Luleå University of Technology

Division of Chemical Engineering:
Biochemical Process Engineering

National Technical University of Athens

School of Chemical Engineering

DIPLOMA THESIS

Evaluation of the inhibitory effect of compounds from halophyte *Salicornia* sp. extract against the SARS-CoV-2 main protease via molecular docking simulations and an *in vitro* assay

Διερεύνηση της ανασταλτικής δράσης συστατικών του εκχυλίσματος από το αλόφυτο *Salicornia* sp. ενάντια στη κύρια πρωτεάση του ιού SARS-CoV-2 μέσω μοριακής πρόσδεσης και *in vitro* δοκιμής

Eleftheria Sapountzaki

Host university supervisor: Prof. Paul Christakopoulos

Home university supervisor: Assoc. Prof. Evangelos Topakas

Supervisor: Assoc. Senior lecturer Io Antonopoulou

Athens, 2021

Acknowledgements

This thesis is the result of the work conducted in the Biochemical Process Engineering research group of Luleå University of Technology (LTU), in collaboration with the National Technical University of Athens through the Erasmus student exchange programme. The project took place between March and August of 2021. The master thesis was held under the Horizon 2020 project: 'AQUACOMBINE: Integrated on-farm Aquaponics systems for co-production of fish, halophyte vegetables, bioactive compounds, and bioenergy'.

To begin with, I would like to thank my home university supervisor, Assoc. Prof. Evangelos Topakas, for initiating my interest in the field of biotechnology and giving me the opportunity to conduct my thesis in LTU as an exchange student. In addition, I would like to express my sincere thanks to my host university supervisor Prof. Paul Christakopoulos, for accepting me into the group and assisting with everything needed throughout my stay in Sweden. Most important of all, I would like to express my deepest appreciation to my main supervisor, Assoc. Senior Lecturer Io Antonopoulou, who helped, guided, encouraged and trusted me throughout the whole process and supported me when any problem came about, lab-related or not. She taught me a lot of new things and her enthusiasm made me even more interested in research in the field of biotechnology. Also, I would particularly like to thank all the members of the laboratory, who created a very welcoming and pleasant environment and were always eager to help.

I am also very grateful to my friends from Greece, who supported me all these months, and to my new friends in Sweden, who, apart from helping me with anything I needed, made the whole experience unforgettable.

Last but not least, I would like to express my gratitude to my family for their constant love, support and encouragement without which this could not have been possible.

Eleftheria Sapountzaki,

Athens, 2021

Εκτεταμένη περίληψη

Από τον Μάρτιο του 2020, η ανθρωπότητα βιώνει μία πανδημία, που εκτός από πολυάριθμους ασθενείς και θύματα, έχει επιφέρει τεράστιες αλλαγές στην καθημερινότητα και την κοινωνική και οικονομική πραγματικότητα. Η πανδημία οφείλεται στην νόσο COVID-19, που εμφανίστηκε τον Δεκέμβριο του 2019 στην Γουχάν της Κίνας, με κύρια συμπτώματα τον ξηρό βήχα, τον πυρετό και την δυσκολία στην αναπνοή και προκαλείται από τον ιό που προκαλεί το σοβαρό οξύ αναπνευστικό σύνδρομο τύπου 2 γνωστό με το διεθνές όνομα SARS-CoV-2. Ο SARS-CoV-2 είναι ένας RNA ιός που ανήκει στην οικογένεια των κορονοϊών, οι οποίοι έχουν και στο παρελθόν προκαλέσει ανησυχία για τη δημόσια υγεία (SARS-CoV, MERS-CoV). Εισέρχεται στα κύτταρα του ξενιστή μέσω πρόσδεσης στον υποδοχέα ACE2 (ένζυμο μετατροπής της αγγειοτενσίνης 2), ο οποίος εκφράζεται κυρίως σε κυψελιδικά επιθηλιακά κύτταρα τύπου ΙΙ, γι' αυτό και οι πνεύμονες είναι το κύριο όργανο που προσβάλλει ο ιός.

Για την αντιμετώπιση του ιού έχουν ήδη αναπτυχθεί εμβόλια, ενώ η ομοιότητα του γονιδιώματός του με αυτό άλλων κορονοϊών έχει διευκολύνει και την χρήση ήδη υπαρχόντων φαρμάκων. Οι κύριοι στόχοι για την ανάπτυξη αντίδικών σκευασμάτων είναι η παρεμπόδιση της εισόδου του ιού στον οργανισμό και κατόπιν της αναπαραγωγής και διάδοσής του σε υγιή κύτταρα. Αυτοί μπορούν να επιτευχθούν στοχεύοντας την καταστολή λειτουργικών και δομικών πρωτεϊνών του ιού. Η πρώτη κατηγορία των λειτουργικών πρωτεϊνών περιλαμβάνει την κύρια πρωτεάση (M^{pro}), την πρωτεάση παπαΐνης (PL^{pro}), την RNA εξαρτώμενη RNA πολυμεράση (RdRp) και την ελικάση (nsp13). Στις δομικές πρωτεΐνες περιλαμβάνονται οι γλυκοζυλιωμένες πρωτεΐνες που βρίσκονται στην επιφάνεια του ιού και έχουν σχήμα αιχμής (spike) (S proteins), οι πρωτεΐνες του περιβλήματος (envelope) (E proteins) και της μεμβράνης (M proteins) καθώς και οι νουκλεοκαψιδικές πρωτεΐνες (N proteins). Άλλοι πιθανοί στόχοι αντιικών φαρμάκων είναι οι παράγοντες μολυσματικότητας, πρωτεΐνες του ιού που καταστέλλουν την ανοσολογική απόκριση του οργανισμού. Τέλος μια άλλη στρατηγική είναι ο περιορισμός πρωτεϊνών του ξενιστή που διευκολύνουν την αναπαραγωγή του ιού, όπως ο υποδοχέας ACE2.

Λόγω του καθοριστικού ρόλου της κύριας πρωτεάσης (M^{pro}) στην αναπαραγωγή του ιού, η παρούσα μελέτη εστιάζει στην παρεμπόδιση του συγκεκριμένου ενζύμου. Η M^{pro} ανήκει στην κατηγορία των υδρολασών και καταλύει τον σχηματισμό μη δομικών πρωτεϊνών, συμπεριλαμβανομένης και της ίδιας, μέσω της διάσπασης πεπτιδικών δεσμών σε πάνω από 11 κέντρα διάσπασης στις δύο πολυπρωτεΐνες pp1a και pp1ab που προκύπτουν από την μετάφραση του ιικού RNA στα κύτταρα του ξενιστή. Η αλληλουχία αμινοξέων που αναγνωρίζει η πρωτεάση για να πραγματοποιήσει την υδρόλυση είναι (Leu-Gln)-(Ser/Ala/Gly), με τον πεπτιδικό δεσμό που διασπάται να βρίσκεται μετά την γλουταμίνη. Η δομή της είναι διαθέσιμη στην βάση δεδομένων Protein Data Bank (PDB), τόσο σε ελεύθερη μορφή, όσο και σε σύμπλεγμα με διάφορους παρεμποδιστές, με την πρώτη δομή που καταχωρήθηκε να είναι η 6LU7. Η πρωτεάση εμφανίζει ενεργότητα σε διμερή μορφή, αποτελούμενη από 2 αλυσίδες μήκους 306 αμινοξέων. Η καθεμία από τις αλυσίδες χωρίζεται σε τρεις τομείς, τον τομέα I (αμινοξέα 8–101), τομέα II (αμινοξέα 102–184) and τομέα III (αμινοξέα 201–303). Το ένζυμο δρα μέσω μια καταλυτικής δυάδας κυστεΐνης-ιστιδίνης, αποτελούμενης από τα αμινοξέα Cys 145 και His 41. Το ενεργό κέντρο αποτελείται από 4 υποκέντρα: τα S1 (περιλαμβάνει τις πλευρικές αλυσίδες των Phe 140, Asn 142, Ser 144, Cys 145, His 163, Glu 166, His 172, και τον σκελετό των Leu 141, Gly 143, His 164), S1' (περιλαμβάνει τις πλευρικές αλυσίδες των Thr 25, His 41, Val 42, Asn 119, Gly 143, Cys 145

και τον σκελετό της Thr 26) , S2 (περιλαμβάνει τις πλευρικές αλυσίδες των His 41, Met 49, Tyr 54, Asp 187 και τον σκελετό της Arg 188) και S4 (περιλαμβάνει τις πλευρικές αλυσίδες των Met 165, Leu 167, Pro 168, Ala 191, Gln 192 και τον σκελετό των Glu 166, Arg 188, Thr 190), εκ των οποίων τα S2 και S4 εμφανίζουν μεγαλύτερη πλαστικότητα, ενώ τα S1 και S1' είναι πολύ βασικά για την κατάλυση, δεδομένου ότι η καταλυτική δυάδα βρίσκεται ανάμεσά τους. Ο καταλυτικός μηχανισμός βασίζεται σε μία αντίδραση ηλεκτρονιόφιλης προσθήκης: αρχικά πραγματοποιείται μεταφορά πρωτονίου από την θειόλη της καταλυτικής κυστεΐνης στο ιμιδαζόλιο της ιστιδίνης και στη συνέχεια η κυστεΐνη αντιδρά με το καρβονύλιο του προς λύση δεσμού σχηματίζοντας ένα ενδιάμεσο θειοημιακετάλης, ενώ η ιστιδίνη προσεγγίζει το άζωτο του πεπτιδικού δεσμού, σχηματίζοντας ένα ενδιάμεσο ακυλοενζύμου. Σημαντικό ρόλο στην καταλυτική αντίδραση έχει και ένα μόριο νερού, που συμβάλει στην σταθεροποίηση του καρβονυλίου της γλουταμίνης, που βρίσκεται δίπλα στον υδρολυόμενο δεσμό, και την διάσπαση του ακυλοενζύμου.

Η ενζυμική παρεμπόδιση (ή αναστολή), δηλαδή η μερική ή ολική αδρανοποίηση ενός ενζύμου, μπορεί να γίνει αντιστρεπτά (μη επηρεάζοντας τη δομή της πρωτεΐνης) ή μη αντιστρεπτά (καταστρέφοντάς την λειτουργικά). Η αντιστρεπτή παρεμπόδιση μπορεί επιπλέον να χωριστεί στις κατηγορίες της συναγωνιστικής, μη-συναγωνιστικής, ασυναγώνιστης και μεικτής αναστολής. Οι δύο τελευταίες κατηγορίες περιλαμβάνουν και την αλλοστερική αναστολή, όπου ο αναστολέας συνδέεται σε άλλη περιοχή του ενζύμου από το ενεργό κέντρο. Η πλειοψηφία των ήδη γνωστών αναστολέων της M^{pro} δρουν αντιστρεπτά και ανταγωνιστικά, ενώ υπάρχουν και κάποιοι που αντιδρούν ομοιοπολικά και προκαλούν αναντίστρεπτη αναστολή. Υπάρχει μία άμεση σχέση μεταξύ της δομής ενός αναστολέα και του πόσο αποτελεσματικός είναι. Για παράδειγμα, κάποια ηλεκτρονιόφιλη ομάδα, όπως μια αλδεϋδομάδα, μπορεί να αποτελεσματική στην θέση P1, δηλαδή στην πλευρική ομάδα του αμινοξέος που βρίσκεται δίπλα στον προς διάσπαση δεσμό προς την κατεύθυνση του N-άκρου της πεπτιδικής αλυσίδας. Λόγω του ότι το υποκέντρο S2 προσελκύει υδρόφοβα μόρια, κάποιο μόριο που περιέχει κάποια τέτοια ομάδα, όπως για παράδειγμα μια αλειφατική αλυσίδα, μπορεί να ευνοήσει την πρόσδεσή του στο ενεργό κέντρο της πρωτεάσης.

Η μελέτη διαφόρων αναστολέων και του πώς προσδένονται στο ενεργό κέντρο έχει αναδείξει κάποια κύρια αμινοξέα τα οποία συμμετέχουν συχνά στο σχηματισμό δεσμών υδρογόνου. Ειδικότερα, δύο τέτοιες ομάδες αποτελούν τα Arg 188 και Thr 190 και Glu 166 και Gln 189, στο υποκέντρο S4. Αντίστοιχα, στο υποκέντρο S2, τα γειτονικά αμινοξέα His 41, Tyr 54 και Asp 187 σχηματίζουν άλλη μια ομάδα, ενώ στα υποκέντρο S1 και S1' τα αμινοξέα πιο επιδεκτικά στον σχηματισμό δεσμών υδρογόνου είναι τα Leu 141, Gly 143, Ser 144 και His 163 και Thr 24 και Thr 45, αντίστοιχα.

Η έρευνα έχει ήδη αναδείξει μόρια που δρουν αποτελεσματικά ως αναστολείς της M^{pro} , για τα οποία υπάρχει και διαθέσιμη κρυσταλλική δομή σε σύμπλεγμα με την πρωτεάση. Οι αναστολείς αυτοί δρουν είτε σχηματίζοντας ομοιοπολικό δεσμό με την πρωτεάση, είτε απλά συνδεόμενοι στο ενεργό της κέντρο και σταθεροποιούμενοι με δεσμούς υδρογόνου και υδρόφοβες αλληλεπιδράσεις. Ο πιο μελετημένος αναστολέας είναι ο N3, ο οποίος μιμείται τη δομή ενός πεπτιδίου και συνεπώς είναι στενά συγγενικός στο πραγματικό υπόστρωμα του ενζύμου. Σχηματίζει ομοιοπολικό δεσμό με την πρωτεάση και η δράση του είναι αναντίστρεπτη. Άλλοι αναστολείς που προσδένονται ομοιοπολικά στην πρωτεάση είναι: η φαρμακευτική ουσία ebselen, τα αμίδια 11a, 11b, 13b και 5h, η μποσεπρεβίρη, τα αντιικά μόρια GC376, GC373, ναρλαπρεβίρη, MI-23, calpeptin, το αντινεοπλασματικό φάρμακο καρμοφούρη, το φυτοχημικό μυρικετίνη και η ουσία MG-132. Λιγότεροι σε αριθμό, αλλά όχι

λιγότερο σημαντικοί, είναι και οι μη ομοιοπολικοί αναστολείς της M^{pro}. Αυτοί περιλαμβάνουν τα μόρια 2-(3-(3-χλωρο-5-προποξυφαινυλ)-2-όξο-2H-[1,3'-διπυριδιν]-5-υλ-βενζονιτρίλιο (compound 5) και 2-(3-(3-χλωρο-5-(κυκλοπροπυλμέθοξυ)φαινυλ)-2-όξο-2H-[1,3'-διπυριδιν]-5-υλ-βενζονιτρίλιο (compound 26), Z1220452176 (compound x0104), Z18197050 (x0161), Z369936976 (x0397), MUT056399, ML 188, Mcule-5948770040 και X77.

Η βιβλιογραφία αναφέρει επίσης και την ύπαρξη δύο αλλοστερικών περιοχών όπου έχει εντοπιστεί ότι συνδέονται αναστολείς της πρωτεάσης. Το ένα αποτελεί μία υδρόφοβη κοιλότητα που απαρτίζεται από τα αμινοξέα Ile 213, Leu 253, Gln 256, Val 297 and Cys 300, ενώ το δεύτερο εντοπίζεται μεταξύ των τομέων I, II και III και η πρόσδεση αναστολέων σε αυτό επιφέρει στην αποσταθεροποίηση του διμερούς ενζύμου και του υποκέντρου S1. Αποτελεσματικοί αναστολείς αναφέρονται το μόριο pelitinib για το πρώτο αλλοστερικό κέντρο και το μόριο AT7519 για το δεύτερο.

Εκτός από την έρευνα για την ανάπτυξη φαρμάκων και συνθετικών ουσιών που μπορούν να χρησιμοποιηθούν ενάντια στον SARS-CoV-2, μία πολλά υποσχόμενη εναλλακτική αποτελούν φυσικές ουσίες που απαντώνται σε φυτά. Για παράδειγμα η μυρικετίνη είναι ένα φλαβονοειδές που έχει πειραματικά αποδειχθεί ότι αναστέλλει τη δράση της πρωτεάσης. Γενικότερα η κατηγορία των φλαβονοειδών περιλαμβάνει πολυάριθμες ουσίες που δείχνουν να έχουν αντιική δράση, όπως η ρουτίνη, η καιμπφερόλη και η κουερσετίνη. Εκτός από τα φλαβονοειδή, τα φαινολικά οξέα είναι ακόμα μία κατηγορία για μέλη της οποίας μελέτες που έχουν γίνει αναφέρουν ενθαρρυντικά αποτελέσματα.

Η παρούσα έρευνα επικεντρώνεται σε φυσικές ουσίες που έχουν εντοπιστεί στο φυτό *Salicornia* (αρμυρήθρα). Το *Salicornia* είναι ένα αλόφυτο που συναντάται σε αλατούχα εδάφη, όπως παραθαλάσσιες περιοχές, σε όλον τον κόσμο εκτός από την Ανταρκτική, την Αυστραλία και την Νότια Αμερική και καταναλώνεται ως τρόφιμο, ενώ χρησιμοποιείται και στην παραδοσιακή ιατρική ενάντια στην υπέρταση, την παχυσαρκία και τον καρκίνο. Το φυτό έχει σημαντικά διατροφικά και ιατρικά οφέλη που μεταξύ άλλων περιλαμβάνουν μείωση του επιπέδου των λιπιδίων στο αίμα, ενίσχυση του ανοσοποιητικού συστήματος αντιοξειδωτική, αντιφλεγμονώδη και αντιική δράση. Ειδικότερα αντιική δράση εντοπίζεται ενάντια στον ιό της γρίπης, του έρπητα, τους αδενοϊούς ADV-3, ADV-8 και ADV-11 και τον αναπνευστικό συγκυτιακό ιό (RSV). Αυτές οι δράσεις αποδίδονται στην ύπαρξη φανολικών οξέων και φλαβονοειδών. Στο εκχύλισμα του φυτού, το οποίο μπορεί να παραχθεί και από τα μη άμεσα αξιοποιήσιμα μέρη του φυτού που συνήθως απορρίπτονται, έχει εντοπιστεί ένα εύρος πιθανά βιοενεργών συστατικών που ανήκουν στις κατηγορίες των υδροξυκινναμικών, υδροξυβενζοϊκών και καφεοϋλκινικών οξέων, φλαβονοειδών και φλαβανονών, στερολών, χρωμονών, λιγνανών και σαπονινών.

Στόχος της παρούσας διπλωματικής εργασίας είναι η αξιολόγηση των συσταστικών του εκχυλίσματος του φυτού *Salicornia* που ανήκουν στις προαναφερθείσες κατηγορίες, ως προς την παρεμποδιστική τους δράση ενάντια στην κύρια πρωτεάση του SARS-CoV-2. Αρχικά, κάνοντας χρήση των εργαλείων της βιοπληροφορικής και ειδικότερα της προσομοίωσης μοριακής πρόσδεσης (molecular docking) αξιολογήθηκαν τα βιοενεργά συστατικά που έχουν αναφερθεί ότι υπάρχουν σε εκχυλίσματα του φυτού *Salicornia* με βάση τη βιβλιογραφία. Στη συνέχεια η παρεμποδιστική δράση των πιο αντιπροσωπευτικών και φαινομενικά δραστικών ουσιών, όπως και η συνολική δράση ενός εκχυλίσματος από το φυτό, αξιολογήθηκαν πειραματικά με μια δοκιμή παρεμπόδισης του ενζύμου *in vitro*.

Αρχικά, για να γίνει δυνατή η αξιολόγηση των αποτελεσμάτων της προσομοίωσης για τα συστατικά του εκχυλίσματος, χρειάστηκε η συλλογή δεδομένων για τους ήδη υπάρχοντες αναστολείς τις πρωτεάσης, όπως και η εκτίμηση της ενέργειας αλληλεπίδρασης του αναστολέα με την πρωτεΐνη και των αλληλεπιδράσεων μεταξύ τους, ώστε να είναι δυνατή η σύγκριση με τα αποτελέσματα της παρούσας προσομοίωσης. Τα δεδομένα αυτά συλλέχθηκαν από την βιβλιογραφία και από τις κρυσταλλογραφημένες δομές του συμπλόκου αναστολέα-ενζύμου που είναι διαθέσιμες στην PDB, ενώ για την οπτικοποίησή τους χρησιμοποιήθηκε το πρόγραμμα YASARA structure, έκδοση 20.12.24.

Η προσομοίωση μοριακής πρόσδεσης υπολογίζει την χαμηλότερη δυνατή ενέργεια (binding energy), και την γεωμετρία που την επιφέρει, για το σύμπλεγμα ενός μακρομορίου (receptor), όπως είναι η πρωτεάση, και ενός μικρότερου μορίου (ligand), όπως στην προκειμένη περίπτωση ο αναστολέας. Για να πραγματοποιηθεί η προσομοίωση χρειάστηκε να γίνει κατάλληλη προετοιμασία των δύο μορίων (καθαρισμός της δομής τους, προσθήκη όλων των ατόμων υδρογόνου, ελαχιστοποίηση της ενέργειάς τους) καθώς και ο ορισμός ενός κατάλληλου κελιού προσομοίωσης, το οποίο επιλέχθηκε να εκτείνεται γύρω από την καταλυτική δυάδα His 41-Cys 145 ως κύβος με ακμή 2-3 Å μεγαλύτερη από το μήκος του ligand. Η προσομοίωση εκτελέστηκε αρχικά για τις ήδη κρυσταλλογραφημένες δομές της πρωτεάσης με τους αντίστοιχους αναστολείς, ώστε να υπάρχει μία αναφορά για τις ενέργειες του συστήματος πρωτεάσης-αναστολέα και να μπορούν να συγκριθούν με τυχόν αντίστοιχες καταγεγραμμένες τιμές στη βιβλιογραφία, καθώς και για να διερευνηθεί η αξιοπιστία του προγράμματος και να οριστεί η μέθοδος που θα ακολουθηθεί στις υπόλοιπες ουσίες. Η προσομοίωση για τα συστατικά του *Salicornia* έγινε με τον ίδιο τρόπο, χρησιμοποιώντας ως receptor την δομή της πρωτεάσης 6LU7. Για την πραγματοποίηση της προσομοίωσης μοριακής πρόσδεσης χρησιμοποιήθηκε το ενσωματωμένο στο YASARA λογισμικό Autodock Vina, το οποίο παράγει 25 δυνατές διαμορφώσεις του συμπλέγματος receptor-ligand, οι οποίες ομαδοποιούνται με βάση την ομοιότητά τους. Από κάθε προκύπτουσα ομάδα επιλέγεται η διαμόρφωση με την χαμηλότερη ενέργεια. Οι προτεινόμενες γεωμετρίες που επιλέγονται από κάθε ομάδα ονομάζονται cluster. Ο αριθμός των cluster που προκύπτουν διαφέρει σε κάθε προσομοίωση, ανάλογα με το πόσες ευνοϊκές διαμορφώσεις προκύπτουν και πόσο αυτές διαφέρουν μεταξύ τους. Εκτός από την ενέργεια πρόσδεσης που δίνεται για κάθε cluster, δίνεται και η τιμή της σταθεράς διάστασης και τα αμινοξέα που αλληλεπιδρούν με τον ligand (contacting residues), ενώ το YASARA επιτρέπει συγκεκριμένα τον υπολογισμό των δεσμών υδρογόνου και των υδροφοβικών και π-π αλληλεπιδράσεων. Για την αξιολόγηση των αποτελεσμάτων της προσομοίωσης των γνωστών αναστολέων, χρησιμοποιήθηκαν ως μέτρο σύγκρισης αρχικά η ρίζα μέσης τετραγωνικής απόκλισης (RMSD) της προκύπτουσας δομής από την κρυσταλλική, για την επιλογή του cluster, και στη συνέχεια οι αλληλεπιδράσεις που προκύπτουν από τη βιβλιογραφία και την κρυσταλλική δομή, για την αξιολόγησης της αξιοπιστίας της μεθόδου. Αναφορικά με τα άγνωστων ιδιοτήτων φυτοχημικά του *Salicornia*, για την επιλογή του cluster βασικό κριτήριο ήταν η χαμηλότερη ενέργεια πρόσδεσης, σε συνδυασμό με την σύγκριση με τη βιβλιογραφία όπου αυτή ήταν δυνατή. Παράγοντας επιλογής ήταν επίσης και τυχόν κοινές γεωμετρίες που εμφανίστηκαν σε ουσίες παρόμοιας δομής. Λόγω αδυναμίας σε πολλές περιπτώσεις να εξαχθεί κάποιο συμπέρασμα για κάποια προτιμώμενη διαμόρφωση, συχνά παραπάνω από ένα cluster λήφθηκαν υπ' όψιν.

Για την πειραματική διερεύνηση της παρεμποδιστικής δράσης των ουσιών χρησιμοποιήθηκε αναλυτικό κιτ για την συγκεκριμένη πρωτεάση (3CL Protease MBP-tagged Assay Kit από την εταιρεία BPS Bioscience (San Diego, CA, USA), το οποίο περιλάμβανε ως μέτρο θετικού

ελέγχου τον αναστολέα GC376. Μετρήθηκε η ενεργότητα του ενζύμου παρουσία του γνωστού αναστολέα και επιλεκτικά κάποιων αντιπροσωπευτικών φυτοχημικών ουσιών που οδήγησαν σε ενθαρρυντικά αποτελέσματα κατά την προσομοίωση, και σε εκχύλισμα του *Salicornia*. Συγκεκριμένα οι ουσίες που μελετήθηκαν είναι: φερουλικό οξύ, ροσμαρινικό οξύ, κινικό οξύ, γαλλικό οξύ, χλωρογενικό οξύ, 3,4-δικαφεοϋλκινικό οξύ, 3,5-δικαφεοϋλκινικό οξύ, κουερσετίνη, ισοκουερσετίνη, ρουτίνη, ισοραμνετίνη, 3-γλυκοζίτης της ισοραμνετίνης, 3-ρουτινοζίτης της ισοραμνετίνης, εσπερετίνη, εσπεριδίνη, ακακετίνη, γκαλανζίνη, μυρικετίνη, απιγενίνη, 3-ρουτινοζίτης της πελαργονιδίνης, καιμπφερόλη, χρυσίνη και κατεχίνη. Τα αρχικά διαλύματα των ουσιών παρασκευάστηκαν σε DMSO, και στη συνέχεια αραιώθηκαν με το ρυθμιστικό διάλυμα ανάλυσης σε πολλαπλές συγκεντρώσεις για την κάθε ουσία, μεταξύ 0.05 και 5000 μM. Αφού η κάθε ουσία επωάθηκε με το διάλυμα του ενζύμου, προστέθηκε το υπόστρωμα και το σύστημα τοποθετήθηκε προς επώαση σε θερμοκρασία περιβάλλοντος για 4-24 h. Ο υπολογισμός της ενεργότητας έγινε μετρώντας την ένταση φθορισμού (διέγερση: 360 nm, εκπομπή: 460 nm). Από την καμπύλη της ενεργότητας σε συνάρτηση με τις διάφορες συγκεντρώσεις για την κάθε ουσία υπολογίστηκε η συγκέντρωση που επιφέρει 50% μείωση της ενεργότητας, IC_{50}.

Από την ανάλυση των δεσμών υδρογόνου που πραγματοποιούν οι αναστολείς με τα αμινοξέα του ενεργού κέντρου της πρωτεάσης, όπως αυτοί φαίνονται από την κρυσταλλική δομή, προκύπτει ότι το Glu 166 σχηματίζει δεσμούς υδρογόνου με την πλειοψηφία των αναστολέων. Άλλα αμινοξέα με συχνές αλληλεπιδράσεις είναι τα Gly 143, Cys 145 και His 163. Κατά πλειοψηφία, τα αμινοξέα με τα οποία οι αναστολείς φαίνονται να αλληλεπιδρούν πιο συχνά είναι αμινοξέα που βρίσκονται στα υποκέντρα S1 και S1', κάτι που θα ήταν αναμενόμενο λόγω του βασικού τους ρόλου κατά την κατάλυση, αφού ο προς λύση δεσμός του πραγματικού πεπτιδικού υποστρώματος τοποθετείται ανάμεσα σε αυτά τα δύο υποκέντρα.

Η προσομοίωση για τους γνωστούς αναστολείς της M^{pro} έδωσε ως αποτέλεσμα προσανατολισμούς των αναστολέων στο ενεργό κέντρο πολύ παρόμοιους με αυτούς που φαίνονται στις αντίστοιχες κρυσταλλικές δομές. Για τον πιο μελετημένο αναστολέα του ενζύμου, N3, η ενέργεια που υπολογίστηκε από το Vina είναι -8.26 kcal/mol, ενώ το RMSD του μορίου από την κρυσταλλική δομή ήταν 3.16 Å. Ένας άλλος αναστολέας ευρείας δράσης, ο οποίος χρησιμοποιείται στη συνέχεια και στην εργαστηριακή ανάλυση, είναι ο GC376, για τον οποίον η ενέργεια πρόσδεσης υπολογίστηκε -7.798 kcal/mol και η γεωμετρία του μορίου ήταν πολύ παρόμοια με αυτήν της κρυσταλλικής δομής, με απόκλιση μόλις 1 Å. Για τα υπόλοιπα μόρια, εκτός από την καρμοφούρη, για την οποία υπολογίστηκε πολύ χαμηλή κατ' απόλυτη τιμή ενέργεια πρόσδεσης (-2.93 kcal/mol), οι τιμές για την ενέργεια πρόσδεσης κυμάνθηκαν από -5.399 kcal/mol έως -9.464 kcal/mol, ενώ το RMSD πήρε τιμές μεταξύ 0.318 και 6.65 Å. Στις περισσότερες περιπτώσεις, η μικρή απόκλιση μεταξύ του συμπλόκου αναστολέα-ενζύμου που απεικονίζεται στην κρυσταλλική δομή και αυτού που προέκυψε από την προσομοίωση, ενισχύει την αξιοπιστία της μοριακής πρόσδεσης ως ένα μέσο πρόβλεψης της συγγένειας ενός μορίου με το ενεργό κέντρο. Παρατηρήθηκε ωστόσο ότι το πρόγραμμα οδήγησε στον υπολογισμό σημαντικά λιγότερων δεσμών υδρογόνου από αυτούς που εμφανίζονταν στις κρυσταλλικές δομές ή αναφέρονταν στην βιβλιογραφία, κάτι που μπορεί να οφείλεται στο γεγονός ότι, αφού το πρόγραμμα υπολογίζει τους δεσμούς με βάση την απόσταση, κάποια λίγο διαφορετική διαμόρφωση στον αναστολέα ή την πρωτεΐνη μπορεί να οδηγήσει σε αύξηση της μεταξύ τους απόστασης και έτσι να μην υπολογιστεί από το πρόγραμμα κάποιος δεσμός που σχηματίζεται στην πραγματικότητα. Μία επιπλέον γενική παρατήρηση είναι ότι τα αποτελέσματα της προσομοίωσης για τους μη ομοιοπολικούς

αναστολείς είναι πιο πιστά στα πραγματικά δεδομένα από τα αντίστοιχα για τους ομοιοπολικούς, κάτι που μπορεί να αποδοθεί στο ότι το πρόγραμμα δεν μπορεί να εκτελέσει προσομοίωση για σχηματισμό ομοιοπολικού δεσμού, επομένως αντιμετωπίζει την κάθε περίπτωση σαν μη ομοιοπολική πρόσδεση.

Όπως είναι αναμενόμενο, οι μεγαλύτεροι σε μέγεθος αναστολείς παρεμποδίζουν περισσότερες από τις υποπεριοχές του ενεργού κέντρου, με τους N3, 5h, Mg-132, 13b, x2705, compound 5, compound 26 και X77 να τις καλύπτουν όλες. Παρατηρείται επίσης ότι οι αναστολείς με συγγενικές δομές εμφανίζουν και πολύ παρόμοιες αλληλεπιδράσεις με το ενεργό κέντρο. Επιπλέον συχνά εμφανιζόμενα μοτίβα είναι η συμμετοχή καρβονυλίων γειτονικών σε αμινομάδες, είτε σε αλειφατικές είτε σε κυκλικές ανθρακικές αλυσίδες, σε δεσμούς υδρογόνου, καθώς και ο σχηματισμός δεσμού μεταξύ των προαναφερθέντων αμινομάδων και καρβονυλίων των αμινοξέων του ενεργού κέντρου. Οι κυκλοπεντανικοί και κυκλοεξανικοί δακτύλιοι των εξεταζόμενων μορίων εμφανίζουν επίσης π-π αλληλεπιδράσεις με τις ιμιδαζόλες των αμινοξέων His 41 και His 163. Επιπλέον, επιβεβαιώνεται από την προσομοίωση το αμινοξύ Glu 166 ως η πιο συχνή επαφή των συνδεόμενων μορίων, ακολουθούμενη από το Gly 143.

Στη συνέχεια, η προσομοίωση μοριακής πρόσδεσης πραγματοποιήθηκε για τα βιοδραστικά συστατικά του εκχυλίσματος των φυτών *Salicornia.* Για τα υδροξυκινναμικά οξέα, υπολογίστηκαν χαμηλές κατ' απόλυτη τιμή ενέργειες πρόσδεσης, μεταξύ -4.98 και -7.41 kcal/mol, χαμηλότερες τόσο από τον αναστολέα N3 όσο και από τον GC376. Ενώ η ενέργεια των απλών οξέων κυμάνθηκε ως επί το πλείστον μεταξύ -5 και -6 kcal/mol, παρατηρήθηκε μία αύξηση στην τιμή της ενέργειας για το ροσμαρινικό οξύ, το οποίο είναι εστέρας του καφεϊκού οξέος και περιέχει δύο φαινολικούς δακτυλίους (-7.409 kcal/mol) . Η επικρατέστερη γεωμετρία των μορίων αυτής της κατηγορίας στο ενεργό κέντρο είναι αυτή κατά την οποία ο φαινολικός δακτύλιος προηγείται της γραμμικής ανθρακικής αλυσίδας, και σταθεροποιείται είτε στην περιοχή S2 είτε μεταξύ των S1 και S1', ενώ το υπόλοιπο μόριο εκτείνεται προς τις περιοχές S2 και S4. Επιπλέον, τα μόρια αυτής της κατηγορίας φαίνονται να αλληλεπιδρούν υδρόφοβα με το αμινοξύ Met 165, ενώ ο φαινολικός τους δακτύλιος εμφανίζει π-π αλληλεπιδράσεις με τις ιστιδίνες His 41 και His 163. Αντίστοιχες μελέτες έχουν γίνει *in silico* για το φερουλικό, το καφεϊκό, το σιναπικό, το κινναμωμικό και το κουμαρικό οξύ και αναφέρουν πολύ παρόμοια αποτελέσματα με τα παραπάνω, με εξαίρεση κάποιες διαφορετικές τιμές για την ενέργεια πρόσδεσης του φερουλικού οξέος, που μπορούν να αποδοθούν στο γεγονός ότι χρησιμοποιήθηκε διαφορετική κρυσταλλική δομή της πρωτεάσης και άλλο λογισμικό για την προσομοίωση.

Τα υδροξυβενζοϊκά οξέα, όντας πολύ παρόμοια σε δομή και μέγεθος, οδήγησαν και σε πολύ παρόμοια μεταξύ τους αποτελέσματα για την ενέργεια πρόσδεσης (-4.826 έως -5.601 kcal/mol). Σχετική έρευνα στην βιβλιογραφία παρέχει δεδομένα για το βανιλλικό, το γαλλικό και το 4-υδροξυβενζοϊκό οξύ, τα οποία είναι πολύ κοντινά με αυτά της παρούσας μελέτης, ενισχύοντας έτσι την εγκυρότητά τους. Βασικά αμινοξέα που συμμετέχουν σε δεσμούς υδρογόνου με τις ουσίες αυτής της κατηγορίας προκύπτουν να είναι τα Glu 166, Gly 143 και Leu 141, ενώ ειδικότερα η Glu 166 αλληλεπιδρά και υδρόφοβα με τα μόρια, και η His 163 προσφέρεται για π-π αλληλεπιδράσεις και σε αυτήν την κατηγορία. Η χαμηλότερη ενέργεια προκύπτει για το 4-υδροξυβενζοϊκό οξύ, το οποίο είναι υποκατεστημένο μόνο στην θέση 4 με ένα υδροξύλιο, ενώ το μόριο που προκύπτει ότι έχει την μεγαλύτερη συγγένεια στο ενεργό κέντρο είναι το γαλλικό οξύ, που είναι υποκατεστημένο με τρία υδροξύλια. Μεταξύ των ουσιών που είναι υποκατεστημένες μόνο με υδροξύλια, παρατηρείται ότι αύξηση του

αριθμού των υδροξυλίων οδηγεί σε αύξηση της συγγένειάς τους με το ενεργό κέντρο της πρωτεάσης.

Τα καφεοϋλκινικά οξέα έδωσαν πολύ ενθαρρυντικά αποτελέσματα για την αναστολή της M^{pro}, με κατά πλειοψηφία καλύτερες ενέργειες πρόσδεσης στο ενεργό κέντρο από τους αναστολείς N3 και GC376. Εκτός από το κινικό οξύ, που είναι ένα μικρό μόριο το οποίο αποτελεί δομική μονάδα των υπολοίπων και η ενέργειά πρόσδεσής του υπολογίστηκε ίση με -5.887 kcal/mol, για τα υπόλοιπα παράγωγα αυτής της κατηγορίας υπολογίστηκαν ενέργειες πρόσδεσης μεταξύ -7.9 και -8.935 kcal/mol, με τις πιο υποσχόμενες ουσίες να είναι το 3,5-δικαφεοϋλκινικό οξύ και ο μεθυλεστέρας του, ο μεθυλεστέρας του 4-καφεοϋλ-3-διυδροκαφεοϋλκινικού οξέος, το 3,4-δικαφεοϋλκινικό οξύ και το 3-καφεοϋλ-5-διυδροκαφεοϋλκινικό οξύ. Γενικότερα παρατηρήθηκε ότι οι υποκαταστάσεις στις θέσεις 3,5 του κινικού οξέος επιφέρουν καλύτερα αποτελέσματα από τις υποκαταστάσεις στις θέσεις 3,4, καθώς επίσης και ότι σε αντίστοιχη θέση, η υποκατάσταση με καφεϊκό οξύ προτιμάται συγκριτικά με το υδροκαφεϊκό. Τα καφεοϋλκινικά οξέα που εξετάστηκαν αλληλεπιδρούν συχνά με την Gly 143 και με αμινοξέα της περιοχής S4 (Arg 188, Gln 189). Επίσης εντοπίζονται δύο κυρίαρχες διαμορφώσεις για τα παράγωγα υποκατεστημένα στις 3,4 και 3,5 θέσεις. Στην πρώτη, η μονάδα του κινικού οξέος βρίσκεται μεταξύ των περιοχών S1 και S1', ενώ οι υποκαταστάσεις καφεϊκού οξέος εκτείνονται κάθετα μεταξύ τους, προς τα πάνω και αριστερά του ενεργού κέντρου, και στην δεύτερη το κινικό οξύ βρίσκεται στην ίδια θέση αλλά οι υποκαταστάσεις εκτείνονται προς την ίδια κατεύθυνση, προς τα αριστερά, και σχεδόν παράλληλα μεταξύ τους. Η βιβλιογραφία παρέχει δεδομένα προς σύγκριση για το κινικό και το χλωρογενικό οξύ. Για το πρώτο η τιμή ενέργειας πρόσδεσης που δίνεται είναι πολύ κοντά στα αποτελέσματα αυτής της έρευνας, ενώ για το δεύτερο δίνονται τιμές που κυμαίνονται σε μεγάλο εύρος μη επιτρέποντας την εξαγωγή συμπερασμάτων.

Μία άλλη πολυπληθής ομάδα, που έδωσε και τα πιο ενθαρρυντικά αποτελέσματα μεταξύ των ενώσεων που εξετάστηκαν και επιπλέον περιλαμβάνει τις ενώσεις που έχουν μελετηθεί περισσότερο στη βιβλιογραφία για την αντιική τους δράση, είναι τα φλαβονοειδή και οι φλαβανόνες. Οι ενέργειες πρόσδεσης των ουσιών υπολογίστηκαν μεταξύ -6.88 και -9.384 kcal/mol με την χαμηλότερη ενέργεια να ανήκει στη ρουτινόζη της ισοραμνετίνης. Η μυρικετίνη, που είναι αναγνωρισμένος αναστολέας της πρωτεάσης, ανήκει σε αυτήν την ομάδα και επομένως η κρυσταλλική της δομή παρουσιάζει ένα σημαντικό, πραγματικό μέτρο σύγκρισης για την αξιολόγηση των αποτελεσμάτων της προσομοίωσης. Η πλειοψηφία των πρώτων cluster για όλα τα φλαβονοειδή είχε προσανατολισμό στο ενεργό κέντρο παρόμοιο με αυτόν της μυρικετίνης, ενώ σε κάποιες περιπτώσεις ο προσανατολισμός αυτός εντοπιζόταν σε κάποιο επόμενο cluster. Εξαίρεση αποτελούν κάποιες ενώσεις όπως η ραμνετίνη, όπου κανένα από τα cluster δεν απεικόνισε κάποια γεωμετρία που να μοιάζει σε αυτήν της μυρικετίνης. Ένα χαρακτηριστικό μοτίβο της σχέσης δομής-συγγένειας με την πρωτεάση είναι η αύξηση αυτής όταν ένα φλαβονοειδές ή φλαβανόνη υποκαθίσταται με κάποιον γλυκοζίτη. Ειδικότερα, για την κουερσετίνη υπολογίστηκε ενέργεια πρόσδεσης -7.396 kcal/mol ενώ για τα παράγωγά της, υποκατεστημένα με γλυκόζη και ρουτινόζη, οι αντίστοιχες ενέργειες ανήλθαν στα -9.114 και -9.166 kcal/mol. Η βιβλιογραφία επιβεβαιώνει την παραπάνω τάση, με τις τιμές που δίνονται για την κουερσετίνη να είναι πολύ κοντινές στην προαναφερθείσα, ενώ για την ρουτίνη η πλειοψηφία των πηγών αναφέρει υψηλότερη ενέργεια πρόσδεσης, της τάξης των -11 kcal/mol, που προκύπτουν όμως από προσομοιώσεις με διαφορετικές παραμέτρους. Η ίδια αυξητική τάση στην ενέργεια πρόσδεσης παρατηρείται και στην ισοραμνετίνη (-7.233 kcal/mol) και τα παράγωγά της, όπου τα αποτελέσματα που προέκυψαν είναι προοδευτικά καλύτερα όταν αυτή υποκατασταθεί στην θέση 3 με

νεοεσπεριδόζη (-7.838 kcal/mol), ραμνοζυλ-αραβινόζη (-8.167 kcal/mol), ραμνόζη (-8.335 kcal/mol), γαλακτόζη (-8.383 kcal/mol), γλυκόζη (-8.613 kcal/mol) και ρουτινόζη (-9.384 kcal/mol). Η ίδια τάση εντοπίζεται σε όλα τα αντίστοιχα παράγωγα των φλαβονοειδών, όπως η εσπερετίνη (-7.09 kcal/mol) με την εσπεριδίνη (-8.636 kcal/mol) και η καιμπφερόλη (-7.752 kcal/mol) με την αστραγαλίνη (-8.989 kcal/mol). Η απιγενίνη (-7.796 kcal/mol), ο 7-γαλακτοζίτης της (-7.834 kcal/mol) και ο 7-γλυκοζίτης της (-8.391 kcal/mol) δείχνουν την ίδια αυξητική τάση, σε πολύ μικρότερο βαθμό όμως, ειδικότερα αφού ο γαλακτοζίτης έχει σχεδόν την ίδια ενέργεια με την μητρική ουσία. Αυτό μπορεί να οφείλεται στο ότι σε αυτήν την περίπτωση η υποκατάσταση βρίσκεται στην θέση 7, σε αντίθεση με την θέση 3 όπου είναι στις υπόλοιπες περιπτώσεις. Από τα παραπάνω, η γαλακτόζη, η γλυκόζη αλλά ιδιαίτερα η ρουτινόζη, είναι σάκχαρα που μπορούν να αυξήσουν την αποτελεσματικότητα των φλαβονοειδών και φλαβανονών ως παρεμποδιστές της πρωτεάσης του SARS-CoV-2 αλλά και την υδροφιλικότητα των μορίων, επιτρέποντας καλύτερη διαλυτοποίηση και διείσδυση σε υδατικά περιβάλλοντα, όπως είναι ο ανθρώπινος οργανισμός. Αναφορικά με τις μη υποκατεστημένες ουσίες, καλύτερα αποτελέσματα παράχθηκαν για την απιγενίνη και την καιμπφερόλη, ακολουθούμενες από τον επιβεβαιωμένο αναστολέα μυρικετίνη, για τον οποίο υπολογίστηκε ενέργεια -7.529 kcal/mol. Από τα παραπάνω προκύπτει ότι θα ήταν ενδιαφέρον να διερευνηθούν περαιτέρω τα παράγωγα των παραπάνω ενώσεων με διάφορους μονοσακχαρίτες ή δισακχαρίτες.

Οι στερόλες που ανιχνεύθηκαν στο *Salicornia* δεν έδωσαν ιδιαίτερα ενθαρρυντικές ενδείξεις προς αντιική δράση αναφορικά με την ενέργεια πρόσδεσής τους, αξίζει ωστόσο να διερευνηθούν περαιτέρω λόγω του ότι αποτελούν ογκώδη μόρια με την ικανότητα να καταλαμβάνουν μεγάλο μέρος του ενεργού κέντρου, συμπεριλαμβανομένης και της περιοχής όπου βρίσκεται η καταλυτική δυάδα. Οι τιμές για την ενέργεια πρόσδεσης είχαν στενό εύρος, μεταξύ -6.56 και -7.45 kcal/mol, με την βέλτιστη τιμή να αφορά το μικρότερο σε μέγεθος από τα μόρια αυτής της ομάδας, την εργοστερόλη. Κατά την αξιολόγηση των cluster λήφθηκε υπ' όψιν ότι εμφανίστηκε μια κοινή διαμόρφωση μεταξύ της πλειοψηφίας των μορίων, με τον σκελετό της γονάνης να βρίσκεται στην περιοχή S1' και την υποκατάσταση στο κυκλοπεντάνιο της γονάνης να εκτείνεται προς την περιοχή S4.

Τα αποτελέσματα για τις χρωμόνες αντίστοιχα δεν είναι ιδιαίτερα ενθαρρυντικά, χωρίς αυτό να αποκλείει ωστόσο τη πιθανή αντιική τους δράση. Πέρα από την 7-γλυκοπυρανοζυλ-6-μεθοξυχρωμόνη, που εμφανίζει την καλύτερη ενέργεια πρόσδεσης (-7.229 kcal/mol), οι υπόλοιπες ουσίες κυμαίνονται μεταξύ -5.30 και 5.94 kcal/mol. Είναι ενδιαφέρον το γεγονός ότι η ουσία που ξεχωρίζει, με κριτήριο την ενέργεια πρόσδεσης, είναι η μόνη ουσία που είναι υποκατεστημένη με ένα μόριο γλυκόζης, εμφανίζοντας σημαντική διαφορά από τις υπόλοιπες ουσίες που έχουν τον ίδιο σκελετό, επιβεβαιώνοντας έτσι το μοτίβο που εμφανίστηκε και στην κατηγορία των φλαβονοειδών και φλαβανών. Αναφορικά με το μοτίβο πρόσδεσης των χρωμονών στο ενεργό κέντρο, παρατηρούνται δύο επικρατέστερες τάσεις, μία όπου ο σκελετός της χρωμόνης βρίσκεται στην υποπεριοχή S2 και μία όπου βρίσκεται μεταξύ των περιοχών S1 και S1'.

Τα αποτελέσματα για τις λιγνάνες σίγουρα αξίζει να ληφθούν υπ' όψιν, καθώς οι ενέργειες πρόσδεσής τους κυμαίνονται από -7.175 έως -7.835 kcal/mol, με εξαίρεση το μεγαλύτερο μόριο αυτής της κατηγορίας, για το οποίο η ενέργεια υπολογίστηκε -6.564 kcal/mol. Γενικότερα το γεγονός ότι οι λιγνάνες είναι ογκώδη μόρια τους δίνει την δυνατότητα να καταλάβουν μεγαλύτερο όγκο στο ενεργό κέντρο, ωστόσο όπως φαίνεται από τα παραπάνω, πέρα από κάποιο όριο, το μέγεθος του μορίου αποτελεί περιοριστικό παράγοντα.

Μία επιπλέον ενδιαφέρουσα κατηγορία ουσιών είναι οι τριτερπενοειδείς σαπωνίνες (με βασικό σκελετό το πεντακυκλικό τριτερπενοειδές ολεανάνη), με ενέργειες πρόσδεσης μεταξύ -7.018 και -8.614 kcal/mol, που ξεπερνούν αυτήν του αναστολέα N3 σε πολλές περιπτώσεις. Η ουσία με τα καλύτερα αποτελέσματα αποτελεί ένα παράγωγο του ακεμπονικού οξέος εστεροποιημένο με μία μονάδα γλουκουρονικού οξέος στην θέση του C3, ενώ παρατηρήθηκε γενικότερα ότι οι σαπωνίνες με υποκατάσταση γλουκουρονικού οξέος στην ίδια θέση οδήγησαν σε καλύτερα αποτελέσματα. Ωστόσο, υπήρξε η ένδειξη ότι η εστεροποίηση με μόριο γλυκόζης στο καρβοξύλιο του C28 οδηγεί σε μείωση της ενέργειας πρόσδεσης. Στις περιπτώσεις όπου εξετάστηκαν και οι μεθυλεστέρες κάποιων σαπωνινών υποκατεστημένες με γλουκουρονικό οξύ, οι εστέρες έδειξαν να έχουν μια ελάχιστα βελτιωμένη συγγένεια με το ενεργό κέντρο. Γενικότερα οι σαπωνίνες που εξετάστηκαν μπορούν να χωριστούν σε τρεις κατηγορίες με βάση δομικές ομοιότητές τους: ακεμπονικό οξύ και τα παράγωγά του, γυψογενίνη και τα παράγωγά της και ολεανολικό οξύ και τα παράγωγά του. Όλες οι σαπωνίνες εμφανίζουν παρόμοιο μοτίβο πρόσδεσης στο ενεργό κέντρο, αλλά ειδικότερα παρατηρείται μεγαλύτερη ομοιότητα στα παράγωγα του ακεμπονικού και του ολεανολικού οξέος. Αξίζει να αναφερθεί επίσης πως στην κατηγορία αυτή συμπεριλαμβάνονται και ουσίες που ανιχνεύθηκαν για πρώτη φορά στο φυτό *Salicornia*: ο 3-O-β-D-γλουκουρονοπυρανοζυλ-28-O-β-D-γλυκοπυρανοζίτης του 3β-υδρόξυ-23-οξο-30-νορολεαναν-12,20(29)-διεν-28-οϊκού οξέος, οι ουσίες Salieuropaea A, Salbige A, Salbige B και ο 28-O-β-D-γλυκοπυρανοζυλ-εστέρας του 3β,-29-διυδρόξυ-ολεαν-12-εν-28-οϊκού οξέος. Όλες οδήγησαν σε ενέργειες πρόσδεσης μεγαλύτερες από -7 kcal/mol, με την καλύτερη τιμή, για το Salieuropaea A, να είναι ιδιαίτερα υψηλή, -8.498 kcal/mol.

Από τις υπόλοιπες ουσίες που εξετάστηκαν, οι οποίες δεν μπόρεσαν να ενταχθούν σε κάποια από τα παραπάνω κατηγορίες, οι περισσότερες δεν έδειξαν να έχουν ιδιαίτερα ευνοϊκές αλληλεπιδράσεις με την M^{pro}, με την πλειοψηφία των τιμών για την ενέργεια πρόσδεσης να είναι μικρότερη από -6.0 kcal/mol. Οι ουσίες με τα πιο ενδιαφέροντα αποτελέσματα ήταν η Pheophorbide A (-7.935 kcal/mol) και τα παράγωγά της (13^2 S)-Hydroxy-pheophorbide A (-7.414 kcal/mol) και (13^2 S)-Hydro-pheophorbide-lactone A (-8.103 kcal/mol), ενώ ξεχωρίζει και το ελλαγικό οξύ, με ενέργεια πρόσδεσης -7.386 kcal/mol. Δεδομένα στην βιβλιογραφία προς σύγκριση υπήρξαν μόνο για την πυρογαλλόλη και το ελλαγικό οξύ, επιβεβαιώνοντας τα παρόντα αποτελέσματα στην πρώτη περίπτωση, αλλά διαφέροντας σημαντικά τόσο όσον αφορά την ενέργεια, όσο και τον προσανατολισμό του μορίου για το ελλαγικό οξύ.

Τα αποτελέσματα της πειραματικής ανάλυσης έδωσαν σημαντικές ενδείξεις ανασταλτικής δράσης ουσιών που μελετήθηκαν παραπάνω. Αρχικά, η ανάλυση έγινε για τον αναστολέα GC376, για τον οποίο υπολογίστηκε η τιμή του IC_{50} (0.454 μM). Ουσίες που έχουν βρεθεί ως συστατικά του φυτού *Salicornia* εμφάνισαν 50% αναστολή του ενζύμου σε τουλάχιστον πενταπλάσια συγκέντρωση, έδειξαν όμως ότι έχουν παρεμποδιστική δράση. Από την κατηγορία των υδροξυκινναμικών οξέων, επιλέχθηκαν το φερουλικό και το ροσμαρινικό οξύ το πρώτο γιατί είναι το κύριο υδροξυκινναμικό οξύ που συναντάται στα κυτταρικά τοιχώματα των φυτών, και το δεύτερο λόγω του ότι είναι παράγωγο του καφεϊκού οξέος, με δύο φαινολικούς δακτυλίους, και εμφανίζει την καλύτερη ενέργεια πρόσδεσης στην πρωτεάση, επομένως αποτελεί μια ένδειξη για την τάση που εμφανίζουν τα παράγωγα των υδροξυκινναμικών οξέων. Πράγματι, το ροσμαρινικό οξύ ανέστειλε αποτελεσματικότερα την δράση της πρωτεάσης, με IC_{50} =801.45 μM έναντι της τιμής 3090.99 μM για το φερουλικό. Λόγω του ότι τα αποτελέσματα της μοριακής προσομοίωσης ήταν πολύ παρόμοια για όλες τις ουσίες στην κατηγορία των υδροξυβενζοϊκών οξέων, επιλέχθηκε μόνο μία από αυτές για *in vitro* ανάλυση, το γαλλικό οξύ, λόγω του ότι είχε ελαφρώς καλύτερα αποτελέσματα

προσομοίωσης. Από την ανάλυση υπολογίστηκε η τιμή του IC_{50} ίση με 4424.22 μM, μεγαλύτερη από την αντίστοιχη για το φερουλικό οξύ, όπως θα ήταν αναμενόμενο λαμβάνοντας υπ' όψιν ότι και από την προσομοίωση προέκυψε χαμηλότερη ενέργεια πρόσδεσης για το γαλλικό οξύ.

Από την κατηγορία των καφεοϋλκινικών οξέων επιλέχθηκε αρχικά το κινικό οξύ ως βασική δομική μονάδα, και στην συνέχεια το χλωρογενικό, το 3,4-δικαφεοϋλκινικό και το 3,5-δικαφεοϋλκινικό οξύ, ως βασικά και εμπορικά διαθέσιμα παράγωγα του. Τα δύο δικαφεοϋλκινικά οξέα οδήγησαν και σε πολύ υψηλές τιμές για την ενέργεια πρόσδεσης κατά την προσομοίωση. Το κινικό οξύ δεν έδειξε να έχει παρεμποδιστική δράση, τα παράγωγά του, ωστόσο, παρεμπόδισαν αποτελεσματικά την πρωτεάση, αναστέλλοντάς την εντελώς στην μεγαλύτερη συγκέντρωση που μελετήθηκε, των 5000 μM. Οι τιμές IC_{50} υπολογίστηκαν ίσες με 546.07, 503.59 και 597.81 μM για το χλωρογενικό, 3,4-δικαφεοϋλκινικό και 3,5-δικαφεοϋλκινικό οξύ, αντίστοιχα.

Επιπλέον, εξετάστηκαν η μυρικετίνη και η καιμπφερόλη, καθώς έχουν ήδη μελετηθεί στην βιβλιογραφία, παρέχοντας έτσι ένα μέτρο αξιολόγησης των αποτελεσμάτων. Σε σύγκριση με τις αναφορές (IC_{50}= 0.22 μM για την μυρικετίνη και IC_{50}=34.46 για την καιμπφερόλη), οι τιμές που υπολογίστηκαν στην παρούσα μελέτη (505.27 μM and 341.85 μM αντίστοιχα) είναι σημαντικά μεγαλύτερες, κάτι που μπορεί οφείλεται σε κάποιο βαθμό και στην χρήση διαφορετικής πειραματικής μεθόδου. Επιβεβαιώνουν όμως σε κάθε περίπτωση την ανασταλτική δράση των δύο ουσιών.

Επιπλέον εξετάστηκε και η κουερσετίνη καθώς και κάποια παράγωγά της (ισοκουερσετίνη, ρουτίνη, ισοραμνετίνη, 3-γλυκοζίτης της ισοραμνετίνης και 3-ρουτινοζίτης της ισοραμνετίνης). Η κουερσετίνη επιλέχθηκε γιατί είναι ένα από τα κύρια φλαβονοειδή, που όπως φαίνεται και από τα παραπάνω έχει πολλά παράγωγα. Τα παράγωγά της επίσης επιλέχθηκαν, ώστε να μπορεί να εξεταστεί το πως η γλυκοζυλίωση με σάκχαρα επιδρά στην παρεμποδιστική δράση των ουσιών. Τα αποτελέσματα επαλήθευσαν το μοτίβο που έγινε ορατό και από την προσομοίωση μοριακής πρόσδεσης, ότι γενικότερα η σύνδεση με σάκχαρα διευκολύνει την σύνδεση των ουσιών στο ενεργό κέντρο της πρωτεάσης. Για την κουερσετίνη η τιμή IC_{50} υπολογίστηκε ίση με 1910.96 μM, ενώ η ισοκουερσετίνη εμφάνισε σημαντικά μειωμένη την αντίστοιχη τιμή (IC_{50}= 605.13 μM). Ακόμα καλύτερα αποτελέσματα προέκυψαν για την ρουτίνη, με IC_{50}=286.93 μM, που ήταν και η χαμηλότερη τιμή που προέκυψε ανάμεσα στις υπό μελέτη ουσίες. Αντίστοιχα, ο γλυκοζίτης (IC_{50}= 586.31 μM) και ο ρουτινοζίτης (351.81 μM) της ισοραμνετίνης είχαν εντονότερη παρεμποδιστική δράση από την ισοραμνετίνη (IC_{50}=1435.99 μM), ειδικότερα με τον ρουτινοζίτη να είναι η ουσία που επιτυγχάνει 100% αναστολή του ενζύμου σε χαμηλότερη συγκέντρωση (1000 μM). Φαίνεται από τα παραπάνω ότι η ρουτινόζη ως υποκατάσταση επιφέρει καλύτερα αποτελέσματα από την γλυκόζη. Ένας ακόμα ρουτινοζίτης που μελετήθηκε είναι αυτό της πελαργονιδίνης, καθώς ήταν και από τις ουσίες με τα καλύτερα αποτελέσματα από την προσομοίωση, τα οποία επαληθεύθηκαν και πειραματικά, με μία χαμηλή τιμή για το IC_{50}=463.92 μM. Στην ανάλυση συμπεριλήφθηκαν και η εσπερετίνη και η εσπεριδίνη, ως ένα αντιπροσωπευτικό ζεύγος φλαβανόνης και αντίστοιχου γλυκοζίτη. Καμία από τις δύο ουσίες δεν έδειξε επιτυχή αναστολή, παρά την μειωτική τάση που εμφάνισε η ενεργότητα του ενζύμου, κάτι που μπορεί να οφείλεται στην περιορισμένη διαλυτότητα των ουσιών σε μεγάλες συγκεντρώσεις. Μελετήθηκαν επίσης η ακακετίνη, η γκαλανζίνη, η απιγενίνη, η χρυσίνη και η κατεχίνη, ως βασικοί σκελετοί φλαβονοειδών και φλαβανονών. Η ακακετίνη, η γκαλανζίνη και η χρυσίνη δεν έδωσαν ενδείξεις δράσης, το γεγονός όμως ότι δεν κατέστη δυνατή η διαλυτοποίησή τους σε

συγκεντρώσεις μεγαλύτερες των 250 μM εμποδίζει κάποιο οριστικό συμπέρασμα για τις ιδιότητες των ουσιών από το να εξαχθεί. Η απιγενίνη και η κατεχίνη έδειξαν να παρεμποδίζουν το ένζυμο σε συγκεντρώσεις μεγαλύτερες των 250 μM, με IC_{50}=604.07 και 928.55 μM, αντίστοιχα. Συνολικά, από τα μη υποκατεστημένα φλαβονοειδή, τα καλύτερα αποτελέσματα προέκυψαν για την καιμπφερόλη, επομένως, όπως προκύπτει από τα υπόλοιπα δεδομένα, θα ήταν ενδιαφέρον να μελετηθούν στο μέλλον παράγωγα της καιμπφερόλης, όπως εστέρες με μονοσακχαρίτες ή δισακχαρίτες. Πολύ θετικό είναι επίσης το γεγονός ότι το εκχύλισμα του φυτού *Salicornia* που εξετάστηκε επίσης εμφάνισε παρεμποδιστική δράση, με IC_{50}=400.66 μM, χαμηλότερο και από την πλειοψηφία των επιμέρους ουσιών που εξετάστηκαν, υποδηλώνοντας ότι τα διαφορετικά συστατικά του μπορούν να δράσουν συνεργιστικά.

Από την πειραματική ανάλυση προκύπτει επίσης συσχέτιση των αποτελεσμάτων της προσομοίωσης μοριακής πρόσδεσης με την τιμή που υπολογίστηκε για το IC_{50} , με συντελεστή συσχέτισης 0.8189. Η συσχέτιση αυτή, παρ' ότι όχι απόλυτη, υπογραμμίζει την χρησιμότητα της προσομοίωσης μοριακής πρόσδεσης ως ένα μέσο για την αρχική εκτίμηση των ιδιοτήτων των ουσιών και την εδραίωση μιας σχέσης δομής-δράσης μεταξύ του ενζύμου και του υπό μελέτη αναστολέα. Συνολικά, τα αποτελέσματα είναι πολύ ενθαρρυντικά και αναδεικνύουν μία πληθώρα φυτοχημικών, καθώς και το εκχύλισμα του φυτού *Salicornia*, ως βιοδραστικές ουσίες με ενδιαφέρουσες προοπτικές στην ενίσχυση του ανοσοποιητικού μας συστήματος ενάντια στον SARS-CoV-2. Το γεγονός ότι η εκχύλιση των ουσιών αυτών μπορεί να πραγματοποιηθεί και από μέρη του *Salicornia* που συνήθως απορρίπτονται προσφέρεις νέες εναλλακτικές και στην βιώσιμη αξιοποίηση της βιομάζας για τη παραγωγή προϊόντων προστιθέμενης αξίας. Σημαντικός παράγοντας για την χρήση των παραπάνω ουσιών και την απορρόφησή τους είναι, όπως φάνηκε και από την παρούσα έρευνα, η διαλυτότητα, η οποία αποτελεί σε κάποιες περιπτώσεις περιοριστικό παράγοντα. Ένας τρόπος να βελτιωθεί αυτή η ιδιότητα είναι η σύζευξη των ουσιών με σάκχαρα, αυξάνοντας την υδροφιλικότητά τους. Η βιοτεχνολογία έχει αναπτύξει βιώσιμες και εκλεκτικές μεθόδους για την πραγματοποίηση τέτοιων αντιδράσεων, με την χρήση ενζύμων, όπως για παράδειγμα οι φερουλικές εστεράσες ή οι λιπάσες, και οι τρανς-γλυκοζυλάσες παρέχοντας πολλές δυνατότητες περαιτέρω μελέτης της δράσης των παραπάνω ουσιών. Από την διπλωματική αυτή εργασία προκύπτουν ενθαρρυντικά δεδομένα για τις προοπτικές των συστατικών του αλόφυτου *Salicornia* ως αναστολείς της κύριας πρωτεάσης του SARS-CoV-2. Όμως για την πλήρη αξιολόγηση των ιδιοτήτων τους, χρειάζεται να γίνει πειραματική ανάλυση και σε κύτταρα, αλλά και *in vivo* μελέτη, ως μέρος μελλοντικής έρευνας.

Abstract

The world is currently going through the second year of a pandemic, which started on March, 2020, and has had numerous cases and victims and enormous consequences on social and economic life. The cause of this pandemic is the newly identified SARS-CoV-2, an RNA virus of the family of *Coronoviridae*. Although various vaccines have been developed and vaccinations are ongoing, the antiviral drugs employed are limited and mostly already known repurposed drugs, so there is a need for additional ways to boost our defense against the virus. Phytochemicals emerge as a possible immune boosting solution that can act synergistically with pharmaceutical products, since many of them have proved to be active against various viruses. Particularly the extract of halophyte *Salicornia* contains a broad variety of compounds (hydroxycinnamic acids, hydroxybenzoic acids, caffeoyl quinic acids and their derivatives, flavonoids and flavanones, sterols, chromones, lignans, oleanane triterpenoid saponins), including molecules with confirmed antiviral properties, among numerous health benefits.

Aim of this thesis is to initially utilize *in silico* methods (molecular docking using the YASARA Structure software), to perform a screening of the contents of the *Salicornia* extract for their inhibitory potential against the main protease of SARS-CoV-2 (M^{pro}), whose vital role in viral replication makes it an ideal target for the development of antiviral agents. As second step, the most promising compounds were tested *in vitro*, using an enzyme inhibition assay, together with an extract from a *Salicornia* plant.

Docking and visualizing already confirmed and co-crystallized M^{pro} inhibitors was done in order to establish the method and obtain additional data on the binding mode mechanisms that result in effective inhibition. The simulation was then performed for the *Salicornia* constituents and resulted in an assessment of binding energies and contacting residues between the protease and each tested compounds. Caffeoylquinic acids and their z derivatives together with flavonoids and flavanones were highlighted as the most promising groups of compounds, with binding energies ranging from -7.9 to -8.935 kcal/ mol for the first group (excluding quinic acid) and -6.88 to -9.384 kcal/mol for the second group. The latter binding energy corresponds to the highest scoring compound, isorhamnetin-3-O-rutinoside. Patterns connecting the structure of the compounds and their binding affinity to the active site of M^{pro} were also detectable, the major one being that glycosylated compounds have a higher binding affinity to the enzyme than their parent structures.

In vitro screening involved a selection of compounds based on the results of the preceding step, their commercial availability and how well they represent the variety of compounds present in the extract. Results were very encouraging, with the majority of the compounds inhibiting the activity of M^{pro} and a correlation between the molecular docking results and the IC_{50} (the concentration of a compound that results in 50% inhibition of the enzyme) calculations being indicated. The compound with the lowest IC_{50} was rutin (IC_{50}=286.93 μM), followed by kaempferol (IC_{50}=341.85 μM) and isorhamnetin-3-rutinoside (IC_{50}=351.81 μM). Very promising results were also yielded for the crude *Salicornia* extract, which showed inhibitory activity with an IC_{50} of 400.66 kcal/mol. The assay results mainly confirmed the molecular docking results, providing useful information on which further investigation, both *in vitro* in cells and *in vivo*, could rely on. From the present findings, it is suggested that *Salicornia* extract and its contents can be valuable nutraceuticals and potential contributors to the fight against the ongoing pandemic.

Contents

1. INTRODUCTION

1.1. The Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) and emerged pandemic

As of the beginning of 2020 up until now, the world is going through a pandemic, which apart from a severe public health crisis, counting more than 219 million cases and more than 4,5 million deaths, has had a tremendous impact on economic and social life. In December 2019, in the city of Wuhan, Hubei province, China, a series of pneumonia cases were reported, exhibiting symptoms such as fever, dry cough, chest discomfort or even dyspnea and bilateral lung infiltration. The first case believed to have appeared on 8 December, and by the end of 2019, 27 confirmed cases where reported. The local outbreak was further investigated and led to the identification of a novel coronavirus, which was later given the name Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2). The disease caused by the virus was also named as COVID-19 (Coronavirus disease 2019) and was widely spread all over the world, resulting in the World Health Organization (WHO) declaring a pandemic on 11 March, 2020 (Hu et al. 2021; Wu et al. 2020)

SARS-CoV-2 is the third coronavirus creating a public health concern in the past 20 years, after the severe acute respiratory syndrome-coronavirus (SARS-CoV) and Middle East respiratory syndrome (MERS-CoV), which created an outbreak in 2002 and 2012, respectively. SARS-CoV-2 shares common genomic sequence by a percentage of 79% with SARS-CoV and 50% with MERS (Stoddard et al. 2020). A coronavirus detected in bats, RaTG13-CoV, hosted by *Rhinolophus affinis*, shows 96.2 % genome sequence identity with SARS-CoV-2. Therefore, it is widely believed that the novel coronavirus originated from RaTG13 and was transmitted to humans through an intermediate host, reminding that the possibility of a viral spillover, with serious effects on humans, is continuously present. Pangolin was suggested as a potential intermediate host, as there have been detected strains of coronavirus in pangolins in the area, having a similar genome sequence with SARS-CoV-2 by around 92 %. However, the evidence is not conclusive, since Pangolin-CoV lacks a peptide needed for the proteolytic cleavage of the spike protein that the virus uses to attack host cells. In addition, the fact that pangolins also exhibit symptoms of disease, due to infection from the coronavirus, suggests that they are not a natural reservoir (Adil et al. 2021; Friend and Stebbing 2021; Hu et al. 2021).

SARS-CoV-2 belongs to the family of *Coronaviridae*, and is a *Betacoronavirus* of the subgenus *Sarbecovirus*. It is a positive-sense single-stranded RNA (+ssRNA) virus, with a genome with a size of about 29.8 kb (Xu et al. 2020). As seen in Figure 1, it has an almost spherical shape and is surrounded by an envelope made of a lipid bilayer, onto which spike proteins are attached. The N-terminal of the spike protein attaches to the host's receptor, angiotensin-converting enzyme 2 (ACE2). The C-terminal contributes to merging the viral and the cellular membrane, with the necessary contribution of a cellular protease of the host cells, particularly transmembrane protease serine 2 (TMPRSS2), which cleaves the spike after it is bound to ACE2. The virus initially infects epithelial cells in the upper respiratory tract, moving then onto epithelial cells in the lungs (Matheson and Lehner 2020). The fact that ACE2 is mainly expressed in alveolar epithelial type II cells can explain why lungs are the main target of the virus, in combination with the presence in the cells of genes that facilitate viral replication. Another contribution to that is the accessibility of the lungs, due to their large surface area. Apart from the lungs, impairment of the function of other organs has also been observed. That can be attributed to the presence of ACE2-expressing cells in the heart, kidney, endothelium

and intestine. Additionally, considerable expression of the enzyme in the lumine suggests that epithelial cells of the intestine can act as receptors too, providing an additional entry point for the virus (Zhang et al. 2020; Adil et al. 2021). People of all ages are subject to infection from the virus, and it is observed that males are more easily attacked, as the virus receptor protein is expressed higher in male cells than female (Wang et al. 2021).

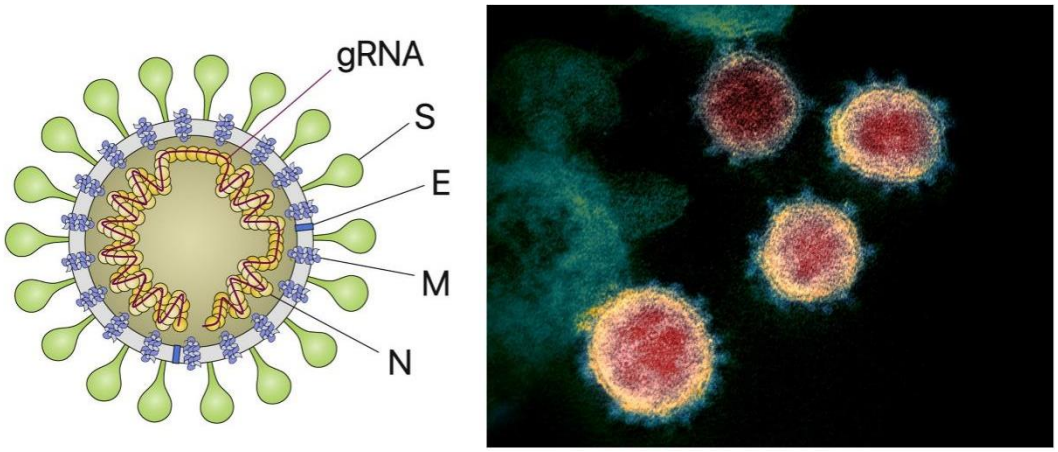


Figure 1: Graphic depiction of SARS-CoV-2 (gRNA: genomic RNA; S: spike protein; E: envelope protein: M: membrane protein, N: nucleocapsid protein) as presented by D. Kim et al (2020) (left) and image of the virus captured on scanning and transmission electron microscope by the National Institute of Allergy and Infectious Diseases (NIAID) Rocky Mountain Laboratories (right).

1.2. Targets for blocking the activity of SARS-CoV-2

The emergence of coronavirus-related diseases highlights the importance of the development of ways of defense and immunity boosting. Eliminating SARS-CoV-2 infection, includes blocking any step within the pathway of viral entry, replication and release of new viruses in the human organism (Figure 2). The phylogenetic similarity with previous coronaviruses has provided a lead in drug development research, as drugs with antiviral activity against SARS-CoV are likely to be able to inhibit SARS-CoV-2, too. Apart from that, ongoing research is being conducted, both towards interfering with the functional and structural proteins of the virus and towards acting on the host, to reinforce immune response or block proteins that assist the viral entry in the cells and replication. The spike protein, a structural protein of the virus, is the main antigen for which the vaccines have been developed (Creech et al. 2021), but repurposed and newly designed drugs target functional proteins, virulence factors or host proteins that are useful to viral reproduction (Gil et al. 2020).

Functional proteins can be a main target to block viral activity. **The main functional proteins of SARS-CoV-2 include the main protease (M^{pro}, $3CL^{pro}$ or nsp5), papain-like protease (PL^{pro}), RNA-dependent RNA polymerase (RdRp) and helicase (nsp13)**. M^{pro} cleaves the formation of non-structural proteins (nsps), including itself, from overlapping viral polyproteins pp1a and pp1ab, at at least 11 cleavage sites. These nsps are essential parts of the replication and transcription complex of the virus. Thus, M^{pro} is necessary for the formation of functional components that SARS-CoV-2 needs to reproduce. Similarly to M^{pro}, PL^{pro} also releases three non-structural proteins from the initial polyprotein, which play a role in correcting the replication of the virus (Arya et al. 2021). RdRp is a structure mainly consisting of catalytic non-structural protein (nsp12) and two assisting nonstructural proteins nsp7 and nsp8, all of which are released by M^{pro}. This complex is key for replicating and transcribing the viral genome, as it catalyzes the polymerization of RNA. The U.S. Food and Drug Administration-approved anti-

SARS-CoV-2 drug remdesivir acts by inhibiting RdRp (Hillen et al. 2020). Nsp13 is another enzyme, conserved amongst coronavirus species, that is needed for replication, as it unwinds double strands of DNA and RNA in a 5'-3' direction, through an NTP-based reaction (Habtemariam et al. 2020).

Attacking the structural proteins of the virus is another strategy that essentially prevents the binding of the virus to the receptor and its self-assembly. **The main structural proteins of SARS-CoV-2 include the spike protein (S), the small envelope protein (E), the membrane protein (M) and nucleocapsid protein (N).** As described above, the spike glycoprotein is responsible for the binding of the virus to the host cells and the fusion of the viral and cellular surfaces. It is bound to the outside part of the viral envelope and has the shape of a spike. The protein consists of three subparts, S1, S2 and S2'. S1 is the domain that recognizes the receptor and binds to it, S2 contributes in merging the membranes and S2' is a fusion peptide. Spike protein is of prominent importance for the entry of the virus into the host cells, as any variations in this protein affect the way and the type of cells the virus attacks. Thus, the spike protein is an antibody target and also the focus of vaccine development. The envelope protein is crucial for the morphogenesis of the virus. It also creates ion channels through which the virion communicates with its environment and regulates protein transfer, consequently being important for the biological functions of the virus. Similarly, the membrane protein, a glycoprotein with three transmembrane domains, has a substantial structural role. Together with the other structural proteins, it provides a frame for viral RNA, in addition to maintaining intracellular equilibrium of metabolite concentrations. The nucleoprotein also plays a role in the assembly of the virus, by assisting with incorporating viral RNA into a nucleocapsid. Moreover, it is necessary in other parts of the life cycle of the virion, such as organization of the cytoskeleton and host cell apoptosis (Shamsi et al. 2021).

Targeting the virus virulence factors is also a defense mechanism against it. More specifically, nsp1 is a virulence factor that destroys host mRNA and blocks the production of type-1 interferon. ORF7 blocks the bone marrow matrix antigen 2 (BST-2), that functions as an inhibitor of the release of new SARS-CoV-2 cells from the already infected ones. Nsp3c is another factor that fights host immunity response by binding to its ADP-ribose (Wu et al. 2020).

Lastly, another way to block viral infection is to block host proteins that are useful to viral reproduction. For example, the binding site of the receptor, ACE2, can be occupied, so that the virus cannot bind to the host cells (Wu et al. 2020). Wang et al. (2021) also mention antiviral activity exhibited by compound S416, a compound that inhibits pyrimidine synthesis, which is essential for viral replication by targeting its rate-limiting enzyme, dihydroorotate dehydrogenase (DHODH). Another crucial part of the viral cell entry is Transmembrane Protease Serine 2 (TMPRSS2), which activates the spike protein and consequently facilitates the binding of the virus to the host cells. TMPRSS2 is a confirmed antiviral target, as already investigated inhibitors prove that it blocks viral cell entry (Hoffmann et al. 2020).

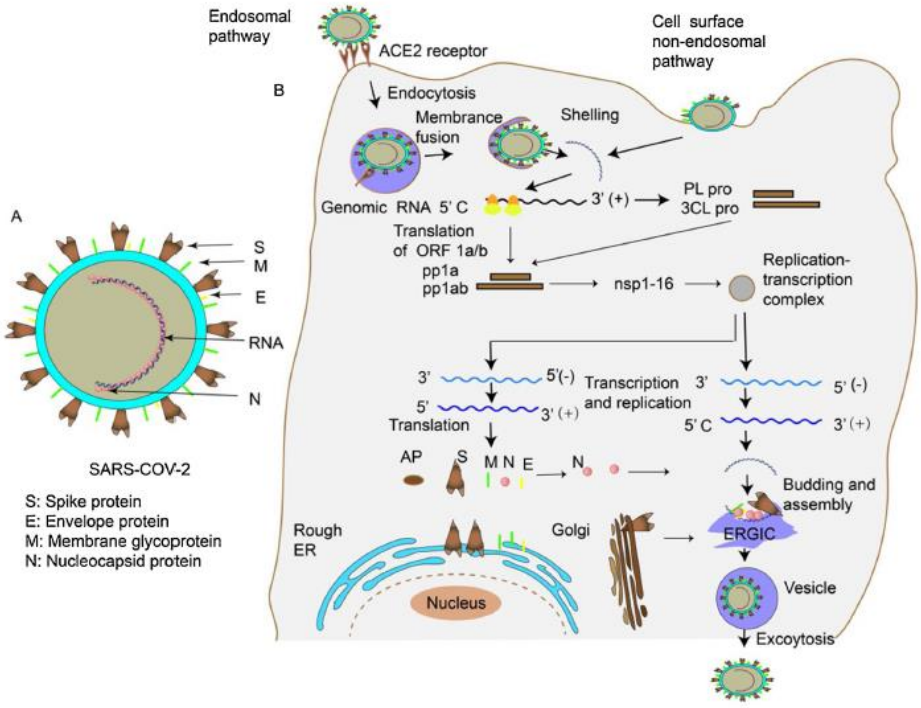


Figure 2: Mechanism of SARS-CoV-2 infection as presented in Huang et al. (2020)

1.3. The main protease of SARS-CoV-2 (M^{pro})

The translation of the viral RNA of SARS-CoV-2, once it enters the host cells, leads to the synthesis of two polyproteins, pp1a and pp1ab. After auto-processing its own N- and C-terminals to release itself from the polyproteins, SARS-CoV-2 M^{pro} cleaves the peptide bonds of pp1a and pp1ab, catalyzing the formation of nonstructural proteins necessary for the construction of the replication transcription complex that the virus needs to synthesize new RNA (Koudelka et al. 2021; Jin et al. 2020a; Hegyi and Ziebuhr 2002). The proteolysis takes place in more than 11 cleavage sites. The amino acid sequence that the enzyme recognizes as a cleavage site is (Leu-Gln)-(Ser/Ala/Gly), with the peptide bond being hydrolyzed after Gln. Koudelka et al. (2021) also mention that SARS-CoV-2 M^{pro} is able to cleave human proteins as well. More specifically, optineurin, a protein that participated in activating innate immune response during viral infection, was found to have two potential cleavage sites where M^{pro} could act. **The vital role of M^{pro} in the reproduction of SARS-CoV-2 and the release of many of its proteins, combined with the fact that its structure and mechanism have been investigated, make it a very appealing target to block viral activity. Moreover, the fact that there is no human enzyme cleaving proteins after the Gln residue, is another advantage of M^{pro} as target for the development of inhibitors to act as antiviral drugs or immune-boosting compounds** (Dai et al. 2020; Mengist et al. 2020; Zhang et al. 2020b; Kneller et al. 2020a; Świderek and Moliner 2020).

1.3.1. Structure

SARS-CoV-2 M^{pro} (EC 3.4.22.69) is a cysteine protease and a member of the PA clan of proteases. Proteases are enzymes that hydrolyze peptide bonds and thus belong to the category of hydrolases. The first crystal structure of SARS-CoV-2 M^{pro} was determined by X-ray diffraction at a resolution of 2.16 Å and was deposited at the Protein Data Bank (PDB) by Liu et al. and released on February 5, 2020, under PDB ID 6LU7. Since then, many structures of the protease have been deposited, including the enzyme co-crystallized with various

inhibitors. The active form of the enzyme is a homodimer (Figure 3). The structure of a single monomer consists of a 306-residue-long polypeptide chain, which can be divided into three domains: domain I (residues 8–101), domain II (residues 102–184) and domain III (residues 201–303). Domains I and II are composed of antiparallel β-barrels and host the active site in a cleft formed between them, whereas domain III consists of 5 α-helices and plays a role in the dimerization of the enzyme. Residues 185-200 form a loop that connects domains II and III (Kneller et al. 2020a; Jin et al. 2020a; Zhang et al. 2020b). The enzyme is active only in as a dimer because the NH_2-terminal of each protomer interacts with residue Glu 166 of the other protomer and contributes to the formation of the S1 subsite of active site (Sacco et al. 2020). This interaction results in the NH_2-terminal of a monomer being positioned between domains II and III of this monomer and domain II of the other. The dimeric structure of the enzyme is regulated through a salt-bridge between residues Glu 290 of one protomer and Arg 4 of the other (Zhang et al. 2020b).

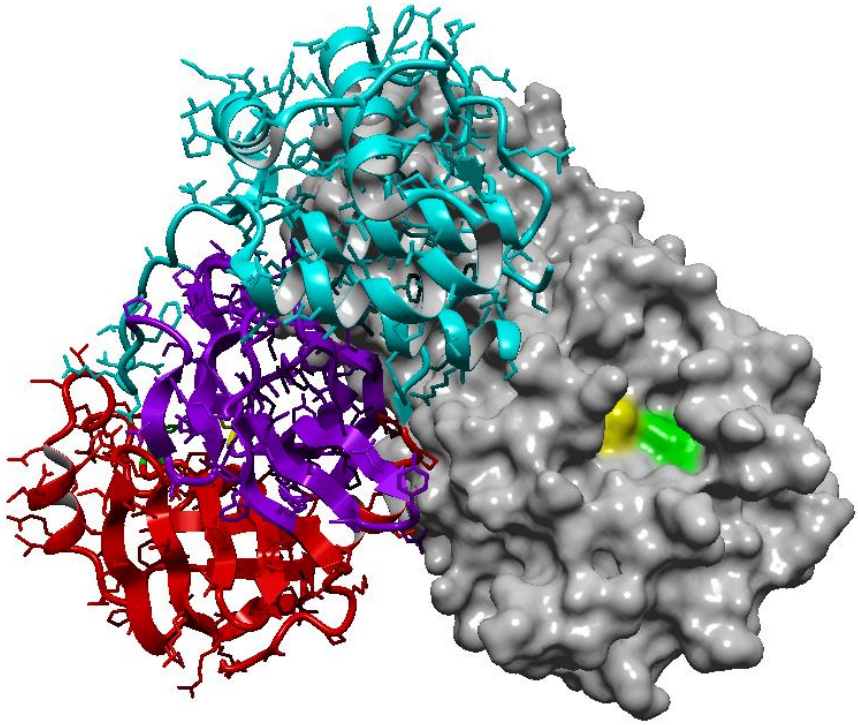


Figure 3: SARS-CoV-2 M^{pro} in the active form of a homodimer (PDB 7JKV). The right monomer is shown as surface while the left portrays the secondary structure and the three domains of the enzyme. Domain I is in red, domain II in purple and domain III in cyan. Catalytic residues His 41 and Cys 145 are highlighted in yellow and green respectively. The molecule was visualized in YASARA Structure.

At its active site, the enzyme has a cysteine-histidine catalytic dyad (Cys 145-His 41). The existence of the stabilizing oxyanion hole, consisting of residues Gly 143, Ser 144 and Cys 145, is also noteworthy. During catalysis, the negative charge of the carbonyl oxygen in the scissile bond of the natural substrate of the protease is being balanced by the oxyanion hole. It is also reported that the oxyanion hole similarly stabilizes inhibitors, as many of them form a hemithioacetal intermediate with a negatively charged oxygen atom and bind to the Cys 145 residue of the protease with a similar geometry as the tetrahedral intermediate formed by the natural substrate (Zhang et al. 2020b; Świderek and Moliner 2020; Kneller et al. 2020a; Kneller et al. 2020b). The catalytic mechanism will be further analyzed below.

Except for catalytic dyad (Cys 145, His 41), the active site of M^{pro} is demarcated by residues Ser 46, Gln 189, Thr 190, Ala 191, Pro 168, Glu 166, Leu 141 and Asn 142 (Kneller, Phillips, et al. 2020b). It consists of four main subsites, S1, S1' S2 and S4 (Figure 4), similar to the active sites of the main proteases of other coronaviruses (Qiao et al. 2021; Dai et al. 2020). More specifically, out of the 306 residues of the protease sequence, only 12 are different between the main proteases of SARS-CoV-2 and SARS-CoV, which corresponds to 96% identity (Griffin 2020). As described by Stoddard et al. (2020), **S1** subsite consists of the side chains of Phe 140, Asn 142, Ser 144, Cys 145, His 163, Glu 166, His 172, and the backbone of Leu 141, Gly 143, His 164 and Met 165. However, Bai et al. (2021) mention Phe 140, Tyr 161, His 162, Glu 166 and His 172 as the key residues forming the S1 subsite whereas Jin et al. (2020a) point out the side chains of Phe 140, Asn 142, His 163, Glu 166, His 172 of one protomer and the backbone of Phe 140 and Leu 141 of the other as the parts of this subsite. **S1'** subsite is formed by the side chains of Thr 25, His 41, Val 42, Asn 119, Gly 143, Cys 145 and the backbone of Thr 26. **S2** is created by the side chains of His 41, Met 49, Tyr 54, Asp 187 and the backbone of Arg 188. According to Jin et al. (2020a) though, S2 subsite is a hydrophobic cleft, formed by the side chains of His 41, Met 49, Met 165 and the alkyl part of the side chain of Asp 187 of the other protomer. **S4** is made up of the side chains of Met 165, Leu 167, Pro 168, Ala 191, Gln 192 and the backbones of Glu 166, Arg 188, Thr 190.

From the above description of the active subsites and the slight differences found in literature, it is obvious that the borders of each subsite are not entirely specific and strictly defined. For example, Stoddard et al. (2020) depict similarly the active subsites (Figure 5a), but also mention an additional accessible cleft named S6. Dai et al. (2020) note the cavities that represent each binding subsite, but do not define clear limits between them (Figure 5b). Other studies mention additional binding pockets. For example, Świderek and Moliner (2020) refer to S3 cleft, located next to S4. Lockbaum et al. (2021) divide the active site in more subsites, including S3 and an additional S2' (Figure 5c). In both cases, no detailed description of the subsites is being given, however they are presented in the following image.

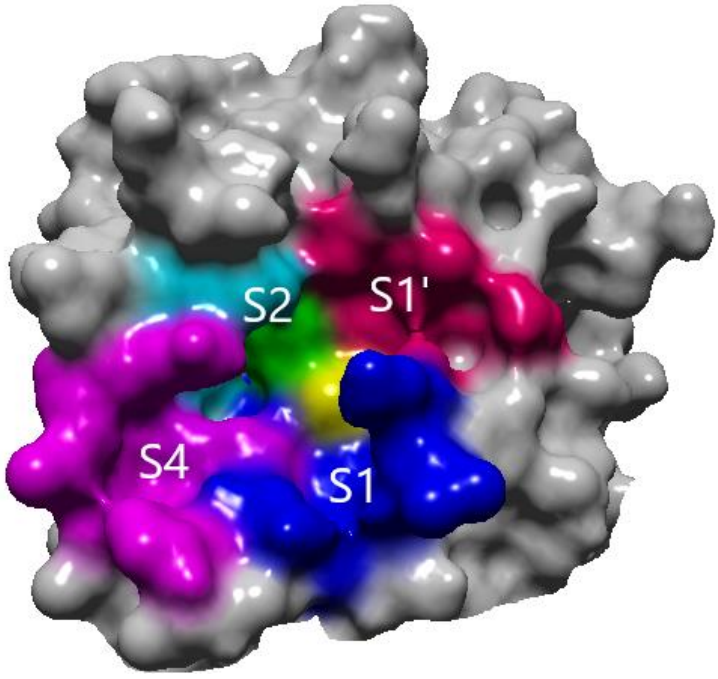
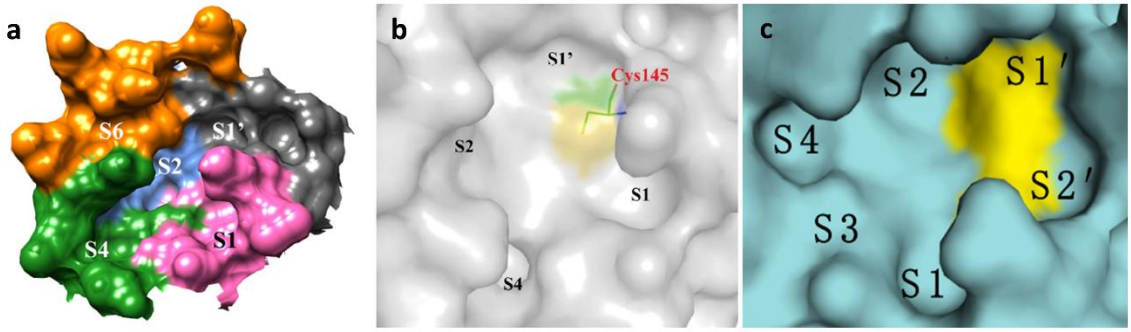


Figure 4: M^{pro} subsites, colored and marked on the image. Catalytic residues are also marked (His 41 in green and Cys 145 in yellow). Active site visualized in YASARA Structure. (PDB:6LU7).



Stoddard et al. (2020) Dai et al. (2020) Lockbaum et al. (2021)

Figure 5: SARS-CoV-2 M^{pro} active site with subsites marked, as presented in different studies

Overall, subsites S1 and S1' exhibit a higher degree of conservation, both in the different types of coronaviruses and when in complex with different substrates, whereas residues in subsites S2 and S4 are more mobile. Shitrit et al. (2020), after superposing crystal structures of SARS-CoV-1 and SARS-CoV-2 M^{pro} with different inhibitors, suggest that residues Gln 189, Met 49 and Asp 142 are residues of the binding site that show the greatest flexibility and variation between different ligands binding. More specifically, as far as residue Met 49 is concerned, its position has an effect on the size of S2 binding pocket, thus playing an important role in the ligand binding (Stoddard et al. 2020). On the other hand, they point out residues His 163 and Glu 166 as residues with which the vast majority of inhibitors form key interactions for inhibition. These conclusions are supported by the work of Gimeno et al. (2020), who claim that residues Met 49 and Arg 188 in the S2 subsite and Met 165 and Gln 189 in the S3 subsite are the most susceptible to displacement in the binding pocket of M^{pro}, while S1' subsite is the most stable part of the active site, followed by S1 where only Ser 1 and Asn 142 side chains exhibit subtle variation.

The malleability of the protein allows it to be arranged in different conformations, depending on the substrate or inhibitor that binds to it. As Stoddard et al. (2020) point out, since the available crystal structures provide an image of the enzyme at a specific moment, usually when it is in complex with an inhibitor and therefore with a certain conformation, molecular docking simulation results may differ considerably depending on the receptor's structure used. Kneller, Phillips, et al. (2020b) describe in detail the changes in the 3D structure of the protein upon ligand binding and specifically when peptide-like inhibitor N3 is bound to the active site (Ligand-free PDB structure: 6WQF; Inhibitor-bound PDB structure: 6LU7). Residues 46-50 that form a small helix close to P2 group of N3 move away from the β-hairpin loop arranged by residues 166-170, while the loop surrounding P5 group (residues 190-194) approaches it. The side chains of residues Met 49 and Met 165 change conformation in order to drift away from P2 group and specifically the leucine it includes, causing a movement of Ser 46 and Leu 50 residues. Also, the C-terminus of the protein (residues 301-306) flips its position by 180° when N3 binds to the enzyme, something that potentially destabilizes the dimer due to reduction of hydrogen bonds. Molecular dynamics simulation has shown great plasticity of P2 helix, P5 loop and the C-terminus, indicating that these regions could be accessible for binding by a greater variety of chemical compounds.

1.3.2. Catalytic mechanism

The mechanism through which proteolytic cleavage is conducted by SARS-CoV-2 M^{pro} is not studied in depth. However, due to the high similarity of the enzyme with the main proteases of other coronaviruses, especially SARS-CoV, very plausible hypothesis about its catalytic mechanism can be made.

Świderek and Moliner (2020) have used computational methods to deduce this catalytic mechanism, which is presented in Figure 6. The cleavage of the peptide bond is suggested to be initiated by a proton transfer from the thiol group of Cys 145 to the imidazole of His 41. Then, a highly reactive nucleophilic ion pair is formed. The Cys residue attacks the carbonyl portion of the scissile peptide bond, forming a thiohemiketal intermediate, while the protonated His attacks the N-atom of the peptide bond, creating the acyl-enzyme complex intermediate. A polypeptide chain is released as the first product of the reaction. Then an active water molecule attacks the carbonyl carbon atom of the Gln residue, whereas His is being reprotonated, no longer maintaining the acyl-enzyme complex. Lastly, Cys 145 is released as the covalent bond with the peptide is broken. The water molecule taking part in the above series of reactions is also part of interactions between residues His 41, His 164 and Asp 187, balancing the polar contacts between them. Kneller, Phillips, et al. (2020a) have pointed out its role, characterizing it a part of a potential non-canonical catalytic triad.

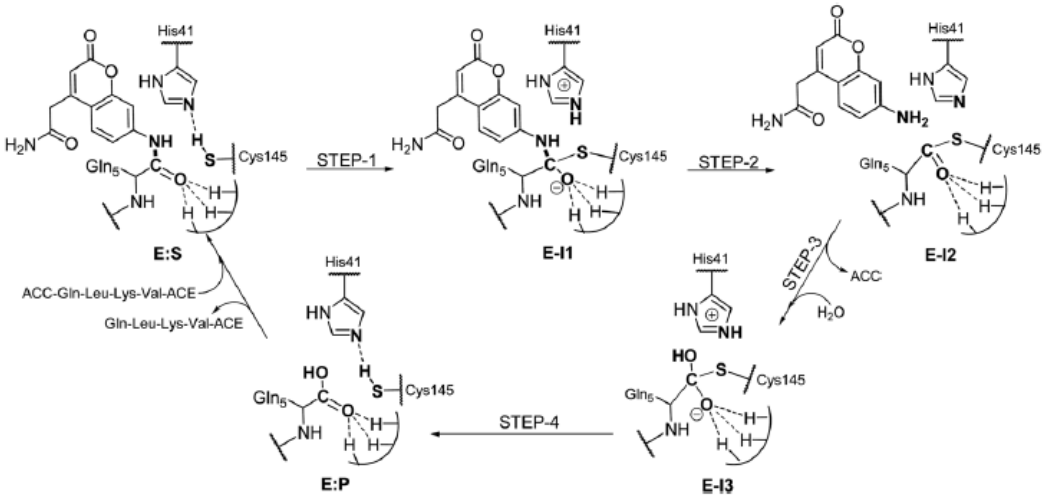


Figure 6: Steps of the mechanism of peptide bond hydrolysis by SARS-CoV-2 M^{pro} (Świderek and Moliner 2020)

1.4. Types of enzyme inhibition

Enzyme inhibition is the partial or complete inactivation of an enzyme by a compound, and can happen in multiple ways. An initial discrimination could be between reversible and irreversible inhibition. Irreversible inhibitors react with the protein to create a covalent bond, usually destroying a functional part of it (Sharma 2012). Although in the majority of the cases irreversible inhibitors bind covalently to the enzyme, they can also interact non-covalently. In any case, the interactions are strong and the complex is hard to separate. Reversible inhibitors, on the other hand, can interact both covalently and non-covalently with the enzyme, but they are quickly disconnected from the protein, leaving it intact (Berg et al. 2002).

Reversible inhibition can be divided into four categories, also depicted in Figure 7:

- Competitive inhibition: The inhibitor and the substrate, which are often in this case structurally similar, compete with each other for binding to the active site. Therefore, a competitive inhibitor acts by allowing less enzyme molecules to host the substrate. As far as the effect on the kinetic constants of the enzyme is concerned, competitive inhibition does not affect the maximum velocity of the enzyme (V_{max}) but results in an increase of the enzyme-substrate dissociation constant K_m (Todhunter 1979; Engelking 2015)

- Uncompetitive inhibition: The inhibitor can only bind to the enzyme-substrate intermediate. That results in a reduction of both kinetic constants V_{max} and K_m, which corresponds to reducing the speed of the catalytic reaction and increasing the binding affinity of the enzyme to the substrate, meaning that it is harder for the enzyme-substrate complex to disassociate and proceed to the formation of the product (Dougall and Unitt 2015; Palmer and Bonner 2011).

- Mixed inhibition: Although the term is often being used as a synonym for non-competitive inhibition, mixed inhibition is defined in many sources as more general form of non-competitive inhibition. The inhibitor is able to bind both to the free enzyme and to the enzyme-substrate complex, but the binding affinity is different to each one of them. It decreases both the enzyme units available for substrate binding and the turnover rate of the enzyme. This is portrayed by an apparent decrease in V_{max} and apparent increase in K_m (Ochs 2000; Saboury 2009; Todhunter 1979).

- Non-competitive or allosteric inhibition: A non-competitive inhibitor binds to the enzyme at a site different from the active site (allosteric site). The affinity of the inhibitor to the enzyme is the same as to the enzyme-substrate complex, so the binding of the substrate and the inhibitor are independent events, but the inhibitor reduces the catalytic activity by altering the structural conformation of the protein. This type of inhibition reduces the turnover rate of the enzyme, which means that catalysis is being slowed down, while the affinity of the enzyme to the substrate remains intact. The terms allosteric and non-competitive inhibition are usually used interchangeably. The difference between mixed and non-competitive inhibition can be seen through the kinetic constants: non-competitive inhibition leads to a reduction of the value of V_{max} but no apparent change in K_m (Aldred et al. 2009; Delaune and Alsayouri 2020).

Another categorization of inhibitors is between covalent and non-covalent inhibitors. Covalent inhibitors generally include a reactive group (e.g., hydroxyl, epoxy, carbonyl) and react with the active site of the enzyme forming a covalent bond with a nucleophilic residue, such as Cys, Ser, Thr or Lys. It is often mentioned that the covalent inhibition is time-dependent, meaning that the covalent bond is not formed immediately and it is preceded by the formation of a non-covalent complex (Awoonor-Williams and Abu-Saleh 2021). Non-covalent inhibition is usually achieved small molecules, whose shape and interactions with the

active site, including hydrogen bonds, van der Waals interactions and salt bridges enhance the ability to block the catalytic center of the enzyme (Aljoundi et al. 2020).

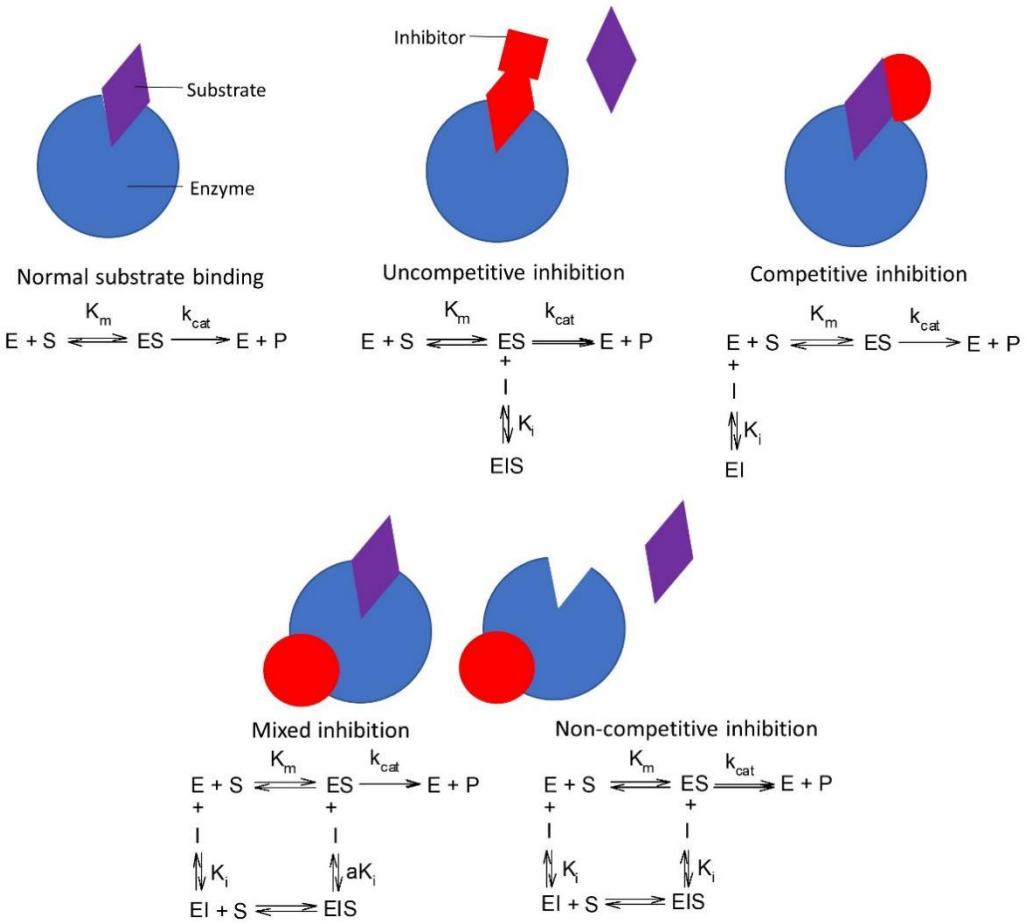


Figure 7: Types of reversible inhibition

1.5. Substance efficacy

There are various measurable quantities that demonstrate how effective a substance is as a virus inhibitor. The quantities more often used in literature describe the inhibitory effect of a compound, as well as its cytotoxicity. More specifically, half-maximal inhibitory concentration (IC_{50}) is the concentration of a substance required to inhibit a biological process by 50% (Aykul and Martinez-Hackert 2016). Specifically for enzyme inhibition, IC_{50} value corresponds to the concentration of an inhibitor that results in 50% reduction of the enzyme's activity. This measure of a substance's efficacy is a function of the concentrations of the enzyme, the substrate and the inhibitor, as well as the experimental conditions. A value that can be more useful for comparisons between different studies is the inhibition constant K_i, for which the substrate plays no role, as it depends only on the enzyme and the inhibitor (Cer et al. 2009). Another relative measure of inhibitory activity is the half-maximal effective concentration (EC_{50}). It is defined as the concentration of a substance that causes an specific effect to reach 50% of its maximum possible value (Neubig et al. 2003). It can be the same as IC_{50} in the sense reduction of the activity of an enzyme to 50% the same as increase of the inhibitory effect of a compound against the enzyme to 50%.

Another indicative property of a substance that is a candidate to be used as a drug or nutraceutical is its cytotoxicity. Half-maximal cytotoxicity concentration (CC_{50}) is a measure of cytotoxicity and is defined as the concentration that reduces cell viability by 50% (Abid et al. 2012). Therefore, a high CC_{50} is desirable, so that a higher concentration of the compound can be used (which usually causes a stronger inhibitory effect) without the host cells being endangered. In studies where the ability of a substance to eliminate infected cells is investigated, a low value for CC_{50} is desirable, however this is not the case in the studies mentioned in this work.

1.6. Inhibition of SARS-CoV-2 M^{pro}

The majority of protease inhibitors act competitively and reversibly, and not by binding to the active site with the same mechanism as the substrate does, but rather by blocking access to it, when interacting with its subsites and catalytic residues. There are also irreversible inhibitors, altering the protease when reacting with it, as well as allosteric inhibitors (Farady and Craik 2010). The way reversible inhibitors of M^{pro} reported bind to its active site is through the reaction of their electrophilic carbon with the sulfur atom of Cys 145, to form a tetrahedral complex, which is often stabilized by interactions with the oxyanion hole residues. Compounds containing aldehyde, thio-, oxy- and amido- methylketone, cyclic ketone, nitrile and 1,2-dicarbonyl moieties have been reported as reversible inhibitors, as well as peptidomimetic compounds with an α-ketoamide active group. Irreversible inhibition also happens through a reaction of nucleophilic addition between the cysteine sulfur and the carbonyl group of the inhibitor. It might however result in S_N2 displacement, leading in the migration of the sulfur atom and deactivation of the enzyme (Hoffman et al. 2020). Multiple covalent and non-covalent inhibitors of M^{pro} have been studied, both *using in silico* and *in vitro* methods. Both types of inhibitors form non-covalent interactions with key residues of the active site, however covalent inhibitors, as mentioned above, react with the protein, and more specifically with catalytic residue Cys 145, through a reaction of nucleophilic addition.

1.6.1. Desired inhibitor characteristics

Apart from the type of protease inhibitors, it is also important to gain insight into the way inhibition occurs and the relationship of the structure of the inhibitor with its binding to the active site of the enzyme, in order to understand how M^{pro} can be blocked and be able to predict inhibitory potential of novel compounds. A common way of approaching the structural analysis of inhibitors is through the system of nomenclature for the peptide substrates of proteases, according to which substrate residues are numbered, beginning from the scissile bond, as P1', P2' etc., to the direction of the C-terminus and as P1, P2 etc. in the direction of the N-terminus (Figure 8). Catalytic residues are located between S1 and S1' subsites, so that they are accessible by the scissile bond.

As mentioned in Dai et al. (2020), an electrophilic moiety, such as an aldehyde, is a good choice for the P1' position, so as to interact with the nucleophilic catalytic cysteine and potentially create a covalent bond, which contributes to stability and specificity of the inhibitor. Apart from that, the goal in order for an inhibitor to be successful is to have as many other, non-covalent interactions that can stabilize the complex as possible. For example, a common occurrence is a (S)-γ- lactam ring in the P1 position, whose oxygen and NH- group make easier the formation of interactions with neighboring residues.

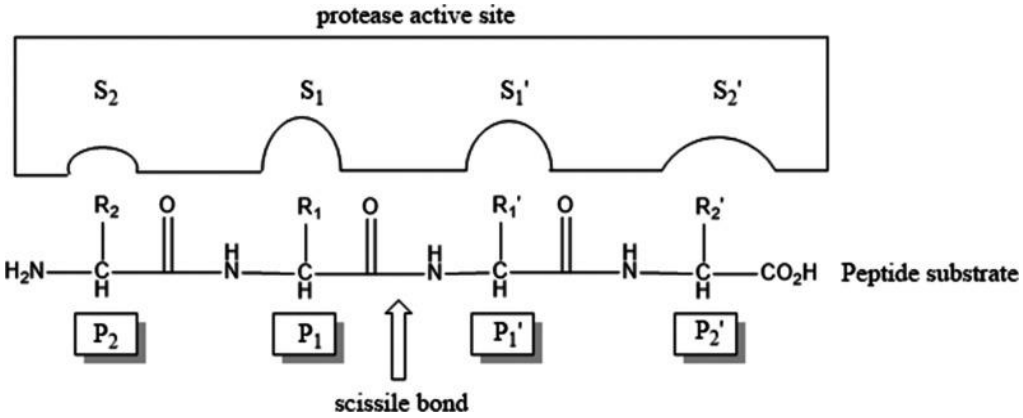


Figure 8: Proteolytic enzyme substrate nomenclature as shown in Hoffman et al. (2020)

Rut et al. (2020) also performed an enzymatic assay to screen a hybrid combinatorial substrate library and compare amino acids with different side chains in different substrate positions and concluded that an isobutyl group is almost two times more reactive than an o- and m-phenylene group in the P2 position. Tert-butyl, primary amine, phenol, guanidine and phenyl groups are reactive in the P3 position, in a series of decreasing reactivity. Lastly, the same happens with methyl, isopropyl, thiazole, o-phenylene and indole groups in the P4 position.

Douangamath et al. (2020) conducted a fragment screening study from which various conclusions about the relationship between the structure of the ligand and the position where it bounds and its orientation can be derived. More specifically, it occurs that a ligand can block the S1 subsite when having a pyridine or other ring containing a nitrogen atom, that interacts with His 163, or an amide or urea group, which forms interactions with Glu 166 through the carbonyl. Moreover, hydrophobic residues, such as Leu, have been shown to more easily bind to S2 pocket. It is observed that the aromatic ring of binding compounds forms hydrophobic interactions with Met 49 or pi-pi interactions with His 41, so it is being stabilized in this cleft. Ligands containing and N-chloroacetyl moiety, especially as part of an N-chloroacetyl piperidinyl-4-carboxamide group (Figure 9a) form multiple hydrogen bonds with oxyanion hole residues Gly 143, Ser 144 and Cys 145 and they are orientated towards the S2 subsite. Screening of compounds containing a N-chloroacetyl-N'-sulfonamido-piperazine group (Figure 9b) shows that they bend, allowing substitutions of the phenyl group to block the S2 subsite. In addition, halophenyl moieties present appear possible to interact with residue Asn 142. The presence of the N-chloroacetyl-N´-carboxamido and N-chloroacetyl-N´-heterobenzyl-piperazine motif (Figure 9c) tend to bind towards the S2 subsite, providing access to S3 pocket as well.

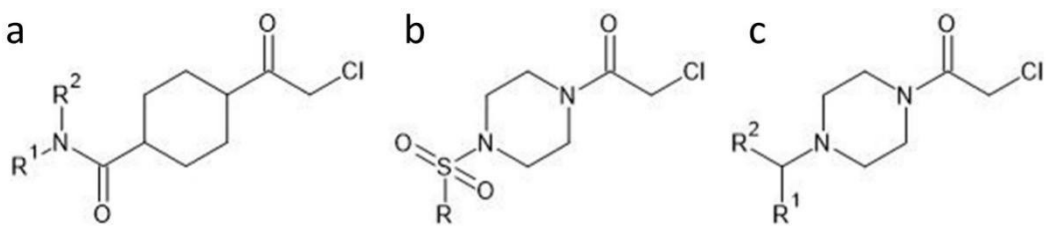


Figure 9: N-chloroacetyl piperidinyl-4-carboxamide motif(a), N-chloroacetyl-N'-sulfonamido-piperazine motif (b) and backbone of N-chloroacetyl-N´-carboxamido and N-chloroacetyl-N´-heterobenzyl-piperazine motifs (c) as described in Douangamath et al. (2020)

Stoddard et al. (2020) have conducted similar research to study the effects of structure on binding affinity. Results showed that the presence or absence of halogens in a ligand does not considerably affect binding affinity. In addition, due to the fact that the active site has several hydrophobic regions, the effect of hydrophobic groups in a ligand was investigated. Results show that aliphatic substitutions increase binding affinity, and the longer they are the greater the increase. Also, there was an edge of meta substitutions over ortho and para, as it gives the ligand an orientation towards S1 and S2 subsites. S2 subsite being a pocket that shows a preference for hydrophobic moieties, the binding affinity is increased if the aliphatic group of the ligand manages to enter deeper into the S2 pocket.

That is also the case with aliphatic ring additions, which also contribute in a great increase in the binding affinity and are preferable to aromatic rings. Overall, the presence of nitrogen atoms increases binding affinity as it creates the potential of the formation of multiple hydrogen bonds or pi-pi stacking interactions. A heterocyclic moiety with a nitrogen atom is more favorable to access the S2 subsite and form a hydrogen bond with Tyr 54 when the nitrogen atom is at the para position.

From the aforementioned study, the existence of five locations accessible to hydrogen bonding to the active site of M^{pro} was described. One is found in the S4 subsite and consists of residues Arg 188 and Thr 190, with the backbones of which many compounds are found to create hydrogen bonds. The binding of a compound to the S2 site can occur through hydrogen bonding with residues Tyr 54, His 41 or Asp 187, which form the second binding location. Also, accessible to hydrogen bonding are residues Glu 166 and Gln 189, creating a third hydrogen bonding hotspot. In the S1' subsite, hydrogen bonding is facilitated by residues Thr 24 and Thr 45. An amine or alcohol group is favorable for such an interaction. S1 subsite is made accessible for ligands to create hydrogen bonds, especially for those with a protonated nitrogen atom, through the side chains of residues, Leu 141, Gly 143, Ser 144, His 163.

In the following paragraphs, specific compounds that have been recognized as SARS-CoV-2 M^{pro} inhibitors and their interactions with M^{pro} will be described, confirming the above-mentioned data resulting from the structural analysis.

1.6.2. Covalent inhibitors

Research has led to the identification of multiple compounds as M^{pro} inhibitors, which include both already known drugs, as well as compounds designed for the specific target. The co-crystallization structure of the inhibitors in complex with the enzyme proves that the majority of identified inhibitors bind covalently to the active site.

N3 is a peptidomimetic compound that successfully inhibits the protease, as it binds to its active site very similarly to the actual substrate. It is a Michael acceptor, and acts as a time-dependent, irreversible inhibitor. Its 50 % cytotoxicity concentration (CC_{50}) is reported to be greater than 133 μM, whereas the half-maximal effective concentration (EC_{50}) is 16.77 μM. In the original publication that provided the crystal structure, the interactions between the enzyme and N3 are described in detail. More specifically, the inhibitor forms a 1.8 Å covalent bond with the sulfur atom of residue Cys 145 of the protein. Moreover, N3 forms one hydrogen bond with each one of residues Gly 143, His 163, His 164, , Gln 189 and Thr 190 and two hydrogen bonds with Glu 166 (Jin et al. 2020a). **Ebselen** is another auspicious drug molecule worth mentioning, as it inhibits the protease with an IC_{50} of 0.67 μM, and an EC_{50} of 4.67 μM, also exhibiting very low cytotoxicity. In the case of ebselen, covalent inhibition is

reinforced by its non-covalent interaction with the active site residues, which are however not described in detail (Jin et al. 2020a; Sies and Parnham 2020).

Two other covalent inhibitors are **11a** and **11b**. As covalent inhibitors, both compounds are covalently bound to the S-atom of Cys 145, with a 1.8Å bond. The enzyme-inhibitor complex is further stabilized with a hydrogen bond between the oxygen of the aldehyde group of 11a and 11b and Cys 145. Additionally, they both form one hydrogen bond with Phe 140, His 163 and His 164 and three with Glu 166. Inhibitor 11b contains an F-atom that forms an additional hydrogen bond with Gln189. The cyclohexyl group of 11a inserts the hydrophobic pocket that makes up S2 subsite, showing hydrophobic interactions with residues His 41, Met 49, Tyr 54, Asp 187 and Arg 188. The indole moiety of the inhibitor also interacts hydrophobically with Pro 168 and Gln 189. As for 11b, the 3-fluorophenyl group interacts with the active site similarly to the cyclohexyl group of 11a, forming hydrophobic interactions with residues His 41, Met 49, Met 165, Val 186, Asp 187, Arg 188. An important role in the stabilization of the inhibitors is played by some water molecules, that form hydrogen bonds with both 11a/11b and the residues of the binding cleft. All the above-described interactions can be summed up in Figure 10. At a concentration of 1 μM, 11a and 11b exhibited 100% and 96% inhibitory activity, respectively. Moreover, the IC_{50} values are promising, equaling 0.053±0.005 μM for 11a and 0.040±0.002 μM for 11b. Between the two inhibitors, results showed that 11a has a greater potential to act as an antiviral compound (Dai et al. 2020).

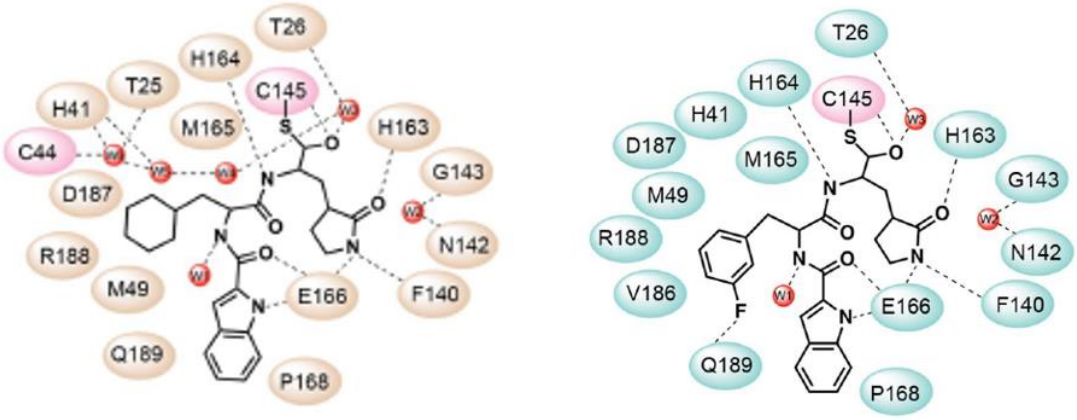


Figure 10: Interactions between SARS-CoV-2 M^{pro} and inhibitors 11a (right) and 11b (left). W1-W6 represent water molecules that play an important role in the binding of the inhibitors since they act as an intermediate for their interactions with the active site residues and stabilize their binding (Dai et al. 2020)

As mentioned above, due to the high conservation of the active site of SARS-CoV-2 M^{pro}, compared to the main proteases of other coronaviruses, a lot of the already tested inhibitors for SARS-CoV or other coronaviruses can be effective against SARS-CoV-2 too. For example, Hattori et al. (2021) tested such compounds and report the significant inhibitory potential of compound **5h**, with a CC_{50} value greater than 100 μM and EC_{50}= 4.2 ± 0.7 μM. Inhibitor 5h forms a reversible covalent bond with Cys 145, via the same nucleophilic addition mechanism that the other covalent inhibitors exhibit. More specifically, the sulfur atom of Cys 145 attacks the carbonyl carbon next to the benzothiazole of 5h. 5h forms two hydrogen bonds with Glu 166, and one with each one of Gly 143, Cys 145, His 164, Gln 189. In this case, too, there are several water molecules that form hydrogen bonds with the inhibitor and the active site

residues and act as intermediates and stabilize the interactions between them. In addition, van der Waals interactions between the hydrophobic residues Leu 27, Met 49, Phe 140, Met 165 and Ala 191 and the inhibitor improve its binding affinity.

Boceprevir is another compound that can inhibit M^{pro} by covalently binding to its active site, exhibiting EC_{50} value of 15.57 μM. The keto carbon of boceprevir is the atom that takes part in the covalent bond. There are also hydrogen bonds formed with residues His 41, Gly 143, Cys 145, His 164 and Glu 166. As for Glu 166, boceprevir forms three hydrogen bonds with that particular residue. Hydrophobic interactions between the inhibitor and the enzyme are mostly found in subsites S2 and S4, and more specifically with residues Met 149, Met 165, Asp 187, Gln 189, Thr 190 and Gln 192 (Fu et al. 2020).

GC376 is a broad-spectrum antiviral medication with a half maximal effective concentration EC_{50}= 0.70 μM against SARS-CoV-2, which is very close to the approved anti-SARS-CoV-2 drug remdesivir (EC_{50}=0.58 μM). In order for GC376 to form a covalent bond, its bisulfite group is removed. The compound forms one hydrogen bond with residues Phe 140, Gly 143, Cys 145, His 163, His 164 and two with Glu 166 and interacts with the hydrophobic pocket residues Arg 40, His 41, Met 49, Tyr 54 and Asp 187 (Fu et al. 2020). Also effective against SARS-CoV-2 is the parent compound of GC376, **GC373**. It shows no toxicity in cell culture and inhibits M^{pro} with an IC_{50} value of 0.40 ±0.05 μM. The inhibition occurs through a reversible reaction of the thiol of Cys 145 with the carbonyl of GC373 resulting in a hemithioacetal. The conformation of the inhibitor in the active site is stabilized with hydrogen bonds with the oxyanion hole residues Gly 143, Ser 144, Cys 145. There is also one hydrogen bond formed with His 163 and two with Glu 166. There are also hydrophobic interactions present, both with S2 pocket residues His 41, Met 49 and Met 165 and His 172 (Vuong et al. 2020).

Narlaprevir is also a potent antiviral compound, with an IC_{50} value of 16.11 μM and EC_{50} value of 7.23 μM (Bai et al. 2021). According to literature, except for the covalent bond, it creates four hydrogen bonds with residues His 41, Asn 142, Gly 143 and His 164 and three hydrogen bonds with Glu 166. It also interacts with residues Leu 141, Ser 144, Met 165, Pro 168, Gln 192 (Bai et al. 2021). Binding to the active site of SARS-CoV-2 M^{pro} in a very similar way to narlaprevir and boceprevir, peptidomimetic compound **telaprevir** acts as an effective inhibitor, with an IC_{50} of 18 μM (Kneller et al. 2020b). More specifically, apart from the covalent bond with Cys 145, telaprevir forms direct hydrogen bonds with His 41, Gly 143, Ser 144, His 164, His 166 (with which there are two interactions) and Gln 189. There is also shown to be a water-mediated hydrogen bond with Gln 192, as well as pi-pi interactions with residues Thr 190 and Ala 191 (Qiao et al. 2021).

Another potent compound is **MI-23**, which has been designed based on telaprevir and exhibits IC_{50} = 7.6 nM. It forms the characteristic 1.8Å covalent bond with Cys 145 and additionally hydrogen bonds with Phe 140, Gly 143, Cys 145, His 163, His 164 and Glu 166. The bicycloproline moiety is located in the hydrophobic S2 subsite, having hydrophobic interactions with residues His 41, Met 49, Met 165, Leu 167, Pro 168, Asp 187, Arg 188 and Gln 189 (Qiao et al. 2021). Alpha-ketoamide **13b** is also a compound that has been found to covalently inhibit SARS-CoV-2 with IC_{50}=0.67±0.18 μM and EC_{50} 4 to 5 μM. Its conformation in the binding site is further stabilized with six hydrogen bonds with compounds His 41, Phe 140, Gly 143, Ser 144, Cys 145, His 163 three hydrogen bonds with Glu 166 (Zhang et al. 2020b).

Another peptidomimetic compound that binds in a similar manner to the binding site of M^{pro} is **calpeptin.** When in contact with the protease, Cys 145 attacks its aldehyde group to form a

thiohemiacetal intermediate. The compound forms two hydrogen bonds, with residues His 164 and Glu 166. In addition, Van der Waals forces are developed between calpeptin and residues Phe 140, Leu 141 and Asn 142. Due to this interactions, the inhibitor successfully blocks part of the active site, showing an EC_{50} value of 72 nM and CC_{50} value greater than 100 μM (Günther et al. 2021).

Carmofur is an antineoplastic drug that has also proved to inhibit M^{pro}. Inhibitory effect and cytotoxicity have been tested on Vero E6 cells and resulted in an EC_{50} value of 24.30 μM and a CC_{50} value of 133.4 μM. Unlike previous inhibitors that occupy multiple subsites of the protease, carmofur only binds to S2 subsite. The fact that this small compound is able to inhibit SARS-CoV-2 provides a good starting point from which more elaborate structures could be designed to inhibit the enzyme even more effectively. The mechanism through which the covalent bond is created is slightly different that the previously described cases, as the sulfur atom of Cys 145 binds to the carbonyl group of the fatty acid tail of carmofur creating a 1.8Å covalent bond, but this reaction results in the release of the 5-Fluorouracil moiety. The tail of carmofur inserts the S2 subsite and forms a hydrogen bond with each of Gly 143 and Cys 145. The conformation of the inhibitor in the active site is also affected by hydrophobic interactions with residues His 41, Met 49, Met 165 and Asp 187 (Jin et al. 2020b).

Myricetin has also been identified by Kuzikov et al. (2021) as a flavonoid that covalently binds to the active site of M^{pro}. As in the case of carmofur, although it does inhibit the action of Mpro, exhibiting IC_{50}= 0.22 μM, it does not fully occupy the active site and therefore could be used as a parent compound for an optimized inhibitor.

Compound **MG-132** is another reversible M^{pro} inhibitor (IC_{50}= 0.36 μM , CC_{50}= 2.9 μM). Its relatively large size allow effective blocking of the subsites of the protein, the precise interactions it forms with the proteins are not described (Kuzikov et al. 2021). Lastly, crystal structures that have been deposited to the Protein Data Bank provide evidence of covalent inhibition of M^{pro} by various fragments. Two of them are **PG-COV-34**, or **x2754**, a small amide (Douangamath et al. 2020), and **x2705**, a more complex compound, for which the supporting paper has not been published. In both cases, there is no documented description of their interactions with the residues of the active site, but the crystal structure itself is an important indication.

The chemical structure of all the covalent inhibitors mentioned is presented in Figure 11.

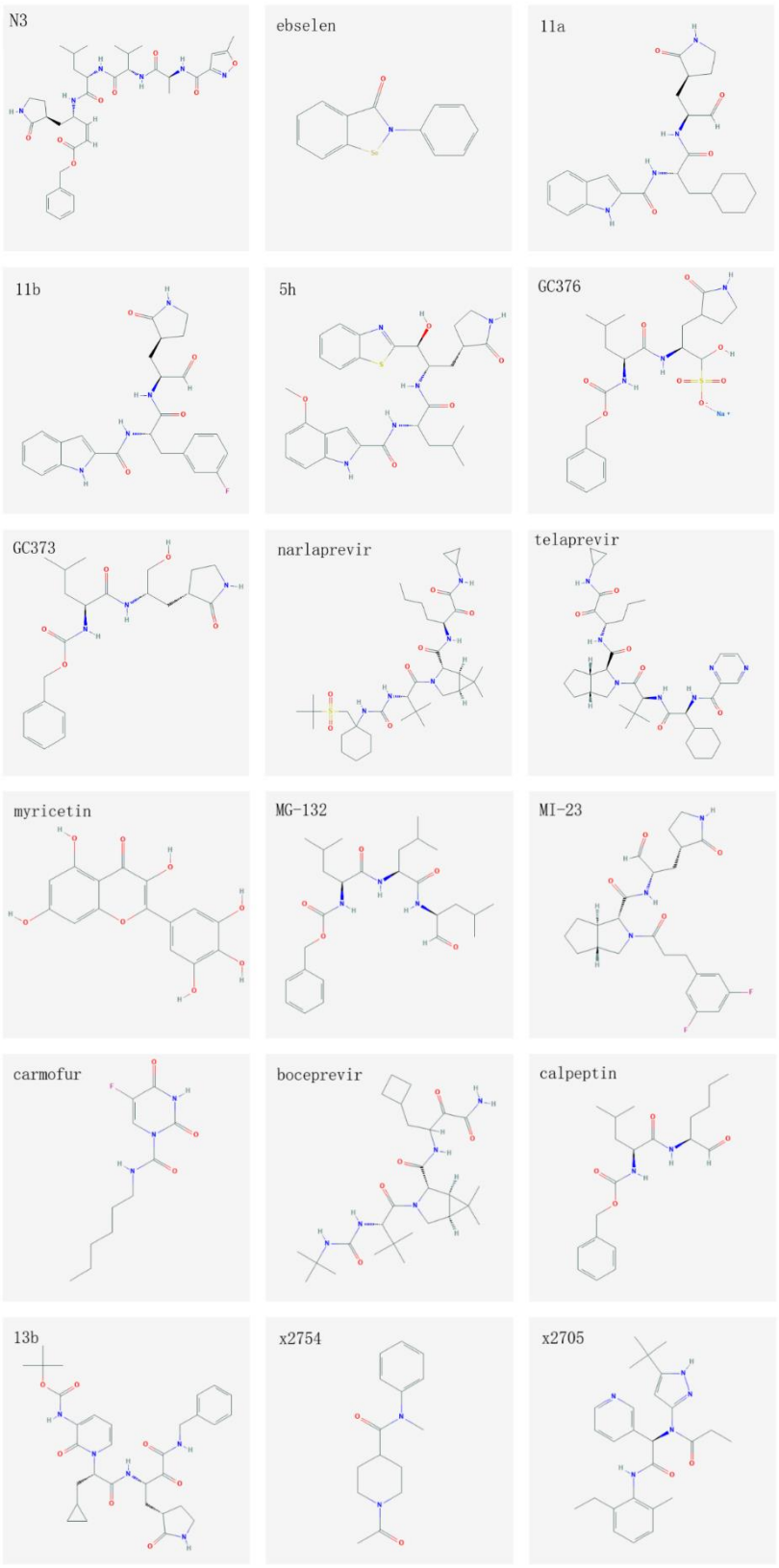


Figure 11: Chemical structures of SARS-CoV-2 M^{pro} covalent inhibitors

1.6.3. Non-covalent inhibitors

Non-covalent inhibitors reported in literature are significantly less than the covalent ones. However, some of them are very promising as antiviral compounds. Zhang et al. (2021) have conducted an interesting study using free-energy perturbation calculations and Vero E6 cell assays to investigate the inhibitory potential and antiviral properties of different compounds, which were designed as an optimized version of perampanel, an anti-epileptic drug. Two of these compounds were the most promising: **compound 5** (2-(3-(3-Chloro-5-propoxyphenyl)-2-oxo-2H-[1,3'-bipyridin]-5-yl)benzonitrile) and **compound 26** (2-(3-(3-Chloro-5-(cyclopropylmethoxy)phenyl)-2-oxo-2H-[1,3'-bipyridin]-5-yl)benzonitrile). The difference in the structure of the two compounds is that the propyl group of compound 5 is replaced by a cyclopropyl group in compound 26. The calculated IC_{50} values for the two compounds were 0.140 ± 0.020 μM and 0.170 ± 0.022 μM respectively, indicating that the replacement of the propyl by the cyclopropyl group leads to an increase of the IC_{50}. The anti-SARS-CoV-2 activity of the two compounds is demonstrated by EC_{50} values of 1.5 and 0.98 μM respectively, as measured with a lower-throughput viral plaque assay. The cytotoxicity of compound 5 was significantly higher than compound 26, as indicated by the CC_{50} values measured in Vero E6 and normal human bronchial epithelial (NHBE) cells, which were as low 22 and 20 μM respectively for compound 5 and higher than 100 μM in both cases for compound 26. Compound 5 was shown to form three hydrogen bonds with active site residues Gly 143, His 163 and Met 165, whereas the detailed interactions of compound 26 are not described.

Useful insight of how the active site of M^{pro} can be inhibited is provided by the fragment screening performed by Douangamath *et al*. (2020). Compound **x0104** (Z1220452176) occupies the S2 subsite of the protease, whereas compound **x0161** (Z18197050) the S3 subsite. An interesting observation is related to the binding of compound **x0397** (Z369936976), which interacts with the two catalytic residues changing their conformation. This alteration changes the shape of S1' subsite and consecutively the one of S1 too. Therefore, this fragment blocks both sites, with its N-methyl group also providing the potential to block S2 and S3 subsites too. Although there is a crystal structure that proves the binding of these inhibitors to the active site of M^{pro}, there have not been in vitro experiments conducted to measure antiviral activity or cytotoxicity.

MUT056399 is another compound that binds non-covalently to the active site, inhibiting it with an EC_{50} of 38.24 μM. It also shows low cytotoxicity, as described by a CC_{50} value greater than 100 μM. Its carboxamide group binds to the S1 subsite, forming hydrogen bonds with residues His 163 and Phe 140. The other end of the molecule, consisting of an ethyl-phenyl moiety, occupies S2 pocket (Günther et al. 2021). A compound reported to also inhibit SARS-CoV Mpro, **ML 188,** binds to the active site of SARS-CoV-2 M^{pro} as well, and inhibits its activity with an IC_{50}= 2.5 ± 0.3 μM. However, apart from pointing out the importance of the interaction with His 41 for the inhibition, the interactions of the ligand with the active site are not described in detail (Lockbaum et al. 2021)**.** Also, among other inhibitors, available crystal structures for two compounds, **Mcule-5948770040** and **X77**, prove their ability to bind to the active site of the protease. The works framing the crystal structures though have not been published, therefore no additional information is available about them.

The chemical structure of the inhibitors described above is presented in Figure 12.

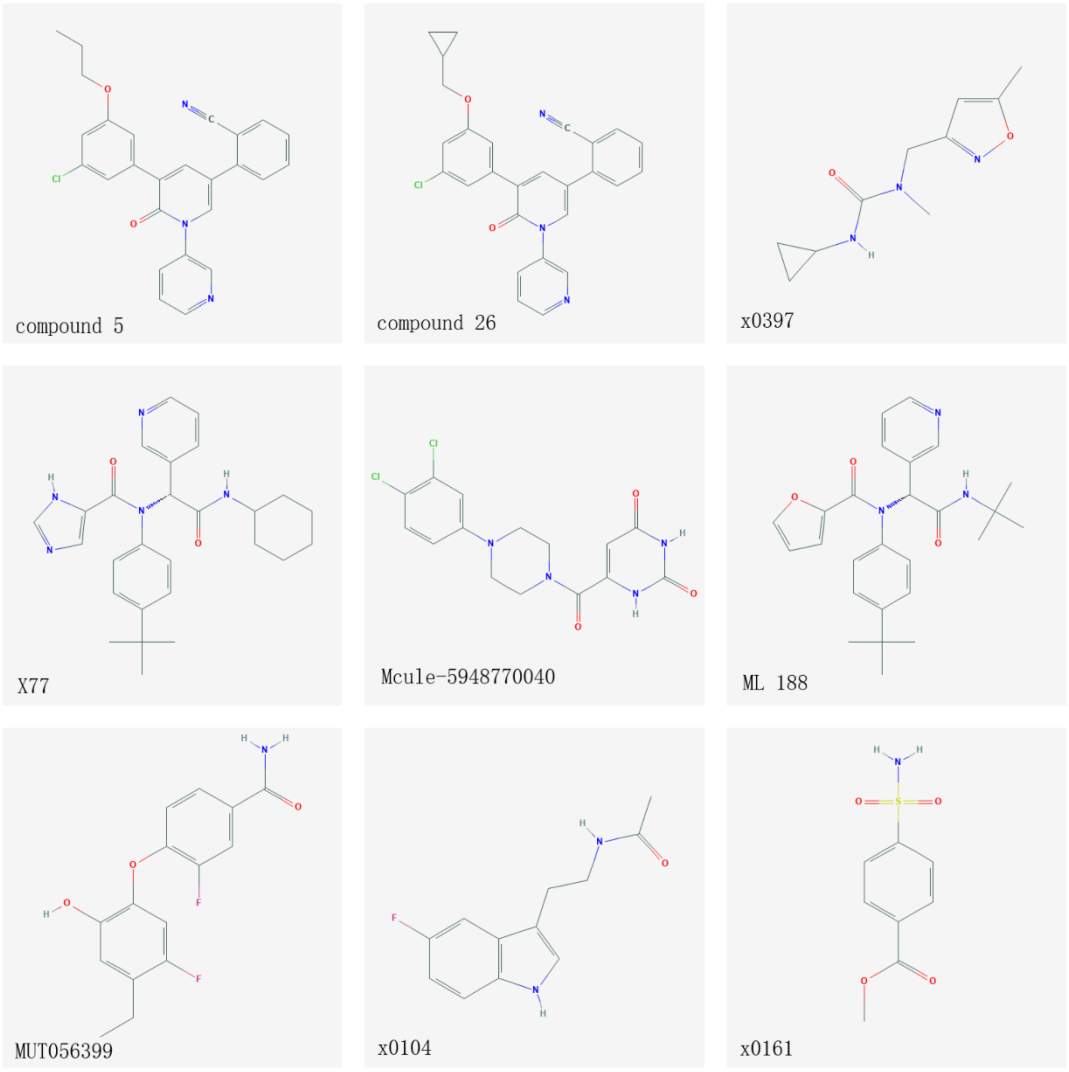


Figure 12: Chemical structures of SARS-CoV-2 M^{pro} non-covalent inhibitors

1.6.4. Allosteric inhibitors

Günther et al. (2021) discovered two regions outside the binding site that act as allosteric binding sites (Figure 14), as well as inhibitors binding to these allosteric sites exhibiting remarkable antiviral activity. Residues Ile 213, Leu 253, Gln 256, Val 297 and Cys 300 form a hydrophobic pocket that serves as the first allosteric binding site. This pocket accommodates the aromatic groups of inhibitors pelitinib, ifenprodil, RS-102895, PD-168568, and tofogliflozin. Among these compounds, pelitinib shows good efficacy potential (EC_{50} = 1.25 µM) but a not very high cytotoxicity of infected cells (CC_{50}=13.96 µM). Although pelitinib does not occupy the canonical active site of M^{pro}, its ethyl ether group interacts with residues Tyr 118 and Asn 142, affecting the S1 pocket. The second allosteric binding pocket is located in the cavity between domains I and II, and domain III. Inhibition through binding to this site is connected to interactions of the inhibitor with residue Arg 298, which plays a critical role in dimerization. Change in the conformation of Arg 298 causes the alteration of the relative position of domains I&II and III and therefore destabilizes the oxyanion hole and the S1

subsite. Inhibitor AT7519 binds to this site, forming Van der Waals contacts with residues Ile 249, Phe 294 through its pyrazole ring. The carbonyl group interacts with Gln 110 with a hydrogen bond and the piperidine group forms a hydrogen bond with Asp 153. The reorientation of Asp 153 is concomitant with a slight disposition of Tyr 154 and its hydrogen-bonding to the inhibitor, as well as the interaction with Arg 298, which is achieved through a salt bridge. The structure of the two inhibitors is presented in Figure 13.

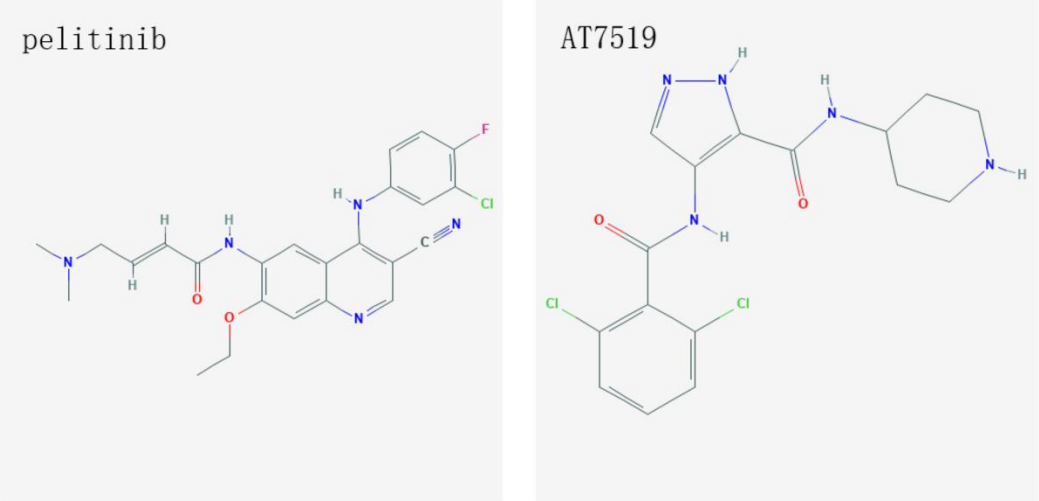


Figure 13: Chemical structure of allosteric inhibitors of SARS-CoV-2 M^{pro}

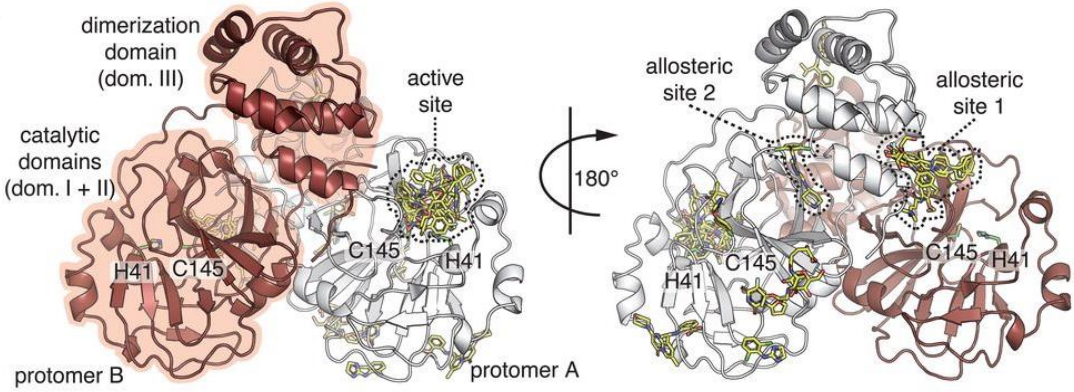


Figure 14: Active site and allosteric sites of M^{pro} as depicted in Günther et al. (2021)

A summary of the antiviral properties of the inhibitors described above is presented in Table 1.

Table 1: Antiviral activity indicators for the inhibitors of SARS-CoV-2 M^{pro}

Inhibitor	PDB ID	IC_{50} (μM)	Assay	EC_{50} (μM)	Assay	CC_{50} (μM)	Assay	Reference
N3	6LU7	-	-	16.77	Plaque reduction assay	133	MTS cell proliferation assays (Promega) on Vero E6 cells	Jin et al 2020a
Covalent								
Ebselen		0.67	FRET[1]-based cleavage assay	4.67 ± 0.80	Plaque reduction assay			Jin et al 2020a
11a	6LZE	0.053±0.005	FRET–based cleavage assay	0.53 ± 0.01	Plaque reduction assay			Dai et al. 2020
11b	6M0K	0.040±0.002	FRET–based cleavage assay	0.72 ± 0.09	Plaque reduction assay			Dai et al. 2020
5h	7JKV	-	-	4.2 ± 0.7	RNA-qPCR quantitative assay on VeroE6 cells	>100	RNA-qPCR quantitative assay on VeroE6 cells	Hattori et. al 2021
GC376	7D1M	0.19 ±0.04	FRET–based cleavage assay	0.92	Plaque reduction assay	>200	CellTiter-Glo assay in Vero E6 cells	Vuong et al. 2020
GC373	6WTK	0.40 ±0.05	FRET–based cleavage assay	1.5	Plaque reduction assay	>200	CellTiter-Glo assay in Vero E6 cells	Vuong et al. 2020
Narlaprevir	7JYC	16.11	FRET–based enzyme activity inhibition assay	7.23	Plaque reduction assay	>200	Cytotoxicity assay on Vero E6 cells.	Bai et al. 2021
Telaprevir	7K6D	18						Kneller et al. 2020b
Myricetin	7B3E	0.22	FRET-based cleavage assay					Kuzikov et al. 2021
Mg-132	7BE7	0.36	CPE assay in Vero E6 cells			2.9	Vero E6 imaging assay	Kuzikov et al. 2021
MI-23	7D3I	7.6 nm	FRET-based cleavage assay			>500	Enzyme inhibition assay (Cell Counting Kit-8 (CCK8) assay)	Qiao et al. 2021
Carmofur	7BUY	1.82 ± 0.06	FRET-based cleavage assay	24.3	qRT-PCR assay in Vero E6 cells	133.4	Cytotoxicity assays in Vero E6 cells	Jin et al 2020b
Boceprevir	7C6S			15.57	Plaque reduction assay			Fu et al. 2020
Calpeptin	7AKU			72 nm	Antiviral activity assay in vero E6 cells	>100	Cytotoxicity assays in Vero E6 cells (Cell Counting Kit-8)	Günther et al. 2021
13b	6Y2G	0.67±0.18	FRET-based cleavage assay	4 to 5	Antiviral activity assay in human Calu-3 lung cells			Zhang et al. 2020b
Non-covalent								
Compound 5	7L11	0.14±0.02	FRET-based cleavage assay	1.5	Plaque reduction assay	22± 7.2	Methylthiazolyl-diphenyl-tetrazolium bromide (MTT) dye assay in Vero E6 cells	Zhang et al. 2021
Compound 26	7L14	0.170± 0.022	FRET-based cleavage assay	0.98	Plaque reduction assay	>100	MTT dye assay in Vero E6 cells	Zhang et al. 2021
ML 188	7L0D	2.5 ± 0.3	FRET-based cleavage assay					Lockbaum et al. 2021
MUT056399	7AP6			38.24	Antiviral activity assay in vero E6 cells	>100	Cytotoxicity assays in Vero E6 cells (Cell Counting Kit-8)	Günther et al. 2021
Allosteric								
Pelitinib	7AXM			1.25	Antiviral activity assay in vero E6 cells	13.96	Cytotoxicity assays in Vero E6 cells (Cell Counting Kit-8)	Günther et al. 2021
AT7519	7AGA	Not determined						

[1]: Fluorescence resonance energy transfer

1.7. Promising phytochemicals with inhibitory effect against SARS-CoV-2 M^{pro}

Apart from drug discovery and repurposing, research has been orientated towards phytochemicals in search for ways to restrain the effect that COVID-19 has on public health that can reinforce the action of antiviral drugs and vaccines, which are much more time-consuming to be developed. Natural compounds found in extracts of plants, may be employed, as a tool for boosting immunity and aid protection against infection. Moreover, knowledge on the beneficial action of bioactive phytochemicals, may enhance preparedness for future viral outbreaks.

To begin with, the above-mentioned M^{pro} inhibitor myricetin (Kuzikov et al. 2021) is a natural compound found in several plants. In addition, aqueous extract of the plant *Scutellaria barbata* D. Don, including flavonoids apigenin, naringenin, scutellarin, baicalein, luteolin and wogonin, has been reported to inhibit both M^{pro} and the transmembrane protease TMPRSS2 of SARS-CoV-2 *in vitro,* as resulted from a FRET assay (Huang et al. 2021). Extract from the Thai medicinal plant *Boesenbergia rotunda,* which is widely used as a culinary herb in Asia, and more specifically its compound panduratin A, hinder SARS-CoV-2 infection as shown in studies in Vero E6 cells and human airway epithelial cells (Calu-3)(Kanjanasirirat et al. 2020).

The number of studies that have examined the antiviral activity of natural compounds *in vitro* is limited, however there are several docking studies screening phytochemicals. For example, compounds found in Indian ginseng, such as flavonoids quercetin-3-rutinoside-7- glucoside and rutin and caffeoylquinic acid, have shown inhibitory potential of M^{pro} (Kushwaha et al. 2021). Additionally to rutin, which is found in several plants, such as apples or tea, flavonoid hesperidin has shown good binding affinity to the active site of M^{pro} (Kiani et al. 2020). *Argania spinosa* L. is mentioned as a plant with potential anti-SARS-CoV-2 activity attributed to its compounds procyanidin B1, kaempferol, betulinic acid, quercetin and luteolin (Mrid et al. 2021). Polyphenolic compounds in sumac (*Rhus* spp.) have also exhibited antiviral potential *in silico* (Sherif et al. 2021). Tahir ul Qamar et al. (2020) report methyl rosmarinate and flavonoid derivatives including myricitrin, myricetin 3-O-beta-D-glucopyranoside, licoleafol and amaranthin as phytochemicals with high binding affinity to the active site of M^{pro}. Moreover, promising bioactive compounds have been detected in ayurvedic medicinal plants: withanoside and somniferine in *Withania somnifera*, tinocordiside in *Tinospora cordifolia* and vicenin, isorientin 4'-O-glucoside 2"-O-p-hydroxybenzoagte and ursolic acid in *Ocimum sanctum*(Shree et al. 2020). Provided all these indications, extracts from plants containing bioactive compounds can be an effective way to modulate the immune system and shield the human body from viral infection.

1.8. The extract from halophyte plant Salicornia as potential antiviral agent

Plants of the genus *Salicornia* L. (family *Amaranthaceae,* subfamily *Salicornioideae*) are annual succulent halophytes often found in wet, saline areas, such as coastlines, salt marshes or salt lakes. *Salicornia* is very close to the genus of *Sarcocornia*, and both of them are known as glassworts, the two however are distinct. *Salicornioideae* grow in every continent apart from Antarctica, but specifically Salicornia also cannot be found in Australia and South America. It has approximately 13 species, the most common of which is *Salicornia herbacea L*.(Figure 15), also known as glasswort (greek: αρμυρήθρα) and Tungtungmadi or Hamcho in Korea (Rhee et al. 2009; Shepherd et al. 2005). Distribution of two species of the plant across Europe and Asia is presented in Figure 16. The aerial parts of the plant are being consumed fresh, as salad, or fermented, as preserves or beverages, and have also been used in traditional medicine as remedies against hypertension, diabetes, obesity and cancer (Kang et al. 2015; Essaidi et al.

2013). It is important to note that the fact that these plants grow and can be cultivated in saline, marginal environments provides a promising future outlet from the dependence of agriculture on the supply of fresh water (Ventura and Sagi 2013). Moreover, the beneficial compounds found in the plant can be extracted from the waste generated from its cultivation, which include stems, branches and deseeded inflorescences, making possible the valorization of organic matter that could otherwise be remained unused (Chaturvedi et al. 2012).

Figure 15: Salicornia herbacea L.

S. herbacea has numerous nutritional and health benefits. Apart from containing proteins, fatty acids and carbohydrates and therefore providing energy, it is a source of minerals, such as Mg, Ca, Fe, K, dietary fibers and bioactive compounds. Bioactive compounds include phenolic acids and flavonoids, sterols, saponins, alkaloids, tannins and micronutrients, such as selenium. The seeds of the plant also contain proteins and unsaturated fatty acids such as linoleic and oleic acids (Loconsole et al. 2019). Extracts of *Salicornia* species have been shown to contain hydroxycinnamic acids, hydroxybenzoic acids, quinic acid derivatives, flavonoids, sterols, chromones, lignans and saponins among other compounds. A summary of the contents of *Salicornia* extracts is presented in Table 2, and their respective chemical structes in Figures 17-25.

The halophyte extract exhibits high antioxidant activity, which is attributed to its high content of phenolic acids and flavonoids. These compounds, in addition to some fatty acids, polysaccharides and the osmotic compound betaine are also contributing to the plants antimicrobial properties (Essaidi et al. 2013). In addition, polysaccharides from *S. herbacea* have exhibited antiproliferative effects against human colon cancer H-29 cells, as well as enhancement of immune response (Patel 2016). The plant also helps reduce levels of lipids in the blood. Studies in mice have shown that it can act against weight gain, hepatic lipid accumulation and diet-induced hyperglycemia and hyperlipidemia (Park et al. 2006; Pichiah and Cha 2015). Its multiple benefits also include immune modulation, protection of the liver, and activity against diabetes, hypertension and inflammation (Rahman et al. 2018). *S. herbacea* has also been reported to have anti-osteoporotic properties, as it can boost osteoblastogenesis (Karadeniz et al. 2014). The plant extract has been proven to inhibit

tyrosinase and melanin synthesis, thus having potential skin whitening effects (Sung et al. 2009).

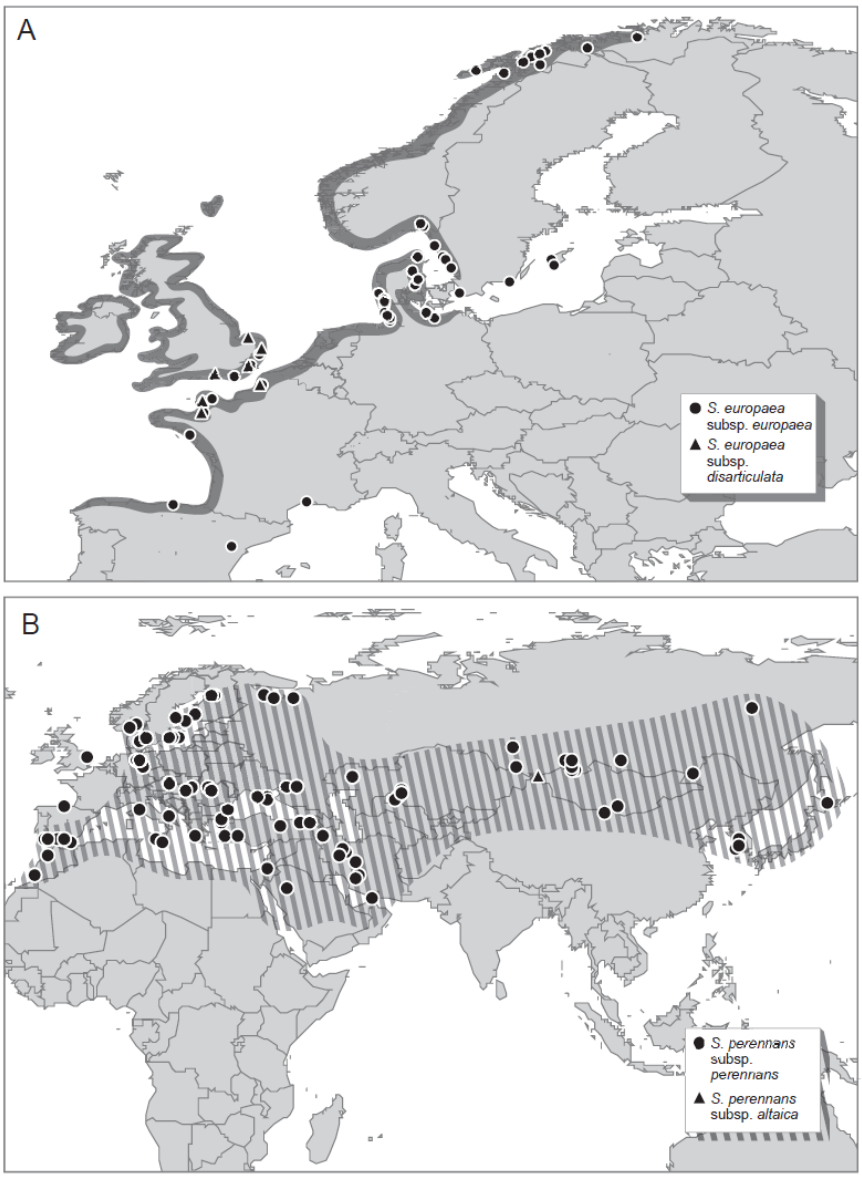


Figure 16: Distribution of Salicornia species (A: S. europaea, B: S. perennans) in Europe and Asia (Kadereit et al. 2012)

Various studies have reported information on the specific phenolic compounds present in *Salicornia* species. A lot of these compounds have documented antiviral potential. Phenolic acids, particularly quinic, chlorogenic, and gallic acid, have been reported to have antiviral activity against RNA virus parainfluenza type 3. In addition to these compounds, also ferulic acid (as a major compound in *Ficus carica* extract), caffeic acid and flavonoids quercetin and apigenin exhibit antiviral activity against herpes simplex DNA virus (HSV-1) (Özçelik et al. 2011; Aref et al. 2011). Quercetin and another flavonoid, catechin, show inhibitory potential against rabies virus (Chávez et al. 2006), whereas quercetin has also been mentioned as potential antiviral against Equine Herpes Virus *1* (EHV-1), which is a DNA virus that causes respiratory disease in horses (Gravina et al. 2011), adenoviruses ADV-3, ADV-8 and ADV-11 (Chiang et al. 2003) and respiratory syncytial virus (RSV), another RNA virus that attacks the respiratory system (Formica and Regelson 1995). In addition, *Salicornia* extract has reported antiviral

properties against Encephalomyocarditis virus, Semliki Forest virus and Hepatitis B virus (Premnathan et al. 1992).

Attention has already been given to natural compounds as potential allies against the current pandemic via *in silico* simulations. The previously mentioned myricetin, a flavonoid that has exhibited anti-SARS-CoV-2 activity (Kuzikov et al. 2021), has been detected in *Salicornia*. Rutin is another compound found in natural products that has shown inhibitory potential of SARS-CoV-2 M^{pro}, as deduced from *in silico* molecular docking studies, where it exhibited a binding energy of -15.63 kcal/mol and multiple interactions with active site residues (Bharadwaj et al. 2021). Ferulic acid interacts with the proteins of SARS-CoV-2 forming hydrogen bonds, as shown in (Salman et al. 2020a), therefore proving that additional research could be fruitful. All this data indicates that Salicornia extracts may have an exceptional ability to block SARS-CoV-2 and particular potentially inhibit M^{pro}.

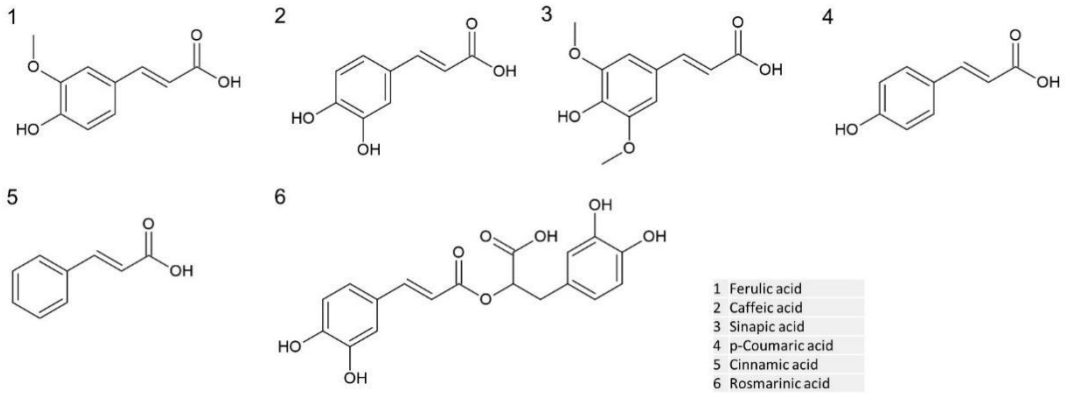


Figure 17: Chemical structures of hydroxycinnamic acids detected in Salicornia sp.

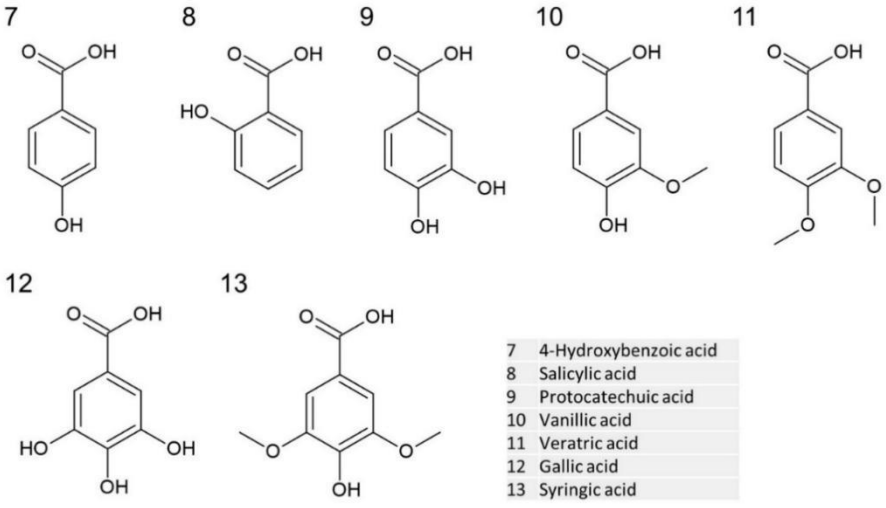


Figure 18: Chemical structures of hydroxybenzoic acids detected in Salicornia sp.

Table 2: Phenolic compounds present in Salicornia species.

	Compound	Salicornia species	Concentration	Analytical method	References
	Hydroxycinnamic acids				
1	Ferulic acid	*S. herbacea*		HPLC-DAD	Essaidi et al. 2013
		S. herbacea	6.87 mg%[1]	UV-Vis ; NMR spectroscopy; HPLC	Oh et al. 2007
		S. europaea	18.2 mg/g[2]	HPLC-UV	Won et al. 2017
		S. neei Lag.	0.39 µg/g[3]	HPLC-UV-Vis	de Souza et al. 2018
		S.herbacea		HPLC-UV	Bi et al. 2012
2	Caffeic acid	*S. herbacea*		HPLC-DAD	Essaidi et al. 2013
		S. herbacea	8.45 mg%[1]	UV-Vis; NMR spectroscopy; HPLC	Oh et al. 2007
		S. patula	0.313 mg/g [4]	HPLC-MS	Sánchez-Gavilán et al. 2021
		S.fruticosa	1.81%[5]	HPLC-UV	Elsebaie et al. 2014
		S. europaea	9.5 mg/g[2]	HPLC-UV	Won et al. 2017
		S. neei Lag.	1.21 µg/g[3]	HPLC-UV-Vis	de Souza et al. 2018
		S.herbacea		HPLC-UV	Bi et al. 2012
		S. europaea	10.07 µg/g[6]	UHPLC-ESI-MS/MS	Zengin et al. 2018
3	Sinapic acid	*S. herbacea*		HPLC-DAD	Essaidi et al. 2013
4	p-Coumaric acid	*S. herbacea*		HPLC-DAD	Essaidi et al. 2013
		S. patula	0.605 mg/g[4]	HPLC-MS	Sánchez-Gavilán et al. 2021
		S. fruticosa	0.42%[5]	HPLC-UV	Elsebaie et al. 2014
		S. europaea	6.8 mg/g[2]	HPLC-UV	Won et al. 2017
		S. europaea	72.06 µg/g[6]	UHPLC-ESI-MS/MS	Zengin et al. 2018
5	Cinnamic acid	*S. herbacea*		HPLC-DAD	Essaidi et al. 2013
		S. patula	0.99 mg/g[4]	HPLC-MS	Sánchez-Gavilán et al. 2021
6	Rosmarinic acid	*S. europaea*	346.41 µg/g[6]	UHPLC-ESI-MS/MS	Zengin et al. 2018
	Hydroxybenzoic acids				
7	4-Hydroxybenzoic acid	*S. fruticosa*	10.92%[5]	HPLC-UV	Elsebaie et al. 2014
8	Salicylic acid	*S. herbacea*		HPLC-DAD	Essaidi et al. 2013
		S. patula	2.92 mg/g[4]	HPLC-MS	Sánchez-Gavilán et al. 2021
9	Protocatechuic acid	*S. herbacea*	1.54 mg%[1]	UV-Vis; NMR spectroscopy; HPLC	Oh et al. 2007
		S. fruticosa	5.16%[5]	HPLC-UV	Elsebaie et al. 2014
		S. europaea	8.4 mg/g[2]	HPLC-UV	Won et al. 2017
		S. neei Lag.	10 µg/g[3]	HPLC-UV-Vis	de Souza et al. 2018
		S.herbacea		HPLC-UV	Bi et al. 2012
10	Vanillic acid	*S. fruticosa*	2.88%[5]	HPLC-UV	Elsebaie et al. 2014
11	Veratric acid	*S. patula*	1.65 mg/g[4]	HPLC-MS	Sánchez-Gavilán et al. 2021
12	Gallic acid	S. neei Lag.	0.64 µg/g[3]	HPLC-UV-Vis	de Souza et al. 2018
		S. europaea	4.24 µg/g[6]	UHPLC-ESI-MS/MS	Zengin et al. 2018
13	Syringic acid	*S. herbacea*		HPLC-DAD	Essaidi et al. 2013
		S. neei Lag.	1.51 µg/g[3]	HPLC-UV-Vis	de Souza et al. 2018
	Caffeoyl quinic acids and derivatives				
14	Quinic acid	*S. europaea*	116.7 µg/g[6]	UHPLC-ESI-MS/MS	Zengin et al. 2018
15	Chlorogenic acid (3-Caffeoylquinic acid)	*S. herbacea*		HPLC-DAD	Essaidi et al. 2013
		S. fruticosa	8.09%[5]	HPLC-UV	Elsebaie et al. 2014
		S. europaea	14.1 mg/g[2]	HPLC-UV	Won et al. 2017
		S. europaea	26.43 µg/g[6]	UHPLC-ESI-MS/MS	Zengin et al. 2018
		S. neei Lag.	3.46 µg/g[3]	HPLC-UV-Vis	de Souza et al. 2018
16	Methyl chlorogenate	*S. herbacea*		NMR; LC-ESI-MS	Cho et al. 2016

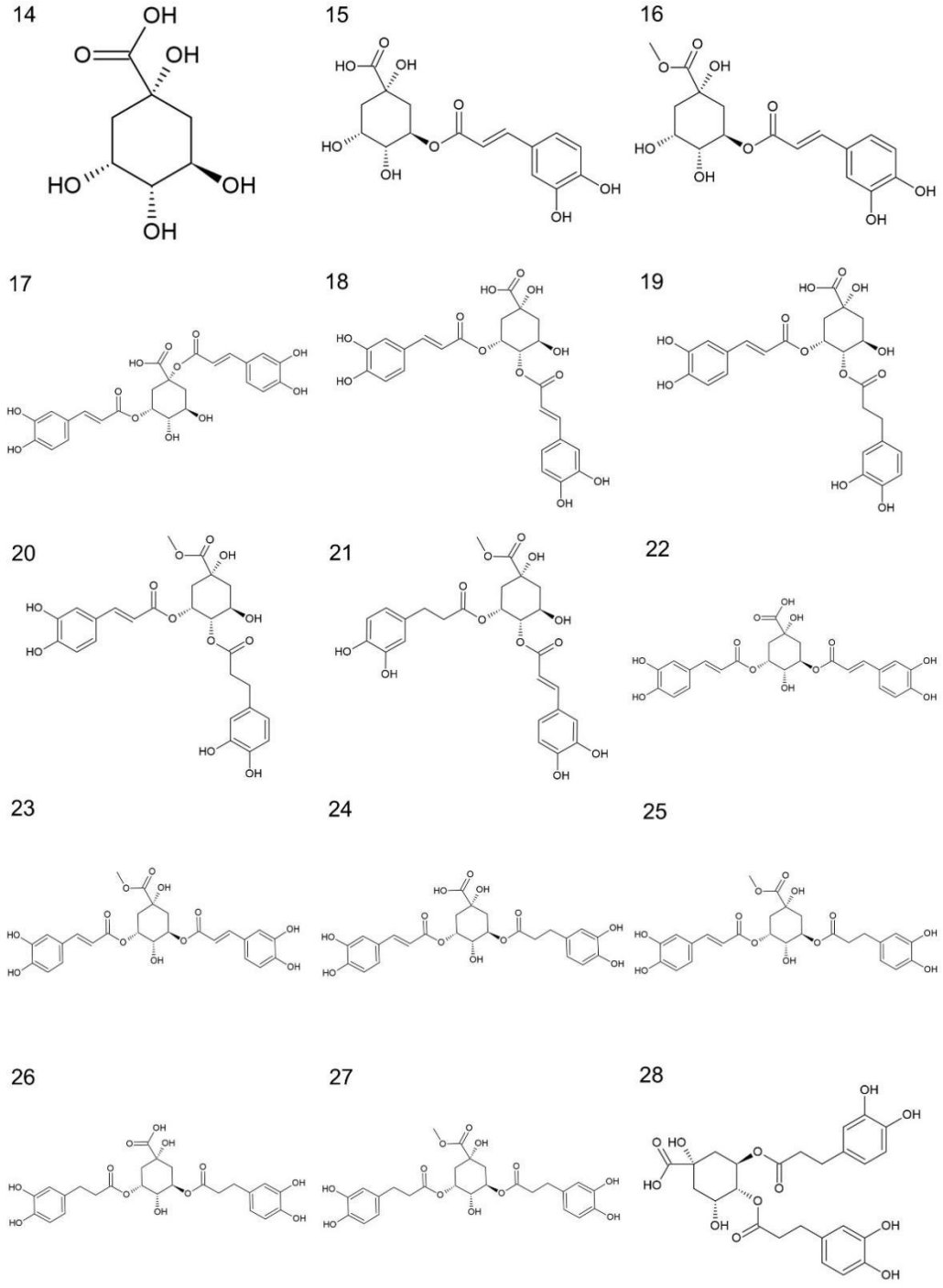
	(3-Caffeoylquinic acid methyl ester)				
17	1,3-di-O-Caffeoyl quinic acid	*S. herbacea*	2.488 mg/g[8]	HR-ESI-MS; NMR	Tuan et al. 2015a
18	3,4-Dicaffeoyl quinic acid	*S. herbacea*		ESI-MS; NMR spectroscopy	Kim et al. 2011
19	Tungtungmadic acid (3-Caffeoyl-4-dihydrocaffeoylquinic acid)	*S. herbacea*		TLC -UV; HRMS; NMR spectroscopy	Chung et al. 2005
		S. herbacea		ESI-MS; NMR spectroscopy	Kim et al. 2011
20	3-Caffeoyl-4-dihydrocaffeoylquinic acid methyl ester	*S. herbacea*	71.9 µg/100 g[9]	NMR; LC-ESI-MS	Cho et al. 2016
21	Methyl 4-caffeoyl-3-dihydrocaffeoyl quinate (salicornate)	*S. herbacea*		ESI-MS; NMR spectroscopy	Kim et al. 2011
22	3,5-Dicaffeoylquinic acid	*S. herbacea*		ESI-MS; NMR spectroscopy	Kim et al. 2011
		S. herbacea	2.930 mg/g[8]	HR-ESI-MS; NMR	Tuan et al. 2015a
23	3,5-di-O-caffeoyl-quinic acid methyl ester	*S. herbacea*	1.765 mg/g[8]	HR-ESI-MS; NMR	Tuan et al. 2015a
	(Methyl 3,5-dicaffeoyl quinate)	*S. herbacea*		ESI-MS; NMR spectroscopy	Kim et al. 2011
24	3-Caffeoyl-5-dihydrocaffeoylquinic acid	*S. herbacea*	75.6 µg/100 g[7]	NMR; LC-ESI-MS	Cho et al. 2016
		S. herbacea	2.225 mg/g[8]	HR-ESI-MS; NMR	Tuan et al. 2015a
25	3-Caffeoyl-5-dihydrocaffeoylquinic acid methyl ester	*S. herbacea*	69.3 µg/100 g[7]	NMR; LC-ESI-MS	Cho et al. 2016
26	3,5-di-O-Dihydrocaffeoyl quinic acid	*S. herbacea*	1.026 mg/g[8]	HR-ESI-MS; NMR	Tuan et al. 2015a
27	3,5-di-Dihydrocaffeoylquinic acid methyl ester	*S. herbacea*	171.9 µg/100 g[7]	NMR; LC-ESI-MS	Cho et al. 2016
28	4,5-di-O-Dihydrocaffeoyl quinic acid	*S. herbacea*	2.059 mg/g[8]	HR-ESI-MS; NMR	Tuan et al. 2015a
	Flavonoids and Flavanones				
29	Quercetin	*S. neei Lag.*	14.8 µg/g[3]	HPLC-UV-Vis	de Souza et al. 2018
		S. herbacea		HPLC-DAD	Essaidi et al. 2013
		S. europaea		MS; NMR	Lyu et al. 2018
		S. europaea	2.5 mg/g[2]	HPLC-UV	Won et al. 2017
30	Quercetin 3',4'-dimethyl ether	*S. fruticosa*		paper chromatography; UV; NMR spectroscopy	Abdel Elatif et al., 2020
31	Isoquercetin	*S. europaea*	3.4 mg/g[2]	HPLC-UV	Won et al. 2017
	(Quercetin-3-O-β-D-glucopyranoside)	*S. herbacea*	3992.49 ppm[9]	LC/MS	Kim et al. 2008
		S. herbacea		column chromatography	Kim and Park 2004
		S. herbacea		ESI-MS; NMR spectroscopy	Kim et al. 2011
		S. europaea		MS; NMR	Geslin and Verbist 1985
		S. europaea		MS; NMR	Lyu et al. 2018
32	Quercetin-3',4'-diglucoside	*S. herbacea*	0.08 ppm[9]	LC/MS	Kim et al. 2008
33	Quercetin-3-O-(6''-O-malonyl)-β-d-glucoside	*S. europaea*		MS; NMR spectroscopy	Geslin and Verbist 1985
34	Isoquercitrin 6''-O-methyloxalate	*S. herbacea*		ESI-MS; NMR spectroscopy	Kim et al. 2011
35	Rutin	*S. herbacea*	2.57 ppm[9]	LC/MS	Kim et al. 2008
	(Quercetin-3-O-rutinoside)	*S. patula*		HPLC-MS	Sánchez-Gavilán et al. 2021
		S. europaea	13.22 µg/g[6]	UHPLC-ESI-MS/MS	Zengin et al. 2018
		S. europaea		MS; NMR	Lyu et al. 2018
36	Rhamnetin (Quercetin 7-methyl ether)	*S. herbacea*		HPLC-DAD	Essaidi et al. 2013
		S. europaea	33.26 µg/g[6]	UHPLC-ESI-MS/MS	Zengin et al. 2018
37	Isorhamnetin (Quercetin 3'-methyl ether)	*S. herbacea*		HPLC-DAD	Essaidi et al. 2013
		S. herbacea	27.81 ppm[9]	LC/MS	Kim et al. 2008
		S. herbacea	6.65 mg%[1]	UV-Vis; NMR spectroscopy; HPLC	Oh et al. 2007
		S. europaea	18.4 mg/g[2]	HPLC-UV	Won et al. 2017
		S. *fruticosa*		paper chromatography; UV; NMR spectroscopy	Abdel Elatif et al., 2020
		S. fruticosa		PC; TLC; UV; MS	Radwan and Nazif 2007
38	Isorhamnetin 3-O-β-D-glucopyranoside	*S. herbacea*		column chromatography; EI-MS; NMR spectroscopy; IR spectroscopy	Lee et al., 2004
		S. herbacea		Column chromatography	Kim and Park 2004
		S. herbacea		ESI-MS; NMR spectroscopy	Kim et al. 2011

		S. europaea	16.2 mg/g[2]	HPLC-UV	Won et al. 2017
		S. fruticosa		paper chromatography; UV; NMR spectroscopy	Abdel Elatif et al., 2020
		S. europaea		MS; NMR	Geslin and Verbist 1985
		S. europaea		MS; NMR	Lyu et al. 2018
		S. herbacea		Column chromatography; MS; NMR	Lee et al. 2004
39	Isorhamnetin-3-O-galactoside	*S. fruticosa*		PC; TLC; UV spectroscopy; MS	Radwan and Nazif 2007
40	Isorhamnetin 3-O-rhamnoside	*S. fruticosa*		Paper chromatography; UV; NMR spectroscopy	Abdel Elatif et al., 2020
41	Isorhamnetin 3-O-rutinoside	*S. fruticosa*		Paper chromatography; UV; NMR spectroscopy	Abdel Elatif et al., 2020
42	Isorhamnetin 3-O-neohesperidoside	*S. fruticosa*		Paper chromatography; UV; NMR spectroscopy	Abdel Elatif et al., 2020
43	Isorhamnetin 3-O-rhamnosyl(1-2)arabinoside	*S. fruticosa*		Paper chromatography; UV; NMR spectroscopy	Abdel Elatif et al., 2020
44	Hesperetin	*S. herbacea*		HPLC-DAD	Essaidi et al. 2013
45	Hesperidin (Hesperetin 7-rutinoside)	*S. europaea*	13.74 µg/g[6]	UHPLC-ESI-MS/MS	Zengin et al. 2018
46	Acacetin	*S. herbacea*		HPLC-DAD	Essaidi et al. 2013
		S. fruticosa		PC; TLC; UV; MS	Radwan and Nazif 2007
47	Galangin	*S. herbacea*		HPLC-DAD	Essaidi et al. 2013
48	Myricetin	*S. herbacea*		HPLC-DAD	Essaidi et al. 2013
49	Apigenin	S. fruticosa		PC; TLC; UV; MS	Radwan and Nazif 2007
50	Apigenin 7-glucoside	*S. Patula*		HPLC-MS	Sánchez-Gavilán et al. 2021
51	Apigenin-7-O-galactoside	*S. fruticosa*		PC; TLC; UV; MS	Radwan and Nazif 2007
52	Pelargonidin-3-O-rutinoside	*S. Patula*		HPLC-MS	Sánchez-Gavilán et al. 2021
53	Kaempferol	*S. herbacea*		HPLC-DAD	Essaidi et al. 2013
		S. patula		HPLC-MS	Sánchez-Gavilán et al. 2021
		S. europaea		MS; NMR	Lyu et al. 2018
54	Kaempferol-3-O-β-D-glucoside (Astragalin)	*S. europaea*		MS; NMR	Lyu et al. 2018
55	Luteolin	*S. europaea*		MS; NMR	Lyu et al. 2018
56	Chrysin	*S. fruticosa*	0.04%[5]	HPLC-UV	Elsebaie et al. 2014
57	Catechin	*S. fruticosa*	10.26%[5]	HPLC-UV	Elsebaie et al. 2014
58	2S-5,2′-Dihydroxy-6,7-methylenedioxyflavanone	*S. herbacea*		HRESIMS; NMR	Tuan et al. 2015b
59	(−)-(2S)-2′-Hydroxy-6,7-methylenedioxyflavanone	*S. europaea*		MS; IR spectroscopy; UV- spectroscopy; NMR	Arakawa et al. 1982
60	2′-Hydroxy-6,7-methylenedioxyisoflavone	*S. europaea*		MS; IR spectroscopy; UV- spectroscopy; NMR	Arakawa et al. 1982
61	2S-2′-Hydroxy-6,7-dimethoxy-flavanone	*S. herbacea*		HRESIMS; NMR	Tuan et al. 2015b
62	2S-2',7- dihydroxy-6-methoxyflavanone	*S. herbacea*		HRESIMS; NMR	Tuan et al. 2015b
63	2′,7-dihydroxy-6-methoxyisoflavone	*S. europaea*		MS; IR spectroscopy; UV- spectroscopy; NMR	Arakawa et al. 1982
64	Irilin B	*S. europaea*		ESI-MS; H- and C- NMR spectroscopy	Kim et al. 2019
	Sterols				
65	β-Sitosterol	*S. herbacea*		column chromatography; EI-MS; NMR spectroscopy spectroscopy; IR spectroscopy	Lee et al., 2004
		S. europaea		MS; NMR	Lyu et al. 2018
66	Stigmasterol	*S. herbacea*		column chromatography; EI-MS; NMR spectroscopy spectroscopy; IR spectroscopy	Lee et al., 2004
		S. herbacea		column chromatography; NMR spectroscopy spectroscopy; ESI-MS	Wang et al. 2013
		S. europaea		MS; NMR	Lyu et al. 2018
67	Ergosterol	*S. herbacea*		column chromatography; NMR spectroscopy spectroscopy; ESI-MS	Wang et al. 2013
68	Cerevisterol	*S. europaea*		MS; NMR	Lyu et al. 2018
69	β-Daucosterol	*S. europaea*		MS; NMR	Lyu et al. 2018
	Chromones				
70	6,7-Methylenedioxychromone	*S. europaea*		IR spectroscopy; MS; NMR	Arakawa et al. 1983
71	6,7-Dimethoxychromone	*S. europaea*		IR spectroscopy; MS; NMR	Arakawa et al. 1983

		S. herbacea	HRESIMS; NMR	Tuan et al. 2015b
72	7-Hydroxy-6-methoxychromone	*S. europaea*		Arakawa et al. 1983
73	7-O-β-d-Glucopyranosyl-6-methoxychromone	*S. europaea*		Arakawa et al. 1983
		S. europaea	MS; NMR	Lyu et al. 2018
74	7-Hydroxy-6,8-dimethoxychromone	*S. herbacea*	HRESIMS; NMR	Tuan et al. 2015b
75	6-Methoxychromanone	*S. herbacea*	HRESIMS; NMR	Tuan et al. 2015b
	Lignans			
76	(−)-Syringaresinol	*S. europaea*	MS; NMR	Lyu et al. 2018
77	Syringaresinol 4-O-beta-D-glucopyranoside	*S. europaea*	Chromatographic techniques; spectroscopic methods (not specified)	Wang et al. 2011
78	Episyringaresinol-4''-O-β-D-glucopyranoside	*S. europaea*	MS; NMR	Lyu et al. 2018
79	Acanthoside B	*S. europaea*	Chromatographic techniques; spectroscopic methods (not specified)	Wang et al. 2011
80	Erythro-1-(4-O-β-d-glucopyranosyl-3,5-dimethoxyphenyl)-2-syringaresinoxyl-propane-1,3-diol	*S. europaea*	Chromatographic techniques; spectroscopic methods (not specified)	Wang et al. 2011
		S. europaea	MS; NMR	Lyu et al. 2018
81	Longifloroside B	*S. europaea*	Chromatographic techniques; spectroscopic methods (not specified)	Wang et al. 2011
		S. europaea	MS; NMR	Lyu et al. 2018
	Oleane Triterpenoid Saponins			
82	Akebonic acid	*S. europaea*	MS; NMR	Lyu et al. 2018
83	Boussingoside A1	*S. europaea*	MS; NMR	Lyu et al. 2018
84	Boussingoside A2	*S. europaea*	MS; NMR	Lyu et al. 2018
85	3β-Hydroxy-23-oxo-30-noroleana-12,20 (29)-diene-28-oic acid 3-O-β-d-glucuronopyranosyl-28-O-β-d-glucopyranoside	*S. herbacea*	NMR	Kim et al. 2012
86	30-Norhederagenin 3-O-β-d-glucuronopyranosyl-28-O-β-d-glucopyranoside	*S. herbacea*	NMR	Kim et al. 2012
87	3-O-[β-D-Glucuronopyranosyl-6'-O-methyl ester]-30-norolean-12,20(29)-dien-28-O-[β-D-glucopyranosyl] ester	*S. europaea*	MS; NMR	Lyu et al. 2018
88	Salieuropaea A	*S. europaea*	MS; NMR	Lyu et al. 2018
89	Salbige A	*S. herbacea*	HRMS; NMR spectroscopy	Zhao et al. 2014
90	Salbige B	*S. herbacea*	HRMS; NMR spectroscopy	Zhao et al. 2014
91	Gypsogenin	*S. herbacea*	HRMS; NMR spectroscopy	Zhao et al. 2014
92	Gypsogenin 3-O-β-d-glucuronopyranoside	*S. herbacea*	NMR	Kim et al. 2012
93	Gypsogenin 3-O-β-d-glucuronopyranosyl-28-O-β-d-glucopyranoside	*S. herbacea*	NMR	Kim et al. 2012
94	Oleanolic acid	*S. europaea*	MS; NMR	Lyu et al. 2018
		S. herbacea	HRMS; NMR spectroscopy	Zhao et al. 2014
95	Oleanolic acid 3-O-β-D-glucopyranoside	*S. europaea*	MS; NMR	Lyu et al. 2018
96	Oleanolic acid 28-O-β-D-glucopyranoside	*S. europaea*	HR-ESI-MS; NMR	Yin et al. 2012
		S. europaea	MS; NMR	Lyu et al. 2018
97	Oleanolic acid 3-O-β-D-glucuronopyranoside (calenduloside E)	*S. europaea*	MS; NMR	Lyu et al. 2018
		S. europaea	HR-ESI-MS; NMR	Yin et al. 2012
98	Oleanolic acid-3-O-6'-O-methyl-β-D-glucuronopyranoside (calenduloside E 6'-methyl ester)	*S. europaea*	MS; NMR	Lyu et al. 2018
		S. europaea	HR-ESI-MS; NMR	Yin et al. 2012
99	Chikusetsusaponin Iva	*S. europaea*	MS; NMR	Lyu et al. 2018
100	Chikusetsusaponin Iva methyl ester	*S. europaea*	MS; NMR	Lyu et al. 2018
		S. europaea	HR-ESI-MS; NMR	Yin et al. 2012

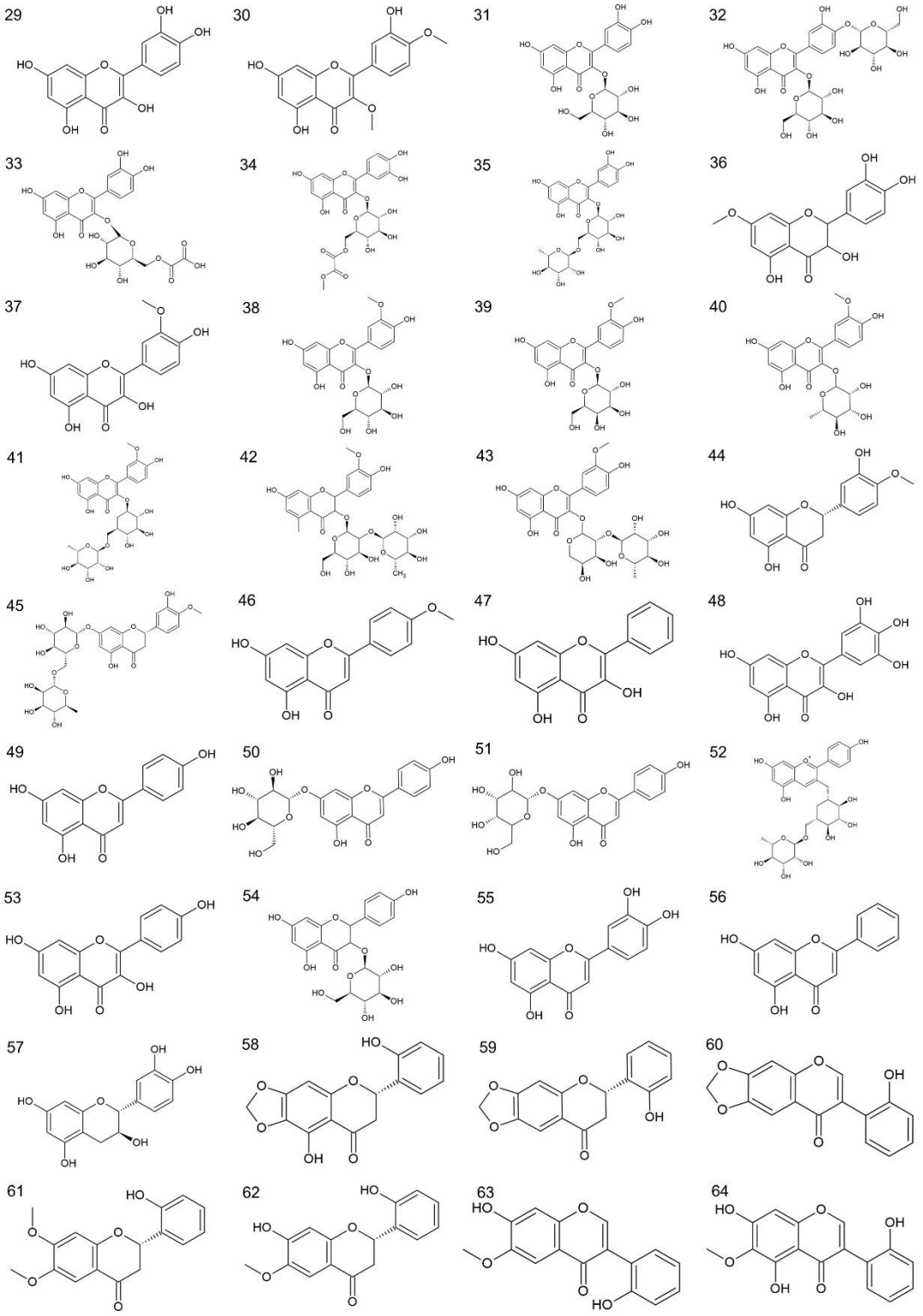
101	3β,29-Dihydroxy-olean-12-en-28-oic acid 28-O-β-D-glucopyranosyl ester	*S. europaea*		HR-ESI-MS; NMR	Yin et al. 2012
102	Zygophyloside K	*S. europaea*		MS; NMR	Lyu et al. 2018
	Others				
103	Pyrogallol	*S. fruticosa*	28.14%[5]	HPLC-UV	Elsebaie et al. 2014
104	Vanillin	*S. neei Lag.*	2.73 μg/g[3]	HPLC-UV-Vis	de Souza et al. 2018
		S. herbacea		column chromatography; NMR spectroscopy; ESI-MS	Wang et al. 2013
105	Uracil	*S. herbacea*		column chromatography; EI-MS; NMR spectroscopy; IR spectroscopy	Lee et al., 2004
106	Caffeine	*S. fruticosa*	4.38%[5]	HPLC-UV	Elsebaie et al. 2014
107	Scopoletin	*S. herbacea*		column chromatography; NMR spectroscopy; ESI-MS	Wang et al. 2013
108	Pentadecyl ferulate	*S. herbacea*		column chromatography; NMR spectroscopy; ESI-MS	Wang et al. 2013
109	Dibutyl phthalate	*S. herbacea*		column chromatography; NMR spectroscopy; ESI-MS	Wang et al. 2013
110	Dioctyl phthalate	*S. herbacea*		column chromatography; NMR spectroscopy; ESI-MS	Wang et al. 2013
111	Icariside B2	*S. europaea*		MS; NMR	Lyu et al. 2018
112	Ellagic acid	*S. fruticosa*	25.41%[5]	HPLC-UV	Elsebaie et al. 2014
113	Pheophorbide A	*S. herbacea*		HRMS; NMR spectroscopy	Zhao et al. 2014
114	(13^2 S)-Hydroxy-pheophorbide A	*S. herbacea*		HRMS; NMR spectroscopy	Zhao et al. 2014
115	(13^2 S)-Hydro-pheophorbide-lactone A	*S. herbacea*		HRMS; NMR spectroscopy	Zhao et al. 2014

[1]: In dry base, as detected in the EtOH extract of the viscozyme-treated plant; [2]: In desalted Salicornia extract; [3]: In Salicornia shoots, dry weight; [4]: In dry weight of the plant, as determined in different samples; [5]: In the methanolic extract of the air part of the plant; [6]: In methanol extract; [7]: In the fresh plant; [8]: In crude plant extract; [9]: In powder sample of the plant.



14	Quinic acid	22	3,5-Dicaffeoylquinic acid
15	Chlorogenic acid	23	3,5-di-O-caffeoyl-quinic acid methyl ester
16	Methyl chlorogenate	24	3-Caffeoyl-5-dihydrocaffeoylquinic acid
17	1,3-di-O-Caffeoyl quinic acid	25	3-Caffeoyl-5-dihydrocaffeoylquinic acid methyl ester
18	3,4-Dicaffeoyl quinic acid	26	3,5-di-O-Dihydrocaffeoyl quinic acid
19	Tungtungmadic acid	27	3,5-di-Dihydrocaffeoylquinic acid methyl ester
20	3-Caffeoyl-4-dihydrocaffeoylquinic acid methyl ester	28	4,5-di-O-Dihydrocaffeoyl quinic acid
21	Methyl 4-caffeoyl-3-dihydrocaffeoyl quinate		

Figure 19: Chemical structures of caffeoylquinic acids and their derivatives detected in Salicornia sp.



29	Quercetin	41	Isorhamnetin 3-O-rutinoside	53	Kaempferol
30	Quercetin 3′,4′-dimethyl ether	42	Isorhamnetin 3-O-neohesperidoside	54	Kaempferol-3-O-β-D-glucoside
31	Isoquercetin	43	Isorhamnetin 3-O-rhamnosyl(1-2)arabinoside	55	Luteolin
32	Quercetin-3′,4′-diglucoside	44	Hesperetin	56	Chrysin
33	Quercetin-3-O-(6″-O-malonyl)-β-d-glucoside	45	Hesperidin	57	Catechin
34	Isoquercitrin 6″-O-methyloxalate	46	Acacetin	58	2S-5,2′-Dihydroxy-6,7- methylenedioxyflavanone
35	Rutin	47	Galangin	59	(−)-(2S)-2′-Hydroxy-6,7- methylenedioxyflavanone
36	Rhamnetin	48	Myricetin	60	2′-Hydroxy-6,7-methylenedioxyisoflavone
37	Isorhamnetin	49	Apigenin	61	2S-2′-Hydroxy-6,7-dimethoxy-flavanone
38	Isorhamnetin 3-O-β-D-glucopyranoside	50	Apigenin 7-glucoside	62	2S-2′,7- dihydroxy-6-methoxyflavanone
39	Isorhamnetin-3-O-galactoside	51	Apigenin-7-O-galactoside	63	2′,7-dihydroxy-6-methoxyisoflavone
40	Isorhamnetin 3-O-rhamnoside	52	Pelargonidin-3-O-rutinoside	64	Irilin B

Figure 20: Chemical structures of flavonoids detected in Salicornia sp.

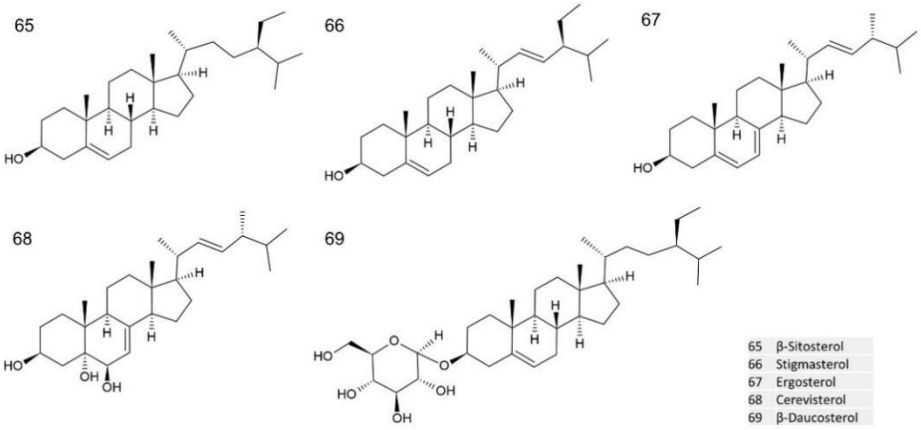


Figure 21: Chemical structures of sterols detected in Salicornia sp.

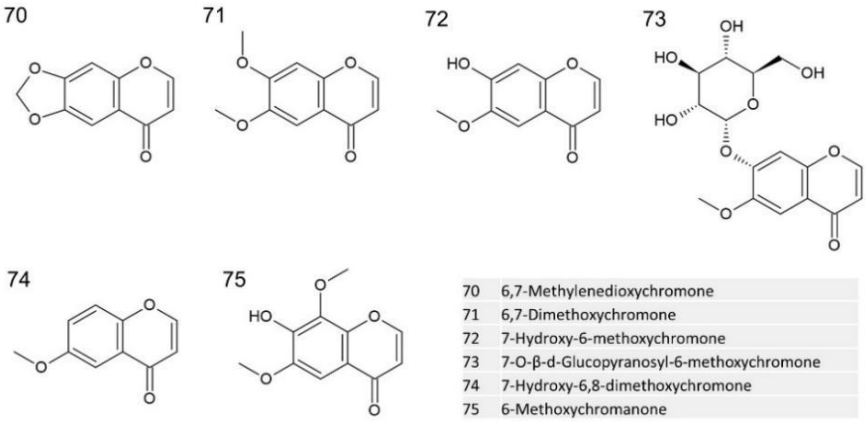


Figure 22: Chemical structures of chromones detected in Salicornia sp.

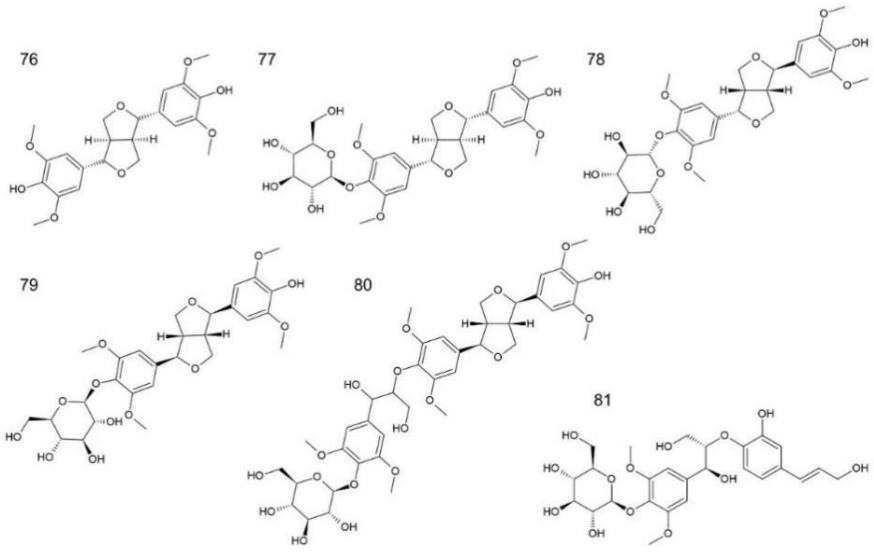
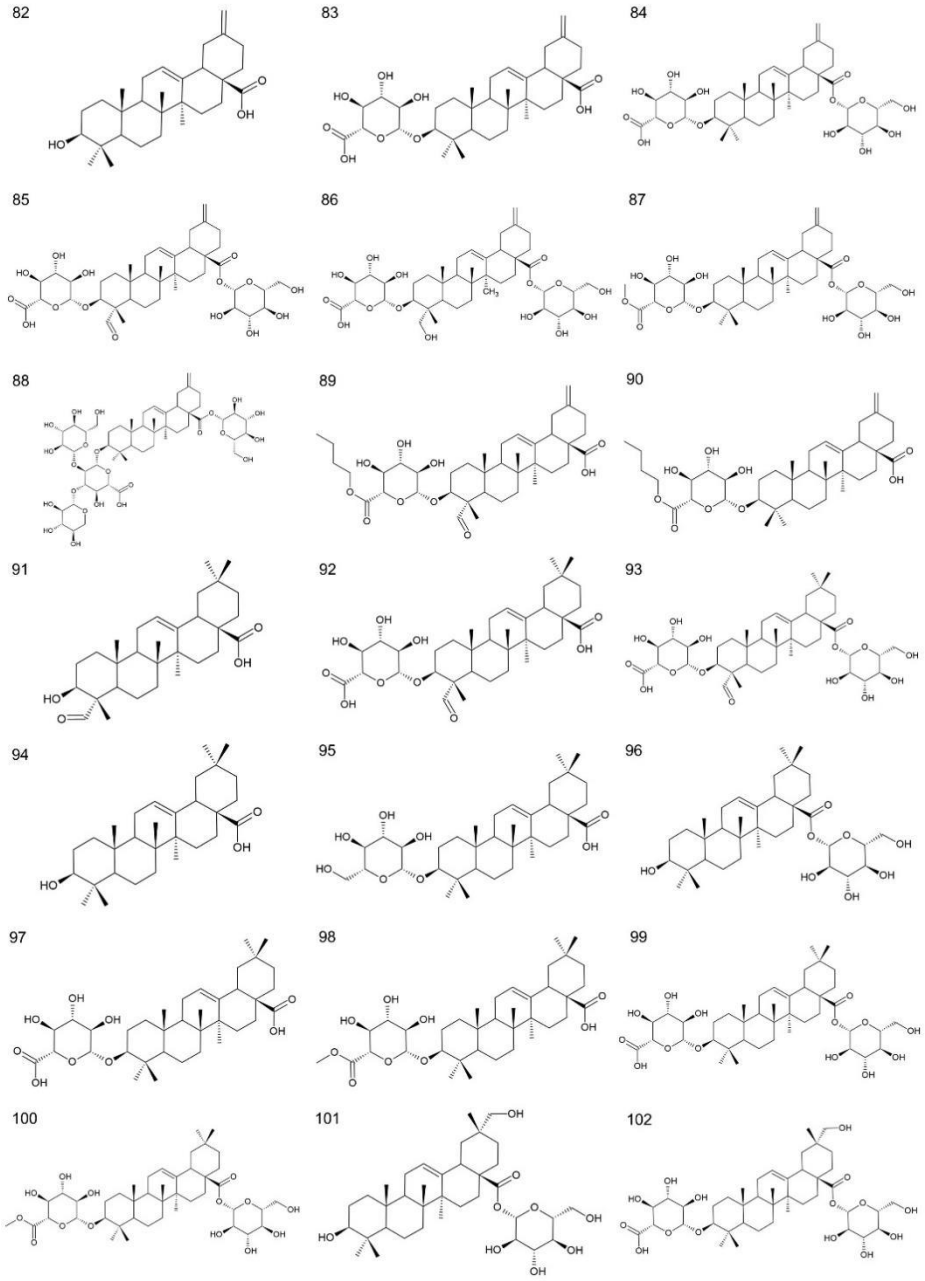
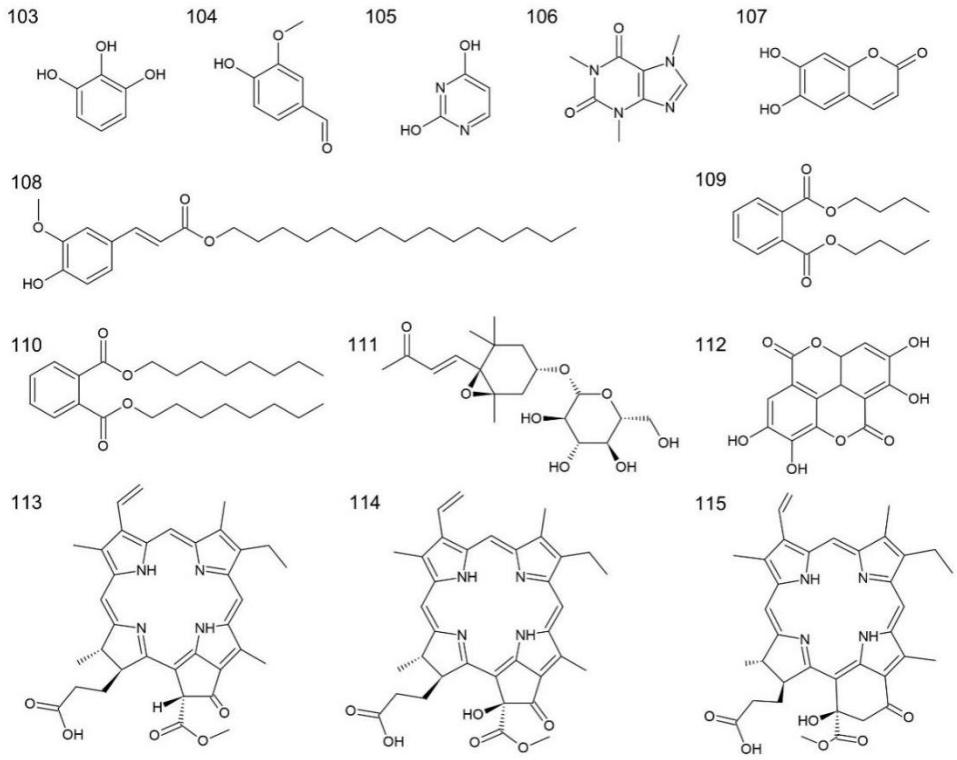


Figure 23: Chemical structures of lignans detected in Salicornia sp.



82	Akebonic acid
83	Boussingoside A1
84	Boussingoside A2
85	3β-Hydroxy-23-oxo-30-noroleana-12,20(29)-diene-28-oic acid 3-O-β-d-glucuronopyranosyl-28-O-β-d-glucopyranoside
86	30-Norhederagenin 3-O-β-d-glucuronopyranosyl-28-O-β-d-glucopyranoside
87	3-O-[β-D-Glucuronopyranosyl-6′-O-methyl ester]-30-norolean-12,20(29)-dien-28-O-[β-D-glucopyranosyl] ester
88	Salieuropaea A
89	Salbige A
90	Salbige B
91	Gypsogenin
92	Gypsogenin 3-O-β-d-glucuronopyranoside
93	Gypsogenin 3-O-β-d-glucuronopyranosyl-28-O-β-d-glucopyranoside
94	Oleanolic acid
95	Oleanolic acid 3-O-β-D-glucopyranoside
96	Oleanolic acid 28-O-β-D- glucopyranoside
97	Oleanolicacid 3-O-β-D-glucuronopyranoside (calenduloside E)
98	Oleanolic acid-3-O-6′-O-methyl-β-D-glucuronopyranoside (calenduloside E 6′-methyl ester)
99	Chikusetsusaponin Iva
100	Chikusetsusaponin Iva methyl ester
101	3β,29-Dihydroxy-olean-12-en-28-oic acid 28-O-β-D-glucopyranosyl ester
102	Zygophyloside K

Figure 24: Chemical structures of saponins detected in Salicornia sp.



103	Pyrogallol	110	Dioctyl phthalate
104	Vanillin	111	Icariside B2
105	Uracil	112	Ellagic acid
106	Caffeine	113	Pheophorbide A
107	Scopoletin	114	(13^2 S)-Hydroxy-pheophorbide A
108	Pentadecyl ferulate	115	(13^2 S)-Hydro-pheophorbide-lactone A
109	Dibutyl phthalate		

Figure 25: Chemical structures of uncategorized compounds detected in Salicornia sp.

2. AIM OF THE THESIS

Aim of this thesis is the evaluation of bioactive compounds found in *Salicornia* extracts as potential inhibitors of SARS-CoV-2 M^{pro}. M^{pro} has been proved to be a main non-structural protein target for the development of drugs and immune boosting agents, aiming the inhibition of SARS-CoV-2 infection in human cells. *Salicornia* extracts have been shown to contain a vast variety of bioactive compounds with high antiviral potential. However, only very few of these compounds, have been described a potential inhibitors of M^{pro} via *in silico* simulations. Thus, there is great potential into valorizing this highly resistant and abundant plant for producing high-added value immunity boosting compounds.

The specific goals of the thesis are:

-The identification of key residues that are involved in M^{pro} inhibition, visualizing the co-crystallization data of the M^{pro} structure with known inhibitors and performing small molecule docking simulations. The study will provide valuable knowledge towards establishing a structure-function relationship for studied inhibitory compounds.

-Evaluation of the different classes of compounds (hydroxycinnamic acids, hydroxybenzoic acids, caffeoylquinic acids and derivatives, flavonoids and derivatives, sterols, chromones, lignans, saponins) found in *Salicornia* extract for their binding potential onto the active site of M^{pro}, via small molecule docking simulations

-Evaluation of the most promising phenolic compounds and a *Salicornia* extract for their inhibitory effect against M^{pro} using an *in vitro* assay.

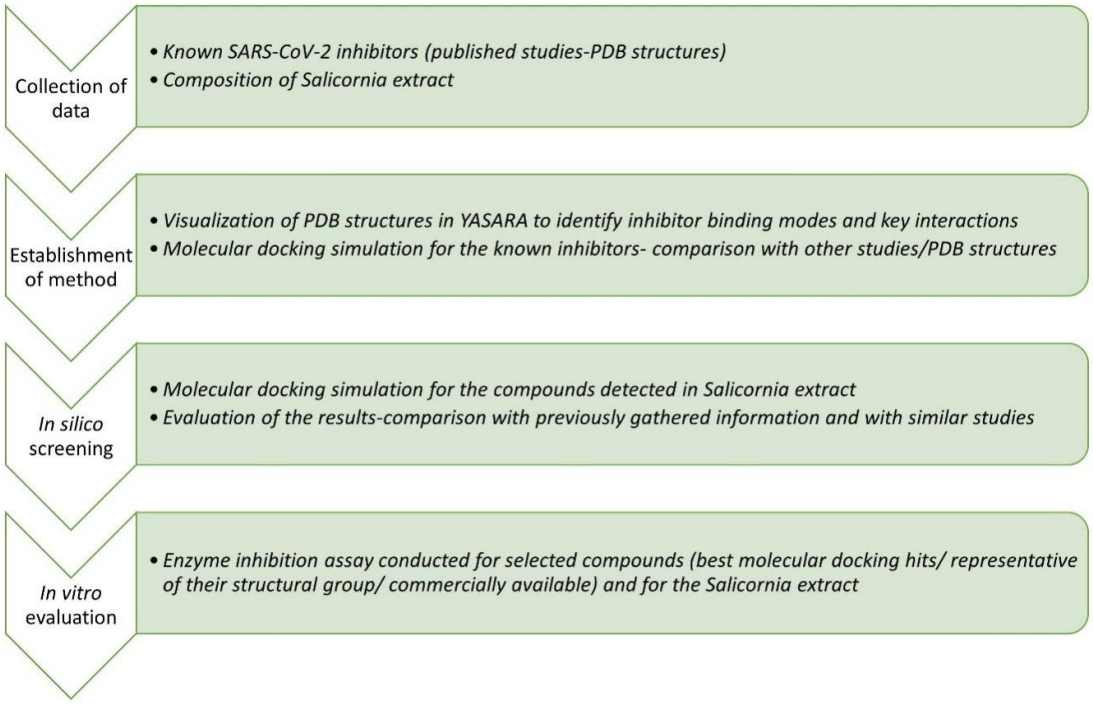


Figure 26: Graphical representation of the steps followed in the present thesis.

3. MATERIALS AND METHODS

3.1. Evaluation of inhibitory effect via *in silico* simulations

3.1.1. Visualization of co-crystallization data for known inhibitors and M^{pro}

Initially, the structures of the co-crystallized inhibitors in complex with M^{pro} that were available in the Protein Data Bank (PDB) (https://www.rcsb.org/) were visualized YASARA Structure version 20.12.24. Its feature that allows demonstration of the hydrogen bonds, hydrophobic and pi-pi interactions formed between molecules was utilized in order to identify the interactions of the inhibitors and therefore the key residues of the active site involved in them.

3.1.2. Ligand and receptor preparation for docking simulations

3.1.2.1. Docking of known inhibitors to verify simulation accuracy

The docking simulation essentially calculates the lowest possible energy and the conformation that leads to it, of the complex between a larger macromolecule (receptor) and a smaller molecule (ligand). For the docking to be executed, the receptor and the ligand need to be defined and prepared. To confirm the repeatability and reliability of the simulation, already confirmed SARS-CoV-2-M^{pro} inhibitors, which have been co-crystallized in complex with the protease, were initially docked to the respective structure. The ligand and receptor structures were taken from the respective PDB file. In the cases where the protease structure was in the form of a dimer, only one of the monomers was used for the docking. Any water or solvent molecules were deleted. The ligands were energy minimized through YASARA, the structure of the receptor was cleaned and then its hydrogen network was optimized. The simulation cell was built as a cube centered in the atoms of the catalytic dyad His 41 and Cys 145 and extended as many Å as needed for its side to be 2-3 Å longer than the length of the ligand (Table 3). This margin allowed for flexibility for the ligand to acquire different conformations onto the active site, however, it was not so large in order to increase the uncertainty of the docking simulation. The hydrogen bonds and other interactions formed between the ligands and the prtease structures were verified with the accompanying literature.

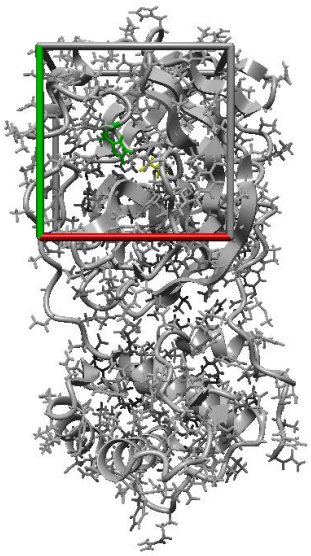


Figure 27: Simulation cell as defined around the catalytic dyad (His 41 in green, Cys 145 in yellow) for ferulic acid (with a 7 Å extension around the catalytic dyad) (PDB ID: 6LU7)

Table 3: PDB structures used for M^{pro} inhibitor docking and simulation cell based on the inhibitor length

Inhibitor	PDB ID	Length (Å)	Simulation cell extension (Å)
N3	6LU7		
covalent			
11a	6LZE	14	8
11b	6M0K	13	8
5h	7JKV	15.5	9
GC376	7D1M	15	9
GC373	6WTK	12	7
Narlaprevir	7JYC	19.3	11
Telaprevir	7K6D	21.2	12
x2754	5RHF	10.5	6.5
x2705	5RH7	12.7	7.5
Myricetin	7B3E	12	7
Mg-132	7BE7	17	9.5
MI-23	7D3I	14	8
Carmofur	7BUY	10	6
Boceprevir	7C6S	16.2	9
Calpeptin	7AKU	12.4	7.5
13b	6Y2G	16.7	9.5
Non-covalent			
Compound 5	7L11	16	9
Compound 26	7L14	16	9
x0397	5RGI	10.5	6.5
x77	6W63	13.5	8
Mcule-5948770040	7LTJ	14.2	8
ML 188	7L0D	13.2	8
MUT056399	7AP6	14	8
x0104	5R7Z	9.3	6
x0161	5R80	10	6

3.1.2.2. Docking of bioactive compounds from Salicornia extracts

After the reliability of the docking method was confirmed, docking of the different classes of compounds identified in *Salicornia* extracts was performed. The compounds are enlisted also in Table 2. In this case, each ligand was constructed in ChemSketch and its 3D structure was optimized through the same program and saved in a .mol file, before being imported and energy minimized in YASARA. The protease structure used for the simulation is the one under PDB ID 6LU7, with a resolution of 2.16 Å. This structure was chosen among the numerous structures deposited in the PDB because it was the first one to be made available and the one that is more often used in other studies featuring molecular docking simulations. The receptor and the simulation cell were prepared in the same method as mentioned above. The simulation cell defined for each ligand is presented in Table 4.

Table 4: Simulation cell as a function of ligand length for the docked phenolic compounds. The M^{pro} structure was PDB: 6LU7

Ligand	Ligand length (Å)	Simulation cell extension (Å)]
Hydroxycinnamic acids		
Ferulic acid	12	7
Caffeic acid	10.4	6.5
Sinapic acid	10.5	6.5
p-Coumaric acid	10.3	6.5
Cinnamic acid	8.6	5.5
Rosmarinic acid	17.4	10
Hydroxybenzoic acids		

4-Hydroxybenzoic acid	7	4.5
Salicylic acid	7	4.5
Protocatechuic acid	7.7	5
Vanillic acid	6.9	4.5
Veratric acid	9.3	6
Gallic acid	7.8	5
Syringic acid	9	5.5
Caffeoyl quinic acids and derivatives		
Quinic acid	7.8	5
Chlorogenic acid	15	8.5
Methyl chlorogenate	17	9.5
1,3-di-O-Caffeoyl quinic acid	23	12.5
3,4-Dicaffeoylquinic acid	16	9
3-Caffeoyl-4-dihydrocaffeoylquinic acid	17	9.5
3-Caffeoyl-4-dihydrocaffeoylquinic acid methyl ester	17	9.5
Methyl 4-caffeoyl-3-dihydrocaffeoyl quinate	17	9.5
3,5-di-O-Caffeoyl quinic acid	22	12
3,5-di-O-Caffeoyl-quinic acid methyl ester	22.2	12
3-O-Caffeoyl-5-O-dihydrocaffeoyl quinic acid	21.3	12
3-Caffeoyl-5-dihydrocaffeoylquinic acid methyl ester	22	12
3,5-di-O-Dihydrocaffeoyl quinic acid	21	11.5
3,5-di-Dihydrocaffeoylquinic acid methyl ester	21.2	11.5
4,5-di-O-Dihydrocaffeoyl quinic acid	15.5	9
Flavonoids		
Quercetin	12	7
Quercetin 3,4’-dimethyl ether	13.7	8
Isoquercetin	12.6	7.5
Quercetin-3’,4’ glucoside	17.2	9.5
Quercetin 3-O-(6”-O-malonyl)-β-d-glucoside	15	8.5
Isoquercitrin 6’’-O-methyloxalate	19.2	11
Rutin	17	10
Rhamnetin	12.7	7.5
Isorhamnetin	11.7	7
Isorhamnetin 3-O-β-D-glucopyranoside	13.2	7.5
Isorhamnetin-3-O-galactoside	12.7	7.5
Isorhamnetin 3-O-rhamnoside	12.5	7.5
Isorhamnetin 3-O-rutinoside	19	10.5
Isorhamnetin 3-O-neohesperidoside	16	9
Isorhamnetin 3-O-rhamnosyl(1-2)arabinoside	16.6	9.5
Hesperetin	13.8	8
Hesperidin	22.6	12.5
Acacetin	14	8
Galangin	11.5	7
Myricetin	12	7
Apigenin	12.4	7
Apigenin 7-glucoside	16.7	9.5
Apigenin-7-O-galactoside	16.7	9.5
Pelargonidin-3-O-rutinoside	17	9.5
Kaempferol	12.3	7
Astragalin	12.8	7.5
Luteolin	12.4	7
Chrysin	11.4	7
Catechin	12.2	7
2S-5,2’-Dihydroxy-6,7- methylenedioxyflavanone	12.5	7.5
(−)-(2S)-2’-Hydroxy-6,7- methylenedioxyflavanone	12.5	7.5
2’-Hydroxy-6,7-methylenedioxyisoflavone	13	7.5
2S-2’-Hydroxy-6,7-dimethoxy-flavanone	13.6	8
2S-2',7- Dihydroxy-6-methoxyflavanone	13.6	8
2’,7-Dihydroxy-6-methoxyisoflavone	13.2	7.5
Irilin B	13.3	7.5
Sterols		
β-Sitosterol	19.2	11
Stigmasterol	17.8	10
Ergosterol	17.7	10
Cerevisterol	17.5	10
β-Daucosterol	23.6	13
Chromones		

6,7-Methylenedioxychromone	8.7	5.5
6,7-Dimethoxychromone	9.5	6
7-Hydroxy-6-methoxychromone	9.5	6
7-O-β-d-Glucopyranosyl-6-methoxychromone	12.2	7
7-Hydroxy-6,8-dimethoxychromone	9.5	6
6-Methoxychromanone	9.5	6
Lignans		
(−)-Syringaresinol	16.5	9.5
Syringaresinol 4-O-beta-D-glucopyranoside	19.5	11
Episyringaresinol-4''-O-β-D-glucopyranoside	19.6	11
Acanthoside B	20	11
Erythro-**1-(4**-O-β-d-glucopyranosyl-**3,5**-dimethoxyphenyl**)-2**-syringaresinoxyl-propane-**1,3**-diol	27.7	15
Longifloroside B	20.4	11.5
Oleane Triterpenoid Saponins		
Akebonic acid	14.6	8.5
Boussingoside A1	18.8	10.5
Boussingoside A2	22.7	12.5
3β-Hydroxy-23-oxo-30-noroleana-12,20(29)-diene-28-oic acid 3-O-β-d-glucuronopyranosyl-28-O-β-d-glucopyranoside	23.6	13
30-Norhederagenin 3-O-β-d-glucuronopyranosyl-28-O-β-d-glucopyranoside	24	13
3-O-**[**β-D-Glucuronopyranosyl-**6'**-O-methyl ester**]-30**-norolean-**12,20(29)**-dien-**28**-O-**[**β-D-glucopyranosyl**]**ester	23.8	13
Salieuropaea A	28.4	15
Salbige A	20	11
Salbige B	20	11
Gypsogenin	14.6	8.5
Gypsogenin 3-O-β-d-glucuronopyranoside	19	10.5
Gypsogenin 3-O-β-d-glucuronopyranosyl-28-O-β-d-glucopyranoside	22.3	12
Oleanolic acid	13.6	8
Oleanolic acid 3-O-β-D-glucopyranoside	18.4	10
Oleanolic acid 28-O-β-D-glucoside	16.8	9.5
Calenduloside E (oleanolicacid 3-O-β-D-glucuronopyranoside)	19	10.5
Calenduloside E 6'-methyl ester (oleanolic acid-3-O-6'-O-methyl-β-D-glucuronopyranoside)	18.8	10.5
Chikusetsusaponin Iva	23.6	13
Chikusetsusaponin Iva methyl ester	21.7	12
3β**,29**-Dihydroxy-olean-**12**-en-**28**-oic acid **28**-O-β-D-glucopyranosyl ester	16.7	9.5
Zygophyloside K	22.2	12
Others		
Pyrogallol	5.7	4
Vanillin	7.6	5
Uracil	5.2	3.5
Caffeine	7.8	5
Scopoletin	8	5
Pentadecyl ferulate	30	16
Dibutyl phthalate	13	7.5
Dioctyl phthalate	15	8.5
Icariside B2	15.5	9
Ellagic acid	10.4	6.5
Pheophorbide A	15	8.5
(13^2S)-Hydroxy-pheophorbide A	14.7	8.5
(13^2S)-Hydro-pheophorbide-lactone A	16	9

3.1.3. Docking simulation and data output

Molecular docking was performed using the embedded macro in YASARA, AutoDock Vina, using the default parameters. The program calculates the binding energy of the possible receptor-ligand complexes taking into consideration steric, hydrophobic and hydrogen bonding interactions (Trott and Olson 2010). During the simulation, the program performs 25 docking runs, which produce 25 possible ligand-receptor binding conformations. Some of the different conformations are arranged around the same hotspot. Amongst these structures, the one that has the lowest energy is saved as a cluster. The number of clusters differs depending on the simulation. The criterion automatically set by VINA to differentiate between

two clusters is for the root-mean square deviation (RMSD) between the different binding positions of a ligand to be greater than 5 Å. RMSD is defined in the YASARA manual as

$$RMSD = \sqrt{\frac{\sum_{i=1}^{n} R_i * R_i}{n}} \quad (1)$$

where n is the number of corresponding atoms of the two selections for which the RMSD is being calculated and R is the vector linking them.

After each simulation, the program generates a report documenting the output data for each run and cluster. This includes the binding energy (in kcal/mol), the dissociation constant (in pM) and the contacting residues for each case (including hydrogen bond, hydrophobic, pi-pi, cation-pi and ionic interactions). The binding energy is calculated by subtracting the energy of the ligand-receptor complex in the bound state from the energy when the ligand is at an infinite distance from the receptor and is given as a positive number. Dissociation constant is calculated through the binding energy and the contacting residues listed are the residues of the receptor that have a distance of less than 4 Å from the ligand. Higher binding energies and lower dissociation constants are indicators of better binding. Apart from this data, VINA produces a YASARA object file for each cluster, in which the exact conformation of the binding complex for each cluster can be visualized and edited. YASARA allows the selection and depiction of individual interactions between the protein and the ligand, including hydrogen bonds, hydrophobic, pi-pi, cation-pi and ionic interactions. An example of the simulation output is presented in Figure 28.

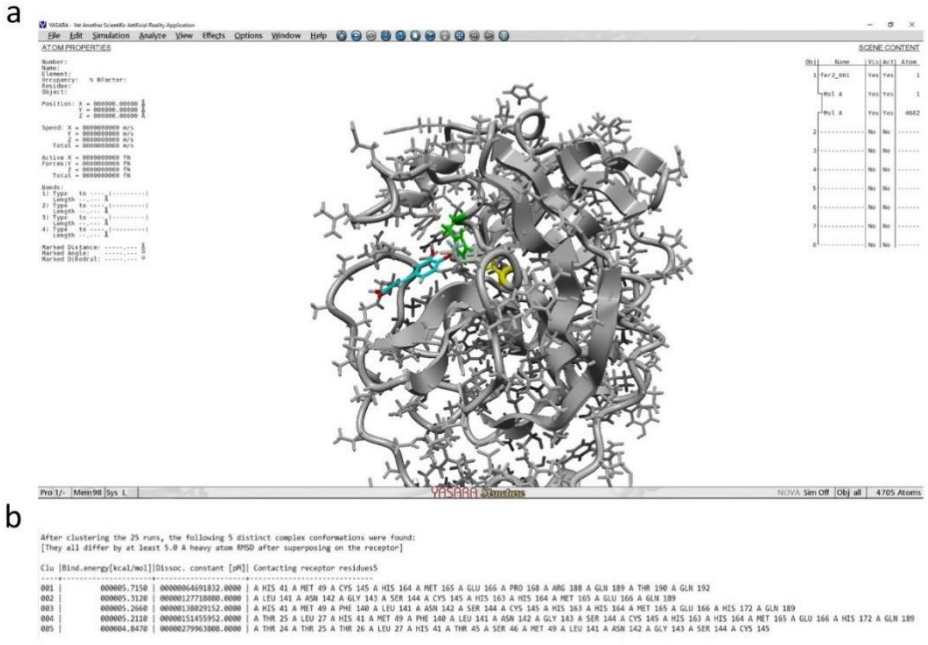


Figure 28: Molecular docking output as produced by YASARA Structure. (a): First cluster (.yob file) generated for the docking of ferulic acid (PDB: 6LU7, His 41 in green, Cys 145 in yellow, ligand in blue); (b): Information on the occurring clusters as given in the data output file.

Regarding selecting the best cluster, in the case of inhibitor docking, the cluster in which the ligand had the lowest RMSD from the ligand in the original crystal structure was saved as best, whereas in the case of *Salicornia* ligands docking, the one that combined the highest binding

energy with the most and most important interactions with the catalytic site and higher similarity with respective published studies, if available, was selected. In the cases where the first cluster was not clearly the preferrable one, all the clusters above the one selected were mentioned.

3.2. Evaluation of SARS-CoV-2 M^{pro} inhibition in vitro

3.2.1. Chemicals and instruments

Chemicals used in the study are:

- **Phenolic acids:** Ferulic acid, rosmarinic acid, gallic acid (Monohydrate, >98%, Sigma Aldrich), quinic acid (98%), chlorogenic acid, 3,4-dicaffeoylquinic acid (≥90%, LC/MS-ELSD), 3,5-dicaffeoylquinic acid (≥95%, LC/MS-ELSD). All phenolic acids and dimethyl sulfoxide (DMSO) were purchased from Sigma-Aldrich (USA).
- **Flavonoids and flavanones:** Myricetin (≥96.0%, crystalline), kaempferol (≥97.0%, HPLC), isorhamnetin-3-O-β-D- glucopyranoside (≥98%), hesperetin (≥95%), hesperidin (≥80%), apigenin, apigenin 7-glucoside (analytical standard), chrysin (97%), catechin, galangin, rutin, acacetin were purchased by AdooQ® Bioscience (USA). Pelargonidin rutinoside was purchased by Carbosynth Ltd. (UK). Quercetin (≥95%,HPLC), isoquercetin, isorhamnetin (≥95.0%, HPLC) and isorhamnetin-3-O-rutinoside (phyproof® Reference Substance) were purchased by Sigma-Aldrich (USA).
- ***Salicornia* extract:** The extract was kindly provided by the company Celabor (Belgium).
- **Commercial 3CL Protease, MBP-tagged (SARS-CoV-2) Assay kit** (BPS Bioscience, USA) containing:

➢ Recombinant 3CL Protease with MBP-tage stock solution (2.1 mg/mL)

➢ 3CL Protease substrate stock solution (10 mM)

➢ 3CL Protease assay buffer

➢ 0.5 M DTT stock solution

➢ Inhibitor GC376

Instrumentation and materials:

- Spectramax microplate fluorimeter (Molecular Devices)
- Vortex mixer (VWR, Sweden)
- Plate incubator (VWR, Sweden)
- Sample tubes (Eppendorf, Germany)
- Laboratory micropipettes and tips
- 384-well black microplate, low binding microtiter plate (Greiner Bio-One, Austria)
- Plate sealing film (Greiner Bio-One, Austria)

3.2.2. In vitro assay for evaluation of M^{pro} inhibition

3.2.2.1. Preparation of solutions

Representative compounds of the screened categories and the compounds with the best hits, focusing on phenolic acids and flavonoids, were further tested for their inhibitory activity against SARS-CoV-2 M^{pro} *in vitro*, using an enzyme inhibition assay. For this purpose, the 3CL protease MBP-tagged Assay Kit by BPS Bioscience (San Diego, CA, USA) was used and the instructions included in the kit were followed. Shortly, the assay buffer was prepared adding DTT to the 3CL Protease assay buffer, achieving a final concentration of 1 mM. 3CL protease was diluted in assay buffer in such manner to achieve a final concentration of 10 µg/mL (150 ng per reaction) in the assay. Serial dilutions of tested compounds were prepared in DMSO, being 100 times more concentrated than the desired final concentration in the assay. Then a 20-fold dilution was made in assay buffer. In that way, the final concentration of DMSO in the assay would not exceed 1%. Serial dilutions of *Salicornia* extract and GC376 were prepared in assay buffer at concentrations 5-fold higher than the desired final concentration in the assay. The 3CL protease substrate solution was prepared by dilution with assay buffer to make a 250 µM solution, would translate to 50 µM concentration in the assay. All assay solutions were prepared fresh.

3.2.2.2. Assay conditions and inhibition detection

To prepare a test sample, 15 µL of 3CL protease was added to the wells of the microplate. Then, 5 µL of appropriately diluted inhibitor was added following subsequent incubation for 30 min at room temperature with slow shaking. The reaction was started by adding 5 µL of the 3CL protease substrate solution to each well. All reactions were carried out in duplicate. The microplate was sealed and incubated at room temperature and slow shaking for 4-24 h. The fluorescence intensity was measure at 360 nm (excitation) and 460 (emission). The test sample, positive control, inhibitor control and blank sample preparation is presented in Table 5.

Table 5: Sample preparation for the enzyme inhibition assay

Component	Concentration in reaction	Positive control	Test sample	Inhibitor control	Blank
3CL protease	10 µg/mL	15 µL	15 µL	15 µL	-
Test inhibitor	0.05-5000 µM (pure compounds) 0.5-10000 µg/mL (extract)	-	5 µL	-	-
GC376	0.005-50 µM	-	-	5 µL	-
Substrate	50 µM	5 µL	5 µL	5 µL	5 µL
Inhibitor buffer (no inhibitor, DMSO containing or not)		5 µL	-	-	5 µL
Assay buffer (with DTT)		-	-	-	15 µL

3.2.2.3. Calculations

Based on the acquired data, the residual activity of M^{pro} and the half maximal inhibitory concentration (IC_{50}) were calculated:

$$Relative\ activity\ (\%) = \frac{Abs_{test\ sample} - Abs_{blank}}{Abs_{positive\ control} - Abs_{blank}} \cdot 100 \quad (2)$$

where $Abs_{test\ sample}$ is the average of the two fluorescence intensity values measured for each concentration sample, $Abs_{positive\ control}$ is the average of the two fluorescence intensity values measured for the two duplicates for the positive control sample (containing only enzyme and substrate) and Abs_{blank} is the average of the two fluorescence intensity values measured by the photometer for the two duplicate blank samples, containing DMSO in the first case and only assay buffer in the second.

For the IC_{50} calculation, the relative activity (%) was plotted against the test compound concentration. A linear curve was fitted, using the R^2 value provided by Microsoft Excel as a measure to evaluate the fitting of the curve, since it expresses the distance between the actual and the fitted values. R^2 values over 0.85 were considered acceptable. The equation provided by Microsoft Excel for the fitted curve was in the form of

$$Relative\ activity\ (\%) = a\ \cdot C\ (\mu M)\ +\ b \quad (3)$$

IC_{50} was calculated by setting that the M^{pro} activity should be 50% of the activity of the positive control.

$$IC_{50} = \frac{(50 - b)}{a} \quad (4)$$

Standard deviation for the fluorescence intenstity values was calculated using the STDEV.P function of Microsoft Excel for the two values produced for each concentration sample. The percent standard deviation was calculated by dividing the standard deviation value with the average of the two measurements and multiplying the result by 100.

4. RESULTS

4.1. Identification of key residues implicated with M^{pro} inhibition based on co-crystallization data

The interactions between M^{pro} and the already confirmed and co-crystallized inhibitors can provide valuable information as to how inhibition occurs and which active site residues stand out when it comes to inhibitor binding. The data collected from the respective studies, as well as from visualization of the PDB structures in YASARA, are presented in Table 6.

The data indicates that Glu 166 is a major interacting residue, as it forms at least one hydrogen bond with the vast majority of the inhibitors studied. Its central position in the active site cavity, between S1 and S4 subsites, makes it easily accessible and available for the formation of hydrogen bonds. Further more, residues Gly 143, Cys 145, His 163 and His 164 are also very common, interacting with more than half of the inhibitors. Gly 143 is located in S1' subsite, while His 163 in the centre of S1 and His 164 is adjacent to the catalytic dyad, in the centre of the active site cavity, indicating that interactions are more often with S1 and S1' subsite residues. It is also observed that Gly 143, Cys 145 and His 163 interact with both covalent and non-covalent inhibitors, while His 164 participates in interactions only with the covalent inhibitors. Further more, His 41, Phe 140 and Gln 189 stand out as interacting residues for the covalent inhibitors. Phe 140 is also located in the S1 subsite, while Gln 189 is the only residue on the left sife of the molecule, over S4 subsite. Overall, there is an indication that inhibitors form hydrogen bonds with residues from S1 and S1' subsites. This could be expected, as it is known that the natural substrate of the enzyme binds to the active site in such a way so that the scissile bond is located between S1 and S1' subsites, therefore their role in catalysis is crucial.

4.2. Docking simulations of known inhibitors onto the M^{pro} crystal structure

Originally, the known inhibitors, that have been co-crystallized in complex with SARS-CoV-2 M^{pro} and for which these crystal structures have been deposited in the PDB, were docked, in order to establish the method for the molecular docking of the bioactive compounds from *Salicornia*. In that way, we could examine and verify the reliability and reproducibility of the results of *in silico* simulations. Most of the crystal structures were accompanied by published work that analyzes their interactions with the protease. However, there were some structures that were deposited in PDB as a result of work that has yet to be published. Even though in these cases the results could not be cross-referenced with a respective publication, the docking of every inhibitor was useful in highlighting binding patterns and residues that play an important role in ligand-receptor interactions, apart from the already known catalytic residues. Moreover, docking of inhibitors onto the respective co-stryctallized M^{pro} structure, offered an insight into the reliability of *in silico* simulations and how well they can predict the naturally occurring ligand-receptor conformation.

The most widely-accepted and analyzed inhibitor of M^{pro} in literature is N3. It is often used as a positive control, to provide some reference values with which the binding energy and interactions of an unknown molecule with M^{pro} can be compared. The binding energy for N3 calculated in this work by Vina is -8.26 kcal/mol, whereas Das et al. (2020) report it to be -7.7 kcal/mol and Ahmed et al. (2020) -7.5 kcal/mol.

Table 6: Hydrogen bond interactions of ligands with M^{pro} residues

Inhibitors	PDB ID	Number of Hydrogen bonds	Contacting residues										
N3	6LU7	7				Gly 143			His 163	His 164	Glu 166[1]	Gln 189	Thr 190
[2]		7				Gly 143			His 163	His 164	Glu 166[1]	Gln 189	Thr 190
11a	6LZE	7		Phe 140				Cys 145	His 163	His 164	Glu 166[3]		
		6						Cys 145	His 163	His 164	Glu 166[3]		
11b	6M0K	8		Phe 140				Cys 145	His 163	His 164	Glu 166[3]	Gln 189	
		6						Cys 145	His 163	His 164	Glu 166[3]		
5h	7JKV	6				Gly 143		Cys 145		His 164	Glu 166[1]	Gln 189	
		6				Gly 143		Cys 145		His 164	Glu 166[1]	Gln 189	
GC376	7D1M	7		Phe 140						His 164	Glu 166[1]		
		7	His 41	Phe 140						His 164	Glu 166	Gln 189[1]	
GC373	6WTK	6				Gly 143	Ser 144	Cys 145	His 163		Glu 166[1]		
		5						Cys 145	His 163	His 164	Glu 166	Gln 189	
Narlaprevir	7JYC	7	His 41		Asn 142	Gly 143				His 164	Glu 166[3]		
		7	His 41			Gly 143		Cys 145		His 164	Glu 166[3]		
Telaprevir	7K6D	7	His 41			Gly 143	Ser 144			His 164	Glu 166[1]	Gln 189	
		7	His 41			Gly 143		Cys 145		His 164	Glu 166[1]	Gln 189	
x2754	5RHF	-											
		2				Gly 143		Cys 145					
x2705	5RH7	-											
		3		Met 49					His 163		Glu 166		
Myricetin	7B3E	-											
		2						Cys 145			Glu 166		
Mg-132	7BE7	-											
		5						Cys 145		His 164	Glu 166[1]	Gln 189	
MI-23	7D3I	5		Phe 140		Gly 143		Cys 145	His 163	His 164	Glu 166		
		4						Cys 145	His 163	His 164	Glu 166		
Carmofur	7BUY	2				Gly 143		Cys 145					
		2				Gly 143		Cys 145					
Boceprevir	7C6S	7	His 41			Gly 143		Cys 145		His 164	Glu 166[3]		
		7	His 41			Gly 143		Cys 145		His 164	Glu 166[3]		
Calpeptin	7AKU	2								His 164	Glu 166		
		2								His 164	Glu 166		
13b	6Y2G	8	His 41	Phe 140		Gly 143		Cys 145	His 163		Glu 166[3]		
		7	His 41	Phe 140				Cys 145	His 163	His 164	Glu 166[1]		
Non-covalent													
Compound 5	7L11	-											
		3						Cys 145	His 163		Glu 166		
Compound 26	7L14	-											
		3						Cys 145	His 163		Glu 166		
x0397	5RGI	-											
		3				Gly 143		Cys 145	His 163				

x77	6W63	-				
		2		Gly 143		Glu 166
Mcule-5948770040	7LTJ	-				
		3		Gly 143	His 163	Glu 166
ML 188	7L0D	-				
		2		Gly 143	His 163	
MUT056399	7AP6	2	Phe 140		His 163	
		2	Phe 140		His 163	
x0104	5R7Z	-				
		1				Glu 166
x0161	5R80	-				
		1				Glu 166

[1]: The ligand forms two hydrogen bonds with this residues; [2]: The first row for each inhibitor corresponds to the interactions described in literature, whereas the second refers to the interactions shown in the crystal structures of the protein-inhibitor complex; [3]: The lignad forms three hydrogen bonds with this residue.

This difference could be attributed to the different pretreatment of the receptor and the ligand structures before the simulation. Superposing of the two structures (the co-crystallized and the docked complex) reveals that the conformation of N3 in the binding site is very similar in the two cases. The major differences can be spotted in the P1' part of N3, which blocks the S1' subsite in the co-crystallized protein-ligand complex, whereas it seems to be rotated outwards in the simulated complex. Moreover, the pentacyclic ring on the other end of the molecule shows a slight divergence, but in both cases the inhibitor takes up the space of the S4 subsite. The total RMSD of the two conformations of N3 (Figure 29, left) is 3.16 Å. In the docking output, the interaction of N3 with Cys 145 is a 2 Å hydrogen bond, formed between pentacyclic ring of N3 and the hydrogen attached to the sulfur atom of the residue whereas in reality, as determined by the co-crystallization data, it is a 1.8 Å covalent bond between the sulfur atoms of Cys 145 and the Cβ atom of the vinyl group of the inhibitor. Although a different part of the inhibitor is bound to the protease, it results in a very similar geometry of the molecule. The additional hydrogen bonds with Glu 166 and Gln 189, as well as hydrophobic interactions with Met 49 and pi-pi interactions with His 41 further stabilize the molecule.

GC376 is another broad spectrum inhibitor, that has recorded activity against the main protease of various coronaviruses (Y. Hu et al. 2021). It is also being used as a positive control in the enzymatic inhibition assay used later in this work. In the present simulation, the binding energy calculated for GC376 was -7.798 kcal/mol, which is higher than the majority of the inhibitors. Additionally, the hydrogen bonds calculated by YASARA for the molecule are more than any other of the docked inhibitors. They are formed with residues His 41, Phe 140, His 163, Glu 166, and Gln 189, the majority of which belong to S1 subsite. The docking output is remarkably similar to the co-crystallized structure, as seen in Figure 30, with the ligands in the two conformations having an RMSD as low as 1.0 Å. In addition, another worked including docking of GC376 to PDB 6LU7 with AutoDock VINA reports a binding energy of -8.1 kcal/mol, which is very close to the result of the present study (Rakib et al. 2021).

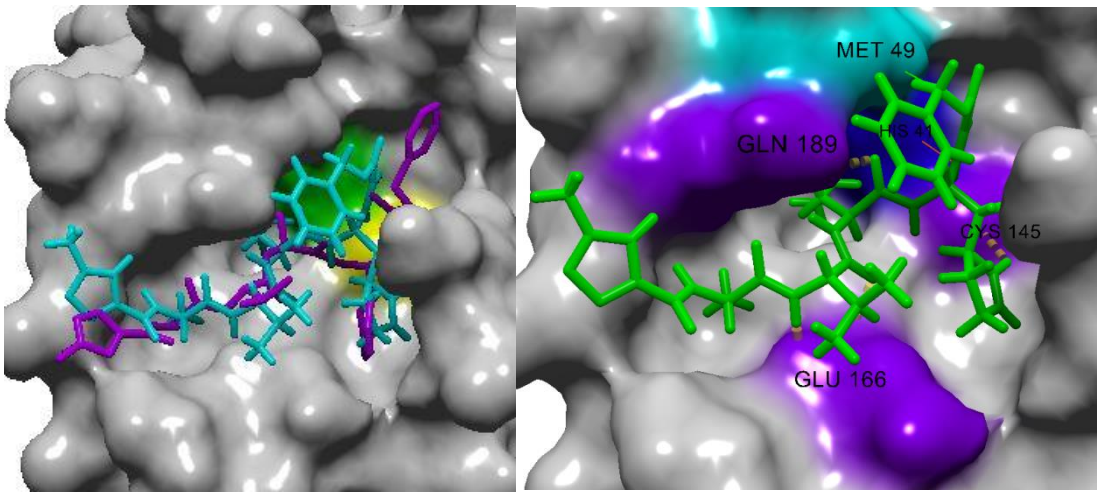


Figure 29: Left: Superimposed structures of N3 in complex with M^{pro} in the PDB co-crystallized structure (purple) and as resulted from the docking simulation (cyan). Catalytic residues are shown in green (His 41) and yellow (Cys 145). Right: Detailed interactions of N3 with M^{pro}. Residues with which N3 forms a hydrogen bond are colored in purple, residues that interact hydrophobically in cyan and residues that form Pi-Pi interactions in blue.

As seen in Figures 30 and 31 and Table 7, the results produced by AutoDock Vina are in accordance with the actual crystal structures for the majority of the rest of the inhibitors. The accuracy of the result varies, as RMSD ranges from 0.318 to 6.65 Å and has an average value

of 2.91 Å. Also, from the comparative data presented in Table 7, it is noticed that the hydrogen bonds given as a docking output are less that the ones in the co-crystallized structure. This is the case not only for N3, but for all the other inhibitors docked. This could lead to the conclusion that the simulation produces similar binding modes to the ones observed in reality, but does not reproduce the exact same interactions. The program does not show potential weak hydrogen bonds (above certain Å), if distances are slightly altered and become longer, less hydrogen bonds will be predicted. Nevertheless, this does not negate the fact that the docking simulation can be a good indicator of the binding affinity of a ligand to the protein. The important is the interactions with catalytic dyad and stabilization of the ligand with interactions with the rest of the residues in the active site. Lastly, an observation that can be made is that the similarity of the docking results to the co-crystallized structures is greater in the case of the non-covalent inhibitors, both in terms of RMSD values, as well as interaction correspondence. In the case of the non-covalent inhibitors, the average RMSD value is 2.1 Å as opposed to the 3.4 Å value for the covalent inhibitors, whereas the contacting residues taking part in hydrogen bonds are 64.6% identical for the non-covalent inhibitors while only 35.6% for the covalent ones. This difference could be due to the fact that Vina cannot reproduce covalent bonds. Therefore, the stabilization of the binding mode that the covalent bond causes and the 3D conformation the inhibitor acquires, when one part of it is covalently bound to the protease, cannot be entirely resembled by this simulator.

Docking results show that the greater in length or bulkier inhibitors block the most M^{pro} subsites, as expected. More specifically, N3, 5h, Mg-132, 13b, x2705, compounds 5 and 26 and X77 block access to all subsites. Other large molecules, such as 11a, 11b, GC376, GC373, bind to the cavities forming S1, S2 and S4 subsites. Although S1' subsite seems accessible in these cases, catalytic residues His 41 and Cys 145 are blocked. Also, Narlaprevir, Telaprevir and Boceprevir have a similar binding conformation, impeding binding to all M^{pro} subsites apart from S1, whereas ML188, Calpeptin and MI-23 leave the S4 subsite uncovered. Lastly, the smaller inhibitors obstruct binding to fewer subsites: x2754 and Carmofur block the S1' subsite, Myricetin, Mcule-5948770040 and MUT056399 the S1 and S2 subsites, x0397 the S1 and S1' subsites and x0104 and x0161 the S2 and S4 subsites.

The study of the detailed interactions of the inhibitors reveals some patterns in the binding of compounds with similar structure or the same functional groups. The major observations are summarized below:

- Inhibitors 11a and 11b, having almost identical structures, form the exact same hydrogen bonds. Two of them are formed between the =O and -NH groups of the pyrrolidone ring. The docking results do not show all the interactions present in the crystal structures. There, it is also visible that the indole moiety that is neighboring with a carbonyl group creates the same two hydrogen bonds with Glu 166. The same bonds are present in inhibitor 5h that has the same indole-carbonyl sequence.

- The carbonyls adjacent to -NH groups, both in linear and cyclic carbon chains, are very susceptible to the formation of hydrogen bonds. Many hydrogen bonds are also formed between the hydrogen atom in the -NH group of the same pattern, and carbonyl groups of M^{pro} residues. The first scenario is observed in both the aforementioned hydrogen bonds of 11a and 11b, in the ones of MG-132 with Glu 166 and Gly 143 or MI-23 with Glu 166 and His 163. The second scenario can be confirmed in the bonds between 11a and 11b and Phe 140, 13b and Leu 141 and Calpeptin and His 164.

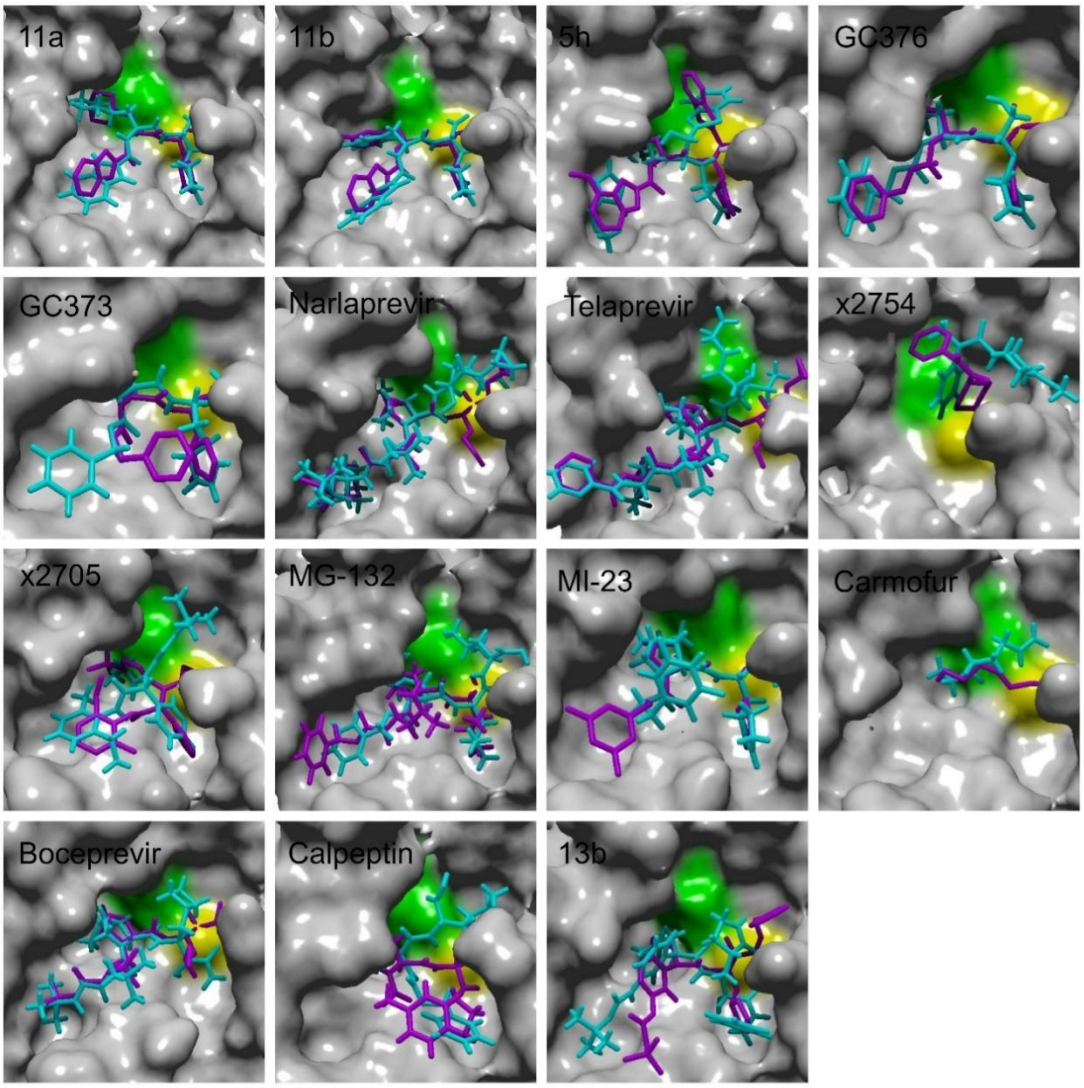


Figure 30: Superimposed structures of covalent inhibitors in complex with M^{pro} in the PDB co-crystallization structure (purple) and as resulted from the docking simulation (cyan). Catalytic residues are shown in green (His 41) and yellow (Cys 145).

- Narlaprevir, telaprevir and boceprevir also have very closely related structures, and almost identical bonds. More specifically, they bind to Glu 166 with more than one hydrogen bonds, which also correspond to the aforementioned category of hydrogen bonds between -NH and =O groups. In addition, in the co-crystallized structures of all three inhibitors, the carbonyl next to the outer -NH group of the molecule forms two hydrogen bonds with the -NH groups of Gly 143 and Cys 145.
- Pi-pi stalking is observed between the pentacyclic or hexacyclic rings of inhibitors and the imidazole moieties present in residues His 41 or His 163.
- The benzene ring that is part of the indole moiety present in inhibitors 11a, 11b and 5h forms hydrophobic interactions in all three cases, with the pyrrolidine of residues Pro 168 in the first two cases and with Ala 191 in the third.

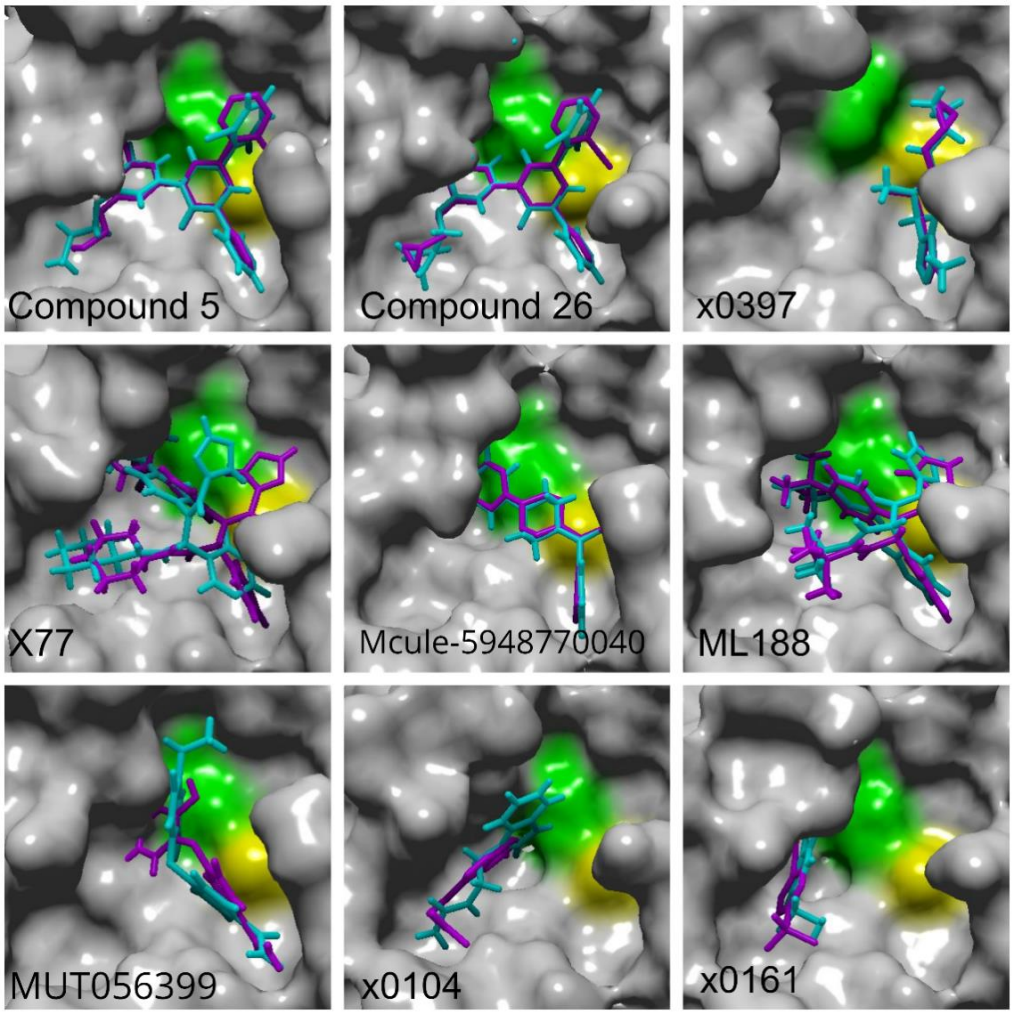


Figure 31: Superimposed structures of non-covalent inhibitors in complex with M^{pro} in the PDB co-crystallization structure (purple) and as resulted from the docking simulation (cyan). Catalytic residues are shown in green (His 41) and yellow (Cys 145).

The hydrogen bonding interactions that are visible through YASARA for the most favorable ligand-receptor structures, which resulted from the molecular docking, show that there are some residues that play in important role in inhibitor binding, apart from the catalytic residues His 41 and Cys 145. More specifically, the residue that formed the most hydrogen bonds with the ligands was Glu 166, present in 15 out of the 26 cases. As seen in Table 9, 25 out of the 26 ligands interact with Glu 166 in some way, even when not forming a hydrogen bond. The next most susceptible residue to hydrogen bonding appears to be Gly 143, forming a bond with 8 out of the 26 ligands and overall interacting with 80% of them. Gln 189, His 163, Cys 145 and Phe 140 form hydrogen bonds with a total of 6, 5, 4 and 3 inhibitors, respectively, whereas His 41, Leu 141, Asn 142, His 164, Glu 167 appear singly. These results partly confirm the study of (Stoddard et al. 2020) in the case of Glu 166 which is referred as a hydrogen bond hotspot, but the rest of the hydrogen bonds that came as an output from the docking simulation are not frequent enough to establish a pattern. This observations aligns with the results from Stoddart et. al. (2020).

Table 7: Docking results and co-crystallization structure data for SARS-CoV-2 M^{pro} inhibitors

	Docking results									Co-crystallization structure data			
Inhibitor	**PDB ID**	**Total clusters**	**Best cluster[1]**	**RMSD[2]**	**Binding energy (kcal/mol)**	**No of H-bonds**	**H-bond[3]**	**Hydrophobic[3]**	**Pi-Pi [3]**	**No of H-bonds**	**H-bond[3]**	**Hydrophobic[3]**	**Pi-Pi[3]**
N3	6LU7	12	1	3.16	-8.260	4	CYS 145, GLU 166 (x2), GLN 189	MET 49	HIS 41	7	GLY 143, HIS 163, HIS 164, GLU 166 (x2), GLN 189, THR 190	PRO 168	-
Covalent													
11a	6LZE	8	1	1.3466	-8.128	4	PHE 140, GLY 143, HIS 163, GLU 166	PRO 168	-	7	PHE 140, CYS 145, HIS 163, HIS 164, GLU 166 (x3)	MET 49	-
11b	6M0K	8	4	1.8313	-7.663	4	PHE 140, GLY 143, HIS 163, GLU 166	PRO 168	HIS 41	8	PHE 140, CYS 145, HIS 163, HIS 164, GLU 166 (x3), GLN 189	MET 49	HIS 41
5h	7JKV	12	2	2.6574	-7.530	2	GLU 166, GLN 189	ALA 191	HIS 41	6	GLY 143, CYS 145, HIS 164, GLU 166 (x2), GLN 189	MET 49	HIS 41
GC376	7D1M	9	1	1.0026	-7.798	5	HIS 41, PHE 140, HIS 163, GLU 166, GLN 189	ASP 187	-	7	HIS 41, PHE 140, HIS 164, GLU 166, GLN 189 (x2)	MET 49	-
GC373	6WTK	7	1	3.0919	-7.350	2	ASN 142, GLU 166	HIS 41	-	5	CYS 145, HIS 163, HIS 164, GLU 166 (x3)	MET 165	-
Narlaprevir	7JYC	6	1	2.5916	-8.571	3	GLU 166 (x3)	ASN 142	-	7	HIS 41, GLY 143, CYS 145, HIS 164, GLU 166 (x3)	LEU 167	-
Telaprevir	7K6D	5	2	3.2027	-8.005	3	GLU 166 (x2), GLN 189	LEU 27	-	7	HIS 41, GLY 143, CYS 145, HIS 164, GLU 166 (x2), GLN 189	GLY 143	-
x2754	5RHF	3	3	4.7475	-5.891	0	-	ASN 142	HIS 41	2	GLY 143, CYS 145	MET 49	-
x2705	5RH7	5	2	4.3271	-7.235	0	-	GLN 189	HIS 41	3	MET 49, HIS 163, GLU 166	PHE 140	HIS 163
Mg-132	7BE7	7	1	6.6545	-7.073	2	GLY 143, GLU 166	MET 165	-	5	CYS 145, HIS 164, GLU 166 (x2), GLN 189	PRO 168	-
MI-23	7D3I	7	1	2.9002	-7.850	2	HIS 163, GLU 166	HIS 41	-	4	CYS 145, HIS 163, HIS 164, GLU 166	HIS 41	-
Carmofur	7BUY	7	7	2.0852	-2.930	0	-	MET 49	-	2	CYS 143, CYS 145	MET 49	-
Boceprevir	7C6S	7	1	2.4926	-8.063	3	GLY 143, GLU 166 (x2)	MET 165	-	7	HIS 41, GLY 143, CYS 145, HIS 164, GLU 166 (x3)	MET 165	-
Calpeptin	7AKU	6	2	4.5275	-6.526	2	HIS 164, GLU 166	GLU 166	-	2	HIS 164, GLU 166	GLN 189	-
13b	6Y2G	7	3	4.6032	-6.924	2	LEU 141, GLY 143	GLN 189	-	7	HIS 41, PHE 140, CYS 145, HIS 163, HIS 164, GLU 166 (x2)	GLN 189	-
Non-covalent													
Compound 5	7L11	10	1	1.6946	-9.376	1	GLU 166	MET 49	HIS 163	3	CYS 145, HIS 163, GLU 166	MET 165	HIS 163
Compound 26	7L14	5	1	1.6844	-9.464	1	GLU 166	GLN 189	HIS 163	3	CYS 145, HIS 163, GLU 166	THR 25	HIS 153
x0397	5RGI	4	1	0.3183	-6.432	3	GLY 143, CYS 145, HIS 163	HIS 163	HIS 163	3	GLY 143, CYS 145, HIS 163	HIS 163	HIS 172
x77	6W63	7	2	2.5763	-8.032	1	GLU 166	PRO 168	HIS 41	2	GLY 143, HIS 163, GLU 166	PHE 140	HIS 41
Mcule-5948770040	7LTJ	5	1	0.4565	-9.329	2	GLY 143, GLU 166	HIS 41	HIS 41	3	GLY 143, HIS 163, GLU 166	HIS 41	HIS 163

ML 188	7L0D	8	2	2.7112	-6.999	1	GLY 143	LEU 27	HIS 163	2	GLY 143, HIS 163	PHE 140	HIS 163
MUT056399	7AP6	3	2	3.6902	-6.449	2	PHE 140, HIS 163	THR 25	HIS 41	2	PHE 140, HIS 163	HIS 41	HIS 41
x0104	5R7Z	3	2	4.2392	-5.608	0		MET 49	HIS 41	1	GLU 166	GLN 189	HIS 41
x0161	5R80	5	3	1.5396	-5.399	1	GLU 166	MET 49	-	1	GLU 166	GLN 189	-

[1]: The criterion for selecting the best cluster is mainly the RMSD of the ligand in the complex resulted from the docking from the ligand in the co-crystallization structure. For cases where RMSDs of various cluster are very close, the binding energy and orientation of the molecule was taken into consideration; [2]: RMSD between the ligands in the docked and co-crystallized ligand-protease complex; [3]: Residues with which the ligand forms a hydrogen bond, hydrophobic or pi-pi interaction respectively, as calculated by YASARA structure.

Table 8: Total contacting residues as calculated from docking of SARS-COV-2 M^{pro} known inhibitors

N3	THR 25		LEU 27	HIS 41			MET 49		TYR 54	PHE 140	LEU 141	ASN 142	GLY 143	SER 144	CYS 145	HIS 163	HIS 164	MET 165	GLU 166	LEU 167	PRO 168	HIS 172		ASP 187	ARG 188	GLN 189	THR 190	ALA 191	GLN 192
Covalent																													
11a				HIS 41	CYS 44		MET 49		TYR 54	PHE 140	LEU 141	ASN 142	GLY 143	SER 144	CYS 145	HIS 163	HIS 164	MET 165	GLU 166		PRO 168	HIS 172		ASP 187	ARG 188	GLN 189	THR 190		
11b				HIS 41			MET 49			PHE 140	LEU 141	ASN 142	GLY 143	SER 144	CYS 145	HIS 163	HIS 164	MET 165	GLU 166	LEU 167	PRO 168	HIS 172	VAL 186	ASP 187	ARG 188	GLN 189			GLN 192
5h	THR 25	THR 26	LEU 27	HIS 41			MET 49		TYR 54	PHE 140	LEU 141	ASN 142	GLY 143	SER 144	CYS 145	HIS 163	HIS 164	MET 165	GLU 166	LEU 167	PRO 168	HIS 172	VAL 186	ASP 187	ARG 188	GLN 189	THR 190	ALA 191	
GC376				HIS 41			MET 49		TYR 54	PHE 140	LEU 141	ASN 142	GLY 143	SER 144	CYS 145	HIS 163	HIS 164	MET 165	GLU 166	LEU 167	PRO 168	HIS 172		ASP 187	ARG 188	GLN 189	THR 190	ALA 191	GLN 192
GC373			LEU 27	HIS 41			MET 49	PRO 52	TYR 54	PHE 140	LEU 141	ASN 142	GLY 143	SER 144	CYS 145	HIS 163	HIS 164	MET 165	GLU 166	LEU 167	PRO 168	HIS 172		ASP 187	ARG 188	GLN 189	THR 190	ALA 191	
Narlaprevir	THR 25	THR 26	LEU 27	HIS 41			MET 49		TYR 54			ASN 142	GLY 143	SER 144	CYS 145		HIS 164	MET 165	GLU 166	LEU 167	PRO 168		VAL 186	ASP 187	ARG 188	GLN 189	THR 190	ALA 191	GLN 192
Telaprevir	THR 25	THR 26	LEU 27	HIS 41	CYS 44	SER 46	MET 49					ASN 142	GLY 143		CYS 145		HIS 164	MET 165	GLU 166	LEU 167	PRO 168		VAL 186	ASP 187	ARG 188	GLN 189	THR 190	ALA 191	GLN 192
x2754	THR 25	THR 26	LEU 27	HIS 41	CYS 44	SER 46	MET 49					ASN 142	GLY 143		CYS 145			MET 165	GLU 166										
x2705	THR 25	THR 26	LEU 27	HIS 41	CYS 44	SER 46	MET 49			PHE 140	LEU 141	ASN 142	GLY 143	SER 144	CYS 145		HIS 164	MET 165	GLU 166	LEU 167	PRO 168			ASP 187	ARG 188	GLN 189	THR 190		GLN 192
Myricetin				HIS 41			MET 49			PHE 140	LEU 141	ASN 142		SER 144	CYS 145	HIS 163	HIS 164	MET 165	GLU 166				VAL 186	ASP 187	ARG 188	GLN 189			GLN 192
Mg-132	THR 25	THR 26	LEU 27	HIS 41			MET 49			PHE 140	LEU 141	ASN 142	GLY 143	SER 144	CYS 145	HIS 163	HIS 164	MET 165	GLU 166	LEU 167	PRO 168	HIS 172		ASP 187	ARG 188	GLN 189	THR 190	ALA 191	GLN 192
MI-23				HIS 41			MET 49		TYR 54	PHE 140	LEU 141	ASN 142	GLY 143	SER 144	CYS 145	HIS 163	HIS 164	MET 165	GLU 166			HIS 172		ASP 187	ARG 188	GLN 189			
Carmofur	THR 25		LEU 27	HIS 41			MET 49								CYS 145		HIS 164	MET 165						ASP 187	ARG 188				
Boceprevir	THR 25			HIS 41			MET 49		TYR 54	PHE 140	LEU 141	ASN 142	GLY 143	SER 144	CYS 145	HIS 163	HIS 164	MET 165	GLU 166	LEU 167	PRO 168			ASP 187	ARG 188	GLN 189	THR 190	ALA 191	GLN 192
Calpeptin	THR 25	THR 26	LEU 27	HIS 41			MET 49		TYR 54	PHE 140	LEU 141	ASN 142	GLY 143	SER 144	CYS 145	HIS 163	HIS 164	MET 165	GLU 166					ASP 187	ARG 188	GLN 189			
13b				HIS 41			MET 49			PHE 140	LEU 141	ASN 142	GLY 143	SER 144	CYS 145	HIS 163	HIS 164	MET 165	GLU 166	LEU 167	PRO 168	HIS 172	VAL 186	ASP 187	ARG 188	GLN 189	THR 190	ALA 191	GLN 192
Non-covalent																													
Compound 5	THR 25	THR 26	LEU 27	HIS 41			MET 49		TYR 54	PHE 140	LEU 141	ASN 142	GLY 143	SER 144	CYS 145	HIS 163	HIS 164	MET 165	GLU 166	LEU 167	PRO 168	HIS 172		ASP 187	ARG 188	GLN 189	THR 190		GLN 192

Compound 26	THR	THR	LEU	HIS			MET		TYR	PHE	LEU	ASN	GLY	SER	CYS	HIS	HIS	MET	GLU	LEU	PRO	HIS		ASP	ARG	GLN	THR	GLN
	25	26	27	41			49		54	140	141	142	143	144	145	163	164	165	166	167	168	172		187	188	189	190	192
x0397	THR	THR	LEU	HIS						PHE	LEU	ASN	GLY	SER	CYS	HIS	HIS	MET	GLU			HIS						
	25	26	27	41						140	141	142	143	144	145	163	164	165	166			172						
x77				HIS	CYS		MET	PRO	TYR	PHE	LEU	ASN	GLY	SER	CYS	HIS		MET	GLU	LEU	PRO			ASP	ARG	GLN	THR	GLN
				41	44		49	52	54	140	141	142	143	144	145	163		165	166	167	168			187	188	189	190	192
Mcule-5948770040				HIS	CYS		MET	PRO	TYR	PHE	LEU	ASN	GLY	SER	CYS	HIS	HIS	MET	GLU			HIS		ASP	ARG	GLN		
				41	44		49	52	54	140	141	142	143	144	145	163	164	165	166			172		187	188	189		
ML 188	THR	THR	LEU	HIS	CYS		MET	PRO	TYR	PHE	LEU	ASN	GLY	SER	CYS	HIS	HIS	MET	GLU			HIS		ASP	ARG	GLN		
	25	26	27	41	44		49	52	54	140	141	142	143	144	145	163	164	165	166			172		187	188	189		
MUT056399	THR			HIS	CYS	SER	MET			PHE	LEU	ASN		SER	CYS	HIS	HIS	MET	GLU			HIS						
	25			41	44	46	49			140	141	142		144	145	163	164	165	166			172						
x0104				HIS		SER	MET										HIS	MET	GLU	LEU	PRO				ARG	GLN	THR	GLN
				41		46	49										164	165	166	167	168				188	189	190	192
x0161				HIS	CYS		MET	PRO	TYR									MET	GLU	LEU	PRO		VAL	ASP	ARG	GLN		GLN
				41	44		49	52	54									165	166	167	168		186	187	188	189		192

Table 9: Most common contacting residues as determined from the docking results and the co-crystallization structures

Docking results			Co-crytallization structures data		
Residues	**Appearance frequency[1]**	**Percentage in the total of inhibitors**	**Residues**	**Appearance frequency[1]**	**Percentage in the total of inhibitors**
HIS 41	26	100	**HIS 41**	25	96.2
MET 165	26	100	**MET 49**	24	92.3
MET 49	25	96.2	**MET 165**	24	92.3
GLU 166	25	96.2	**MET 166**	24	92.3
CYS 145	24	92.3	**GLY 143**	23	88.5
ARG 188	23	88.4	**SER 144**	23	88.5
HIS 164	23	88.4	**CYS 145**	23	88.5
ASN 142	23	88.4	**HIS 164**	23	88.5
ASP 187	22	84.6	**ASN 142**	22	84.6
GLN 189	22	84.6	**ASP 187**	22	84.6
GLY 143	21	80.8	**ARG 188**	22	84.6
SER 144	21	80.8	**GLN 189**	22	84.6
PHE 140	20	76.9	**LEU 141**	20	76.9
LEU 141	20	76.9	**HIS 163**	20	76.9
HIS 163	19	73.1	**PHE 140**	18	69.2
PRO 168	17	65.4	**HIS 172**	18	69.2
LEU 167	16	61.5	**TYR 54**	16	61.5
THR 190	16	61.5	**LEU 27**	15	57.7
HIS 172	15	57.7	**PRO 168**	15	57.7
THR 25	15	57.7	**THR 25**	13	50
TYR 54	15	57.7	**MET 167**	13	50
LEU 27	14	53.8			
GLN 192	14	53.8			

[1]: Number of inhibitors with which the residues interact, in total of 26 inhibitors docked.

More distinguishable patterns, can be observed from the contacting residues between each ligand and protease structure, which are calculated by YASARA structure based on their distance from the ligand in the binding complex. In Table 8, the residues that are common contacts for more than 50% of the inhibitors studied are presented, both for the simulation results as well as the co-crystallized structures. From data extracted from the docking results, it can be assumed that His 41, Met 165, Met 49, Glu 166, Cys 145, Arg 188, His 164, Asn 142, Asp 187, Gln 189, Gly 143, Ser 144, Phe 140, Leu 141, His 163, Pro 168, Leu 167, Thr 190, His 172, Thr 25, Tyr 54, Leu 27, Gln 192 are key residues, since they appear in more than half of the cases, as seen in Table 9. The respective data from the co-crystallization structures are almost identical, with His 41 emerging as the most common interaction in this case, too, and Met 49, Met 165 and Glu 166 also interacting with over 90% of the inhibitors. These residues are spread across the active site cavity and represent all the subsites, therefore no dominant subsite emerges from this data. In addition, it is observed that although catalytic residues are very common interactions for the inhibitors studied, neither of them form hydrogen bonds. However, the interactions of the inhibitors with the protein show that His 41 is a common residue in pi-pi interactions.

4.3. Molecular docking simulation for the bioactive compounds found in *Salicornia* extracts

After the simulation method was established and more information regarding the key residues involved in binding to the protease became available, the bioactve compounds reported to be present in *Salicornia* species (Table 2) were docked. The compounds screened can be divided in nine categories: Hydroxycinnamic acids and derivatives, hydroxybenzoic acids, caffeoylquinic acids and derivatives, flavonoids and flavanones, chromones, sterols, lignans,

oleanane triterpenoid saponins and other compounds. Overall, the most promising results were yielded for caffeoylquinic acids and their derivatives and flavonoids and flavanones, followed by some encouraging results for specific saponins.

4.3.1. Hydroxycinnamic acids and derivatives

The results show a range of binding energies for hydroxycinnamic acids, that are relatively low compared to both native inhibitor N3 (-8.26 kcal/mol) and GC376 (-7.798 kcal/mol). However, other confirmed inhibitors had lower binding energies, proving that binding energy alone is not sufficient to judge the efficiency of a compound as a potential inhibitor. Cinnamic acid, that has no substitutions it its phenolic ring, exhibits the lowest binding energy (-4.98 kcal/mol), whereas rosmarinic acid, an ester of caffeic acid possessing two phenolic rings, has the highest binding energy (-7.41 kcal/mol), followed by sinapic acid (-6.1 kcal/mol) that has two methoxy substitutions in the phenolic ring. On the whole, it is observed that the bulkier the ligand, the higher the binding energy. Additionally, a pattern is observed in the interactions of hydroxycinnamic acids with the protease in terms of hydrophobic and pi-pi interactions. More specifically, residue Met 165 is a hydrophobic interaction hotspot, whereas the phenolic ring present in the ligands forms pi-pi stalking with the pentacyclic ring of residues His 41 and His 163. The orientation of the molecules in the binding site, presented in Figure 31, shows that hydroxycinnamic acids, with the exception of caffeic acid, bind to S2 subsite. Caffeic acid binds in the space between S1 and S1' subsites, and rosmarinic acid practically blocks the entire active site of the protease. Rosmarinic acid is the one whose orientation seems to inhibit access to the catalytic dyad more effectively.

In the cases of ferulic (1), caffeic (2) and sinapic acid (3), the first clusters have a similar orientation, with the phenolic ring on the right sight and the carbon chain extending behind it (Figure 32). However, for coumaric acid, the first cluster has a different orientation, with the phenolic ring on the left side and the tail extending to the left. This can be justified by the smaller size of the molecule compared to the rest previously mentioned hydroxycinnamic acid, which allows for greater mobility in the active site and perhaps more possible conformations. The fact that the second cluster, which has a small difference in binding energy (-5.188 as opposed to -5.329 kcal/mol) has the same orientation as the other hydroxycinnamic acids leads to the consideration of it as potentially more dominant. The same is the case for cinnamic acid, for which the first cluster has the phenolic ring inserted into the S2 subsite and extends to the left, while the second is orientated into the S2 site in the opposite direction, but both clusters have similar interactions and comparable binding energies (-4.978 and -4.705 respectively). As far as rosmarinic acid is concerned, the first two clusters have the same binding affinity to the active site. Other than the fact that cluster 2 seems to be forming more hydrogen bonds, there is no criterion to choose between the two clusters, since the orientation of the hydroxycinnamoyl group is also different from the dominant one in the previous clusters and can only acids in the case of the second cluster (where the phenolic ring is located into the S2 cavity), it can be compared to the one of the first clusters of p-coumaric and cinnamic. Therefore, both clusters are presented as no conclusion can be drawn for a more likely conformation. The second clusters are not taken into consideration for ferulic, caffeic and sinapic acids because they have considerably lower binding energy compared to the first clusters. The collective data is presented in Table 10.

Screening of natural compounds performed in corresponding studies provides data for the binding energies of hydroxycinnamic acids, but not for their binding mode and orientations in the active site, so as to provide a reference to which the docking poses produced by this

simulation could be compared to. However, binding energies can also be a comparable measure to contribute in validating the results of the present study. In particular, a study on M^{pro} inhibition using 6LU7 as a PDB structure reports a binding energy of -4.7 kcal/mol for ferulic acid, but very different hydrogen bond interactions, with residues Glu 14, Met 17 and Gly 71, indicating binding to a different site of the enzyme (Salman et al. 2020b). On the contrary, other docking studies on ferulic acid calculated binding energies of -4.91 (Kundu Debanjan et al. 2021) and -5.7 kcal/mol (Vicidomini, Roviello, and Roviello 2021). Although both of these studies were performed with 6Y84 as a PDB protein structure and AutoDock and AutoDock Vina respectively, they yield considerably different results, which might be due to different simulation parameters (e.g. simulation cell) and pretreatment of the receptor and ligand. Between the two, the second study reports a result almost identical to the present one. The same work mentions results comparable to the present ones also for caffeic and sinapic acid (-5.7 and -6.1 kcal/mol respectively). Lastly, further studies conducted with the same PDB structure as a receptor and AutoDock Vina as the simulation software also produced results in consonance with this work: Murugesan et al (2021). reported binding energies of -6.1 kcal/mol for sinapic and caffeic acids and -6.0 kcal/mol for cinnamic and coumaric acids, Patil et al. (2020) -5.3 and -5.4 kcal/mol for ferulic and caffeic acid respectively, Mohapatra et al. (2020) -5.5 kcal/mol for caffeic acid and Umar et al. (2021) -5.6 kcal/mol for caffeic acid, supplemented by the description of contacting residues, which are to a great extent common with the ones in this work.

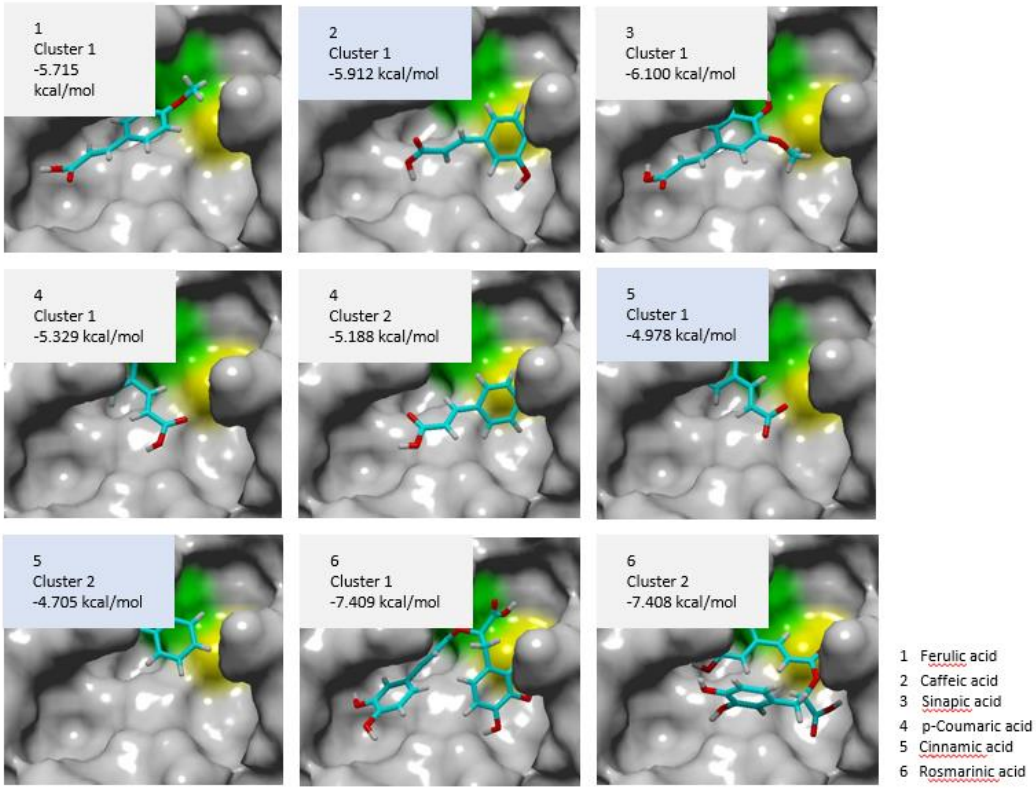


Figure 32: Binding conformations of hydroxycinnamic acids to the active site of SARS-CoV-2 M^{pro}, as resulted from the docking simulation

Table 10: Molecular docking results for hydroxycinnamic acids and derivatives (M^{pro} structure PDB: 6LU7)

Compound		Total clusters	Cluster	Binding energy (kcal/mol)	No of interactions			Total contacting residues
No.	**Name**				**H-bonds**	**Hydrophobic**	**Pi-pi**	
1	Ferulic acid	5	1	-5.715	1 (THR 190)	1 (MET 165)	1 (HIS 41)	HIS 41, MET 49, CYS 145, HIS 164, MET 165, GLU 166, PRO 168, ARG 188, GLN 189, THR 190, GLN 192
2	Caffeic acid	4	1	-5.912	2 (LEU 141, GLU 166)	1 (MET 165)	1 (HIS 163)	PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, HIS 172, GLN 189
3	Sinapic acid	5	1	-6.1	1 (ARG 188)	1 (MET 165)	1 (HIS 41)	HIS 41, MET 49, TYR 54, ASN 142, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, PRO 168, ASP 187, ARG 188, GLN 189, THR 190, GLN 192
4	p-Coumaric acid	5	1	-5.329	1 (TYR 54)	1 (MET 165)	1 (HIS 41)	HIS 41, CYS 44, MET 49, PRO 52, TYR 54, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
			2	-5.188	0	1 (MET 165)	1 (HIS 163)	PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, GLN 189
5	Cinnamic acid	5	1	-4.978	0	1 (MET 165)	1 (HIS 41)	LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, GLN 189
			2	-4.705	0	1 (HIS 41)	1 (HIS 41)	HIS 41, CYS 44, MET 49, PRO 52, TYR 54, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
6	Rosmarinic acid	8	1	-7.409	1 (GLN 189)	1 (GLN 189)	1 (HIS 163)	THR 25, LEU 27, HIS 41, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, GLN 192
			2	-7.408	3 (LEU 141, GLY 143, GLU 166	1 (HIS 41)	1 (HIS 41)	HIS 41, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, ASP 187, ARG 188, GLN 189

Table 11: Molecular docking results for hydroxybenzoic acids (M^{pro} structure PDB: 6LU7)

Compound		Total clusters	Cluster	Binding energy (kcal/mol)	No of interactions			Total contacting residues
No.	**Name**				**H-bonds**	**Hydrophobic**	**Pi-pi**	
7	4-Hydroxybenzoic acid	4	1	-4.826	3 (LEU 141, GLY 143, GLU 166)	1 (GLU 166)	1 (HIS 163)	PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, HIS 172, GLN 189
8	Salicylic acid	3	1	-5.342	2 (LEU 141, GLY 143)	1 (GLU 166)	1 (HIS 163)	PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, HIS 172, GLN 189
9	Protocatechuic acid	1	1	-5.456	1 (GLU 166)	1 (GLU 166)	1 (HIS 163)	PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, HIS 172, GLN 189
10	Vanillic acid	4	1	-5.303	2 (GLY 143, GLU 166)	1 (HIS 163)	1 (HIS 163)	LEU 27, HIS 41, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, GLN 189
11	Veratric acid	4	1	-5.237	2 (GLY 143, GLU 166)	1 (GLU 166)	1 (HIS 163)	PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, GLN 189
12	Gallic acid	4	1	-5.601	2 (LEU 141, GLU 166)	1 (GLU 166)	1 (HIS 163)	HIS 41, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, GLN 189
13	Syringic acid	4	1	-5.347	3 (GLY 143, SER 144, GLU 166)	1 (MET 165)	0	HIS 41, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, GLN 189
			2	-5.278	2 (LEU 141, GLY 143)	1 (CYS 145)	0	THR 26, LEU 27, HIS 41, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, HIS 172, GLN 189

4.3.2. Hydroxybenzoic acids

The results for the hydroxybenzoic acids, in terms of binding affinity to the active site, are less encouraging than hydroxycinnamic acids (Table 11). The binding energies are overall lower and the fact that all the compounds in this category are highly related and are of very similar size results in the range of their binding energies being also narrower. Corresponding docking simulation studies report very similar binding energy for some of the hydroxybenzoic acids: Murugesan et al. (2021) calculated -5.4 kcal/mol for vanillic and -5.3 kcal/mol for gallic acid, while Patil et al. (2020) report -4.5 kcal/mol for 4-hydroxybenzoic and -5.1 kcal/mol for gallic acid and Mohapatra et al. (2020) calculated -4.68 kcal/mol for 4-hydroxybenzoic and -5.307 kcal/mol for gallic acid, therefore providing an encouraging indication towards the validity of the present results. No information on the conformation of the ligands in the binding site is provided, to which the present results could be compared to, but the binding complexes that occurred from the simulation reveal very similar binding for all the compounds, as their phenolic ring is stabilized between S1 and S1' subsites, below residue Cys 145. The main difference observed from molecule to molecule is the orientation of the carboxyl group, which is in some cases towards the left side of the active site and in some cases towards the right. That is the reason why both first clusters are considered in the case of syringic acid, since their binding energy is very similar. The phenolic ring is located in almost the same position, but the substitutions have opposite orientations in each case.

All docked hydroxybenzoic acids exhibit similar binding, as seen in Figure 33, ranging from -4.8 to -5.6 kcal/mol, with Glu 166, Gly 143 and Leu 141 emerging as key residues for hydrogen bonding, in this order of appearance frequency. Moreover, the O atom of the double bond in the -COOH group of the hydroxybenzoic acids is, in all the cases, involved in a hydrogen bond. Glu 166 is also a hydrophobic interaction hotspot, while His 163 is interacts through pi-pi stalking with six of the seven compounds of this group.

4-hydroxybenzoic acid, having only a -OH substitution in the benzene ring, has the lowest binding energy, whereas the highest binding energy is achieved by gallic acid, which has 3 hydroxy substitutions in the benzene ring. In this group, the hydroxy substitutions seem to be affecting binding more positively than methoxy substitutions. For example, protocatechuic (9), vanillic (10) and veratric acid (11) each have two substitutions in the benzene ring, starting with two hydroxy groups in 9, one hydroxy and one methoxy group in 10 and two methoxy groups in 11. As observed, the binding energy slightly drops as the methoxy substitutions increase. Additionally, is noted that 4-hydroxybenzoic (7) and salicylic acid (8), which are positional isomers, have almost identical interactions with the active site, whereas the transfer of the substitution from the 4- to the 2- position results in a 0.5 kcal/mol increase in the binding energy.

4.3.3. Caffeoyl quinic acids and derivatives

Caffeoyl quinic acids and their derivatives yielded very promising results for the inhibition of M^{pro} (Table 12). Apart from quinic acid, which is the smallest, parent-compound and has a low binding affinity of -5.887 kcal/mol, the other compounds have binding energies very close to -8 kcal/mol and the majority of them resulted in a better binding energy than native inhibitor N3 and enzymatic assay positive control GC376. The top five best hits were 3,5-dicaffeoyl quinic acid (22), 3,5-dicaffeoyl-quinic acid methyl ester (23), methyl 4-caffeoyl-3-dihydrocaffeoyl quinate (21), 3,4-dicaffeoylquinic acid (18) and 3-Caffeoyl-5-dihydrocaffeoylquinic acid (24). Another compound that stands out is 3-Caffeoyl-4-dihydrocaffeoylquinic acid (tungtungmadic acid), which has a good binding energy among the

rest of the caffeoylquinic acids (-8.396 kcal/mol) and was isolated from *Salicornia* from the first time (Chung et al. 2005).

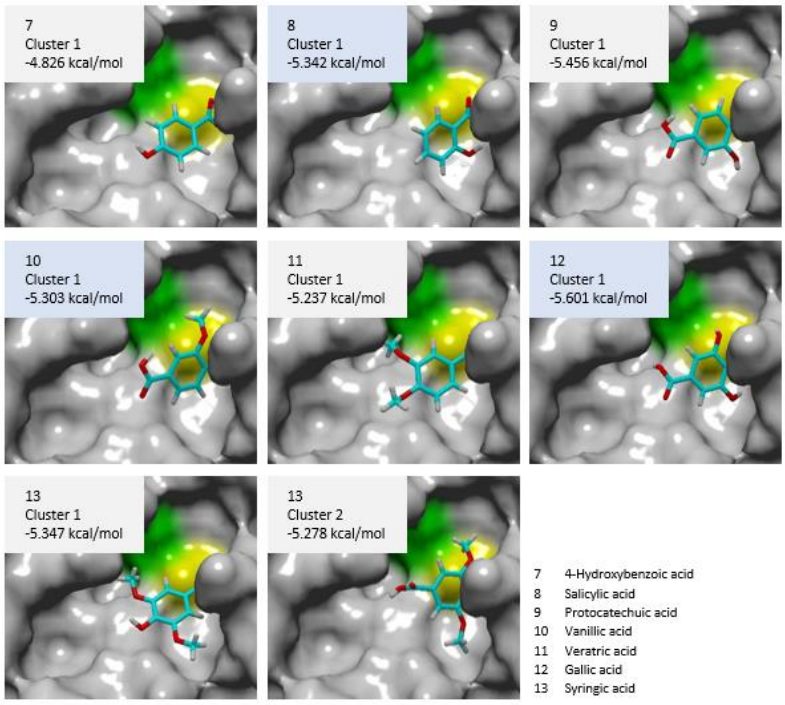


Figure 33: Binding conformations of hydroxybenzoic acids to the active site of SARS-CoV-2 M^{pro}, as resulted from the docking simulation

The results indicate that 3,5 substitutions on the quinic acid results in better binding compared to the respective 3,4 substitution. Also, caffeoyl substitutions appear to contribute in greater affinity to the protease than the hydrocaffeoyl substitutions, as in all the cases where the number and position of the substitutions is the same, more caffeoyl ones result in a better binding score. It is also noteworthy that in most cases, the methyl esters of the caffeoylquinic acids have almost identical binding energy with the respective acids. The only exception is chlorogenic acid, the methyl esterification of which results in a slightly better binding affinity, and the case of 3-caffeoyl-5-dihydrocaffeoylquinic acid, that exhibits a better binding score than its ester.

In terms of intramolecular interactions, it is observed that very few of the compounds form pi-pi interactions with the protease. As far as hydrogen bonds are concerned, there is no distinguishable pattern apart from the quite often bonding to residue Gly 143 and residues of the S4 subsite (Arg 188, Gln 189) and the fact that the majority of the hydrogen bonds are formed between -OH groups of the caffeoyl moieties of the ligands and the acti5.ve site.

Although there is no certainty as to whether the ligand-receptor complexes produced by the docking simulation are close to reality, the accuracy of the data can be reinforced by the observation of patterns in the orientation of the ligands that have structural affinity and by comparison with corresponding information in literature.

Quinic and chlorogenic acid have been studied in literature and yielded a binding score of -5.7 and -7.1 kcal/mol respectively, with molecular docking using AutoDock Vina and M^{pro} structure with the PDB ID 6Y84 as a receptor. These results are quite close to the ones produced by the simulation in this work, although lower in the case of chlorogenic acid. Murugesan et al. (2021)

also calculated a lower binding energy for chlorogenic acid (-6.9 kcal/mol). Regarding chlorogenic acid in particular, the analysis of its interactions with the protease show that the hydroxyl groups of the caffeoyl part interact with residues Leu 141, Cys 145 and His 163, while the ones in the quinic part interact with Thr 26 (Vicidomini et al. 2021). This suggests an orientation in the active site opposite to the first cluster of this simulation, but closer to the third, which has, however, considerably lower binding energy (by -1.136 kcal/mol). Another in silico study on chlorogenic acid as an M^{pro} inhibitor, using the same software and receptor as the present work, describes binding of the ligand to the protease in a different orientation, sideways, with the caffeoyl part interacting with residues Phe 140, Leu 141 and Glu 166 closer to the S1 subsite and the quinic moiety interacting with Thr 190 in S4, which does not correspond to any of the first clusters given as output of this simulation. Moreover, in this case the binding energy calculated is low, -6.0 kcal/mol (Patil et al. 2020). A similar configuration for chlorogenic acid is described in Mohapatra et al. (2020), with the binding affinity of the ligand being calculated however as high as -8.43 kcal/mol. All the above-mentioned information is inconclusive and highlights how diverse the results can be depending on the method used. Therefore, all three first clusters produced by the docking software will be presented, both for chlorogenic acid as well as for its methyl ester, taking into consideration that it is very likely that the two have a similar configuration in the binding site. For example, this is confirmed by the comparison of their first clusters: the conformation is indistinguishable, while the quinic part of the molecule is in the same position as the quinic acid, when docked to the active site of the protease on its own. In addition, the hydrogen bonds of the two compounds are identical, formed between the =O atom of the carboxylic group and residue Glu 166 and the -OH in the 5- position of the quinic acid aromatic ring and residue His 163.

Among the 3,4 and 3,5 caffeoylquinic acid derivatives, two dominant conformations are observed in the first clusters, as seen in Figure 34: the one involves the quinic moiety being located between the S1 and S1' subsites with the caffeoyl groups extending perpendicularly to each other and the other is characterized by the quinic moiety in the same position or higher in the S1' subsite with the two caffeoyl substitutions being towards the same direction, almost parallel to each other. The last arrangement is more common in the first clusters of the 3,4 caffeoylquinic derivatives, while the second seems to clearly prevail in the 3,5 derivatives, with the only exception being compound 27. The fact that almost exclusively the perpendicular conformation is observed in the 3,5 derivatives, while both are seen in the first clusters of the 3,4 derivatives could either mean that perhaps the first arrangement is more likely or that the different position of the substitutions creates different binding patterns, being an additional factor to be taken into consideration apart from the similarities in the structure and functional groups. Both these geometries are also found in compound 17, where the first two clusters have the same binding energy, but there is no clear criterion to rule one of the two out, so both are taken into consideration. Lastly in the case of compound 28, the first cluster does not resemble any of the above. Only the third one resembles the second of the geometries described above, while its binding energy does not drop significantly compared to the first clusters. Therefore, all three clusters will be presented. The results are summed up in Table 12.

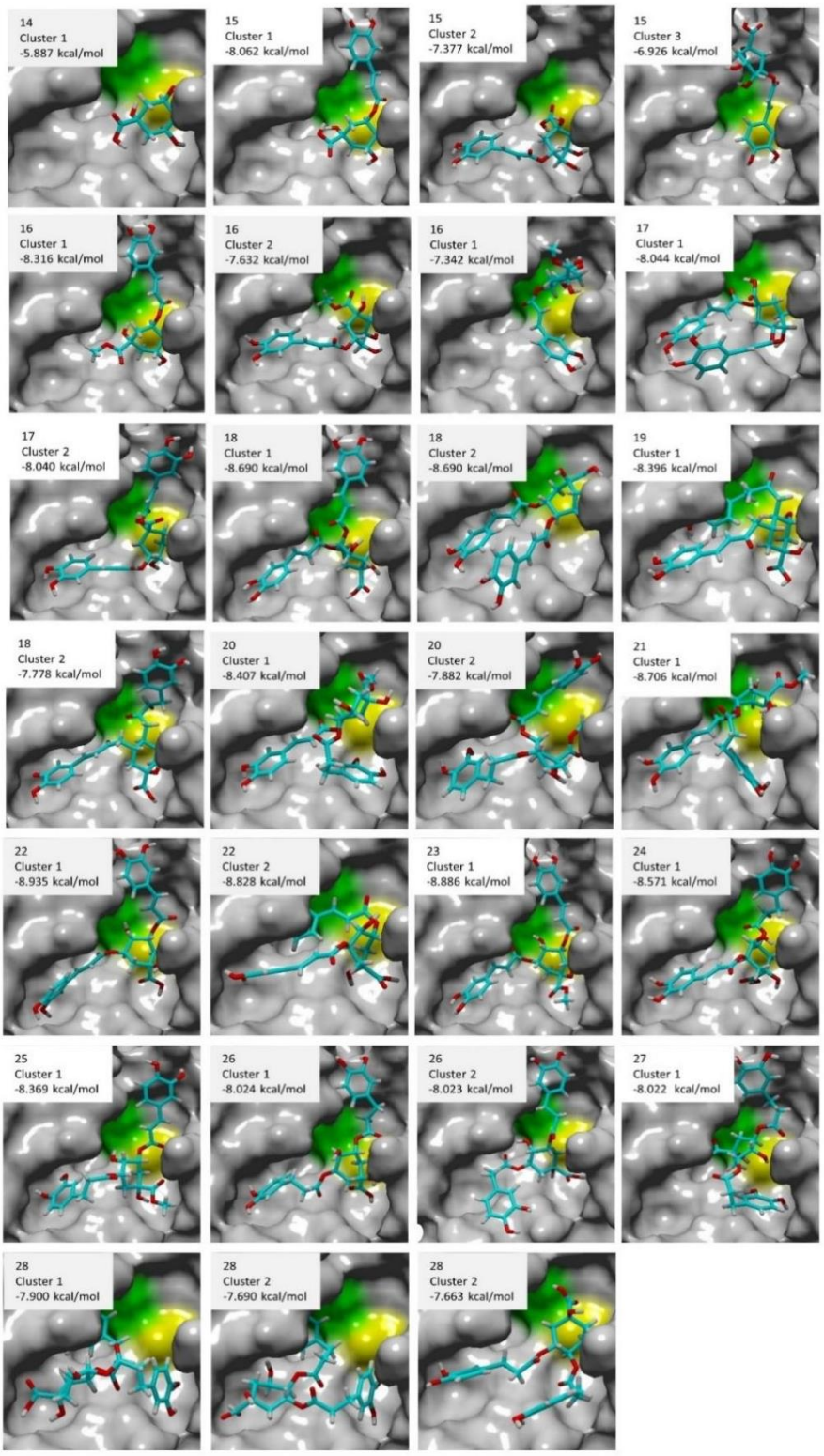


Figure 34: Binding conformations of caffeoylquinic acids to the active site of SARS-CoV-2 M^{pro}, as resulted from the docking simulation

Table 12: Molecular docking results for caffeoylquinic acids 2222222and derivatives (M^{pro} structure PDB: 6LU7)

Compound		Total cluster	Cluster	Binding energy (kcal/mol)	No of interactions			Contacting residues
No.	**Name**				**H-bonds**	**Hydrophobic**	**Pi-pi**	
14	Quinic acid	2	1	-5.887	1 (SER 144)	1 (GLU 166)	0	HIS 41, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, GLN 189
15	Chlorogenic acid	4	1	-8.062	3 (SER 46, HIS 163, GLU 166)	1 (THR 25)	0	THR 25, THR 26, LEU 27, HIS 41, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, GLN 189
			2	-7.377	1 (LEU 141)	1 (GLN 189)	0	HIS 41, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
			3	-6.926	0	1 (THR 25)	1 (HIS 163)	THR 24, THR 25, THR 26, LEU 27, HIS 41, CYS 44, THR 45, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, HIS 172, GLN 189
16	Methyl chlorogenate	7	1	-8.316	2 (HIS 163, GLU 166)	1 (THR 25)	0	THR 24, THR 25, THR 26, HIS 41, CYS 44, THR 45, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, GLN 189
17	1,3-di-O-Caffeoyl quinic acid	10	1	-8.044	2 (THR 190, GLN 192)	1 (GLU 166)	0	PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, GLN 192
			2	-8.040	1 (THR 24)	1 (THR 25)	1 (HIS 41)	THR 24, THR 25, THR 26, LEU 27, HIS 41, THR 45, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, LEU 167, PRO 168, ARG 188, GLN 189, THR 190, GLN 192
18	3,4-Dicaffeoylquinic acid	7	1	-8.69	2 (THR 24, THR 190)	1 (MET 49)	0	THR 24, THR 25, THR 26, HIS 41, THR 45, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
			2	-8.69	1 (ARG 188)	1 (MET 165)	0	THR 26, LEU 27, HIS 41, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, ARG 188, GLN 189, THR 190, GLN 192
19	3-Caffeoyl-4-dihydrocaffeoylquinic acid	9	1	-8.396	2 (GLY 143, ARG 188)	1 (HIS 41)	1 (HIS 41)	HIS 41, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, PRO 168, HIS 172, ASP 187, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
			2	-7.778	2 (CYS 145 x2)	1 (MET 165)	0	THR 24, THR 25, THR 26, HIS 41, THR 45, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
20	3-Caffeoyl-4-dihydrocaffeoylquinic acid methyl ester	9	1	-8.407	2 (GLY 143, GLN 189)	1 (GLU 166)	0	THR 25, THR 26, LEU 27, HIS 41, VAL 42, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, GLN 192
			2	-7.882	2 (THR 26, ARG 188)	1 (MET 49)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
21	Methyl 4-caffeoyl-3-dihydrocaffeoyl quinate	1	1	-8.706	1 (GLN 189)	1 (GLN 189)	1 (HIS 163)	THR 25, THR 26, LEU 27, HIS 41, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, GLN 192
22	3,5-Dicaffeoyl quinic acid	10	1	-8.935	0	1 (MET 165)	0	THR 24, THR 25, THR 26, HIS 41, THR 45, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
			2	-8.828	3 (LEU 141, GLY 143, THR 190)	1 (MET 165)	1 (HIS 41)	HIS 41, CYS 44, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, HIS 172, VAL 186, ASP 187, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
23	3,5-di-O-Caffeoyl-quinic acid methyl ester	7	1	-8.886	2 (GLY 143, GLN 189)	1 (GLN 189)	0	THR 24, THR 25, THR 26, HIS 41, THR 45, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, ALA 191, GLN 192

24	3-Caffeoyl-5-dihydrocaffeoyl quinic acid	6	1	-8.571	2 (THR 24, THR 190)	1 (GLN 189)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
25	3-Caffeoyl-5-dihydrocaffeoylquinic acid methyl ester	8	1	-8.369	2 (GLY 143, CYS 145)	1 (MET 49)	0	THR 24, THR 25, THR 26, HIS 41, THR 45, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
26	3,5-di-O-Dihydrocaffeoyl quinic acid	5	1	-8.024	4 (THR 25, GLY 143, GLU 166, ARG 188)	1 (THR 25)	0	THR 24, THR 25, THR 26, LEU 27, CYS 44, THR 45, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, PRO 168, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
			2	-8.023	3 (SER 46, LEU 141, GLU 166)	1 (MET 49)	0	THR 24, THR 25, THR 26, HIS 41, THR 45, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, GLY 170, GLN 189
27	3,5-di-dihydrocaffeoylquinic acid methyl ester	9	1	-8.022	3 (SER 46, GLY 143, GLN 189)	1 (GLU 166)	0	THR 24, THR 25, THR 26, HIS 41, THR 45, SER 46, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, ASP 187, ARG 188, GLN 189
28	4,5-di-O-dihydrocaffeoyl quinic acid	9	1	-7.9	3 (HIS 163, GLU 166, GLN 189)	1 (GLU 166)	1 (HIS 41)	HIS 41, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ASP 187, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
			2	-7.69	3 (HIS 163, GLU 166, THR 190)	1 (GLU 166)	1 (HIS 41)	HIS 41, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, VAL 186, ASP 187, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
			3	-7.66	1 (ARG 188)	1 (GLU 166)	0	THR 25, LEU 27, HIS 41, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, HIS 172, ARG 188, GLN 189, THR 190, ALA 191, GLN 192

4.3.4. Flavonoids and flavanones

This group of compounds is the one to which the best hits of this screening belong to, with the binding energies ranging from -6.880 to -9.384 kcal/mol) and the majority of compounds having a better binding energy than GC376. Fewer, but still many of the flavonoids and flavanones also exceeded the binding affinity of inhibitor N3, as calculated from the present simulation. The docking simulation resulted in the most favorable binding energy for compound 38, isorhamnetin-3-O-rutinoside, while numerous compounds had a binding energy close to or above -9.0 kcal/mol.

Regarding the review of the clusters produced by the simulation, since the backbone is the same for all the flavonoids, the binding orientation of myricetin is being used as an indication. Overall, the majority of the flavonoids had an orientation similar to myricetin in their first cluster (Figure 35). This fact reinforces the hypothesis that this could be the actual conformation of the molecules in nature and validates the use of myricetin as a point of reference for the selection of the best cluster. However, in the case of quercetin, all the first three clusters have essentially identical binding energy and the third cluster is the one that is closer to myricetin, with the chromone moiety being positioned in the cavity of the S2 subsite. Similarly, cluster 2 was taken into consideration for compound 33 even though in both the first two clusters the chromone group was in the S2 subsite, because in the second case the phenyl substitution was orientated towards S1' subsite as in myricetin and not in the opposite direction, as it happened with the first cluster. It was also taken into consideration that the two clusters had a minor difference in binding energy (0.2 kcal/mol). In the case of rhamnetin, although it has a high structural resemblance to myricetin, none of the clusters have a very similar orientation.

This led to the selection of the first cluster, as it had the highest binding energy and the closest conformation. This deviation can be due to the fact that rhamnetin has a methoxy substitution at the 7- position of the chromone, which increases the volume of the molecule on that side and might not allow it to enter the S2 cavity. The case is similar for hesperetin as well, with the only difference being that the first three clusters have very similar binding energy, therefore they will all be considered. There are also other molecules which do not have the same conformation as myricetin in none of their first clusters, such as isorhamnetin-3-O-neohesperidoside or hesperidin. This can also be justified by the significantly larger size of the molecule.

The data, gathered in Table 13, also designates a pattern of structure-binding affinity relationship: all the sugar derivatives of flavonoids and flavanones have a higher binding affinity than their parent compound. For example, **quercetin** has a binding energy of -7.396 kcal/mol, while for its glucoside, **isoquercetin**, the binding energy increases by 1.5 kcal/mol (-8.952 kcal/mol) and for the malonyl-glucoside substituted derivative the increase is even higher (-9.114 kcal/mol). The greatest improvement is observed in **rutin**, the rutinoside derivative of quercetin (-9.166 kcal/mol). These findings are supported by other recent anti-SARS-CoV-2 *in silico* studies, in which both quercetin and rutin have drawn attention. Molecular docking performed for quercetin using structures 6LU7, 6Y2E and 6Y2F as receptors resulted in binding energies between -6.9 and -7.5 kcal/mol, which are very close to the binding energy calculated in the present study (Murugesan et al. 2021; Abian et al. 2020). As far as interactions are concerned, Met 165 is highlighted as a key residue in literature, while in this study it appears only as a hydrophobic contact for the first cluster. Also, Abian et al. (2020) report three hydrogen bonds with residues Asn 142, Ser 144 and Met 165, as opposed

to the no hydrogen bonds calculated by Vina in this work. The increase in the binding energy for rutin is also confirmed in literature. Das et. Al (2020) calculated a binding energy of -8.8 kcal/mol through molecular docking to M^{pro} (PDB:6LU7). Both the binding energy and the contacting residues calculated between rutin and the protease (Thr25, Thr26, Leu 27, His41, Phe 140, Leu 141, Asn 142, Gly 143, Ser 144, Cys 145, His 163, His 164, Met 165, Glu 166, Arg 188, Gln 189, Thr 190) are very similar to the ones produced by this study, reinforcing the validity of the results. In a different study, the binding energy for **rutin**, where the same PDB structure was used as the receptor, was calculated to be much higher, -11.33 kcal/mol. The contacting residues were to a great extent common, but the hydrogen bonds shown were more (10 bonds, with residues Tyr 54, Phe 140, Cys 145, His 163, His 164, Glu 166, Gln 192) (Shivanika et al. 2020). Another *in silico* study on the same receptor reports a binding energy of -11.187 kcal/mol (Teli, Shah, and Chhabria 2021), while molecular docking using PDB structure 6Y2E also produces a binding score of the same range (-11.8 kcal/mol) (Gajjar, Dhameliya, and Shah 2021). These deviations could be due to the different program used for the simulation (AutoDock 4.2.6, extra precision GLIDE docking module of Maestro), docking parameters (e.g. simulation cell, number of runs) and the different treatment of the receptor and the ligand. In any case, they results indicate a very good potential of rutin as a SARS-CoV-2 M^{pro} inhibitor.

Rhamnetin and **isorhamnetin**, two methyl esters of quercetin, have lower binding affinity compared to quercetin (-7.155 and -7.233 kcal/mol respectively). However, the above-described pattern can be observed in the derivatives of isorhamnetin as well. The neohesperidoside resulted in the least evident improvement (-7.838 kcal/mol), followed by the rhamnosyl arabinoside (-8.167), the galactoside and the rhamnoside increased significantly and to the same extent the binding affinity of the compound (-8.383 and -8.335 kcal/mol), the glucoside produced an even more encouraging result (-8.613 kcal/mol) and in this case too the rutinoside boosted the ligands binding affinity the most, by more than 2 kcal/mol (-9.384 kcal/mol). Respective studies report very similar data for i**sorhamnetin and isorhamnetin 3-O-β-D-glucopyranoside**. The binding affinity to M^{pro} (PDB ID: 6LU7) was calculated equal to -7.3 and -8.7 kcal/mol respectively., verifying the tendency of glucosides to have higher binding energy than their parent compounds. Specifically for isorhamnetin-3-O-β-D-glucopyranoside the contacting residues were also described, and are to a great extent common to the results of this work (Met 165, Thr 26, His 41, Tyr 54, Met 49, Phe 140, Asn 142, Gly 143, Ser 144, Cys145, His 163, His 172, Glu 186, Asp 187, Arg 188, Gln 189) (Das et al. 2020). Another study performed with Autodock Vina and the use of the protease structure with the PDB ID 6Y84 resulted in calculation of the same binding energy for isorhamnetin **(**-7.3 kcal/mol) but considerably lower for its glucopyranoside -7.5 kcal/mol. Analysis of the interactions of the ligand with the proteases revealed hydrogen bonds of isorhamnetin with Thr 26, Asn 142 and Gln 189 and of 38 with Thr 24, Thr 26, Leu 141, Asn 142, Gly 143 and Gln 189 (Vicidomini et al. 2021).

This tendency is confirmed by the rest of the pairs of flavonoids and flavanones and their glucosides (or other sugar derivatives), such as **hesperetin** (-7.09 kcal/mol) and its rutinoside, **hesperidin** (-8.636 kcal/mol). Published data with which the present results can be juxtaposed are available only for hesperidin, which has also been tested *in silico*, using PDB structure 6Y84 as a receptor, and resulted in a quite lower binding score than the one in this work (-5.8 kcal/mol) (Tomic et al. 2020). The difference can be attributed to the difference of the protein structures used, since it has been highlighted that the conformation of the protein plays an

important role in the docking simulation because some parts are plastic and can considerably change depending on the interacting compound.

Apigenin (-7.796 kcal/mol) also seems to have a less encouraging inhibitory potential compared to its glucoside (-8.391 kcal/mol) and its galactoside (-7.834 kcal/mol), although in this case the increase is very slight in the case of **apigenin-7-O-galactoside** and also lower in the case of **apigenin-7-O-glucoside** compared to the respective increase for the quercetin or the isorhamnetin derivatives, indicating that perhaps the 7- substitution is not as favorable as the 3- substitution. From this subgroup of compounds, only apigenin has been screened in relevant studies, and resulted in less promising binding energies (-6.7 and -7.090 kcal/mol), which correspond to values closer to the third cluster of this study than the first (Murugesan et al. 2021; Teli et al. 2021).

Another parent compound whose derivative exhibits better inhibitory potential, as depicted in the binding energy values, is **kaempferol** (-7.752 kcal/mol). Its glucoside, **astragalin**, has an increased binding affinity to M^{pro} of -8.989 kcal/mol. Kaempferol has been included in *in silico* screenings for potential inhibitors performed in various studies, with results for its binding energy varying from -6.4 to -7.8 kcal/mol. It is interesting that one of the studies reports a considerable binding affinity of kaempferol to the spike protein of SARS-CoV-2, apart from the main protease (Tallei et al. 2020; Teli et al. 2021; Murugesan et al. 2021; Khan et al. 2021). Astragalin is also a previously investigated compound, with results from different studies varying substantially, even though the simulation was executed on the same receptor. Reported values include -5.8 kcal/mol (calculated with Autodock Vina) (Murugesan et al. 2021), -7.9 kcal/mol (calculated with Autodock Vina) (Vicidomini et al. 2021), -8.5 kcal/mol (calculated with Autodock 4.2.6) (Adejoro et al. 2020) and -9.120 kcal/mol (calculated with extra precision module of GLIDE) (Teli et al. 2021), while the use of GLIDE and 6Y2E as a receptor gave a binding energy of -7.6 kcal/mol as an output (Gajjar et al. 2021). Particularly the study of Murugesan et al. (2021) provides an image of the binding complex of astragalin with the protease, which is very similar to the first cluster of the present molecular docking study.

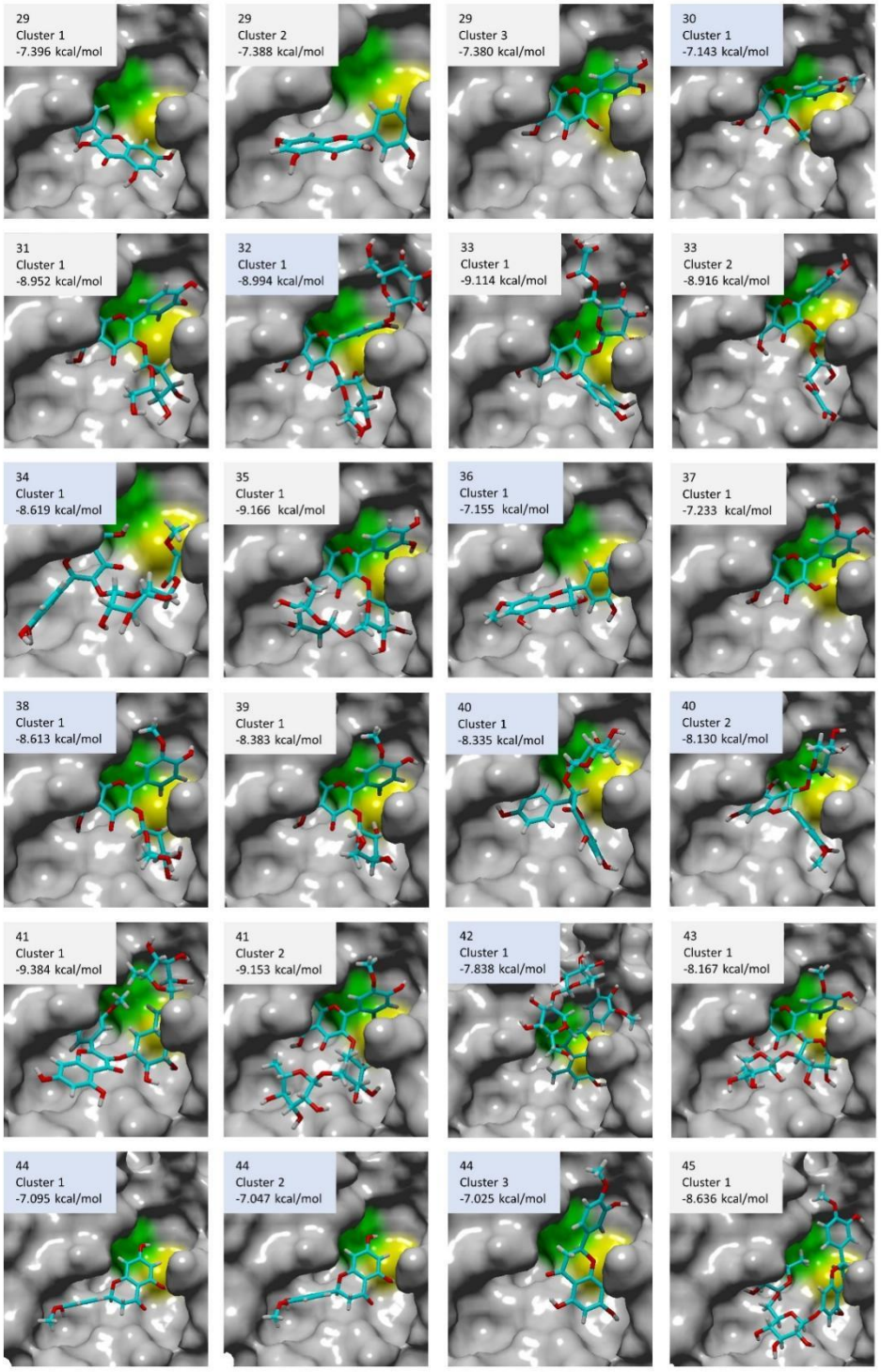
Overall, the aforementioned data implies that glycosylation of flavonoids and flavanones increases their binding affinity to M^{pro}, with galactosides, glucosides and rutinosides having an increasingly stronger improving effect.

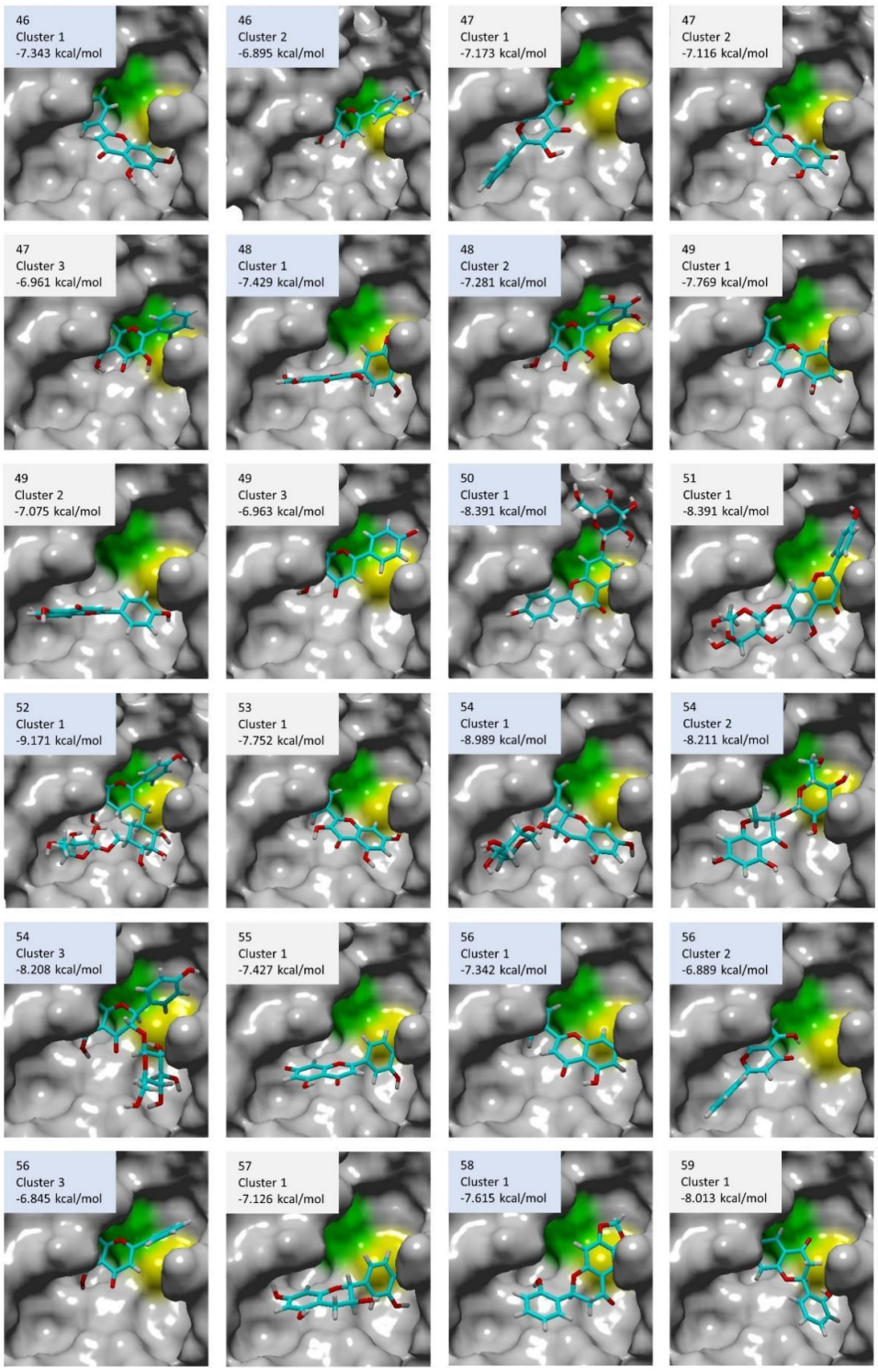
Among the unsubstituted flavonoids, apigenin and kaempferol had the highest binding energy, thus it would be interesting to further test their derivatives for potential antiviral activity. Moreover, binding energy calculated for **myricetin** was -7.529 kcal/mol. Although the value is not particularly high, myricetin is an *in vitro* confirmed and co-crystallized inhibitor of SARS-CoV-2 M^{pro} as mentioned previously. This confirms that the binding energy alone is not a clear indicator of the inhibitory potential of a compound. The binding energy of myricetin to M^{pro} has also been calculated in other studies, with values between -6.5 and -7.311 kcal/mol. Although they slightly deviate from each other, they are comparable to the results of the present study (Murugesan et al. 2021; Teli et al. 2021).

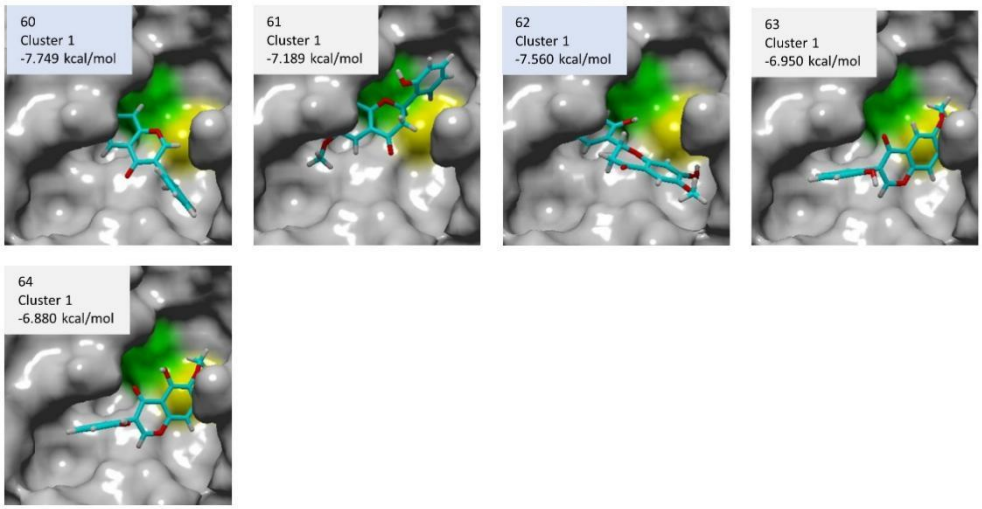
Luteolin, acacetin and chrysin also have highly comparable binding energy values (-7.427, -7.343 ad -7.342 kcal/mol respectively), which are also very similar to the ones for quercetin. The result for chrysin is validated by the similar value published in Teli, Shah and Chhambria's work (2021) (-7.162 kcal/mol). The binding scores for galangin and catechin are also alike (-7.173 and -7.126 kcal/mol respectively), although marginally lower than the rest of the

flavonoids and flavanones. The results for catechin are confirmed by other studies, which published binding energy values of -6.709 and -7.1 kcal/mol (Teli et al. 2021; Ghosh et al. 2020). Available data for galangin report a lower value of -6.65 kcal/mol, which occurred, however, from a simulation with AutoDock and 6Y84 as a protein structure (Kundu Debanjan et al. 2021).

Lastly, is it important to mention that compounds 58, 59 and 60, whose structure involves a methylenedioxy substitution, resulted in higher binding scores comparted to the previously mentioned non-glycosylated flavonoids and flavanones.




46
Cluster 1
-7.343 kcal/mol
46
Cluster 2
-6.895 kcal/mol
47
Cluster 1
-7.173 kcal/mol
47
Cluster 2
-7.116 kcal/mol
47
Cluster 3
-6.961 kcal/mol
48
Cluster 1
-7.429 kcal/mol
48
Cluster 2
-7.281 kcal/mol
49
Cluster 1
-7.769 kcal/mol
49
Cluster 2
-7.075 kcal/mol
49
Cluster 3
-6.963 kcal/mol
50
Cluster 1
-8.391 kcal/mol
51
Cluster 1
-8.391 kcal/mol
52
Cluster 1
-9.171 kcal/mol
53
Cluster 1
-7.752 kcal/mol
54
Cluster 1
-8.989 kcal/mol
54
Cluster 2
-8.211 kcal/mol
54
Cluster 3
-8.208 kcal/mol
55
Cluster 1
-7.427 kcal/mol
56
Cluster 1
-7.342 kcal/mol
56
Cluster 2
-6.889 kcal/mol
56
Cluster 3
-6.845 kcal/mol
57
Cluster 1
-7.126 kcal/mol
58
Cluster 1
-7.615 kcal/mol
59
Cluster 1
-8.013 kcal/mol



29 Quercetin
30 Quercetin 3′,4′-dimethyl ether
31 Isoquercetin
32 Quercetin-3′,4′-diglucoside
33 Quercetin-3-O-(6″-O-malonyl)-β-d-glucoside
34 Isoquercitrin 6″-O-methyloxalate
35 Rutin
36 Rhamnetin
37 Isorhamnetin
38 Isorhamnetin 3-O-β-D-glucopyranoside
39 Isorhamnetin-3-O-galactoside
40 Isorhamnetin 3-O-rhamnoside
41 Isorhamnetin 3-O-rutinoside
42 Isorhamnetin 3-O-neohesperidoside
43 Isorhamnetin 3-O-rhamnosyl(1-2)arabinoside
44 Hesperetin
45 Hesperidin
46 Acacetin
47 Galangin
48 Myricetin
49 Apigenin
50 Apigenin 7-glucoside
51 Apigenin-7-O-galactoside
52 Pelargonidin-3-O-rutinoside
53 Kaempferol
54 Kaempferol-3-O-β-D-glucoside
55 Luteolin
56 Chrysin
57 Catechin
58 2S-5,2′-Dihydroxy-6,7- methylenedioxyflavanone
59 (−)-(2S)-2′-Hydroxy-6,7- methylenedioxyflavanone
60 2′-Hydroxy-6,7-methylenedioxyisoflavone
61 2S-2′-Hydroxy-6,7-dimethoxy-flavanone
62 2S-2',7- dihydroxy-6-methoxyflavanone
63 2′,7-dihydroxy-6-methoxyisoflavone
64 Irilin B

Figure 35: Binding conformations of flavonoids and flavanones to the active site of SARS-CoV-2 M^{pro}, as resulted from the docking simulation

Table 13: Molecular docking results for flavonoids and flavanones (M^{pro} structure PDB: 6LU7)

Compound		Total cluster	Cluster	Binding energy (kcal/mol)	No of interactions			Total contacting residues
No	Name				H-bonds	Hydrophobic	Pi-pi	
29	Quercetin	6	1	-7.396	**0**	1 (MET 165)	1 (HIS 163)	HIS 41, MET 49, PRO 52, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
			2	-7.388	**0**	1 (GLN 189)	1 (HIS 163)	PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, HIS 172, ARG 188, GLN 189, THR 190, GLN 192
			3	-7.380	0	1 (GLN 189)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, CYS 44, MET 49, PRO 52, TYR 54, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
30	Quercetin 3,4'-dimethyl ether	8	1	-7.143	0	1 (GLN 189)	1 (HIS 41)	THR 26, LEU 27, HIS 41, CYS 44, MET 49, PRO 52, TYR 54, LEU 141, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
31	Isoquercetin	7	1	-8.952	1 (GLU 166)	1 (HIS 41)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, ASP 187, ARG 188, GLN 189
32	Quercetin-3',4' diglucoside	9	1	-8.994	1 (THR 24)	1 (GLN 189)	1 (HIS 41)	THR 24, THR 25, THR 26, LEU 27, HIS 41, CYS 44, MET 49, PRO 52, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, ASP 187, ARG 188, GLN 189
33	Quercetin 3-O-(6″-O-malonyl)-β-d-glucoside	11	1	-9.114	1 (LEU 141)	1 (MET 49)	1 (HIS 41)	THR 24, THR 25, THR 26, LEU 27, HIS 41, THR 45, SER 46, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, VAL 186, ASP 187, ARG 188, GLN 189, GLN 192
			2	-8.916	1 (CYS 145, GLU 166)	1 (HIS 41)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, HIS 172, ASP 187, ARG 188, GLN 189
34	Isoquercitrin 6''-O-methyloxalate	11	1	-8.619	0	1 (GLN 189)	1 (HIS 41)	THR 26, LEU 27, HIS 41, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, ASP 187, ARG 188, GLN 189
35	Rutin	10	1	-9.166	2 (CYS 145, GLU 166)	1 (HIS 41)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, MET 49, PRO 52, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, ASP 187, ARG 188, GLN 189
36	Rhamnetin	6	1	-7.155	1 (SER 144)	1 (MET 165)	1 (HIS 163)	PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, GLN 192
37	Isorhamnetin	4	1	-7.233	1 (ASP 187)	1 (HIS 41)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, MET 49, PRO 52, TYR 54, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
38	Isorhamnetin 3-O-β-D-glucopyranoside	7	1	-8.613	1 (THR 26)	1 (HIS 41)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, CYS 44, MET 49, PRO 52, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, ASP 187, ARG 188, GLN 189
39	Isorhamnetin-3-O-galactoside	7	1	-8.383	1 (ASP 187)	1 (HIS 41)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, MET 49, PRO 52, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, ASP 187, ARG 188, GLN 189
40	Isorhamnetin 3-O-rhamnoside	9	1	-8.335	1 (PHE 140)	1 (MET 165)	1 (HIS 163)	THR 25, THR 26, LEU 27, HIS 41, VAL 42, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, ASP 187, ARG 188, GLN 189, THR 190, GLN 192
			2	-8.130	1 (LEU 141)	1 (GLN 189)	1 (HIS 163)	THR 24, THR 25, THR 26, LEU 27, HIS 41, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, ASP 187, ARG 188, GLN 189, THR 190
41	Isorhamnetin 3-O-rutinoside	11	1	-9.384	3 (THR 26, LEU 141, GLY 143)	1 (GLN 189)	0	THR 25, THR 26, LEU 27, HIS 41, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, VAL 186, ASP 187, ARG 188, GLN 189
			2	-9.153	1 (THR 26)	1 (GLN 189)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, CYS 44, MET 49, PRO 52, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, GLY 170, HIS 172, ASP 187, ARG 188, GLN 189
42	Isorhamnetin 3-O-neohesperidoside	11	1	-7.838	3 (THR 24, LEU 141, GLY 143)	1 (ASN 142)	1 (HIS 163)	THR 24, THR 25, THR 26, HIS 41, CYS 44, THR 45, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, GLN 189
43	Isorhamnetin 3-O-rhamnosyl(1-2)arabinoside	11	1	-8.167	0	1 (MET 49)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, CYS 44, MET 49, PRO 52, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, PRO 168, HIS 172, ASP 187, ARG 188, GLN 189
44	Hesperetin	6	1	-7.095	0	1 (HIS 41)	1 (HIS 41)	HIS 41, CYS 44, MET 49, PRO 52, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189

			2	-7.047	1 (GLN 189)	1 (MET 165)	1 (HIS 41)	LEU 27, HIS 41, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, LEU 167, PRO 168, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
			3	-7.025	2 (GLU 166, GLN 189)	1 (MET 49)	1 (HIS 163)	THR 24, THR 25, THR 26, LEU 27, HIS 41, THR 45, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, HIS 172, GLN 189
45	Hesperidin	9	1	-8.636	3 (THR 26 x2, GLU 166)	1 (MET 49)	1 (HIS 163)	THR 25, THR 26, HIS 41, CYS 44, THR 45, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, ASP 187, ARG 188, GLN 189, THR 190
46	Acacetin	7	1	-7.343	0	1 (HIS 41)	1 (HIS 41)	HIS 41, CYS 44, MET 49, PRO 52, TYR 54, PHE 140, LEU 141, ASN 142, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
			2	-6.895	0	1 (GLN 189)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, CYS 44, MET 49, PRO 52, TYR 54, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
47	Galangin	6	1	-7.173	1 (GLU 166)	1 (MET 165)	1 (HIS 41)	HIS 41, MET 49, TYR 54, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, ASP 187, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
			2	-7.116	1 (SER 144)	1 (GLN 189)	1 (HIS 163)	MET 49, TYR 54, PHE 140, LEU 141, ASN 142, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
			3	-6.961	0	1 (GLN 189)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, MET 49, PRO 52, TYR 54, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
48	Myricetin	6	1	-7.429	3 (SER 144, HIS 163, ARG 188)	1 (GLN 189)	1 (HIS 163)	HIS 41, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, HIS 172, ARG 188, GLN 189, THR 190, GLN 192
			2	-7.281	0	1 (HIS 41)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, CYS 44, MET 49, PRO 52, TYR 54, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
49	Apigenin	6	1	-7.796	1 (GLU 166)	1 (MET 165)	1 (HIS 41)	HIS 41, MET 49, PRO 52, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
			2	-7.075	3 (HIS 163, GLU 166, ARG 188)	1 (GLN 189)	1 (HIS 163)	PHE 140, LEU 141, ASN 142, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, ARG 188, GLN 189, THR 190, GLN 192
			3	-6.963	1 (THR 26)	1 (HIS 41)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, CYS 44, MET 49, PRO 52, TYR 54, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
50	Apigenin 7-glucoside	6	1	-8.391	1 (GLY 143)	1 (MET 165)	1 (HIS 41)	THR 24, THR 25, THR 26, LEU 27, HIS 41, CYS 44, THR 45, SER 46, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, ARG 188, GLN 189, THR 190, GLN 192
51	Apigenin-7-O-galactoside	5	1	-7.834	0	1 (GLU 166)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
52	Pelargonidin-3-O-rutinoside	4	1	-9.171	1 (TYR 54)	1 (GLN 189)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, CYS 44, MET 49, PRO 52, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ASP 187, ARG 188, GLN 189, THR 190, ALA 191
53	Kaempferol	5	1	-7.752	0	1 (MET 165)	1 (HIS 41)	HIS 41, CYS 44, MET 49, PRO 52, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
54	Astragalin	10	1	-8.989	1 (SER 144)	1 (HIS 41)	1 (HIS 41)	HIS 41, MET 49, TYR 54, PHE 140, LEU 141, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ASP 187, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
			2	-8.211	1 (GLY 143)	1 (MET 165)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, MET 49, TYR 54, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
			3	-8.208	2 (THR 26, CYS 145)	1 (GLN 189)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, CYS 44, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, ASP 187, ARG 188, GLN 189
55	Luteolin	6	1	-7.427	1 (LEU 141)	1 (GLN 189)	1 (HIS 163)	PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, HIS 172, ARG 188, GLN 189, THR 190, GLN 192
56	Chrysin	5	1	-7.342	2 (LEU 141, GLU 166)	1 (MET 165)	1 (HIS 163)	HIS 41, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
			2	-6.889	1 (HIS 164)	1 (HIS 41)	1 (HIS 41)	HIS 41, MET 49, TYR 54, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, ASP 187, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
			3	-6.845	0	1 (HIS 41)	1 (HIS 41)	THR 26, LEU 27, HIS 41, CYS 44, MET 49, PRO 52, TYR 54, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189

57	Catechin	6	1	-7.126	2 (LEU 141 x2)	1 (GLN 189)	1 (HIS 163)	PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, ARG 188, GLN 189, THR 190, GLN 192
58	2S-5,2′-Dihydroxy-6,7-methylenedioxyflavanone	5	1	-7.615	1 (SER 144)	1 (GLU 166)	1 (HIS 41)	THR 26, LEU 27, HIS 41, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, GLN 189
59	(−)-(2S)-2′-Hydroxy-6,7-methylenedioxyflavanone	6	1	-8.013	0	1 (GLU 166)	1 (HIS 41)	HIS 41, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, ASP 187, ARG 188, GLN 189
60	2′-Hydroxy-6,7-methylenedioxyisoflavone	5	1	-7.749	1 (GLU 166)	1 (MET 165)	1 (HIS 41)	HIS 41, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, ASP 187, ARG 188, GLN 189
61	2S-2′-Hydroxy-6,7-dimethoxy-flavanone	7	1	-7.189	0	1 (THR 25)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, CYS 44, MET 49, PRO 52, TYR 54, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189, THR 190, GLN 192
62	2S-2',7- Dihydroxy-6-methoxyflavanone	6	1	-7.56	0	1 (MET 165)	1 (HIS 41)	HIS 41, TYR 54, PHE 140, LEU 141, ASN 142, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
63	2′,7-dihydroxy-6-methoxyisoflavone	4	1	-6.95	1 (GLY 143)	1 (GLU 166)	1 (HIS 163)	THR 26, LEU 27, HIS 41, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, GLN 189
64	Irilin B	6	1	-6.88	3 (LEU 141, GLY 143, GLN 189)	1 (GLN 189)	1 (HIS 163)	THR 26, LEU 27, HIS 41, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, ARG 188, GLN 189, THR 190

4.3.5. Sterols

Although all the sterols screened are bulky molecules capable of blocking multiple subsites of the active site of M^{pro}, the binding scores for this category of ligands are not particularly high, ranging from -6.56 to -7.45 kcal/mol for the first clusters (Table 14). Compound 67, ergosterol, is the smallest among the sterols and yields the best results in terms of binding energy.

A similar orientation was observed among the first clusters of the compounds of this category (Figure 36), with the gonane being situated in the S1' subsite and the carbon chain substitution of the pentacyclic ring extending to the S4 subsite. Therefore, this pattern can serve as an indication for the cluster review. In the case of cerevisterol, where this conformation was observed in the second cluster, the results for the first one is also being displayed. The case is the same for β-daucosterol (β-sitosterol glucoside), with the only difference being that the first cluster is not considered at all since the ligand was located in a different binding site. Only in the case of β-sitosterol is this orientation not among the first clusters. It can be observed in the fourth cluster, where the binding energy, however, drops to significantly lower levels. Lastly, as far as interactions are concerned, no distinctive pattern apart from the pi-pi interactions between the cyclohexanic ring of the sterols containing the double bond and His 41 is observed. Overall, the results produced by the docking simulation are not exceptional, but are definitely worth further investigation, especially taking into consideration the size of the ligands and the fact that they seem to be positioned in a way that effectively blocks the catalytic dyad.

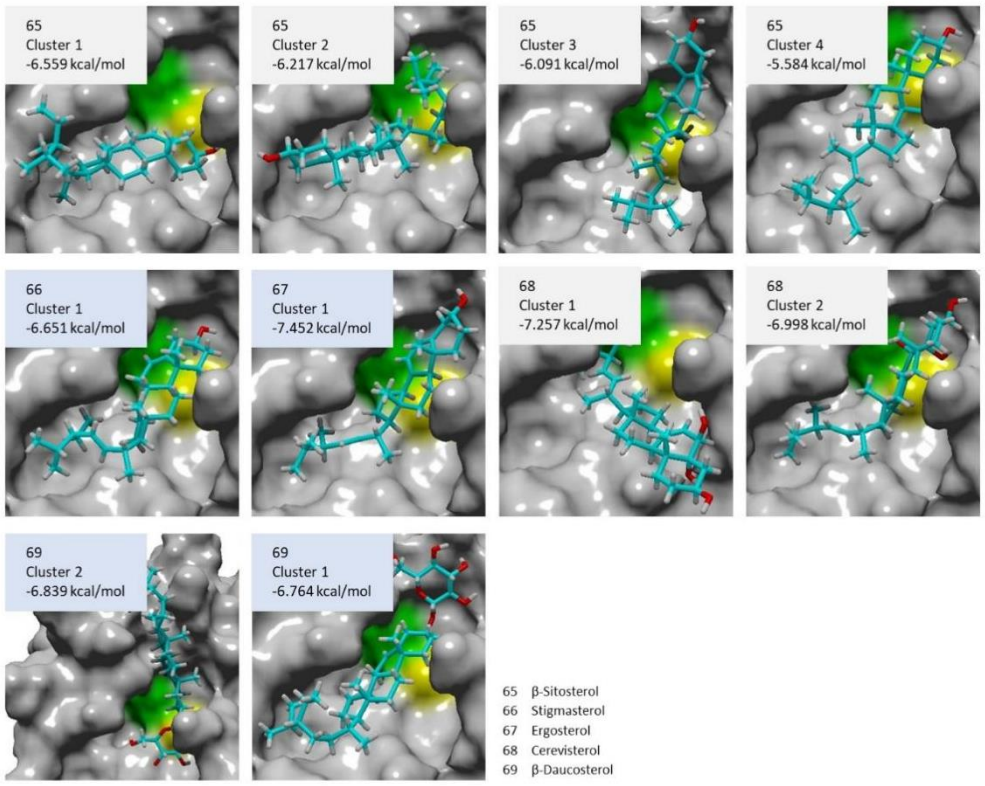


Figure 36 : Binding conformations of sterols to the active site of SARS-CoV-2 M^{pro}, as resulted from the docking simulation

4.3.6. Chromones

As derived from Table 15, among the chromones detected in extracts of *Salicornia* plants, 7-O-β-d-Glucopyranosyl-6-methoxychromone (73) has the highest binding affinity to the active

stie of M^{pro} (-7.229 kcal/mol), forming a hydrogen bond with catalytic residue Cys 145 and being further stabilized through hydrophobic and pi-pi interactions with key residues Glu 166 and His 163, respectively. It is noticed that this compound, that is conjugated with glucose, is the one that has a significantly better binding energy compared to the rest of the compounds, that have a similar binding score between -5.30 and -5.94 kcal/mol. This same tendency, of glycosylation increasing the binding affinity of a compound, is observed in the other categories of compounds, too. The simulations results are presented in Table 15.

Regarding the analysis of the clusters and the different conformations of the molecules, there is no clear binding conformation observed in all the clusters. As resulted from the docking output (Figure 37), it is quite common among chromones for the chromone moiety to be stabilized in the S2 subsite, as it happens in the first cluster of compounds 70 and 75.

Another tendency observed is for the chromone moiety to be positioned between S1 and S1' subsites. In these cases, its orientation is either sideways or more vertical. For example, for compound 70, the first two clusters involve the ligand being bound to S2 subsite, in two different, reversed orientations, which have almost identical interactions and very similar binding energy. In the third cluster, the ligand is orientated vertically. In all the cases, residues His 41 is a pi-pi staking formation hotspot. For compound 71, the first two clusters have a vertical conformation, each in a different direction, while in the third one the ligand is located in the S2 subsite. Another observation is that compounds 72 and 74, which only differ by a hydroxy substitution, have almost identical conformations in their first clusters. As far compound 73 is concerned, in the first cluster, the configuration of the molecule does not resemble any of the ligand in this group, but in the second cluster, the pattern of the chromone moiety inserted in S2 subsite can be observed. Since there is again no consistent pattern nor a point of reference to compare the results, in all the cases where clusters were close in terms of binding energy or resembled one of the described patterns, the cluster is being taken into consideration.

4.3.7. Lignans

The binding scores for the group of lignans (Table 16) are quite promising. Even though they are lower than the native inhibitor N3, they are within a very close range, in which other confirmed inhibitors also belong to. Apart from compound 80, which has the lowest, in absolute value, binding energy (-6.564 kcal/mol), the rest of the lignans have binding energies between -7.175 and -7.835 kcal/mol. Being large molecules, they are able to take up space of the entire active site, potentially being effective inhibitors. It is observed, however, that the lowest binding energy is calculated for the largest of the molecules, indicating that above a certain size, the bulkiness of a molecule prevents from being able to arrange itself in a favorable way in the active site.

Selecting a preferable cluster among the ones calculated through the docking simulation is not always easy, as the criteria is often not clear. For most of the compounds, their interactions with the active site and their orientation in its cavity were similar. The fact that the first clusters of the lignans generally exhibited a similar binding pattern, combined with the binding energy drop in the subsequent clusters, led to the presentations of the first clusters in this discussion (Figure 38). One exception is compound 76, for which all three first clusters will be considered, because they have similar binding energies and orientations that slightly deviate from the pattern of the rest of the lignans. The selection was also more complicated in the case of compound 81, as the first cluster has the best binding energy (-

7.835 kcal/mol) and interacts with the protease with two hydrogen bonds, hydrophobically with one residue and through pi-pi staking with another residue, while the second cluster has a lower binding affinity (-7.569 kcal/mol) but forms 5 hydrogen bonds and one hydrophobic interaction with the enzyme. Since there is no evidence of what the real-life conformation of the molecule is in order to compare it to the docking results, both clusters will be presented.

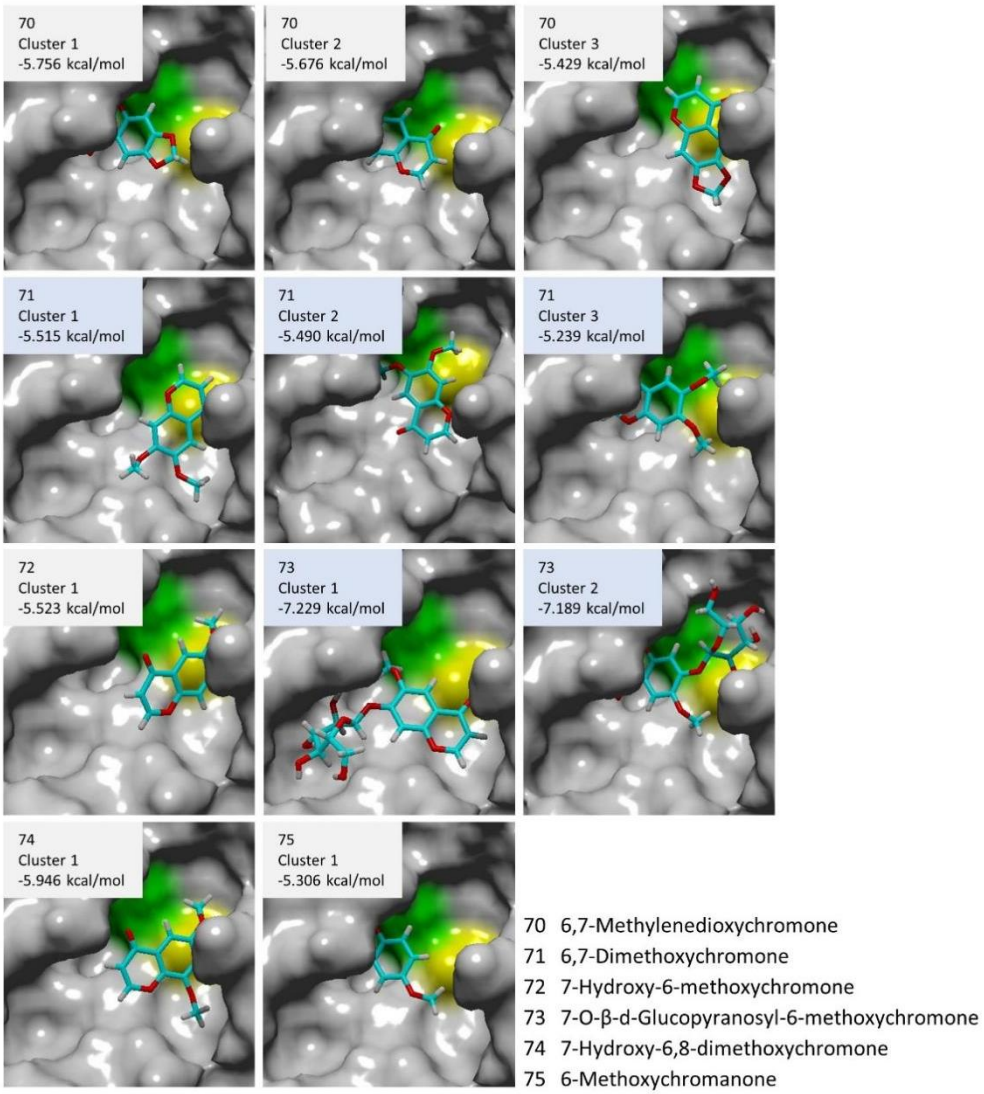


Figure 37: Binding conformations of chromones to the active site of SARS-CoV-2 M^{pro}, as resulted from the docking simulation

Table 14: Molecular docking results for sterols (M^{pro} structure PDB: 6LU7)

Compound		Total cluster	Cluster	Binding energy (kcal/mol)	No of interactions			Total contacting residues
					H-bonds	Hydrophobic	Pi-pi	
65	β-Sitosterol	9	1	-6.559	1 (LEU 141)	1 (GLN 189)	0	LEU 50, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, ALA 191
			2	-6.217	0	1 (MET 49)	0	THR 25, LEU 27, HIS 41, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, PRO 168, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
			3	-6.091	0	1 (THR 25)	0	THR 24, THR 25, THR 26, HIS 41, THR 45, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, GLN 189
			4	-5.584	0	1 (PRO 168)	0	THR 25, THR 26, LEU 27, HIS 41, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, GLY 170, GLN 189
66	Stigmasterol	6	1	-6.651	1 (THR 26)	1 (HIS 41)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, MET 49, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
67	Ergosterol	7	1	-7.452	0	1 (THR 25)	1 (HIS 41)	THR 24, THR 25, THR 26, LEU 27, HIS 41, VAL 42, MET 49, LEU 141, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, ARG 188, GLN 189, THR 190, GLN 192
68	Cerevisterol	7	1	7.257	1 (ASN 142)	1 (HIS 41)	0	HIS 41, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
			2	-6.998	2 (THR 26, GLY 143)	1 (MET 49)	1 (HIS 41)	HIS 41, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
69	β-Daucosterol	7	2	-6.839	2 (LEU 141, GLY 143)	1 (MET 49)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, CYS 44, THR 45, SER 46, GLU 47, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, GLN 189
			3	-6.764	1 (THR 24)	1 (MET 165)	1 (HIS 41)	THR 24, THR 25, THR 26, HIS 41, CYS 44, THR 45, SER 46, MET 49, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, GLN 189, THR 190, ALA 191, GLN 192

Table 15: Molecular docking results for chromones (M^{pro} structure PDB: 6LU7)

Compound		Total cluster	Cluster	Binding energy (kcal/mol)	No of interactions			Total contacting residues
No	Name				H-bonds	Hydrophobic	Pi-pi	
70	6,7-Methylenedioxychromone	4	1	-5.756	0	1 (HIS 41)	1 (HIS 41)	HIS 41, CYS 44, MET 49, PRO 52, TYR 54, ASN 142, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
			2	- 5.676	0	1 (HIS 41)	1 (HIS 41)	HIS 41, MET 49, PRO 52, TYR 54, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
			3	-5.429	1 (GLY 143)	1 (GLU 166)	1 (HIS 41)	HIS 41, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, GLN 189
71	6,7-Dimethoxychromone	3	1	-5.515	2 (GLY 143, GLU 166)	1 (PHE 140)	1 (HIS 163)	HIS 41, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, HIS 172, GLN 189
			2	-5.490	1 (GLU 166)	1 (CYS 145)	1 (HIS 41)	THR 25, LEU 27, HIS 41, MET 49, LEU 141, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, GLN 189
			3	-5.239	0	1 (MET 165)	1 (HIS 41)	LEU 27, HIS 41, MET 49, TYR 54, ASN 142, GLY 143, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189

72	7-Hydroxy-6-methoxychromone	3	1	-5.523	2 (LEU 141, GLY 143)	1 (GLU 166)	1 (HIS 163)	THR 26, LEU 27, HIS 41, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, GLN 189
73	7-O-β-d-Glucopyranosyl-6-methoxychromone	5	1	-7.229	1 (CYS 145)	1 (GLU 166)	1 (HIS 163)	HIS 41, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
			2	-7.189	2 (THR 26, GLY 143)	1 (MET 49)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, CYS 44, MET 49, PRO 52, TYR 54, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
74	7-Hydroxy-6,8-dimethoxychromone	4	1	-5.946	1 (GLY 143)	1 (CYS 145)	1 (HIS 41)	THR 26, LEU 27, HIS 41, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, HIS 172, GLN 189
75	6-Methoxychromanone	4	1	-5.306	0	1 ((HIS 41)	1 (HIS 41)	HIS 41, CYS 44, MET 49, TYR 54, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189

Table 16: Molecular docking results for lignans (M^{pro} structure PDB: 6LU7)

Compound		**Total cluster**	**Cluster**	**Binding energy (kcal/mol)**	**No of interactions**			**Total contacting residues**
No.	Name				**H-bonds**	**Hydrophobic**	**Pi-pi**	
	(−)-Syringaresinol	4	1	-7.175	2 (GLY 143, SER 144)	1 (LEU 167)	1 (HIS 163)	LEU 27, HIS 41, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
			2	-7.113	2 (GLY 143, THR 190)	1 (LEU 167)	1 (HIS 163)	LEU 27, HIS 41, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
			3	-7.091	0	1 (MET 49)	1 (HIS 41)	THR 24, THR 25, THR 26, LEU 27, HIS 41, CYS 44, THR 45, SER 46, MET 49, PRO 52, TYR 54, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, VAL 186, ASP 187, ARG 188, GLN 189
77	Syringaresinol 4-O-beta-D-glucopyranoside	3	1	-7.684	1 (GLU 166)	1 (MET 49)	1 (HIS 41)	THR 24, THR 25, THR 26, LEU 27, HIS 41, CYS 44, THR 45, SER 46, MET 49, ASN 119, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ASP 187, ARG 188, GLN 189, THR 190, GLN 192,
78	Episyringaresinol-4''-O-β-D-glucopyranoside	4	1	-7.499	2 (GLU 166, ARG 188)	1 (THR 25)	1 (HIS 41)	THR 24, THR 25, THR 26, LEU 27, HIS 41, CYS 44, THR 45, SER 46, MET 49, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, ASP 187, ARG 188, GLN 189, THR 190, GLN 192
79	Acanthoside B	4	1	-7.699	3 (THR 24, GLU 166, THR 190)	1 (THR 25)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, CYS 44, THR 45, SER 46, MET 49, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, ASP 187, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
80	Erythro-1-(4-O-β-d-glucopyranosyl-3,5-dimethoxyphenyl)-2-syringaresinoxyl-propane-1,3-diol	7	1	-6.564	3 (GLY 143, SER 144, GLN 189)	1 (GLU 166)	0	THR 24, THR 25, THR 26, HIS 41, THR 45, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, PRO 168, HIS 172, GLN 189
81	Longifloroside B	4	1	-7.835	2 (CYS 145, ARG 188)	1 (LEU 27)	1 (HIS 41)	THR 24, THR 25, THR 26, LEU 27, HIS 41, VAL 42, CYS 44, THR 45, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, PRO 168, ARG 188, GLN 189, THR 190, GLN 192
			2	- 7.569	5 (THR 26, LEU 141 x2, GLY 143, CYS 145)	1 (THR 26)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, THR 45, SER 46, MET 49, TYR 118, ASN 119, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, GLN 189

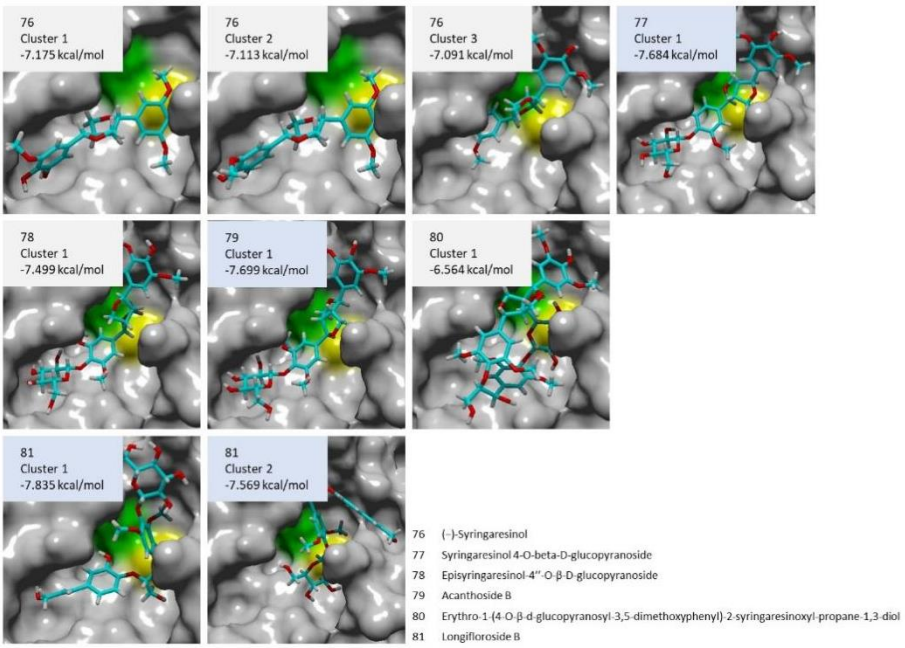


Figure 38: Binding conformations of lignans to the active site of SARS-CoV-2 M^{pro}, as resulted from the docking simulation

4.3.8. Oleanane Triterpenoid Saponins

The saponins found in *Salicornia* extracts are all large molecules, especially in comparison to the other ligands screened. Their size contributes to their taking up almost the entire space of the active site cavity, indicating the possibility of effective inhibition. Indeed, saponins resulted in binding energies ranging from -7.018 to -8.614 kcal/mol, with a lot of them being better than the one of native inhibitor N3. The compound with the highest affinity to the active site (-8.614 kcal/mol) is compound 83, Boussingoside A1, a glucuronate of akebonic acid, while compound 88, another akebonic acid derivative substituted with a glucopyranosyl-xylopyranosyl-glucuronic acid moiety and another single glucuronic acid group, being the largest in size ligand of this group, also resulted in a promising binding energy of -8.498 kcal/mol. The second-best hit was is compound 98(-8.559 kcal/mol), an oleanolic acid derivative substituted by a methyl ester of glucuronic acid, while the respective acid is also among the top three compounds in terms of binding energy. The binding energy and contacting residues for each compound can be found in Table 17.

In all the cases investigated, compounds substituted with a glucuronic acid moiety had better binding affinity to the active site compared to their parent compounds. On the contrary, additional glycosylation at the carboxyl of the triterpenoid moiety resulted in a reduction of the binding affinity. There are two cases where compounds and their methyl esters were screened (compounds 97-98 and 99-100) in both of which methylation increased the binding score of the compound, although not significantly. This tendency was also observed in the group of caffeoylquinic acids and derivatives, but there were also examples contradicting this hypothesis. Another general observation is that compounds if this group do not form pi-pi interactions with the residues of M^{pro}. Regarding the spatial arrangement of the ligands in the active site, all of them obstruct access to the catalytic dyad, extending to the S1' subsite. The majority of them also block S4 subsite, while some are shifted towards the side of S1. Further analysis of the ligand-protein complexes produced by the simulation follows.

All the saponins are structurally very closely related. However, they could be divided into three groups based on their common pentacyclic triterpenoid backbone. Similar behavior in the binding to the active site could be expected among all the saponins, but dividing them to subgroups can be a starting point to facilitate the analysis of the clusters. To begin with, one category is formed by **akebonic acid and its derivatives**. Two configurations are more often observed: one with the triterpenoid group being found in S1' subsite and extending upwards, out of the cavity of the active site, such as in the case of the first cluster of compound 82, and one with the triterpenoid group covering the vast part of the active site cavity, being located in the middle, diagonally between S1' and S4 subsites, as in the third cluster of the same compound. In this example, the third cluster has a lower binding energy by 0.8 kcal/mol. On the other hand, for its glucuronidated form, compound 83, the first conformation is found in the third cluster, which has a considerably lower binding energy, while the second one in the first. In the second cluster for this molecule, another orientation is observed, which involves the glucuronic acid moiety being located in the center of the active site, where its oxygen and hydroxy groups allow the formation of hydrogen bonds with residues Leu 141, Cys 145 and Gln 189, and the triterpenoid part extending upwards. This pattern, of glucuronic acid or the respective substitution in the same position, stabilized between S1 and S1' subsites and the rest of the molecule facing vertically upwards, or sideways, is found in one of the first clusters of all the other ligands. Apart from compound 84, for which the first cluster has a conformation that is not comparable to the rest of the compounds, the second resembles the second cluster of compound 82, and only the third mimics the previously described configuration, all the first clusters of the rest of the ligand of this subgroup exhibit this tendency. Therefore, taking additionally into consideration the reduction of the binding energy in the rest of the clusters, only the first ones will be presented for these compounds, with the exception of compound 87 for which the second cluster has a very similar binding energy and conformation.

Gypsogenin and its derivatives, compounds 91-93, could also be grouped based on their structural similarity. The derivatives do not have the same direction when bound to the protease as any of the first clusters for gypsogenin. Only the second cluster of compound 92 resembles the first cluster of compound 91, but its binding affinity is lower by 1 kcal/mol, which is a considerable amount. Although there are not great similarities between the compounds with one another, the binding complexes of the first three clusters of compound 91 show relevance to the ones for compound 82. Similarly, compounds 92 and 93 acquire orientations comparable to the ones of compounds 83 and 84. Another observation is that the derivatives of gypsogenin, substituted with glucuronic acid in one case and both glucuronic acid and glucose in the second, have improved binding energy compared to gypsogenin (-7.833 and -7.740 kcal/mol for compounds 92 and 93 respectively, as opposed to -7.061 kcal/mol for gypsogenin).

Table 17: Molecular docking results for oleanane triterpenoid saponins (M^{pro} structure PDB: 6LU7)

Compound		Total cluster	Cluster	Binding energy (kcal/mol)	No of interactions			Total contacting residues
No.	**Name**				**H-bonds**	**Hydrophobic**	**Pi-pi**	
82	Akebonic acid	7	1	-7.512	1 (GLY 143)	1 (MET 165)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, THR 45, SER 46, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, GLN 189
			2	-7.207	1 (THR 26)	1 (THR 25)	0	THR 25, THR 26, LEU 27, HIS 41, MET 49, LEU 141, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, GLN 189
			3	-6.739	0	1 (PRO 168)	0	HIS 41, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, GLY 170, GLN 189, THR 190, ALA 191
83	Boussingoside A1	7	1	-8.614	1 (THR 24)	1 (PRO 168)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, CYS 44, THR 45, SER 46, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, GLN 189, THR 190, ALA 191
			2	-8.158	3 (LEU 141, CYS 145, GLN 189)	1 (MET 49)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, VAL 42, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, GLN 189
			3	-7.636	3 (ASN 142, GLY 143, THR 24)	1 (MET 49)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, THR 45, SER 46, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, GLN 189
84	Boussingoside A2	7	1	-8.35	2 (GLU 166, THR 190)	1 (THR 25)	0	THR 24, THR 25, THR 26, LEU 27, ASN 28, HIS 41, SER 46, MET 49, TYR 118, ASN 119, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
			2	-7.658	1 (GLY 143, THR 190)	1 (THR 25)	0	THR 25, THR 26, LEU 27, ASN 28, HIS 41, MET 49, TYR 118, ASN 119, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, GLN 189, THR 190, ALA 191, GLN 192
			3	-7.445	1 (GLY 23, CYS 145)	1 (MET 49)	0	GLN 19, THR 21, GLY 23, THR 24, THR 25, THR 26, MET 49, LEU 67, ASN 119, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, GLN 189
85	3β-Hydroxy-23-oxo-30-noroleana-12,20(29)-diene-28-oic acid 3-O-β-d-glucuronopyranosyl-28-O-β-d-glucopyranoside	12	4	-7.108	3 (THR 26 x2, CYS 145)	1 (GLU 166)	0	THR 25, THR 26, LEU 27, HIS 41, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, THR 169, GLY 170, GLN 189
86	30-Norhederagenin 3-O-β-d-glucuronopyranosyl-28-O-β-d-glucopyranoside	12	2	-7.018	1 (GLU 166)	1 (MET 49)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, THR 45, SER 46, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, GLN 189
87	3-O-[β-D-Glucuronopyranosyl-6′-O-methyl ester]-30-norolean-12,20(29)-dien-28-O-[β-D-glucopyranosyl]ester	9	1	-7.443	0	1 (HIS 41)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, THR 45, SER 46, GLU 47, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, GLN 189
			2	-7.289	0	1 (MET 49)	0	THR 21, GLY 23, THR 24, THR 25, THR 26, LEU 27, HIS 41, VAL 42, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, GLN 189
88	Salieuropaea A	7	1	-8.498	2 (GLY 23, GLU 166)	1 (CYS 145)	0	THR 21, GLY 23, THR 24, THR 25, THR 26, LEU 27, HIS 41, SER 46, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ASP 187, ARG 188, GLN 189, THR 190, ALA 191, GLN 192

89	Salbige A	8	1	-8.073	2 (LEU 141, GLN 189)	1 (GLN 189)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, THR 45, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, GLN 192
90	Salbige B	9	1	-8.105	2 (SER 144, GLN 189)	1 (ASN 142)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, VAL 42, CYS 44, SER 46, MET 49, PRO 52, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189
91	Gypsogenin	6	1	-7.061	0	1 (PRO 168)	0	THR 25, HIS 41, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, GLY 170, GLN 189, THR 190, ALA 191, GLN 192
			2	-6.685	1 (GLY 143)	1 (MET 49)	0	THR 24, THR 25, THR 26, HIS 41, THR 45, SER 46, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, GLN 189
			3	-6.638	2 (LEU 141, CYS 145)	1 (GLY 143)	0	THR 25, THR 26, LEU 27, HIS 41, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, GLN 189, THR 190
92	Gypsogenin 3-O-β-d-glucuronopyranoside	5	1	-7.833	1 (GLN 189)	1 (MET 49)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, THR 45, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, GLN 189
			2	-6.846	1 (CYS 145)	1 (GLU 166)	0	THR 25, THR 26, LEU 27, HIS 41, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, GLN 189
93	Gypsogenin 3-O-β-d-glucuronopyranosyl-28-O-β-d-glucopyranoside	11	1	-7.740	2 (THR 26, GLN 189)	1 (THR 25)	1 (HIS 41)	THR 24, THR 25, THR 26, LEU 27, HIS 41, SER 46, MET 49, TYR 118, ASN 119, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, ARG 188, GLN 189, THR 190, GLN 192
			2	-7.167	0	1 (GLY 143)	1 (HIS 41)	THR 24, THR 25, THR 26, LEU 27, HIS 41, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, GLN 189
94	Oleanolic acid	7	1	-7.087	0	1 (PRO 168)	0	HIS 41, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, GLN 189, THR 190, ALA 191, GLN 192
			2	-7.040	1 (GLN 189)	1 (GLY 143)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, MET 49, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, GLN 189
			3	-6.761	2 (LEU 141, CYS 145)	1 (MET 165)	0	THR 25, HIS 41, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, GLN 189, THR 190
95	Oleanolic acid 3-O-β-D-glucopyranoside	6	1	-7.987	0	1 (MET 49)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, THR 45, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, GLN 189
			2	-7.859	2 (THR 24, THR 26)	1 (PRO 168)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, CYS 44, THR 45, SER 46, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, GLN 189, THR 190, ALA 191, GLN 192
96	Oleanolic acid 28-O-β-D-glucopyranoside	10	1	-8.194	0	1 (MET 49)	0	THR 25, THR 26, LEU 27, HIS 41, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, GLN 189, THR 190, ALA 191, GLN 192
97	Calenduloside E	7	1	-8.425	0	1 (PRO 168)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, CYS 44, THR 45, SER 46, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, GLY 170, GLN 189, THR 190, ALA 191, GLN 192

			2	-8.300	0	1 (MET 49)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, THR 45, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, GLN 189
98	Calenduloside E 6'-methyl ester	8	1	-8.559	1 (THR 26)	1 (MET 49)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, CYS 44, THR 45, SER 46, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, GLY 170, GLN 189, THR 190, ALA 191, GLN 192
			2	-8.542	0	1 (MET 49)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, THR 45, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, GLN 189
99	Chikusetsusaponin IVa	13	1	-7.773	3 (THR 26, ASN 119, GLU 166)	1 (THR 25)	0	THR 24, THR 25, THR 26, HIS 41, SER 46, MET 49, ASN 119, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
				-7.530	1 (HIS 163)	1 (THR 26)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, THR 45, SER 46, MET 49, ASN 119, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, GLN 189
				-7.439	1 (HIS 164)	1 (THR 26)	0	GLN 19, THR 21, GLY 23, THR 24, THR 25, THR 26, HIS 41, SER 46, MET 49, LEU 67, GLN 69, ASN 119, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, GLN 189
100	Chikusetsusaponin IVa methyl ester	9	1	-7.920	1 (GLY 23)	1 (HIS 172)	0	THR 21, GLY 23, THR 24, THR 25, THR 26, LEU 27, HIS 41, VAL 42, SER 46, MET 49, ASN 119, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, GLN 189
			2	-7.781	2 (THR 26, ARG 188)	1 (THR 25)	0	THR 24, THR 25, THR 26, HIS 41, SER 46, MET 49, TYR 118, ASN 119, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, ARG 188, GLN 189, THR 190, GLN 192
101	3β,29-Dihydroxy-olean-12-en-28-oic acid 28-O-β-D-glucopyranosyl ester	10	1	-7.994	0	1 (ASN 142)	0	THR 25, THR 26, LEU 27, HIS 41, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, GLN 189, THR 190, ALA 191, GLN 192
102	Zygophyloside K	9	1	-7.496	2 (THR 26, ASN 119)	1 (THR 25)	0	THR 24, THR 25, THR 26, HIS 41, SER 46, MET 49, TYR 118, ASN 119, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, ARG 188, GLN 189, THR 190, GLN 192
			2	-7.185	4 (THR 24, GLU 166, GLN 189, THR 190)	1 (MET 49)	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, MET 49, ASN 119, PHE 140, LEU 141, ASN 142, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ARG 188, GLN 189, THR 190, ALA 191, GLN 192
			3	-7.02	0	1 (THR 26)		GLN 19, THR 21, GLY 23, THR 24, THR 25, THR 26, HIS 41, MET 49, ASN 119, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, GLN 189

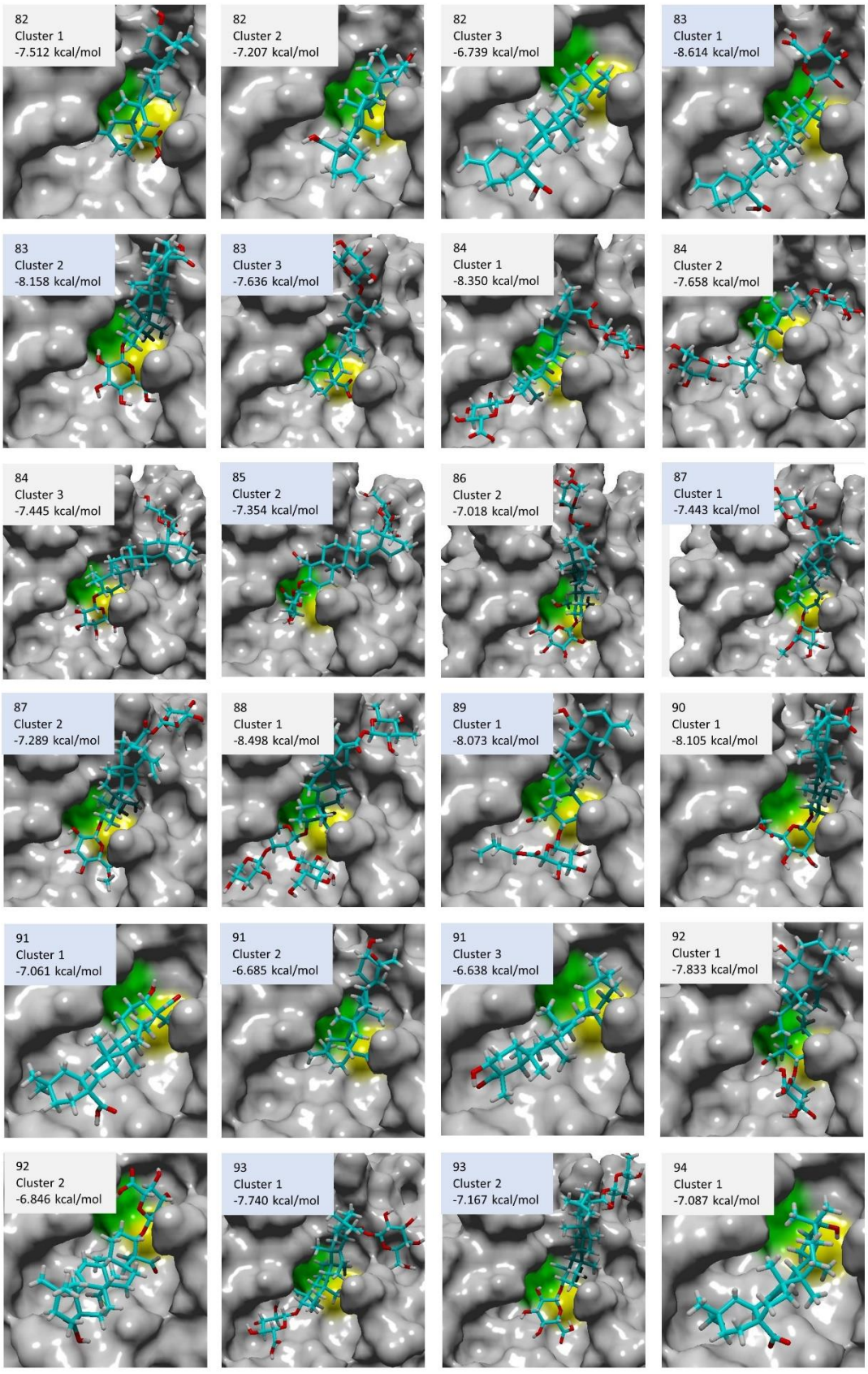
Another distinctive category consists of **oleanolic acid and its derivatives**, compounds 94-100. Again, the unsubstituted compound shows similar orientations in the first clusters as compounds 82 and 91. Moreover, compounds 95, 97 and 98, along with compound 92, which is also substituted with one glucuronic acid, resulted in remarkably similar configurations. Respectively, the first two clusters of compounds 99 and 100 are closely related, although they appear in reversed order. Lastly, compounds 101 and 102 also have a slightly different backbone, but appear to block the active site in an analogous manner, both when compared to one another but also when compared to structurally alike molecules (e.g. compounds 99 and 102). In this subgroup of compounds, too, the pattern of increased binding affinity for the substituted derivatives as opposed to the parent compound is observed. More specifically, while oleanolic acid has a binding energy of -7.087 kcal/mol in the first cluster, which is exceeded by its glucosides, compounds 95 (-7.987 kcal/mol) and 96 (-8.194 kcal/mol), its glucuronides, compounds 97 (-8.425 kcal/mol) and 98 (-8.559 kcal/mol) and its derivatives that combined both these substitutions, compounds 99 (-7.773 kcal/mol) and 100 (-7.920 kcal/mol). The data also suggests that the bulkier, double-substituted derivatives have a lower binding affinity to the active site compared to the single-substituted ones, indicating that the larger size of a molecule, although seemingly complimentary to its inhibitory potential, can be limitating for its binding after a certain point.

Overall, all saponins show inhibitory potential worth of further investigation. Some of them are also specific to *Salicornia* species. Particularly compound 85 has been detected in *Salicornia* for the first time (Kim et al. 2012), but exhibits a moderate binding energy of -7.108 kcal/mol. Moreover, compound 88, with a very encouraging antiviral activity indication (-8.498 kcal/mol) as resulted from this initial screening, has only been detected in *S. europaea* (Lyu et al. 2018). Promising results were also produced for compounds 89, 90 and 101(-8.073, -8.105 and -7.994 kcal/mol respectively), which have been isolated from στη halophyte for the first time (Zhao et al. 2014; Yin et al. 2012).

The clusters presented below (Figure 39) are the ones that fall into one of the patterns observed, while not having a major difference in the binding affinity compared to the first cluster.

4.3.9. Other compounds

This category includes molecules that could not be categorized elsehow. The majority of the compounds in this group, especially the smaller in size (106, 107, 111, 113, 114) resulted in quite low binding affinity, below -6 kcal/mol, as seen in Table 18. The only point of reference for the compounds of this groups is for pyrogallol and ellagic acid, which have been also tested with the same software and receptor and resulted in binding energies equal to -4.9 and -8.4 kcal/mol respectively (Murugesan et al. 2021). The reported binding energy for pyrogallol is almost identical to the one calculated in this work, though ellagic acid appears to have a higher binding affinity in literature. For ellagic acid in particular, the orientation of the molecule inside the active site has been depicted. It appears to be located in the S1 subsite, almost vertical to the surface of the protease, however none of the resulting clusters of the present simulation is comparable (Murugesan et al. 2021). Therefore, the clusters that were considered for the compounds of these category where the first ones, except for the cases were the difference in the binding affinity in the first two clusters was marginal.


82
Cluster 1
-7.512 kcal/mol
82
Cluster 2
-7.207 kcal/mol
82
Cluster 3
-6.739 kcal/mol
83
Cluster 1
-8.614 kcal/mol
83
Cluster 2
-8.158 kcal/mol
83
Cluster 3
-7.636 kcal/mol
84
Cluster 1
-8.350 kcal/mol
84
Cluster 2
-7.658 kcal/mol
84
Cluster 3
-7.445 kcal/mol
85
Cluster 2
-7.354 kcal/mol
86
Cluster 2
-7.018 kcal/mol
87
Cluster 1
-7.443 kcal/mol
87
Cluster 2
-7.289 kcal/mol
88
Cluster 1
-8.498 kcal/mol
89
Cluster 1
-8.073 kcal/mol
90
Cluster 1
-8.105 kcal/mol
91
Cluster 1
-7.061 kcal/mol
91
Cluster 2
-6.685 kcal/mol
91
Cluster 3
-6.638 kcal/mol
92
Cluster 1
-7.833 kcal/mol
92
Cluster 2
-6.846 kcal/mol
93
Cluster 1
-7.740 kcal/mol
93
Cluster 2
-7.167 kcal/mol
94
Cluster 1
-7.087 kcal/mol

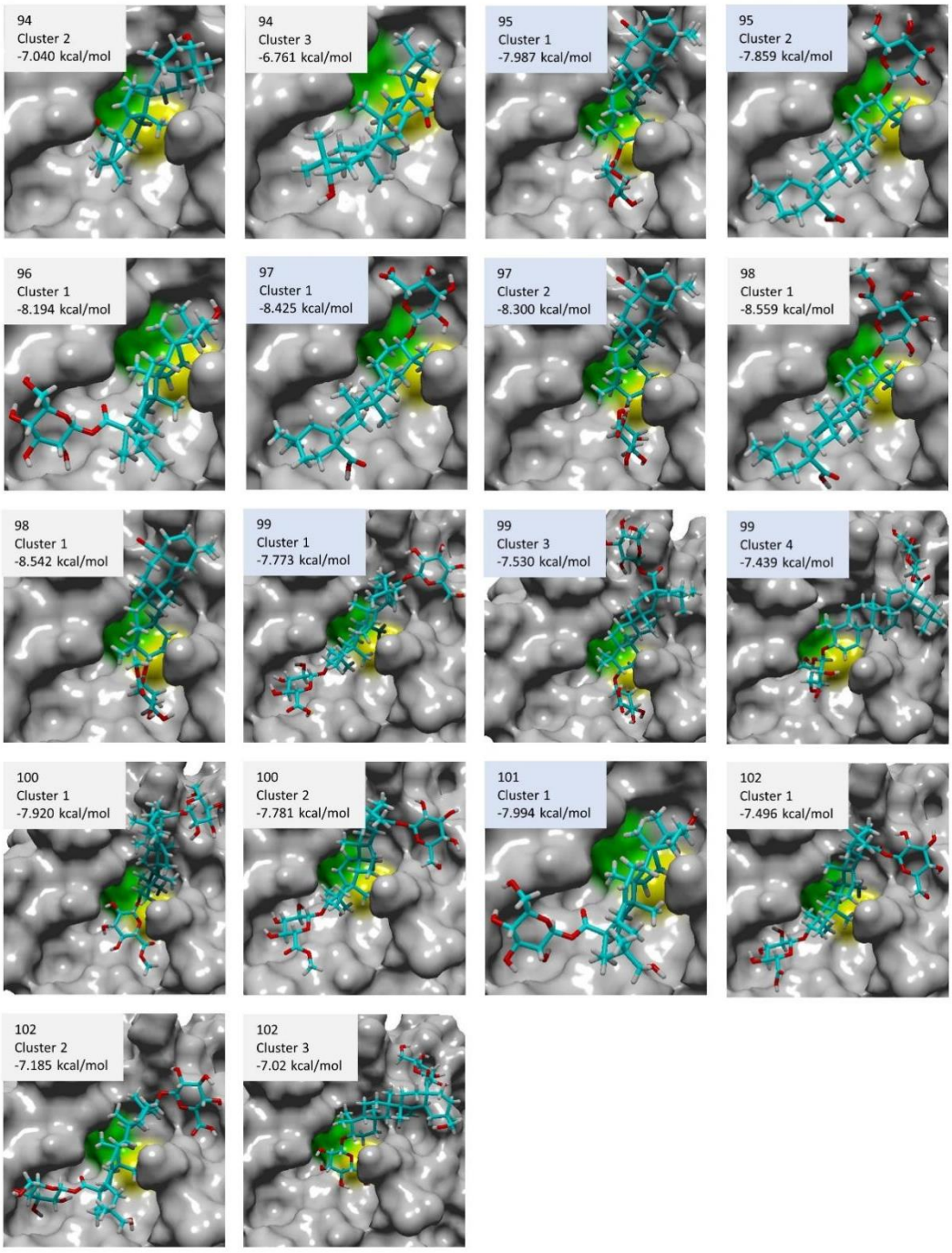


Figure 39: Binding conformations of oleanane triterpenoid saponins to the active site of SARS-CoV-2 M^{pro}, as resulted from the docking simulation

Pheophorbide A and its derivatives (compounds 113-115) yielded the best results within this category, with compound 115 having a binding energy of -8.103 kcal/more and therefore the highest score. The three compounds are structurally closely related, something that is also depicted in their complexes with M^{pro}. In the case of compound 114, its first two clusters have essentially the same binding energy, and represent two slightly different orientations, one of which is closer to compound 113 and one to 115. Regarding the rest of the compounds, the lack of distinct structural similarities does not allow many comparisons. Only pyrogallol and vanillin have quite similar structures, possessing a main substituted phenolic ring, and result in very similar binding energies, as well as positions in the active site cavity (-4.887 and -4.954

kcal/mol respectively), stabilized through pi-pi interactions between the phenolic ring and His 163. Compounds 109 and 110 are also alike, but the difference in their size causes remarkably different conformations, as given by the simulation output. The only other observation that can be made is that, overall, the largest molecules result in better binding energies. Indications towards this hypothesis were also present in the previous categories, with the data showing that increase in size improves binding affinity, until a certain point when perhaps the size of the molecule does not allow it to properly enter the binding site cavity.

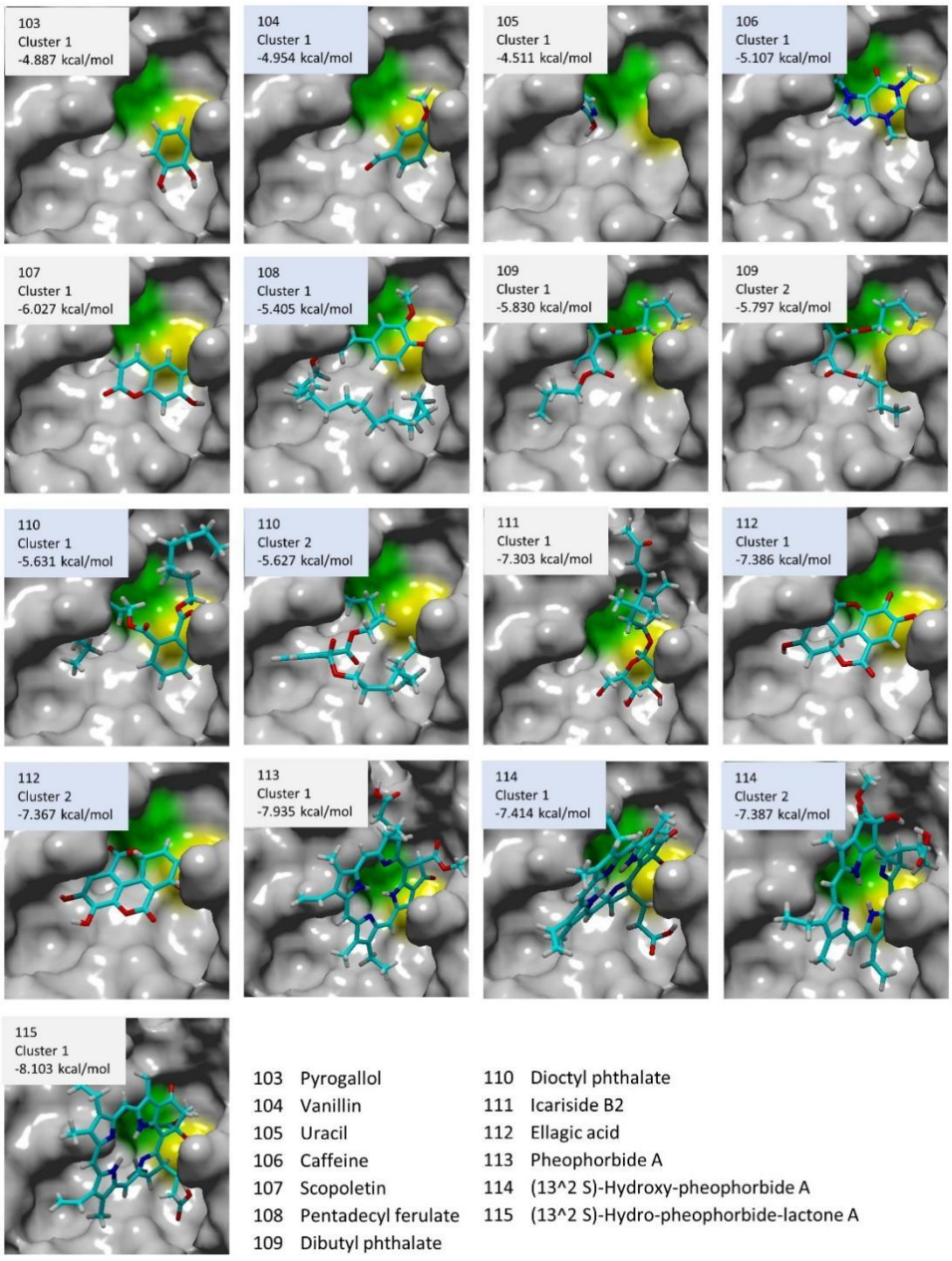


Figure 40: Binding conformations of the remaining uncategorized Salicornia compounds to the active site of SARS-CoV-2 M^{pro}, as resulted from the docking simulation

Table 18: Molecular docking results for uncategorized structures (M^{pro} structure PDB: 6LU7)

Compound		Total cluster	Cluster	Binding energy (kcal/mol)	No of interactions			Total contacting residues
No.	Name				H-bonds	Hydrophobic	Pi-pi	
103	Pyrogallol	3	1	-4.887	0	1 (CYS 145)	1 (HIS 163)	PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, HIS 172, GLN 189
104	Vanillin	4	1	-4.954	2 (GLY 143, GLU 166)	1 (ASN 142)	1 (HIS 163)	PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, HIS 172
105	Uracil	4	1	-4.511	0	1 (MET 49)	1 (HIS 41)	HIS 41, CYS 44, MET 49, PRO 52, TYR 54, HIS 164, MET 165, PHE 181, ASP 187, ARG 188, GLN 189
106	Caffeine	2	1	-5.107	1 (GLY 143)	1 (MET 49)	1 (HIS 41)	LEU 27, HIS 41, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, GLN 189
107	Scopoletin	2	1	-6.027	1 (LEU 141)	1 (GLU 166)	1 (HIS 163)	PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, MET 165, GLU 166, HIS 172, GLN 189
108	Pentadecyl ferulate	7	1	-5.405	1 (GLY 143)	1 (HIS 41)	1 (HIS 41)	THR 25, LEU 27, HIS 41, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, LEU 167, PRO 168, HIS 172, ASP 187, ARG 188, GLN 189, THR 190, GLN 192
109	Dibutyl phthalate	6	1	-5.830	0	1 (MET 165)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, MET 49, TYR 54, GLY 143, CYS 145, HIS 164, MET 165, GLU 166, LEU 167, PRO 169, ASP 187, ARG 188, GLN 189, THR 190, GLN 192
			2	-5.797	0	1 (MET 165)	1 (HIS 41)	THR 25, THR 26, LEU 27, HIS 41, VAL 42, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, ASP 187, ARG 188, GLN 189
110	Dioctyl phthalate	7	1	-5.631	1 (GLY 143)	1 (HIS 41)	1 (HIS 163)	THR 24, THR 25, THR 26, LEU 27, HIS 41, MET 49, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, VAL 186, ASP 187, ARG 188, GLN 189, THR 190
			2	-5.627	1 (GLU 166)	1 (MET 165)	0	HIS 41, CYS 44, MET 49, PRO 52, TYR 54, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, ASP 187, ARG 188, GLN 189, THR 190, GLN 192
111	Icariside B2	7	1	-7.303	2 (GLY 143, SER 144)	0	0	THR 24, THR 25, THR 26, LEU 27, HIS 41, CYS 44, THR 45, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, GLN 189
112	Ellagic acid	5	1	-7.386	1 (GLU 166)	1 (MET 165)	1 (HIS 41)	LEU 27, HIS 41, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, ASP 187, ARG 188, GLN 189, THR 190, GLN 192
			2	-7.367	0	1 (MET 165)	1 (HIS 41)	HIS 41, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 164, MET 165, GLU 166, ARG 188, GLN 189, THR 190, GLN 192
113	Pheophorbide A	7	1	-7.935	0	1 (GLU 166)	0	THR 24, THR 25, THR 26, LEU 27, CYS 44, THR 45, SER 46, MET 49, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, MET 165, GLU 166, GLN 189
114	(13^2 S)-Hydroxy-pheophorbide A	7	1	-7.414	1 (LEU 141)	1 (LEU 27)	0	THR 25, THR 26, LEU 27, HIS 41, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145, HIS 163, HIS 164, MET 165, GLU 166, HIS 172, GLN 189
			2	-7.387	1 (ASN 142)	1 (GLU 166)	1 (HIS 163)	THR 24, THR 25, THR 26, LEU 27, HIS 41, THR 45, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143 , SER 144, CYS 145, HIS 163, MET 165, GLU 166, HIS 172, GLN 189
115	(13^2 S)-Hydro-pheophorbide-lactone A	8	1	-8.103	1 (HIS 41)	1 (LEU 27)	0	THR 25, THR 26, LEU 27, HIS 41, SER 46, MET 49, PHE 140, LEU 141, ASN 142, GLY 143, SER 144, CYS 145 , HIS 163, HIS 164, MET 165, GLU 166, HIS 172, GLN 189

4.4. Evaluation of inhibitory effect against M^{pro} *in vitro*

4.4.1. Phenolic acids

An enzyme inhibition assay was performed in order to further investigate the anti-SARS-CoV-2 potential of the phytochemicals detected in *Salicornia* sp. Since it was not possible for all the substances to be tested, a selection was made taking into consideration the binding score of the compounds and how well they represent the structural group they belong to. Another major limiting factor was their commercial availability. The positive control substance provided in the kit was known inhibitor GC376. The results for the inhibitor are presented in Figure 41, and led to the calculation of an IC_{50} value of 0.454 μM, which is remarkably lower than the lowest IC_{50} calculated for the screened compounds.

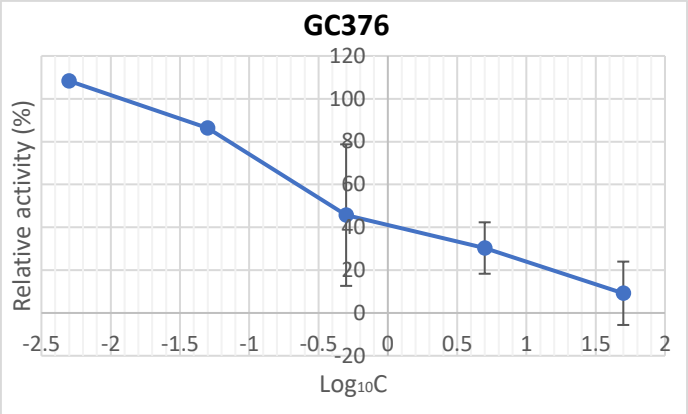


Figure 41: Enzyme inhibition assay results for inhibitor GC376: Relative activity of the enzyme as a function of the log10 of the different concentrations (C in μM)

Considering all the above, ferulic and rosmarinic acid were selected from the group of hydroxycinnamic acids, because ferulic acid is the main hydroxycinnamic acid found in plant cell walls (Mathew and Abraham 2004) and also has high binding energy compared to the other hydroxycinnamic acids, while rosmarinic acid is a bulkier hydroxycinnamic acid derivative that exhibits a significantly higher binding affinity to M^{pro} than the rest of the compounds of the group. The relative activity of M^{pro} for the different concentrations of the substances tested is presented in Figure 40. The results confirm the indication provided by the docking simulation that the substituted hydroxycinnamic acid has better inhibitory potential than a smaller, unsubstituted one. Ferulic acid shows no inhibitory effect for concentrations lower than 250 μM, but starts effectively reducing the activity of the enzyme to almost 25% at the higher concentration tested (5000 μM). Through equation (3), the IC_{50} value was calculated equal to 3090.99 μM. Rosmarinic acid exhibits inhibitory effect even for concentrations exceeding 100 μM and has a significantly lower IC_{50} value of 801.45 μM.

The docking simulation results for hydroxybenzoic acids showed very similar binding energy for all the compounds, therefore, only one compound was selected amongst them. Gallic acid was chosen as the one with the best docking score. As seen in Figure 42, it has moderate inhibitory effect, with an IC_{50} value of 4424.22 μM. The highest concentration tested, 5000 μM, resulted in inhibition slightly above 50%, therefore higher concentrations would be needed to acquire the full inhibitory curve and have a broader perspective of its activity. It is also observed that gallic acid had the lowest binding energy among the compounds tested and resulted in the highest IC_{50} value.

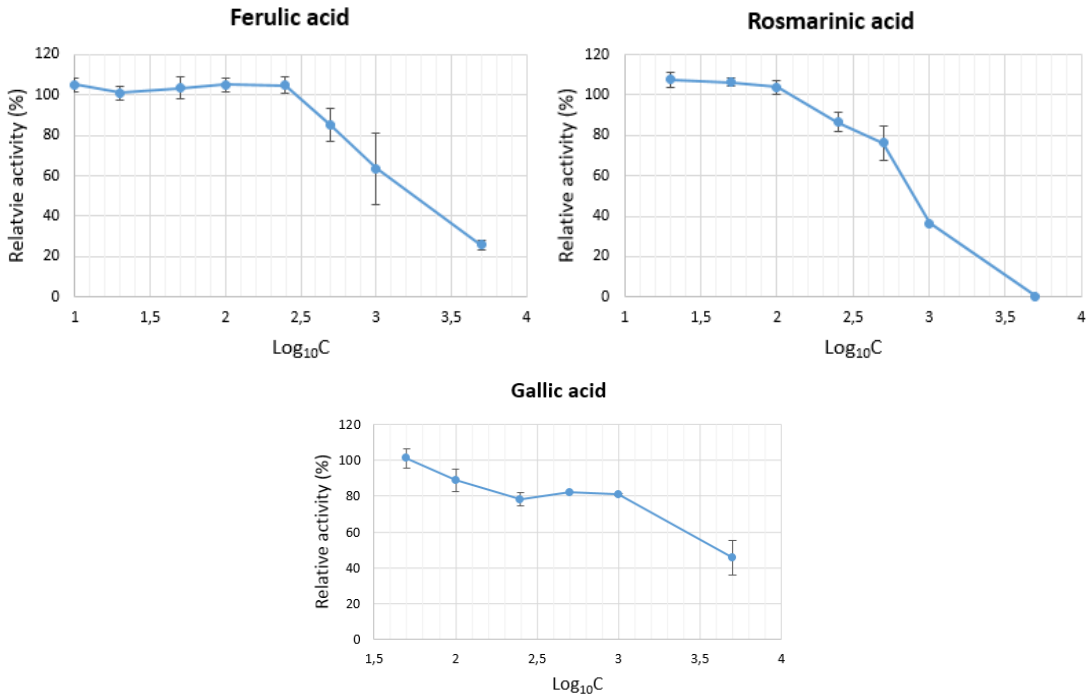


Figure 42: Enzyme inhibition assay results for hydroxycinnamic and hydroxybenzoic acids: Relative activity of the enzyme as a function of the log_{10} of the different concentrations (C in μM)

Regarding caffeoylquinic acids and their derivatives, quinic acid was selected as a basic parent compound for the rest of the compounds, although its binding energy is not high. Chlorogenic acid, 3,4,-Dicaffeoyl quinic and 3,5-Dicaffeoyl quinic acids were decided to be tested because they are also some basic quinic and caffeic acid derivatives, with the two dicaffeoyl quinic acids also exhibiting a very high binding energy (3,5-Dicaffeoyl quinic acid is the best hit among the group in terms of binding energy).

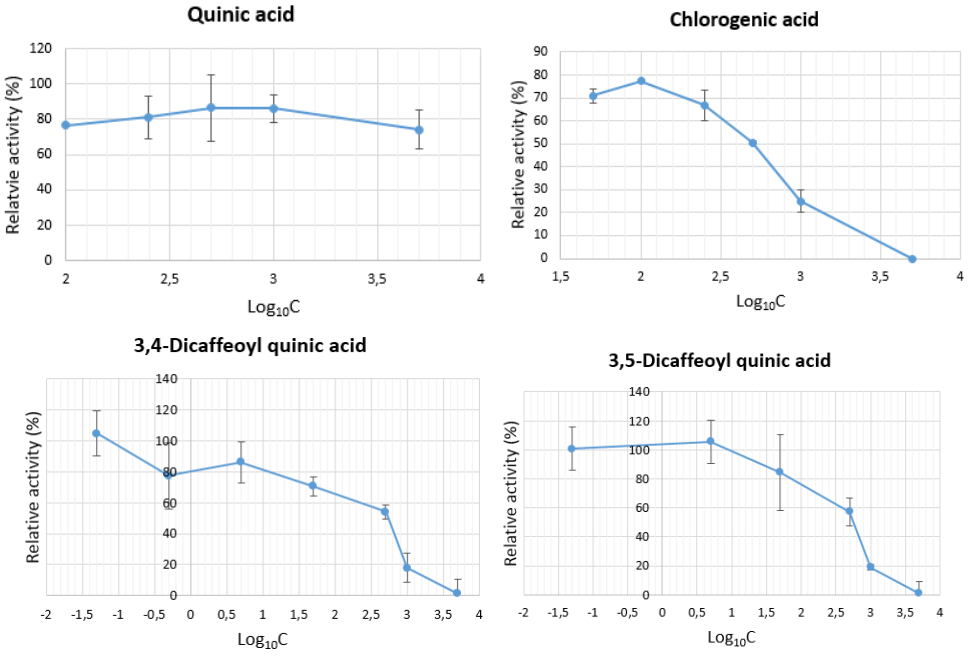


Figure 43: Enzyme inhibition assay results for caffeoylquinic acids: Relative activity of the enzyme as a function of the log_{10} of the different concentrations (C in μM)

Quinic acid showed no substantial inhibitory effect for the concentrations tested. It showed to inhibit around 20% of the enzymatic activity for all the concentrations tested, thus testing at higher concentrations would be needed to confirm the tendency of its inhibitory effect. Chlorogenic, 3,4- and 3,5- dicaffeoyl quinic acids showed good antiviral potential, as also portraited in the IC_{50} values calculated, 546.07, 503.59 and 597.81 μM, respectively. All three compounds completely blocked the action of M^{pro} at the highest concentration of 5000 μM, while chlorogenic acid showed inhibitory effect even at the lowest concentration tested.

4.4.2. Flavonoids and flavanones

To begin with, myricetin and kaempferol were selected to be tested in vitro since they have also been tested in other studies and therefore can be a measure of comparison. Particularly myricetin is identified as a covalent inhibitor of M^{pro} with IC_{50}= 0.22 μM (Kuzikov et al. 2021) while for kaempferol IC_{50}=34.46, as calculated through a CPE inhibition assay on Vero E6 cells (Khan et al. 2021). The values for the IC_{50} calculated in this study are considerably higher, 505.27 μM and 341.85 μM respectively. However, a valid comparison cannot be made since as IC_{50} depends on the concentration of the substrate and a different assay was used, so the parameters affecting the calculation of the IC_{50} are different. From this work it can be deduced that kaempferol has a better inhibitory potential than myricetin, since the enzyme has lower relative activity when kaempferol is used as inhibitor for the respective concentrations (Figure 44). It is also remarkable how kaempferol has the second lowest IC_{50} value among the screened phytochemicals, while being smaller in size compared to the other flavonoids and flavanones that yielded promising results. Taking into consideration the appearing increase of a compound's inhibitory efficacy when it is substituted by sugars, which is also reinforced by the results for other compounds mentioned below, it could be interesting to further investigate sugar derivatives of kaempferol for their inhibitory effect against the protease.

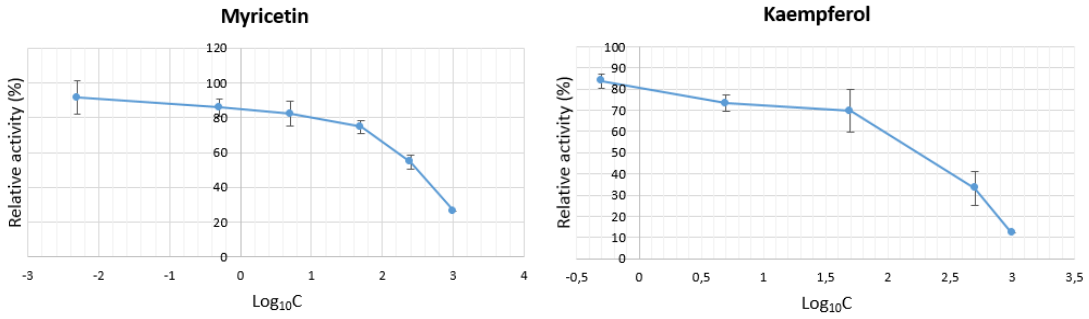


Figure 44: : Enzyme inhibition assay results for already tested inhibitors myricetin and kaempferol: Relative activity of the enzyme as a function of the log_{10} of the different concentrations (C in μM)

Quercetin was also decided to be part of the *in vitro* investigation, since it is a major flavonoid with many derivatives detected in *Salicornia*. Thus, testing it and some of its derivatives (isoquercetin, rutin, isorhamnetin , isorhamnetin-3-glucoside and isorhamnetin-3-rutinoside) could also provide insight of the structure-activity relationship of the compounds. Quercetin showed no particular activity at concentrations below 500 μM, but a decline in the enzyme activity was caused at concentrations above that (IC_{50}= 1910.96 μM). Isoquercetin started showing a limitation of the enzyme's activity already for concentrations above 50 μM, and had a much lower IC_{50} value, equal to 605.13 μM, confirming the allegation that the docking simulation also indicated, that glycosylation increases the binding affinity of a compound to M^{pro}. Rutin resulted in a smooth inhibition curve from which an IC_{50} of 286.93 μM was calculated, which is the lowest among all the compounds tested. Rutin is a rhamnoglucoside

derivative of quercetin, and achieves even better inhibition than isoquercetin, which is a single sugar derivative. The same pattern is observed for isorhamnetin (IC_{50} = 1435.99 µM), in which case the respective glucoside and rutinoside show progressively better inhibitory potential , with isorhamnetin 3-glucoside having an IC_{50} equal to 586.31 and isorhamnetin 3-rutinoside equal to 351.81 µM. The latter is also the only compound among the ones tested that achieved such a quick rate of reduction of the activity of M^{pro} as the concentration of the inhibitor increased, reaching 100% inhibition for 1000 µM of inhibitor.

Pelargonidin-3-rutinoside was also included in the screening, since it resulted in a very high binding energy in the molecular docking simulation (-9.171 kcal/mol) and was also commercially available. Although it was not possible to compare pelargonidin-3-rutinoside to its parent compound, pelargonidin, the molecule had a quite low IC_{50} , equal to 463.92 µM. It is worth pointing out that out of the top five hits of the assay, in terms of IC_{50}, three are flavonoid rutinosides.

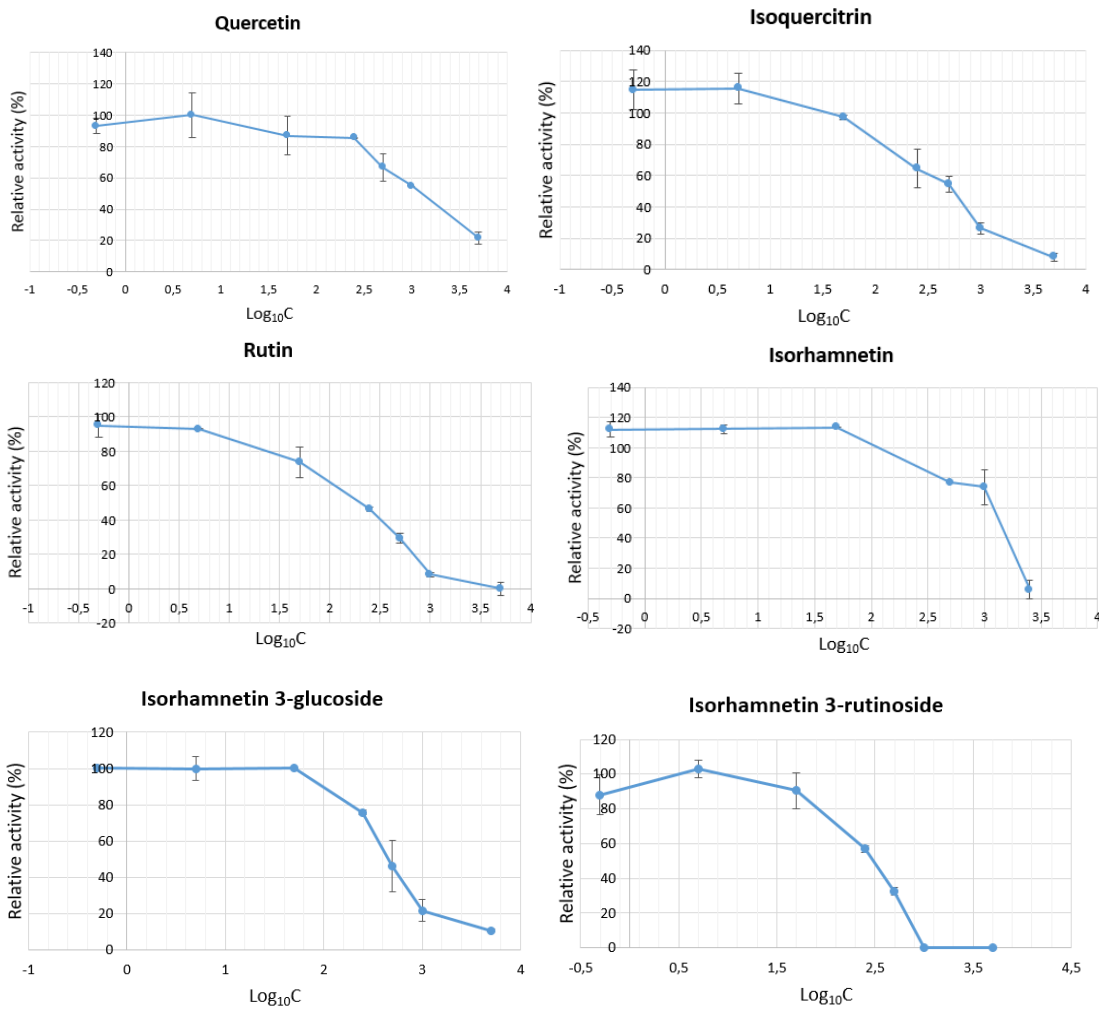


Figure 45: : Enzyme inhibition assay results for quercetin and derivatives: Relative activity of the enzyme as a function of the log_{10} of the different concentrations (C in µM)

Hesperetin and hesperidin were selected as another commercially available pair of parent compound and derivative whose activity could be correlated, while acacetin, galangin, apigenin, chrysin and catechin represent other flavonoid backbones to which the rest can be

compared. Hesperetin and hesperetin had a propensity towards inhibition, but did not manage to reduce the enzymatic activity more than 40% within the concentration range tested. Taking into consideration that the highest concentrations were not easily dissolved in DMSO for the assay to be conducted, it would perhaps not be possible to evaluate a potential better inhibitory effect of the compound at higher concentrations.

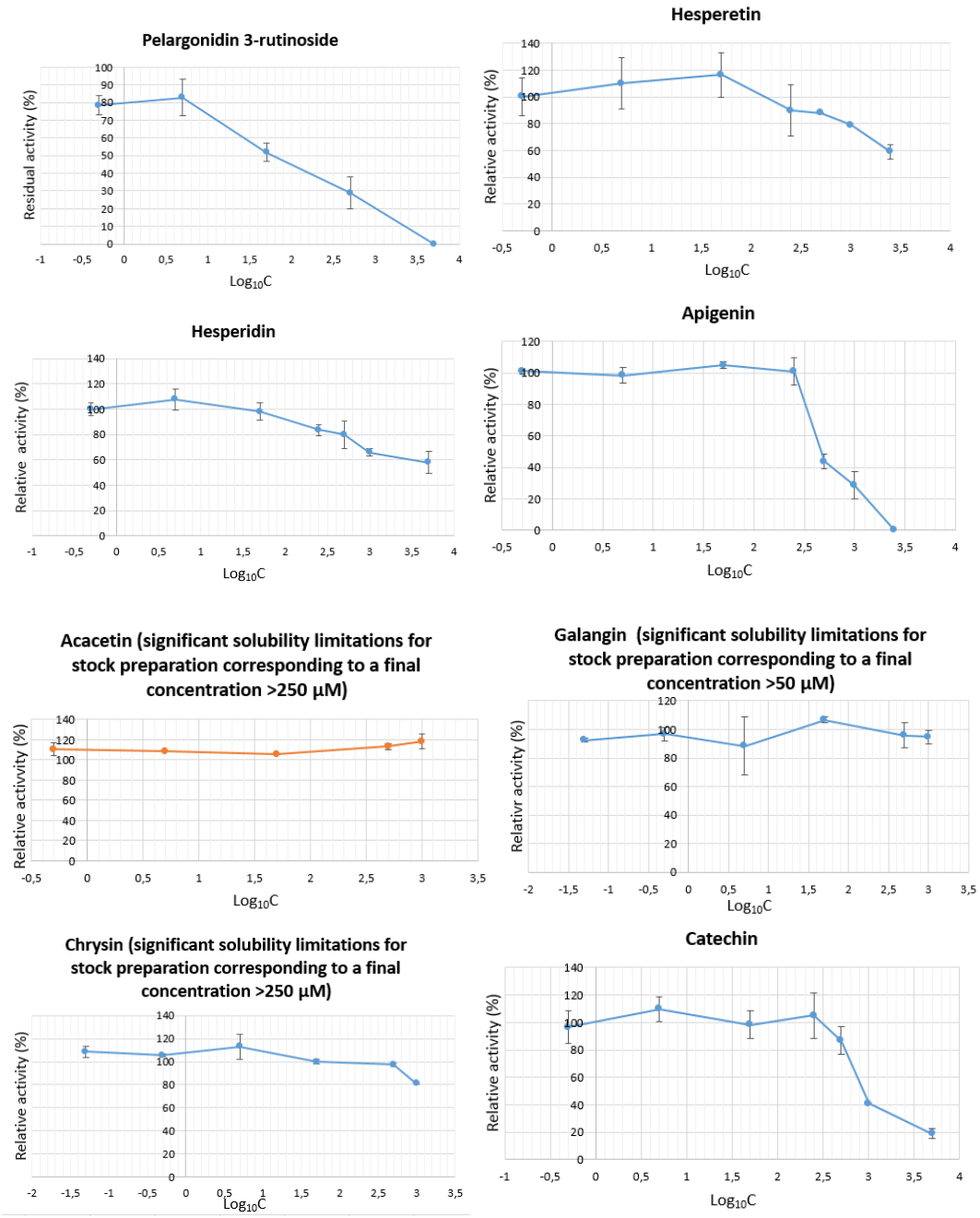


Figure 46: : Enzyme inhibition assay results for other flavonoids and flavanones: Relative activity of the enzyme as a function of the log10 of the different concentrations (C in µM)

Acacetin, galangin and chrysin are the three of the investigated flavonoids that showed no inhibitory activity. As seen in Figure 46, the activity of the enzyme remains the same regardless of the increase of the inhibitor concentration. Only in the case of chrysin is there a tendency to reduced enzyme activity, but that would require further experimentation with higher concentrations in order to be confirmed. An explanation for these results can be that the

compounds could not be well solubilized in order to fit the assay requirements. Therefore, it is highly probably that the actual concentration of the inhibitor was lower than the target one, and the inhibitory effect of the compound could not be effectively evaluated because a considerable amount of the inhibitor might not have been available to interact with the enzyme in the incubation volume.

Both apigenin and catechin show no inhibitory effect for concentrations up to 250 µM, above which the activity of the enzyme is subject to a steep decline. The IC_{50} values calculated are 604.07 µM for apigenin and 928.55 µM for catechin. Among the non-derivatized flavonoids, kaempferol shows the strongest M^{pro}-inhibitory effect, followed by myricetin and apigenin. Combining this pattern with the increase of the antiviral activity observed in glucosides and rutinosides of flavonoids, it would be interesting to investigate the potential of a glucoside or rutinoside of kaempferol, myricetin or apigenin.

4.4.3. Salicornia extract

The extract obtained from *Salicornia* plants, of unknown exact composition, also inhibits the SARS-CoV-2 protease. From the inhibition curve obtained from the assay results, the inhibition percentage seems to be reaching a plateau for concentrations higher than 500 µg/mL (Figure 45). In order for the behavior of the inhibitor to be verified, concentrations exceeding the range of this study (higher than 10 mg/mL) should be tested. In the case that the activity of the enzyme remains indeed stable above a certain concentration limit, this could be justified by the limited solubility of the constituents of the extract, that does not allow them to come in contact with the enzyme in the reaction volume. In any case, the *Salicornia* extract showed considerable inhibitory activity (IC_{50}= 400.66 µg/mL). This result is very encouraging, since the extract is the direct product of the utilization of the plant that can be made available as a potential immune-boosting nutraceutical, as opposed to the pure compounds, which would be more difficult to isolate.

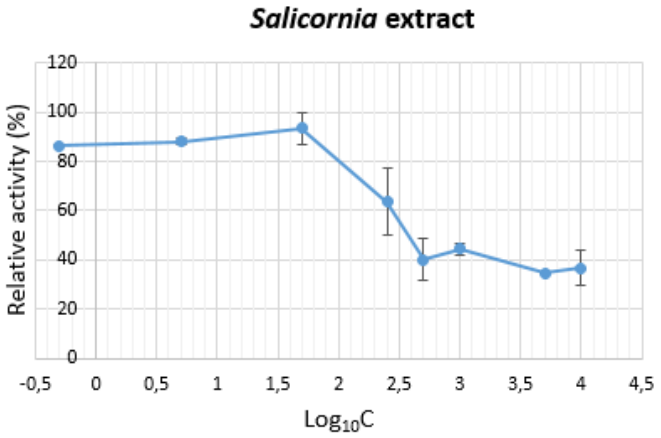


Figure 47: Enzyme inhibition assay results for the Salicornia extract: Relative activity of the enzyme as a function of the log_{10} of the different concentrations (C in µM)

4.4.4. Collective results

On the whole, most of the compounds tested showed an inhibitory effect to some extent, with flavonoids emerging as a very potent group of compounds against SARS-CoV-2. Caffeoylquinic acids also provided encouraging results on which a broader screen of structurally related compounds could be based on. A comparative overview of the in vitro assay results is presented in Figure 46 and Table 12.

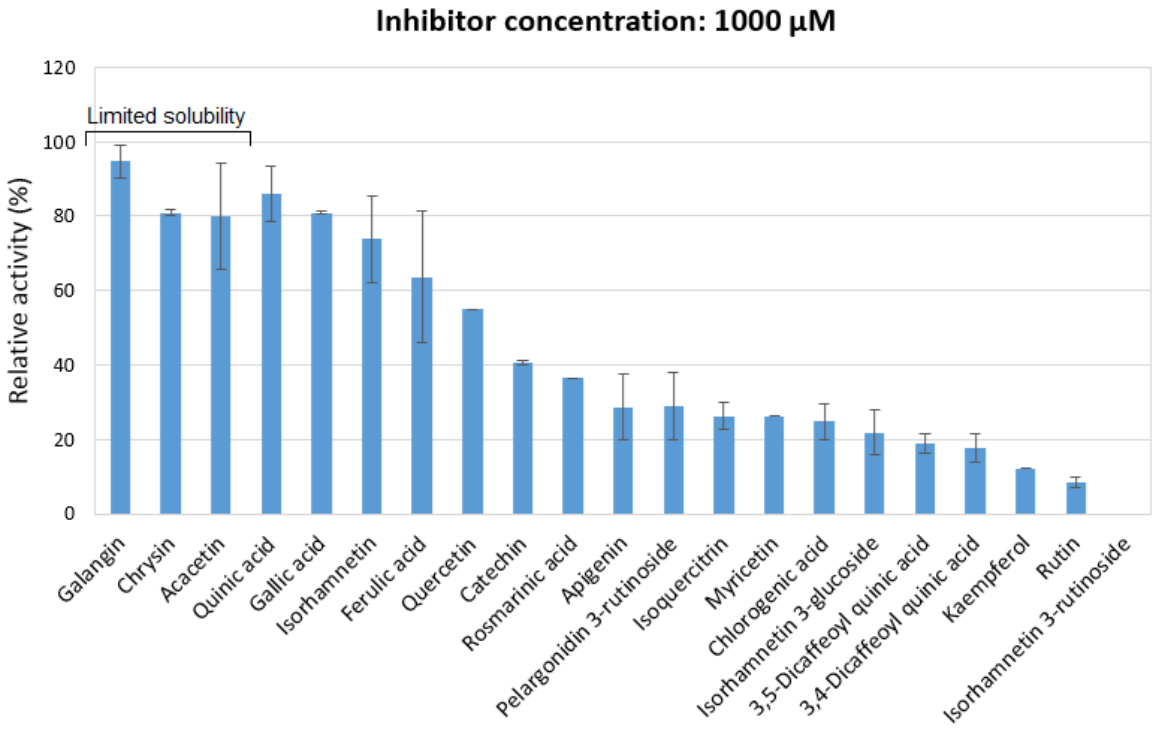


Figure 48: Comparison of the activity of M^{pro} when incubated with various inhibitors at a concentration of 1000 µM.

As seen in Figure 48, the majority of the compounds reduce the activity of M^{pro} by more than half at a concentration of 1000 µM. The order in which the compounds appear is different from Table 19, meaning that compounds that achieve a higher level of inhibition at 1000 µM are not necessarily the ones with the lowest IC_{50}. This is due to the fact that the inhibitory effect is affected differently by the concentration for each compound. For example, pelargonidin-3-rutinoside, which has an almost linear inhibition curve above 5 µM, has a lower inhibition percentage at 1000 µM compared to 3,4-dicaffeoyl quinic acid even though its IC_{50} is lower, since the latter causes a sharp decrease in the enzyme's activity right below 1000 µM.

As shown by the IC_{50} values, the most potent phytochemicals belong to the category of flavonoids and flavanones, followed by caffeoylquinic acids. The fact that the *Salicornia* extract has a lower IC_{50} than the majority of the compounds suggests a potential synergistic effect of the different compounds present.

Table 19: IC_{50} values calculated for the screened compounds, from lowest to highest (a,b are the parameters of equation (3), used to calculate IC_{50} and R^2 is the fitting of the linear curve from the equation of which a and b were derived).

Compound	a	b	R^2	IC_{50} (µM)
Rutin	-0.0648	68.593	0.9251	286.929
Kaempferol	-0.0626	71.4	0.9748	341.853
Isorhamnetin 3-rutinoside	-0.1465	101.54	0.9562	351.8089
***Salicornia* extract**	-0.1174	97.038	0.9829	400.6644
Pelargonidin 3-rutinoside	-0.0406	68.835	0.709	463.9163
3,4-Dicaffeoyl quinic acid	-0.0619	81.172	0.9599	503.5864
Myricetin	-0.0524	76.476	0.9358	505.2672
Chlorogenic acid	-0.0577	81.508	0.9949	546.0659

Isorhamnetin 3-glucoside	-0.081	97.491	0.9463	586.3086
3,5-Dicaffeoyl quinic acid	-0.079	97.227	0.9626	597.8101
Apigenin	-0.1105	116.75	0.8776	604.0724
Isoquercetin	-0.0686	91.512	0.9225	605.1312
Rosmarinic acid	-0.0761	110.99	0.9926	801.4455
Catechin	-0.0613	106.92	0.9104	928.5481
Isorhamnetin	-0.0414	109.45	0.965	1435.99
Quercetin	-0.0094	67.963	0.9752	1910.957
Ferulic acid	-0.0141	93.583	0.8605	3090.993
Gallic acid	-0.009	89.818	0.8635	4424.222

5. DISCUSSION

Molecular docking tools allowed for an initial screening of the contents of *Salicornia* extracts to determine their binding affinity and therefore inhibitory potential against SARS-CoV-2 M^{pro}. The results indicated binding affinity comparable to established inhibitors for the majority of the compounds tested, with some compounds showing a higher binding affinity even compared to native SARS-CoV-2 M^{pro} inhibitor N3. The most promising and representative compounds were further evaluated using an *in vitro* assay, the results of which supported the evidence for antiviral activity, with only few exceptions. The data (presented in table 20) demonstrates a considerable connection between the binding energy and the IC_{50} values calculated for each compound, as suggested by the correlation coefficient of 0.8189. The binding energy considered for each compound is the one of the first cluster, since there is no solid indication as to which cluster is the one that most accurately depicts the reality. It is observed that compounds with high (in absolute value) binding energies (close to 9 kcal/mol) generally have IC_{50} values below 500 μM, while for compounds with a binding energy between -7.0 and -7.5 kcal/mol there is an evident increase of the IC_{50} value, which is even greater when the binding energy drops below 6.0 kcal/mol (in terms of absolute value). This correlation supports the effectiveness of the molecular docking simulation as a useful tool able to provide an initial estimation of the inhibitory activity of a compound. which can be used to accelerate screening and save on resources when the number of potential active substances is large.

Table 20: Correlation between binding energy and IC_{50} values for the in vitro tested compounds

Compound	Binding energy (kcal/mol)	IC_{50} (μM)
Rutin	-9.166	286.929
Isorhamnetin 3-rutinoside	-9.384	351.809
Kaempferol	-7.752	341.853
Pelargonidin 3-rutinoside	-9.171	463.916
3,4 Di-caffeoyl quinic acid	-8.690	503.586
Myricetin	-7.429	505.267
Chlorogenic acid	-7.793	546.066
Isorhamnetin 3-glucoside	-8.613	586.309
3,5 Di-caffeoyl quinic acid	-8.935	597.810
Apigenin	-7.796	604.072
Isoquercetin	-8.952	605.131
Rosmarinic acid	-7.409	801.446
Catechin	-7.126	928.548
Isorhamnetin	-7.155	1435.99

Quercetin	-7.396	1910.96
Ferulic acid	-5.715	3090.99
Gallic acid	-5.601	4424.22

Moreover, the findings of the study confirm the initial hypothesis, which has also been supported in other recent works regarding antiviral strategies, that phytochemicals have remarkable properties that can be of great benefit to human health and, specifically in this case, emerge as valuable allies for fighting the ongoing pandemic. They certainly cannot replace antiviral drugs or vaccines, but the fact that they can be found in aliments and be easily consumed in the form of a supplement or extract, makes them appealing immune boosting agents that can support existing or future pharmaceutical products. In that context, the findings of this work indicate that the *Salicornia* extract is a very promising candidate, since it contains a wide variety of compounds with inhibitory properties, including compounds that have been detected exclusively in this plant species. The fact that it exhibited such properties also supports alternative routes of valorization of waste biomass, since the plant extract can be obtained from parts of the plant that are not otherwise commercially utilized.

The present thesis evaluates the anti-SARS-CoV-2 M^{pro} properties of pure compounds and the *Salicornia* extract when they are directly brought in contact with the enzyme by co-incubation. Whereas, in reality the route followed from ingestion to absorption of the substance and its transportation to infected cells is much more complex. A relevant parameter playing an important role in absorption of the active substances, that aroused as a challenge in this work as well, is the solubility of the extract and its constituents. For instance, phenolic acid, flavonoids and flavanones show limited solubility to both oil and aqueous media. At higher tested concentrations, certain substances could not be diluted and therefore could not react with the enzyme. Derivatization of the compounds could be a way to improve their solubility facilitating their access to the human body and cells. Glycosylation, in particular, could result in increased hydrophilicity and thus better solubility of the molecules, while also significantly increasing their inhibitory effect, as deduced from the present results. A selective and sustainable route to achieve this could be enzymatic modification. Various enzymes have been reported to achieve (trans)glycosylation of compounds. For example, rutinosides, which proved to have increased inhibitory properties compared to other glycosides, can be enzymatically synthesized utilizing the catalytic activity of rutinases (Katayama et al. 2013). Moreover, esterification with a wide variety of substitutions can be performed employing lipases or feruloyl esterases (Antonopoulou et al. 2019; Kumar et al. 2016; Schär and Nyström 2016). The possibilities are very broad and the indications regarding the inhibitory activity of *Salicornia* extracts could be a starting point for further investigation.

This work is an initial evaluation that provides some mechanistic insight into the inhibitory effect of *Salicornia* extract and its constituents against SARS-CoV-2 M^{pro} and is not adequate on its own to characterize the compounds as antiviral compounds. For the evaluation to be complete and conclusive, there would need to be an *in vitro* assay performed on infected cells, as well as tests conducted *in vivo*.

REFERENCES

Abian, Olga, David Ortega-Alarcon, Ana Jimenez-Alesanco, Laura Ceballos-Laita, Sonia Vega, Hugh T Reyburn, Bruno Rizzuti, and Adrian Velazquez-Campoy. 2020. "Structural Stability of SARS-CoV-2 3CLpro and Identification of Quercetin as an Inhibitor by Experimental Screening." *International Journal of Biological Macromolecules* 164 (December): 1693–1703. https://doi.org/10.1016/j.ijbiomac.2020.07.235.

Abid, Nabil, Zyed Rouis, Mohamed Lassoued, Souad Sfar, and Mahjoub Aouni. 2012. "Assessment of the Cytotoxic Effect and in Vitro Evaluation of the Anti-Enteroviral Activities of Plants Rich in Flavonoids." *Journal of Applied Pharmaceutical Science* 2 (May): 74–78. https://doi.org/10.7324/JAPS.2012.2532.

Adejoro, Isaiah A., Damilare D. Babatunde, and Gideon F. Tolufashe. 2020. " Molecular Docking and Dynamic Simulations of Some Medicinal Plants Compounds against SARS-CoV-2: An in Silico Study ." *Journal of Taibah University for Science* 14 (1): 1563–70. https://doi.org/10.1080/16583655.2020.1848049.

Adil, Md Tanveer, Rumana Rahman, Douglas Whitelaw, Vigyan Jain, Omer Al-Taan, Farhan Rashid, Aruna Munasinghe, and Periyathambi Jambulingam. 2021. "SARS-CoV-2 and the Pandemic of COVID-19." *Postgraduate Medical Journal* 97 (1144): 110–16. https://doi.org/10.1136/postgradmedj-2020-138386.

Ahmed, Sabri, Cherrak Id, Hafida Merzouk, and Nassima Mokhtari-soulimane. 2020. "Potential Bioactive Glycosylated Flavonoids as SARS-CoV-2 Main Protease Inhibitors : A Molecular Docking and Simulation Studies" 2: 1–14. https://doi.org/10.1371/journal.pone.0240653.

Aldred, Elaine M, Charles Buck, and Kenneth Vall. 2009. "Chapter 19 - Pharmacodynamics: How Drugs Elicit a Physiological Effect." In , edited by Elaine M Aldred, Charles Buck, and Kenneth B T - Pharmacology Vall, 137–43. Edinburgh: Churchill Livingstone. https://doi.org/https://doi.org/10.1016/B978-0-443-06898-0.00019-0.

Aljoundi, Aimen, Imane Bjij, Ahmed El Rashedy, and Mahmoud E.S. Soliman. 2020. "Covalent Versus Non-Covalent Enzyme Inhibition: Which Route Should We Take? A Justification of the Good and Bad from Molecular Modelling Perspective." *Protein Journal* 39 (2): 97–105. https://doi.org/10.1007/s10930-020-09884-2.

Antonopoulou, Io, Laura Iancu, Peter Jütten, Alexander Piechot, Ulrika Rova, and Paul Christakopoulos. 2019. "Screening of Novel Feruloyl Esterases from Talaromyces Wortmannii for the Development of Efficient and Sustainable Syntheses of Feruloyl Derivatives." *Enzyme and Microbial Technology* 120: 124–35. https://doi.org/https://doi.org/10.1016/j.enzmictec.2018.08.007.

Arakawa, Yoshihito, Yo-zo Asada, Hideaki Ishida, Hideyuki Chiji, and Masao Izawa. 1982. "Structures of New Two Isoflavones and One Flavanone from Glasswort." *Journal of the Faculty of Agriculture, Hokkaido University* 61 (1): 1–12.

Arakawa, Yoshihito, Hideyuki Chiji, and Masao Izawa. 1983. "Structural Elucidation of Two New Chromones Isolated from Glasswort (Salicornia Europaea L.)." *Agricultural and Biological Chemistry* 47 (9): 2029–33. https://doi.org/10.1271/bbb1961.47.2029.

Aref, Houda Lazreg, Badii Gaaliche, Abdelwaheb Fekih, Massoud Mars, Mahjoub Aouni, Jean Pierre Chaumon, and Khaled Said. 2011. "In Vitro Cytotoxic and Antiviral Activities of Ficus Carica Latex Extracts" 6419. https://doi.org/10.1080/14786419.2010.528758.

Arya, Rimanshee, Shweta Kumari, Bharati Pandey, Hiral Mistry, Subhash C Bihani, Amit Das, Vishal Prashar, Gagan D Gupta, Lata Panicker, and Mukesh Kumar. 2021. “Structural Insights into SARS-CoV-2 Proteins.” *Journal of Molecular Biology* 433 (2): 166725. https://doi.org/10.1016/j.jmb.2020.11.024.

Awoonor-Williams, Ernest, and Abd Al Aziz A. Abu-Saleh. 2021. “Covalent and Non-Covalent Binding Free Energy Calculations for Peptidomimetic Inhibitors of SARS-CoV-2 Main Protease.” *Physical Chemistry Chemical Physics* 23 (11): 6746–57. https://doi.org/10.1039/d1cp00266j.

Aykul, Senem, and Erik Martinez-Hackert. 2016. “Determination of Half-Maximal Inhibitory Concentration Using Biosensor-Based Protein Interaction Analysis.” *Analytical Biochemistry* 508: 97–103. https://doi.org/https://doi.org/10.1016/j.ab.2016.06.025.

Bai, Yu, Fei Ye, Yong Feng, Hanyi Liao, Hao Song, Jianxun Qi, George Fu Gao, Wenjie Tan, Lifeng Fu, and Yi Shi. 2021. “Structural Basis for the Inhibition of the SARS-CoV-2 Main Protease by the Anti-HCV Drug Narlaprevir.” *Signal Transduction and Targeted Therapy* 6 (1): 2020–22. https://doi.org/10.1038/s41392-021-00468-9.

Berg, JM, JL Tymoczko, and L. Stryer. 2002. “Enzymes Can Be Inhibited by Specific Molecules.” In *Biochemistry*, 5th editio. New York. https://www.ncbi.nlm.nih.gov/books/NBK22530/#A1068.

Bharadwaj, Shiv, Amit Dubey, Umesh Yadava, Sarad Kumar Mishra, Sang Gu Kang, and Vivek Dhar Dwivedi. 2021. “Exploration of Natural Compounds with Anti-SARS-CoV-2 Activity via Inhibition of SARS-CoV-2 Mpro.” *Briefings in Bioinformatics* 22 (2): 1361–77. https://doi.org/10.1093/bib/bbaa382.

Bi, Wentao, Minglei Tian, and Kyung Ho Row. 2012. “Separation of Phenolic Acids from Natural Plant Extracts Using Molecularly Imprinted Anion-Exchange Polymer Confined Ionic Liquids.” *Journal of Chromatography A* 1232: 37–42. https://doi.org/10.1016/j.chroma.2011.08.054.

Cer, R Z, U Mudunuri, R Stephens, and F J Lebeda. 2009. “IC50-to-Ki: A Web-Based Tool for Converting IC50 to Ki Values for Inhibitors of Enzyme Activity and Ligand Binding.” *Nucleic Acids Research* 37 (suppl_2): W441–45. https://doi.org/10.1093/nar/gkp253.

Chaturvedi, Tanmay, Joao M Uratani, Mette H Thomsen, and Jorge Rodríguez. 2012. “Evaluation of Pre-Treatment Conditions of Biomass Waste from the Halophyte Salicornia Bigelovii Cultivated in Sea Water.” *IWA Congress Anaerobic Digestion*, 10–13.

Chávez, Juliana H., Paulo C. Leal, Rosendo A. Yunes, Ricardo J. Nunes, Célia R.M. Barardi, Aguinaldo R. Pinto, Cláudia M.O. Simões, and Carlos R. Zanetti. 2006. “Evaluation of Antiviral Activity of Phenolic Compounds and Derivatives against Rabies Virus.” *Veterinary Microbiology* 116 (1–3): 53–59. https://doi.org/10.1016/j.vetmic.2006.03.019.

Chiang, L C, W Chiang, M C Liu, and C C Lin. 2003. “In Vitro Antiviral Activities of Caesalpinia Pulcherrima and Its Related Flavonoids.” *Journal of Antimicrobial Chemotherapy* 52: 194–98. https://doi.org/10.1093/jac/dkg291.

Cho, Jeong-Yong, Jin Y Kim, Yu G Lee, Hyoung J Lee, Hyun J Shim, Ji H Lee, Seon-Jae Kim, Kyung-Sik Ham, and Jae-Hak Moon. 2016. “Four New Dicaffeoylquinic Acid Derivatives from Glasswort (Salicornia Herbacea L.) and Their Antioxidative Activity.” *Molecules* . https://doi.org/10.3390/molecules21081097.

Chung, Chul Young, Hyo Kon Chun, Jae Young Yang, Ji Young Kim, Eun Hee Han, Yung Hee Kho, and Hye Gwang Jeong. 2005. “Tungtungmadic Acid, a Novel Antioxidant, from Salicornia Herbacea.” *Archives of Pharmacal Research* 28 (10): 1122–26. https://doi.org/10.1007/bf02972972.

Creech, C Buddy, Shannon C Walker, and Robert J Samuels. 2021. “SARS-CoV-2 Vaccines.” *JAMA* 325 (13): 1318–20. https://doi.org/10.1001/jama.2021.3199.

Dai, Wenhao, Bing Zhang, Xia Ming Jiang, Haixia Su, Jian Li, Yao Zhao, Xiong Xie, et al. 2020. “Structure-Based Design of Antiviral Drug Candidates Targeting the SARS-CoV-2 Main Protease.” *Science* 368 (6497): 1331–35. https://doi.org/10.1126/science.abb4489.

Das, Pratik, Ranabir Majumder, Mahitosh Mandal, and Piyali Basak. 2020. “In-Silico Approach for Identification of Effective and Stable Inhibitors for COVID-19 Main Protease (Mpro) from Flavonoid Based Phytochemical Constituents of Calendula Officinalis.” *Journal of Biomolecular Structure and Dynamics* 0 (0): 1–16. https://doi.org/10.1080/07391102.2020.1796799.

Delaune, Keith P., and Khalid Alsayouri. 2020. “Physiology, Noncompetitive Inhibitor.” In *StatPearls*. https://www.ncbi.nlm.nih.gov/books/NBK545242/#article-25953.s7.

Douangamath, Alice, Daren Fearon, Paul Gehrtz, Tobias Krojer, Petra Lukacik, C. David Owen, Efrat Resnick, et al. 2020. “Crystallographic and Electrophilic Fragment Screening of the SARS-CoV-2 Main Protease.” *Nature Communications* 11 (1): 1–11. https://doi.org/10.1038/s41467-020-18709-w.

Dougall, Iain G, and John Unitt. 2015. “Chapter 2 - Evaluation of the Biological Activity of Compounds: Techniques and Mechanism of Action Studies.” In , edited by Camille Georges Wermuth, David Aldous, Pierre Raboisson, and Didier B T - The Practice of Medicinal Chemistry (Fourth Edition) Rognan, 15–43. San Diego: Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-417205-0.00002-X.

Elsebaie, EM, SYA Elsanat, MS Gouda, and KM Elnemr. 2014. “Utilization of Salicornia Fruticosa Herb for Producing Antioxidants.” *Bangladesh Journal of Scientific and Industrial Research* 49 (1): 53–58. https://doi.org/10.3329/bjsir.v49i1.18856.

Engelking, Larry R. 2015. “Chapter 6 - Enzyme Kinetics.” In , edited by Larry R B T - Textbook of Veterinary Physiological Chemistry (Third Edition) Engelking, 32–38. Boston: Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-391909-0.50006-2.

Essaidi, Ismahen, Zeineb Brahmi, Ahmed Snoussi, Hayat Ben Haj Koubaier, Hervé Casabianca, Naoki Abe, Abdelfatteh El Omri, Mohamed Moncef Chaabouni, and Nabiha Bouzouita. 2013. “Phytochemical Investigation of Tunisian Salicornia Herbacea L., Antioxidant, Antimicrobial and Cytochrome P450 (CYPs) Inhibitory Activities of Its Methanol Extract.” *Food Control* 32 (1): 125–33. https://doi.org/10.1016/j.foodcont.2012.11.006.

Farady, Christopher J, and Charles S Craik. 2010. “Mechanisms of Macromolecular Protease Inhibitors.” *Chembiochem : A European Journal of Chemical Biology* 11 (17): 2341–46. https://doi.org/10.1002/cbic.201000442.

Formica, J. V., and W. Regelson. 1995. “Review of the Biology of Quercetin and Related Bioflavonoids.” *Food and Chemical Toxicology*. Pergamon. https://doi.org/10.1016/0278-6915(95)00077-1.

Friend, Thomas, and Justin Stebbing. 2021. “What Is the Intermediate Host Species of SARS-CoV-2?” *Future Virology* 16 (3): 153–56. https://doi.org/10.2217/fvl-2020-0390.

Fu, Lifeng, Fei Ye, Yong Feng, Feng Yu, Qisheng Wang, Yan Wu, Cheng Zhao, et al. 2020. "Both Boceprevir and GC376 Efficaciously Inhibit SARS-CoV-2 by Targeting Its Main Protease." *Nature Communications* 11 (1): 1–8. https://doi.org/10.1038/s41467-020-18233-x.

Gajjar, Normi D, Tejas M Dhameliya, and Gaurang B Shah. 2021. "In Search of RdRp and Mpro Inhibitors against SARS CoV-2: Molecular Docking, Molecular Dynamic Simulations and ADMET Analysis." *Journal of Molecular Structure* 1239: 130488. https://doi.org/https://doi.org/10.1016/j.molstruc.2021.130488.

Geslin, Michel, and Jean-François Verbist. 1985. "Flavonoides de Salicornia Europaea." *Journal of Natural Products* 48 (1): 111–13. https://doi.org/10.1021/np50037a020.

Ghosh, Rajesh, Ayon Chakraborty, Ashis Biswas, and Snehasis Chowdhuri. 2020. "Evaluation of Green Tea Polyphenols as Novel Corona Virus (SARS CoV-2) Main Protease (Mpro) Inhibitors–an in Silico Docking and Molecular Dynamics Simulation Study." *Journal of Biomolecular Structure and Dynamics* 0 (0): 1–13. https://doi.org/10.1080/07391102.2020.1779818.

Gil, Carmen, Tiziana Ginex, Inés Maestro, Vanesa Nozal, Lucía Barrado-Gil, Miguel Ángel Cuesta-Geijo, Jesús Urquiza, et al. 2020. "COVID-19: Drug Targets and Potential Treatments." *Journal of Medicinal Chemistry* 63 (21): 12359–86. https://doi.org/10.1021/acs.jmedchem.0c00606.

Gravina, H D, N F Tafuri, A Silva Júnior, J L R Fietto, T T Oliveira, M A N Diaz, and M R Almeida. 2011. "In Vitro Assessment of the Antiviral Potential of Trans-Cinnamic Acid, Quercetin and Morin against Equid Herpesvirus 1." *Research in Veterinary Science* 91 (3): e158–62. https://doi.org/https://doi.org/10.1016/j.rvsc.2010.11.010.

Griffin, Jeddidiah W D. 2020. "Since January 2020 Elsevier Has Created a COVID-19 Resource Centre with Free Information in English and Mandarin on the Novel Coronavirus COVID- 19 . The COVID-19 Resource Centre Is Hosted on Elsevier Connect , the Company ' s Public News and Information ," no. January.

Günther, Sebastian, Patrick Y A Reinke, Yaiza Fernández-García, Julia Lieske, Thomas J Lane, Helen M Ginn, Faisal H M Koua, et al. 2021. "X-Ray Screening Identifies Active Site and Allosteric Inhibitors of SARS-CoV-2 Main Protease." *Science* 372 (6542): 642 LP – 646. https://doi.org/10.1126/science.abf7945.

Habtemariam, Solomon, Seyed Fazel Nabavi, Maciej Banach, Ioana Berindan-Neagoe, Kasturi Sarkar, Parames C. Sil, and Seyed Mohammad Nabavi. 2020. "Should We Try SARS-CoV-2 Helicase Inhibitors for COVID-19 Therapy?" *Archives of Medical Research* 51 (7): 733–35. https://doi.org/10.1016/j.arcmed.2020.05.024.

Hattori, Shin ichiro, Nobuyo Higashi-Kuwata, Hironori Hayashi, Srinivasa Rao Allu, Jakka Raghavaiah, Haydar Bulut, Debananda Das, et al. 2021. "A Small Molecule Compound with an Indole Moiety Inhibits the Main Protease of SARS-CoV-2 and Blocks Virus Replication." *Nature Communications* 12 (1): 1–12. https://doi.org/10.1038/s41467-021-20900-6.

Hegyi, Annette, and John Ziebuhr. 2002. "Conservation of Substrate Specificities among Coronavirus Main Proteases." *Journal of General Virology* 83 (3): 595–99. https://doi.org/10.1099/0022-1317-83-3-595.

Hillen, Hauke S., Goran Kokic, Lucas Farnung, Christian Dienemann, Dimitry Tegunov, and Patrick Cramer. 2020. "Structure of Replicating SARS-CoV-2 Polymerase." *Nature* 584

(7819): 154–56. https://doi.org/10.1038/s41586-020-2368-8.

Hoffman, Robert L, Robert S Kania, Mary A Brothers, Jay F Davies, Rose A Ferre, Ketan S Gajiwala, Mingying He, et al. 2020. “Discovery of Ketone-Based Covalent Inhibitors of Coronavirus 3CL Proteases for the Potential Therapeutic Treatment of COVID-19.” *Cite This: J. Med. Chem* 63: 12725–47. https://doi.org/10.1021/acs.jmedchem.0c01063.

Hoffmann, Markus, Hannah Kleine-Weber, Simon Schroeder, Nadine Krüger, Tanja Herrler, Sandra Erichsen, Tobias S Schiergens, et al. 2020. “SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor.” *Cell* 181 (2): 271-280.e8. https://doi.org/https://doi.org/10.1016/j.cell.2020.02.052.

Hu, Ben, Hua Guo, Peng Zhou, and Zheng Li Shi. 2021. “Characteristics of SARS-CoV-2 and COVID-19.” *Nature Reviews Microbiology* 19 (3): 141–54. https://doi.org/10.1038/s41579-020-00459-7.

Hu, Yanmei, Chunlong Ma, Tommy Szeto, Brett Hurst, Bart Tarbet, and Jun Wang. 2021. “Boceprevir, Calpain Inhibitors II and XII, and GC-376 Have Broad-Spectrum Antiviral Activity against Coronaviruses.” *ACS Infectious Diseases* 7 (3): 586–97. https://doi.org/10.1021/acsinfecdis.0c00761.

Huang, Sheng-Teng, Yeh Chen, Wei-Chao Chang, Hsiao-Fan Chen, Hsiang-Chun Lai, Yu-Chun Lin, Wei-Jan Wang, et al. 2021. “Scutellaria Barbata D. Don Inhibits the Main Proteases (Mpro and TMPRSS2) of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Infection.” *Viruses* . https://doi.org/10.3390/v13050826.

Jin, Zhenming, Xiaoyu Du, Yechun Xu, Yongqiang Deng, Meiqin Liu, Yao Zhao, Bing Zhang, et al. 2020. “Structure of Mpro from SARS-CoV-2 and Discovery of Its Inhibitors.” *Nature* 582 (7811): 289–93. https://doi.org/10.1038/s41586-020-2223-y.

Jin, Zhenming, Yao Zhao, Yuan Sun, Bing Zhang, Haofeng Wang, Yan Wu, Yan Zhu, et al. 2020. “Structural Basis for the Inhibition of SARS-CoV-2 Main Protease by Antineoplastic Drug Carmofur.” *Nature Structural and Molecular Biology* 27 (6): 529–32. https://doi.org/10.1038/s41594-020-0440-6.

Kadereit, Gudrun, Mikko Piirainen, Jacques Lambinon, and Alain Vanderpoorten. 2012. “Cryptic Taxa Should Have Names: Reflections in the Glasswort Genus Salicornia (Amaranthaceae).” *Taxon* 61 (6): 1227–39. https://doi.org/10.1002/tax.616005.

Kang, Smee, Mi Ri Kim, Maehee Chiang, and Jungil Hong. 2015. “Evaluation and Comparison of Functional Properties of Freshwater-Cultivated Glasswort (Salicornia Herbacea L.) with Naturally-Grown Glasswort.” *Food Science and Biotechnology* 24 (6): 2245–50. https://doi.org/10.1007/s10068-015-0299-1.

Kanjanasirirat, Phongthon, Ampa Suksatu, Suwimon Manopwisedjaroen, Bamroong Munyoo, Patoomratana Tuchinda, Kedchin Jearawuttanakul, Sawinee Seemakhan, et al. 2020. “High-Content Screening of Thai Medicinal Plants Reveals Boesenbergia Rotunda Extract and Its Component Panduratin A as Anti-SARS-CoV-2 Agents.” *Scientific Reports* 10 (1): 19963. https://doi.org/10.1038/s41598-020-77003-3.

Karadeniz, Fatih, Jung-Ae Kim, Byul-Nim Ahn, Myeong S Kwon, and Chang-Suk Kong. 2014. “Effect of Salicornia Herbacea on Osteoblastogenesis and Adipogenesis in Vitro.” *Marine Drugs* . https://doi.org/10.3390/md12105132.

Katayama, Shigeru, Fumiaki Ohno, Yuki Yamauchi, Miyuki Kato, Hidefumi Makabe, and Soichiro Nakamura. 2013. “Enzymatic Synthesis of Novel Phenol Acid Rutinosides Using

Rutinase and Their Antiviral Activity in Vitro." *Journal of Agricultural and Food Chemistry* 61 (40): 9617–22. https://doi.org/10.1021/jf4021703.

Khan, Abbas, Wang Heng, Yanjing Wang, Jingfei Qiu, Xiaoyong Wei, Shaoliang Peng, Shoaib Saleem, Mazhar Khan, Syed Shujait Ali, and Dong Qing Wei. 2021. "In Silico and in Vitro Evaluation of Kaempferol as a Potential Inhibitor of the SARS-CoV-2 Main Protease (3CLpro)." *Phytotherapy Research* 35 (6): 2841–45. https://doi.org/10.1002/ptr.6998.

Kiani, Aysha Karim, Kristjana Dhuli, Kyrylo Anpilogov, Simone Bressan, Astrit Dautaj, Munis Dundar, Tommaso Beccari, Mahmut C Ergoren, and Matteo Bertelli. 2020. "Natural Compounds as Inhibitors of SARS-CoV-2 Endocytosis: A Promising Approach against COVID-19." *Acta Bio-Medica : Atenei Parmensis* 91 (13-S): e2020008–e2020008. https://doi.org/10.23750/abm.v91i13-S.10520.

Kim, Dongwan, Lee, Joo-Yeon, Yang, Jeong-Sun, Kim, Jun Won, Kim, V. Narry, Chang, Hyeshik. 2020. "The Architecture of SARS-CoV-2 Transcriptome". *Cell* 181 (4): 914-921.e10. https://doi.org/10.1016/j.cell.2020.04.011.

Kim, Ha Song, Young Seung Yoon, and Jai Woo Cho. 2008. "Quantitative Analysis of Flavonoids from Salicomia Herbacea L. Extracst by LC-MS." *Korean Journal of Medicinal Crop Science* 16 (4): 231–37.

Kim, Jin Young, Jeong Yong Cho, Young Kyu Ma, Keun Young Park, Sang Hyun Lee, Kyung Sik Ham, Hyoung Jae Lee, Keun Hyung Park, and Jae Hak Moon. 2011. "Dicaffeoylquinic Acid Derivatives and Flavonoid Glucosides from Glasswort (Salicornia Herbacea L.) and Their Antioxidative Activity." *Food Chemistry* 125 (1): 55–62. https://doi.org/10.1016/j.foodchem.2010.08.035.

Kim, Joonsoo, Govindarajan Karthivashan, Mee-Hyang Kweon, Deuk-Hoi Kim, and Dong-Kug Choi. 2019. "The Ameliorative Effects of the Ethyl Acetate Extract of *Salicornia Europaea* L. and Its Bioactive Candidate, Irilin B, on LPS-Induced Microglial Inflammation and MPTP-Intoxicated PD-Like Mouse Model." *Oxidative Medicine and Cellular Longevity* 2019: 6764756. https://doi.org/10.1155/2019/6764756.

Kim, Kwan-Su, and Si-Hyung Park. 2004. "Isolation and Identification of Antioxidant Flavonoids from Salicornia Herbacea L." *Applied Biological Chemistry*.

Kim, You Ah, Chang-Suk Kong, Jung Im Lee, Hojun Kim, Hee Yeon Park, Hyi-Seung Lee, Chulhyun Lee, and Youngwan Seo. 2012. "Evaluation of Novel Antioxidant Triterpenoid Saponins from the Halophyte Salicornia Herbacea." *Bioorganic & Medicinal Chemistry Letters* 22 (13): 4318–22. https://doi.org/10.1016/j.bmcl.2012.05.017.

Kneller, Daniel W., Gwyndalyn Phillips, Hugh M. O'Neill, Robert Jedrzejczak, Lucy Stols, Paul Langan, Andrzej Joachimiak, Leighton Coates, and Andrey Kovalevsky. 2020a. "Structural Plasticity of SARS-CoV-2 3CL Mpro Active Site Cavity Revealed by Room Temperature X-Ray Crystallography." *Nature Communications* 11 (1): 7–12. https://doi.org/10.1038/s41467-020-16954-7.

———. 2020b. "Structural Plasticity of SARS-CoV-2 3CL Mpro Active Site Cavity Revealed by Room Temperature X-Ray Crystallography." *Nature Communications* 11 (1): 7–12. https://doi.org/10.1038/s41467-020-16954-7.

Kneller, Daniel W, Stephanie Galanie, Gwyndalyn Phillips, Hugh M O Neill, Leighton Coates, Andrey Kovalevsky, Daniel W Kneller, et al. 2020. "Malleability of the SARS-CoV-2 3CL M pro Active-Site Cavity Facilitates Binding of Clinical Antivirals." *Structure/Folding and Design* 28 (12): 1313-1320.e3. https://doi.org/10.1016/j.str.2020.10.007.

Koudelka, Tomas, Juliane Boger, Alessandra Henkel, Robert Schönherr, Stefanie Krantz, Sabine Fuchs, Estefanía Rodríguez, Lars Redecke, and Andreas Tholey. 2021. "N-Terminomics for the Identification of In Vitro Substrates and Cleavage Site Specificity of the SARS-CoV-2 Main Protease." *Proteomics* 21. https://doi.org/10.1002/pmic.202000246.

Kumar, Vinod, Firdaus Jahan, Richi V Mahajan, and Rajendra Kumar Saxena. 2016. "Efficient Regioselective Acylation of Quercetin Using Rhizopus Oryzae Lipase and Its Potential as Antioxidant." *Bioresource Technology* 218: 1246–48. https://doi.org/https://doi.org/10.1016/j.biortech.2016.06.057.

Kundu Debanjan, Umesh, Chandrabose Selvaraj, Sanjeev Kumar Singh, and Vikash Kumar Dubey. 2021. "Identification of New Anti-NCoV Drug Chemical Compounds from Indian Spices Exploiting SARS-CoV-2 Main Protease as Target." *Journal of Biomolecular Structure and Dynamics* 39 (9): 3428–34. https://doi.org/10.1080/07391102.2020.1763202.

Kushwaha, Prem Prakash, Atul Kumar Singh, Kumari Sunita Prajapati, Mohd Shuaib, Sanjay Gupta, and Shashank Kumar. 2021. "Phytochemicals Present in Indian Ginseng Possess Potential to Inhibit SARS-CoV-2 Virulence: A Molecular Docking and MD Simulation Study." *Microbial Pathogenesis* 157: 104954. https://doi.org/https://doi.org/10.1016/j.micpath.2021.104954.

Kuzikov, Maria, Elisa Costanzi, Jeanette Reinshagen, Francesca Esposito, Laura Vangeel, Markus Wolf, Bernhard Ellinger, et al. 2021. "Identification of Inhibitors of SARS-CoV-2 3CL-Pro Enzymatic Activity Using a Small Molecule in Vitro Repurposing Screen." *ACS Pharmacology and Translational Science*. https://doi.org/10.1021/acsptsci.0c00216.

Lockbaum, Gordon J., Archie C. Reyes, Jeong Min Lee, Ronak Tilvawala, Ellen A. Nalivaika, Akbar Ali, Nese Kurt Yilmaz, Paul R. Thompson, and Celia A. Schiffer. 2021. "Crystal Structure of SARS-CoV-2 Main Protease in Complex with the Non-Covalent Inhibitor ML188." *Viruses* 13 (2). https://doi.org/10.3390/v13020174.

Loconsole, Danilo, Giuseppe Cristiano, and Barbara De Lucia. 2019. "Glassworts: From Wild Salt Marsh Species to Sustainable Edible Crops." *Agriculture (Switzerland)* 9 (1). https://doi.org/10.3390/agriculture9010014.

Lyu, Hui, Xin Ma, Fuqin Guan, Yu Chen, Qizhi Wang, and Xu Feng. 2018. "30-Noroleanane Triterpenoid Saponins from Salicornia Europaea Linn. and Their Chemotaxonomic Significance." *Biochemical Systematics and Ecology* 78: 106–9. https://doi.org/https://doi.org/10.1016/j.bse.2018.04.007.

Matheson, By Nicholas J, and Paul J Lehner. 2020. "How Does SARS-CoV-2 Cause COVID-19?" 369 (6503): 510–12.

Mathew, Sindhu, and Tanya Abraham. 2004. "Ferulic Acid: An Antioxidant Found Naturally in Plant Cell Walls and Feruloyl Esterases Involved in Its Release and Their Applications." *Critical Reviews in Biotechnology* 24 (February): 59–83. https://doi.org/10.1080/07388550490491467.

Mengist, Hylemariam Mihiretie, Xiaojiao Fan, and Tengchuan Jin. 2020. "Designing of Improved Drugs for COVID-19: Crystal Structure of SARS-CoV-2 Main Protease Mpro." *Signal Transduction and Targeted Therapy* 5 (1): 2–3. https://doi.org/10.1038/s41392-020-0178-y.

Mrid, Reda Ben, Najat Bouchmaa, Imad Kabach, Mansour Sobeh, Abdelmajid Zyad,

Mohamed Nhiri, and Abdelaziz Yasri. 2021. “In Silico Screening of Moroccan Medicinal Plants with the Ability to Directly Inhibit the Novel Coronavirus, SARS-CoV-2.” *Research Square*. https://doi.org/10.21203/rs.3.rs-38104/v1.

Murugesan, Selvakumar, Sanjay Kottekad, Inchara Crasta, Sivakumar Sreevathsan, Dandamudi Usharani, Madan Kumar Perumal, and Sandeep Narayan Mudliar. 2021. “Targeting COVID-19 (SARS-CoV-2) Main Protease through Active Phytocompounds of Ayurvedic Medicinal Plants – Emblica Officinalis (Amla), Phyllanthus Niruri Linn. (Bhumi Amla) and Tinospora Cordifolia (Giloy) – A Molecular Docking and Simulation Study.” *Computers in Biology and Medicine* 136 (Murugesan, S., Kottekad, S., Crasta, I., Sreevathsan, S., Usharani, D., Perumal, M. K., Mudliar, S. N. (2021). Targeting COVID-19 (SARS-CoV-2) main protease through active phytocompounds of ayurvedic medicinal plants – Emblica officinalis (Amla), Phylla): 104683. https://doi.org/https://doi.org/10.1016/j.compbiomed.2021.104683.

Neubig, Richard R, Michael Spedding, Terry Kenakin, and Arthur Christopoulos. 2003. “International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification. XXXVIII. Update on Terms and Symbols in Quantitative Pharmacology.” *Pharmacological Reviews* 55 (4): 597–606. https://doi.org/10.1124/pr.55.4.4.

Ochs, Raymond S. 2000. “Understanding Enzyme Inhibition.” *Journal of Chemical Education* 77 (11): 1453. https://doi.org/10.1021/ed077p1453.

Oh, Ji Hae, Eun Ok Kim, Sung Kwon Lee, Mee Hee Woo, and Sang Won Choi. 2007. “Antioxidant Activities of the Ethanol Extract of Hamcho (Salicornia Herbacea L.) Cake Prepared by Enzymatic Treatment.” *Food Science and Biotechnology*.

Özçelik, Berrin, Murat Kartal, and Ilkay Orhan. 2011. “Cytotoxicity, Antiviral and Antimicrobial Activities of Alkaloids, Flavonoids, and Phenolic Acids.” *Pharmaceutical Biology* 49 (4): 396–402. https://doi.org/10.3109/13880209.2010.519390.

Palmer, Trevor, and Philip L Bonner. 2011. “8 - Enzyme Inhibition.” In , edited by Trevor Palmer and Philip L B T - Enzymes (Second Edition) Bonner, 126–52. Woodhead Publishing. https://doi.org/https://doi.org/10.1533/9780857099921.2.126.

Park, Sang Hyun, Sung Kwon Ko, Jin Gyu Choi, and Sung Hyun Chung. 2006. “Salicornia Herbacea Prevents High Fat Diet-Induced Hyperglycemia and Hyperlipidemia in ICR Mice.” *Archives of Pharmacal Research* 29 (3): 256–64. https://doi.org/10.1007/BF02969402.

Patel, Seema. 2016. “Salicornia: Evaluating the Halophytic Extremophile as a Food and a Pharmaceutical Candidate.” *3 Biotech* 6 (1): 1–10. https://doi.org/10.1007/s13205-016-0418-6.

Patil, Rajkumar Sanjay, Nayeem A. Khatib, Vishal Shivalingappa Patil, and Shailendra Sanjay Suryawanshi. 2020. “Chlorogenic Acid May Be a Potent Inhibitor of Dimeric SARS-CoV-2 Main Protease 3CLpro: An in Silico Study.” *Traditional Medicine Research*, 1–12. https://doi.org/10.12032/TMR20201208211.

Pichiah, P B Tirupathi, and Youn-Soo Cha. 2015. “Salicornia Herbacea Prevents Weight Gain and Hepatic Lipid Accumulation in Obese ICR Mice Fed a High-Fat Diet.” *Journal of the Science of Food and Agriculture* 95 (15): 3150–59. https://doi.org/https://doi.org/10.1002/jsfa.7054.

Premnathan, M., K. Kathiresan, S. K. Bajpai, and K. Chandra. 1992. “A Survey of Some Indian Marine Plants for Antiviral Activity.” *Botanica Marina* 35 (4): 321–24.

https://doi.org/10.1515/botm.1992.35.4.321.

Qiao, Jingxin, Yue-shan Li, Rui Zeng, Feng-liang Liu, Rong-hua Luo, Chong Huang, Yi-fei Wang, et al. 2021. “SARS-CoV-2 Mpro Inhibitors with Antiviral Activity in a Transgenic Mouse Model” 1378 (March): 1374–78.

Radwan, HM, and NM Nazif. 2007. “Phytochemical Investigation of Salicornia Fruticosa (L.) And Their Biological Activity.” *Research Journal Of* 2 (2): 72–78. http://www.aensionline.com/rjmms/rjmms/2007/72-78.pdf.

Rahman, Md Mahbubur, Myung Jin Kim, Jin Hyoung Kim, Sok Ho Kim, Hyeon Kyu Go, Mee Hyang Kweon, and Do Hyung Kim. 2018. “Desalted Salicornia Europaea Powder and Its Active Constituent, Trans-Ferulic Acid, Exert Anti-Obesity Effects by Suppressing Adipogenic-Related Factors.” *Pharmaceutical Biology* 56 (1): 183–91. https://doi.org/10.1080/13880209.2018.1436073.

Rakib, Ahmed, Zulkar Nain, Saad Ahmed Sami, Shafi Mahmud, Ashiqul Islam, Shahriar Ahmed, Adnan Bin Faisul Siddiqui, et al. 2021. “A Molecular Modelling Approach for Identifying Antiviral Selenium-Containing Heterocyclic Compounds That Inhibit the Main Protease of SARS-CoV-2: An in Silico Investigation.” *Briefings in Bioinformatics* 22 (2): 1476–98. https://doi.org/10.1093/bib/bbab045.

Rhee, Man Hee, Hwa Jin Park, and Jae Youl Cho. 2009. “Salicornia Herbacea: Botanical, Chemical and Pharmacological Review of Halophyte Marsh Plant.” *Journal of Medicinal Plants Research* 3 (8): 548–55.

Saboury, A. A. 2009. “Enzyme Inhibition and Activation: A General Theory.” *Journal of the Iranian Chemical Society* 6 (2): 219–29. https://doi.org/10.1007/BF03245829.

Sacco, Michael Dominic, Chunlong Ma, Panagiotis Lagarias, Ang Gao, Julia Alma Townsend, Xiangzhi Meng, Peter Dube, et al. 2020. “Structure and Inhibition of the SARS-CoV-2 Main Protease Reveals Strategy for Developing Dual Inhibitors against Mpro and Cathepsin L.” *BioRxiv*, no. December. https://doi.org/10.1101/2020.07.27.223727.

Salman, Saad, Fahad H. Shah, Jawaria Idrees, Fariha Idrees, Shreya Velagala, Johar Ali, and Abid A. Khan. 2020a. “Virtual Screening of Immunomodulatory Medicinal Compounds as Promising Anti-SARS-CoV-2 Inhibitors.” *Future Virology* 15 (5): 267–75. https://doi.org/10.2217/fvl-2020-0079.

Salman, Saad, Fahad H Shah, Jawaria Idrees, Fariha Idrees, Shreya Velagala, Johar Ali, and Abid A Khan. 2020b. “Virtual Screening of Immunomodulatory Medicinal Compounds as Promising Anti-SARS-COV-2 Inhibitors.” *Future Virology*, May, 10.2217/fvl-2020–0079. https://doi.org/10.2217/fvl-2020-0079.

Sánchez-Gavilán, Irene, Esteban Ramírez, and Vicenta de la Fuente. 2021. “Bioactive Compounds in Salicornia Patula Duval-Jouve: A Mediterranean Edible Euhalophyte.” *Foods* 10 (2): 1–12. https://doi.org/10.3390/foods10020410.

Schär, Aline, and Laura Nyström. 2016. “Enzymatic Synthesis of Steryl Ferulates.” *European Journal of Lipid Science and Technology* 118 (10): 1557–65. https://doi.org/10.1002/ejlt.201500586.

Shamsi, Anas, Taj Mohammad, Saleha Anwar, Samreen Amani, Mohd Shahnawaz Khan, Fohad Mabood Husain, Md Tabish Rehman, Asimul Islam, and Md Imtaiyaz Hassan. 2021. “Potential Drug Targets of SARS-CoV-2: From Genomics to Therapeutics.” *International Journal of Biological Macromolecules* 177: 1–9.

https://doi.org/10.1016/j.ijbiomac.2021.02.071.

Sharma, Rakesh. 2012. “Enzyme Inhibition: Mechanisms and Scope.” In *ENZYME INHIBITION AND BIOAPPLICATIONS*, edited by Rakesh Sharma. InTech. https://books.google.se/books?hl=en&lr=&id=mcqgDwAAQBAJ&oi=fnd&pg=PA3&dq=types+of+enzyme+inhibition+&ots=NPPAta3QuI&sig=GXbt2RLVxzlIEIjRtMUv39LtBeE&redir_esc=y#v=onepage&q=types of enzyme inhibition&f=false.

Shepherd, K. A., T. D. Macfarlane, and T. D. Colmer. 2005. “Morphology, Anatomy and Histochemistry of Salicornioideae (Chenopodiaceae) Fruits and Seeds.” *Annals of Botany* 95 (6): 917–33. https://doi.org/10.1093/aob/mci101.

Sherif, Yousery E, Sami A Gabr, Nasser M Hosny, Ahmad H Alghadir, and Rayan Alansari. 2021. “Phytochemicals of *Rhus* Spp. as Potential Inhibitors of the SARS-CoV-2 Main Protease: Molecular Docking and Drug-Likeness Study.” Edited by Weicheng Hu. *Evidence-Based Complementary and Alternative Medicine* 2021: 8814890. https://doi.org/10.1155/2021/8814890.

Shitrit, Alina, Daniel Zaidman, Ori Kalid, Itai Bloch, Dvir Doron, Tali Yarnizky, Idit Buch, et al. 2020. “Conserved Interactions Required for Inhibition of the Main Protease of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2).” *Scientific Reports* 10 (1): 1–11. https://doi.org/10.1038/s41598-020-77794-5.

Shivanika, C., S Deepak Kumar, Ragunathan Venkataraghavan, Tiwari Pawan, A. Sumitha, and B. D. Brindha Devi. 2020. “Molecular Docking, Validation, Dynamics Simulations, and Pharmacokinetic Prediction of Natural Compounds against the SARS-CoV-2 Main-Protease.” *Journal of Biomolecular Structure and Dynamics* 0 (0): 1–27. https://doi.org/10.1080/07391102.2020.1815584.

Shree, Priya, Priyanka Mishra, Chandrabose Selvaraj, Sanjeev Kumar Singh, Radha Chaube, Neha Garg, and Yamini Bhusan Tripathi. 2020. “Targeting COVID-19 (SARS-CoV-2) Main Protease through Active Phytochemicals of Ayurvedic Medicinal Plants – Withania Somnifera (Ashwagandha), Tinospora Cordifolia (Giloy) and Ocimum Sanctum (Tulsi) – a Molecular Docking Study.” *Journal of Biomolecular Structure and Dynamics*, August, 1–14. https://doi.org/10.1080/07391102.2020.1810778.

Sies, Helmut, and Michael J Parnham. 2020. “Potential Therapeutic Use of Ebselen for COVID-19 and Other Respiratory Viral Infections,” no. January.

Souza, Manuel M. de, Carlos Rafael Mendes, Kennia B. Doncato, Eliana Badiale-Furlong, and César S.B. Costa. 2018. “Growth, Phenolics, Photosynthetic Pigments, and Antioxidant Response of Two New Genotypes of Sea Asparagus (Salicornia Neei Lag.) to Salinity under Greenhouse and Field Conditions.” *Agriculture (Switzerland)* 8 (7): 1–18. https://doi.org/10.3390/agriculture8070115.

Stoddard, Shana V., Serena D. Stoddard, Benjamin K. Oelkers, Kennedi Fitts, Kellen Whalum, Kaylah Whalum, Alexander D. Hemphill, et al. 2020. “Optimization Rules for SARS-CoV-2 Mpro Antivirals: Ensemble Docking and Exploration of the Coronavirus Protease Active Site.” *Viruses* 12 (9). https://doi.org/10.3390/v12090942.

Sung, Jong Hyuk, So Hyun Park, Dong Hwan Seo, Ju Hee Lee, Sang Won Hong, and Soon Sun Hong. 2009. “Antioxidative and Skin-Whitening Effect of an Aqueous Extract of Salicornia Herbacea.” *Bioscience, Biotechnology and Biochemistry* 73 (3): 552–56. https://doi.org/10.1271/bbb.80601.

Świderek, Katarzyna, and Vicent Moliner. 2020. “Revealing the Molecular Mechanisms of

Proteolysis of SARS-CoV-2 Mproby QM/MM Computational Methods." *Chemical Science* 11 (39): 10626–30. https://doi.org/10.1039/d0sc02823a.

Tahir ul Qamar, Muhammad, Safar M. Alqahtani, Mubarak A. Alamri, and Ling Ling Chen. 2020. "Structural Basis of SARS-CoV-2 3CLpro and Anti-COVID-19 Drug Discovery from Medicinal Plants." *Journal of Pharmaceutical Analysis* 10 (4): 313–19. https://doi.org/10.1016/j.jpha.2020.03.009.

Tallei, Trina Ekawati, Sefren Geiner Tumilaar, Nurdjannah Jane Niode, Fatimawali, Billy Johnson Kepel, Rinaldi Idroes, Yunus Effendi, Shahenur Alam Sakib, and Talha Bin Emran. 2020. "Potential of Plant Bioactive Compounds as SARS-CoV-2 Main Protease (M^{pro}) and Spike (S) Glycoprotein Inhibitors: A Molecular Docking Study." Edited by Chiara Riganti. *Scientifica* 2020: 6307457. https://doi.org/10.1155/2020/6307457.

Teli, Divya M., Mamta B. Shah, and Mahesh T. Chhabria. 2021. "In Silico Screening of Natural Compounds as Potential Inhibitors of SARS-CoV-2 Main Protease and Spike RBD: Targets for COVID-19." *Frontiers in Molecular Biosciences* 7 (January): 1–25. https://doi.org/10.3389/fmolb.2020.599079.

Todhunter, John A. 1979. "Reversible Enzyme Inhibition." *Methods in Enzymology* 63 (C): 383–411. https://doi.org/10.1016/0076-6879(79)63017-3.

Tomic, Nikolina, Lejla Pojskic, Abdurahim Kalajdzic, Jasmin Ramic, Naida Lojo Kadric, Tarik Ikanovic, Milka Maksimovic, and Naris Pojskic. 2020. "Screening of Preferential Binding Affinity of Selected Natural Compounds to SARS-CoV-2 Proteins Using in Silico Methods." *Eurasian Journal of Medicine and Oncology* 4 (4): 319–23. https://doi.org/10.14744/ejmo.2020.72548.

Trott, Oleg, and Arthur J Olson. 2010. "AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading." *Journal of Computational Chemistry* 31 (2): 455–61. https://doi.org/10.1002/jcc.21334.

Tuan, Nguyen Q, Wonhwa Lee, Joonseok Oh, Soyoung Kwak, Hyun Gyu Lee, Daneel Ferreira, Jong-Sup Bae, and MinKyun Na. 2015. "Quinic Acid Derivatives from Salicornia Herbacea Alleviate HMGB1-Mediated Endothelial Dysfunction." *Journal of Functional Foods* 15: 326–38. https://doi.org/https://doi.org/10.1016/j.jff.2015.03.044.

Ventura, Yvonne, and Moshe Sagi. 2013. "Halophyte Crop Cultivation: The Case for Salicornia and Sarcocornia." *Environmental and Experimental Botany* 92 (August): 144–53. https://doi.org/10.1016/j.envexpbot.2012.07.010.

Vicidomini, Caterina, Valentina Roviello, and Giovanni N Roviello. 2021. "In Silico Investigation on the Interaction of Chiral Phytochemicals from Opuntia Ficus-Indica with SARS-CoV-2 Mpro." *Symmetry* . https://doi.org/10.3390/sym13061041.

Vuong, Wayne, Muhammad Bashir Khan, Conrad Fischer, Elena Arutyunova, Tess Lamer, Justin Shields, Holly A. Saffran, et al. 2020. "Feline Coronavirus Drug Inhibits the Main Protease of SARS-CoV-2 and Blocks Virus Replication." *Nature Communications* 11 (1): 1–8. https://doi.org/10.1038/s41467-020-18096-2.

Wang, Shilei, Jinlei Ye, Zhichao Kang, Hongmei Peng, Vienna Mackey, and Lichun Sun. 2021. "The COVID-19 Pandemic and the Potential Treatment of the Novel Coronavirus SARS-CoV-2." *American Journal of Translational Research* 13 (3): 871–81.

Wang, Xiang-yun, Xu Feng, Ming Wang, Yu Chen, Yun-fa Dong, You-yi Zhao, and Hao Sun. 2011. "[Studies on the chemical constituents of Salicornia europaea]." *Zhong yao cai =*

Zhongyaocai = Journal of Chinese medicinal materials 34 (1): 67–69.

Wang, Xiaomin, Min Zhang, Yuhui Zhao, Hui Wang, Tianxing Liu, and Zhihong Xin. 2013. "Pentadecyl Ferulate, a Potent Antioxidant and Antiproliferative Agent from the Halophyte Salicornia Herbacea." *Food Chemistry* 141 (3): 2066–74. https://doi.org/10.1016/j.foodchem.2013.05.043.

Won, Kyung Jong, Kang Pa Lee, Suji Baek, Long Cui, Mee Hyang Kweon, Seung Hyo Jung, Yun Kyoung Ryu, et al. 2017. "Desalted Salicornia Europaea Extract Attenuated Vascular Neointima Formation by Inhibiting the MAPK Pathway-Mediated Migration and Proliferation in Vascular Smooth Muscle Cells." *Biomedicine and Pharmacotherapy* 94: 430–38. https://doi.org/10.1016/j.biopha.2017.07.108.

Wu, Canrong, Yang Liu, Yueying Yang, Peng Zhang, Wu Zhong, Yali Wang, Qiqi Wang, et al. 2020. "Analysis of Therapeutic Targets for SARS-CoV-2 and Discovery of Potential Drugs by Computational Methods." *Acta Pharmaceutica Sinica B* 10 (5): 766–88. https://doi.org/10.1016/j.apsb.2020.02.008.

Xu, Jiabao, Shizhe Zhao, Tieshan Teng, Abualgasim E Abdalla, Wan Zhu, Longxiang Xie, Yunlong Wang, and Xiangqian Guo. 2020. "Systematic Comparison of Two Animal-to-Human Transmitted Human Coronaviruses: SARS-CoV-2 and SARS-CoV." *Viruses* . https://doi.org/10.3390/v12020244.

Yin, Min, Xiangyun Wang, Ming Wang, Yu Chen, Yunfa Dong, Youyi Zhao, and Xu Feng. 2012. "A New Triterpenoid Saponin and Other Saponins from Salicornia Europaea." *Chemistry of Natural Compounds* 48 (2): 258–61. https://doi.org/10.1007/s10600-012-0216-2.

Zengin, Gokhan, Zaahira Aumeeruddy-Elalfi, Adriano Mollica, Mustafa Abdullah Yilmaz, and Mohamad Fawzi Mahomoodally. 2018. "In Vitro and in Silico Perspectives on Biological and Phytochemical Profile of Three Halophyte Species—A Source of Innovative Phytopharmaceuticals from Nature." *Phytomedicine* 38: 35–44. https://doi.org/https://doi.org/10.1016/j.phymed.2017.10.017.

Zhang, Chun Hui, Elizabeth A. Stone, Maya Deshmukh, Joseph A. Ippolito, Mohammad M. Ghahremanpour, Julian Tirado-Rives, Krasimir A. Spasov, et al. 2021. "Potent Noncovalent Inhibitors of the Main Protease of SARS-CoV-2 from Molecular Sculpting of the Drug Perampanel Guided by Free Energy Perturbation Calculations." *ACS Central Science*. https://doi.org/10.1021/acscentsci.1c00039.

Zhang, Haibo, Josef M. Penninger, Yimin Li, Nanshan Zhong, and Arthur S. Slutsky. 2020. "Angiotensin-Converting Enzyme 2 (ACE2) as a SARS-CoV-2 Receptor: Molecular Mechanisms and Potential Therapeutic Target." *Intensive Care Medicine* 46 (4): 586–90. https://doi.org/10.1007/s00134-020-05985-9.

Zhang, Linlin, Daizong Lin, Xinyuanyuan Sun, Ute Curth, Christian Drosten, Lucie Sauerhering, Stephan Becker, Katharina Rox, and Rolf Hilgenfeld. 2020. "Crystal Structure of SARS-CoV-2 Main Protease Provides a Basis for Design of Improved a-Ketoamide Inhibitors." *Science* 368 (6489): 409–12. https://doi.org/10.1126/science.abb3405.

Zhao, Yuhui, Xiaomin Wang, Hui Wang, Tianxing Liu, and Zhihong Xin. 2014. "Two New Noroleanane-Type Triterpene Saponins from the Methanol Extract of Salicornia Herbacea." *Food Chemistry* 151: 101–9. https://doi.org/https://doi.org/10.1016/j.foodchem.2013.11.030.