Acta virologica **58**: 359 – 363, 2014 doi:10.4149/av_2014_04_359

Association of host tropism of Middle East syndrome coronavirus with the amino acid structure of host cell receptor dipeptidyl peptidase 4

M. KANDEEL[1,2], M. A. ELAIZIZ[1], A. KANDEEL[3,4], A. A. ALTAHER[1], Y. KITADE[5]

[1]Department of Physiology, Biochemistry and Pharmacology, King Faisal University, Alhofuf, Alahsa, Saudi Arabia; [2]Department of Pharmacology, Faculty of Veterinary Medicine, Kafrelshikh University, Kafrelshikh 33516, Egypt; [3]Department of Surgery, Faculty of Veterinary Medicine, Zagazig University, Zagazig, Egypt; [4]Department of Biology, College of Science and Arts, Elkamel Branch, King Abdulaziz University, Saudi Arabia; [5]Department of Biomolecular Science, Gifu University, Gifu, Japan

Received January 15, 2014; accepted November 3, 2014

Summary. – The Middle East syndrome coronavirus (MERS-CoV) is a recently emerging betacoronavirus with high fatality. Recently, dipeptidyle peptidase (CD26, DPP4) was identified as the host cell receptor for MERS-CoV. Interestingly, despite of common presence of DPP4 receptors the binding and infection of various cells shows imminent variability. In this report, we provide a tool for prediction of the host tropism of the virus based on the host receptor binding interface. We found out that, in the binding of MERS-CoV to cells the amino acid residues in lancets 4 and 5 of DPP4 receptor, namely K267, Q286, T288, R317, R336, Q344 A291, L294, and I295 are involved. Changes in these residues correspond to profound decrease in virus binding to cells. The nine residues at the interface between the virus spikes and the lancets 4 and 5 of host DPP4 can be used as a predictive tool for the host tropism and virus affinity to host cell receptors.

Keywords: MERS coronavirus; dipeptidyl peptidase 4; amino acid; mutation; virus susceptibility

Introduction

Coronaviruses (CoV) are enveloped single-stranded RNA viruses that infect human and wide variety of animals causing severe respiratory or enteric symptoms (Chang *et al*., 2012; Perlman and Netland, 2009). CoVs that infect humans include human alfa, beta and gamma CoV. Severe acute respiratory syndrome (SARS) has been associated with *Betacornovirus* genus.

CoVs are characterized by high recombination frequencies. The large size of virus genome, unique viral replication, the low fidelity of coronavirus-encoded polymerases and high recombination accounts for unexpected viral evolution of other host infection, changes in clinical signs and resistance to therapy or vaccination. The human SARS-CoV OC43 has evolved from bovine CoV. Furthermore, porcine respiratory CoV has evolved from a gastrointestinal ancestor (Chang *et al*., 2012; Laude *et al*., 1993).

MERS-CoV was initially identified in the Arabian Peninsula in 2012. MERS-CoV was assigned to the *Betacoronavirus* genus. The genome of CoV encodes 4 major structural proteins; nucleocapsid (N protein), spikes (S protein), membrane (M) and small envelope proteins (E). The S protein is a glycoprotein essential for viral attachment to the cell surface receptors. The S protein is cleaved in host cells into S1 and S2 subunits. S1 protein binds the host receptor, while S2 receptor mediates membrane fusion (Wang *et al*., 2013).

The virus replicates in different hosts using DPP4 as a functional receptor (Ohnuma *et al*., 2013). DPP4 is proved to be the only essential receptor for MERS-CoV spikes binding to host cells. Therefore, DPP4 constitutes a unique binding site for MERS-CoV which differs from the binding receptor of SARS-CoV (ACE2, (Wang *et al*., 2013)). Interestingly, the presence of DPP4 receptor in a host cell does not warrant the binding and infection with the MERS-CoV. For instance, MERS-CoV can replicate in cells of humans, pigs, rabbits and non-human primates (Chan *et al*., 2013). In contrast, despite of the expression of DPP4, replication of MERS-CoV was not

E-mail: mahmoud.kandeal@vet.kfs.edu.eg, mkandeel@kfu.edu.sa; phone: +966568918734.

Abbreviations: CoV = coronavirus; DPP4 = dipeptidyl peptidase 4; MERS-CoV = the Middle East syndrome coronavirus; SARS = Severe acute respiratory syndrome

possible in hamsters (de Wit *et al*., 2013) and ferrets (Raj *et al*., 2013). Furthermore, transfection of ferret kidney cells with functional human DPP4 receptors rendered the cells susceptible to infection with MERS-CoV (Raj *et al*., 2013).

In this study, bioinformatics approaches combined with the predetermined critical factors for binding and infection with MERS-CoV were combined to predict the host tropism of the newly emerged MERS-CoV. Here, we identify a group of amino acids at virus-host receptor interface as a marker for efficient binding and subsequent infection with MERS-CoV. The provided predictive tool can be used to predict the host tropism of MERS-CoV in the surrounding environment of an infected region.

Materials and Methods

The full-length DPP4 sequences were retrieved from the nucleotides depository at the National Center for Biotechnology Information (NCBI). The potential protein domain contents of the retrieved sequences were analyzed at the domain search tools at NCBI. BLAST search was used to identify the high homology hits. The retrieved sequences of various species were exported and analyzed for differences in amino acid composition. Multiple sequence alignment was constructed by using Clustal omega tool at the European Bioinformatics Institute (EBI). The output file was retrieved and manually edited by GeneDoc software. The similarity and homology percentage was calculated by Ugene 1.12.2 for mac computer. The output alignments from Clustal omega were examined by Dendroscope software for creation of phylogenetic trees. The phylogenetic tree was created by neighbor-joining method in output formats of radial phylogram or circular cladogram.

Results and Discussion

The bioinformatics of proteins is a useful strategy in determining the biomolecular interactions (Kandeel, 2014; Kandeel and Kitade, 2013a,b), host restrictions of virus infection (Tonnessen *et al*., 2013) and viral evolution (Cui *et al*., 2013; Liu *et al*., 2012). In this report, the bioinformatic tools were adopted to predict the MERS-CoV host tropism. The criteria of prediction were based on: first, the alignment of protein sequences of lancets 4 and 5 of DPP4 as well as the phylogenetic relations with the human DPP4 (Fig. 1a and b). The second is the alignment of amino acid sequences at the interface of interaction (Fig. 1c) between the lancets 4 and 5 of DPP4 and the spike protein of MERS-CoV.

The structure of DPP4 shows N-terminal hydrolase and C-terminal β–propeller domain composed of 8 lancets. Lancets 4 and 5 were found to be the site for binding of the S protein of MERS-CoV. Replacement of the lancets 4 and 5 or mutational changes led to drastic effect on the binding and infection with MERS-CoV.

Phylogentic analysis of lancets 4 and 5 of DPP4 showed that non-human primates (more than 98% homology, Table 1) and rabbits are highly related to the human DPP4 followed by bats and rodents as guinea pig, hamster and rat. The most divergent DPP4 was that of birds and alligator (homology was less than 60%, Table 1).

The binding interface between S protein and DPP4 is composed of polar contacts from hydrophilic residues K267, Q286, T288, R317, R336 and Q344 surrounding a hydrophobic center formed by A291, L294 and I295 (Fig. 1c). Disruption of the mentioned residue interaction resulted in profound decrease of virus entry (Wang *et al*., 2013). In this report, we assume that MERS-CoV binding and replication in a specific host depends on the status of the above mentioned nine residues. The high replication of MERS-CoV in non-human primates coincides with conservation of all of the above mentioned residues. Rabbits and to a lesser extent pigs showed high residue conservation pattern. Therefore, infection with MERS-CoV was possible in cells of these animals. Camelids showed a high conservation profile, indicating a potential incrimination of camels as a host for the virus. In this context, neutralizing antibodies against MERS-CoV were detected in camels from Middle East (Perera *et al*., 2013; Reusken *et al*., 2013). Similar high conservation pattern was evident in the sequences from farm animals as sheep, goats and cattle. Although bats were highly divergent from human DPP4, they showed little changes in the described 9 residues. This clarifies the possible role of bats in the transfer of MERS-CoV. However, an estimated MERS-CoV bat-infection rate was at least 3 folds lower than that of SARS-CoV (Memish *et al*., 2013). Cats and rodents showed amino acid replacements in at least half of the above mentioned residues. This may explain the low viral load in their cell cultures compared with primates (Chan *et al*., 2013). Compared to the human DPP4, birds showed the highest divergence (Fig. 1b). Furthermore, birds showed the greatest changes in the above described marker residues (Fig. 1a). In agreement with our assumption, chicken-derived cell culture did not support the replication of MERS-CoV (Chan *et al*., 2013).

In brief, the resistance to MERS-CoV replication was associated with significant changes in the amino acid residues at the interface of interaction between S protein and lancets 4 and 5 of DPP4. Experimental measures are needed for confirmation of our predictive model. The degree of conservation of the above mentioned residues can be used to predict the host tropism of MERS-CoV. These predictions might be of a value in prevention and control programs, in which the sensitivity and resistance to MERS-CoV infection in the surrounding environment can be anticipated.

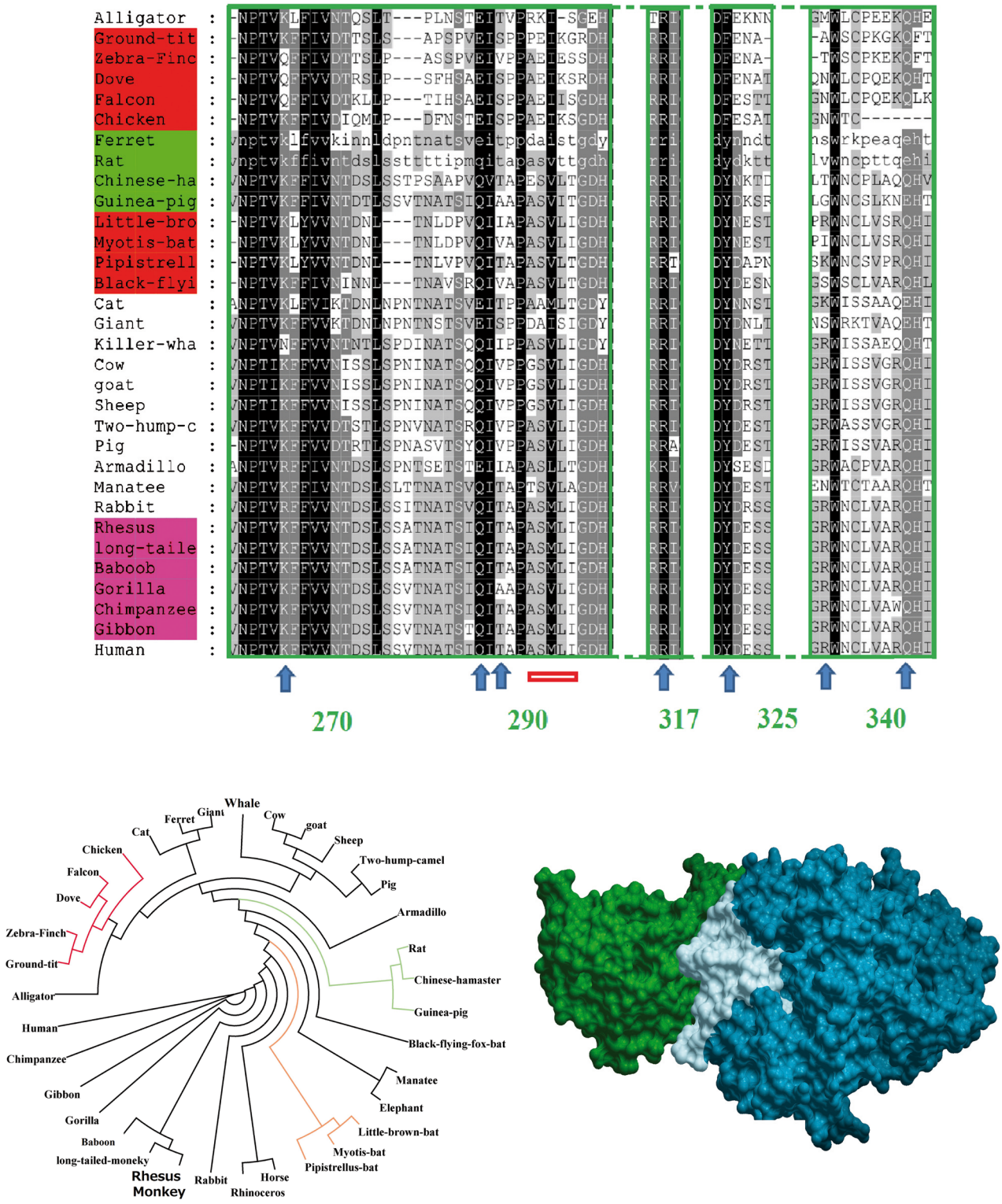

Fig. 1

(a) Amino acid sequence alignment of the lancets 4 and 5 of β-propeller domain of DPP4 in different species, (b) phylogenetic tree of various animal species based on amino acid sequence of DPP4, (c) structure of MERS-CoV complexed with human DPP4

The receptor binding domain of MERS-CoV spike (green), extracellular part of DPP4 (turquoise) and interaction interface (white) are shown.

Table 1. Characteristics of amino acid sequences of DPP4 in different species

Scientific name	Common name	Total score	Query cover	E value	Identity %	Acc. No.
Homo sapiens	Humans	197	100%	4.00E-57	100%	CAA43118.1
Nomascus leucogenys	White-cheeked gibbon	195	100%	2.00E-56	99%	XP_003266219.1
Gorilla gorilla	Gorilla	193	100%	8.00E-56	98%	XP_004032754.1
Macaca mulatta	Rhesus monkey	193	100%	9.00E-56	98%	NP_001034279.1
Macaca fascicularis	Long-tailed Macaque	193	100%	9.00E-56	98%	XP_005573375.1
Papio anubis	Baboon	193	100%	1.00E-55	98%	XP_003907588.1
Pan troglodytes	Chimpanzee	193	100%	1.00E-55	99%	XP_515858.2
Oryctolagus cuniculus	Rabbit	188	100%	5.00E-54	94%	XP_002712206.1
Equus caballus	Horse	179	100%	7.00E-51	88%	XP_005601601.1
Ceratotherium simum	Rhinoceros	175	100%	2.00E-49	87%	XP_004428321.1
Loxodonta Africana	Elephant	174	100%	2.00E-48	85%	XP_003406047.1
Trichechus manatus	See cow	164	100%	2.00E-45	81%	XP_004375482.1
Bos Taurus	Cow	155	100%	6.00E-44	77%	DAA32742.1
Cavia porcellus	Guinea pig	160	100%	1.00E-43	80%	XP_003478612.2
Capra hircus	Goat	155	100%	1.00E-42	77%	XP_005676104.1
Cricetulus griseus	Hamster	154	100%	1.00E-42	73%	EGW01899.1
Myotis lucifugus	Brown bat	156	98%	2.00E-42	78%	XP_006083275.1
Ovis aries	Sheep	156	100%	2.00E-42	77%	XP_004004709.1
Dasypus novemcinctus	Armadillo	154	100%	6.00E-42	76%	XP_004464464.1
Camelus ferus	Camel	153	100%	2.00E-41	75%	XP_006176870.1
Myotis brandtii	Vesper bat	153	98%	3.00E-41	77%	EPQ03437.1
Pipistrellus pipistrellus	Common pipistrelle bat	151	98%	2.00E-40	75%	AGF80256.1
Sus scrofa	Pig	148	98%	2.00E-39	74%	NP_999422.1
Orcinus orca	Killer whale	147	100%	3.00E-39	73%	XP_004283669.1
Felis catus	Cat	146	100%	8.00E-39	71%	NP_001009838.1
Ailuropoda melanoleuca	Panda	144	100%	9.00E-38	70%	XP_002924912.1
Mustela putorius furo	Ferret	130	100%	5.00E-33	63%	ABC72084.1
Columba livia	Pigeon	107	98%	1.00E-24	55%	XP_005498754.1
Falco cherrug	Falcon	105	98%	4.00E-24	51%	XP_005443040.1
Ovophis okinavensis	Pit viper	104	100%	6.00E-24	54%	BAN82157.1
Alligator sinensis	Alligator	100	98%	1.00E-22	52%	XP_006037514.1
Pseudopodoces humilis	Ground tit	100	98%	1.00E-22	51%	XP_005520053.1
Taeniopygia guttata	Zebra finch	99	98%	5.00E-22	50%	XP_004176799.1
Gallus gallus	Fowl	94	82%	4.00E-20	56%	NP_001026426.1

Acknowledgements. This work was supported by the grant No141006 from the deanship of scientific research, King Faisal University to MK.

References

Chan JF, Chan KH, Choi GK, To KK, Tse H, Cai JP, Yeung ML, Cheng VC, Chen H, Che XY, Lau SK, Woo PC, Yuen KY (2013): Differential cell line susceptibility to the emerging novel human betacoronavirus 2c EMC/2012: implications for disease pathogenesis and clinical manifestation. J. Infect. Dis. 207, 1743–1752. http://dx.doi.org/10.1093/infdis/jit123

Chang HW, Egberink HF, Halpin R, Spiro DJ, Rottier PJ (2012): Spike protein fusion peptide and feline coronavirus virulence. Emerg. Infect. Dis. 18, 1089–1095. http://dx.doi.org/10.3201/eid1807.120143

Cui J, Eden JS, Holmes EC, Wang LF (2013): Adaptive evolution of bat dipeptidyl peptidase 4 (dpp4): implications for the origin and emergence of Middle East respiratory syndrome coronavirus. Virol. J. 10, 304. http://dx.doi.org/10.1186/1743-422X-10-304

de Wit E, Prescott J, Baseler L, Bushmaker T, Thomas T, Lackemeyer MG, Martellaro C, Milne-Price S, Haddock E, Haagmans BL, Feldmann H, Munster VJ (2013): The Middle East respiratory syndrome coronavirus (MERS-CoV) does not replicate in Syrian hamsters. PloS One 8, e69127. http://dx.doi.org/10.1371/journal.pone.0069127

Kandeel M (2014): Bioinformatics analysis of the recent MERS-CoV with special reference to the virus-encoded Spike protein. Mol. Enzymol. Drug Target 1, 1–10.

Kandeel M, Kitade Y (2013a): Computational analysis of siRNA recognition by the Ago2 PAZ domain and identification of the determinants of RNA-induced gene silencing. PloS One 8, e57140. http://dx.doi.org/10.1371/journal.pone.0057140

Kandeel M, Kitade Y (2013b): In silico molecular docking analysis of the human Argonaute 2 PAZ domain reveals insights into RNA interference. J. Comput. Aided Mol. Des. 27, 605–614. http://dx.doi.org/10.1007/s10822-013-9665-3

Kandeel M, Noguchi Y, Oh-Hashi K, Kim H-S, Kitade Y (2013): Molecular dynamics and energetic perceptions of substrate recognition by thymidylate kinase. J. Therm. Anal. Cal. 115, 2089–2097. http://dx.doi.org/10.1007/s10973-013-3319-5

Laude H, Van Reeth K, Pensaert M (1993): Porcine respiratory coronavirus: molecular features and virus-host interactions. Vet. Res. 24, 125–150.

Liu Y, Wang J, Ji J, Chang S, Xue C, Ma J, Bi Y, Xie Q (2012): Phylogenetic diversity and genotypic complexity of H1N1 subtype swine influenza viruses isolated in mainland China. Virol. J. 9, 289. http://dx.doi.org/10.1186/1743-422X-9-289

Memish ZA, Mishra N, Olival KJ, Fagbo SF, Kapoor V, Epstein JH, Alhakeem R, Durosinloun A, Al Asmari M, Islam A, Kapoor A, Briese T, Daszak P, Al Rabeeah AA, Lipkin WI (2013): Middle East respiratory syndrome coronavirus in bats, Saudi arabia. Emerg. Infect. Dis. 19, 1819–1823. http://dx.doi.org/10.3201/eid1911.131172

Ohnuma K, Haagmans BL, Hatano R, Raj VS, Mou H, Iwata S, Dang NH, Bosch BJ, Morimoto C (2013): Inhibition of Middle East respiratory syndrome coronavirus infection by anti-CD26 monoclonal antibody. J. Virol. 87, 13892–13899. http://dx.doi.org/10.1128/JVI.02448-13

Perera RA, Wang P, Gomaa MR, El-Shesheny R, Kandeil A, Bagato O, Siu LY, Shehata MM, Kayed AS, Moatasim Y, Li M, Poon LL, Guan Y, Webby RJ, Ali MA, Peiris JS, Kayali G (2013): Seroepidemiology for MERS coronavirus using microneutralisation and pseudoparticle virus neutralisation assays reveal a high prevalence of antibody in dromedary camels in Egypt, June 2013. Euro Surveill. 18, pii=20574.

Perlman S, Netland J (2009): Coronaviruses post-SARS: update on replication and pathogenesis. Nat. Rev. Microbiol. 7, 439–450. http://dx.doi.org/10.1038/nrmicro2147

Raj VS, Smits SL, Provacia LB, van den Brand JM, Wiersma L, Ouwendijk WJ, Bestebroer TM, Spronken MI, van Amerongen G, Rottier PJ, Fouchier RA, Bosch BJ, Osterhaus AD, Haagmans BL (2013): Adenosine deaminase acts as a natural antagonist for dipeptidyl peptidase 4 mediated entry of the Middle East respiratory syndrome coronavirus. J. Virol. 88, 1834–1838. http://dx.doi.org/10.1128/JVI.02935-13

Reusken CB, Haagmans BL, Muller MA, Gutierrez C, Godeke GJ, Meyer B, Muth D, Raj VS, Smits-De Vries L, Corman VM, Drexler JF, Smits SL, El Tahir YE, De Sousa R, van Beek J, Nowotny N, van Maanen K, Hidalgo-Hermoso E, Bosch BJ, Rottier P, Osterhaus A, Gortazar-Schmidt C, Drosten C, Koopmans MP (2013): Middle East respiratory syndrome coronavirus neutralising serum antibodies in dromedary camels: a comparative serological study. Lancet Infect. Dis. 13, 859–866. http://dx.doi.org/10.1016/S1473-3099(13)70164-6

Tonnessen R, Hauge AG, Hansen EF, Rimstad E, Jonassen CM (2013): Host restrictions of avian influenza viruses: in silico analysis of H13 and H16 specific signatures in the internal proteins. PloS One 8, e63270. http://dx.doi.org/10.1371/journal.pone.0063270

Wang N, Shi X, Jiang L, Zhang S, Wang D, Tong P, Guo D, Fu L, Cui Y, Liu X, Arledge KC, Chen YH, Zhang L, Wang X (2013): Structure of MERS-CoV spike receptor-binding domain complexed with human receptor DPP4. Cell Res. 23, 986–993. http://dx.doi.org/10.1038/cr.2013.92