid sequence of the spike protein as depicted in figure 17 (residues 358-588, comprising the essential receptor binding domain) is used for the preparation of a therapeutic or prophylactic peptide, preferably for inclusion in said pharmaceutical composition. Also preferably, a protein comprising the amino acid sequence of the spike protein as depicted in figure 17 (residues 358-588) is used for the preparation of a sub-unit vaccine. Furthermore, the nucleocapsid of Coronaviruses, as depicted in figure 5, is known to be particularly useful for eliciting cell-mediated immunity against Coronaviruses and can be used for the preparation of a sub-unit vaccine. The invention also comprises a pharmaceutical composition comprising a virus according to the invention, and/or a nucleic acid according to the invention, a vector according to the invention, a host cell according the invention, a proteinaceous molecule or fragment thereof according to the invention, an antigen according to the invention, or an antibody according to the invention.

The invention also provides a method for the treatment or prevention of a Betacoronavirus, Lineage C virus infection or for the treatment or prevention of atypical pneumonia comprising providing a mammal, preferably a human individual with a pharmaceutical composition according to the invention . Also, the invention provides a method for the treatment or prevention of atypical pneumonia and/or renal failure comprising providing an individual with a pharmaceutical composition according to the invention. In a preferred embodiment, a method for the treatment or prevention of a MERS-CoV infection is provided comprising providing a mammal with a pharmaceutical composition according the invention, preferably wherein said mammal is a rabbit. The invention also provides a method for in vivo determining of parameters of MERS-CoV infection, preferably for determining parameters of MERS-CoV-DPP4 interaction in an animal experiment, comprising providing a mammal with a pharmaceutical composition according to the invention, and/or with a virus according to the

invention, and/or with a nucleic acid according to the invention, and/or with a vector according to the invention, and/or with a host cell according to the invention, and/or with a proteinaceous molecule or fragment thereof according to the invention, and/or with an antigen according to the invention, and/or with an antibody according to the invention, preferably wherein said mammal is a rabbit. It is herein found that rabbits have several advantages over other experimental animals in that they have a remarkably similar target sequence for MERS-CoV-DPP4-receptor interaction, resulting in proficient infection of a rabbit with MERS-CoV and thus ample chance to study various aspects and parameters of MERS-CoV-DPP4-receptor interaction that resemble those in humans, giving the rabbit experimental animal model a distinct advantage over other animal models, such as the ferret animal model. Phylogenetic analysis of the MERS-CoV binding region of DPP4 indicated that human, macaque, horse and rabbit DPP4 cluster together as do DPP4's from cattle, pig and bats, that are somewhat more distantly related. Small animals including ferret, mice and most likely hamsters, shown to resist MERS-CoV infection, are more divergent in the DPP4 virus binding region, which at least in the case of ferrets has consequences for MERS-CoV binding. Besides macaques, rabbits indeed are a potential animal model for MERS-CoV infection; ex vivo inoculation of rabbit lung and kidney tissues revealed susceptibility to MERS-CoV. Similarly, the invention provides a method for in vivo determining of parameters of MERS-CoV infection, preferably for determining parameters of protection against MERS-CoV-infection, comprising providing a mammal with a pharmaceutical composition according to the invention, and/or with a virus according to the invention, and/or with a nucleic acid according to the invention, and/or with a vector according to the invention, and/or with a host cell according to the invention, and/or with a proteinaceous molecule or fragment thereof according to the invention, and/or with an antigen according to the invention, and/or with an antibody according to the invention, preferably wherein said mammal is a rabbit. In particular, a rabbit model is a model of choice for testing a pharmaceutical composition comprising a subunit peptide vaccine comprising part of the amino acid sequence of the spike protein as depicted in figure 17 (fragments of residues 358-588, comprising the essential receptor binding domain) which is used for the preparation of a therapeutic or prophylactic peptide for the preparation of a sub-unit vaccine. Vaccinating or immunizing rabbits with variant peptide vaccines and then challenging vaccinated and control rabbits with MERS-CoV that allows rapid infection and measurement of essential parameters such as development of (neutralizing) antibodies in experimental and control rabbits, development of protection against MERS-CoV infection or against MERS-CoV transmission allows for relatively inexpensive and rapid vaccine development studies, thereby allowing rapid vaccine development against human MERS-CoV infections. Attenuation of the virus by serial passage of MERS-CoV can now preferably achieved in rabbits by established methods developed for this purpose, including but not limited to the use of related viruses of other species, serial passages through other laboratory animals or/and tissue/cell cultures, serial passages through cell cultures at temperatures below 37C (cold-adaption), site directed mutagenesis of molecular clones and exchange of genes or gene fragments between related viruses.

Now, as herein provided, a new human coronavirus was isolated from a patient with pneumonia. The virus was isolated from sputum of a male patient aged 60 years old presenting with pneumonia associated with acute renal failure. The virus grows readily on Vero cells and LLC-MK2 cells producing CPE in the form of rounding and syncytia formation and uses dipeptidyl peptidase 4 (DPP4) as a viral receptor for entry into cells establishing infection.

The clinical isolate was initially tested for influenza virus A, influenza virus B, parainfluenza virus, enterovirus and adenovirus, with negative results. Testing with a pancoronavirus RT-PCR yielded a band at a molecular weight appropriate for a coronavirus. The virus RNA was tested and it was confirmed to be a new member of the beta group of coronaviruses, closely related to bat coronaviruses. The invention relates to a new previously undescribed Coronavirus isolated from cases of unexplained disease in September 2012 and identified

herein as belonging to a newly recognized and previously undescribed species of human Corona Virus (HCoV), herein identified as HCoV-SA1. In particular the nucleic acid and/or amino acid sequences of the HCoV-SA1 genome and sequences encoding (parts of) viral proteins are provided. Further, the invention relates to diagnostic means and methods, prophylactic means and methods and therapeutic means and methods to be employed in the diagnosis, prevention and/or treatment of disease, in particular of respiratory disease and/or renal failure (atypical pneumonia), in particular of mammals, more in particular in humans.

In particular diagnostic tests for example useful in PCR and serology with nucleic acids (primers) and antibodies and other reagents that are specifically targeted at the nucleic acid or amino acid sequences of the HCoV-SA1 genome are herein provided. The invention also provides vectors, such as bacterial and viral vectors based on nucleic acid or amino acid sequences of the HCoV-SA1 genome. In addition, the invention also provides antigenic polypeptides based amino acid sequences of the HCoV-SA1 genome are herein provided.

Also, the invention provides vaccines against HCoV-SA1 (based on nucleic acid or amino acid sequences or antigenic polypeptides of the HCoV-SA1 genome, and the invention provides use of antiviral drugs directed against nucleic acid or amino acid sequences or polypeptides of the HCoV-SA1 genome, as herein provided.

As for yet it is not known if there is a cure for the disease. Several antiviral therapies have been applied, but with various results. Also, for being able to prevent spread of the disease, it is of great importance to be able to recognize the disease in an early stage. Only then sufficient measures can be taken to isolate patients and initiate quarantine precautions. At this moment there is not yet a diagnostic tool in place. Thus, there is great need in developing diagnostic tools and therapies for this disease.

As further described in the detailed description herein, the isolated essentially mammalian positive-sense single stranded RNA virus classifiable as belonging to the Order: Nidovirales; Family: Coronaviridae; Subfamily: Coronavirinae; Genus: Betacoronavirus; and non-Lineage A, non-Lineage B or non-Lineage D, human betacoronavirus here provided was isolated from a patient with pneumonia. The virus was isolated from sputum of a male patient aged 60 years old presenting with pneumonia associated with acute renal failure. The virus grows readily on Vero cells and LLC-MK2 cells producing CPE in the form of rounding and syncytia formation. It was classified as an isolated essentially mammalian positive-sense single stranded RNA virus classifiable as belonging to the Order: Nidovirales; Family: Coronaviridae; Subfamily: Coronavirinae; Genus: Betacoronavirus; Lineage C human betacoronavirus by comparison of its RNA sequences. It is remarkable that now, at about 9 years after the isolation of the SARS-virus (also related to bat coronavirus) another betacoronavirus has been isolated from humans.

The invention also provides an isolated essentially mammalian positive-sense single stranded RNA virus comprising one or more of the nucleic acid or comprising one or more of amino acid sequences selected from figure 3 depicting the partial open reading frame of HCoV-SA1, figure 5 file N.rtf depicting the nucleocapsid (N) protein, figure 6 file M.rtf depicting the matrix (M) protein, figure 7 file E.rtf depicting the small envelope (E) protein, figure 8 file NS3d.rtf depicting the non-structural gene NS3d, figure 9 file NS3c.rtf depicting the non-structural gene NS3c, figure 10 file NS3b.rtf depicting the non-structural gene NS3b, figure 11 file NS3a.rtf depicting the non-structural gene NS3a, figure 12 file S.rtf depicting the spike surface glycoprotein (S), figure 13 file Orf1ab.rtf encoding many enzymatic products among which the replicase, figure 14 file HCoV-SA1.rtf depicting isolate HCoV-SA1 or figure 15 HCoV-SA1.rtf depicting its 3 translation frames, or comprising amino acid sequences having at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with a nucleic acid or amino acid sequence depicted in said figure 3, or 5 to 15.

In particular, the invention provides an isolated positive-sense single stranded RNA virus belonging to the Coronaviruses, genus Betacoronavirus having an amino acid sequence of its conserved replicase domain that is at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical with the amino acid sequence of the conserved replicase domain of an isolated essentially mammalian positive-sense single stranded RNA virus classifiable as belonging to the Order: Nidovirales; Family: Coronaviridae; Subfamily: Coronavirinae; Genus: Betacoronavirus; Lineage C isolatable from humans and comprising one or more of the sequences selected from any of figures 3 or 5 to 15, preferably wherein said conserved replicase domain comprises ORF1ab, 90% identity being the species definition of the betacoronaviridae . Virus provided by the invention is herein also called HCoV-SA1 virus-like virus. The invention also provides an isolated positive-sense single stranded RNA virus belonging to the Coronaviruses, genus Betacoronavirus and identifiable as phylogenetically corresponding thereto by determining a nucleic acid or amino acid sequence of said virus and testing it in phylogenetic tree analyses wherein maximum likelihood trees are generated, preferably with 100 bootstraps and 3 jumbles, and finding it to be more closely phylogenetically corresponding to a virus isolate or nucleic acid having the sequences as depicted in any of figures 3 or 5 to 15 than it is corresponding to a bat coronavirus virus HKU4 or HKU5.

The invention also provides an isolated positive-sense single stranded RNA virus belonging to the Coronaviruses, genus Betacoronavirus and identifiable as phylogenetically corresponding thereto by determining a nucleic acid or amino acid sequence of said virus and testing it in phylogenetic tree analyses wherein maximum likelihood trees are generated, preferably with 100 bootstraps and 3 jumbles, and finding it to be more closely phylogenetically corresponding to a virus isolate or nucleic acid having the sequences as depicted in any of the oligonucleotide or amino acid sequences submitted to GENBANK under accession JX869059 (http://www.ncbi.nlm.nih.gov/nuccore/JX869059) than it is corresponding to any of the oligonucleotide or amino acid sequences of bat coronavirus virus HKU4 or HKU5.

Although phylogenetic analyses provide a convenient method of identifying a virus as a Betacoronavirus; Lineage C virus several other possibly more straightforward albeit somewhat more coarse methods for identifying said virus or viral proteins or nucleic acids from said virus are herein also provided. As a rule of thumb a Betacoronavirus; Lineage C virus can be identified by the percentages of homology of the virus, proteins or nucleic acids to be identified in comparison with viral proteins or nucleic acids identified herein in or in Genbank accession JX869059by sequence. It is generally known that virus species, especially RNA virus species, often constitute a quasi species wherein a cluster of said viruses displays heterogeneity among its members. Thus it is expected that each isolate may have a somewhat different percentage relationship with the sequences of the isolate as provided herein.

The invention in particular provides an isolated positive-sense single stranded RNA virus belonging to the Coronaviruses and identifiable as phylogenetically corresponding thereto by determining a nucleic acid sequence or amino acid sequence of said virus and testing it in phylogenetic tree analyses wherein maximum likelihood trees are generated and finding it to be more closely phylogenetically corresponding to a virus isolate or nucleic acid having the sequences as depicted in any of the oligonucleotide or amino acid sequences submitted to GENBANK under accession JX869059 (http://www.ncbi.nlm.nih.gov/nuccore/JX869059) than it is corresponding to any of the oligonucleotide or amino acid sequences of human coronavirus virus isolate HCoV-HKU1 or HCoV-OC43 or SARS-CoV.

The invention in particular provides an isolated positive-sense single stranded RNA virus belonging to the Coronaviruses and identifiable as phylogenetically corresponding thereto by determining a nucleic acid sequence or amino acid sequence of said virus and testing it in phylogenetic tree analyses wherein maximum likelihood trees are generated and finding it to be more closely phylogenetically corresponding to a virus

isolate isolatable from humans having the sequences as depicted in any of figures 3 or 5 to 15 than it is corresponding to a human coronavirus virus isolate HCoV-HKU1 or HCoV-OC43 or SARS-CoV.

The invention also provides a virus according to the invention wherein its positive-sense single stranded RNA nucleic acid sequence comprises an open reading frame (ORF) encoding a viral protein of said virus, preferably selected from the group of ORFs encoding the spike surface glycoprotein (S), the non-structural genes NS3a, NS3b, NS3c, NS3d, the small envelope (E) protein, the matrix (M) protein, and the nucleocapsid (N) protein. With the provision of the sequence information of this MERS-CoV, the invention provides diagnostic means and methods, prophylactic means and methods and therapeutic means and methods to be employed in the diagnosis, prevention and/or treatment of disease, in particular of respiratory disease (atypical pneumonia), in particular of mammals, more in particular in humans. In virology, it is most advisory that diagnosis, prophylaxis and/or treatment of a specific viral infection is performed with reagents that are most specific for said specific virus causing said infection. In this case this means that it is preferred that said diagnosis, prophylaxis and/or treatment of a Betacoronavirus; Lineage C virus infection is performed with reagents that are most specific for Betacoronavirus; Lineage C virus. This by no means however excludes the possibilities that less specific, but sufficiently cross-reactive reagents are used instead, for example because they are more easily available and sufficiently address the task at hand. The invention for example provides a method for virologically diagnosing a Betacoronavirus; Lineage C infection of an animal, in particular of a mammal, more in particular of a human being, comprising determining in a sample of said animal the presence of a viral isolate or component thereof by reacting said sample with a Betacoronavirus; Lineage C specific nucleic acid or antibody according to the invention, and a method for serologically diagnosing a Betacoronavirus; Lineage C infection of a mammal comprising determining in a sample of said mammal the presence of an antibody specifically directed against a Betacoronavirus; Lineage C virus or component thereof by reacting said sample with a Betacoronavirus; Lineage C MERS-CoV-specific proteinaceous molecule or fragment thereof or an antigen according to the invention. The invention also provides a diagnostic kit or other system for diagnosing a Betacoronavirus; Lineage C infection comprising a Betacoronavirus; Lineage C virus, a Betacoronavirus; Lineage C MERS-CoV-specific nucleic acid, proteinaceous molecule or fragment thereof, antigen and/or an antibody according to the invention, and preferably a means for detecting said Betacoronavirus; Lineage C virus, Betacoronavirus; Lineage C MERS-CoV-specific nucleic acid, proteinaceous molecule or fragment thereof, antigen and/or an antibody, said means for example comprising an excitable group such as a fluorophore or enzymatic detection system used in the art (examples of suitable diagnostic kit format comprise IF, ELISA, neutralization assay, RT-PCR assay). To determine whether an as yet unidentified virus component or synthetic analogue thereof such as nucleic acid, proteinaceous molecule or fragment thereof can be identified as Betacoronavirus; Lineage C -MERS-CoV-specific, it suffices to analyze the nucleic acid or amino acid sequence of said component, for example for a stretch of said nucleic acid or amino acid, preferably of at least 10, more preferably at least 25, more preferably at least 40 nucleotides or amino acids (respectively), by sequence homology comparison with the provided Betacoronavirus; Lineage C viral sequences and with known non- Betacoronavirus; Lineage C viral sequences (SARS is preferably used) using for example phylogenetic analyses as provided herein. Depending on the degree of relationship with said Betacoronavirus; Lineage C or non- Betacoronavirus; Lineage C viral sequences, (herein also called HCoV-SA1 virus-like virus sequences) the component or synthetic analogue can be identified. The invention also provides a virus according to the invention that is isolatable from a human with atypical pneumonia. Also, isolated or recombinant nucleic acid or MERS-CoV-specific fragments thereof are obtainable, derived or obtained from a virus according to the invention, as are a vector comprising a nucleic acid according to the invention, and a host cell comprising a nucleic acid or vector according to the invention.

The invention also provides an isolated or recombinant proteinaceous molecule or MERS-CoV-specific fragment thereof encoded by a nucleic acid according to the invention. In a preferred embodiment, the invention provides a proteinaceous molecule or MERS-CoV-specific viral protein or fragment thereof encoded by a nucleic acid according to the invention. Useful proteinaceous molecules are for example derived from any of the genes or genomic fragments derivable from a virus according to the invention. Such molecules, or antigenic fragments thereof, as provided herein, are for example useful in diagnostic methods or kits and in pharmaceutical compositions such as sub-unit vaccines and inhibitory peptides. Particularly useful is the viral polymerase protein, the spike protein, the nucleocapsid or antigenic fragments thereof for inclusion as antigen or subunit immunogen, but inactivated whole virus can also be used. Particularly useful are also those proteinaceous substances that are encoded by recombinant nucleic acid fragments that are identified for phylogenetic analyses, of course preferred are those that are within the preferred bounds and metes of ORFs useful in phylogenetic analyses, in particular for eliciting HCoV-SA1 virus-like virus specific antibodies, whether in vivo (e.g. for protective purposes or for providing diagnostic antibodies) or in vitro (e.g. by phage display technology or another technique useful for generating synthetic antibodies). Similarly, the invention provides an antigen comprising a proteinaceous molecule or MERS-CoV-specific fragment thereof according to the invention, or reactive with an antibody according to the invention.

Also provided herein are antibodies, be it natural polyclonal or monoclonal, or synthetic (e.g. (phage) library-derived binding molecules) antibodies that specifically react with an antigen comprising a proteinaceous molecule or HCoV-virus-like MERS-CoV-specific fragment thereof according to the invention. A person skilled in the art will be able to develop (monoclonal) antibodies using isolated virus material and/or recombinantly expressed viral proteins. Sui et al. (Proc. Natl. Acad. Sci. 101(8), 2536-2541, 2004) have transiently expressed fragments of the spike protein and found several antibodies through phage display methods. Such antibodies are also useful in a method for identifying a viral isolate as a HCoV-SA1 virus-like virus comprising reacting said viral isolate or a component thereof with an antibody as provided herein. This can for example be achieved by using purified or non-purified HCoV-SA1 virus-like virus or parts thereof (proteins, peptides) using ELISA, RIA, FACS or similar formats of antigen detection assays (Current Protocols in Immunology). Alternatively, infected cells or cell cultures may be used to identify viral antigens using classical immunofluorescence or immunohistochemical techniques. Specifically useful in this respect are antibodies raised against HCoV-SA1 virus-like virus proteins which are encoded by a nucleotide sequence comprising one or more of the fragments disclosed herein.

The invention also provides method for identifying a viral isolate as a Betacoronavirus, Lineage C comprising reacting said viral isolate or a component thereof with a nucleic acid according to the invention. Other methods for identifying a viral isolate as a HCOV-SA1 virus or MERS-CoV comprise reacting said viral isolate or a component thereof with a virus specific nucleic acid according to the invention

In this way the invention provides a viral isolate identifiable with a method according to the invention as a mammalian virus taxonomically corresponding to a positive-sense single stranded RNA virus identifiable as likely belonging to the HCOV-SA1 or MERS-CoV virus genus within the family of Coronaviruses.

The method is useful in a method for virologically diagnosing a HCOV-SA1 or MERS-CoV virus infection of a mammal, said method for example comprising determining in a sample of said mammal the presence of a viral isolate or component thereof by reacting said sample with a nucleic acid or an antibody according to the invention.

Methods of the invention can in principle be performed by using any nucleic acid amplification method, such as the. Polymerase Chain Reaction (PCR; Mullis 1987, U.S. Pat. No. 4,683,195, 4,683,202, en 4,800,159) or by using amplification reactions such as Ligase Chain Reaction (LCR; Barany 1991, Proc. Natl. Acad. Sci. USA 88:189-193; EP Appl. No., 320,308), Self-Sustained Sequence Replication (3SR; Guatelli et al., 1990, Proc. Natl. Acad. Sci. USA 87:1874-1878), Strand Displacement Amplification (SDA; U.S. Pat. Nos. 5,270,184, en 5,455,166), Transcriptional Amplification System (TAS; Kwoh et al., Proc. Natl. Acad. Sci. USA 86:1173-1177), Q-Beta Replicase (Lizardi et al., 1988, Bio/Technology 6:1197), Rolling Circle Amplification (RCA; U.S. Pat. No. 5,871,921), Nucleic Acid Sequence Based Amplification (NASBA), Cleavage Fragment Length Polymorphism (U.S. Pat. No. 5,719,028), Isothermal and Chimeric Primer-initiated Amplification of Nucleic Acid (ICAN), Ramification-extension Amplification Method (RAM; U.S. Pat. Nos. 5,719,028 and 5,942,391) or other suitable methods for amplification of nucleic acids.

In order to amplify a nucleic acid with a small number of mismatches to one or more of the amplification primers, an amplification reaction may be performed under conditions of reduced stringency (e.g. a PCR amplification using an annealing temperature of 38.degree. C., or the presence of 3.5 mM MgCl2). The person skilled in the art will be able to select conditions of suitable stringency.

The primers herein are selected to be "substantially" complementary (i.e. at least 65%, more preferably at least 80% perfectly complementary) to their target regions present on the different strands of each specific sequence to be amplified. It is possible to use primer sequences containing e.g. inositol residues or ambiguous bases or even primers that contain one or more mismatches when compared to the target sequence. In general, sequences that exhibit at least 65%, more preferably at least 80% homology with the target DNA or RNA oligonucleotide sequences are considered suitable for use in a method of the present invention. Sequence mismatches are also not critical when using low stringency hybridization conditions.

The detection of the amplification products can in principle be accomplished by any suitable method known in the art. The detection fragments may be directly stained or labeled with radioactive labels, antibodies, luminescent dyes, fluorescent dyes, or enzyme reagents. Direct DNA stains include for example intercalating dyes such as acridine orange, ethidium bromide, ethidium monoazide or Hoechst dyes.

Alternatively, the DNA or RNA fragments may be detected by incorporation of labeled dNTP bases into the synthesized fragments. Detection labels which may be associated with nucleotide bases include e.g. fluorescein, cyanine dye or BrdUrd. When using a probe-based detection system, a suitable detection procedure for use in the present invention may for example comprise an enzyme immunoassay (EIA) format (Jacobs et al., 1997, J. Clin. Microbiol. 35, 791-795). For performing a detection by manner of the EIA procedure, either the forward or the reverse primer used in the amplification reaction may comprise a capturing group, such as a biotin group for immobilization of target DNA PCR amplicons on e.g. a streptavidin coated microtiter plate wells for subsequent EIA detection of target DNA-amplicons (see below). The skilled person will understand that other groups for immobilization of target DNA PCR amplicons in an EIA format may be employed.

Probes useful for the detection of the target DNA as disclosed herein preferably bind only to at least a part of the DNA sequence region as amplified by the DNA amplification procedure. Those of skill in the art can prepare suitable probes for detection based on the nucleotide sequence of the target DNA without undue experimentation as set out herein. Also the complementary nucleotide sequences, whether DNA or RNA or chemically synthesized analogs, of the target DNA may suitably be used as type-specific detection probes in a

method of the invention, provided that such a complementary strand is amplified in the amplification reaction employed.

Suitable detection procedures for use herein may for example comprise immobilization of the amplicons and probing the DNA sequences thereof by e.g. southern blotting. Other formats may comprise an EIA format as described above. To facilitate the detection of binding, the specific amplicon detection probes may comprise a label moiety such as a fluorophore, a chromophore, an enzyme or a radio-label, so as to facilitate monitoring of binding of the probes to the reaction product of the amplification reaction. Such labels are well-known to those skilled in the art and include, for example, fluorescein isothiocyanate (FITC), beta-galactosidase, horseradish peroxidase, streptavidin, biotin, digoxigenin, 35S or 125I. Other examples will be apparent to those skilled in the art.

Detection may also be performed by a so called reverse line blot (RLB) assay, such as for instance described by Van den Brule et al. (2002, J. Clin. Microbiol. 40, 779-787). For this purpose RLB probes are preferably synthesized with a 5' amino group for subsequent immobilization on e.g. carboxyl-coated nylon membranes. The advantage of an RLB format is the ease of the system and its speed, thus allowing for high throughput sample processing.

The use of nucleic acid probes for the detection of RNA or DNA fragments is well known in the art. Mostly these procedures comprise the hybridization of the target nucleic acid with the probe followed by post-hybridization washings. Specificity is typically the function of post-hybridization washes, the critical factors being the ionic strength and temperature of the final wash solution. For nucleic acid hybrids, the Tm can be approximated from the equation of Meinkoth and Wahl, Anal. Biochem., 138: 267-284 (1984): Tm=81.5.degree. C.+16.6 (log M)+0.41 (% GC)-0.61 (% form)-500/L; where M is the molarity of monovalent cations, % GC is the percentage of guanosine and cytosine nucleotides in the nucleic acid, % form is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. The Tm is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridizes to a perfectly matched probe. Tm is reduced by about 1.degree. C. for each 1% of mismatching; thus, the hybridization and/or wash conditions can be adjusted to hybridize to sequences of the desired identity. For example, if sequences with >90% identity are sought, the Tm can be decreased 10.degree. C. Generally, stringent conditions are selected to be about 5.degree. C. lower than the thermal melting point (Tm) for the specific sequence and its complement at a defined ionic strength and pH. However, severely stringent conditions can utilize hybridization and/or wash at 1, 2, 3, or 4.degree. C. lower than the thermal melting point (Tm); moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or 10.degree. C. lower than the thermal melting point (Tm); low stringency conditions can utilize a hybridization and/or wash at 11, 12, 13, 14, 15, or 20.degree. C. lower than the thermal melting point (Tm). Using the equation, hybridization and wash compositions, and desired Tm, those of ordinary skill will understand that variations in the stringency of hybridization and/or wash solutions are inherently described. If the desired degree of mismatching results in a Tm of less than 45.degree. C. (aqueous solution) or 32.degree. C. (formamide solution) it is preferred to increase the SSC concentration so that a higher temperature can be used. An extensive guide to the hybridization of nucleic acids is found in Tijssen, Laboratory Techniques in Biochemistm and Molecular Biology--Hybridization with Nucleic Acid Probes, Part I, Chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier. New York (1993); and Current Protocols in Molecular Biology, Chapter 2, Ausubel, et al., Eds., Greene Publishing and Wiley-Interscience, New York (1995).

In another aspect, the invention provides oligonucleotide probes for the generic detection of target RNA or DNA. The detection probes herein are selected to be "substantially" complementary to one of the strands of the double stranded nucleic acids generated by an amplification reaction of the invention. Preferably the probes are substantially complementary to the immobilizable, e.g. biotin labelled, antisense strands of the amplicons generated from the target RNA or DNA.

It is allowable for detection probes of the present invention to contain one or more mismatches to their target sequence. In general, sequences that exhibit at least 65%, more preferably at least 80% homology with the target oligonucleotide sequences are considered suitable for use in a method of the present invention. Antibodies, both monoclonal and polyclonal, can also be used for detection purpose in the present invention, for example, in immunoassays in which they can be utilized in liquid phase or bound to a solid phase carrier. In addition, the monoclonal antibodies in these immunoassays can be detectably labeled in various ways. A variety of immunoassay formats may be used to select antibodies specifically reactive with a particular protein (or other analyte). For example, solid-phase ELISA immunoassays are routinely used to select monoclonal antibodies specifically immunoreactive with a protein. See Harlow and Lane, Antibodies, A Laboratory Manual, Cold Spring Harbor Publications, New York (1988), for a description of immunoassay formats and conditions that can be used to determine selective binding. Examples of types of immunoassays that can utilize antibodies of the invention are competitive and non-competitive immunoassays in either a direct or indirect format. Examples of such immunoassays are the radioimmunoassay (RIA) and the sandwich (immunometric) assay. Detection of the antigens using the antibodies of the invention can be done utilizing immunoassays that are run in either the forward, reverse, or simultaneous modes, including immunohistochemical assays on physiological samples. Those of skill in the art will know, or can readily discern, other immunoassay formats Antibodies can be bound to many different carriers and used to detect the presence of the target molecules. Examples of well-known carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, agaroses and magnetite. The nature of the carrier can be either soluble or insoluble for purposes of the invention. Those skilled in the art will know of other suitable carriers for binding monoclonal antibodies, or will be able to ascertain such using routine experimentation.

The invention also provides a method for serologically diagnosing a MERS-CoV infection of a mammal comprising determining in a sample of said mammal the presence of an antibody specifically directed against a MERS-CoV or component thereof by reacting said sample with a proteinaceous molecule or fragment thereof or an antigen according to the invention

Methods and means provided herein are particularly useful in a diagnostic kit for diagnosing a MERS-CoV infection, be it by virological or serological diagnosis. Such kits or assays may for example comprise a virus, a nucleic acid, a proteinaceous molecule or fragment thereof, an antigen and/or an antibody according to the invention.

Herewith, the invention provides a method for virologically diagnosing a Betacoronavirus, Lineage C infection of a mammal comprising determining in a sample of said mammal the presence of a viral isolate or component thereof by reacting said sample with a nucleic acid according to the invention or an antibody according to the invention or determining in a sample of said mammal the presence of an antibody specifically directed against a Betacoronavirus, Lineage C virus or component thereof by reacting said sample with a proteinaceous molecule or fragment thereof according to the invention or an antigen according to the invention.

The invention also provides diagnostic kit for diagnosing a Betacoronavirus, Lineage C infection comprising a virus according to the invention, a nucleic acid according to the invention, a proteinaceous molecule

or fragment thereof according to the invention, an antigen according to the invention and/or an antibody according to the invention.

The invention also provides use of a MERS-CoV according to the invention, a nucleic acid according to the invention, a vector according to the invention, a host cell according to the invention, a proteinaceous molecule or fragment thereof according to the invention, an antigen according to the invention, or an antibody according to the invention for the production of a pharmaceutical composition, preferably for the production of a pharmaceutical composition for the treatment or prevention of a Betacoronavirus, Lineage C virus infection, preferably a human infection, or for the production of a pharmaceutical composition for the treatment or prevention of atypical pneumonia and/or renal failure, preferably wherein said atypical pneumonia and/or renal failure is a human disease.

The invention also provides pharmaceutical composition comprising a virus according to the invention, a nucleic acid according to the invention, a vector according to the invention, a host cell according to the invention, a proteinaceous molecule or fragment thereof according to the invention, an antigen according to the invention, or an antibody according to the invention.

A pharmaceutical composition comprising a virus, a nucleic acid, a proteinaceous molecule or fragment thereof, an antigen and/or an antibody according to the invention can for example be used in a method for the treatment or prevention of a MERS-CoV infection and/or a respiratory illness comprising providing an individual with a pharmaceutical composition according to the invention. This is most useful when said individual comprises a human. Antibodies against MERS-CoV proteins, especially against the spike protein of MERS-CoV, preferably against the amino acid sequence as depicted herein are also useful for prophylactic or therapeutic purposes, as passive vaccines. It is known from other coronaviruses that the spike protein is a very strong antigen and that antibodies against spike protein can be used in prophylactic and therapeutic vaccination.

The invention also provides method to obtain a modulator or an antiviral agent useful in the treatment of atypical pneumonia comprising establishing a cell culture or experimental animal comprising a virus according to the invention, treating said culture or animal with a candidate antiviral agent, and determining the effect of said modulator or agent on said virus or its infection of said culture or animal. An example of such an antiviral agent comprises a MERS-CoV virus-neutralizing antibody, or functional component thereof, as provided herein, but antiviral agents of other nature, such as ADA or adenosine are obtained as well. The invention also provides use of a modulator or an antiviral agent according to the invention for the preparation of a pharmaceutical composition, in particular for the preparation of a pharmaceutical composition for the treatment of atypical pneumonia, specifically when caused by a MERS-CoV infection, and provides a pharmaceutical composition comprising an antiviral agent according to the invention, useful in a method for the treatment or prevention of a MERS-CoV infection or atypical pneumonia, said method comprising providing an individual with such a pharmaceutical composition.

The invention also provides a method for the treatment or prevention of a Betacoronavirus, Lineage C virus infection or for the treatment or prevention of atypical pneumonia comprising providing an individual, preferably a human individual with a pharmaceutical composition according to the invention. In particular individual MERS-CoV virus-like polypeptide are provided herein as well, such as the viral replicase encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof, as depicted in figure 13. A viral spike protein encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof, as depicted in figure 12, a viral non-structural gene protein encoded by an RNA or DNA or cDNA sequence or fragments or

homologues thereof, as depicted in any of figures 8, 9, 10 or 11, a small envelope (E) protein encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof, as depicted in figure 7, a matrix (M) protein encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof, as depicted in figure 6, a nucleocapsid (N) protein encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof, as depicted in figure 5, a nucleic acid sequence which comprises one or more of the sequences of HCoV-SA1 as depicted in figures 3, or 5 to 15 or a nucleic acid sequence which can hybridize with any of these sequences under stringent conditions.

With the provision of the sequence information of this MERS virus, MERS-CoV, the invention provides diagnostic means and methods, prophylactic means and methods and therapeutic means and methods to be employed in the diagnosis, prevention and/or treatment of disease, in particular of respiratory disease (atypical pneumonia), in particular of mammals, more in particular in humans. In virology, it is most advisory that diagnosis, prophylaxis and/or treatment of a specific viral infection is performed with reagents that are most specific for said specific virus causing said infection. In this case this means that it is preferred that said diagnosis, prophylaxis and/or treatment of a MERS virus infection is performed with reagents that are most specific for MERS virus. This by no means however excludes the possibilities that less specific, but sufficiently cross-reactive reagents are used instead, for example because they are more easily available and sufficiently address the task at hand.

The invention for example provides a method for virologically diagnosing a MERS infection of an animal, in particular of a mammal, more in particular of a human being, comprising determining in a sample of said animal the presence of a viral isolate or component thereof by reacting said sample with a MERS specific nucleic acid or antibody according to the invention, and a method for serologically diagnosing a MERS infection of a mammal comprising determining in a sample of said mammal the presence of an antibody specifically directed against a MERS virus or component thereof by reacting said sample with a MERS-CoV-specific proteinaceous molecule or fragment thereof or an antigen according to the invention. Suitable MERS-CoV specific nucleic acid can for example is provided herein as well, such as the RNA or DNA or cDNA sequence or fragments or homologues thereof, as depicted in figure 13, or as depicted in figure 12, or in any of figures 8, 9, 10 or 11, in figure 7, in figure 6, in figure 5, a nucleic acid sequence which comprises one or more of the sequences of HCoV-SA1 or a MERS-CoV specific nucleic acid sequence which can hybridize with sequences in any of figures as depicted in figures 3, or 5 to 15 under stringent conditions, or a MERS-CoV specific nucleic acid sequence, such as an RNA or a DNA or preferably a cDNA, which has at least 65%, preferably at least 75%, more preferably at least 85%, most preferably at least 95% homology or are substantially, at least 65%, preferably at least 75%, more preferably at least 85%, most preferably at least 95%, complementary with a nucleotide sequence as depicted in figures 3, or 5 to 15. For MERS CoV nucleic acid diagnosis, short nucleotide stretches of 10 to 40, preferably 12 to 30, more preferably 15 to 25 nucleotides long, commonly called “primers” are provided herein that preferably are MERS-CoV specific or at least substantially complementary to MERS virus nucleic acid as depicted in figures 3, or 5-15 and have stretches of at least 10, preferably at least 12, more preferably at least 15, most preferably at least 18 or 19 nucleotides that are 100% complementary to at least a fragment of a nucleotide sequence as depicted in figures 3, or 5 to 15. The term "nucleotide sequence homology" as used herein denotes the presence of homology between two (poly) nucleotides, such as a RNA or a DNA or a cDNA sequence. Polynucleotides have "homologous" sequences if the sequence of nucleotides in the two sequences is the same when aligned for maximum correspondence. Sequence comparison between two or more polynucleotides is generally performed by comparing portions of the two sequences over a comparison window to identify and compare local regions of sequence similarity. The comparison window is generally from about 20 to 200 contiguous nucleotides. The "percentage of sequence homology" for polynucleotides, such as 50, 60, 70, 80, 90, 95, 98, 99 or 100 percent sequence homology may be determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may include additions or deletions (i.e. gaps) as

compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Nucleotide or base G is homologous to G, C is homologous to C, A is homologous to A and nucleotides T or U are homologous to T or U, to calculate overall homology or complementarities between DNA and RNA. The percentage is calculated by: (a) determining the number of positions at which the identical nucleic acid base occurs in both sequences to yield the number of matched positions; (b) dividing the number of matched positions by the total number of positions in the window of comparison; and (c) multiplying the result by 100 to yield the percentage of sequence homology. Optimal alignment of sequences for comparison may be conducted by computerized implementations of known algorithms, or by inspection. Readily available sequence comparison and multiple sequence alignment algorithms are, respectively, the Basic Local Alignment Search Tool (BLAST) (Altschul, S. F. et al. 1990. J. Mol. Biol. 215:403; Altschul, S. F. et al. 1997. Nucleic Acid Res. 25:3389-3402) and ClustalW programs both available on the internet. Other suitable programs include GAP, BESTFIT and FASTA in the Wisconsin Genetics Software Package (Genetics Computer Group (GCG), Madison, Wis., USA).

As used herein, "substantially complementary" means that two nucleic acid sequences have at least about 65%, preferably about 70%, more preferably about 80%, even more preferably 90%, and most preferably about 98%, sequence complementarities to each other. This means that the primers and probes must exhibit sufficient complementarity to their template and target nucleic acid, respectively, to hybridize under stringent conditions. Therefore, the primer sequences as disclosed in this specification need not reflect the exact sequence of the binding region on the template and degenerate primers can be used. A substantially complementary primer sequence is one that has sufficient sequence complementarity to the amplification template to result in primer binding and second-strand synthesis.

The term "hybrid" refers to a double-stranded nucleic acid molecule, or duplex, formed by hydrogen bonding between complementary nucleotides. The terms "hybridize" or "anneal" refer to the process by which single strands of nucleic acid sequences form double-helical segments through hydrogen bonding between complementary nucleotides, according to a strict rule called base-pairing defined by the complementary structures of the nucleotides or bases (b). Typically, in two nucleic acid strands, nucleotide guanine (G) is complementary to nucleotide cytosine (C), G and C pair wise capable of forming three hydrogen bonds, and nucleotide adenine (A) is complementary to nucleotides thymine (T) or uracil (U), A and T or A and U pair wise capable of forming two hydrogen bonds, thus G pairs with C and A pairs with T or U. Conventionally, in depicting a nucleic acid sequence, T is commonly identified as uracil (U) to identify RNA (ribonucleic acid), and as thymine (T) when identifying DNA (deoxyribonucleic acid) or cDNA (complementary or copy DNA). A DNA polymerase is a cellular or viral polymerase enzyme that synthesizes DNA molecules from their nucleotide building blocks. DNA polymerases are essential for DNA replication, and usually function in pairs while copying one double-stranded DNA molecule into two double-stranded DNAs in a process termed DNA replication. RNA viruses commonly use an RNA-dependent RNA-polymerase to replicate their RNA. DNA can be used to produce RNA by the actions of a transcriptase; RNA can be used to produce DNA or cDNA by the actions of a reverse transcriptase. A transcriptase is a polymerase that catalyzes the formation of RNA from a DNA template in the process of transcription. Reverse transcriptase (RT) is a polymerase enzyme used to generate complementary DNA (cDNA) from an RNA template, a process termed reverse transcription.

The term "oligonucleotide" refers to a short sequence of nucleotide monomers (usually 6 to 100 nucleotides) joined by phosphorous linkages (e.g., phosphodiester, alkyl and aryl-phosphate, phosphorothioate), or non-phosphorous linkages (e.g., peptide, sulfamate and others). An oligonucleotide may contain modified nucleotides having modified bases (e.g., 5-methyl cytosine) and modified sugar groups (e.g., 2'-O-methyl ribosyl 2'-O-methoxyethyl ribosyl, 2'-fluoro ribosyl, 2'-amino ribosyl, and the like). Oligonucleotides may be naturally-

occurring or synthetic molecules of double- and single-stranded DNA and double- and single-stranded RNA with circular, branched or linear shapes and optionally including domains capable of forming stable secondary structures (e.g., stem-and-loop and loop-stem-loop structures).

The term "primer" as used herein also refers to an oligonucleotide which is capable of annealing to the amplification target allowing a DNA polymerase to attach thereby serving as a point of initiation of DNA synthesis when placed under conditions in which synthesis of primer extension product which is complementary to a nucleic acid strand is induced, i.e., in the presence of nucleotides and an agent for polymerization such as DNA polymerase and at a suitable temperature and pH. The (amplification) primer is preferably single stranded for maximum efficiency in amplification. Preferably, the primer is an oligodeoxyribonucleotide. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the agent for polymerization. The exact lengths of the primers will depend on many factors, including temperature and source of primer. A "pair of bi-directional primers" as used herein refers to one forward and one reverse primer as commonly used in the art of RNA or DNA amplification such as in PCR amplification.

The term "probe" refers to a single-stranded oligonucleotide sequence that will recognize and form a hydrogen-bonded duplex with a complementary sequence in a target nucleic acid sequence analyte or its cDNA derivative.

The terms "stringency" or "stringent hybridization conditions" refer to hybridization conditions that affect the stability of hybrids, e.g., temperature, salt concentration, pH, formamide concentration and the like. These conditions are empirically optimized to maximize specific binding and minimize non-specific binding of primer or probe to its target nucleic acid sequence. The terms as used include reference to conditions under which a probe or primer will hybridize to its target sequence, to a detectably greater degree than other sequences (e.g. at least 2-fold over background). Stringent conditions are sequence dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. Generally, stringent conditions are selected to be about 5. degree. C. lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridizes to a perfectly matched probe or primer. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M Na+ ion, typically about 0.01 to 1.0 M Na+ ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30.degree. C. for short probes or primers (e.g. 10 to 50 nucleotides) and at least about 60.degree. C. for long probes or primers (e.g. greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Exemplary low stringent conditions or "conditions of reduced stringency" include hybridization with a buffer solution of 30% formamide, 1 M NaCl, 1% SDS at 37.degree. C. and a wash in 2.times.SSC at 40.degree. C. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37.degree. C., and a wash in 0.1.times.SSC at 60.degree. C. Hybridization procedures are well known in the art and are described in e.g. Ausubel et al, Current Protocols in Molecular Biology, John Wiley & Sons Inc., 1994.

The invention for example provides a method for virologically diagnosing a MERS infection of an animal, in particular of a mammal, more in particular of a human being, comprising determining in a sample of said animal the presence of a viral isolate or component thereof by reacting said sample with a MERS specific nucleic acid or antibody according to the invention, and a method for serologically diagnosing a MERS infection of a mammal comprising determining in a sample of said mammal the presence of an antibody specifically directed against a MERS virus or component thereof by reacting said sample with a MERS MERS-CoV-specific proteinaceous molecule or fragment thereof or an antigen according to the invention. Suitable MERS specific

proteinaceous molecules or MERS virus specific fragment thereof is provided herein as well, such as the viral replicase encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof, as depicted in figure 13, a viral spike protein encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof, as depicted in figure 12, a viral non-structural gene protein encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof, as depicted in any of figures 8, 9, 10 or 11, a small envelope (E) protein encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof, as depicted in figure 7, a matrix (M) protein encoded by an RNA pr DNA sequence or fragments or homologues thereof, as depicted in figure 6, a nucleocapsid (N) protein encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof, as depicted in figure 5, a nucleic acid sequence which comprises one or more of the sequences of HCoV-SA1 as depicted in figures 3, or 5 to 15 or a nucleic acid sequence which can hybridize with any of these sequences under stringent conditions.

Suitable MERS CoV specific antibodies directed against MERS CoV specific proteinaceous molecules or MERS CoV specific fragment thereof is provided herein as well, such as antibodies raised against a viral replicase encoded by an RNA sequence or fragments or homologues thereof, as depicted in figure 13, raised against a viral spike protein encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof, as depicted in figure 12, raised against a viral non-structural gene protein encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof, as depicted in any of figures 8, 9, 10 or 11, raised against a small envelope (E) protein encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof, as depicted in figure 7, raised against a matrix (M) protein encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof, as depicted in figure 6, raised against a nucleocapsid (N) protein encoded by an RNA or DNA or cDNA sequence or fragments or homologues thereof, as depicted in figure 5, a nucleic acid sequence which comprises one or more of the sequences of HCoV-SA1 as depicted in figures 3, or 5 to 15 or a nucleic acid sequence which can hybridize with any of these sequences under stringent conditions.

The term "antibody" includes reference to antigen binding forms of antibodies (e. g., Fab, F(ab)2). The term "antibody" frequently refers to a polypeptide substantially encoded by an immunoglobulin gene or immunoglobulin genes, or fragments thereof which specifically bind and recognize an analyte (antigen). However, while various antibody fragments can be defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized de novo either chemically or by utilizing recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments such as single chain Fv, chimeric antibodies (i. e., comprising constant and variable regions from different species), humanized antibodies (i. e., comprising a complementarity determining region (CDR) from a non-human source) and heteroconjugate antibodies (e. g., bispecific antibodies).

The invention also provides a diagnostic kit for diagnosing a MERS-CoV infection comprising a MERS Corona virus, or a MERS-CoV-specific nucleic acid, or a MERS-CoV-specific proteinaceous molecule or fragment thereof, a MERS-CoV-specific antigen and/or an a MERS-CoV-specific antibody according to the invention, and preferably a means for detecting said MERS-CoV, MERS-CoV-specific nucleic acid, said proteinaceous molecule or fragment thereof, said antigen and/or said antibody, said means for example comprising an excitable group such as a fluorophore or enzymatic detection system used in the art (examples of suitable diagnostic kit format comprise IF, ELISA, neutralization assay, RT-PCR assay). To determine whether an as yet unidentified virus component or synthetic analogue thereof such as nucleic acid, proteinaceous molecule or fragment thereof can be identified as MERS-CoV-specific, it suffices to analyze the nucleic acid or amino acid sequence of said component, for example for a stretch of said nucleic acid or amino acid, preferably of at least 10, more preferably at least 25, more preferably at least 40 nucleotides or amino acids (respectively), by sequence homology comparison with the herein provided MERS viral sequences and with known non-MERS viral sequences

(HUK4 or HUK5 are preferably used) using for example phylogenetic analyses as provided herein. Depending on the degree of relationship with said MERS or non-MERS viral sequences, the component or synthetic analogue can be identified.

The sequence of the first isolate of MERS-CoV is also deposited in Genbank under:

LOCUS JX869059 30119 bp RNA linear VRL 04-DEC-2012
DEFINITION Human betacoronavirus 2c EMC/2012, complete genome.
ACCESSION JX869059
VERSION JX869059.2 GI:409052551
KEYWORDS
SOURCE Human betacoronavirus 2c EMC/2012
ORGANISM Human betacoronavirus 2c EMC/2012
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirinae; Betacoronavirus; unclassified
Betacoronavirus.
REFERENCE 1 (bases 1 to 30119)
AUTHORS van Boheemen,S., de Graaf,M., Lauber,C., Bestebroer,T.M., Raj,V.S., Zaki,A.M., Osterhaus,A.D., Haagmans,B.L., Gorbalenya,A.E., Snijder,E.J. and Fouchier,R.A.
TITLE Genomic characterization of a newly discovered coronavirus
associated with acute respiratory distress syndrome in humans
JOURNAL MBio 3 (6), e00473-12 (2012)
PUBMED 23170002

The present invention in particular also relates to the spike (S) protein of a Coronavirus that utilizes DPP4 as a virus receptor and fragments thereof as for example depicted in figures 16, 17 and 32.

The invention in a further embodiment also provides a proteinaceous substance comprising, preferably having been provided, with at least a fragment of a viral protein, preferably an isolated fragment, herein a fragment obtained by recombinant means is preferably used as the probe protein, but a synthetic peptide obtained by chemical peptide synthesis can also be used, also called a first fragment and having also been provided with at least a fragment of an N terminal dipeptidyl peptidase, herein surprisingly found to be the target protein specifically reacting with the used viral probe and herein also called a second fragment.

In describing a proteinaceous substance herein, reference is made to protein containing material, such as an organism or a part thereof, microbial organism, virus, tissue, cell, cell culture, cell culture precipitate, cell culture supernatant, cell content such as cytoplasm, nucleoplasm, nuclei, nucleoli, cell organelles, mitochondria, ribosome, tubuli, plasma, blood, serum, lymph, drainage fluid, and to a protein containing preparation, such as a buffer, dilution, precipitate, extraction, pull down sample, test sample, spray, chromatographic sample, or a crystal.

Surprisingly, in pull down binding experiments with a fragment of a newly discovered coronavirus, it was found that the first isolated fragment, comprising an ectodomain of the spike protein of the virus, bound to at least the ectodomain of a prolyl peptidase, an N terminal dipeptidyl peptidase, the identity of which was confirmed by mass spectrometric analyses of tryptic peptide digests. No binding interaction between an N terminal dipeptidyl peptidase and a viral protein has been found before, in particular, not wherein the peptidase is acting as a receptor for the virus, allowing viral entry and replication in a cell. Blocking DPP4 with specific anti DPP4 antibodies indeed abolished viral infection.

The invention also relates to ten protease families that are unique to higher organisms (16 protease families can be identified in the genomes of all forms of cellular life). Within this core group of ten protease

families, a multitude of proteases evolved to yield intra and extra cellular processes. Dipeptidyl peptidase 4 (DPP4; Dipeptidyl Peptidase IV (DPPIV)) is a member of this large family of proteases (peptidases). DPP4 is a serine protease of family S9. DPP4 is a 240 kDa homodimeric, multi functional type II membrane bound glycoprotein, widely distributed in all mammalian tissues, but highly expressed in kidney, liver and endothelium. DPPIV is also known as DPP4, CD26, adenosine deaminase complexing protein 2 or adenosine deaminase binding protein (ADAbp). DPP4 consists of a short cytoplasmic domain of six amino acids, followed by a hydrophobic transmembrane domain (amino acids 7 28) and an extracellular (ectodomain) sequence of 739 amino acids. DPP4 is a highly specific aminopeptidase and releases dipeptides from the amino terminus of peptides with a Pro or Ala in the penultimate position. N terminal degradation of the substrate peptides may result in the activation, inactivation or modulation of their activity. Besides its well known exopeptidase activity, DPPIV also exhibits endopeptidase activity toward denatured collagen. Expression of DPPIV is associated with cell adhesion and is a co stimulant during T cell activation and proliferation.

DPPIV (DPP4, CD26) is a member of the class of proteases known as prolyl peptidases, which cleave proteins after proline residues. DPPIV, a serine dipeptidyl peptidase, cleaves the N terminal X Ala or X Pro from target polypeptides, such as chemokines (e.g., CXCL11) and peptide hormones (e.g., GLP 1, PACAP, VIP, BNP). DPPIV possesses a transmembrane region and a very short cytoplasmic domain, but is often cleaved and released as a soluble, circulating fragment. Serine proteases are grouped into 43 families. Protease family S9 is divided into four subfamilies: S9A (type prolyl oligopeptidase), S9B (DPP4), S9C (acylaminoacyl peptidase), and S9D (glutamyl endopeptidase).

In humans, members of the subfamily S9B include DPP4, fibroblast activation protein alpha (FAPα), dipeptidyl peptidase 8 (DPP8), and dipeptidyl peptidase 9 (DPP9). DPP4 is also known as adenosine deaminase binding protein (ADBP) or T cell activation antigen CD26. DPP4 is a serine exopeptidase that catalyzes the release of an N terminal dipeptide provided that the next to last residue is proline, hydroxyproline, dehydroproline or alanine.

Only oligopeptides in the trans conformation are able to bind to the active site of DPP4. It also has non peptidase functions: through its interaction with adenosine deaminase (ADA) and extracellular matrix components, it influences T cell activation and proliferation. It is thought to play roles in diabetes, cancer, and autoimmune diseases, making it a target for drug discovery. In particular, cleavage of GLP 1 (7 36) amide, an incretin hormone that stimulates insulin biosynthesis and secretion, into GLP 1 (9 36) amide by DPPIV reverses the glucoregulatory actions of GLP 1. Therefore, DPPIV inhibitors are attractive targets for stimulating insulin production in type II diabetes. Several specific DPPIV inhibitors have been approved by the FDA for type II diabetes.

The invention also provides a proteinaceous substance comprising, preferably having been provided with, an isolated first fragment of a viral protein and an isolated second fragment of an N terminal dipeptidyl peptidase protein.

Repeating binding experiments with a recombinant, isolated, fragment of the peptidase indeed confirmed the identification of the peptidase as a receptor of the MERS CoV and of the HKU4 CoV, allowing binding of the virus to a mammalian cell (both bat DPP4 as well as human DPP4 were tested), and entry of the MERS CoV leading to abundant replication of that virus in COS7 cells having been provided with the isolated second fragment, whereas COS7 cells not having been provided with the second fragment remain essentially impervious for infection with the virus.

In a further embodiment, the invention thus provides a proteinaceous substance having been provided with the isolated first probing fragment, preferably a recombinant fragment, of a viral protein and an isolated second fragment, preferably recombinant fragment, of an N terminal dipeptidyl peptidase protein,

establishing a probe identified target pair of binding proteins that may be used for binding or affinity studies and preferably also for methods to identify modulators of the interaction of the binding pair.

In a further embodiment, the invention provides a proteinaceous substance comprising, preferably having been provided with, a first fragment of a viral protein and an isolated second, preferably recombinant, fragment of an N terminal dipeptidyl peptidase protein (a fragment obtained by regular peptide synthesis may also be use as first or second fragment). Such a substance provided, in particular, is useful in identifying further binding sites of viral proteins, and fragments thereof, e.g., for narrowing down of specific binding site sequences.

In a preferred embodiment, the invention provides a proteinaceous substance having been provided with at least a fragment of a viral protein, preferably an isolated fragment, wherein the viral protein comprises an ectodomain of a spike protein or of an envelope protein, the ectodomain being the most preferred site for virus cell receptor interaction.

It is preferred that a substance according to the invention comprises coronaviral protein, preferably wherein the coronaviral protein is derived from the HCoV EMC 1 virus, preferably wherein the first fragment is derived from the S1 region of a coronavirus. In a particular embodiment, the first fragment comprises residues 1 747 of the viral spike protein of HCoV EMC 1 as depicted in figure 16.

The invention also provides a substance according to the invention wherein the first fragment comprises, preferably consists of, at least 10, preferably of at least 50, preferably of at least 100 residues derived from the S1 region of a coronavirus. Using smaller fragments from distinct locations in the viral sequence allows for further identifying minimal essential sequences, and thereby narrowing down on the binding site, necessary for binding with the peptidase.

In particular, a substance according to the invention is provided wherein the first fragment is derived from the S1 region of a coronavirus, for example, comprising residues 1 747 as depicted in 161. Examples of such selected fragments are also found in figure 17, preferably the invention provides a substance with a first fragment consisting of residues 1 357, or of residues 1 588, or of residues 358 588, or of residues 358 747, or of residues 588 747 as depicted in figure 16 or figure 17, or of residues 363 593 of the spike protein of HKU4 CoV as shown in figure 17.

Even more in particular, a substance according to the invention is provided wherein the first fragment is derived from the S1 region of a coronavirus, which fragment then is subjected to limited proteolysis after which the protease resistant domains are identified by MS, and the interaction between probe and target is studied further.

The invention also provides a substance according to the invention wherein the peptidase belongs to protease family S9, preferably subfamily S9B, preferably wherein the peptidase is a dipeptidyl peptidase 4 (DPP4), preferably human DPP4, and preferably wherein the fragment is derived from the ectodomain of dipeptidyl peptidase. In one embodiment, it is provided that the second fragment comprises residues 39 766 of human DPP4 as depicted in figure 18.

The invention also provides a substance according to the invention wherein the second fragment comprises, preferably consists of, at least 10, preferably of at least 50, preferably of at least 100 residues derived from the ectodomain of dipeptidyl peptidase, such as wherein the second fragment is derived from the ectodomain of human DDP4 comprising residues 39 766 as depicted in figure 18.

Examples of such selected fragments are also found in figure 3, preferably the invention provides a substance with a second fragment consisting of residues 1 6, or of residues 1 28, or of residues 29 38, or of residues 39 51, or of residues 506 766 as depicted in figure 18.

Even more in particular, a substance according to the invention is provided wherein the second fragment is derived from the ectodomain of a peptidase, which fragment is then subjected to limited proteolysis

after which the protease resistant domains are identified by MS, and the interaction between probe and target is studied further.

The invention also provides a substance according to the invention wherein at least one of the isolated fragments has been provided with an affinity tag, preferably a tag having affinity to binding with Protein A or a tag having affinity for binding with streptavidin.

The invention also provides a substance according to the invention consisting essentially of an isolated first fragment of a viral protein and an isolated second fragment of an N terminal peptidase protein. In a preferred embodiment, the viral protein is a coronaviral protein, preferably derived from a virus capable of infecting a human cell, whereas the peptidase protein is a DPP4 protein, preferably a human DPP4 protein.

Furthermore, a substance according to the invention is herein provided that has been subjected to crystallization, preferably a substance comprising a crystal consisting essentially of an isolated first fragment of a viral protein and an isolated second fragment of an N terminal peptidase protein. In a preferred embodiment, the viral protein is a coronaviral protein, preferably derived from a virus capable of infecting a human cell, whereas the peptidase protein is a DPP4 protein, preferably a human DPP4 protein.

The invention also provides a method for identifying a binding site comprising subjecting a crystal consisting essentially of an isolated first fragment of a viral protein and an isolated second fragment of an N terminal peptidase protein to X ray or neutron diffraction analysis. This is, for example, in order to determine the three dimensional structure of fragments of DPPIV and coronaviral protein and, in particular, to assist in the identification of its active site where fragments may bind. Knowledge of the binding site region allows rational design and construction of ligands including inhibitors. Crystallization and structural determination of fragments of DPPIV mutants and/or viral protein mutants having altered bioactivity allows the evaluation of whether such changes are caused by general structure deformation or by side chain alterations at the substitution site.

The invention also provides a container with a substance according to the invention, such as container provided with a virus according to the invention, and/or a nucleic acid according to the invention, and/or a vector according to the invention, and/or a host cell according to the invention, and/or a proteinaceous molecule according the invention, and/or an antigen according to the invention, and/or or an antibody according to the invention and/or a pharmaceutical composition according to the invention. In describing a container herein, reference is made to a test device, test tube (commonly Eppendorf tubes are used), test vessel, pipette, pipette tip, reaction device, cell culture vessel, cell culture well, reaction chamber, cover slip, crystallization chamber, crystallization device, crystallization well, microplate well, crystallization plate well, gel, column wherein, on or under a proteinaceous substance according to the invention may be placed or contained or that are useful for storing, shipping, testing or handling a proteinaceous substance provided herein.

The invention also provides a method of identifying a candidate modulator as an agent that modulates the function of a dipeptidyl peptidase, the method comprising providing a substance with a first and a second fragment according to the invention in the presence and absence of the candidate modulator under conditions permitting binding of the first fragment with the second fragment. Measuring binding of the first fragment to the second fragment, wherein a decrease or increase in binding in the presence of the candidate modulator, relative to binding in the absence of the candidate modulator, identifies the candidate modulator as an agent that modulates the function of a dipeptidyl peptidase.

The invention further provides a method of detecting, in a sample, the presence of an agent that modulates the function of a dipeptidyl peptidase, said method comprising providing a substance with a first and a second fragment according to the invention in the presence and absence of said sample under conditions permitting binding of said first fragment with said second fragment. Measuring binding of said first fragment to said second fragment, wherein a decrease or increase in binding in the presence of said sample, relative to binding

in the absence of said sample, identifies said sample as comprising an agent that modulates the function of a dipeptidyl peptidase.

The invention further provides a method of identifying a candidate modulator as an agent that modulates the function of a dipeptidyl peptidase, said method comprising providing a substance with a first and a second fragment according to the invention in the presence and absence of said candidate modulator under conditions permitting determining enzymatic activity of a peptidase. Measuring enzymatic activity of a peptidase, wherein a decrease or increase in enzymatic in the presence of said candidate modulator, relative to binding in the absence of said candidate modulator, identifies said candidate modulator as an agent that modulates the function of a dipeptidyl peptidase.

The invention further provides a method of detecting, in a sample, the presence of an agent that modulates the function of a dipeptidyl peptidase, said method comprising providing a substance with a first and a second fragment according to the invention in the presence and absence of said sample under conditions permitting determining enzymatic activity of a peptidase. Measuring enzymatic activity of a peptidase, wherein a decrease or increase in enzymatic in the presence of said sample, relative to binding in the absence of said sample, identifies said sample as comprising an agent that modulates the function of a dipeptidyl peptidase.

In a preferred embodiment, the invention further provides a method of identifying a candidate modulator as an that modulates the function of a dipeptidyl peptidase or a provides a method of detecting, in a sample, the presence of an agent that modulates the function of a dipeptidyl peptidase wherein said first fragment and/or said second fragment is detectably labeled, preferably wherein said first fragment and/or said second fragment is detectably labeled with a moiety selected from the group consisting of a radioisotope, a fluorophore, a quencher of fluorescence, an enzyme, and an affinity tag. It is also provided that said substance comprises a cell expressing said first fragment and/or said second fragment.

The invention also provides use of a substance, a container or a method according to the invention for identifying an agent that modulates the function of a peptidase or a viral protein, use of an isolated fragment of a viral protein as an agent that modulates the function of an N-terminal dipeptidyl peptidase, and use of an isolated fragment of a N-terminal dipeptidyl peptidase as an agent that modulates the function of a viral protein.

The invention further provides use of an inhibitor, preferably adenosine, or a functional equivalent thereof, of N-terminal dipeptidyl peptidase cell-surface expression on a cell, as a modulator or antiviral agent for inhibition of replication of a virus in said cell, in particular wherein said peptidase is DPP4, preferably human DPP4, preferably wherein said virus is a Coronavirus.

Also, the invention provides vaccines against HCoV SA1 (based on nucleic acid or amino acid sequences or antigenic polypeptides of the HCoV SA1 genome, and the invention provides use of antiviral drugs directed against nucleic acid or amino acid sequences or polypeptides of the HCoV SA1 (herein also called MERS HCoV) genome, as herein provided. At this time, it is not known if there is a cure for the disease. Several antiviral therapies have been applied, but with various results. Also, for being able to prevent spread of the disease, it is of great importance to be able to recognize the disease in an early stage. Only then, sufficient measures can be taken to isolate patients and initiate quarantine precautions.

The invention also provides an isolated or recombinant proteinaceous molecule or MERS-CoV-specific fragment thereof encoded by a nucleic acid according to the invention. In a preferred embodiment, the invention provides a proteinaceous molecule or corona MERS-CoV-specific viral protein or fragment thereof encoded by a nucleic acid according to the invention. Useful proteinaceous molecules are, for example, derived from any of the genes or genomic fragments derivable from a virus according to the invention. Such molecules or antigenic fragments thereof, as provided herein, are, for example, useful in diagnostic methods or kits and in pharmaceutical compositions such as sub unit vaccines and inhibitory peptides. Particularly useful is the viral

polymerase protein, the spike protein, the nucleocapsid or antigenic fragments thereof for inclusion as antigen or subunit immunogen, but inactivated whole virus can also be used.

Particularly useful are also those proteinaceous substances that are encoded by recombinant nucleic acid fragments that are identified for phylogenetic analyses, of course, preferred are those that are within the preferred bounds and metes of ORFs useful in phylogenetic analyses, in particular, for eliciting MERS-CoV-specific antibodies, whether in vivo (e.g., for protective purposes or for providing diagnostic antibodies) or in vitro (e.g., by phage display technology or another technique useful for generating synthetic antibodies). Similarly, the invention provides an antigen comprising a proteinaceous molecule or MERS-CoV-specific fragment thereof according to the invention, or reactive with an antibody according to the invention.

Also provided herein are antibodies, be it natural polyclonal or monoclonal, or synthetic (e.g., (phage) library derived binding molecules) antibodies that specifically react with an antigen comprising a proteinaceous molecule or HCoV virus like MERS-CoV-specific fragment thereof according to the invention. A person skilled in the art will be able to develop (monoclonal) antibodies using isolated virus material and/or recombinantly expressed viral proteins. Sui et al. (Proc. Natl. Acad. Sci. 101(8):2536 2541, 2004) have transiently expressed fragments of the spike protein and found several antibodies through phage display methods. Such antibodies are also useful in a method for identifying a viral isolate as a MERS HCoV virus like virus comprising reacting the viral isolate or a component thereof with an antibody as provided herein. This can, for example, be achieved by using purified or non purified HCoV SA1 virus like virus or parts thereof (proteins, peptides) using ELISA, RIA, FACS or similar formats of antigen detection assays (Current Protocols in Immunology). Alternatively, infected cells or cell cultures may be used to identify viral antigens using classical immunofluorescence or immunohistochemical techniques. Specifically useful in this respect are antibodies raised against MERS HCoV virus like virus proteins that are encoded by a nucleotide sequence comprising one or more of the fragments disclosed herein.

In particular, MERS HCoV virus like polypeptide or fragments are provided herein as well, such as those provided in figure 16 or figure 17, in particular, fragments derived from a viral spike protein, preferably the S1 spike protein, in particular, fragments of the S1 protein, such as fragment 1 357, or fragment 358 747, or fragment 358–588, or homologues thereof, as depicted in figure 17, or fragment 363 593 of the spike protein of HKU4 Co, as shown in figure 32, are herein provided. Also, isolated or recombinant nucleic acid, or MERS-CoV-specific fragments thereof that are obtainable from a MERS HCoV virus are provided, such as nucleic acid encoding fragments of the S1 protein, such as fragment 1 357, or fragment 358 747, or fragment 358–588, or homologues thereof, as depicted in figure 17, as are a vector or plasmid comprising a nucleic acid according to the invention, and a cell, such as host cell, such as a 293T cell comprising a nucleic acid or vector (vector comprising plasmid herein) according to the invention.

The invention also provides an isolated or recombinant proteinaceous molecule or MERS-CoV-specific fragment thereof encoded by a nucleic acid according to the invention. In a preferred embodiment, the invention provides a proteinaceous molecule or MERS-CoV-specific viral protein or fragment thereof encoded by a nucleic acid according to the invention for use in a vaccine. Useful proteinaceous molecules are, for example, derived from any of the genes or genomic fragments derivable from a virus or fragment thereof according to the invention. Such molecules, or antigenic fragments thereof, as provided herein, are, for example, useful in diagnostic methods or kits and in pharmaceutical compositions such as sub unit vaccines and inhibitory peptides.

Particularly useful are the viral polymerase protein, the spike protein, the nucleocapsid or antigenic fragments thereof for inclusion in a vaccine as antigen or subunit immunogen, in particular, fragments derived from a viral spike protein, preferably the S1 spike protein is provide for use in a vaccine, in particular, fragments of the S1 protein, such as fragment 1 357, or fragment 358 747, or fragment 358–588, or homologues thereof, as depicted in figure 17 that were are shown herein to interact with DPP4 and to elicit neutralizing

antibodies, or fragment 363 593 of the spike protein of HKU4 Co remarkably interacting with DPP4 as well, as shown in figure 32, but inactivated whole virus can also be used in a vaccine. Particularly useful are those proteinaceous substances that are encoded by recombinant nucleic acid fragments that are identified for phylogenetic analyses, of course, preferred are those that are within the preferred bounds and metes of ORFs useful in phylogenetic analyses, in particular, for eliciting MERS-CoV specific antibodies, whether in vivo (e.g., for protective purposes such as by vaccination or for providing diagnostic antibodies) or in vitro (e.g., by phage display technology or another technique useful for generating synthetic antibodies). Similarly, the invention provides an antigen comprising a proteinaceous molecule or MERS-CoV-specific fragment thereof according to the invention, reactive with an antibody according to the invention. Such an antibody as herein provided is preferably reactive with a fragment of the S1 protein, such as fragment 1 357, or fragment 358 747, preferably fragment 358–588, of MERS-CoV or homologues thereof, as depicted in figure 17 and 32.

The invention also provides a pharmaceutical composition comprising a virus, a nucleic acid, a proteinaceous molecule or fragment thereof, preferably consisting of the amino acid sequence 358 588 of MERS CoV or of the sequence 363 593 of the spike protein of HKU4 CoV, more preferably having at least a part of the amino acid sequence 358 588 of MERS CoV or of the sequence 363 593 of the spike protein of HKU4 CoV as depicted herein.

An antigen and/or an antibody according to the invention can, for example, be used in a method for the treatment or prevention of a MERS HCoV infection and/or a respiratory illness comprising providing an individual with a pharmaceutical composition according to the invention, for example as a vaccination against useful against infection with corona viruses that use DPP4 as a virus receptor such as seen with MERS-CoV infection or HKU4-CoV infection. This is most useful when the individual comprises a human. Antibodies directed against MERS HCoV proteins, especially against the spike protein of MERS HCoV, preferably against the amino acid sequence 358 588 or the sequence 363 593 of the spike protein of HKU4 CoV, or more preferably directed against at least a part of the amino acid sequence 358 588 of MERS CoV or of the sequence 363 593 of the spike protein of HKU4 Co are herein also provided and are useful for prophylactic or therapeutic purposes, as passive vaccines or part of an anti-serum useful to protect against infection with corona viruses that use DPP4 as a virus receptor, such as MERS-CoV or HKU4-CoV. It is known from other coronaviruses that the spike protein is a very strong antigen and that antibodies against spike protein can be used in prophylactic and therapeutic treatment.

The invention also proteinaceous substance having been provided with a isolated or recombinant proteinaceous molecule or MERS-CoV-specific fragment thereof encoded by a nucleic acid according to the invention and additionally comprising at least a fragment of an N-terminal dipeptidyl peptidase protein. In a preferred embodiment, the invention provides a proteinaceous substance having been provided with a proteinaceous molecule or MERS-CoV-specific viral protein or fragment thereof encoded by a nucleic acid according to the invention. Useful proteinaceous molecules are for example derived from any of the genes or genomic fragments or open reading frames (ORFs) derivable from a virus according to the invention. Particularly useful are the viral polymerase protein, the spike protein, the nucleocapsid or antigenic fragments thereof, but inactivated whole virus can also be used. Particularly useful are also those proteinaceous substances that are encoded by recombinant nucleic acid fragments that are identified for phylogenetic analyses, of course preferred are those that are within the preferred bounds and metes of ORFs useful in phylogenetic analyses, in particular for eliciting MERS-CoV specific antibodies,

The invention also proteinaceous substance having been provided with a isolated or recombinant proteinaceous molecule according to the invention or MERS-CoV-specific fragment thereof wherein said proteinaceous molecule comprises an ectodomain of a spike protein, said ectodomain preferably derived from the

S1 region of a coronavirus. In another preferred embodiment, the invention also proteinaceous substance having been provided with a isolated or recombinant proteinaceous molecule or MERS-CoV-specific fragment thereof wherein said peptidase protein is a dipeptidyl peptidase 4 (DPP4), preferably human DPP4, or a fragment of DPP4.

Typically the invention provides a proteinaceous substance having been provided with an isolated or recombinant proteinaceous molecule or MERS-CoV-specific fragment thereof encoded by a nucleic acid according to the invention and additionally comprising at least a fragment of an N-terminal dipeptidyl peptidase protein, said substance having been subjected to crystallization. The invention also provides a container provided with a proteinaceous substance having been provided with a isolated or recombinant proteinaceous molecule according to the invention or fragment thereof encoded by a nucleic acid according to the invention and additionally having been provided with or comprising at least a fragment of an N-terminal dipeptidyl peptidase protein.

The invention also provides a method of identifying a candidate modulator as an agent that modulates the function of a dipeptidyl peptidase, said method comprising: providing a proteinaceous substance according to the invention in the presence and absence of said candidate modulator under conditions permitting binding of said proteinaceous molecule of first fragment of viral fragment with said fragment of said peptidase protein, measuring binding of said molecule to said fragment, wherein a decrease or increase in binding in the presence of said candidate modulator, relative to binding in the absence of said candidate modulator, identifies said candidate modulator as an agent that modulates the function of a dipeptidyl peptidase. It is preferred that said molecule and/or said fragment is detectably labeled, preferably with a moiety selected from the group consisting of a radioisotope, a fluorophore, a quencher of fluorescence, an enzyme, and an affinity tag.

The invention also provides use of at least a fragment of a viral protein as an agent that modulates the function of an N-terminal dipeptidyl peptidase, for example such use is provided herein in a method according to the invention. Similarly, the invention provides use of a fragment of an N-terminal dipeptidyl peptidase as an agent that modulates the function of a viral protein for example such use is provided herein in a method according to the invention. The invention also provides use of an inhibitor of N-terminal dipeptidyl peptidase, such as ADA, or a functional equivalent thereof, in a method for detecting inhibition of replication of a virus in a cell, preferably wherein said peptidase is DPP4, more preferably wherein said virus is a Coronavirus. The invention also provides use of an inhibitor of N-terminal dipeptidyl peptidase cell-surface expression, such as adenosine, or a functional equivalent thereof, in a method for detecting inhibition of replication of a virus in a cell, preferably wherein said peptidase is DPP4, more preferably wherein said virus is a Coronavirus.

The invention thus further provides a method of detecting, in a sample, the presence of an agent that modulates the function of a dipeptidyl peptidase, the method comprising providing a substance with a first and a second fragment according to the invention in the presence and absence of the sample under conditions permitting binding of the first fragment with the second fragment. Measuring binding of the first fragment to the second fragment, wherein a decrease or increase in binding in the presence of the sample, relative to binding in the absence of the sample, identifies the sample as comprising an agent that modulates the function of a dipeptidyl peptidase.

The invention further provides a method of identifying a candidate modulator as an agent that modulates the function of a dipeptidyl peptidase, the method comprising providing a substance with a first and a second fragment according to the invention in the presence and absence of the candidate modulator under conditions permitting determining enzymatic activity of a peptidase. Measuring enzymatic activity of a peptidase, wherein a decrease or increase in enzymatic activity in the presence of the candidate modulator, relative to binding in the absence of the candidate modulator, identifies the candidate modulator as an agent that modulates the function of a dipeptidyl peptidase.

The invention further provides a method of detecting, in a sample, the presence of an agent that modulates the function of a dipeptidyl peptidase, the method comprising providing a substance with a first and a second fragment according to the invention in the presence and absence of the sample under conditions permitting determining enzymatic activity of a peptidase. Measuring enzymatic activity of a peptidase, wherein a decrease or increase in enzymatic activity in the presence of the sample, relative to binding in the absence of the sample, identifies the sample as comprising an agent that modulates the function of a dipeptidyl peptidase.

In a preferred embodiment, the invention further provides a method of identifying a candidate modulator as an agent that modulates the function of a dipeptidyl peptidase or provides a method of detecting, in a sample, the presence of an agent that modulates the function of a dipeptidyl peptidase, wherein the first fragment and/or the second fragment is detectably labeled, preferably wherein the first fragment and/or the second fragment is detectably labeled with a moiety selected from the group consisting of a radioisotope, a fluorophore, a quencher of fluorescence, an enzyme, and an affinity tag. It is also provided that the substance comprises a cell expressing the first fragment and/or the second fragment. The invention also provides a method of identifying a candidate modulator as an agent that modulates the function of a dipeptidyl peptidase, said method comprising providing a proteinaceous substance according to the invention in the presence and absence of said candidate modulator under conditions permitting binding of a first fragment derived from a virus with a second fragment derived from a peptidase protein, and measuring binding of said first to said second fragment, wherein a decrease or increase in binding in the presence of said candidate modulator, relative to binding in the absence of said candidate modulator, identifies said candidate modulator as an agent that modulates the function of a dipeptidyl peptidase, it is preferred that said first and/or said second fragment is detectably labeled, preferably with a moiety selected from the group consisting of a radioisotope, a fluorophore, a quencher of fluorescence, an enzyme, and an affinity tag. The invention also provides a method of identifying a candidate antiviral agent as an agent that modulates the binding of a virus to dipeptidyl peptidase, said method comprising providing a proteinaceous substance according to the invention in the presence and absence of said candidate antiviral agent under conditions permitting binding of a first fragment derived from a virus with a second fragment derived from a peptidase protein, measuring binding of said first to said second fragment, wherein a decrease or increase in binding in the presence of said antiviral agent, relative to binding in the absence of said candidate modulator, identifies said antiviral agent as an agent that modulates the function of a dipeptidyl peptidase. It is preferred that said first and/or second fragment is detectably labeled, preferably with a moiety selected from the group consisting of a radioisotope, a fluorophore, a quencher of fluorescence, an enzyme, and an affinity tag.

The invention also provides use of a substance, a container or a method according to the invention for identifying an agent that modulates the function of a peptidase or a viral protein, use of an isolated fragment of a viral protein, preferably of a viral spike protein as provided herein, or recombinant or synthetic peptide derived thereof as provided herein, as an agent that modulates the function of an N terminal dipeptidyl peptidase, and use of an isolated fragment of a N terminal dipeptidyl peptidase, preferably of a soluble fragment of said peptidase as provided herein, or recombinant or synthetic peptide derived thereof as provided herein, as an agent that modulates the function of a viral protein. The invention further provides use of an inhibitor, preferably ADA, or a functional equivalent thereof, of N terminal dipeptidyl peptidase activity of a cell, for inhibition of replication of a virus in a cell, in particular, wherein the peptidase is DPP4, preferably human DPP4, preferably wherein the virus is a Coronavirus. The invention further provides use of an inhibitor, preferably adenosine, or a functional equivalent thereof, of N terminal dipeptidyl peptidase cell surface expression on a cell, for inhibition of replication of a virus in a cell, in particular, wherein the peptidase is DPP4, preferably human DPP4, preferably wherein the virus is a Coronavirus.

In describing protein or peptide composition, structure and function herein, reference is made to amino acids. In the present specification, amino acid residues are expressed by using the following abbreviations.

Also, unless explicitly otherwise indicated, the amino acid sequences of peptides and proteins are identified from N terminal to C terminal, left terminal to right terminal, the N terminal being identified as a first residue. Ala: alanine residue; Asp: aspartate residue; Glu: glutamate residue; Phe: phenylalanine residue; Gly: glycine residue; His: histidine residue; Ile: isoleucine residue; Lys: lysine residue; Leu: leucine residue; Met: methionine residue; Asn: asparagine residue; Pro: proline residue; Gln: glutamine residue; Arg: arginine residue; Ser: serine residue; Thr: threonine residue; Val: valine residue; Trp: tryptophane residue; Tyr: tyrosine residue; Cys: cysteine residue. The amino acids may also be referred to by their conventional one letter code abbreviations; A=Ala; T=Thr; V=Val; C=Cys; L=Leu; Y=Tyr; I=Ile; N=Asn; P=Pro; Q=Gln; F=Phe; D=Asp; W=Trp; E=Glu; M=Met; K=Lys; G=Gly; R=Arg; S=Ser; and H=His.

Figure legends

Figure 1. Light microscopy images of LLC-MK2 cells (A, B) and VERO cells (C, D) inoculated with phosphate-buffered saline (A, C) or novel human coronavirus HCoV-SA1 (B, D) 5 days after inoculation.

Figure 2. Results of pan-coronavirus PCR. Various samples (numbered 1-12) of cell culture supernatants infected with HCoV-SA1 reacted with a pair of primers specific for the coronavirus family. A positive control virus (HCoV-NL63) was also reactive.

Figure 3. Partial open reading frame of HCoV-SA1.

Figure 4. Maximum Likelihood tree of partial polymerase gene sequences of representative coronaviruses. HCoC-SA1 is shown in the cluster on the right hand side of the tree, labeled as "New HCoV". The cluster of viruses at the top represents viruses in the genus alphacoronavirus. The Beluga whale coronavirus (BWCoV) represents a gammacoronavirus, while the Bulbul-CoV and IBV represent a proposed new genus of the coronavirinae.

Figure 5 file N.rtf nucleocapsid (N) protein

Figure 6 file M.rtf matrix (M) protein

Figure 7 file E.rtf small envelope (E) protein

Figure 8 file NS3d.rtf non-structural gene NS3d

Figure 9 file NS3c.rtf non-structural gene NS3c

Figure 10 file NS3b.rtf non-structural gene NS3b

Figure 11 file NS3a.rtf non-structural gene NS3a

Figure 12 file S.rtf spike surface glycoprotein (S)

Figure 13 file Orf1ab.rtf encoding many enzymatic products among which the replicase

Figure 14 file HCoV-SA1.rtf

Figure 15 HCoV-SA1.rtf translation 3 frames

Figure 16 Amino acid sequence of the spike protein of HCoV EMC (HCoV SA1). Panel A, schematic presentation of the HCoV EMC S and S1 Fc fusion protein. Position of the predicted N glycosylation sites (Ψ; predicted by the NetNGlyc server) and TM domain (yellow bar; predicted by the TMHMM server) are indicated in the full length S protein. The border between the S1 and S2 subunits is marked by the presence of a predicted furin cleavage site (red triangle; predicted by the ProP 1.0 server). Residues 1 747 comprise the N terminal region. Panel B, amino acid sequence of the spike protein with the S1 region indicated in red.

Figure 17 Amino acid sequence and domain structure of residues 1 747 of the S1 spike protein of HCoV EMC (HCoV SA1). RBD = Receptor Binding Region.

Figure 18 Domain structure and amino acid sequence of residues 1 766 of human DPP IV, domain borders based on crystal structure (Rasmussen, Nat. Struct. Biol. 2003, herein included by reference).

Figure 19 Binding of HCoV EMC S1 is correlated to infection with HCoV EMC in vero cells (Panel A), Cos 7 cells (Panel B) Huh7 cells (Panel C) and bat cells (Panel D). Shown on the left is the FACS analysis of HCoV EMC S1 binding (red line), a feline CoV S1 protein as control (blue line) and non stained cells (black line). In the middle panels, HCoV EMC infected cells are visualized using an antiserum that recognizes the NSp4 protein and on the left, supernatants of the infected cells are tested by Taqman for the presence of viral transcripts at 0, 20 and 40 hours after infection.

Figure 20 Immunoprecipitation with S1 on Huh7cells and mass spec analysis reveals cd26 as the interacting protein.

Figure 21 Peptides identified in fraction 2 are indicated in red and relate to the fragment or topological domain involving residues 29 766 comprising the extracellular region (ectodomain) of the membrane bound DPP4 (Uniprot identifier P27487) but do not relate to the cytoplasmic domain (residues 1 6) nor to the helical Signal anchor for type II membrane protein domain (residues 7 28) of membrane bound DPP4. Soluble DPP4 runs from residue 39 to residue 766.

Figure 22 HCoV EMC and SARS CoV S1 Fc proteins (2.5 µg) were mock incubated or incubated with 12.5 µg soluble DPP IV (sDPP IV) or soluble ACE2 (sACE2) in 100 µl PBS. Precipitates were washed thrice with lysis buffer and once with PBS, and subjected to a NOVEX® 4 12% Tris Glycine gradient gel (Invitrogen) under non reducing conditions.

Figure 23 Cells were washed twice with ice cold PBS, scraped off the plastic with a rubber policeman and suspended into single cells by pipetting cells up and down. S1 binding of cells was measured by incubating 2.5 x 105 cells with 15 µg/ml of S1 Fc followed by incubation with the fluorescent dye Alexa488 labeled goat anti human IgG antibody and analyzed by flow cytometry.

Figure 24 Inhibition of HCoV EMC replication in Huh7 cells by antibodies to DPP4. Huh7 cells were incubated with 20 µg/ml goat polyclonal antiserum against DPP4, a goat antiserum against ACE2, normal goat serum or left untreated. After 1 hour incubation, the cells were infected with HCoV EMC at a multiplicity of infection of 0.01 and left for 1 hour. Cells were washed twice and again incubated with medium containing the respective antibodies. Supernatant collected at 2 hours (open bars) and 20 hours (closed bars) was tested for presence of HCoV using a Taqman assay. Results are shown as Δ Ct. HCoV EMC infection of Huh7 cells is inhibited by antibodies to DPP4 but not by the other antibodies tested.

Figure 25 Cos7 cells transfected with plasmids encoding human DPP4 (hDPP4) or bat DPP4 (bDPP4), a control plasmid (pcDNA) or left untreated were infected with HCoV EMC at a multiplicity of infection of 1 and left for 1 hour. Cells were washed twice and supernatant collected at 2 hours (open bars), 20 hours (blue bars) and 40 hours (red bars) was tested for presence of HCoV using a Taqman assay. Results are shown as Δ Ct.

Figure 26 Blocking of DPP4 –S1 binding by antibodies directed against S1 serum from a macaque infected with HCoV EMC inhibits binding of recombinant S1 to Huh7 cells. Serum at a dilution of 1:20, obtained from macaques at day 0 (blue line) and day 14 (red line) after infection with 5 x 107 TCID50 HCoV EMC, was preincubated for 1 hour at room temperature with 1.25 µg/ml recombinant S1 protein that was biotinylated and subsequently incubated on Huh7 cells. After treatment with FITC-labeled streptavidin, cells were analyzed for fluorescence. In gray background, binding using a control biotinylated protein is shown.

Figure 27 Inhibition of HCoV EMC replication in Huh7 cells by soluble adenosine deaminase (ADA). Huh7 cells were incubated with different concentrations of recombinant soluble ADA (closed bars) or recombinant soluble ACE2 (open bars). After 1 hour incubation, the cells were infected with HCoV EMC at a multiplicity of infection of 0.01. After 8 hours, cells were fixed and stained with a rabbit antiserum against HCoV EMC nsp4 and cells were counted. Results are shown as number of infected cells per well. Infection of Huh7 cells is inhibited by recombinant soluble ADA but not by recombinant soluble ACE2.

Figure 28 Inhibition of HCoV EMC replication in Huh7 cells by soluble DPP4. Different concentrations of recombinant soluble DPP4 (open bars) or recombinant soluble ACE2 (closed bars) were incubated with HCoV EMC for 1 hour at 37°C and used to infect Huh7 cells. After 8 hours, cells were fixed and stained with a rabbit antiserum against HCoV EMC nsp4 and cells were counted. Results are shown as number of infected cells per well. Infection of Huh7 cells is inhibited by recombinant soluble DPP4 but not by recombinant soluble ACE2.

Figure 29 Receptor binding domains in betacoronavirus spike proteins and S1 Fc expression constructs. Panel a), schematic representation of the betacoronaviruses SARS CoV, hCoV EMC S and MHV (strain A59) spike (S) protein sequence (drawn to scale) aligned at the S1 S2 junction. The known receptor binding domain in the S1 subunit of MHV and SARS CoV S proteins and their corresponding homologous regions in hCoV EMC S as defined by ClustalW alignment are indicated. Positions of the transmembrane domain (yellow bar; predicted by the TMHMM server) and of the predicted N glycosylation sites (Ψ; predicted by the NetNGlyc server, only shown for the hCoV EMC S) are indicated. The border between the S1 and S2 subunits of the spike protein is represented by a vertical white line. Panel b), upper panel, schematic presentation of the hCoV EMC S1 subunit (residues 1 751) sequence. Cysteine positions in S1 subunit are indicated by vertical white lines with corresponding amino acid positions on top. Positions of cysteines highly conserved among betacoronaviruses S1 proteins are in bold. Predicted disulfide bond connections inferred from the structure of the SARS CoV receptor binding domain are presented as connecting black lines underneath. Lower panel, domains of the hCoV EMC S1 subunit expressed as Fc chimeras.

Figure 30 The DPP4 binding domain is located within residues 358 588 of the hCoV EMC spike protein and efficiently elicits neutralizing antibodies. Panel a), S1 Fc chimeric proteins and soluble DPP4 (sDPP4) receptor were expressed from HEK 293T cells and purified from the culture supernatant. S1 Fc proteins were mixed with sDPP4 followed by protein A sepharose affinity isolation, analyzed on a NOVEX® 4 12% Tris Glycine gradient gel under non reducing conditions, and stained with GelCodeBlue reagent. Position of the S1 Fc proteins, running as dimers under non reducing conditions due to an Fc interchain disulphide bond, and sDPP4 as well as the sizes of the marker proteins are indicated. Individual proteins were loaded as controls. Panel b), binding of hCoV EMC S1 Fc proteins to DPP4 expressing cells. 2.5 x 105 HEK 293T cells transfected with control pCAGGS (grey shaded area) or with pCAGGS DPP4 (black line) expression plasmid were incubated with 15 µg/ml of the indicated S1 Fc followed by incubation with DyLight488 labeled goat anti human IgG antibody and analysis by flow cytometry. An Fc chimera containing the S1 of infectious bronchitis virus (IBV S1 Fc) was taken along as a negative control. Panel c), neutralization of hCoV EMC infection by rabbit antisera raised against the S1 Fc 1 747, 1 357 and 358 588 variants. Virus (100 pfu) was premixed 1:1 with serial dilutions of sera obtained (open bars) or after immunization (closed bars) prior to inoculation onto VERO cells and virus infection was monitored by the occurrence of CPE at 72 hours post infection. Virus neutralization titers (VNT) were determined in quadruplicate as the highest serum dilutions that completely prevent CPE. The experiment was carried out twice and the data of one representative experiment are shown.

Figure 31 Localization of receptor binding domains in coronavirus spike proteins. Schematic presentation of the spike proteins of the alphacoronaviruses TGEV and hCoV NL63 and of the betacoronaviruses SARS CoV, hCoV EMC and MHV (drawn to scale), aligned at the S1 S2 junction. Blue boxes represent the receptor binding domains (RBD) and indicate the engaged receptor. The RBD of TGEV, hCoV NL63, SARS CoV and MHV have been confirmed by crystallography (12, 15, 22, 26). Grey boxes indicate the transmembrane domain. Sequence IDs: TGEV (ABG89335.1), hCoV NL63 (NC_005831.2), SARS CoV (NP_828851.1), hCoV EMC (AFS88936.1), MHV (NC_001846.1).

Figure 32 Residues 363 593 of the spike protein of HKU4 CoV bind to human DPP4. Shown is the binding ability of different S1 Fc proteins to DPP4 expressing cells. 2.5 x 105 HEK 293T cells transfected with

control pCAGGS (grey shaded area) or with pCAGGS DPP4 (black line) expression plasmid were incubated with 15 µg/ml of the hCoV EMC S1 Fc followed by incubation with DyLight488 labeled goat anti human IgG antibody and analysis by flow cytometry. S1 Fc protein chimeras were tested containing the hCoV EMC S1 subunit (residues 1 747), the hCoV EMC spike receptor binding domain (RBD; residues 358 588) or the hCoV EMC RBD homologous regions of the spike proteins of HKU4 CoV (residues 363 593) and HKU5 CoV (residues 366 586). Mock incubated cells (mock) or cells incubated with an Fc chimera containing the S1 of feline infectious peritonitis virus (FIPV S1 Fc) was taken along as negative controls.

HKU4 CoV spike (S) protein ID [YP_001039953.1].

HKU5 CoV spike (S) protein ID [YP_001039962.1].

Region in S homologous to hCoV EMC RBD highlighted in yellow.

Figure 33. Characterization of the functional MERS-CoV DPP4 receptor S1 binding site.

A, Different plasmids encoding either full length human DPP4, ferret DPP4 or human-ferret DPP4 chimera's (human-ferret-human and ferret-human ferret, HFH and FHF respectively) were constructed. B, DPP4 expression and S1 binding to cells transfected with different DPP4 constructs as analysed by FACS analysis. C, MERS-CoV RNA levels in supernatants of DPP4 transfected cells infected with MERS-CoV at 2 and 20 h after infection

using a TaqMan assay, expressed as genome equivalents (GE; half maximal tissue-culture infectious dose (TCID50) per ml). D, S1 binding to cells transfected with different hDPP4 mutants. E, MERS-CoV infection of cells transfected with different hDPP4 constructs. Data in panel a and b were corrected for DPP4 expression of the different constructs.

Detailed description

Novel human coronavirus HCoV-SA1

Classification:

Order: Nidovirales

Family: Coronaviridae

Subfamily: Coronavirinae

Genus: Betacoronavirus

Lineage: C

Example 1

Virus was isolated from a 60-year old man with acute pneumonia and acute renal failure in Saudi Arabia.

Virus was isolated from sputum specimen in VERO cells and LLC-MK2 cells.

Five days after inoculation, cytopathic effects were observed, consisting of rounding of the cells, detachment of cells, and syncytium formation (Figure 1).

Cells in the original sputum sample and infected cultured cells were also tested with specific antibodies against influenza A and B viruses, parainfluenza viruses types 1-3, respiratory syncytial virus, and adenovirus, but such tests yielded negative results. Sputum specimens and infected cell culture supernatants did not react in PCR-based assays specific for paramyxoviruses, enteroviruses, and adenoviruses. However, these samples did react with PCR-based assays to detect all coronaviruses. A 251 nucleotide fragment was amplified with one such test (Vijgen, L., E. Moes, E. Keyaerts, S. Li, and M. Van Ranst. 2008. A pancoronavirus RT-PCR assay for detection of all known coronaviruses. Methods Mol Biol 454:3-12). A second PCR-based assay to detect all coronaviruses (Drosten C, Günther S, Preiser W, van der Werf S, Brodt HR, Becker S, Rabenau H, Panning M,

Kolesnikova L, Fouchier RA, Berger A, Burguière AM, Cinatl J, Eickmann M, Escriou N, Grywna K, Kramme S, Manuguerra JC, Müller S, Rickerts V, Stürmer M, Vieth S, Klenk HD, Osterhaus AD, Schmitz H, Doerr HW. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N Engl J Med. 348, 1967-76 (2003)) also yielded positive results (Figure 2).

Example 2

Viral RNA was isolated from infected cell culture supernatants using a High Pure RNA Isolation Kit (Roche). Extracted RNA was copied to cDNA by reverse transcriptase using random hexamers. Pan-coronavirus polymerase chain reaction (PCR) was used to amplify a conserved region of open reading frame 1b (Drosten C, Günther S, Preiser W, van der Werf S, Brodt HR, Becker S, Rabenau H, Panning M, Kolesnikova L, Fouchier RA, Berger A, Burguière AM, Cinatl J, Eickmann M, Escriou N, Grywna K, Kramme S, Manuguerra JC, Müller S, Rickerts V, Stürmer M, Vieth S, Klenk HD, Osterhaus AD, Schmitz H, Doerr HW. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N Engl J Med. 348, 1967-76 (2003)). The PCR fragments of the pan-coronavirus PCRs were sequenced. To this end, PCR products were purified from the gel and sequenced using a BigDye Terminator v3.1 Cycle sequencing kit (Applied Biosystems, Nieuwerkerk a/d IJssel, the Netherlands) and a 3130XL genetic analyzer (Applied Biosystems), according to the instructions of the manufacturer. The sequence clearly corresponded with conserved region of open reading frame 1b of a coronavirus (Figure 3).

Example 3

Reference coronavirus genome sequences were downloaded from GenBank and the part of the genomes that corresponded with the amplified fragment of HCoV-SA1 were aligned. A Maximum Likelihood tree was constructed to infer the phylogenetic relationships (Figure 4). This phylogenetic tree showed that the new HCoV-SA1 belongs to lineage C of the genus Betacoronavirus, along with the bat coronaviruses HKU4 and HKU5. The Betacoronavirus genus contains 3 additional lineages (A, B, D). HCoV-HKU1 and HCoV-OC43 belong to lineage A while SARS-CoV belongs to lineage B. Lineage D does not contain any human pathogens, and is represented in the tree by Rousettus bat coronavirus HKU9.

HCoV-SA1 is thus clearly distinct from previously known human coronaviruses. HCoV-NL63 and HCoV-229E are even more distinct from HCoV-SA1, since these two human pathogens belong to a different genus, the Alphacoronavirus genus.

Example 4

To further characterize the virus genome, viral RNA was extracted from infected cell culture supernatant using the High Pure RNA Isolation Kit (Roche). RNA was subjected to reverse transcriptase using circular permuted primers (Welsh, J. & McClelland, M. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res. 18, 7213–7218 (1990)) that were extended with random hexamer sequences. The amount of DNA was amplified by polymerase chain reaction (PCR), using the circular permuted primers. The randomly amplified fragments were sequenced using the 454/Roche GS-FLX sequencing platform. A fragment library was created according to the manufacturer's protocol without DNA fragmentation (GS FLX Titanium Rapid Library Preparation, Roche). The emPCR (Amplification Method Lib-L) and GS junior sequencing run was performed according to instructions of the manufacturer (Roche). The sequence reads were trimmed at 30 nucleotides from the 3' and 5' ends to remove all primer sequences. Sequence reads from the GS-FLX sequencing data were assembled into contigs using CLC Genomics software 4.6.1. Using this "deep-sequencing" approach on the 454-sequencing platform, approximately 80% of the virus genome sequence was obtained. Subsequently, specific primers were designed to amplify 30 overlapping fragments of approximately 1500 basepairs by PCR. Each of these PCR products was sequenced using conventional Sanger sequencing. To this end, PCR products were

purified from the gel and sequenced using a BigDye Terminator v3.1 Cycle sequencing kit (Applied Biosystems, Nieuwerkerk a/d IJssel, the Netherlands) and a 3130XL genetic analyzer (Applied Biosystems), according to the instructions of the manufacturer. The nearly full-length sequence is presented in file HCoV-SA1.rtf. This sequence contains some uncertainties within the extreme 50 nucleotides of both ends. However, this information is not required to classify the coronavirus. The same figure also displays the full coding potential of HCoV-SA1. As a minimum, the HCoV-SA1 virus genome encodes the open reading frames common to the virus of the betacoronavirus genus, including orf1ab that encodes many enzymatic products, the spike surface glycoprotein (S), the non-structural genes NS3a, NS3b, NS3c, NS3d, the small envelope (E) protein, the matrix (M) protein, and the nucleocapsid (N) protein. Open reading frames are presented in files Orf1ab.rtf, S.rtf, NS3a.rtf, NS3b.rtf, NS3c.rtf, NS3d.rtf, E.rtf, M.rtf, N.rtf. Other open reading frames may be present.

Example 5

Comparison of the Orf1ab gene product of HCoV-SA1 with those of the other members of the Betacoronavirus genus, HKU4 and HKU5 was used to test if HCoV-SA1 belongs to one of these known species or represents a new species within the genus. The International Committee on the Taxonomy of Viruses (ICTV) considers viruses that share more than 90% aa sequence identity in the conserved replicase domains to belong to the same species. This 90% identity threshold serves as the sole species demarcation criterion. Since amino acid sequence identity of Orf1ab between HCoV-SA1 and HKU4 and HKU5 is below 74% (Table 1), we conclude that HCoV-SA1 represents a novel species of the Betacoronavirus genus, although such classification requires ICTV approval.

Table 1. Percentage amino acid sequence identity between ORF1ab of HCoV-SA1, HKU4 (Genbank accession numbers EF065505-EF065508) and HKU5 (accession numbers EF065509-EF065512)

	HCoV-SA1	HKU4	HKU5
HCoV-SA1	100%	72%	74%
HKU4	72%	99-100%	77%
HKU5	74%	77%	99-100%

The present invention in particular also relates to the spike (S) protein of a coronavirus and fragments thereof as depicted in figures 16 and 17.

The present invention also relates to a member of the S9 family of human proteases known as dipeptidyl peptidase IV (DPPIV, figure 18), and fragments thereof.

Protein Expression

Example 6

A plasmid encoding HCoV EMC S1 Fc was generated by ligating a fragment encoding the S1 region (residues 1 747) into the pCAGGS expression vector as an N terminal fusion with the fragment encoding the Fc domain of human IgG (figures 1 and 2). Likewise, an S1 Fc expression plasmid was made for the SARS coronavirus S1 subunit (strain Urbani: residues 1 676) and the FIPV S1 subunit (strain 79 1146; residues 1 788). S1 Fc proteins were expressed by transfection of the expression plasmids into 293T cells and affinity purified from the culture supernatant using protein A sepharose beads.

Example 7

A plasmid encoding the ectodomain of human DPP4 (figure 18) was generated by ligating a fragment encoding residues 39 766 of human DPP4 into a pCD5 expression vector encoding the signal sequence of CD5 and a OneSTrEP affinity tag (IBA GmbH). Soluble DPP4 ectodomain was expressed by transfection of the

expression plasmid into 293T cells and affinity purified from the culture supernatant using Streptactin sepharose beads (IBA GmbH).

Example 8

A plasmid encoding HCoV EMC S1 Fc was generated by ligating a fragment encoding the S1 region (residues 1 747) into the pCAGGS expression vector as an N terminal fusion with the fragment encoding the Fc domain of human IgG separated by a thrombin cleavage site. Likewise, an Fc expression plasmid was made for the SARS coronavirus S1 subunit (isolate CUHK W1: residues 1 676), the FIPV S1 subunit (isolate 79 1146; residues 1 788) and the ectodomain of human ACE2 (sACE2; residues 1 614). Fc chimeric proteins were expressed by transfection of the expression plasmids into 293T cells and affinity purified from the culture supernatant using protein A sepharose beads (GE Healthcare). Purified ACE2 Fc was cleaved with thrombin and soluble ACE2 was purified by gel filtration.

Example 9

A plasmid encoding the ectodomain of human DPP IV (sDPP IV) was generated by ligating a fragment encoding residues 39 766 of human DPP IV into a pCD5 expression vector encoding the signal sequence of CD5 and the OneSTrEP tag (IBA GmbH). Soluble DPP IV ectodomain was expressed by transfection of the expression plasmid into HEK 293T cells and affinity purified from the culture supernatant using Strep Tactin Sepharose beads (IBA GmbH).

Pull down; immunoprecipitation and detection of DPP4

Example 10

The immunoprecipitation protocol was essentially carried out as described before with some modifications (Liet al., 2003, Nature 426:450, included herein by reference). In short, Huh 7 cells were washed twice with ice cold PBS, scraped off the plastic with a rubber policeman, pelleted and lysed in ice cold lysis buffer (0.3% DDM in PBS) containing protease inhibitors (Roche Complete Mini) at a final density of ~2.5 x 107 cells/mL. Cell lysates were precleared with protein A sepharose beads after which 10 micrograms of probe S1 Fc was added to 1 ml of cell lysate and incubated for 1 hour at 4°C under rotation. Precipitates were washed thrice with lysis buffer and once with PBS and subjected to NOVEX® 4 12% Tris Glycine gradient gel (Invitrogen) under reducing and non reducing conditions. A distinct 110 kDa band precipitated with EMC S1 Fc was visualized by GelCodeBlue staining, excised from the gel, incubated with trypsin and analyzed by MS. Results are shown in figure 20 and results of target analyses are shown in figure 21.

Example 11

DPP4 cell surface expression was measured using S1 Fc proteins. Cells were washed twice with ice cold PBS, scraped off the plastic with a rubber policeman and suspended into single cells by pipetting cells up and down. S1 binding of cells was measured by incubating 2.5 x 105 cells with 15 µg/ml of S1 Fc followed by incubation with the fluorescent dye Alexa488 labeled goat anti human IgG antibody and analyzed by flow cytometry. Results are shown in figure 23.

RNA extraction and quantitative RT PCR

Example 12

RNA from 200 µl of supernatant was isolated with the Magnapure LC total nucleic acid isolation kit (Roche) external lysis protocol and eluted in 100 µl. HCoV EMC RNA was quantified on the ABI prism 7700, with use of the Taqman Reverse Transcription Reagents and Taqman PCR Core Reagent kit (Applied Biosystems), using 20 µl isolated RNA, 1× Taqman buffer, 5.5 mM MgCl2, 1.2 mM dNTPs, 0.25 U Amplitaq gold DNA polymerase, 0.25 U Multiscribe reverse transcriptase, 0.4 U RNAse inhibitor, 200 nM primers, and 100 nM probe. Amplification parameters were 30 minutes at 48°C, 10 minutes at 95°C, and 40 cycles of 15 seconds at 95°C, and 1

minute at 60°C. RNA dilutions isolated from a HCoV EMC stock were used as a standard. Results are shown in figures 17, 24, 25 and 26.

Example 13

HCoV EMC and SARS CoV S1 Fc proteins (2.5 µg) were mock incubated or incubated with 12.5 µg soluble DPP IV (sDPP IV) or soluble ACE2 (sACE2) in a total volume of 100 µl PBS. Precipitates were washed thrice with lysis buffer and once with PBS, and subjected to a NOVEX® 4 12% Tris Glycine gradient gel (Invitrogen) under non reducing conditions. Results are shown in figure 22.

Identification of DPP4 using mass spec analysis of peptide fragments

Example 14

1D SDS PAGE gel lanes were cut into ~1 mm slices (indicated as nr. 2 in figure 3) using an automatic gel slicer and subjected to in gel reduction with dithiothreitol, alkylation with chloroacetamide and digestion with trypsin (Promega, sequencing grade), essentially as described by Van den Berg et al. (Cell Stem Cell 6:369, included herein by reference). Alternatively, immunoprecipitated proteins were reduced and alkylated on beads similarly as described above. Nanoflow LC MS/MS was performed on either an 1100 series capillary LC system (Agilent Technologies) coupled to an LTQ Orbitrap XL mass spectrometer (Thermo), or an EASY nLC coupled to a Q Exactive mass spectrometer (Thermo), operating in positive mode and equipped with a nanospray source. Peptide mixtures were trapped on a ReproSil C18 reversed phase column (Dr Maisch GmbH; column dimensions 1.5 cm × 100 µm, packed in house) at a flow rate of 8 µl/minute. Peptide separation was performed on ReproSil C18 reversed phase column (Dr Maisch GmbH; column dimensions 15 cm × 50 µm, packed in house) using a linear gradient from 0 to 80% B (A = 0.1% formic acid; B = 80% (v/v) acetonitrile, 0.1% formic acid) in 70 or 120 minutes and at a constant flow rate of 200 nl/minute. The column eluent was directly sprayed into the ESI source of the mass spectrometer. Mass spectra were acquired in continuum mode; fragmentation of the peptides was performed in data dependent mode by CID or HCD. Peak lists were automatically created from raw data files using the Mascot Distiller software (version 2.3; MatrixScience) or Proteome Discoverer (version 1.3; Thermo). The Mascot algorithm (version 2.2; MatrixScience, UK) was used for searching against a Uniprot database (release 2012_10.fasta, taxonomy: Homo sapiens, or Macaca mulatta, or Myotis lucifugus, or Chlorocebus sabaeus, or Felis catus, included herein by reference). The peptide tolerance was set to 10 ppm and the fragment ion tolerance was set to 0.8 Da for CID spectra (LTQ Orbitrap) or to 20 mmu for HCD (Q Exactive) spectra). A maximum number of two missed cleavages by trypsin were allowed and carbamidomethylated cysteine and oxidized methionine were set as fixed and variable modifications, respectively. Results are shown in figure 21.

Inhibition of HCoV EMC replication in Huh7 cells by antibodies to DPP4

Example 15

Huh7 cells were incubated with 20 µg/ml goat polyclonal antiserum against DPP4, a goat antiserum against ACE2, normal goat serum or left untreated. After 1 hour incubation, the cells were infected with HCoV EMC at a multiplicity of infection of 0.01 and left for 1 hour. Cells were washed twice and again incubated with medium containing the respective antibodies. Supernatant collected at 2 hours (open bars) and 20 hours (closed bars) was tested for presence of HcoV using a Taqman assay. Results are shown as Δ Ct in figure 25.

Blocking of DPP4 –S1 binding by antibodies directed against S1

Example 16

Serum from a macaque infected with HCoV EMC inhibits binding of recombinant S1 to Huh7 cells. Serum at a dilution of 1:20, obtained from macaques at day 0 (blue line) and day 14 (red line) after infection with 5 x 107 TCID50 HCoV EMC, was preincubated for 1 hour at room temperature with 1.25 µg/ml recombinant S1

protein that was biotinylated and subsequently incubated on Huh7 cells. After treatment with FITC labeled streptavidin, cells were analyzed for fluorescence. In gray background, binding using a control biotinylated protein is shown (figure 26).

Crystallization and crystals comprising a DPP fragment and a viral protein fragment

Example 17

One aspect of the present invention relates to methods for forming crystals comprising fragments of DPP and viral protein as well as crystals comprising fragments of DPP and viral protein. Crystallization of DPP is essentially known from, for example, U.S. Patent 7,344,852 or U.S. Patent Publication 2005/0260723 that are included herein by reference.

In one embodiment of the present invention, a method for forming crystals comprising fragments of DPPIV and viral protein is provided comprising forming a crystallization volume comprising fragments of DPPIV and viral protein, one or more precipitants, optionally a buffer, optionally a monovalent and/or divalent salt and optionally an organic solvent; and storing the crystallization volume in a container under conditions suitable for crystal formation.

In yet another embodiment, a method for forming crystals comprising fragments of DPPIV and viral protein is provided comprising forming a crystallization volume comprising fragments of DPPIV and viral protein in solution comprising PEG precipitant listed hereinbelow; and storing the crystallization volume in a container under conditions suitable for crystal formation. PEG precipitant 5 50% w/v of precipitant, wherein the precipitant comprises one or more members of the group consisting of PEG MME having a molecular weight range between 300 10000, and PEG having a molecular weight range between 100 10000 pH 5 9. Buffers that may be used include, but are not limited to, tris, bicine, cacodylate, acetate, citrate, MES and combinations thereof. Additives optionally 0.05 to 0.8 M additives wherein the additives comprises sarcosine or 0.5% to 25% additives wherein the additives comprises xylitrol; Protein Concentration 1 mg/ml 50 mg/ml; Temperature 1°C to 25°C.

In yet another embodiment, a method for forming crystals comprising fragments of DPPIV and viral protein is provided comprising forming a crystallization volume comprising fragments of DPPIV and viral protein; introducing crystals comprising fragments of DPPIV and viral protein as nucleation sites, and storing the crystallization volume under conditions suitable for crystal formation.

Crystallization experiments may optionally be performed in volumes commonly used in the art, for example, typically 15, 10, 5, or 2 microliters or less. It is noted that the crystallization volume optionally has a volume of less than 1 microliter, optionally 500, 250, 150, 100, 50 or less nanoliters.

It is also noted that crystallization may be performed by any crystallization method including, but not limited to, batch, dialysis and vapor diffusion (e.g., sitting drop and hanging drop) methods. Micro and/or macro seeding of crystals may also be performed to facilitate crystallization.

In one variation, crystals comprising DPPIV are formed by mixing a substantially pure DPPIV fragment and a substantially pure S1 HCoV EMC fragment with an aqueous buffer containing a precipitant at a concentration just below a concentration necessary to precipitate the proteinaceous substance. One suitable precipitant for crystallizing fragments of DPPIV and viral protein is polyethylene glycol (PEG), which combines some of the characteristics of the salts and other organic precipitants (see, for example, Ward et al., J. Mol. Biol. 98:161, 1975, and McPherson, J. Biol. Chem. 251:6300, 1976.

During a crystallization experiment, water is removed by diffusion or evaporation to increase the concentration of the precipitant, thus creating precipitating conditions for the protein. In one particular variation, crystals are grown by vapor diffusion in hanging drops or sitting drops. According to these methods, a protein/precipitant solution is formed and then allowed to equilibrate in a closed container with a larger aqueous

reservoir having a precipitant concentration for producing crystals. The protein/precipitant solution continues to equilibrate until crystals grow.

By performing submicroliter volume sized crystallization experiments, as detailed in U.S. Patent No. 6,296,673, effective crystallization conditions for forming crystals of fragments of DPPIV and viral protein complex are obtained. In order to accomplish this, systematic broad screen crystallization trials are performed on a DPPIV/viral protein fragment complex using the sitting drop technique.

One skilled in the art will recognize that the crystallization conditions provided herein can be varied and still yield protein crystals comprising fragments of DPPIV and viral protein. As the conditions for the crystallization, physical and chemical factors such as a hydrogen ion concentration (pH), a kind of buffer used and a concentration thereof, a kind of a precipitant added and a concentration thereof, protein concentration, salt concentration, temperature and the like can be involved. A method for controlling and investigating the factors includes batch methods, dialysis methods, vapor diffusion methods (hanging drop method, sitting drop method and the like) and the like, described, for instance, in T.L. Blundell et al., PROTEIN CRYSTALLOGRAPHY, 59 82 (1976), published by Academic Press, or the like.

The method for crystallization includes the batch methods, dialysis methods, vapor diffusion methods and the like. By the above method, physical and chemical factors such as a hydrogen ion concentration (pH), a kind and a concentration of the buffer used, and a kind and a concentration of the precipitant used, and physical and chemical factors such as protein concentration, salt concentration and temperature can be also determined.

The hydrogen ion concentration (pH) can be adjusted with a buffer. It is desired that the buffer is a buffer having buffering action in a broad range of pH, and being capable of suppressing precipitation of a non proteinous crystal between the co existing ion in the solution used during crystallization and the precipitant or the like. The buffer includes Tris hydrochloric acid buffer, phosphate buffer, cacodylate buffer, acetate buffer, citrate buffer, glycine buffer and the like.

The precipitant may be a substance capable of elevating an effective concentration of the protein or changing a hydrogen ion concentration (pH) of the protein solution. Generally, the precipitant includes salts such as ammonium sulfate, sodium sulfate, sodium phosphate, potassium phosphate, sodium citrate, ammonium citrate, sodium chloride, potassium chloride and ammonium chloride; polyethylene glycols having various average molecular weights of about 200, about 1000, about 2000, about 4000, about 6000, about 8000, about 20000 or the like; organic solvents such as 2 methyl 2,4 pentadiol, methanol, ethanol, isopropanol, butanol and acetone, and the like.

The protein concentration may be a concentration suitable for crystallization, and it is desired that the protein concentration is, for example, 1 to 50 mg/ml, preferably 5 to 20 mg/ml, more preferably 7 to 15 mg/ml.

It is desired that the temperature conditions are 3°C to 25°C., preferably 12°C to 22°C.

In the case where the crystallization is carried out by the batch method, the crystallization can be carried out by gradually adding a precipitant solution comprising a precipitant, buffer and the like, so as to form a layer on the top layer of the solution containing the dipeptidyl peptidase to give a mixture, or by gradually adding the solution comprising the DPPIV/viral protein fragment complex, so that the solution is an upper layer of the precipitant solution to give a mixture. Here, the mixture is allowed to stand in a tightly closed vessel or container.

In the case where the crystallization is carried out by the dialysis method, the crystallization can be carried out by placing a solution comprising DPPIV/viral protein fragment complex in a size exclusion semi permeable membrane, and placing a precipitant solution outside of the size exclusion semi permeable membrane as a reservoir solution, thereby diffusing the reservoir solution to the solution comprising the DPPIV/viral protein fragment complex via the semi permeable membrane.

In the case where the crystallization is carried out by the hanging drop method in the vapor diffusion method, the crystallization can be carried out by placing a mixed solution of a solution comprising the DPPIV/viral protein fragment complex and a precipitant solution in a closed vessel allowing to be hanged at a position above the upper space of a reservoir in which the precipitant solution is contained as a reservoir solution, wherein the vapor pressure of the reservoir solution in the reservoir is set to be lower than that of the mixed solution.

In the case where the crystallization is carried out by the sitting drop method in the vapor diffusion method, the crystallization can be carried out by placing a mixed solution comprising a solution comprising the DPPIV/viral protein fragment complex and a precipitant solution in a closed vessel at a position higher than the liquid surface of a reservoir in which the precipitant solution is contained as a reservoir solution, wherein the vapor pressure of the reservoir solution in the reservoir is set to be lower than that of the mixed solution.

The crystallization can be carried out by the sitting drop method from the viewpoint of obtaining excellent quality and large crystals.

Crystals comprising fragments of DPPIV and viral protein have a wide range of uses. Such crystals may, for example, be used to perform X ray or neutron diffraction analysis in order to determine the three dimensional structure of fragments of DPPIV and viral protein and, in particular, to assist in the identification of its active site where fragments may bind. Knowledge of the binding site region allows rational design and construction of ligands including inhibitors. Crystallization and structural determination of fragments of DPPIV mutants and/or viral protein mutants having altered bioactivity allows the evaluation of whether such changes are caused by general structure deformation or by side chain alterations at the substitution site.

Example 18

Because DPPIV protein levels may not always accurately reflect the levels of active DPPIV enzyme, it may be useful to measure DPPIV enzymatic activity in proteinaceous substances instead. Use of a test system that is tested for DPPIV assay in proteinaceous substances as diverse as plasma, serum, urine, saliva, tissue, live cells and cell extracts, and exudates is recommended. Such a test system may be the DPPIV/CD26 Activity Assay for Biological Samples provided by ENZO® life sciences (on the World Wide Web at enzolifesciences.com). A known DPPIV inhibitor, such as P32/98 (Ki=130 nM) is preferably included for use as a control.

Example 19

To examine if cytokines decrease susceptibility to HCoV EMC infection through an effect on cell surface DPP4 expression, we analyzed DPP4 expression after treatment with different cytokines.

All treatments were done in quadruplets (96 well experiments) or triplicate (6 well and 24 well experiments). Cell cultures were grown for 24 to 48 hours and then changed to medium containing 1% newborn calf serum, and treated with recombinant human (r hu) IL 4 (BD Pharmingen), r hu IFN γ, r hu TNF α, r hu IL 13, r hu IL 10, r hu IL 1, r hu TGF beta (Peprotech Inc.) and r hu IFN α (Roche) at a concentration of 10 ng/ml, 48 hours before infection for a further evaluation of changes in DPPIV surface protein expression and changes in susceptibility to HCoV EMC infection. In a first experiment, r hu TGF beta down regulates DPP4 expression and reduces the cells' susceptibility to virus infection and reduces virus replication.

Example 20

To examine if a compound decreases susceptibility to HCoV EMC infection through an effect on cell surface DPP4 expression, we analyze DPP4 expression after treatment with different compounds. Huh 7 cells are grown in Dulbecco's modified Eagle's medium (DMEM), supplemented with 10% fetal bovine serum (FBS), sodium bicarbonate and 20 mM HEPES buffer. All treatments are done in quadruplets (96 well experiments) or triplicate (6 well and 24 well experiments). Cultures are grown for 24 to 48 hours and then changed to medium containing 1% newborn calf serum, and treated with compound, i.e., adenosine (300 μM) or control vehicle for a further 48 hour evaluation of changes in DPPIV surface protein expression and changes in susceptibility to HCoV EMC

infection. In a first experiment, adenosine down regulates DPP4 expression and reduces the cells' susceptibility to virus infection and reduces virus replication.

In a second experiment, inhibition of HCoV EMC replication in Huh7 cells by soluble adenosine deaminase (ADA) was demonstrated where inhibition with ACE2 was negative. Huh7 cells were incubated with different concentrations of recombinant soluble ADA or recombinant soluble ACE2. After 1 hour incubation, the cells were infected with HCoV EMC at a multiplicity of infection of 0.01. After 8 hours, cells were fixed and stained with a rabbit antiserum against HCoV EMC nsp4 and cells were counted. Results are shown as number of infected cells per well. Infection of Huh7 cells is inhibited by recombinant soluble ADA but not by recombinant soluble ACE2. The results are shown in figure 27.

In a third experiment, inhibition of HCoV EMC replication in Huh7 cells by soluble DPP4 was demonstrated. Different concentrations of recombinant soluble DPP4 or recombinant soluble ACE2 were incubated with HCoV EMC for 1 hour at 37°C and used to infect Huh7 cells. After 8 hours, cells were fixed and stained with a rabbit antiserum against HCoV EMC nsp4 and cells were counted. Results are shown as number of infected cells per well. Infection of Huh7 cells is inhibited by recombinant soluble DPP4 but not by recombinant soluble ACE2. The results are shown in figure 28.

Example 21

The spike (S) protein of the recently emerged human coronavirus (MERS CoV) mediates infection by binding to the cellular receptor dipeptidyl peptidase 4 (DPP4). Here, we mapped the receptor binding domain in the S protein to a 231 amino acid fragment (residues 358 588) by evaluating the interaction of spike truncation variants with receptor expressing cells and soluble DPP4.

Antibodies to this domain much less so to the preceding N terminal region efficiently neutralize MERS CoV infection. It is herein now also shown by co immunoprecipitation and FACS analyses that an internal region of the S1 of hCoV EMC consisting of 231 amino acids is sufficient to bind its receptor, DPP4. It was also shown that the region elicits the most neutralizing antibodies against the virus. Those results identified the receptor binding region of the S protein by convincing methods and the region contains major neuralization epitopes.

Additionally, the inventors herein further map the receptor binding domain (RBD) in the spike protein of the novel coronavirus EMC (hCoV EMC, now MERS CoV). Based on data obtained with bioinformatic tools they designed truncation variants of the S1 portion of hCoV EMC S (EMC S) and showed that the S1 variant harboring residues 358 588 i) co purifies with recombinant CD26 (the hCoV EMC receptor), binds to cellular CD26 in a FACS based assay and elicits neutralizing antibodies in immunized rabbits with higher efficiency than the wt S1 subunit.

Just 10 years following the outbreak of the severe respiratory acute syndrome coronavirus (SARS CoV), the world is confronted with yet another deadly human coronavirus. The virus, first provisionally called human coronavirus EMC (hCoV EMC) but now named MERS CoV, referring to its emergence in the Middle East and to the respiratory syndrome it causes, belongs to the betacoronavirus genus lineage 2c. It has thus far been identified in over 50 patients from or linked to the Arabian Peninsula, approximately half of them being fatal. Like with SARS CoV, patients affected by MERS CoV suffer from severe and often lethal lower respiratory tract infection. The epidemiology of MERS CoV is still enigmatic, but the geographical distribution of epidemiologically unlinked individuals points to intermittent, zoonotic transmission from a – so far unknown animal source, whereas a number of reported clusters indicate limited human to human spread.

The main determinant of coronavirus tropism is the viral spike (S) protein as it mediates binding to a cell surface receptor. The MERS CoV S protein, a 1353 amino acid type I membrane glycoprotein, assembles into trimers that constitute the spikes or peplomers on the surface of the enveloped coronavirus particle. The

protein combines the two essential entry functions, namely that of host receptor binding and membrane fusion, which are attributed to the N terminal (S1, residues 1 751) and C terminal (S2, residues 752 1353) half of the S protein, respectively. Recently, we have identified dipeptidyl peptidase 4 (DPP4, also known as CD26), expressed in the human lung, as a functional receptor for MERS CoV. Importantly, MERS CoV can also use the evolutionary conserved DPP4 of other species, most notably that of bats.

Coronaviruses bind to receptors via independently folded, generally about 150 300 residues long, receptor binding domains (RBD) present in their S1 subunit, of which the location within S1 can vary. Thus, for the betacoronavirus mouse hepatitis virus (MHV), the binding to its CEACAM receptor has been mapped to the N terminal ~300 amino acids of the spike protein whereas for the SARS CoV of the same genus binding to the ACE2 receptor maps to residues 323 502 of S1. Identification of the RBD can hence help the development of monoclonal antibodies and vaccines for the treatment and prevention of infection. The RBD is the most important target for neutralizing antibodies preventing virus receptor interaction.

We previously used the S1 domain of MERS CoV fused to the Fc region of human IgG to demonstrate the interaction of S1 with DPP4 expressing cells and with soluble, i.e., non membrane anchored DPP4. To identify the receptor binding domain in the MERS CoV S1 subunit, we generated S1 Fc protein chimeras with truncations at the C terminus and N terminus of the S1 domain. We considered a three domain structure of the MERS CoV S1 protein (residues 1 357, 358 588 and 589 747) based on the predicted location and structure of the RBD of two other betacoronaviruses, MHV and SARS CoV, of which the homologous regions for MERS CoV S map to the residues 18 351 and 379 580, respectively. In addition, a soluble form of human DPP4 (residues 39 766) was made, which was C terminally tagged with the Fc region. These proteins were expressed in HEK 293T cells after transfection of the expression plasmids and subsequently affinity purified from the cell culture supernatant using protein A sepharose beads as described. The Fc region of purified sDPP4 Fc was proteolytically removed using trypsin (data not shown). First, we analyzed the S1 Fc proteins and C terminal S1 truncations thereof for their ability to interact with sDPP4 using a co purification assay. sDPP4 was efficiently co purified by the S1 Fc variants encompassing residues 1 588 and 1 747, whereas the 1 357 S1 Fc variant was unable to bind sDPP4. We next generated an S1 Fc variant comprising residues 358 588, a region homologous to the ACE2 receptor binding domain in SARS CoV S1. This S1 Fc truncation variant efficiently bound soluble DPP4, indicating that the DPP4 receptor binding domain is located within the 358 588 residues domain of the MERS CoV spike protein.

We subsequently tested the ability of these S1 Fc variants to bind to HEK 293T cells transiently expressing DPP4 by using flow cytometry. The S1 Fc variants encompassing residues 1 588 and 358 588 bound to DPP4 expressing HEK 293T cells with efficiencies comparable to the full length S1 protein, whereas no binding was observed with the 1 357 S1 Fc variant. These data show the 358 588 amino acids S1 region to be essential and sufficient for binding to DPP4 expressing cells, consistent with the results of the sDPP4 interaction study.

Finally polyclonal antibodies were raised in rabbits against the 1 747, 1 357 and 358 588 S1 Fc variants (Davids Biotechnology GmbH, Germany). The sera, which displayed equal ELISA titers towards its antigen (1:300,000, data not shown), were tested for their ability to neutralize virus infectivity. Antibodies raised against the 358 588 S1 Fc variant efficiently neutralized virus infectivity, superior to those raised against the 1 747 and 1 357 S1 Fc variants. This indicates that neutralizing epitopes within S1 are primarily localized to the RBD region. The elicited antibodies are likely to block the interaction of the spike protein with DPP4 thereby neutralizing MERS CoV infectivity. The results demonstrate the preferred potential of S1 protein and of the 358 588 S1 polypeptide or functional fragments thereof reactive with the MERS CoV neutralizing antibody for use as subunit vaccines with a high biosafety profile compared to vaccines based on inactivated viruses or live attenuated virus.

Except for the betacoronavirus MHV, which binds to its CEACAM receptor through a domain in the N terminal part of its S1 protein, the RBDs of all other coronaviruses that engage protein receptors and that have been mapped occur in the C terminal portion of the S1 subunit. Examples also include the alphacoronaviruses binding to ACE2 (hCoV NL63) and APN (e.g., TGEV, hCoV 229E). In this study, we have experimentally mapped the RBD of MERS CoV to a 231 amino acid fragment (residues 358 588) within the spike protein. This domain nicely corresponds with the S1 region recently anticipated to interact with the DPP4 receptor on the basis of theoretical S1 structure predictions. The RBD in the MERS CoV S1 protein localizes in the same region where the SARS CoV S protein interacts with its ACE2 receptor. The SARS CoV RBD structure displays a five stranded ß sheet core structure (ß1 4 and ß7) maintaining the overall domain conformation, and a long extended loop containing two anti parallel ß sheets (ß5 and ß6) responsible for receptor binding{0}. Intriguingly, compared to SARS CoV, the RBD of MERS CoV contains a relatively conserved core domain but a highly variable loop region, tentatively explaining the differential receptor usage. Crystallization and structure analysis of this MERS CoV RBD region in complex with DPP4 will give detailed insight into the virus receptor binding interface.

Example 22

Dipeptidyl peptidase 4 receptor determinants of respiratory MERS-coronavirus infection

Here we show that MERS coronavirus (MERS-CoV) replicates in cells of different species using dipeptidyl peptidase 4 (DPP4) as a functional receptor. This suggests a broad host species tropism allowing zoonotic transmission from many animal species. Here we show contrasting DPP4 receptor functionality in different animal species. Resistance of ferrets to MERS-CoV infection was due to the inability to bind MERS-CoV as a result of amino acid variation in the ferret DPP4 ß-propeller region. In contrast, DPP4 expressing respiratory epithelial cells in the lower - but not upper - respiratory tract of cynomolgus macaques were targeted by MERS-CoV, which resulted in relatively mild disease. Variable DPP4 expression and adenosine deaminase (ADA) - shown to act as a natural antagonist for MERS-CoV infection - may potentially modulate MERS-CoV infection. Our findings illuminate the role of DPP4 sequence and expression variability in host range restriction and outcome of respiratory MERS-CoV infection and lead us to conclude that MERS-CoV receptor sequence and expression variability determine host range restriction of lower respiratory MERS-CoV infection.

Coronaviruses (CoVs) usually cause common colds in humans but zoonotic transmission occasionally introduces more pathogenic viruses into the human population as was demonstrated by the severe acute respiratory syndrome (SARS) outbreak. In 2012 a previously unknown human coronavirus (CoV), now named Middle East respiratory syndrome CoV (MERS-CoV), was isolated from the sputum of a 60-year-old man in Saudi Arabia who presented with acute pneumonia with a fatal outcome. To date, several infection clusters have been reported over a one-year period with around 50% of the reported human cases being fatal. Although limited human-to-human transmission has been observed, it is feared that by acquiring additional mutations MERS-CoV may spread more easily.

MERS-CoV represents a novel betacoronavirus species with the closest known relatives being clade 2c bat CoVs, detected in diverse species of bats but not yet in any animal species from the Arabian Peninsula. Although MERS-CoV replicates in cells of different species including bats, pigs and (non-) human primates, its ability to infect different animal species may be restricted given the fact that hamsters were shown to resist MERS-CoV infection. Therefore, a further understanding of factors that determine host restriction and viral transmission need to be revealed by studies in different animal species.

Herein we identified dipeptidyl peptidase 4 (DPP4) as a functional MERS-CoV receptor in human and bat cells. To further analyse DPP4 usage by MERS-CoV in vivo, ferrets (n = 4), known to be susceptible to several respiratory viruses including SARS-CoV and influenza viruses, were inoculated intratracheally with

MERS-CoV. The animals did not seroconvert and only low levels of virus were detected by RT-qPCR in respiratory swabs at 1-2 days post infection (dpi). In vitro, ferret primary kidney cells could not be infected with MERS-CoV despite DPP4 surface expression, while transfection of these cells with human DPP4 (hDPP4) rendered the cells susceptible, suggesting that ferret DPP4 (fDPP4) does not efficiently bind MERS-CoV. Consistently, MDCK cells transfected with fDPP4 did not bind to synthetic MERS-CoV spike (S1) protein and were not infected by the virus (Fig 33B,C). DPP4 is an ectoenzyme that cleaves dipeptides from hormones, chemokines and cytokines by its conserved C-terminal α/ß-hydrolase domain of the protein, while its N-terminal eight-blade ß-propeller domain contains more sequence variability. By constructing DPP4 chimeras we observed that the blades 4 and 5 containing hDPP4 domain (residues 246-505) could confer to ferret DPP4 the ability to bind to S1 and to mediate MERS-CoV infection when expressed in non-susceptible cells(Fig33B,C). A Quick Change site-directed mutagenesis kit (Stratagene) was used to construct different hDPP4 point mutants. The presence of the correct mutations and absence of undesired mutations was confirmed by sequencing analysis. Plasmids were transfected into MDCK cells in triplicate, after 24 h incubation individual wells were split to determine DPP4 cell surface expression, S1-binding and susceptibility to MERS-CoV infection on the same transfected cell culture. Consistently, substitution of selected solvent exposed residues present in blades 4 and 5 of hDPP4 by those occurring at these positions in fDPP4, abrogated DPP4's capacity to bind to S1 and to mediate MERS-CoV cell susceptibility upon transfection, suggesting that these residues are involved in MERS-CoV binding and entry (Fig33D,E). Reciprocal substitutions of these amino acids in fDPP4 however, did not confer S1 binding, demonstrating the complexity of the interaction in the face the highly polymorphic nature of these two blades. The identified residues also are critical in binding the human enzyme adenosine deaminase (ADA), a natural DPP4 ligand that is involved in the development and maintenance of the immune system. Using recombinant ADA, significant inhibition of MERS-CoV infection and spike protein binding was demonstrated revealing a natural occurring antagonist able to block MERS-CoV infection. The data on the co-crystallization of the receptor binding domain of S1 and DPP4 are in line with the data presented. Phylogenetic analysis of the virus binding region of DPP4 indicated that human, macaque, horse and rabbit DPP4 cluster together as do DPP4's from cattle, pig and bats, that are somewhat more distantly related. Small animals including ferret, mice and most likely hamsters, shown to resist MERS-CoV infection, are more divergent in the DPP4 virus binding region, which at least in the case of ferrets has consequences for MERS-CoV binding.

Considering the highly conserved virus binding region in macaque DPP4 as compared to hDPP4, we first confirmed the use of cynomolgus macaque DPP4 as a functional MERS-CoV receptor. DPP4 antibodies blocked MERS-CoV infection of macaque primary kidney cells in vitro. Besides macaques, rabbits may be a potential animal model for MERS-CoV infection; ex vivo inoculation of rabbit lung and kidney tissues revealed susceptibility to MERS-CoV. We subsequently inoculated ten young adult cynomolgus macaques intratracheally with MERS-CoV and euthanized them at 1 (n = 4, macaques 1-4), 4 (n = 4, macaques 5-8) and 28 dpi (n = 2, macaques 9 and 10). All animals remained free of severe clinical signs and maintained a rhythmic pattern of body temperatures fluctuating between 35°C (night) and 39°C (day) that seemed slightly elevated after inoculation. Neutralizing antibodies with titers 40-80 were detected in the two MERS-CoV infected macaques that were euthanized at 28 dpi. Upon necropsy, there were a few mild focal red-grey slightly depressed areas affecting less than 5 % of the lung tissue, although one lobe of macaque 7 had a dark red rim with evidence of suppurative bronchopneumonia, consistent with the detection of Escherichia coli bacteria in this lobe. MERS-CoV mRNA was detected at highly variable levels in pharyngeal and nasal swabs on 1 to 11 dpi and at low levels in rectal swabs on 2 and 3 dpi. In addition, MERS-CoV was detected by RT-qPCR in the lungs, nasal septum, serum and spleen and in one animal - macaque 1 - also in the kidney, liver, colon and urine at 1 dpi. Infectious virus was detected only in one pharyngeal swab sample and to a limited extent in the lungs. Using a probe that targets the MERS-CoV nucleocapsid gene, hybridization was observed in epithelial cells in bronchioles, and in moderate numbers of type

2 and few type 1 pneumocyte-resembling cells in the alveoli at 1 dpi while at 4 dpi very few cells were found positive. Consistent with activation of cytokines like CCL3, the lungs showed mild alveolitis, characterized by thickening of the alveolar septa with infiltration of few neutrophils and macrophages and moderate type 2 hypertrophy and hyperplasia at 4 dpi. In the alveolar lumina there were increased numbers of alveolar macrophages and occasionally small amounts of edematous fluid with fibrin and few neutrophils. Consistent with the capacity of the virus to induce syncytia in vitro, syncytial cells were seen. By applying a technique that enables successive staining of the same tissue section, tropism of MERS-CoV for cells expressing DPP4 in vivo was demonstrated. Thus, the experimental infection of young adult macaques with MERS-CoV revealed that macaque DPP4 positive cells in the lower respiratory tract can be infected with MERS-CoV but the associated pathological changes are relatively mild, indicating that young adult macaques are at best a suboptimal MERS-CoV animal model for the often fatal MERS-CoV infection in humans.

Abundant ACE2 expression in the respiratory tract has been suggested to facilitate rapid spread of SARS-CoV, a critical factor in the rapid induction of innate immune responses that drive the acute respiratory distress syndrome. In non-infected macaques DPP4 expression was restricted to non-ciliated cells, type 2 cells and endothelial cells whereas no staining was observed in ciliated epithelial cells of the (upper) respiratory tract. The absence of DPP4 on the upper respiratory tract epithelial cells, consistent with the inability to detect viral antigen in these cells, therefore may limit efficient virus transmission through the upper respiratory route. Kidneys, liver, intestine, and sub mucosal glands of the upper respiratory tract were found to contain varying levels of DPP4, which mainly localized to the apical side of the cells. Enhanced DPP4 expression was observed in the lungs of the bacterial co-infected macaque 7, which excreted infectious virus in the pharyngeal swab and displayed a higher body temperature. We observed that LPS stimulation of in vitro differentiated macrophages enhanced DPP4 expression. Attempts to infect these cells were unsuccessful, likely due to ADA production by these cells. Interestingly, DPP4 was virtually absent in the lower respiratory tract epithelium of ferrets but could be visualized in the kidneys of these animals. Contrastingly, relatively strong DPP4 expression was observed on different cell types in human lungs, including a MERS-CoV infected individual. In several pathological conditions such as viral infections and type 2 diabetes increased levels of (soluble) DPP4 have been demonstrated. Thus, relatively low levels of DPP4 expression in the lungs of young adult macaques could partly explain the mild infection observed after MERS-CoV infection but further studies need to reveal the role of varying DPP4 and ADA expression levels in regulating MERS-CoV replication in vivo.

Our findings demonstrate that the host range potential of the emerging novel human MERS-CoV is primarily determined by the MERS-CoV binding to and tissue distribution of DPP4. The co-localisation of DPP4 with MERS-CoV in the lower respiratory tract of MERS-CoV infected non-human primates (in bronchioles and alveoli), and the inability to infect ferrets further supports the sole involvement of DPP4 as a functional receptor in MERS-CoV entry. Variable levels of DPP4 expression in the lower respiratory tract may impose MERS-CoV host range restriction and explain why studies in rhesus macaques have not been successful to reproduce the severe disease seen in humans. Future studies need to unravel the significance of variable DPP4 expression in MERS-CoV patients, for example as a result of co morbidities like microbial infections, type 2 diabetes or aging.

Material and Methods

Cloning of human and ferret DPP4. The hDPP4 cDNA was obtained as described. Total RNA was isolated from ferret primary kidney cells using RNeasy mini kit (Qiagen) and cDNAs were synthesized by using the Superscript reverse transcriptase (Life Technologies). The complete DPP4 genes were amplified using Pfu Ultra II fusion HS DNA polymerase (Stratagene) and cloned into the pcDNA 3.1 expression vector (Life Technologies). Human to ferret DPP4 mutants of cDNA constructs were made by utilizing unique restriction enzyme sites shared by human and ferret DPP4. Pst I can cut human and ferret DPP4 into three fragments (human, amino acid 1-246, 247-504 and 505-766 and ferret, amino acid 1-245, 246-503 and 504-765). The middle

fragment of human and ferret DPP4 was exchanged between human and ferret, the final plasmid constructs contained different combinations of fragments: human-ferret-human (HFH) or ferret-human-ferret (FHF). A Quick Change site-directed mutagenesis kit (Stratagene) was used to construct different hDPP4 point mutants. The presence of the correct mutations and absence of undesired mutations was confirmed by sequencing analysis. Plasmids were transfected into MDCK cells in triplicate, after 24 h incubation individual wells were split to determine DPP4 cell surface expression, S1-binding and susceptibility to MERS-CoV infection on the same transfected cell culture. S1 binding and infection were corrected for DPP4 cell surface expression as determined by the goat polyclonal antiserum against DPP4 (R&D systems), a secondary FITC conjugated rabbit anti goat serum followed by FACS analysis.

Phylogenetic analysis of DPP4. Sequence alignment was performed by using ClustalW in the MEGA5.0 software package (www.megasoftware.net), and the trees were constructed by using the neighbor-joining method with p-distance (gap/missing data treatment; complete deletion) and 1,000 bootstrap replicates as in MEGA version 5.0.

Protein expression and S1 binding assay. A plasmid encoding MERS-CoV S1-Fc was generated by ligating a fragment encoding the S1 domain (residues 1-747) 3'-terminally to a fragment encoding the Fc domain of human IgG into the pCAGGS expression vector. Likewise, an S1-Fc expression plasmid was made the FIPV S1 domain (isolate 79-1146; residues 1-788). Fc chimeric proteins were expressed by transfection of the expression plasmids into HEK-293T cells and affinity purified from the culture supernatant using Protein A Sepharose beads (GE Healthcare). S1 binding of cells was measured by incubating 105 cells with 15 mg/ml of S1–Fc followed by incubation with FITC or DyLight-488-labelled goat-anti-human IgG antibody and analysis by flow cytometry.

Virus infection experiments. Virus stocks of MERS-CoV (EMC isolate) were prepared. Transfected COS-7 cells, Huh-7 and primary ferret and macaque kidney cells were inoculated with MERS-CoV for 1 h with high MOI. After washing the cells were incubated with medium containing 1% fetal bovine serum. Alternatively we used thin cut slices from the lungs and kidneys of rabbits that were incubated in culture medium with virus for 24h. At 8 or 24 h after infection cells were fixed with formaldehyde and stained using rabbit-anti-SARS-CoV NSP4 antibodies that are cross-reactive for hCoV-EMC, according to standard protocols using a FITC conjugated swine-anti-rabbit antibody as a second step. Primary ferret or macaque kidney cells were pre-incubated with antibodies to DPP4 (polyclonal goat-anti DPP4 immunoglobulin, R&D systems) at 20 μg/ml to block MERS-CoV infection. Recombinant human ADA (R&D systems) was preincubated with hDPP4 transfected cells or Huh7 cells for 1 h after which the cells were infected with MERS-CoV for 8 h and processed.

Animal studies. Ten cynomolgus macaques (Macaca fascicularis), 3-5 years old with active temperature transponders in the peritoneal cavity (n = 3), were inoculated with 5 × 106 TCID50 of MERS-CoV via the intranasal and intratracheal route. In addition, four ferrets (Mustello fuoris) were inoculated with 1 × 106 TCID50 of MERS-CoV via the intranasal and intratracheal route. Animals were checked daily for clinical signs. Just before infection and at different dpi, animals were anesthetized with ketamine and nasal, pharyngeal, and rectal swabs were taken, which were placed in 1 ml Dulbecco's modified Eagle's medium supplemented with 100 IU penicillin/ml and 100 □g of streptomycin/ml (virus transport medium) and frozen at -70°C until RT-PCR analysis. The animals were euthanized at different days (Day 1, 4 or 28) p.i. by exsanguination under ketamine anesthesia. Approval for animal experiments was obtained from the Institutional Animal Welfare Committee (nr EMC 2808).

Necropsies were performed according to a standard protocol. For semi-quantitative assessment of gross pathology, the percentage of affected lung tissue from each lung lobe was determined at necropsy, recorded on a schematic diagram of the lung and the area of affected lung tissue was subsequently calculated (gross pathology score). One lung of each monkey was inflated with 10% neutral-buffered formalin by intrabronchial

intubation and suspended in 10% neutral-buffered formalin overnight. Samples were collected in a standard manner (from the cranial, medial and caudal parts of the lung), embedded in paraffin, cut at 3 □m and used for immunohistochemistry (see below) or stained with hematoxylin and eosin (H&E). The lung, liver, spleen, kidney, intestine, trachea, and tracheobronchial lymphnode H&E sections were examined by light microscopy.

In situ hybridization. The ISH probes targeting the nucleocapsid gene of MERS-CoV were designed by Advanced Cell Diagnostics (Hayward, CA) and ISH was performed according to the manufacturer's instructions and ISH staining was visualized using substrate Fast Red (pink). Controls included probes against SARS-CoV nucleocapsid protein and tissues from non infected animals.

Imunohistochemistry. Family consent was granted for limited postmortem tissue retrieval from a MERS-CoV patient in the UK, consisting of a 20-cm-long midline incision in lower chest and upper abdomen, from which tissue samples were collected from both lungs. Archival paraffin-embedded human tissue sections were obtained from the Department of Pathology, Erasmus MC. Four historic macaque controls were used as mock (PBS) infected. For histological analysis, samples were placed in 10% neutral-buffered formalin and further processed for routine immunohistochemistry. Serial 3 µm lung sections were stained using according to standard protocols using antibodies to DPP4 (polyclonal goat-anti DPP4 immunoglobulin, R&D systems. For phenotyping to test DDP4 expression of MERS-CoV infected cells, we used a destaining-restaining technique. Briefly, the precipitate used for visualization of MERS-CoV antigen staining was dissolved in graded 100%-70% alcohols. To detach the primary antibody and red immunohistochemistry signals, slides were soaked in eluding buffer (5ml 0.1M HCl, 95 ml 0.1M NaCl containing 1M glycine) for 2 hours. The sections were treated with two 5 min intervals heating in citric acid buffer pH 6.0 to denature any undetached primary antibody. The slides were then incubated with antibodies against DPP4 in PBS/0.1% BSA for 1 hour at RT. After washing, sections were incubated with horseradishperoxidase labeled anti-goat IgG 1/100 in PBS/0.1% BSA for 1 hour at RT. Peroxidase activity was revealed by incubating slides in 3,3'–diaminobenzidine-tetrachlorhydrate (DAB) (Sigma) for 3-5 minutes, resulting in a brown precipitate, followed by counterstaining with hematoxylin.

RNA-extraction and quantitative RT-PCR. Samples were analysed with the upE PCR and confirmed by a nucleocapsid specific PCR. RNA from 200 □l of culture supernatant was isolated with the Magnapure LC total nucleic acid isolation kit (Roche) and eluted in 100 □l. MERS-CoV RNA was quantified on the ABI prism 7700, with the TaqMan® Fast Virus 1-Step Master Mix (Applied Biosystems) using 20 □l isolated RNA, 1× Taqman mix, 0.5U uracil-N-glycosylase, 45 pmol forward primer (5'-GGGTGTACCTCTTAATGCCAATTC-3'), 45 pmol reverse primer (5'-TCTGTCCTGTCTCCGCCAAT-3') and 5 pmol probe (5'-FAM-ACCCCTGCGCAAAATGCTGGG-BHQ1-3') . Amplification parameters were 5 min at 50°C, 20 sec at 95°C, and 45 cycles of 3 s at 95°C, and 30 sec at 60°C. RNA dilutions isolated from an MERS-CoV stock were used as a standard.

Lung tissue samples (0.3-0.5 gram) were taken for RT-PCR and microarray analysis in RNA-later (Ambion, Inc.). RNA was isolated from homogenized post mortem tissue samples using Trizol Reagent (Invitrogen) and the RNeasy mini kit (Qiagen). cDNA synthesis was performed with ~1 □g total RNA and Superscript III RT (Invitrogen) with oligo(dT), according to the manufacturer's instructions. Semi-quantitative RT-PCR was performed as described previously to detect MERS-CoV and to validate cellular gene expression changes as detected with microarrays of CCL3 (Applied Biosystems). Differences in gene expression are represented as the fold change in gene expression relative to a calibrator and normalized to a reference. GAPDH (glyceraldehydes-3-phosphate dehydrogenase) was used as endogenous control to normalize quantification of the target gene. The samples from the mock-infected macaques were used as a calibrator. Average results (□ s.e.m.) for groups were expressed as fold change compared to PBS-infected animals.

Macrophage cultures. Monocytes isolated from peripheral blood mononuclear cells were cultured with GM-CSF for 5 days to generate macrophages. Subsequently cells were stimulated with LPS at 1µg/ml for 24h and processed for DPP4 staining and FACS analysis.

Statistical analysis. Data were compared using one way ANOVA with post-test Bonferroni. Statistical analysis was performed with Prism 4.0 (Graphpad).

References to example 22

1. J. S. Peiris, Y. Guan, K. Y. Yuen, *Nat. Med.* **10**, S88-97 (2004).
2. A. M. Zaki, S. van Boheemen, T. M. Bestebroer, A. D. Osterhaus, R. A. Fouchier, N. *Engl. J. Med.* **367**, 1814-1820 (2012).
3. S. van Boheemen *et al., MBio* **3**, 00473-12 (2012).
4. World Health Organisation. Middle East respiratory syndrome coronavirus (MERS-CoV)-update, http://www.who.int/csr/don/2013_08_01/en/index.html (8 August 2013).
5. A. Assiri et al., *N. Engl. J. Med.* **369**, 407-416 (2013).
6. B. Guery *et al., Lancet* **381**, 2265-2272 (2013).
7. C. B. Reusken *et al., Vector Borne Zoonotic Dis* **10**, 785-791 (2010).
8. A. Annan *et al., Emerg Infect Dis* **19**, 456-459 (2013).
9. M. A. Muller *et al., MBio* **3**, 00515-12 (2012).
10. J. F. Chan *et al., J. Infect. Dis.* **207**, 1743-1752 (2013).
11. E. de Wit *et al., PLoS One* **8,** e69127 (2013).
12. V. S. Raj *et al., Nature* **495**, 251-254 (2013).
13. B. E. Martina *et al., Nature* **425**, 915 (2003).
14. J. A. Belser, J. M. Katz, T. M. Tumpey, *Dis. Model. Mech.* **4**, 575-579 (2011).
15. E. Boonacker, C. J. Van Noorden, *Eur. J. Cell. Biol.* **82**, 53-73 (2003).
16. W. A. Weihofen, J. Liu, W. Reutter, W. Saenger, H. Fan, *J. Biol. Chem.* **279**, 43330-43335 (2004).
17 R.P. Dong *et al, J. Immunol.* **156**, 1349-1355 (1996).
18. G. Lu *et al., Nature* doi:nature12328 (2013).
19. N. Wang *et al., Cell.* Res. doi:cr201392 (2013).
20. B. Rockx *et al., J. Virol.* **83**, 7062-7074 (2009).
21. S. L. Smits *et al., PLoS Pathog.* **6**, e1000756 (2010).
22. J. Zhong *et al., Diabetes* **62**, 149-157 (2013).
23. B. A. Conlon, W. R. Law, *Clin. Exp. Immunol.* **138**, 14-20 (2004).
24. S. A. Lee et al., J. *Clin. Endocrinol. Metab.* **98**, 2553-2561 (2013).
25. T. Andrieu *et al., J. Clin. Virol.* **27**, 59-68 (2003).
26. V. J. Munster, E. de Wit, H. Feldmann, *N. Engl. J. Med.* **368**, 1560-1562 (2013).
27. S. L. Smits *et al., J. Virol.* **85**, 4234-4245 (2011).
28. V. M. Corman *et al., Euro. Surveil.*l **17**, pii: 20334 (2012).

Claims

1. An essentially mammalian positive-sense single stranded RNA virus, comprising one or more of the nucleic acid or comprising one or more of the amino acid sequences selected from figure 3 depicting the partial open reading frame of HCoV-SA1, figure 5 file N.rtf depicting the nucleocapsid (N) protein, figure 6 file M.rtf depicting the matrix (M) protein, figure 7 file E.rtf depicting the small envelope € protein, figure 8 file NS3d.5 rtf depicting the non-structural gene NS3d, figure 9 file NS3c.rtf depicting the non-structural gene NS3c, figure 10 file NS3b.rtf depicting the non-structural gene NS3b, figure 11 file NS3a.rtf depicting the non-structural gene NS3a, figure 12, file S.rtf depicting the spike surface glycoprotein (S), figure 13 file Orf1ab.rtf, encoding many enzymatic products among which the replicase, figure 14 file HCoV-SA1.rtf depicting isolate HCoV-SA1 or figure 15 HCoVSA1. rtf depicting its 3 translation frames, or comprising amino acid sequences having at least 90% identical with a nucleic acid or amino acid sequence depicted in said figure 3 or figure 5 to 15.

2. A virus according to claim 1 isolatable from humans.

3. A virus according to any one of claims 1 to 2 having an amino acid sequence of its conserved replicase domain that is at least 90%, most preferably at least 95% identical with the amino acid sequence of the conserved replicase domain comprising ORF 1ab.

4. A virus according to any of claims 1 to 3 belonging to the Coronaviruses, genus Betacoronavirus and identifiable as phylogenetically corresponding thereto by determining the amino acid sequence of the conserved replicase domain of said virus to have at least 90% identity with the Orf1AB amino acid sequence as depicted in Fig. 13.

5. A nucleic acid, preferably a cDNA, obtainable from a virus according to anyone of claims 1 to 4, said nucleic acid having at least an identity of 90% with a nucleic acid encoding a protein selected from any of the figure 3 or 5 - 15.

6. A vector comprising a nucleic acid according to claim 5.

7. A cell comprising a virus according to anyone of claims 1 to 4, a nucleic acid according to claim 5 or a vector according to claim 6.

8. A protein encoded by a nucleic acid according to claim 5.

9. An antigen comprising a protein according to claim 8.

10. An antibody specifically directed against an antigen according to claim 9.

11. A method for identifying a viral isolate as a MERS-CoV comprising reacting said viral isolate with a nucleic acid according to claim 5 and/or with an antibody according to claim 10.

12. A method for virologically diagnosing a MERS-CoV infection of a mammal comprising determining in a sample of said mammal the presence of a viral isolate by reacting said sample with a nucleic acid according to claim 5 or an antibody according to claim 10.

13. A method for serologically diagnosing a MERS-CoV infection of a mammal comprising determining in a sample of said mammal the presence of an antibody specifically directed against a Betacoronavirus, preferably Lineage C virus or component thereof by reacting said sample with a protein or fragment thereof according to claim 8 or an antigen according to claim 9.

14. A diagnostic kit for diagnosing a MERS-CoV infection comprising a virus according to anyone of claims 1 to 4, and/or a nucleic acid according to claim 5, and/or a protein according to claim 8, and/or an antigen according to claim 9 and/or an antibody according to claim 10.

15. A virus according to any one claims 1 to 4, and/or a nucleic acid according to claim 5, and/or a vector according to claim 6, and/or a cell according to claim 7, and/or a protein or fragment thereof according to claim 8, and/or an antigen according to claim 9 and/or or an antibody according to claim 10 for use in the treatment or prevention of an infection with a MERS-Cov.

16. A pharmaceutical composition comprising a virus according to any one claims 1 to 4, and/or a nucleic acid according to claim 5, and/or a vector according to claim 6, and/or a cell according to claim 7, and/or a protein or fragment thereof according to claim 8, and/or an antigen according to claim 9, and/or or an antibody according to claim 10.

17. A pharmaceutical composition according to claim 16 for use in the treatment or prevention of a MERS-CoV infection in a mammal, and/or in the treatment or prevention of atypical pneumonia and/or renal failure caused by such a MERS-CoV infection in a mammal.

18. A pharmaceutical composition for use according to claim 17, wherein said mammal is a human.

19. A proteinaceous substance, having been provided with a protein according to claim 8 and additionally comprising an N-terminal dipeptidyl peptidase protein.

20. A proteinaceous substance according to claim 19, wherein said substance is crystallized.

21. A proteinaceous substance according to claim 19 or 20, wherein said dipeptidyl peptidase protein is a dipeptidyl peptidase 4 (DPP4).

22. A proteinaceous substance according to any of claims 19 – 21, wherein said proteinaceous molecule comprises an ectodomain of a spike protein.

23. A proteinaceous substance according to claim 22, wherein said ectodomain is derived from the S1 region of a coronavirus.

24. A container provided with a virus according to any one of claims 1 to 4, and/or a nucleic acid according to claim 5, and/or a vector according to claim 6, and/or a cell according to claim 7, and/or a protein or fragment thereof according to claim 8, and/or an antigen according to claim 9, and/or an antibody according to

claim 10, and/or a pharmaceutical composition according to claim 16 or 17, and/or a substance according to any of claims 19 – 23.

25. A method of identifying a candidate modulator or a candidate antiviral agent as an agent that modulates the function of or the binding of a virus to a dipeptidyl peptidase, said method comprising:

a. Providing a substance according to any of claims 19 to 23 in the presence and absence of said candidate modulator or said candidate antiviral agent under conditions permitting binding of a virus with a peptidase protein.

b. Measuring binding of said virus to said protein, wherein a decrease or increase in binding in the presence of said candidate modulator or said antiviral agent, relative to binding in the absence of said candidate modulator, identifies said candidate modulator as an agent that modulates the function of a dipeptidyl peptidase or identifies said antiviral agent as an agent that modulates that function of a dipeptidyl peptidase, preferably wherein said virus and/or said protein is detectably labeled, preferably with a moiety selected from the group consisting of a radioisotope, a fluorophore, a quencher of flurescence, an enzyme, and an affinity tag.