Molecular Modeling 2020
Lecture 24

SARS-2 lifecycle. ACE2 in mammals. Searching for an animal model.
MOE tools for MSAs

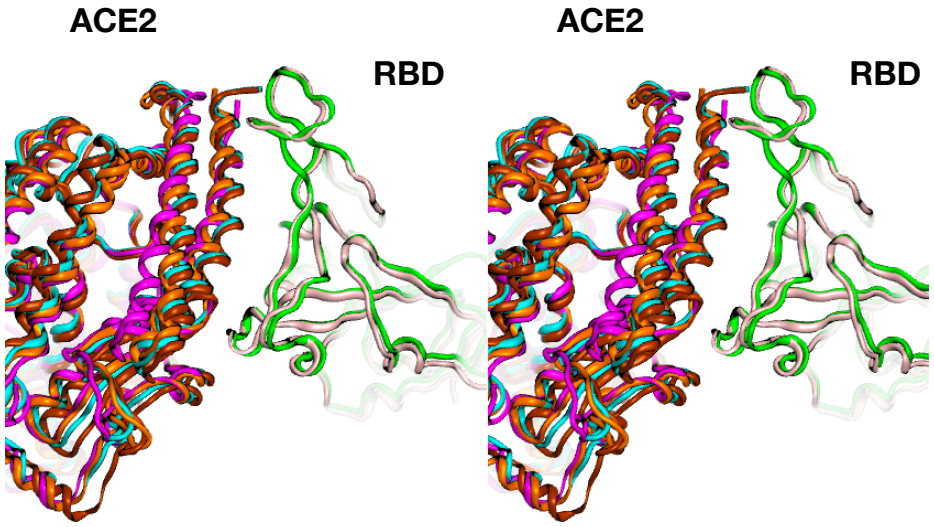
Youtube video on SARS-2

- Find link on course web page

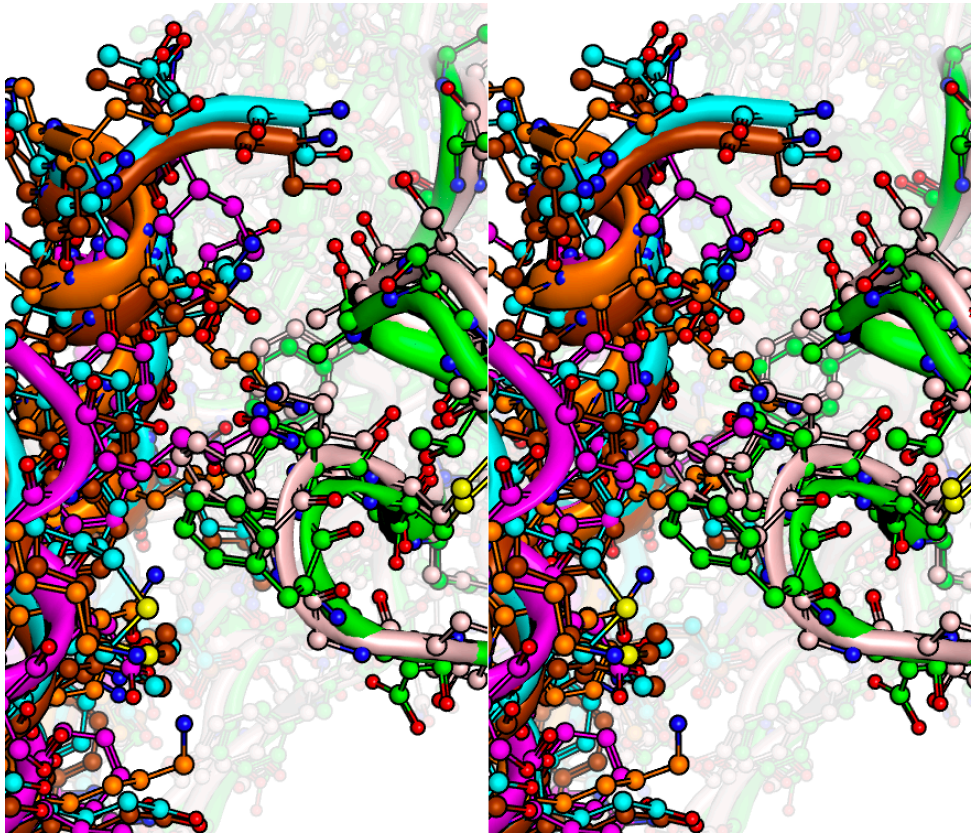
What mammalian species can be infected by SARS-2?

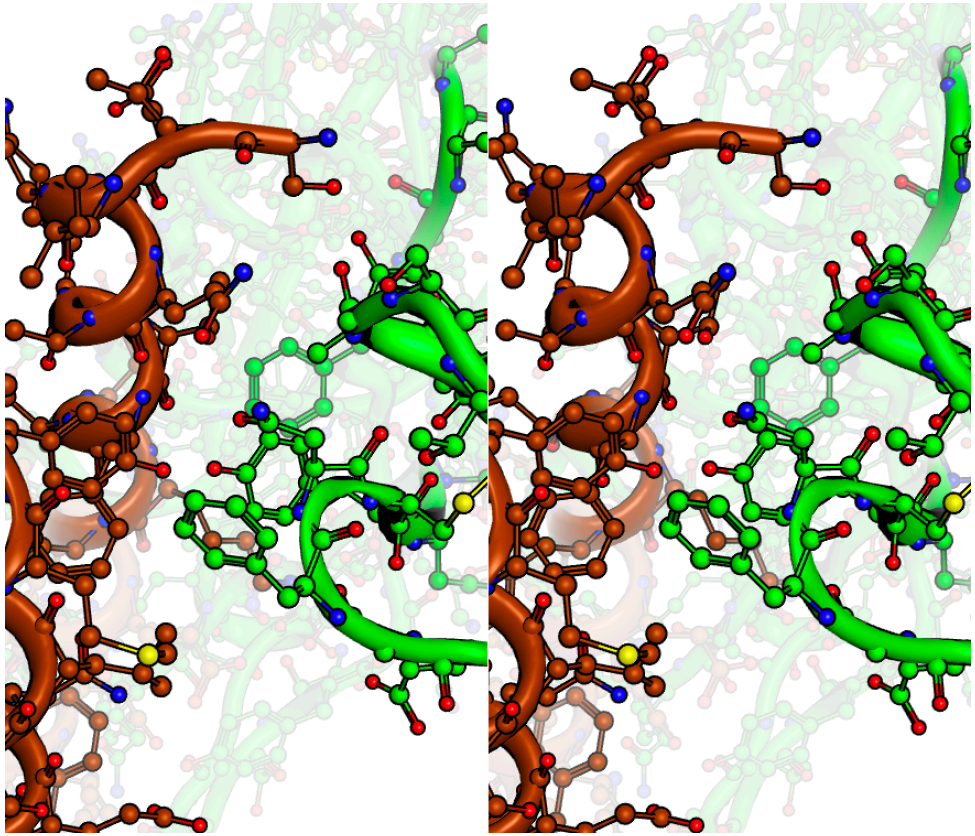
Looking at the ACE2·RBD interface

ACE2 = angiotensin converting enzyme 2, the cellular receptor for SARS-2
RBD = receptor binding domain of the SARS-2 spike glycoprotein

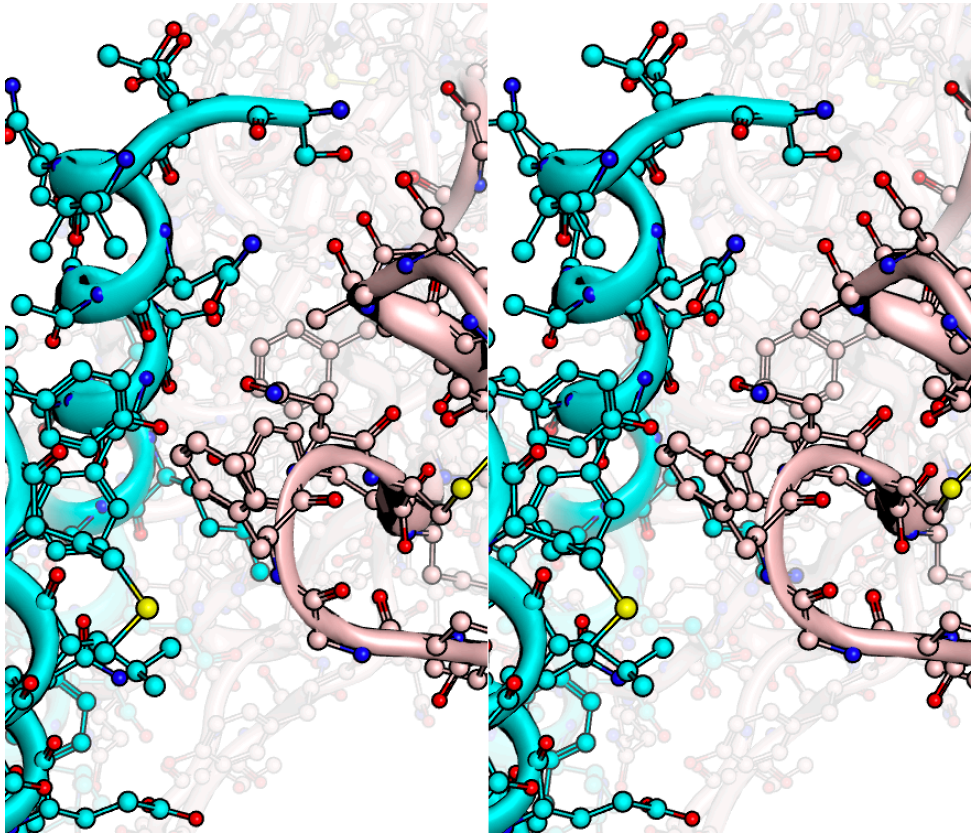


Multiple copies , structures of ACE2 and multiple copies of RBD are in basic agreement.

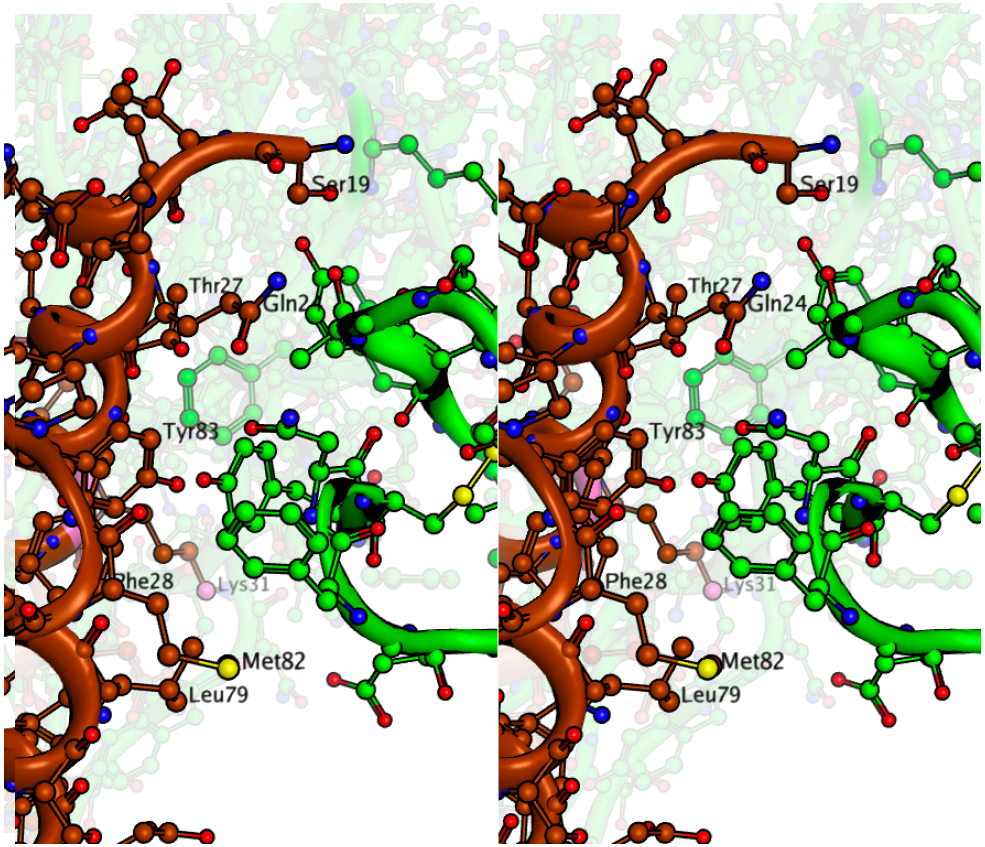




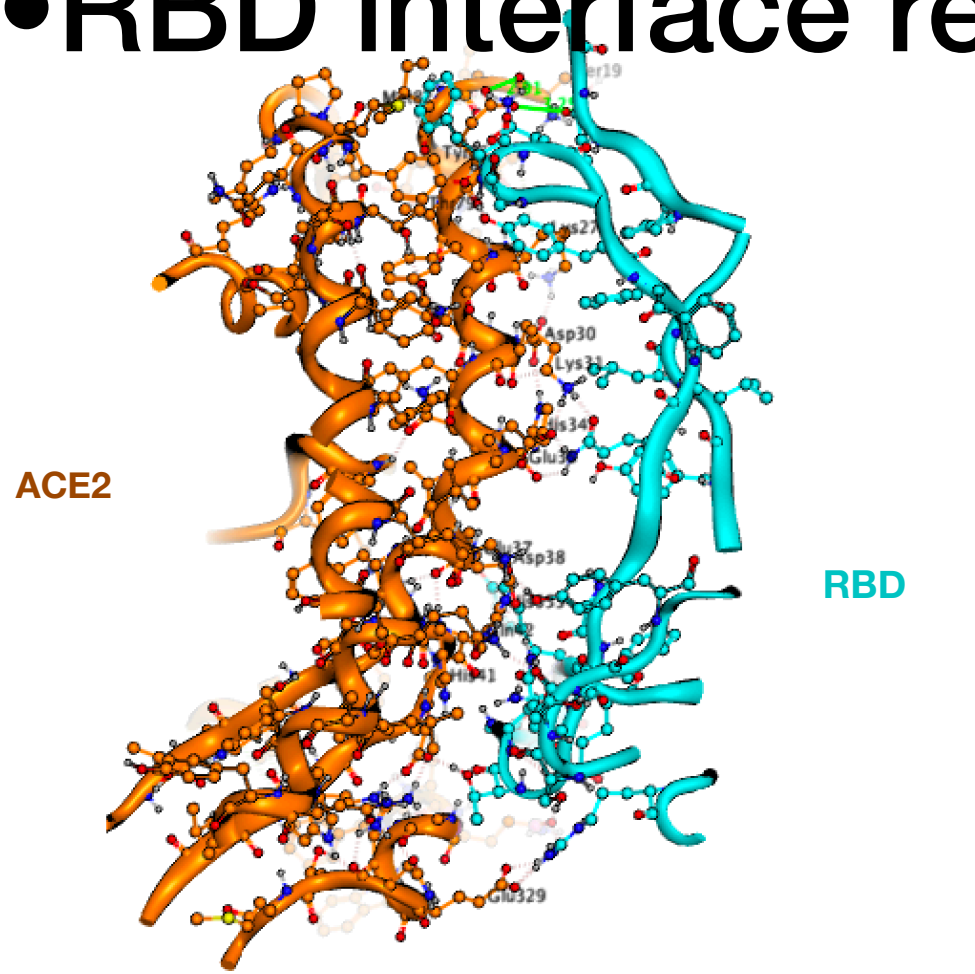
6VW1_BF



6VW1_AE


Ser19
Thr27
Gln2
Tyr83
Phe28
Lys31
Met82
Leu79
Ser19
Thr27
Gln24
Tyr83
Phe28
Lys31
Met82
Leu79

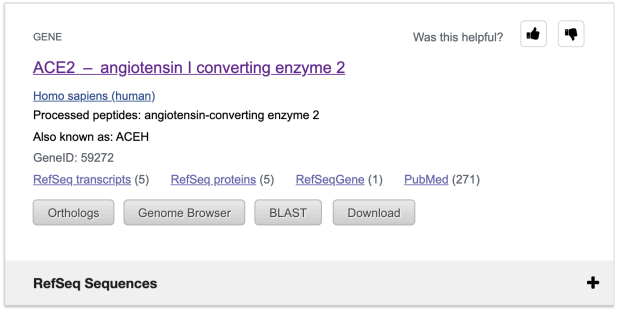
ACE2•RBD interface residues



Exercise 24.1 Phyogenetic analysis of a protein-protein interface

We are asking: Which species will favor this ACE2•RBD interface? Which species will not?

1. Search NCBI Proteins for "ACE2 human". See figure. Select Orthologs. Search mammals. Choose a subgroup of placentals, such as rodents or carnivores. Select individual species or all. Download RefSeq Proteins (fasta). Open in MOE.



2. Open 6vw1 (ACE2•RBD). Align all sequences, except RBD.

3. Select all interface residues on ACE2 by picking. SiteView it.

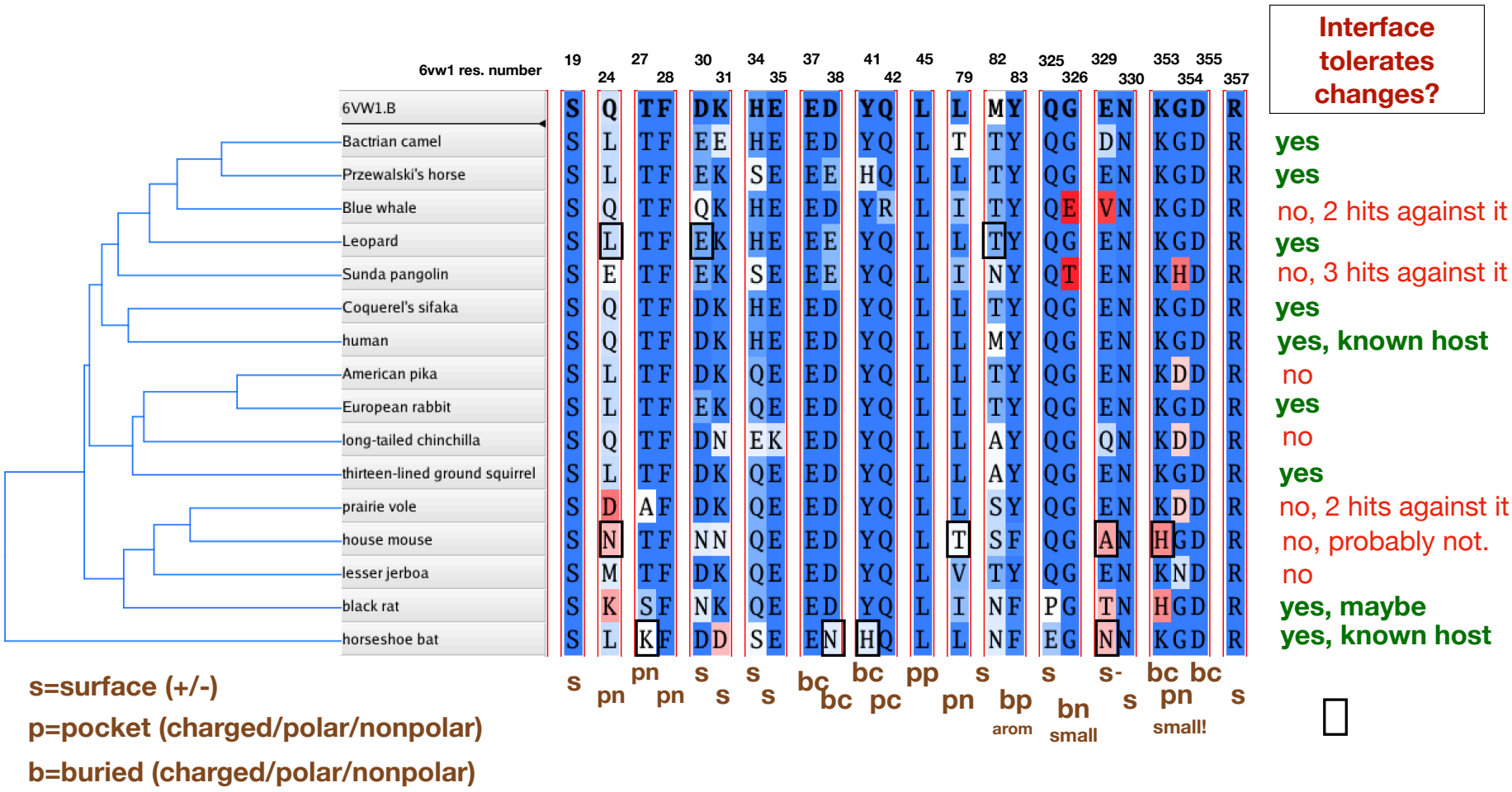
5. With interface residues still selected, **SEQ: Ruler bar | right-mouse | hide | unselected.**
(now you see all interface residues only, all species)

3. SEQ: Select all chains of ACE2. Select all interface residues. **SEQ: Alignment | similarity, tree**

4. Unselect (don't delete) redundant species. Keep interesting ones. (The tree updates)

5. Select each residue in SEQ. Go to MOE to inspect it. As you do, ask: Can the interface tolerate the mutations? For which species?

Interface residues of representative mammals



Interface tolerates changes?

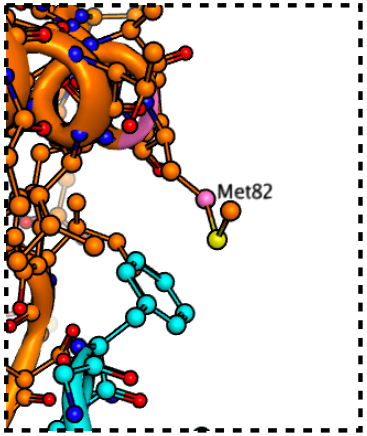
- yes
- yes
- no, 2 hits against it
- yes
- no, 3 hits against it
- yes
- yes, known host
- no
- yes
- no
- yes
- no, 2 hits against it
- no, probably not.
- no
- yes, maybe
- yes, known host

s=surface (+/-)

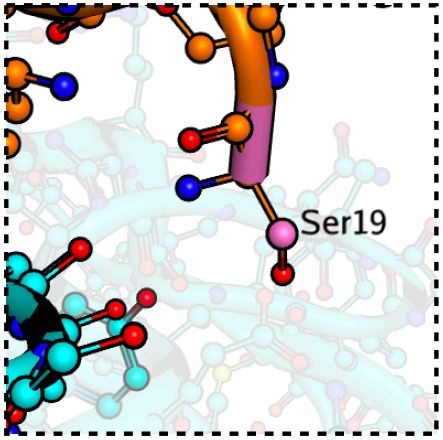
p=pocket (charged/polar/nonpolar)

b=buried (charged/polar/nonpolar)

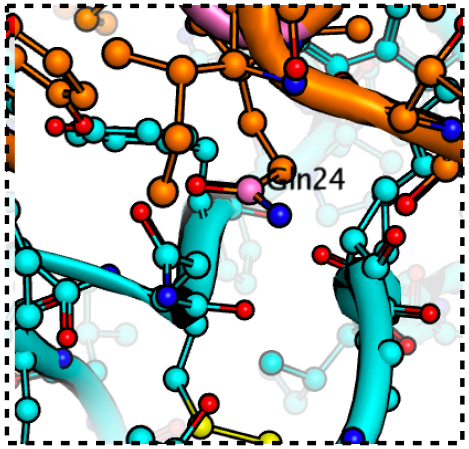
Interface residues on ACE2



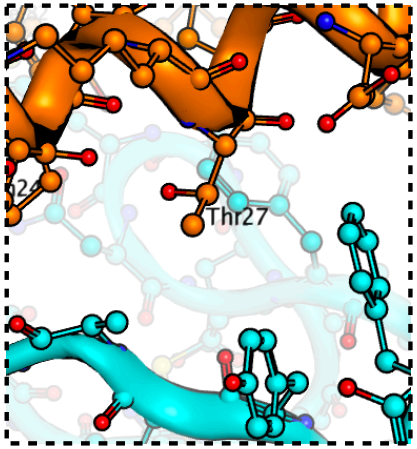
M82 is [MTAS] in mammals, is also not doing much in the structure. So anything goes, but M<A<T,S



S19 is highly exposed. Anything goes here.

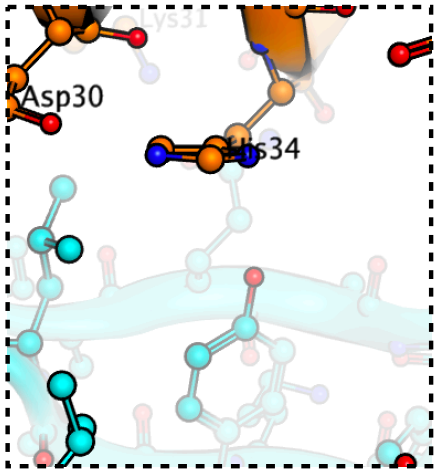


Q24 is partially buried. [LDMQ] in mammals. Probably wouldn't like D (*prairie vole*) because of the absence of a cation nearby.

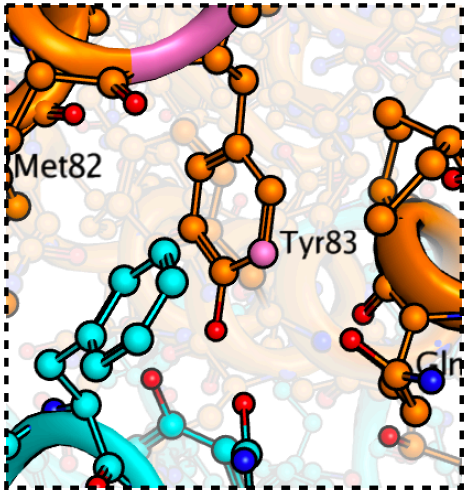


T27 is buried. Can't be bigger than a T. Only [TA] appear. Ala is fine here.

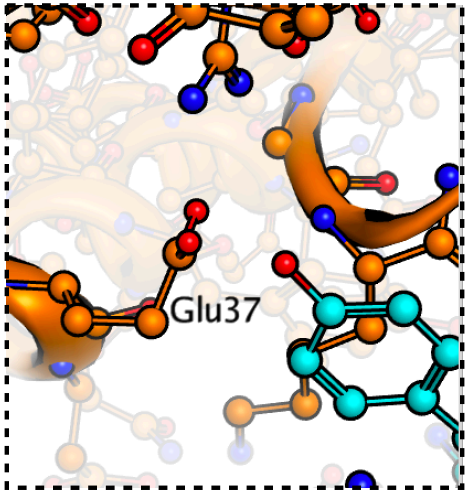
Interface residues on ACE2



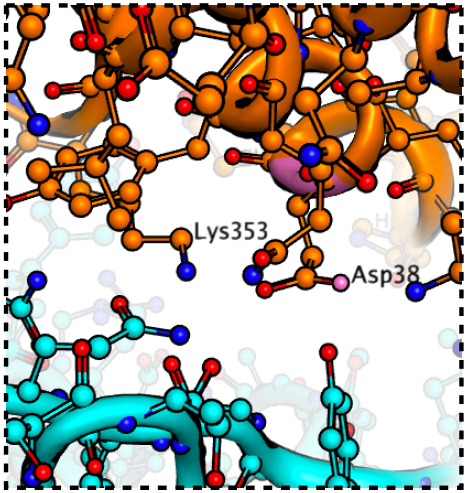
H34 is [HQSY] in mammals. All of those sidechains fit in the wide pocket.



Y83 makes key interactions with RBD, but is strictly conserved in mammals. Does not help us.

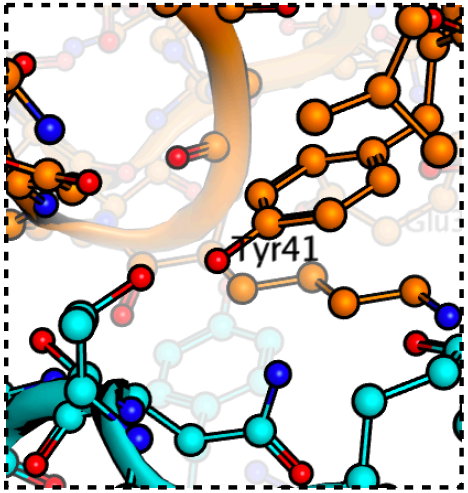


E37 makes 1 H-bond with RBD, and is strictly conserved in mammals. Does not help us.

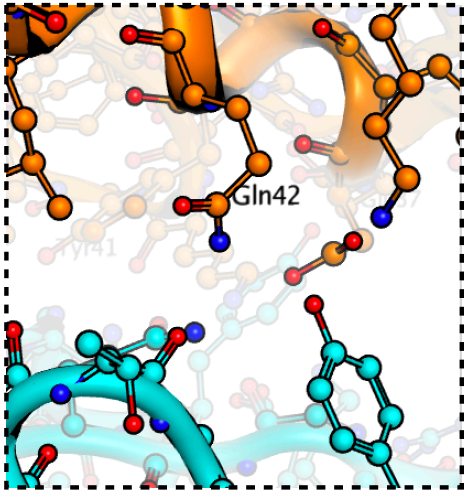


D38 teams up with **K353** to make a H-bond network with RBD. Mutations here could matter. **E** in *horse, cat* is too long, chooses a different rotamer. Binding lost? **K353** is **N** in *tarsier*. H-bonds lost.

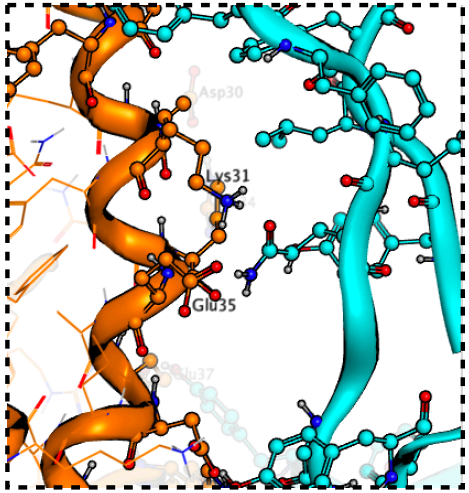
Interface residues on ACE2



Y41 is [H] in *marmoset, horse, tarsier*. Very tight, buried. **H** would fit.

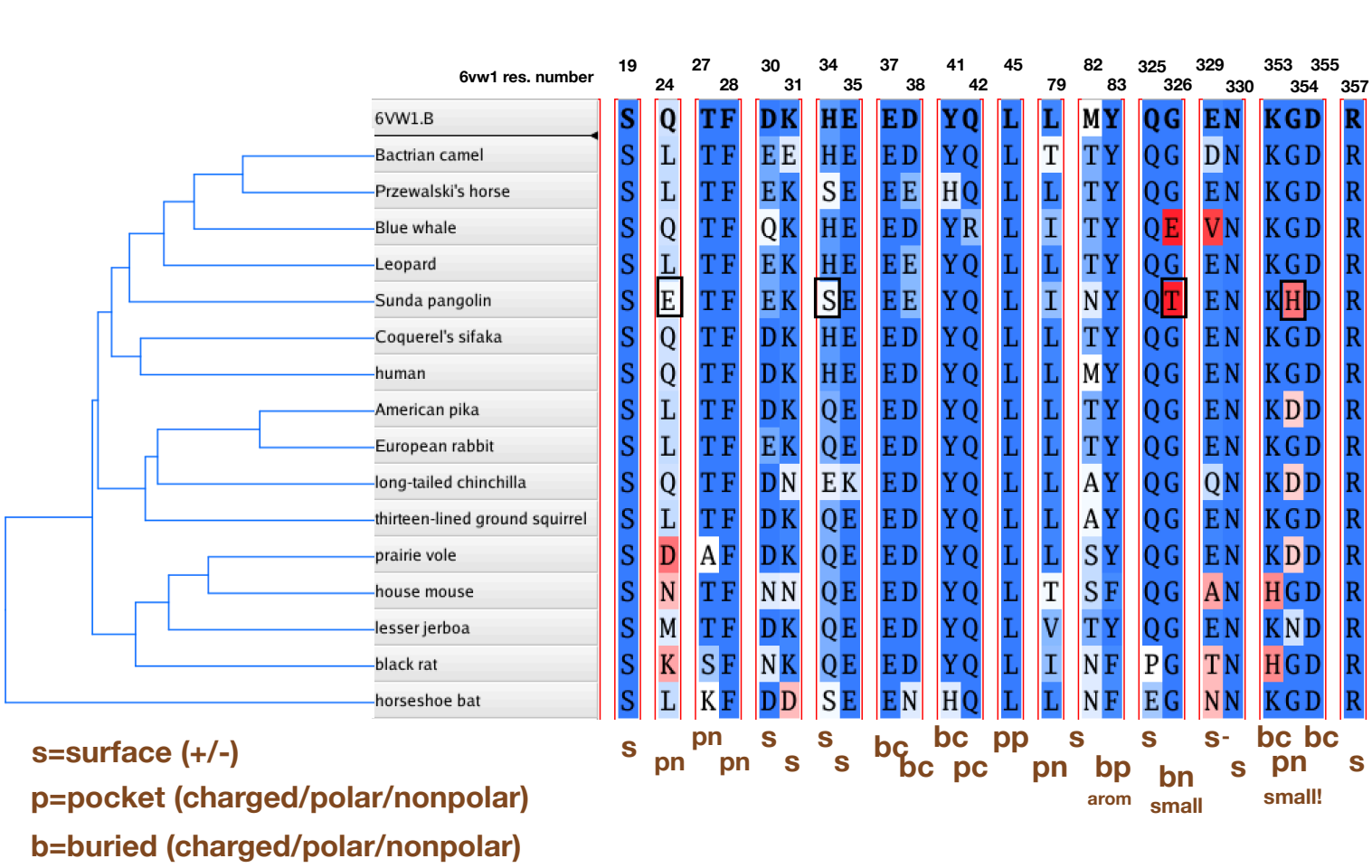


Q42 is [**E**] in *marmoset*. Probably unfavorable due to local negative charges.



K31 and **E35** combine to form a H-bond network. Both are strictly conserved in near-non-primate mammals.

Interface residues of representative mammals



Interface tolerates changes?

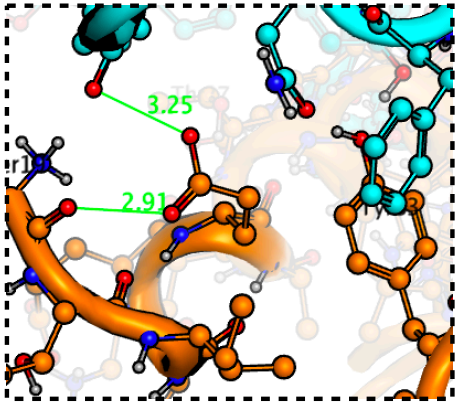
- yes
- yes
- no, 2 hits against it
- yes
- no, 3 hits against it
- yes
- yes, known host
- no
- yes
- no
- yes
- no, 2 hits against it
- no, probably not.
- no
- yes, maybe
- yes, known host

s=surface (+/-)

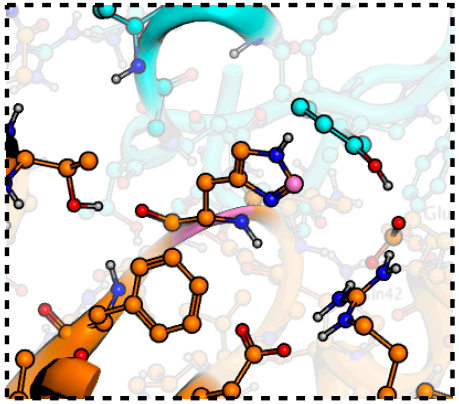
p=pocket (charged/polar/nonpolar)

b=buried (charged/polar/nonpolar)

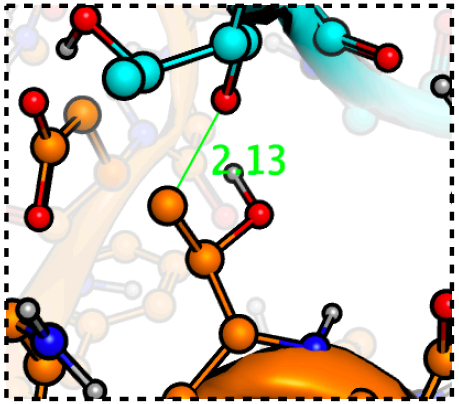
Pangolin , 3 hits against.



Q24E has too many H-bond acceptors, few H-bond donors.

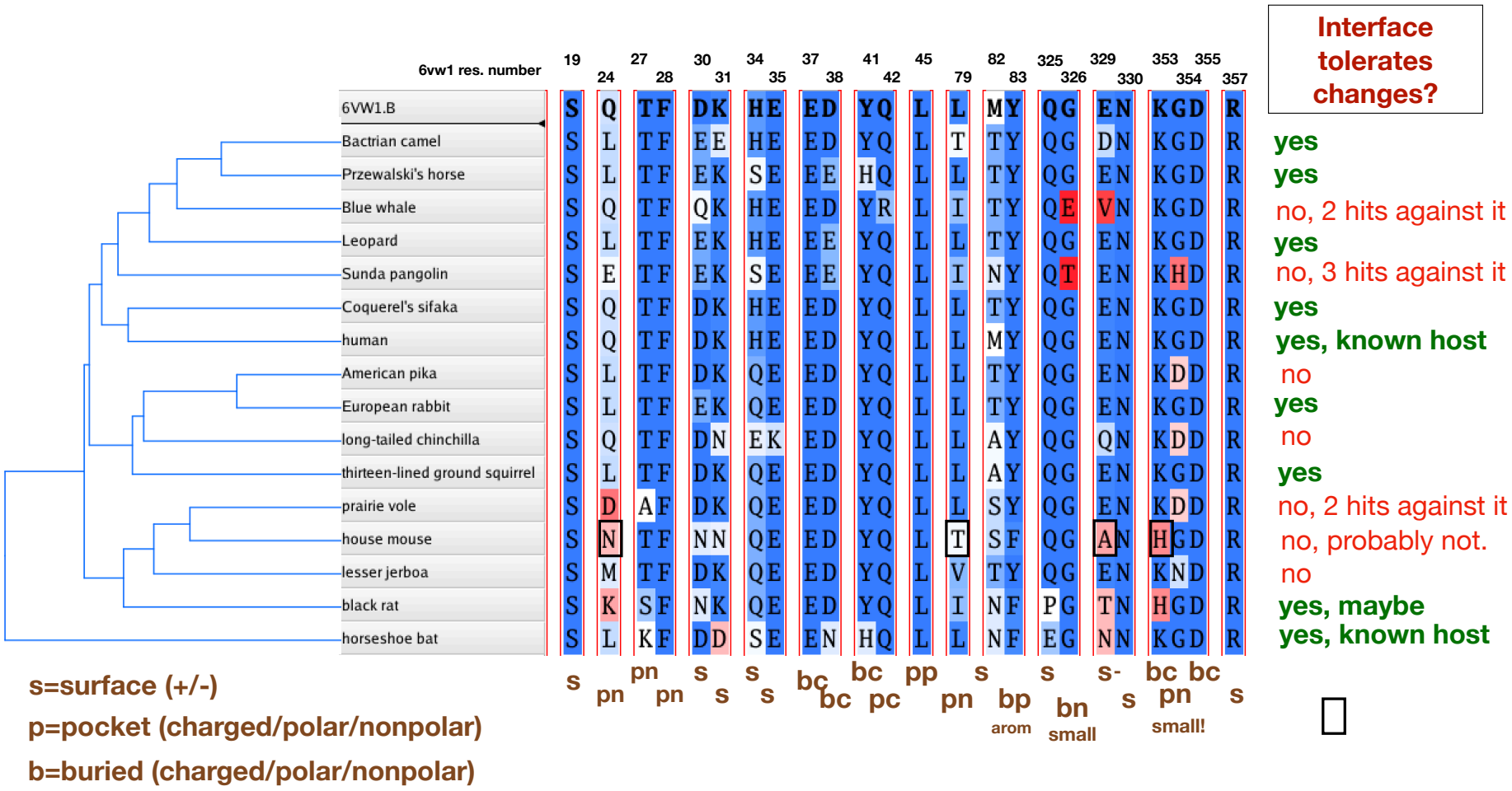


G354H is too close to Arg.



G326T creates unresolvable collisions in a tight place.

Interface residues of representative mammals



Interface tolerates changes?

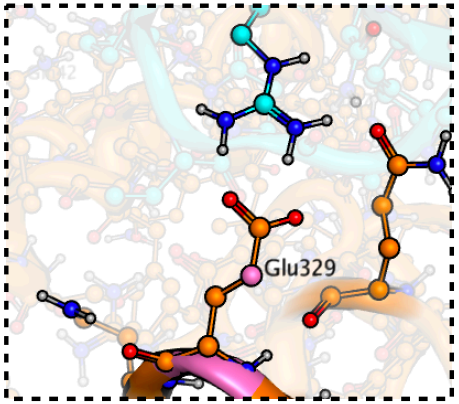
- yes
- yes
- no, 2 hits against it
- yes
- no, 3 hits against it
- yes
- yes, known host
- no
- yes
- no
- yes
- no, 2 hits against it
- no, probably not.
- no
- yes, maybe
- yes, known host

s=surface (+/-)

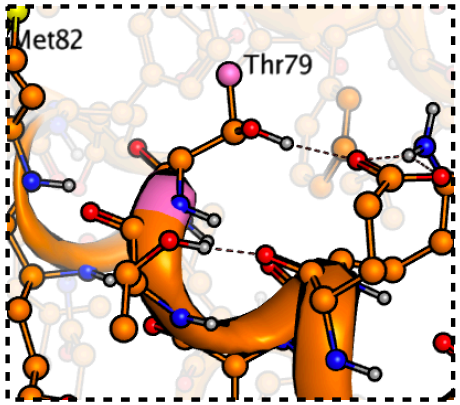
p=pocket (charged/polar/nonpolar)

b=buried (charged/polar/nonpolar)

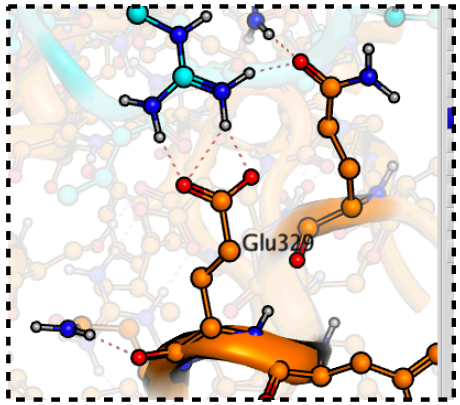
Mouse, probably not a host.



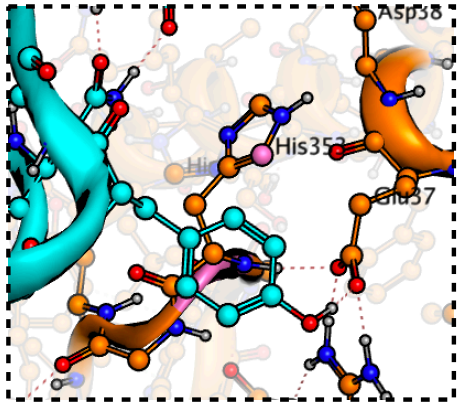
E329A loses a surface salt bridge.



L79T loses hydrophobic interaction, but packs well.

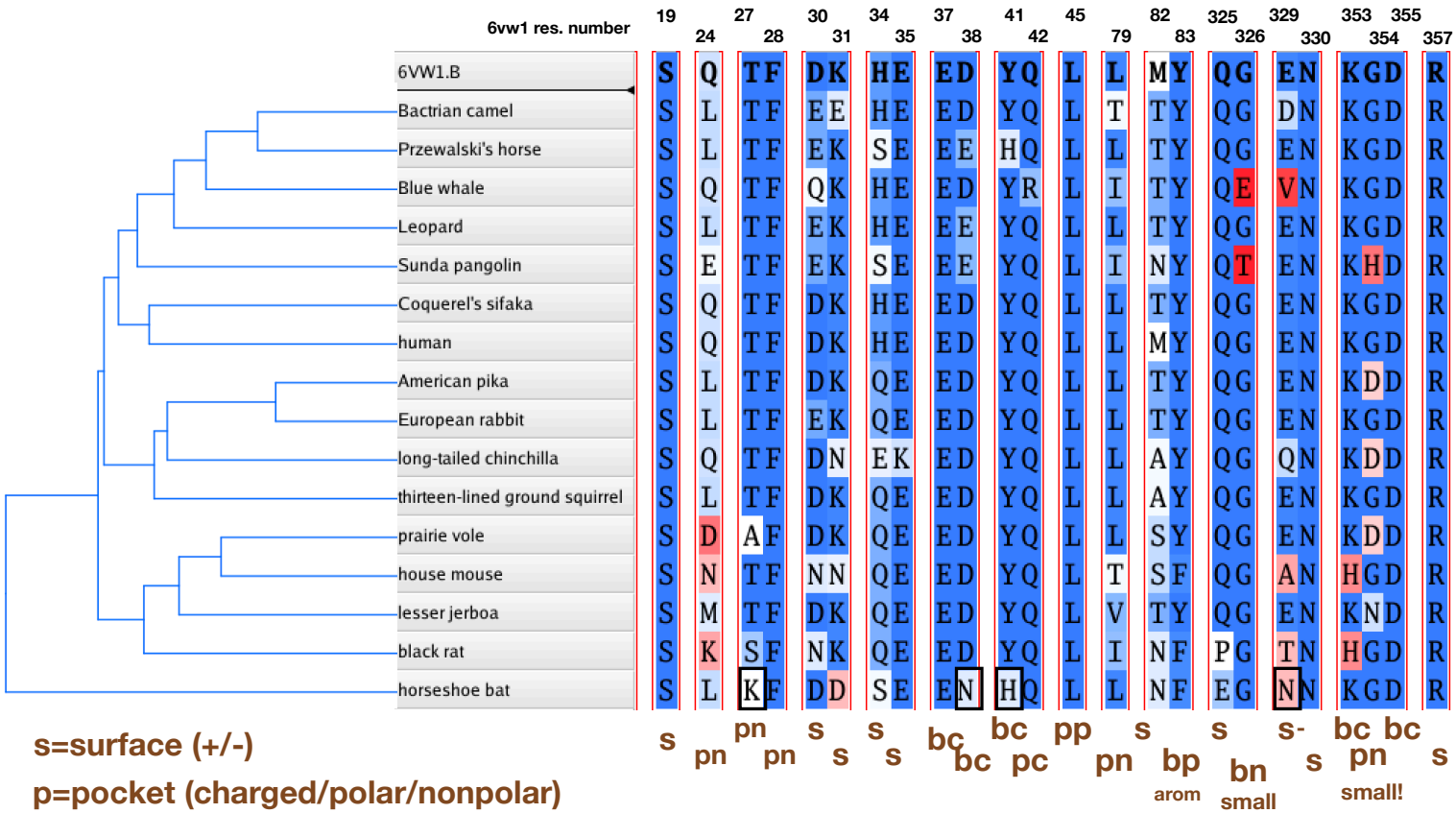


E329A loses a salt bridge.



K353H packs well, but H-bonds are not satisfied.

Interface residues of representative mammals



Good fit?

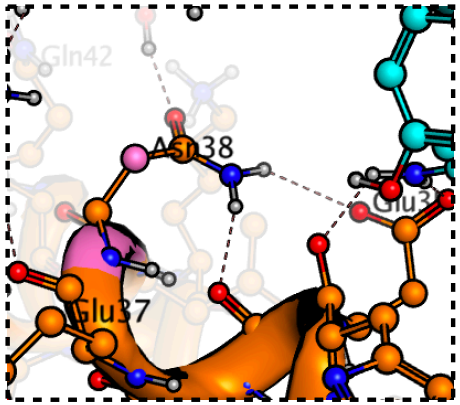
yes
yes
no, 2 hits against it
yes
no, 3 hits against it
yes
yes
no
yes
no
yes
no, 2 hits against it
no
no
no, 1 hit
yes

s=surface (+/-)

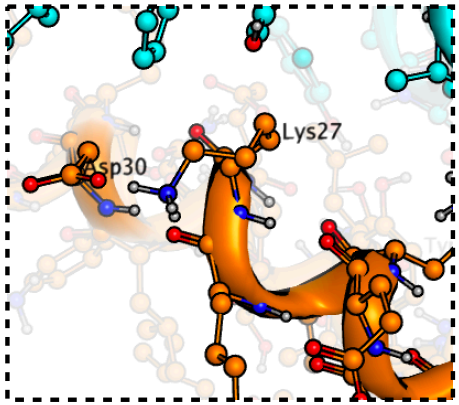
p=pocket (charged/polar/nonpolar)

b=buried (charged/polar/nonpolar)

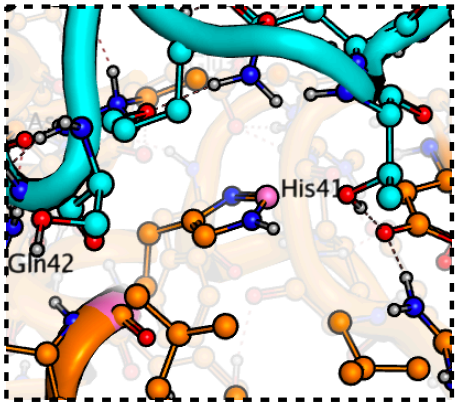
Horseshoe bat, true host for SARS-2



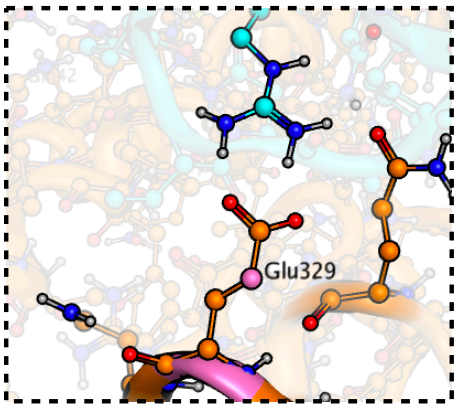
D38N makes better H-bonds with a better rotamer.



T27K actually adds a salt bridge.



Y41H accomodates the smaller side chain and leaves room for waters.

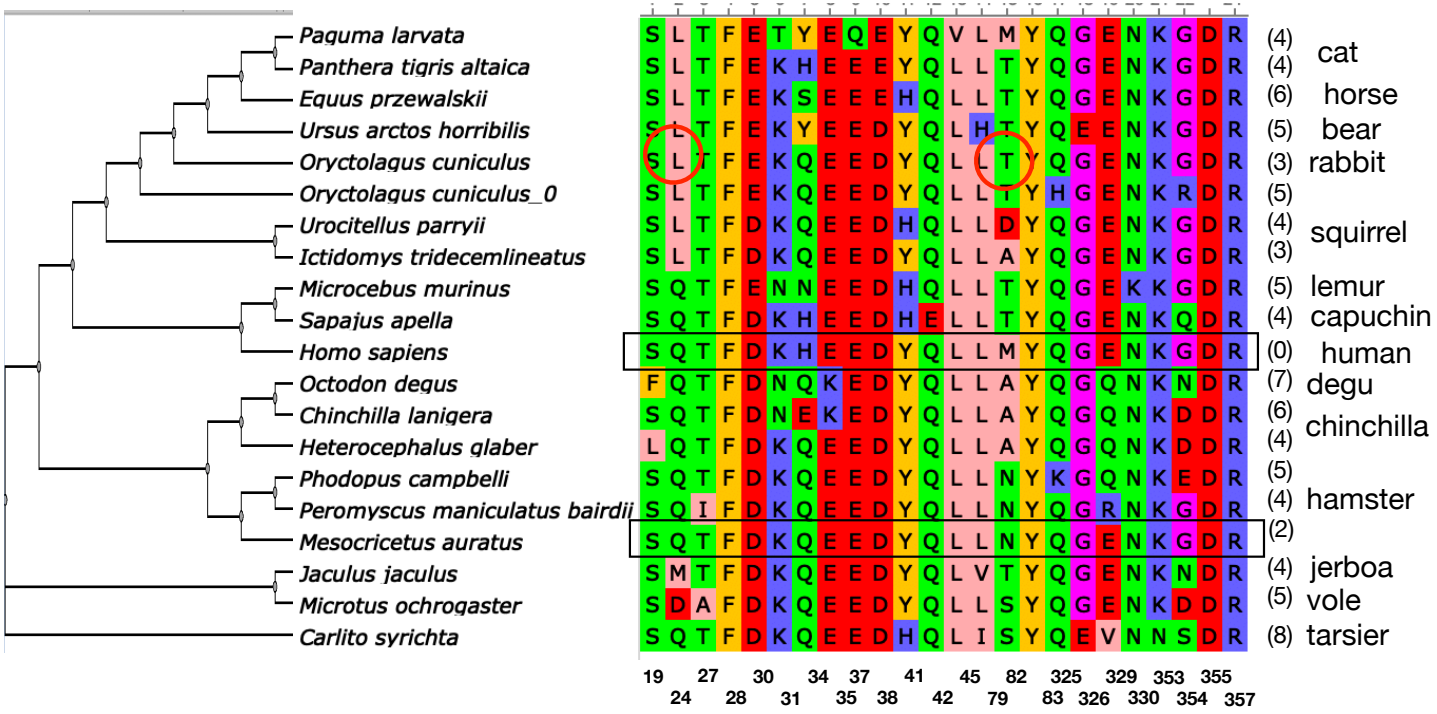


E329N loses a surface salt bridge.

Bat ACE2 has lots of mutations relative to human, but many of these are acceptable or favorable to RBD binding.

Other mammals

animal model search

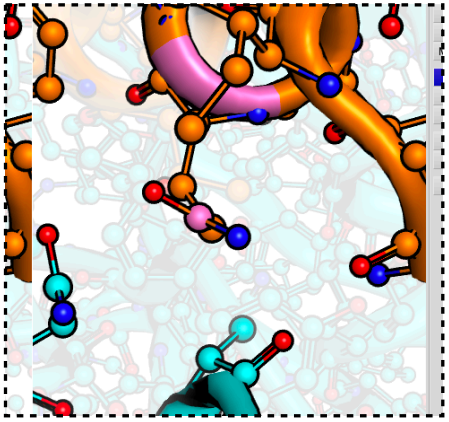


UGENE alignment

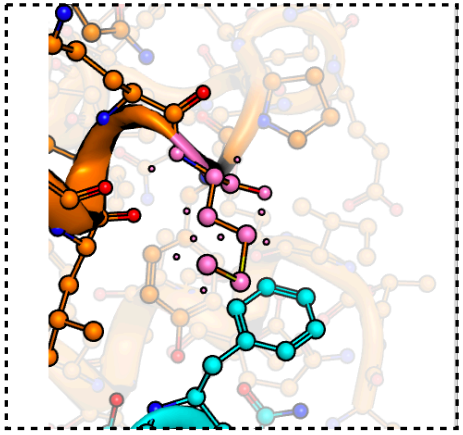
Maybe golden hamster is a good model.

Prediction: Civit ACE2 may bind RBD

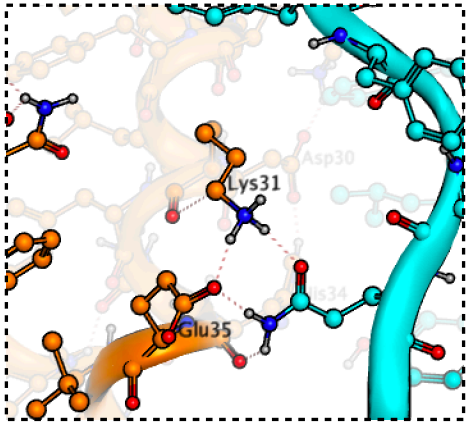
homo sapiens	S	Q	T	F	D	K	H	E	E	D	Y	Q	L	L	M	Y	Q	G	E	N	K	G	D	R
paguma larvata	S	L	T	F	E	T	Y	E	Q	E	Y	Q	V	L	M	Y	Q	G	E	N	K	G	D	R



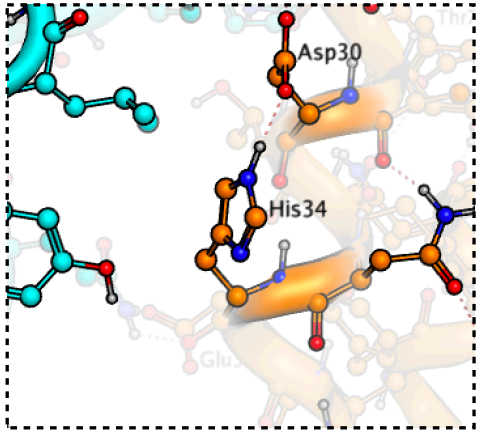
Q24L is in a hydrophobic pocket.



M82T is on the surface.



K31T is in a deep pocket. Smaller side chain leaves space for waters.

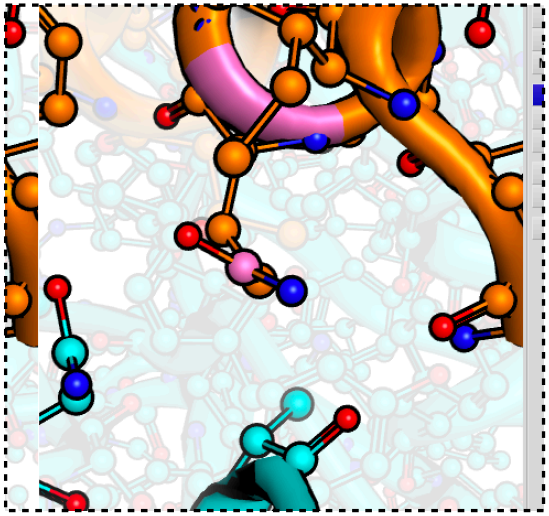


H34Y breaks one H-bond.

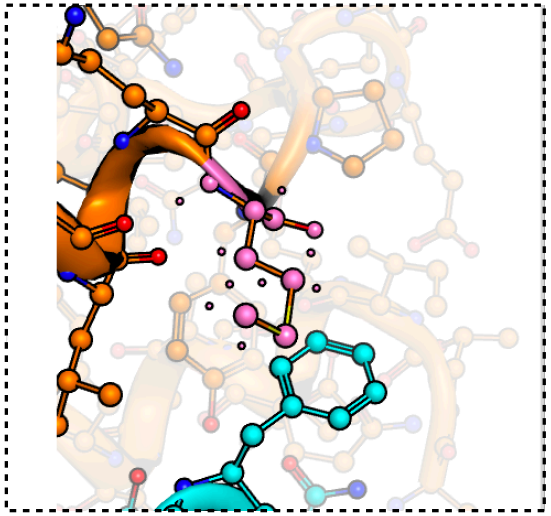
Some loss of hydrophobic effect. But no strong negatives. Probably binds RBD.

Prediction: Tiger ACE2 will bind RBD

homo sapiens	S	Q	T	F	D	K	H	E	E	D	Y	Q	L	L	M	Y	Q	G	E	N	K	G	D	R
Panthera tigris	S	L	T	F	E	K	H	E	E	E	Y	Q	L	L	T	Y	Q	G	E	N	K	G	D	R



Q24L fits a hydrophobic pocket well.



M82T is on the surface.

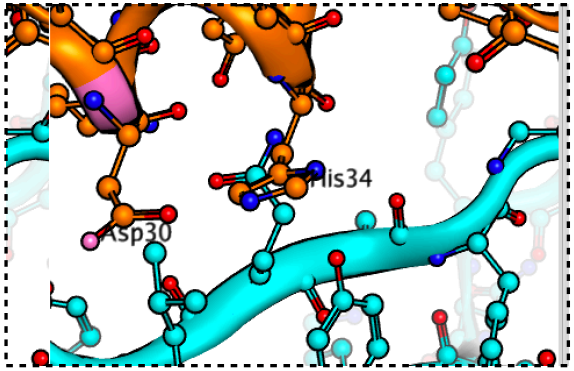
Only 2 meaningful mutations in the interface region!

Both of the mutations are energetically acceptable.

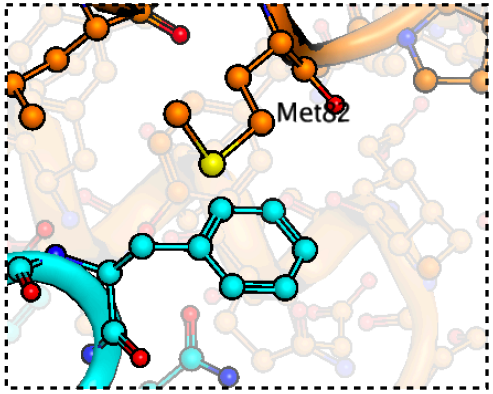
A tiger has tested positive.

Prediction: Hamster ACE2 will bind RBD

homo sapiens S Q T F D K H E E D Y Q L L M Y Q G E N K G D R
miscricetus auratus S Q T F D K Q E E D Y Q L L N Y Q G E N K G D R



H34Q fits in a polar pocket.



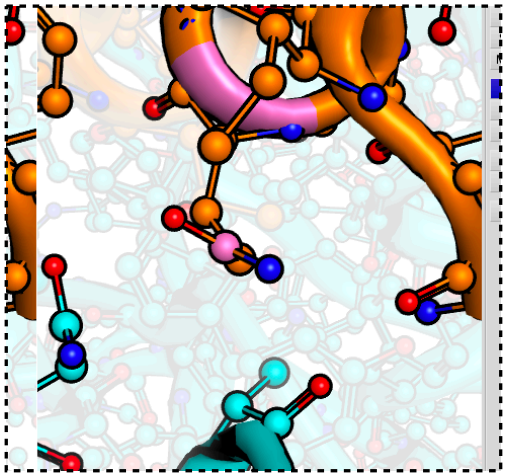
M82N is exposed to solvent and inconsequential.

Both of the mutations are energetically acceptable

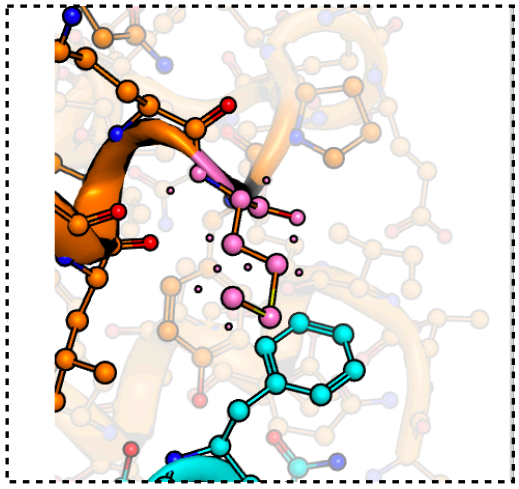
Prediction: Squirrel ACE2 will bind RBD

homo sapiens S Q T F D K H E E D Y Q L L M Y Q G E N K G D R

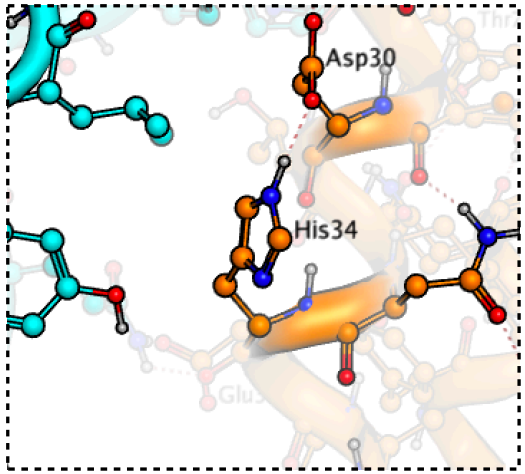
ictidomys tridecemliniatus S L T F D K Q E E D Y Q L L A Y Q G E N K G D R



Q24L is in a hydrophobic pocket



M82A is on the surface



H34Q is acceptable.

All three mutations are energetically acceptable when modeled on 6VW1

What coronaviral species can infect humans?

If you want to know, repeat the process from Exercise 24.1
Then isolate the RBD residues at the interface. Which viruses have mutations that the interface will tolerate?

term project

Docking and Design 1

- Dock using ribbons. Make sure 3-fold symmetry is maintained.
- Refine docked pose with backbone atoms only. Make sure CA-CA distances are at least 4Å, 3.5Å if one of them is a glycine.
- Identify interface residues on one monomer of the ligand by picking. Optionally hide other columns or color those residues in SEQ window.
- Extend selection to near residues twice. Unfix. Select | invert. Fix. Now backbone and sidechains are unfixed only in the interface region.
- Turn on gizmin. For each interface residue, mutate (Protein | Protein builder). Try to remove clashes, maintain good geometry, fill space loosely, satisfy charges and H-bonds.
- Add waters where space and H-bond partners allow.

Docking and Design 2

- When you can't find any more mutations, stop mutating. Turn off gizmin.
- Group the designed monomer and all of its waters under one **tag**.
- Copy the tag twice. Name the new tags "copy 2" and "copy 3". Align sequences.
- Superpose "copy 2" on the second monomer in the template trimer (Unfix "copy 2". Fix template second monomer. Set all other chains to "i", ignore.Superpose, moving all atoms in tag together.). Superpose "copy 3" on the third monomer. Delete the original 2nd and 3rd monomers. Now you have made the compete trimer of the designed monomer.
- Go through the structure and energy minimize locally, using EPUSEIPF. When you have finshed, select the whole trimer with its waters and extend the selection to near residues (the RBD). Unfix. Fix everything else. Perform a 100 ps molecular dynamics simulation at low temperature (t=50). Browse the results. Change to higher or lower T and repeat if necessary. Check geometry again. Use a surface to look for any new locations for waters. Finally, energy minimize.