



Structure-function relationships among selected human coronaviruses

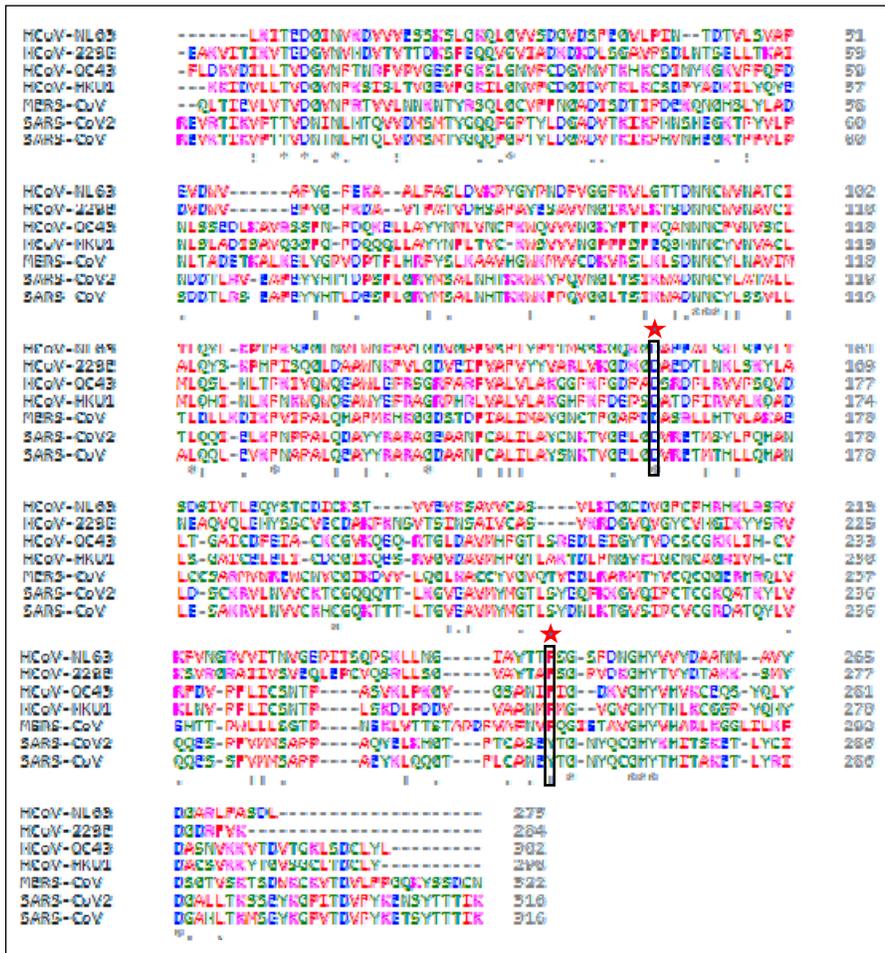
Esther Jamir^{1,2}, Kikrusenuo Kiewhuo^{1,2}, Lipsa Priyadarsinee^{1,2},
Himakshi Sarma¹, Selvaraman Nagamani^{1,2} & G Narahari Sastry^{1,2*}

¹Advanced Computation and Data Sciences Division, CSIR-North East Institute of Science and Technology, Jorhat-785 006, Assam, India

²Academy of Scientific and Innovative Research (AcSIR), Ghaziabad-201 002, Uttar Pradesh, India

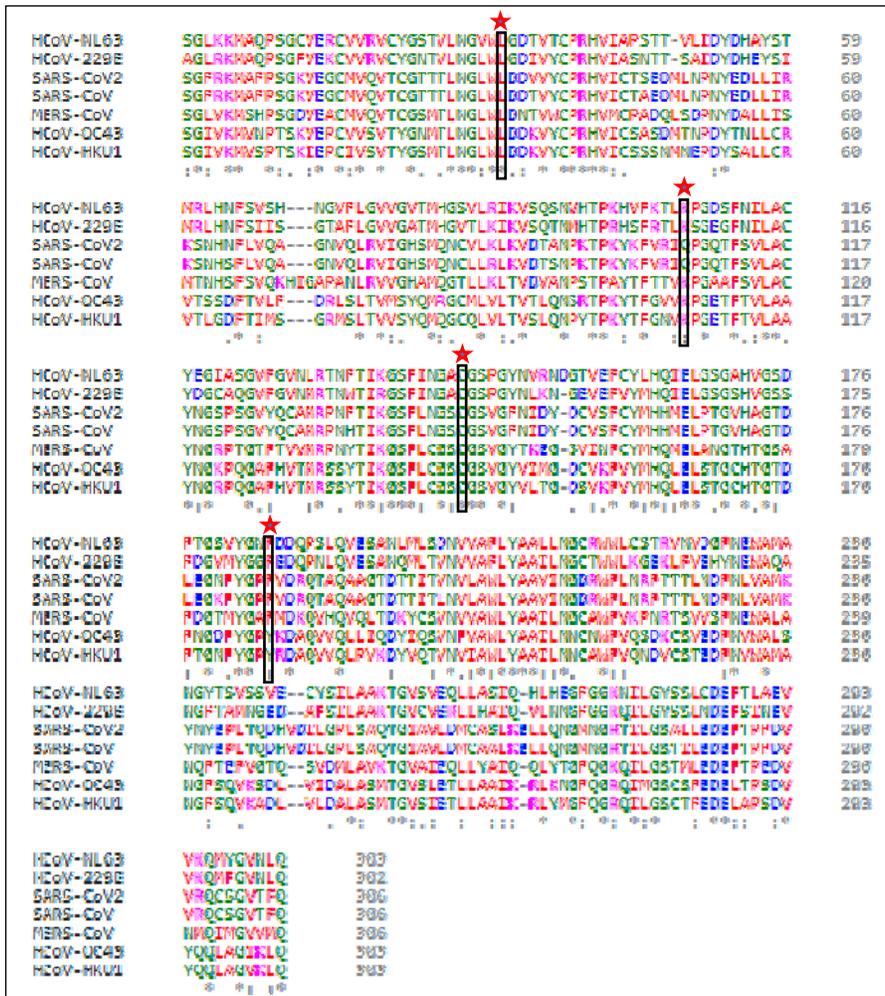
Received 15 February 2022; revised 19 May 2022

Supplementary Data



“*” Fully conserved “.” Conservation between groups of amino acids with strongly similar properties “.” Conservation between groups of amino acids with weakly similar properties “*” Fully conserved “.” Conservation between groups of amino acids with strongly similar properties “.” Conservation between groups of amino acids with weakly similar properties

Figure S1(a). Multiple sequence alignment of NSP3 (PI-pro) protein of Human coronavirus OC43, Human coronavirus HKU1, Human coronavirus NL63, Human coronavirus 229E, MERS -CoV, SARS-CoV2 and SARS-CoV. The analysis was performed using Clustal Omega. The active sites residues present in the conserved regions are highlighted in box with star marks.



“*” Fully conserved “:” Conservation between groups of amino acids with strongly similar properties “.” Conservation between groups of amino acids with weakly similar properties “*” Fully conserved “:” Conservation between groups of amino acids with strongly similar properties “.” Conservation between groups of amino acids with weakly similar properties

Figure S1(b). Multiple sequence alignment of NSP5 (Cl-pro) protein of Human coronavirus OC43, Human coronavirus HKU1, Human coronavirus NL63, Human coronavirus 229E, MERS -CoV, SARS-CoV2 and SARS-CoV. The analysis was performed using Clustal Omega. The active sites residues present in the conserved regions are highlighted in box with star marks.

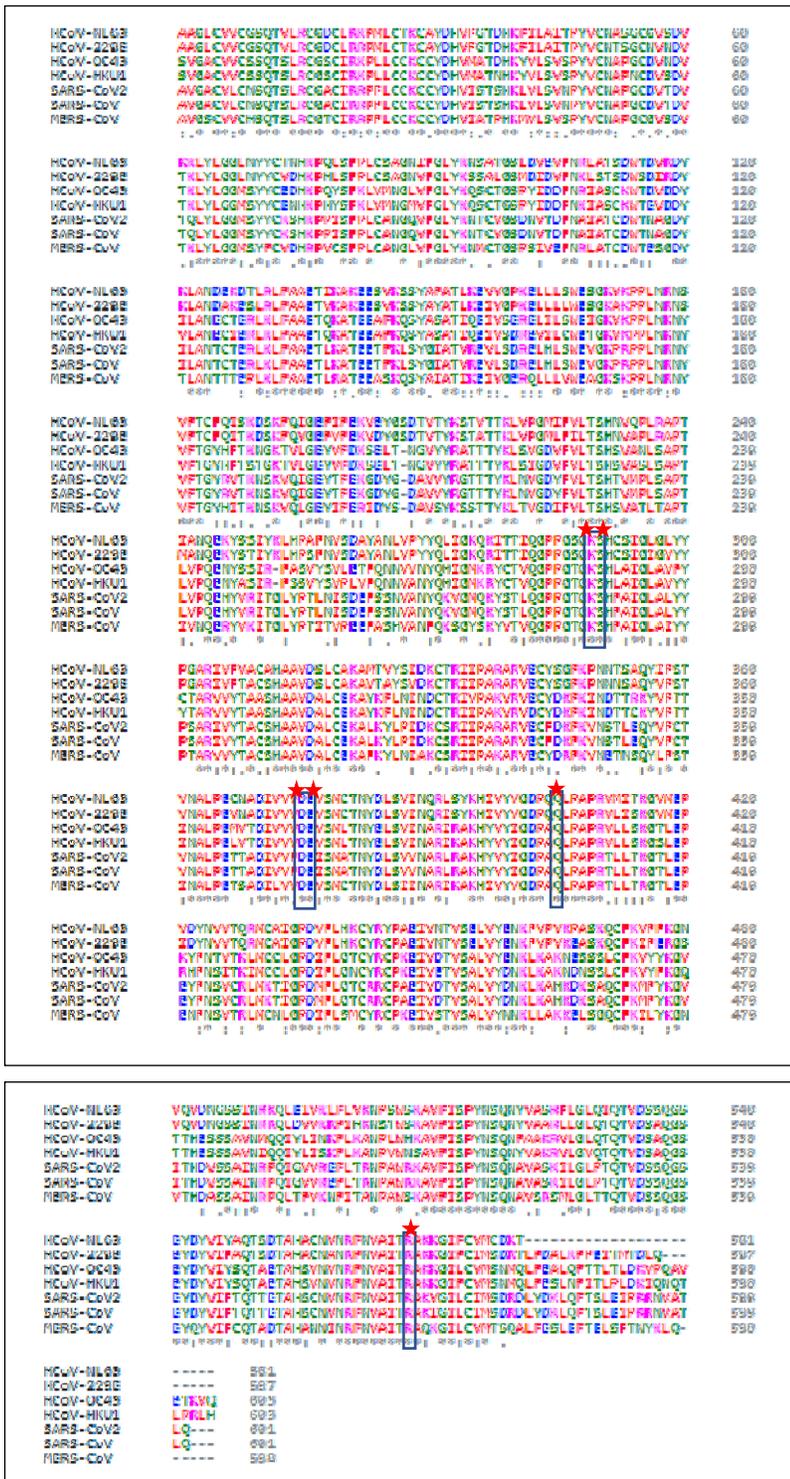
HCov-NL63	--SYDYSYLNRRAG-SSAARLEPCN-QTIDKCYRAPPDIY--NKNVSPLEKCLRMNCVR	53
HCov-229E	--SFUSYLNRRAG-SSAARLEPCN-GIDIDYCVRRAPDY--NNDASLHGNLRKNCVR	53
HCov-OC43	--SKDTNPLNRYRGSVDARLVPCASGLSDVQLRAPDIY--NAGVAGDGLMLKVNCCR	55
HCov-HKU1	--SKDLNPLNRYRGTSMARLVPCASGLSDVQLRAPDGC--NTRRAGDGLYYKVNCCR	55
MERS-CoV	--SFLUNPLNRYRGSIVNARLSPCYSMLLDLYVYHRRPDCYKCAKWAQDQRYKINICR	57
SARS-CoV2	--SADAGSFLNRYCG-VSAARLTCGTTGTSDVYVYRAPPDIY--NDKVGAFKFLKTNCCR	55
SARS-CoV	NGSADASTFLNRYCG-VSAARLTCGTTGTSDVYVYRAPPDIY--NEKVGAFKFLKTNCCR	57
	: :	
HCov-NL63	FRNADI---LKDQYTVIKRCTHSYVHEHCGSNYLLNFSGALAEHDFPTWHDGRVYIYGNVS	109
HCov-229E	FRNVDI---KDDAFYIVKRCIHSYVDHEQSNYLLKGCNAAVAKHDFRTIHGRTIYGNVS	109
HCov-OC43	PQWDENDKDLQFVYVYKTDLTIVNREMYCYRYNDCKFYRKHOPPTFDVYSSRYPHIV	115
HCov-HKU1	RQSDDDGKDLDFVYVYKTDLTIVNREMYCYRYNDCKFYRKHOPPTFDVYSSRYPHIV	115
MERS-CoV	PVELDQGHLLDSYVYVYKTDLTIVNREMYCYRYNDCKFYRKHOPPTFDVYSSRYPHIV	117
SARS-CoV2	RQSDDDGKDLDSYVYVYKTDLTIVNREMYCYRYNDCKFYRKHOPPTFDVYSSRYPHIV	115
SARS-CoV	RQSDDDGKDLDSYVYVYKTDLTIVNREMYCYRYNDCKFYRKHOPPTFDVYSSRYPHIV	117
	: :	
HCov-NL63	RKHLTKYTVNQLVYVYKTDLTIVNREMYCYRYNDCKFYRKHOPPTFDVYSSRYPHIV	169
HCov-229E	RKQLTRYTVNQLLFAKRNKUPKDCGYFKEILVLTIGCCNTDYPENRMMFPPIERSDIHWY	169
HCov-OC43	RKDLTRYTVNQLLFAKRNKUPKDCGYFKEILVLTIGCCNTDYPENRMMFPPIERSDIHWY	175
HCov-HKU1	RKDLTRYTVNQLLFAKRNKUPKDCGYFKEILVLTIGCCNTDYPENRMMFPPIERSDIHWY	175
MERS-CoV	RKDLTRYTVNQLLFAKRNKUPKDCGYFKEILVLTIGCCNTDYPENRMMFPPIERSDIHWY	176
SARS-CoV2	RKDLTRYTVNQLLFAKRNKUPKDCGYFKEILVLTIGCCNTDYPENRMMFPPIERSDIHWY	175
SARS-CoV	RKDLTRYTVNQLLFAKRNKUPKDCGYFKEILVLTIGCCNTDYPENRMMFPPIERSDIHWY	177
	: :	
HCov-NL63	ASLQKIVAAAILKCVALLCDAAVYKQVYVGLTLDNQDLNGLNYDFGDFVSLPMMQPCCT	229
HCov-229E	AALGKVVAAAILKCVAFCDERWLGKVGWGLTLDNQDLNGLNYDFGDFVSLPMMQPCCT	229
HCov-OC43	KILGPIFNALVSAATEFADLVEVGLVGLVGLTLDNQDLNGLNYDFGDFVSLPMMQPCCT	235
HCov-HKU1	KILGPIFNALLNTVSPADTLVYVGLVGLVGLTLDNQDLNGLNYDFGDFVSLPMMQPCCT	235
MERS-CoV	HLGSRVQKDLNLYKPCDHYVYVGLVGLVGLTLDNQDLNGLNYDFGDFVSLPMMQPCCT	239
SARS-CoV2	ANLGRVQKDLNLYKPCDHYVYVGLVGLVGLTLDNQDLNGLNYDFGDFVSLPMMQPCCT	235
SARS-CoV	ANLGRVQKDLNLYKPCDHYVYVGLVGLVGLTLDNQDLNGLNYDFGDFVSLPMMQPCCT	237
	: :	
HCov-NL63	SYYSYVWPDINGLTNCLASECFVKSDFGSDIFGDFKTFDLKLYDFTBHKELNLFNKYFKMHSFDY	289
HCov-229E	SYVAYVWPDINGLTNCLASECFVKSDFGSDIFGDFKTFDLKLYDFTBHKELNLFNKYFKMHSFDY	289
HCov-OC43	SYYSYVWPDINGLTNCLASECFVKSDFGSDIFGDFKTFDLKLYDFTBHKELNLFNKYFKMHSFDY	290
HCov-HKU1	SYYSYVWPDINGLTNCLASECFVKSDFGSDIFGDFKTFDLKLYDFTBHKELNLFNKYFKMHSFDY	290
MERS-CoV	SYYSYVWPDINGLTNCLASECFVKSDFGSDIFGDFKTFDLKLYDFTBHKELNLFNKYFKMHSFDY	295
SARS-CoV2	SYYSYVWPDINGLTNCLASECFVKSDFGSDIFGDFKTFDLKLYDFTBHKELNLFNKYFKMHSFDY	294
SARS-CoV	SYYSYVWPDINGLTNCLASECFVKSDFGSDIFGDFKTFDLKLYDFTBHKELNLFNKYFKMHSFDY	296
	: :	
HCov-NL63	HPNCSDCYDDMCVICHANFNITLFAITIPGTAFGLCRKRVFDGVPVLTAGYHFKQLGLV	349
HCov-229E	HPDCVDCHEMCIHCSNFNITLFAITIPGTAFGLCRKRVFDGVPVLTAGYHFKQLGLV	349
HCov-OC43	HPNTVDCQDDRCIHCANFNILFSAVLPNTCFGLVLRQIFWGDVFPVVSIGYHYKELGIV	350
HCov-HKU1	HPNTVDCQDDRCIHCANFNILFSAVLPNTCFGLVLRQIFWGDVFPVVSIGYHYKELGIV	350
MERS-CoV	HANVCNCTDDRCVLCANFNILFSAVLPNTCFGLVLRQIFWGDVFPVVSIGYHYKELGIV	355
SARS-CoV2	HPNVCNCTDDRCVLCANFNILFSAVLPNTCFGLVLRQIFWGDVFPVVSIGYHYKELGIV	354
SARS-CoV	HPNVCNCTDDRCVLCANFNILFSAVLPNTCFGLVLRQIFWGDVFPVVSIGYHYKELGIV	356
	: :	
HCov-NL63	WNKDVNTHSVRLTETELLOFVTDPSLIIASSPALVDQRTICFSVAALSTGLTNQVAKPGH	409
HCov-229E	WNKDVNTHSTRLTETELLOFVTDPTLIWASSPALVDKRTVCFSSVAALSTGLTSQVAKPGH	409
HCov-OC43	NNNDVTHRYRLSLKDLLYAADPRLHVASASALYDLRTECCFSAVATSGVAFQVAKPGN	410
HCov-HKU1	NNLDVTHRYRLSLKDLLYAADPRLHVASASALLDRTECCFSAVATSGVAFQVAKPGN	410
MERS-CoV	NNNDVTHRRRLSLKDLLYAADPRLHVASASALLDRTECCFSAVATSGVAFQVAKPGN	415
SARS-CoV2	HNQDVNTHSSRLSFLKLLVYAADPRLHVASASGMLLLDKRTTCFSAVATSNVAFQVAKPGN	414
SARS-CoV	HNQDVNTHSSRLSFLKLLVYAADPRLHVASASGMLLLDKRTTCFSAVATSNVAFQVAKPGN	416
	: :	
HCov-NL63	FNIEFYNFLRLRGFFDEGSELTLEKHPFAQNSDAVQDFDYRYNKPITLIDICQARVYTK	469
HCov-229E	FNKEFYDFLRSQGFDEGSELTLEKHPFAQNSDAVQDFDYRYNKPITLIDICQARVYTK	469
HCov-OC43	FNQDFYDFVLSKGLLEKGSVDLKHPPFTQDGHAAITDYNYKYRLPTNCDIROLFLVLE	470
HCov-HKU1	FNQDFYDFVLSKGLLEKGSVDLKHPPFTQDGHAAITDYNYKYRLPTNCDIROLFLVLE	470
MERS-CoV	FNQDFYDFVLSKGFLEKGSVTLKHPPFAQDGHAAITDYNYKYRLPTNCDIROLFLVLE	475
SARS-CoV2	FNKDFYDFVLSKGFLEKGSVTLKHPPFAQDGHAAITDYNYKYRLPTNCDIROLFLVLE	474
SARS-CoV	FNKDFYDFVLSKGFLEKGSVTLKHPPFAQDGHAAITDYNYKYRLPTNCDIROLFLVLE	476
	: :	
HCov-NL63	IVSRYPDIYEGGCIKACVVVNLNHSAGMPLNKRKASLYYESIYEEQDALFALTIRN	529
HCov-229E	VAAYTYDCTCCGCIIRKIVVYVNLNHSAGMPLNKRKALYYCALSYCCQDASILLTKRN	529
HCov-OC43	VVYNYFSEIYDGGCIKASQVIVNNDKASAGYFNKRKARLYYKALSYEEQDEIYAYTIRN	530
HCov-HKU1	VVYNYFSEIYDGGCIKASQVIVNNDKASAGYFNKRKARLYYKALSYEEQDEIYAYTIRN	530
MERS-CoV	VVYNYFSEIYDGGCIKASQVIVNNDKASAGYFNKRKARLYYKALSYEEQDEIYAYTIRN	535
SARS-CoV2	VVYNYFSEIYDGGCIKASQVIVNNDKASAGYFNKRKARLYYKALSYEEQDEIYAYTIRN	534
SARS-CoV	VVYNYFSEIYDGGCIKASQVIVNNDKASAGYFNKRKARLYYKALSYEEQDEIYAYTIRN	536
	: :	
HCov-NL63	VLPTITQNLKYAISAKNIAATVYGVSIKSTWITNQHQLKLSIAATNGATVYIGTSKF	593
HCov-229E	VLPTITQNLKYAISAKNIAATVYGVSIKSTWITNQHQLKLSIAATNGATVYIGTSKF	593
HCov-OC43	VLPTITQNLKYAISAKNIAATVYGVSIKSTWITNQHQLKLSIAATNGATVYIGTSKF	593
HCov-HKU1	VLPTITQNLKYAISAKNIAATVYGVSIKSTWITNQHQLKLSIAATNGATVYIGTSKF	593
MERS-CoV	VLPTITQNLKYAISAKNIAATVYGVSIKSTWITNQHQLKLSIAATNGATVYIGTSKF	594
SARS-CoV2	VLPTITQNLKYAISAKNIAATVYGVSIKSTWITNQHQLKLSIAATNGATVYIGTSKF	594
SARS-CoV	VLPTITQNLKYAISAKNIAATVYGVSIKSTWITNQHQLKLSIAATNGATVYIGTSKF	596
	: :	





“*” Fully conserved “:” Conservation between groups of amino acids with strongly similar properties “.” Conservation between groups of amino acids with weakly similar properties “*” Fully conserved “:” Conservation between groups of amino acids with strongly similar properties “.” Conservation between groups of amino acids with weakly similar properties

Figure S1(c). Multiple sequence alignment of NSP12 protein of Human coronavirus OC43, Human coronavirus HKU1, Human coronavirus NL63, Human coronavirus 229E, MERS -CoV, SARS-CoV2 and SARS-CoV. The analysis was performed using Clustal Omega. The active sites residues present in the conserved regions are highlighted in box with star marks.



“*” Fully conserved “:” Conservation between groups of amino acids with strongly similar properties “.” Conservation between groups of amino acids with weakly similar properties “*” Fully conserved “:” Conservation between groups of amino acids with strongly similar properties “.” Conservation between groups of amino acids with weakly similar properties
Figure S1(d). Multiple sequence alignment of NSP13 protein of Human coronavirus OC43, Human coronavirus HKU1, Human coronavirus NL63, Human coronavirus 229E, MERS -CoV, SARS-CoV2 and SARS-CoV. The analysis was performed using Clustal Omega. The active sites residues present in the conserved regions are highlighted in box with star marks.

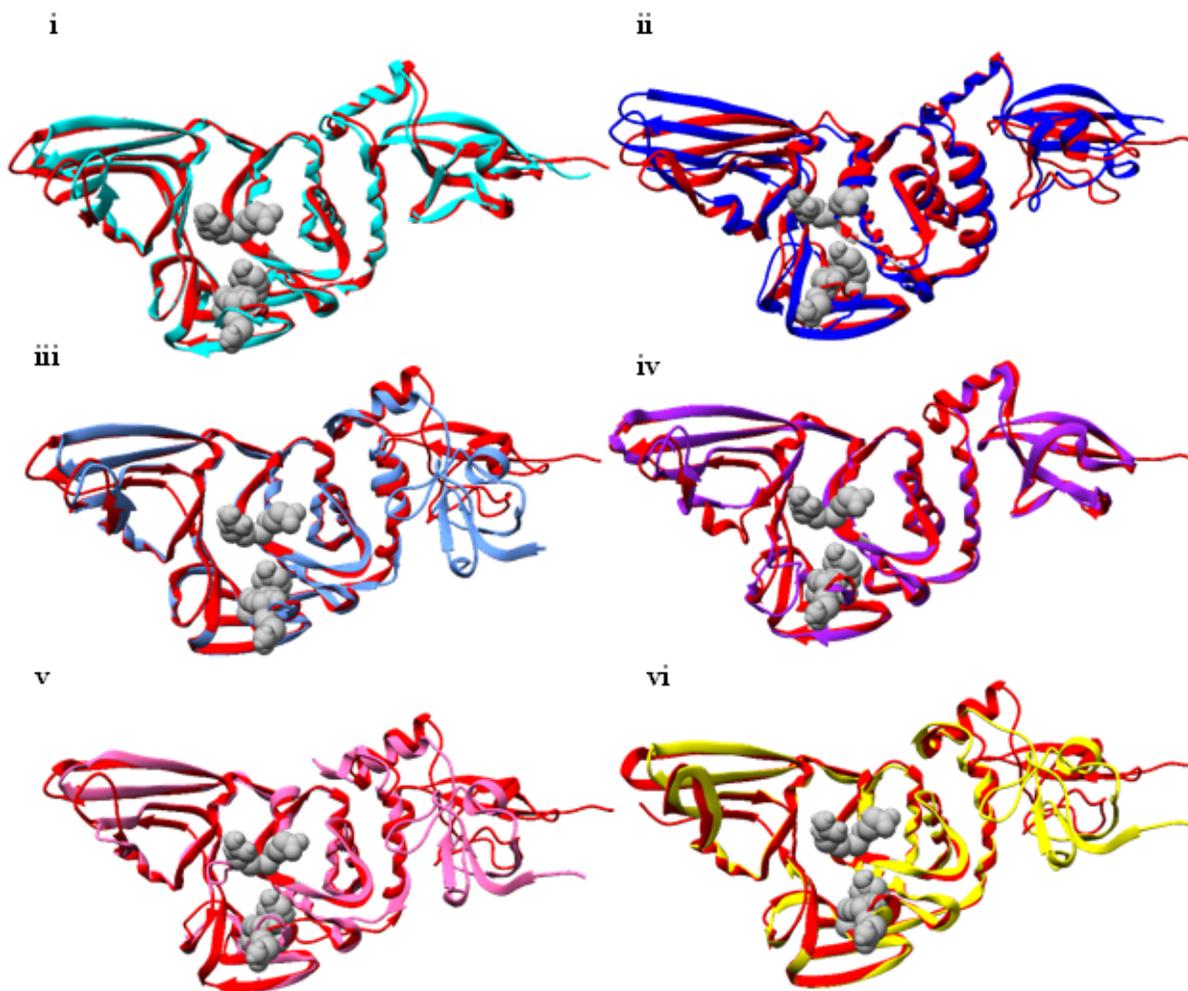


Figure S2(a). Superimposed structures of NSP3 (PI-pro) of i) SARS-CoV2 vs SARS-CoV ii) SARS-CoV2 vs MERS-CoV iii) SARS-CoV2 vs HCoV-HKU1 iv) SARS-CoV2 vs HCoV-NL63 v) SARS-CoV2 vs HCoV-229E vi) SARS-CoV2 vs HCoV-OC43 using TM-Align and UCSF Chimera.

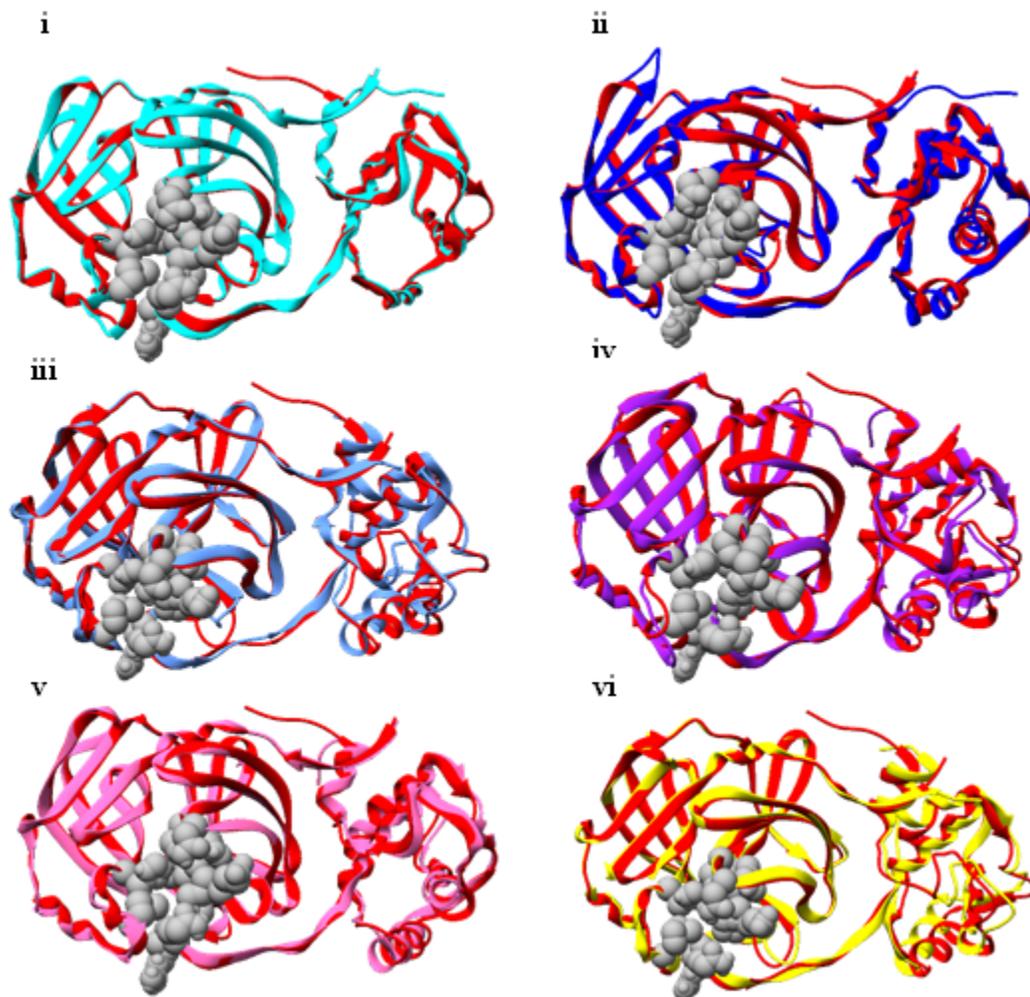


Figure S2 (b). Superimposed structures of NSP5 (3Cl-pro) of i) SARS-CoV2 vs SARS-CoV ii) SARS-CoV2 vs MERS-CoV iii) SARS-CoV2 vs HCoV-HKU1 iv) SARS-CoV2 vs HCoV-NL63 v) SARS-CoV2 vs HCoV-229E vi) SARS-CoV2 vs HCoV-OC43 using TM-Align and UCSF Chimera.

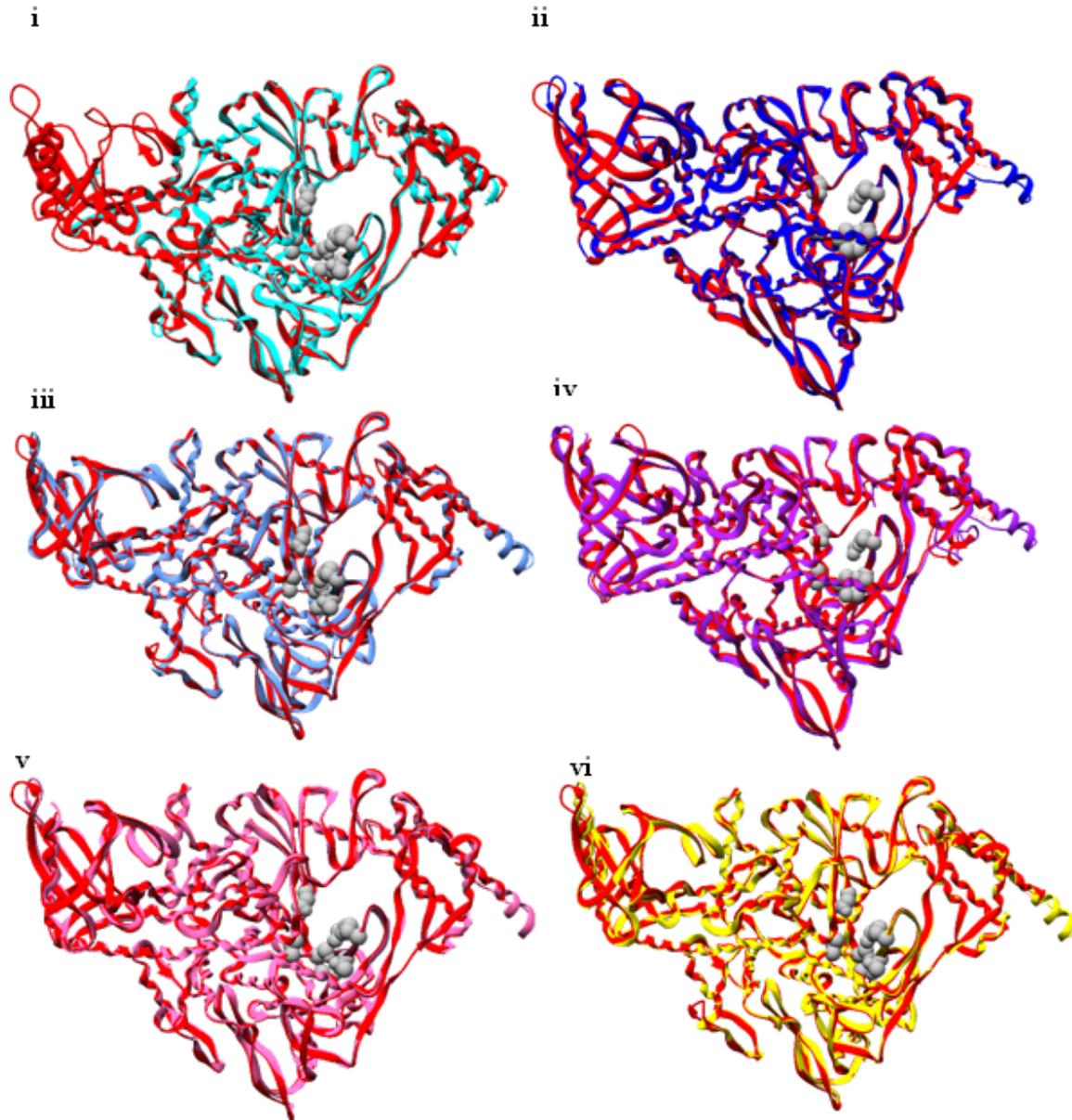


Figure S2 (c). Superimposed structures of NSP12 of i) SARS-CoV2 vs SARS-CoV ii) SARS-CoV2 vs MERS-CoV iii) SARS-CoV2 vs HCoV-HKU1 iv) SARS-CoV2 vs HCoV-NL63 v) SARS-CoV2 vs HCoV-229E vi) SARS-CoV2 vs HCoV-OC43 using TM-Align and UCSF Chimera.

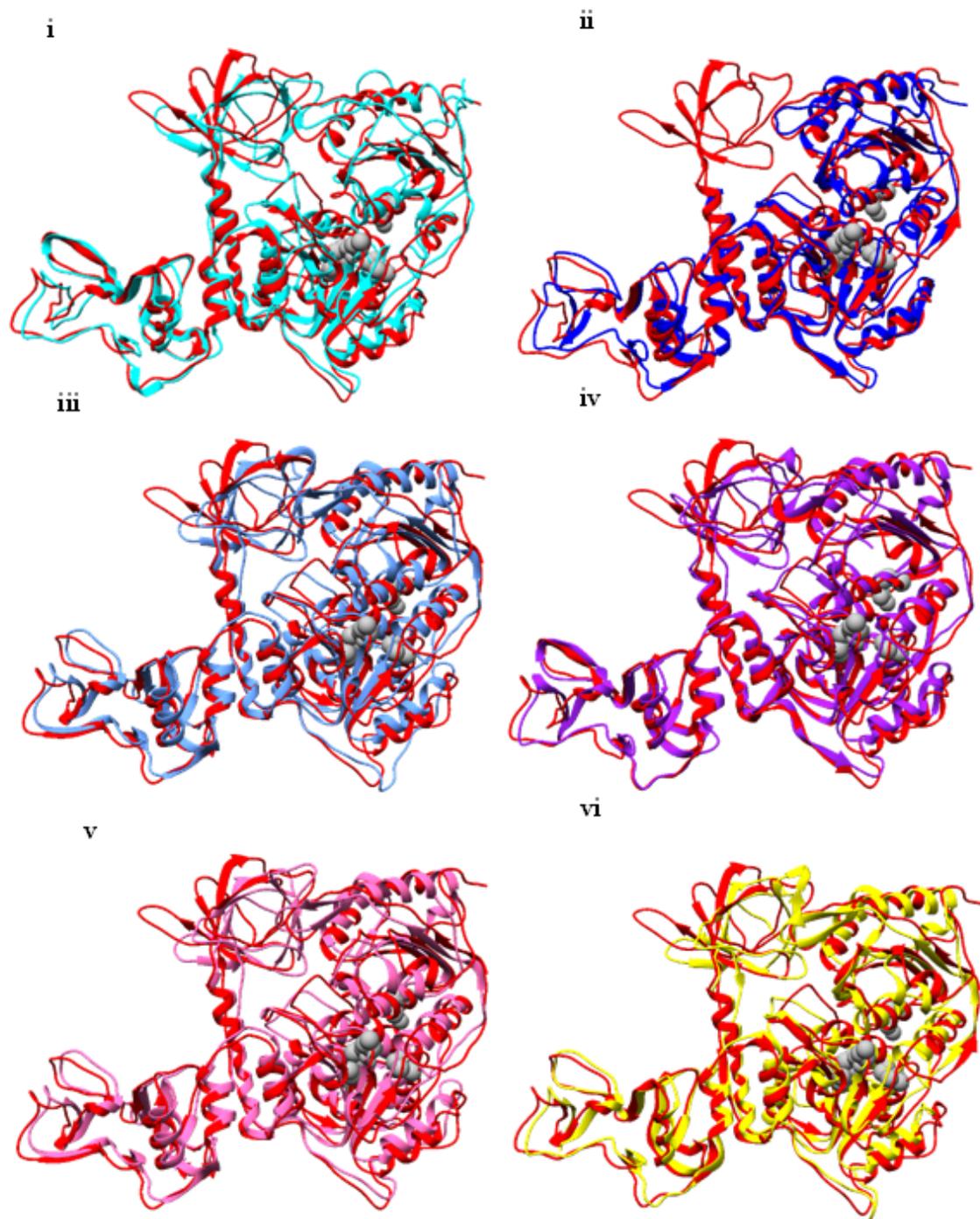


Figure S2 (d). Superimposed structures of NSP13 of i) SARS-CoV2 vs SARS-CoV ii) SARS-CoV2 vs MERS-CoV iii) SARS-CoV2 vs HCoV-HKU1 iv) SARS-CoV2 vs HCoV-NL63 v) SARS-CoV2 vs HCoV-229E vi) SARS-CoV2 vs HCoV-OC43 using TM-Align and UCSF Chimera.

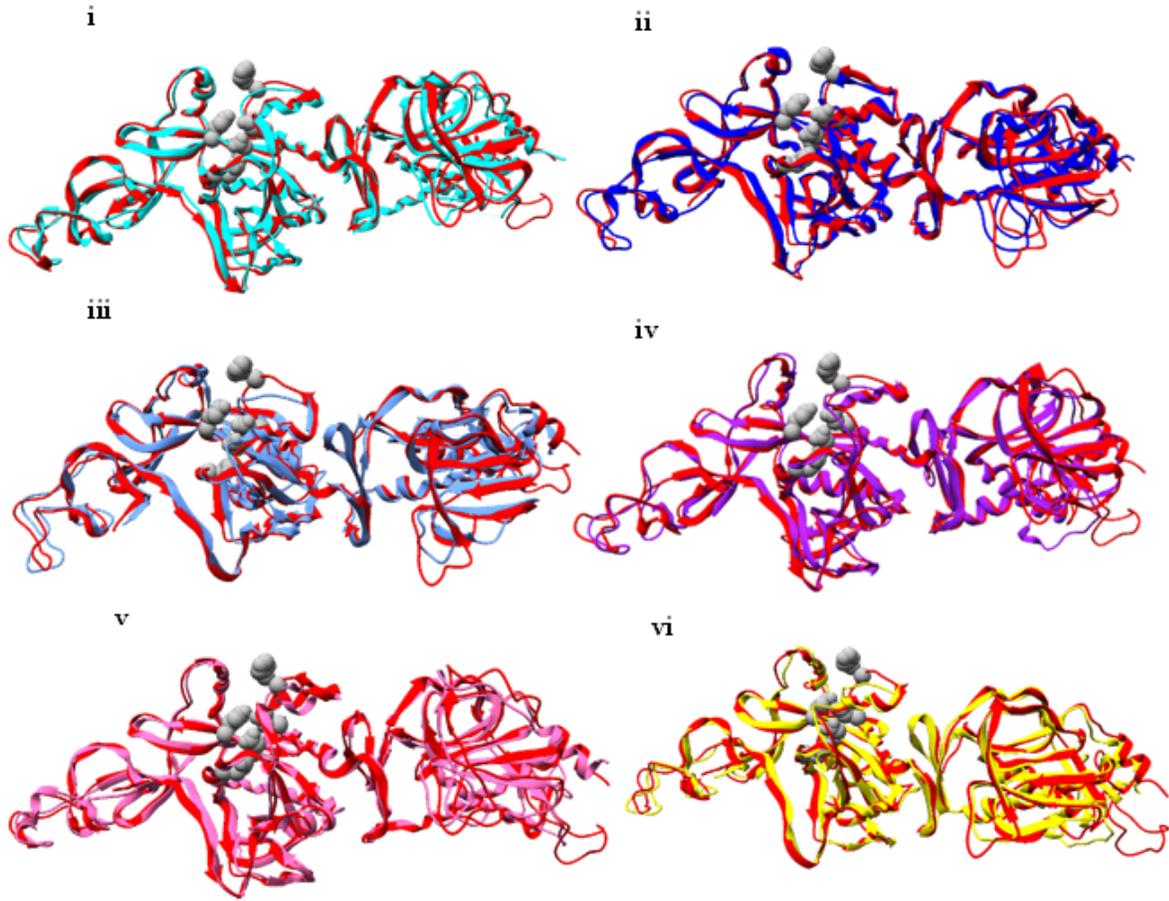


Figure S2 (e). Superimposed structures of NSP14 of i) SARS-CoV2 vs SARS-CoV ii) SARS-CoV2 vs MERS-CoV iii) SARS-CoV2 vs HCoV-HKU1 iv) SARS-CoV2 vs HCoV-NL63 v) SARS-CoV2 vs HCoV-229E vi) SARS-CoV2 vs HCoV-OC43 using TM-Align and UCSF Chimera.

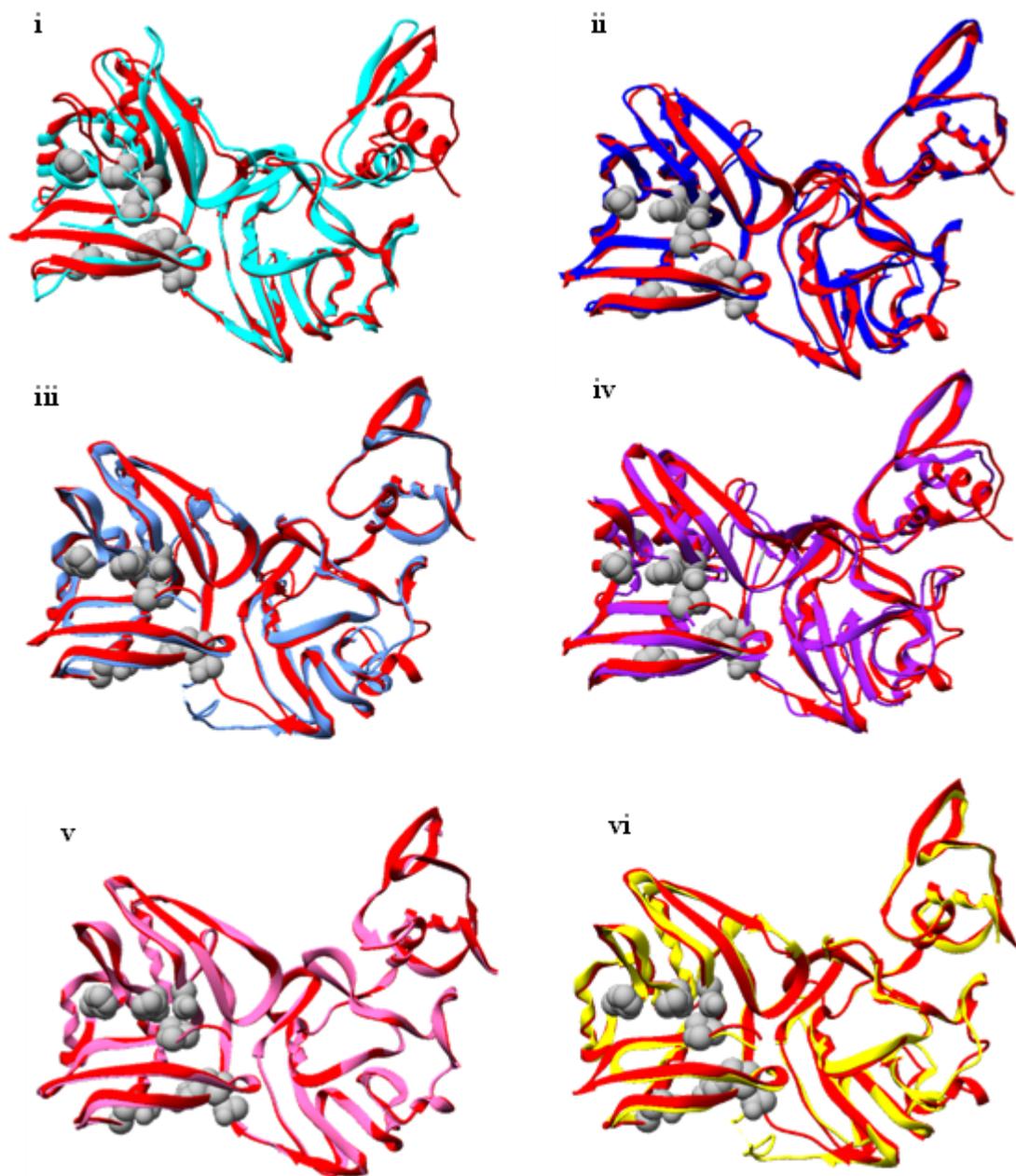


Figure S2 (f). Superimposed structures of NSP15 of i) SARS-CoV2 vs SARS-CoV ii) SARS-CoV2 vs MERS-CoV iii) SARS-CoV2 vs HCoV-HKU1 iv) SARS-CoV2 vs HCoV-NL63 v) SARS-CoV2 vs HCoV-229E vi) SARS-CoV2 vs HCoV-OC43 using TM-Align and UCSF Chimera.

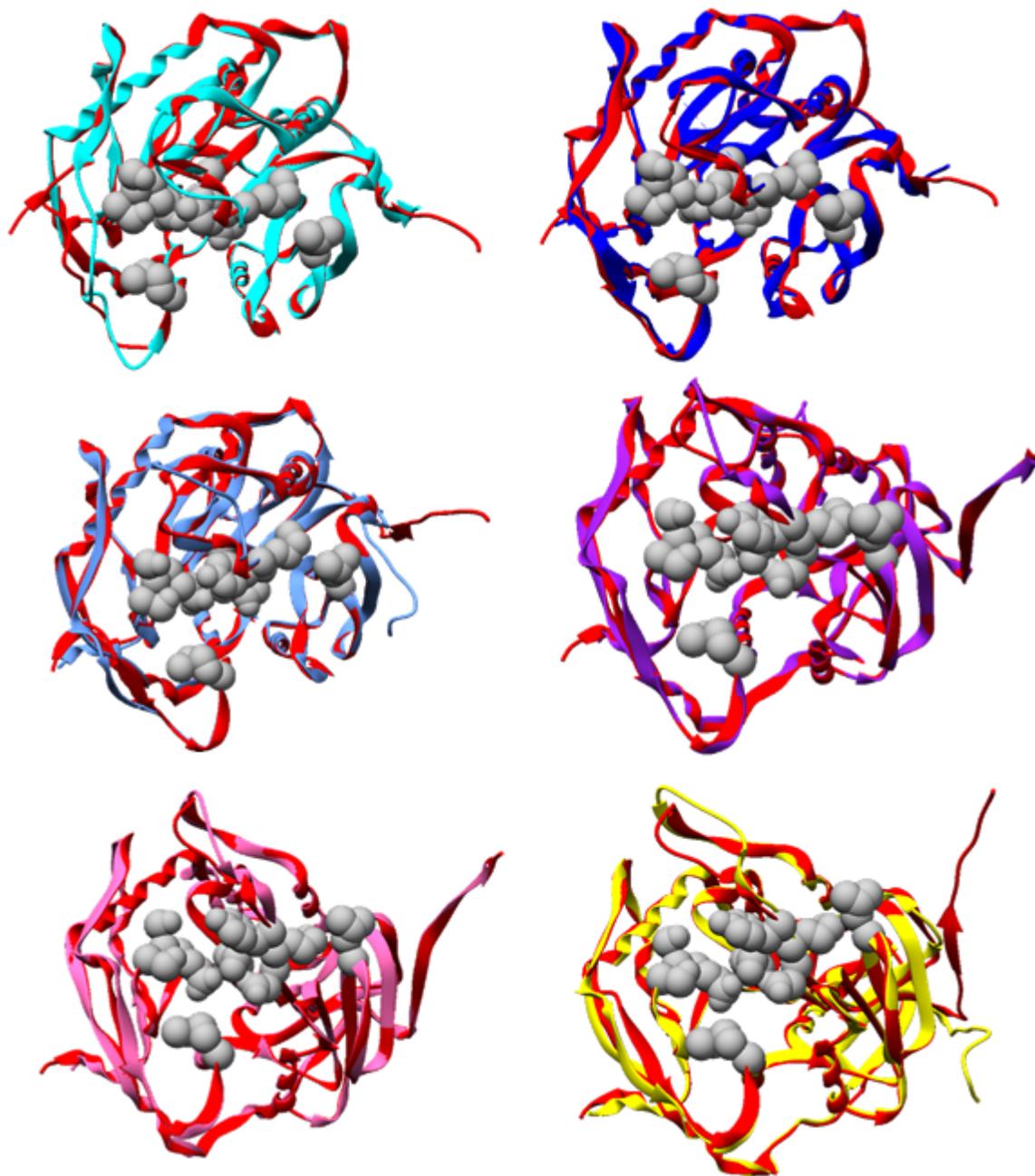


Figure S2 (g). Superimposed structures of NSP16 of i) SARS-CoV2 vs SARS-CoV ii) SARS-CoV2 vs MERS-CoV iii) SARS-CoV2 vs HCoV-HKU1 iv) SARS-CoV2 vs HCoV-NL63 v) SARS-CoV2 vs HCoV-229E vi) SARS-CoV2 vs HCoV-OC43 using TM-Align and UCSF Chimera.

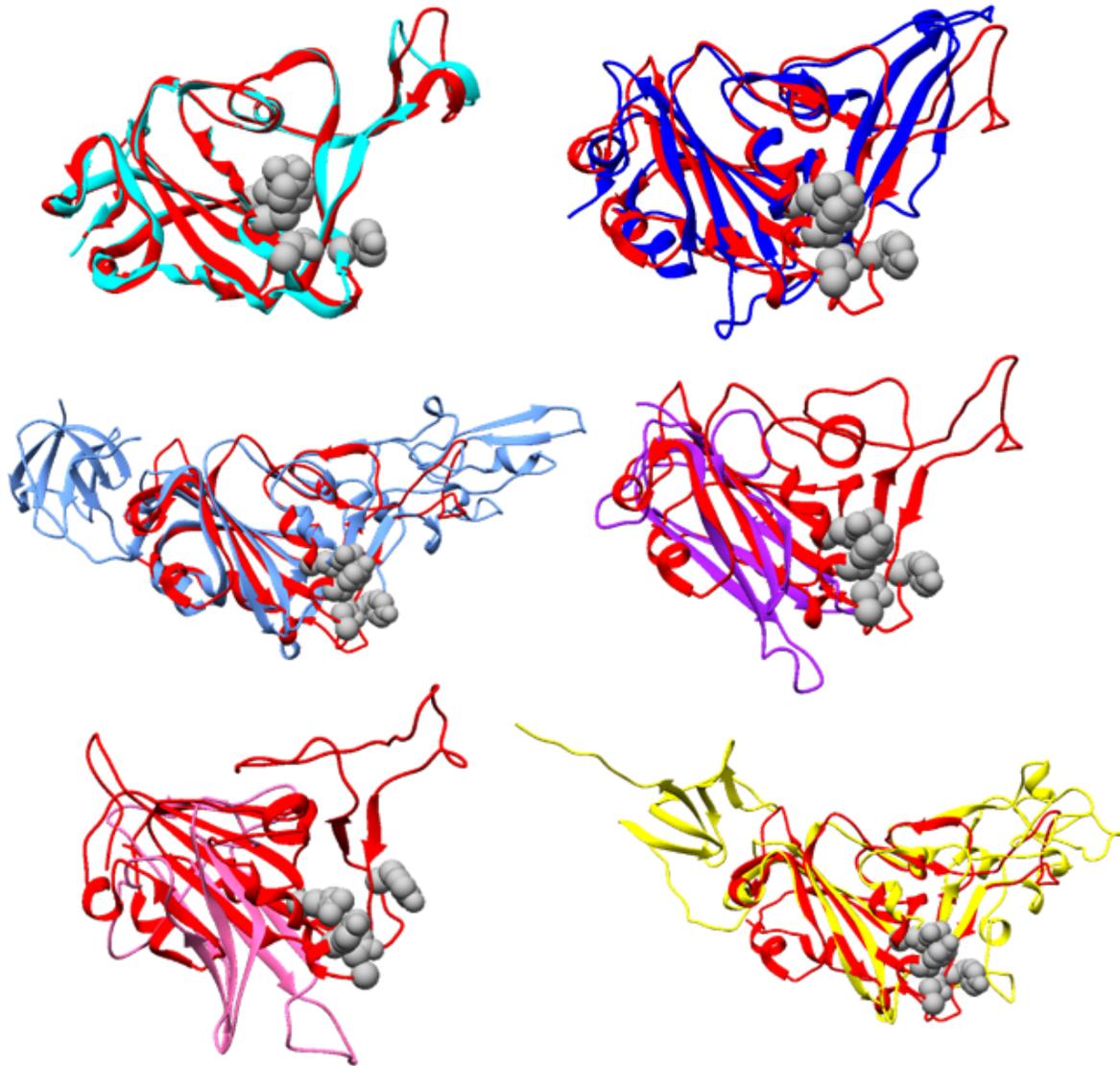


Figure S2(h). Superimposed structures of Spike RBD of i) SARS-CoV2 vs SARS-CoV ii) SARS-CoV2 vs MERS-CoV iii) SARS-CoV2 vs HCoV-HKU1 iv) SARS-CoV2 vs HCoV-NL63 v) SARS-CoV2 vs HCoV-229E vi) SARS-CoV2 vs HCoV-OC43 using TM-Align and UCSF Chimera

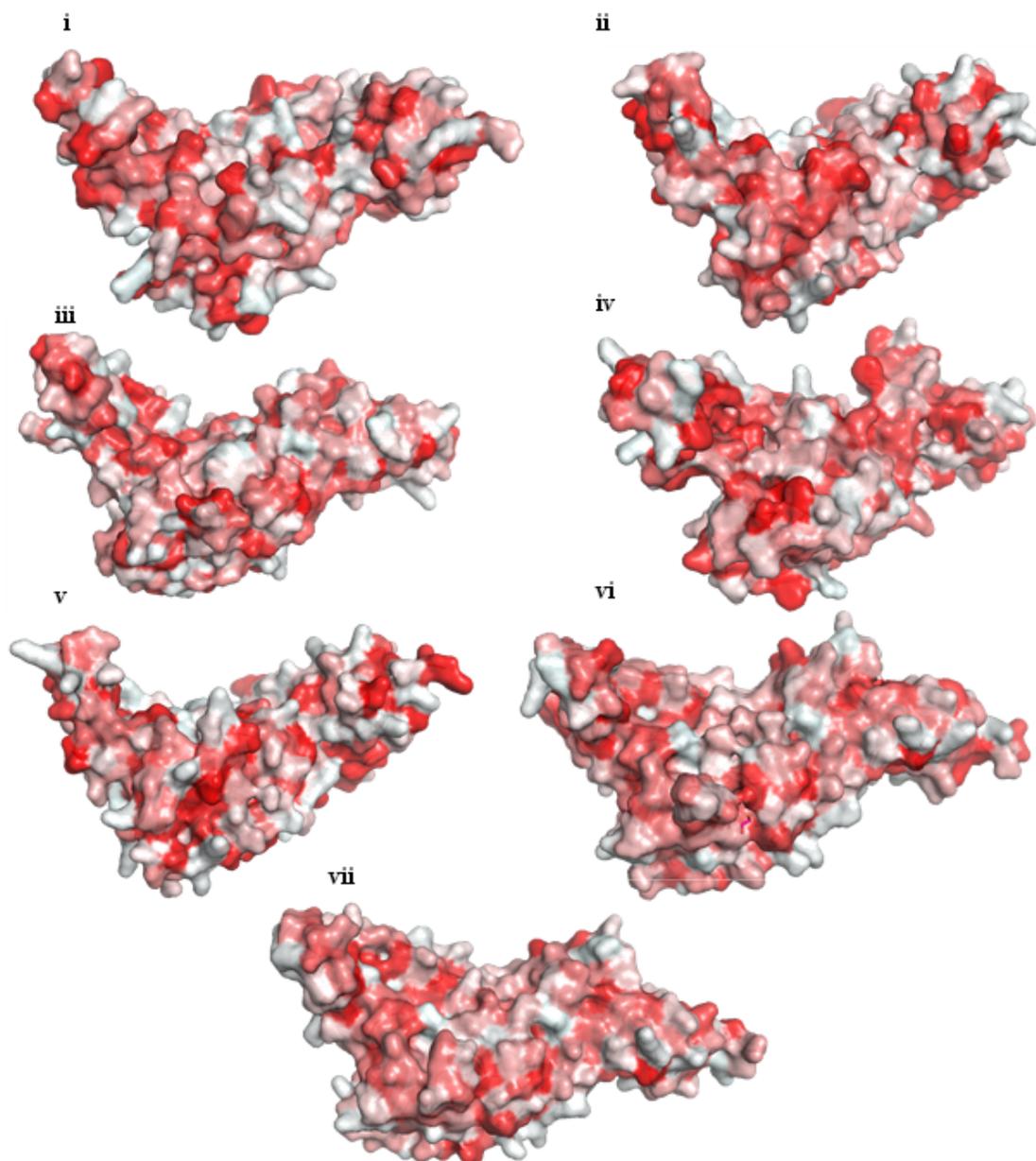


Figure S3(a). Hydrophobicity map of NSP3 protein surface of all 7 types of corona viruses (i) 229E, (ii) HKU1 (iii) MERS CoV (iv) NL63 (v) OC43 (vi) SARS CoV and (vii) SARS CoV2. Red colour regions indicate hydrophobicity and white colour regions indicate hydrophilicity.

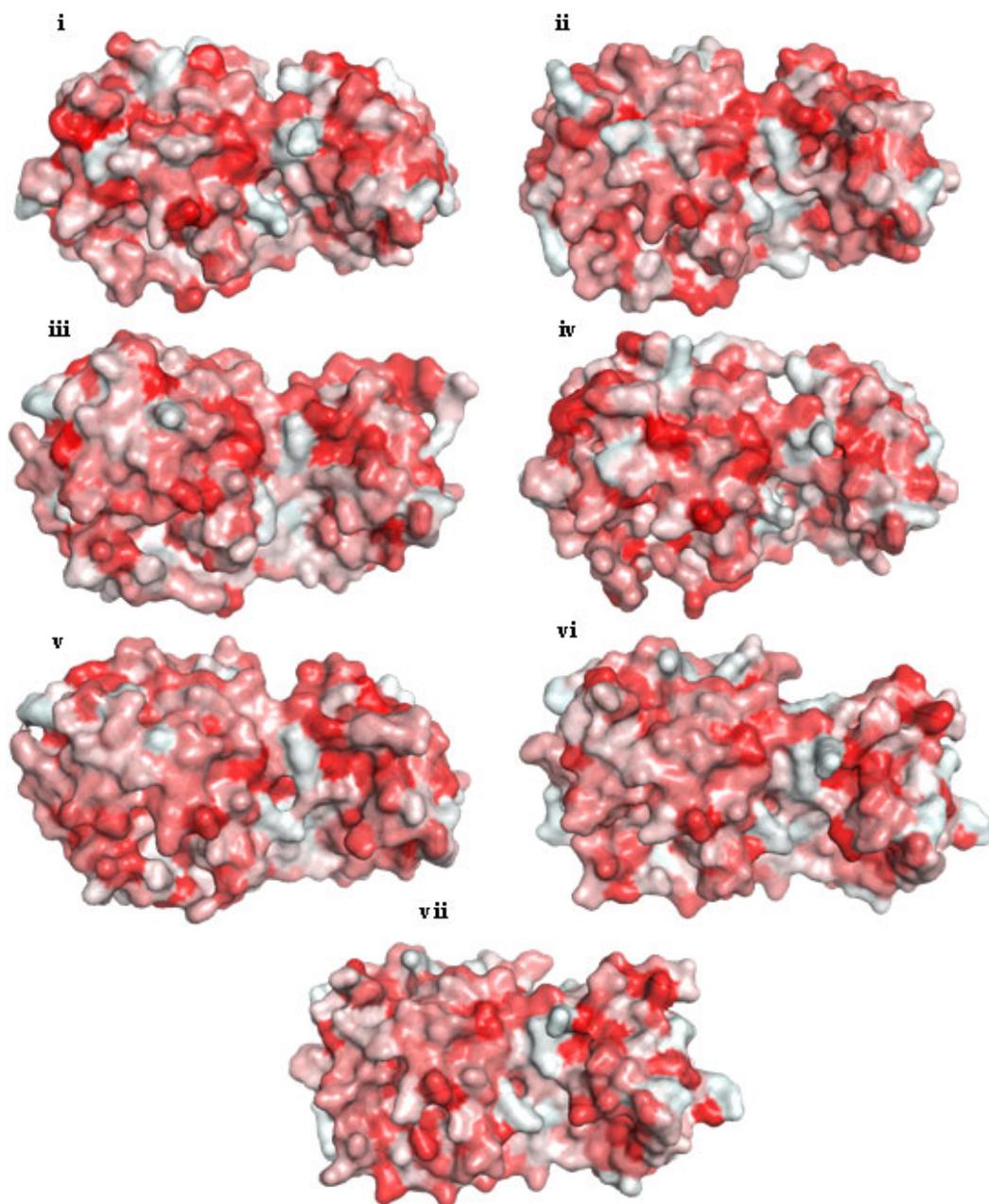


Figure S3 (b). Hydrophobicity map of NSP5 protein surface of all 7 types of corona viruses (i) 229E, (ii) HKU1 (iii) MERS CoV (iv) NL63 (v) OC43 (vi) SARS CoV and (vii) SARS CoV2. Red colour regions indicate hydrophobicity and white colour regions indicate hydrophilicity.

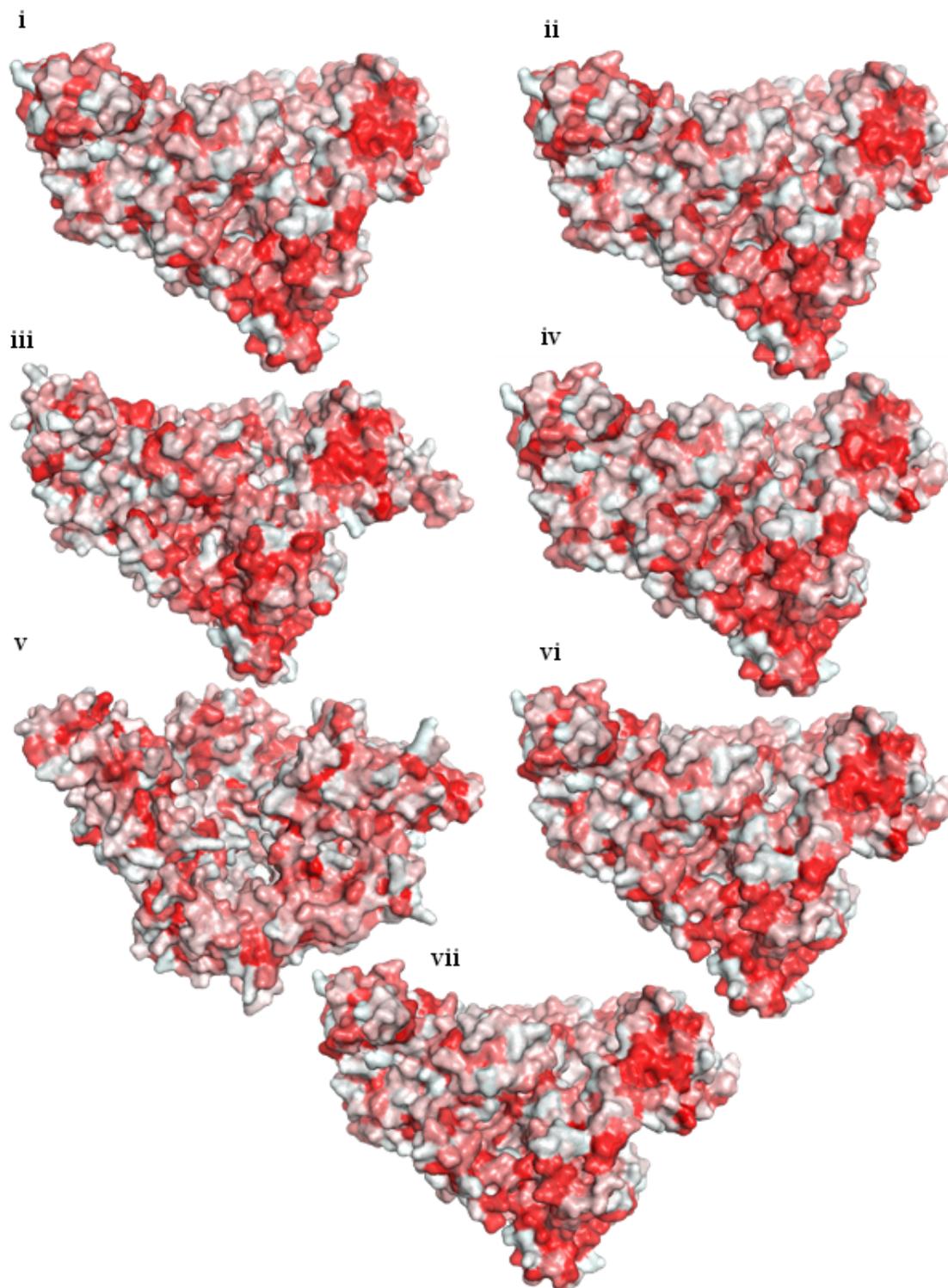


Figure S3(c). Hydrophobicity map of NSP12 protein surface of all 7 types of corona viruses (i) 229E, (ii) HKU1 (iii) MERS CoV (iv) NL63 (v) OC43 (vi) SARS CoV and (vii) SARS CoV2. Red colour regions indicate hydrophobicity and white colour regions indicate hydrophilicity

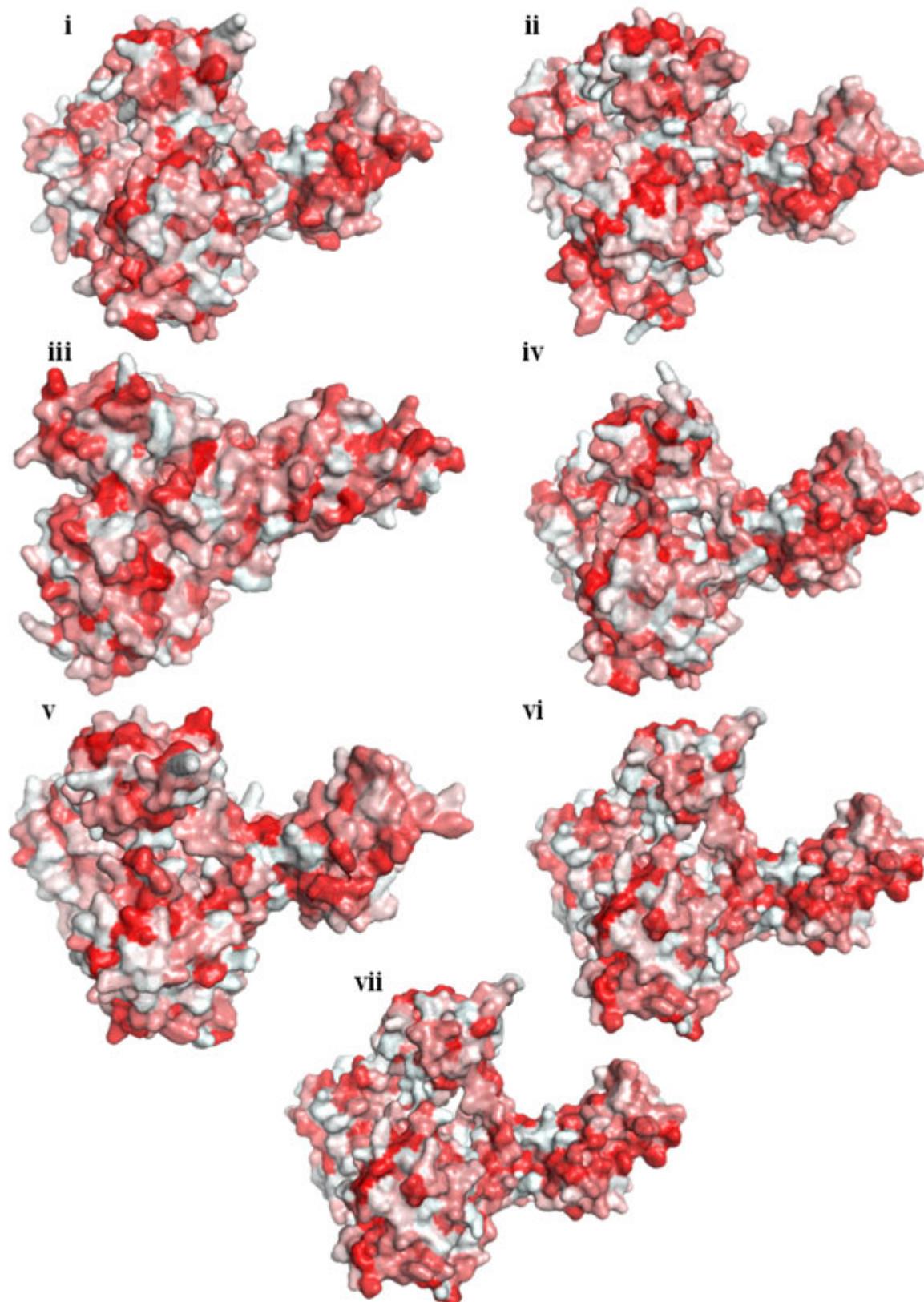


Figure S3(d). Hydrophobicity map of NSP13 protein surface of all 7 types of corona viruses (i) 229E, (ii) HKU1 (iii) MERS CoV (iv) NL63 (v) OC43 (vi) SARS CoV and (vii) SARS CoV2. Red colour regions indicate hydrophobicity and white colour regions indicate

hydrophilicity

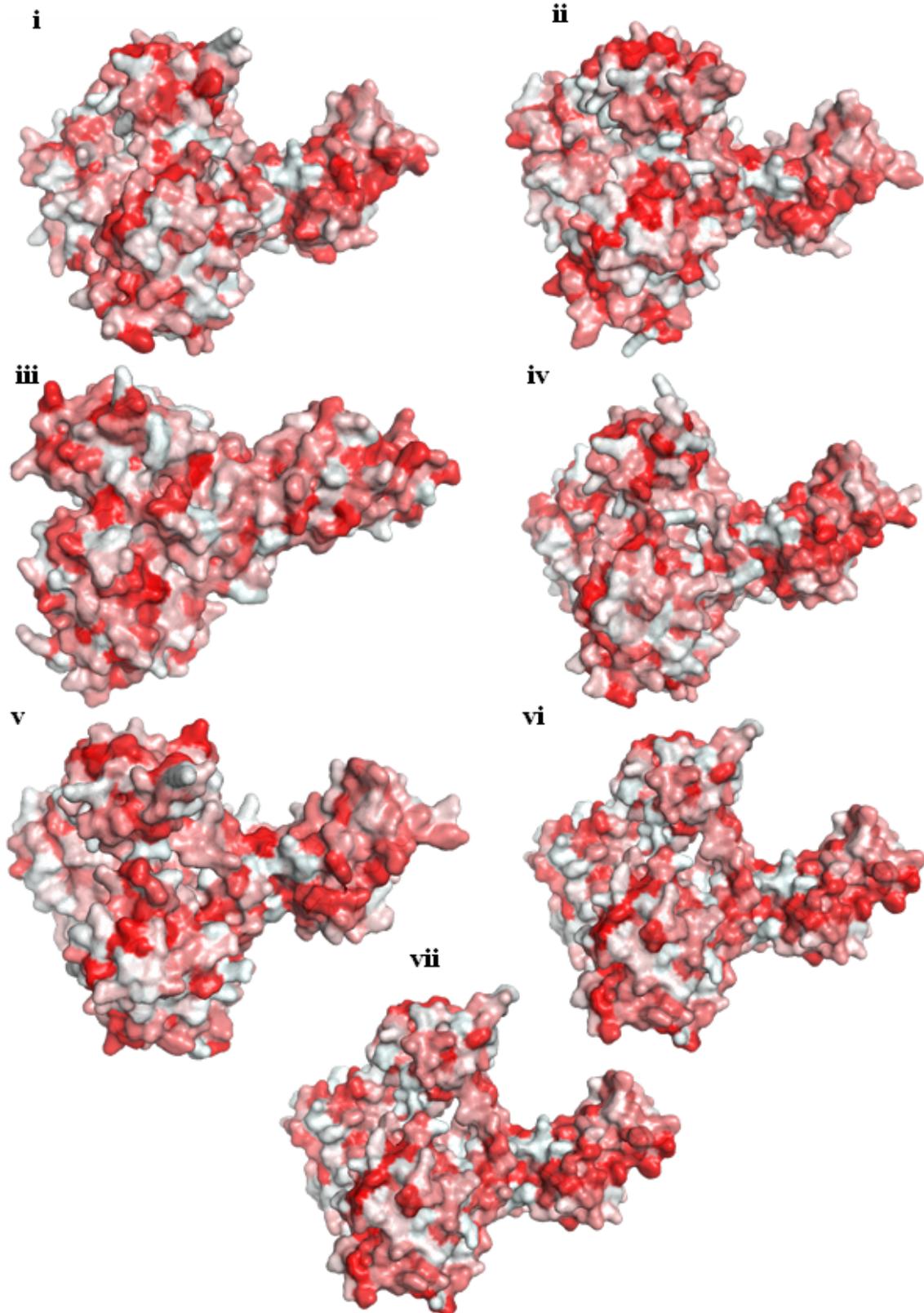


Figure S3(e). Hydrophobicity map of NSP 14 protein surface of all 7 types of corona viruses (i) 229E, (ii) HKU1 (iii) MERS CoV (iv)

NL63 (v) OC43 (vi) SARS CoV (vii) SARS CoV2. Red colour regions indicate hydrophobicity and white colour regions indicate hydrophilicity.

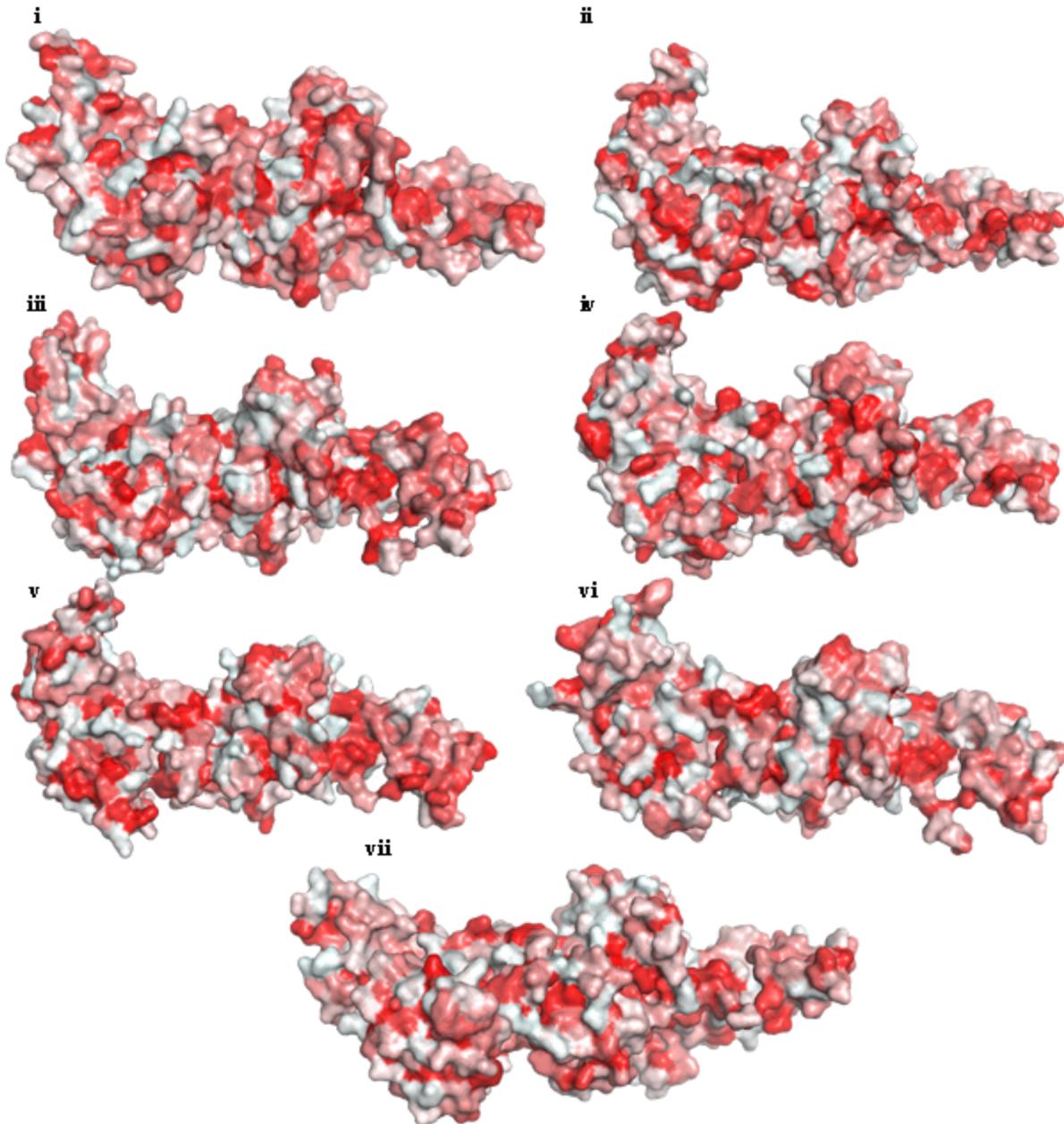


Figure S3(f). Hydrophobicity map of NSP 15 protein surface of all 7 types of corona viruses (i) 229E, (ii) HKU1 (iii) MERS CoV (iv) NL63 (v) OC43 (vi) SARS CoV and (vii) SARS CoV2. Red colour regions indicate hydrophobicity and white colour regions indicate hydrophilicity.

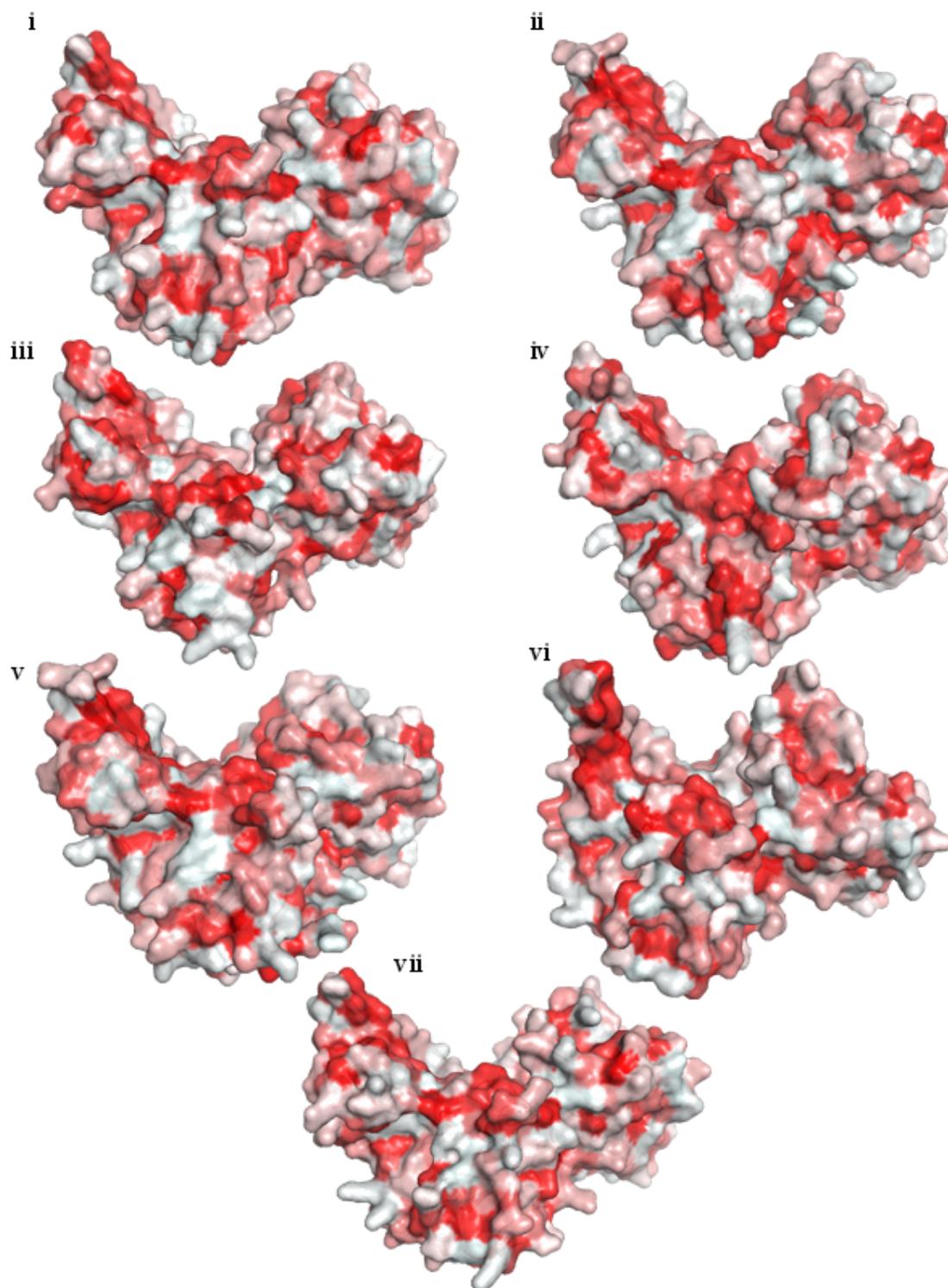


Figure S3(g). Hydrophobicity map of NSP 16 protein surface of all 7 types of corona viruses (i) 229E, (ii) HKU1 (iii) MERS CoV (iv) NL63 (v) OC43 (vi) SARS CoV and (vii) SARS CoV2. Red colour regions indicate hydrophobicity and white colour regions indicate hydrophilicity.

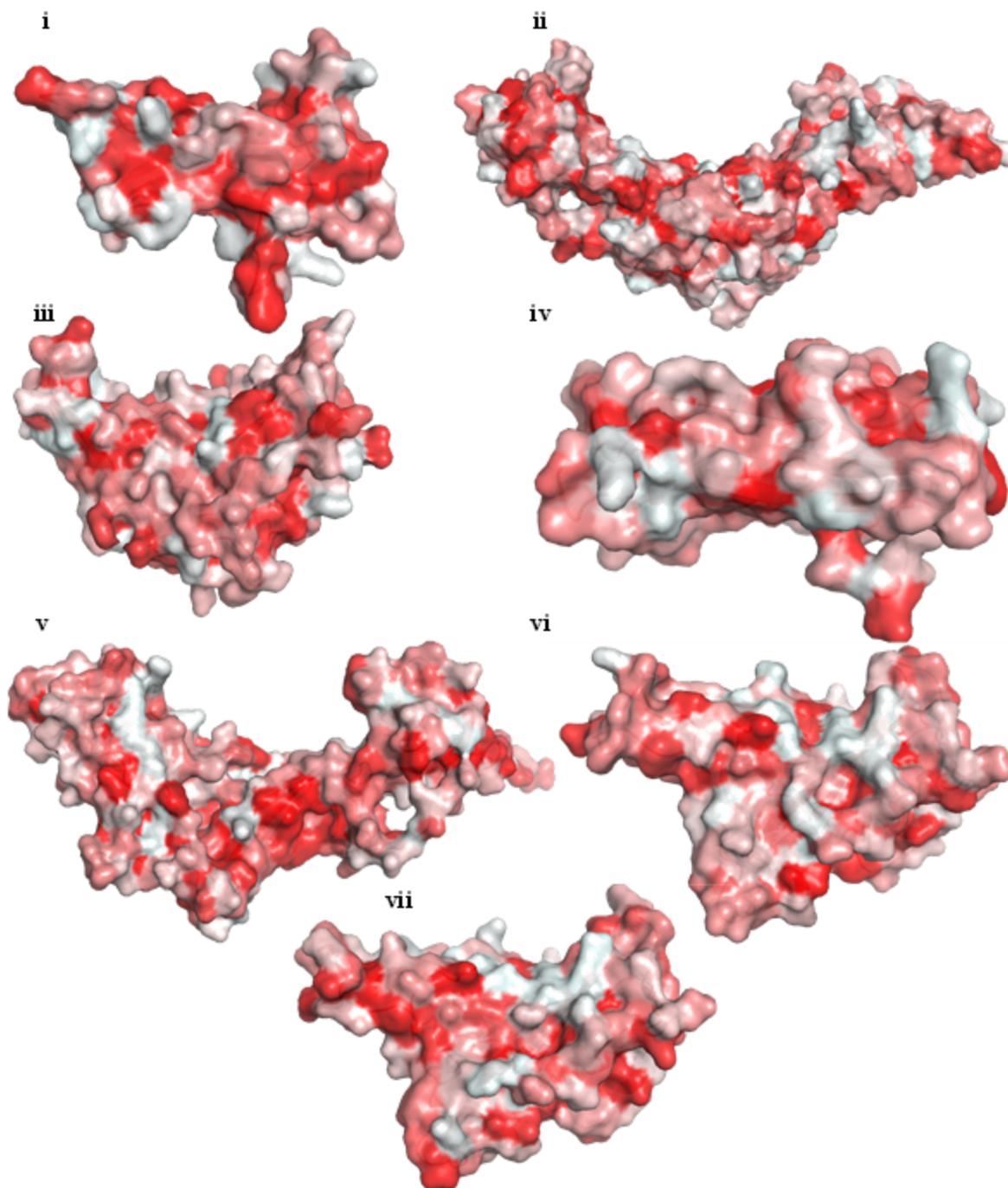
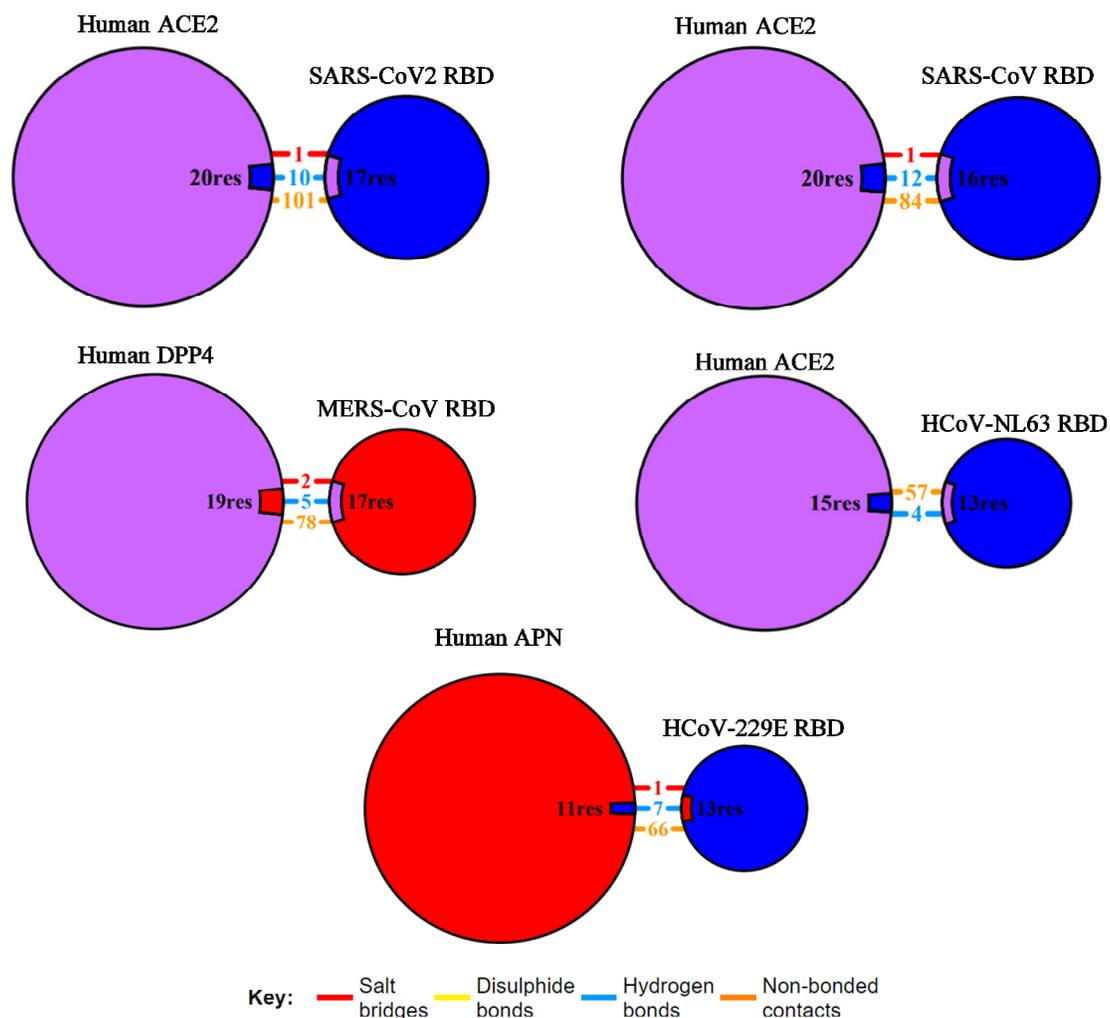


Figure S3(h): Hydrophobicity map of Spike protein surface of all 7 types of corona viruses (i) 229E, (ii) HKU1 (iii) MERS CoV (iv) NL63 (v) OC43 (vi) SARS CoV and (vii) SARS CoV2. Red colour regions indicate hydrophobicity and white colour regions indicate hydrophilicity



*ACE2 (Angiotensin-converting enzyme 2), APN (Aminopeptidase N), DPP4 (Dipeptidyl peptidase 4)

Figure S4. The protein-protein interface between the viral spike receptor binding domain and host proteins. The interactions are represented for SARS-CoV2 with ACE2, SARS-CoV with ACE2, MERS-CoV with DPP4, HCoV-NL63 with ACE2, HCoV-229E with APN. The diagrammatic representation indicates the interaction interface between the proteins such as salt bridges, hydrogen bonds and non-bonded contacts.