

**UNIVERSITÀ DEGLI STUDI DI PADOVA**  
**DIPARTIMENTO DI SCIENZE CHIMICHE**

**CORSO DI LAUREA IN CHIMICA**

**RIBOTACs: Degradazione selettiva dell'RNA di  
SARS-CoV-2 mediante l'utilizzo di piccole  
molecole**

Relatore: Prof. Alice Susic

Laureando: Daniele Pozzebon

ANNO ACCADEMICO 2023/2024



# Indice

|   |           |
|---|-----------|
| <b>Abstract</b> . . . . .   | <b>v</b>  |
| <b>1 Introduzione</b> . . . . .                                       | <b>1</b>  |
| 1.1 Struttura di SARS-CoV-2 . . . . .                                 | 1         |
| 1.2 Replicazione di SARS-CoV-2 . . . . .                              | 3         |
| 1.3 Inibitori di M <sup>pro</sup> . . . . .                           | 5         |
| 1.4 Frameshifting . . . . .   | 6         |
| <b>2 Strategie con RNA come target</b> . . . . .                      | <b>9</b>  |
| 2.1 ASOs . . . . .  | 9         |
| 2.2 siRNA . . . . .   | 10        |
| 2.3 RIBOTACs . . . . .  | 10        |
| <b>3 Utilizzo dell'approccio RIBOTACs contro SARS-CoV-2</b> . . . . . | <b>13</b> |
| 3.1 AbsorbArray . . . . .   | 13        |
| 3.2 C5 . . . . .  | 15        |
| 3.3 C5-PA . . . . .   | 17        |
| 3.4 C5-RIBOTAC . . . . .  | 18        |
| <b>4 Conclusioni</b> . . . . .  | <b>21</b> |
| <b>Bibliografia</b> . . . . .   | <b>23</b> |



# Abstract

La degradazione selettiva dell'RNA è una nuova strategia farmaceutica; in particolare, è molto promettente per la terapia contro il cancro e nelle terapie antivirali.

L'approccio RIBOTACs, acronimo di "Ribonuclease-Targeting Chimeras" utilizza una piccola molecola per indurre la degradazione selettiva dell'RNA grazie alla selettività del legame formato tra la piccola molecola e l'RNA.

L'obiettivo di questa tesi è quello di illustrare un possibile utilizzo della strategia RIBOTACs nella terapia antivirale per il SARS-CoV-2.



# 1. Introduzione

SARS-CoV-2 appartiene alla famiglia *Coronaviridae*, che al suo interno comprende anche SARS-CoV e MERS-CoV (Middle East respiratory syndrome coronavirus). SARS-CoV-2 si diffonde più rapidamente dei suoi predecessori, ma presenta una mortalità più bassa.

Per contrastare la diffusione di SARS-CoV-2, è già presente un farmaco approvato da FDA (Food and Drug Administration), tuttavia è ancora in corso una ricerca di possibili target e di possibili farmaci. Questa tesi si prefigge l'obiettivo di illustrare una possibile strategia farmaceutica, mediante l'uso di piccole molecole bifunzionali chiamate RIBOTAC, che usa l'RNA virale come target.

## 1.1 Struttura di SARS-CoV-2

Il virione SARS-CoV-2 è formato da diverse proteine strutturali<sup>1</sup>, ed esse sono rappresentate in figura 1.1.

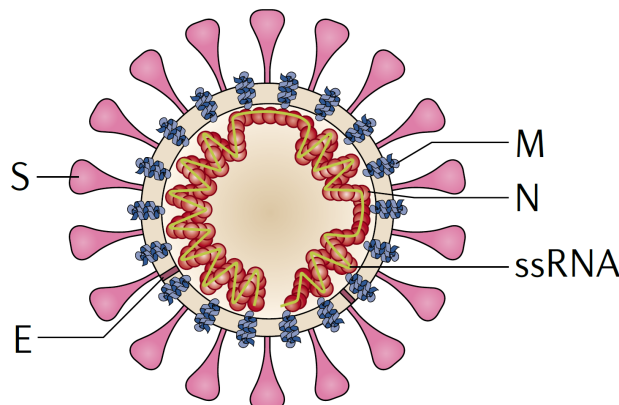


Fig. 1.1: In figura sono rappresentate le proteine strutturali di SARS-CoV-2; esse sono la proteina di membrana (M), la proteina spike (S), la proteina envelope (E), e la proteina nucleocapside (N). (V'kovski, P; et al. 2021)

Esse hanno diverse funzioni:

- la proteina M facilita la sintesi delle altre strutture proteiche all'interno della cellula.
- la proteina S è una glicoproteina ed è un trimero. Questa proteina riconosce e si fonde con la membrana cellulare della cellula ospite, consentendo all'RNA virale di entrare all'interno. Il gene che identifica questa proteina codifica 1300 amminoacidi e richiede

l'intervento di una proteasi per poter essere attivata. Il taglio proteolitico avviene nella parte N-terminale della subunità S1, che viene separata dalla subunità S2. Le subunità S1 e S2 sono eterodimeri e, polimerizzandosi tra loro, formano la proteina S.<sup>2</sup>

- la proteina E è una proteina transmembrana che aiuta il rilascio del virus appena codificato dal citoplasma allo spazio extracellulare.<sup>2</sup>
- la proteina N è la proteina più presente nel SARS-CoV-2. Essa lega l'RNA virale e lo impacchetta, formando il complesso ribonucleoproteico. Inoltre la proteina N diminuisce la risposta antivirale, riducendo l'efficienza dell'RNAi.<sup>2</sup>

Sono presenti altre 16 proteine, dette non-strutturali (nsp), tra cui nsp5, detta M<sup>pro</sup> o *main-protease*. Altre proteine non-strutturali importanti sono nsp7 e nsp8, che coordinano il processo di replicazione virale, e nsp12, che contiene la polimerasi RNA dipendente (RdRp) e che è una componente critica per i processi di replicazione/trascrizione di SARS-CoV-2.<sup>3</sup>

La famiglia dei Coronavirus possiede un singolo filamento di RNA, detto ssRNA (single stranded RNA), che è scritto in senso positivo, quindi può essere direttamente letto dai ribosomi per codificare proteine. L'ssRNA è suddiviso in più Open Reading Frames (ORFs); un Open Reading Frame è una sequenza nucleotidica non contiene codoni di stop all'interno del frame. SARS-CoV-2 ha un genoma che contiene due grandi ORFs, ovvero ORF1a e ORF1ab.

Questi due grandi ORF contengono le informazioni genetiche che codificano le proteine pp1a e pp1ab che, in seguito all'azione proteolitica di M<sup>pro</sup>, formano le proteine non-strutturali.<sup>4</sup> I due ORF con le rispettive proteine sintetizzate sono rappresentati in figura 1.2.

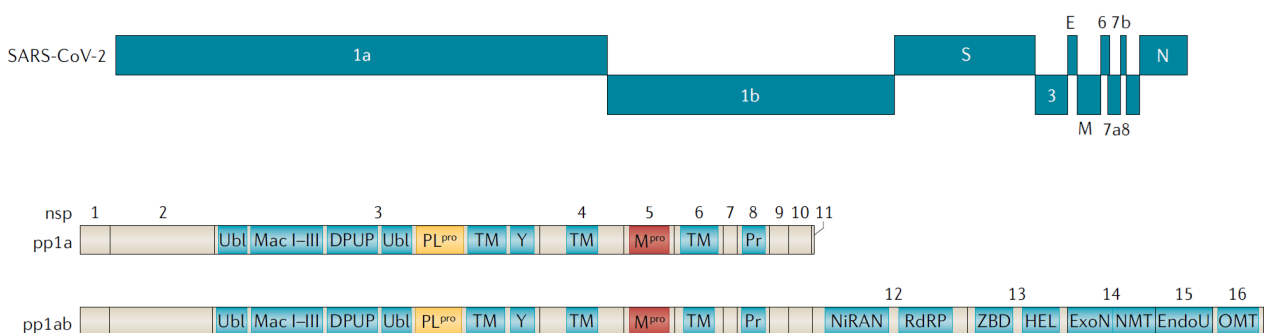


Fig. 1.2: Rappresentazione di ORF1a, ORF1ab e delle rispettive proteine pp1a e pp1ab. (V'kovski, P; et al. 2021)



## 1.2 Replicazione di SARS-CoV-2

SARS-CoV-2 ha bisogno di una cellula ospite per potersi replicare, in quanto è un virus. Una rappresentazione schematica del ciclo virale è esposta in figura 1.3.

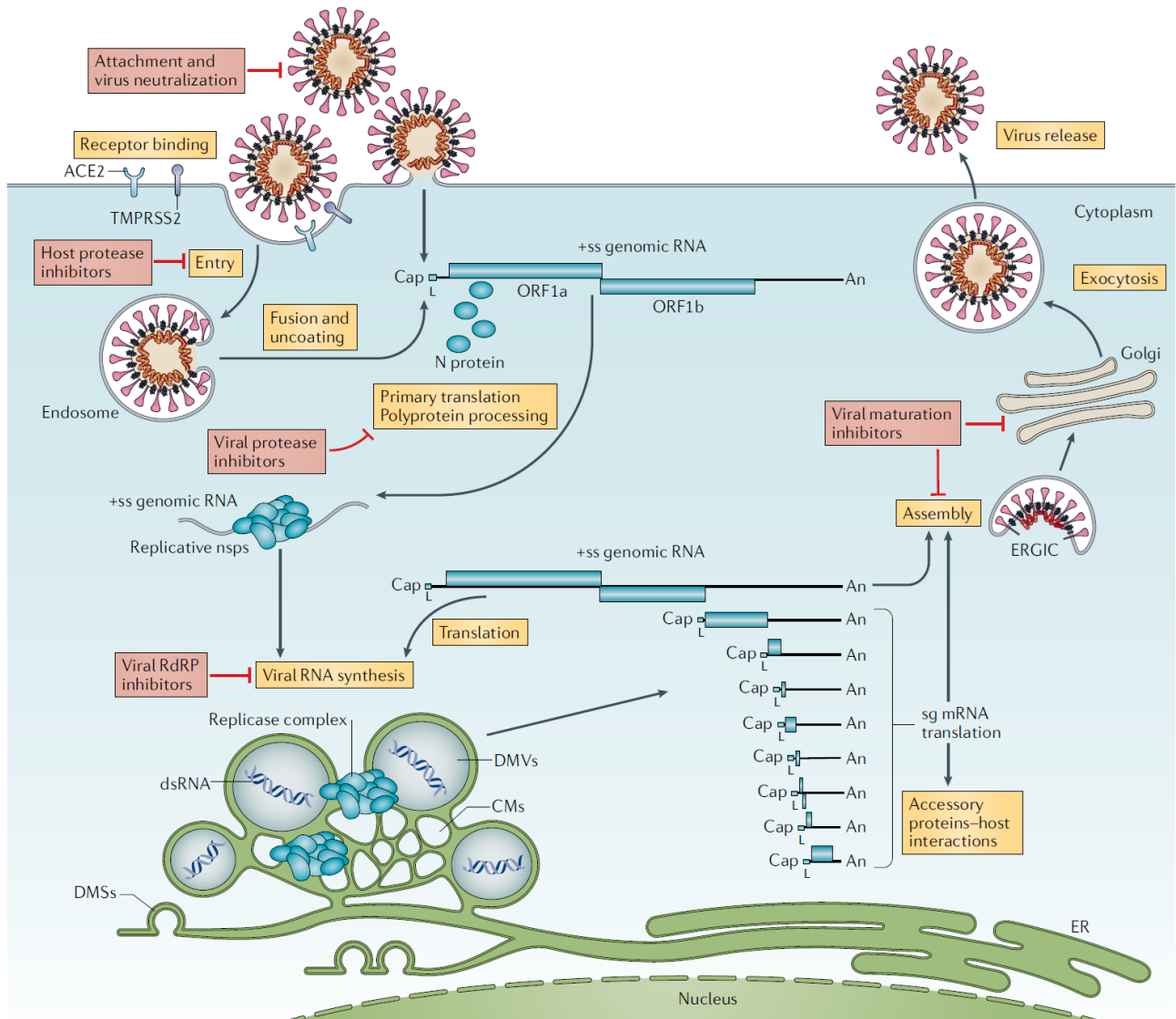


Fig. 1.3: Rappresentazione del ciclo virale di SARS-CoV-2.  
(V'kovski, P; et al. 2021)

La replicazione di SARS-CoV-2 viene iniziata mediante la formazione del legame tra la proteina S e il suo recettore, ovvero ACE2 (angiotensin-converting enzyme 2). Il legame dipende da due fattori: il taglio dei siti S1/S2 di TMPRSS2 (surface transmembrane protease

serine 2) e da una seconda proteasi endogena, chiamata catepsina L, che media la fusione di SARS-CoV-2 alla membrana cellulare. In seguito, l'RNA virale viene rilasciato nel citosol, dove viene codificato per formare le proteine pp1a e pp1ab. Queste proteine in seguito vengono tagliate da M<sup>pro</sup> e formano le proteine non-strutturali nsp.<sup>5</sup> In particolare, tra le proteine non-strutturali è presente RdRp, che consente la replicazione dell'RNA di SARS-CoV-2; questa proteina richiede che altre proteine virali siano presenti, come nsp7 e nsp8.

I due filamenti di RNA sono complementari e vengono separati nella componente positiva e nella componente negativa. (+)ssRNA viene utilizzato come template per la codifica di (-)ssRNA e di (sg)RNA, detto anche subgenomic RNA. La traslazione di (sg)RNA comporta la sintesi delle proteine strutturali S, M, E ed N; le componenti strutturali del virus vengono assemblate in una zona vicina all'apparato di Golgi, chiamata ERGIC. In seguito, viene assemblato il virus tramite l'incorporazione di (+)ssRNA nella struttura formata con le proteine strutturali.<sup>5</sup> Infine, il virione viene secreto dalla cellula per esocitosi.

RdRp contiene, nell'estremità N-terminale, un'estensione del dominio che adotta una struttura detta NiRAN, mentre nell'estremità C-terminale ha una forma a "mano destra" ed è formata da tre subdomini, ovvero un dominio a dito, uno a palmo e uno a pollice. I due domini sono collegati da un  $\beta$ -hairpin, ovvero una struttura che appaia due  $\beta$ -strands in senso antiparallelo. nsp7 e nsp8 vanno a legarsi al dominio a pollice e una seconda unità nsp8 va a legarsi al dominio del dito. Il  $\beta$ -hairpin si va ad inserire nel solco formato dal dominio a palmo e dalla struttura NiRAN.

Quando è presente del RNA virale, il complesso nsp12-nsp7-nsp8 lo lega e forma RdRp-RNA. Questa struttura è simile RdRp libera con alcune caratteristiche peculiari. La subunità nsp12 lega il primo giro di RNA virale duplex, tra il suo dominio a palmo e quello a dito. Il sito attivo si trova nel dominio del palmo ed è formato da cinque elementi, che vengono chiamati motifs A-E. Il motif C lega l'estremità 3' dell'RNA virale. I motif F e G si trovano localizzati nel dominio a dito e consentono il posizionamento del template di RNA. Quando l'RNA duplex esce dal solco, forma un secondo giro che si estende dalla superficie di nsp12. In RdRp non sono presenti elementi che limitano l'estensione dell'RNA duplex. Nell'unità nsp8, l'estremità N-terminale forma un'estensione a forma di  $\alpha$ -elica, e possiede dei residui caricati positivamente che consentono l'interazione con il backbone formato da gruppi fosfato dell'RNA appena sintetizzato. Questo consente a RdRp di essere una proteina specifica per la sintesi di RNA rispetto alla sintesi di DNA.<sup>3,6</sup>

RdRp è il target di un farmaco approvato da FDA, ovvero Remdesivir. Questa molecola si lega covalentemente all'RNA primer e blocca l'allungamento di catena.<sup>7</sup>

### 1.3 Inibitori di M<sup>pro</sup>

M<sup>pro</sup> è una proteina non-strutturale che consente il taglio proteolitico delle poliproteine pp1a e pp1ab in 16 proteine non-strutturali. Questa proteina è formata da due omodimeri, detti protomeri A e B; ogni protomero contiene tre domini, indicati come dominio I, dominio II e dominio III.<sup>3</sup> M<sup>pro</sup> è rappresentata in figura 1.4.

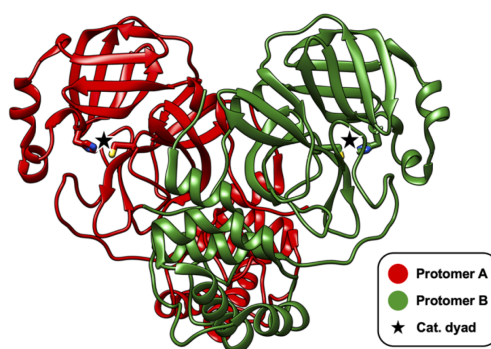


Fig. 1.4: Nell'immagine è rappresentata la main-protease. Sono evidenziati i due siti catalitici, che comprendono l'anello imidazolico dell'istidina e il gruppo tiolico della cisteina. Ogni protomero consiste in tre domini; nel caso di SARS-CoV e di SARS-CoV-2, il dominio I comprende i residui 9-101, il dominio II i residui 102-184 ed entrambi includono un  $\beta$ -barrel che presenta delle similitudini alle proteasi seriniche. Il dominio II è collegato al dominio III (residui 201-306) attraverso una regione di loop. Il dominio III è formato da cinque  $\alpha$ -eliche. I protomeri si legano tra di loro attraverso l'estremità N-terminale (residui 1-7), che sono localizzati in mezzo ai domini II e III. L'estremità N-terminale è coinvolta nella formazione del legame tra il substrato e la proteina.

(Ullrich, S; et al. 2020)

M<sup>pro</sup> è, quindi, una proteasi cisteinica, dove il sito attivo possiede una diade catalitica formata dagli aminoacidi cisteina (Cys145) e istidina (His41). Solitamente le altre proteasi cisteiniche e istidiniche hanno un terzo aminoacido presente nel sito attivo, invece M<sup>pro</sup> possiede una molecola d'acqua al posto di questo aminoacido.

Il processo proteolitico segue un meccanismo multi-step. Inizialmente, il protone della cisteina viene rimosso da un anello imidazolico dell'istidina; si forma quindi un nucleofilo, che quindi attacca il legame ammidico del substrato. Il prodotto di questo attacco nucleofilo riarrangia, e mediante, il trasferimento del protone inizialmente scambiato, forma il prodotto C-terminale, con il ripristino della diade catalitica.<sup>8</sup> Il monomero M<sup>pro</sup> è praticamente inattivo, mentre il suo omodimero è molto attivo.<sup>8</sup>

Questa proteina viene utilizzata come target da un farmaco, Paxlovid, in quanto è essenziale

nella replicazione virale. Paxlovid è un farmaco formato da due piccole molecole, ovvero nirmatrelvir e ritonavir. Solo nirmatrelvir presenta un'effettiva attività farmaceutica, mentre ritonavir viene utilizzato per aumentare la biodisponibilità di nirmatrelvir.<sup>9</sup>

Il meccanismo d'azione di nirmatrelvir prevede la formazione di un legame covalente con la Cys145 nella diade catalitica, inibendo l'azione catalitica della proteasi. Quindi viene a bloccarsi il meccanismo di taglio proteolitico di pp1a e di pp1ab, inibendo la replicazione virale.<sup>9</sup>

## 1.4 Frameshifting

Vista la necessità di nuovi target per SARS-CoV-2, si vuole provare a usare l'RNA come bersaglio. Nell'RNA di SARS-CoV-2, ORF1a e ORF1ab sono collegati tra loro da un elemento di frameshift. Il frameshifting avviene quando vi è un errore nella lettura delle basi azotate nella sintesi proteica.<sup>10</sup>

Sono presenti tre diversi frameshifting: quando il frameshift avviene e il numero di basi saltate è pari a tre, o un suo multiplo, si ha il frameshift in-frame. Questo tipo di frameshift non altera la sequenza amminoacidica complessiva, in quanto i codoni codificati successivamente al frameshift vengono letti correttamente. Questo frameshift prende il nome di *frameshift 0-frame*. Se, invece, il numero di basi saltate è diverso da tre, allora si ha il frameshift out-of-frame. In questo caso, la sequenza amminoacidica successiva viene alterata significativamente. Il frameshift out-of-plane può essere di due tipologie: il *frameshift -1-frame* e il *frameshift +1-frame*. Nel primo caso, si ha uno spostamento del ribosoma dalla posizione iniziale, definita 0, alla posizione -1, quindi legge due volte la stessa base. Nel secondo caso, il ribosoma si sposta alla posizione +1, quindi una base non viene letta.<sup>11</sup>

Alcune proteine prevedono l'utilizzo del frameshift per essere codificate; questo fenomeno viene chiamato Programmed Frameshift (PRF). In particolare, la sintesi proteica viene alterata, con l'effetto che due proteine possono essere codificate dallo stesso filamento di RNA. SARS-CoV-2 utilizza il frameshift per separare i due ORFs, in quanto è un virus con un singolo filamento di RNA.<sup>12</sup>

Nell'RNA di SARS-CoV-2, l'ORF1a è codificato nell'estremità 5', quindi è il primo ad essere letto. Questo ORF codifica le proteine necessarie per prendere il sopravvento sulla cellula ospite in tre modi differenti: assicurandosi l'ingresso nel ribosoma, diminuendo la risposta immunitaria cellulare e tagliando le poliproteine in proteine individuali. ORF1a codifica le prime 11 nsp, mentre ORF1ab codifica le rimanenti 7. Per codificare ORF1ab è necessario un frameshift programmato out-of-plane; nello specifico un -1-PRF.<sup>12</sup> Una rappresentazione schematica degli ORFs e del frameshift element (FSE) di SARS-CoV-2 si trova in figura 1.5.

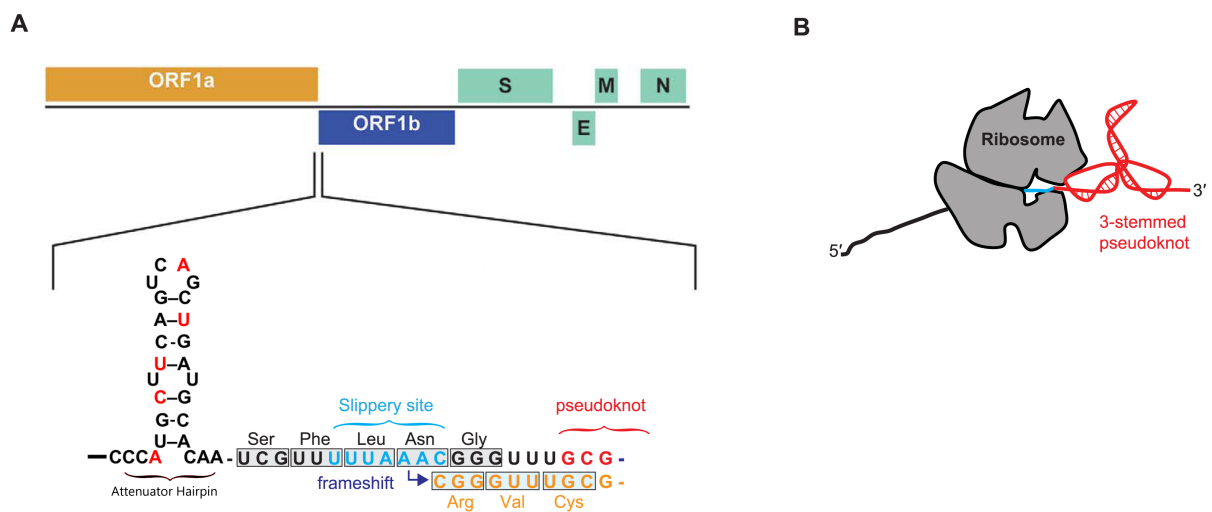


Fig. 1.5: (a) Rappresentazione schematica del ORF di SARS-CoV-2. Nel dettaglio, sono rappresentate le triplette di basi azotate che compongono il genoma e con esse l'amminoacido corrispondente. Si vede quindi che l'ultimo codone (AAC) codifica Asn e, in seguito a -1-PRF, si ha una seconda lettura della base C, nella sequenza CGG, che inizia a codificare un secondo amminoacido, ovvero Arg. Inoltre è rappresentata la struttura dell'hairpin di SARS-CoV-2. (b) Rappresentazione schematica dello pseudoknot (rosso) e dello slippery site (azzurro) all'interno di un ribosoma. (Kelly J. A., et al. 2020; Haniff H. S., et al. 2020)

La zona di RNA virale che causa il -1-PRF viene definita Frameshift Element (FSE), ed è composta da tre elementi:

- un slippery site formato da 7 basi azotate
- uno spacer
- un pseudoknot

Lo slippery site è composto da una sequenza N NNW WWH, dove NNN rappresentano tre basi uguali, WWW rappresentano tre A oppure tre U e H rappresenta una base qualsiasi, ad esclusione di G; gli spazi vuoti nella sequenza rappresentano gli spacer presenti nella sequenza. Il pseudoknot è formato da diversi ripiegamenti di RNA; è una struttura complessa, che tuttavia è stabile. Nel caso di SARS-CoV-2, -1-PRF avviene nella sequenza dello slippery site a causa della presenza del pseudoknot. Il frameshift avviene con elevata efficienza ed è compreso tra il 25 % e il 75 %.<sup>12,13</sup> L'efficienza del frameshifting viene misurata comparando le proteine sintetizzate una volta avvenuto il frameshift con tutte le proteine codificate da un singolo filamento di RNA<sup>10</sup>, ovvero

$$\% \text{ Frameshifting} = 100 \cdot \frac{\text{Out-of-frame}}{\text{In-frame} + \text{Out-of-frame}} \quad (1.1)$$

SARS-CoV e SARS-CoV-2 presentano un ulteriore elemento, chiamato *attenuator hairpin* (AH), che è localizzato prima dello slippery site. Durante la sintesi proteica, il ribosoma trasla sopra l'AH e inizia ad avvicinarsi allo slippery site; quando questo entra all'interno, l'AH viene tagliato dal ribosoma e, con i frammenti di RNA, permette la formazione dello pseudoknot. La sua formazione consente al ribosoma di compiere il -1-PRF.<sup>12</sup> SARS-CoV e il SARS-CoV-2 posseggono entrambi questa regione, come mostrato in figura 1.6

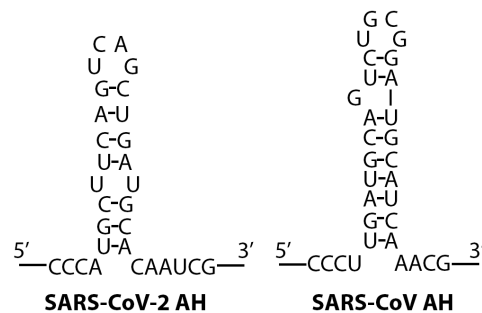


Fig. 1.6: Hairpin (AH) nel RNA del SARS-CoV-2 e di SARS-CoV.

Vi sono tuttavia delle differenze, in quanto nel caso del SARS-CoV-2, si è visto che è presente una piccola tasca, formata da due serie di nucleotidi, ovvero 5'CUU e 3'GUA, detta anche loop UU, mentre nel caso del SARS-CoV non è presente una tasca, in quanto le due sequenze nucleotidiche sono 5'AUG e 3'UAC. Nel caso del SARS-CoV è presente un rigonfiamento formato da 5'AGU e 3'U\_A.

In seguito sono spiegate alcune strategie che utilizzano l'RNA come target.

## 2. Strategie con RNA come target

L'RNA può essere utilizzato come target; sono state sviluppate diverse tecnologie con questo fine. Negli ultimi anni, si è visto in ambito farmaceutico, un interesse sempre più importante sull'utilizzo di grandi molecole a discapito delle piccole molecole, specialmente nel caso delle terapie geniche; le grandi molecole in questione sono gli oligonucleotidi antisenso, detti anche ASO, e i small interfering RNA, detti anche siRNA.

### 2.1 ASOs

Negli ultimi anni, si è visto in ambito farmaceutico, un interesse sempre più importante sull'utilizzo di grandi molecole a discapito delle piccole molecole, specialmente nel caso delle terapie geniche; le grandi molecole in questione sono gli oligonucleotidi antisenso, detti anche ASO.

Gli oligonucleotidi antisenso sono dei singoli filamenti di DNA artificiale contenenti dalle 13 alle 25 basi azotate, che sono codificate in senso opposto, ovvero 3'→5' rispetto al normale senso di trascrizione, ovvero 5'→3'.<sup>14</sup>

I frammenti di DNA vanno quindi ad appaiarsi all'mRNA, inducendone la degradazione. In questo modo è possibile limitare la sintesi proteica di una specifica proteina, in modo molto selettivo.

Gli oligonucleotidi antisenso si dividono in due categorie:

- RNAsi dipendenti
- RNAsi indipendenti

I primi si legano al frammento di RNA corrispondente e inducono la degradazione del del complesso ASO-RNA per via endogena, ovvero tramite l'enzima ribonucleasi. I secondi, invece, fungono da ingombro sterico, impedendo la traduzione dell'informazione genetica in proteine.

Vi sono alcuni vantaggi nell'utilizzo della prima categoria di ASO: in primis, si ha un'efficienza molto alta, pari a circa 80-95%, mentre in secondo luogo, l'RNAsi può degradare qualsiasi sequenza nucleotidica, mentre, utilizzando un bloccante sterico, si possono solo legare determinate regioni spaziali.<sup>14</sup>

L'utilizzo di ASO presenta diversi vantaggi rispetto alle piccole molecole, in particolare nello studio e nell'ottimizzazione della farmacocinetica e della farmacodinamica, in quanto le piccole

molecole non sono determinabili in modo predittivo, ma devono essere verificate con studi in vitro e in vivo.<sup>14</sup>

## **2.2 siRNA**

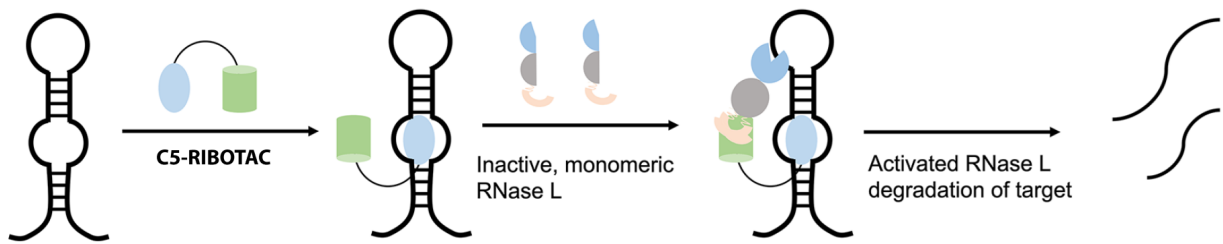
Uno sviluppo della tecnologia ASOs è rappresentato dai siRNA (small interfering RNA), che consiste nell'utilizzare dei frammenti di RNA non solo per controllare, ma addirittura per silenziare l'espressione genica di un singolo gene. I siRNA sono simili ai miRNA, tuttavia mentre i primi si legano ad una specifica porzione di gene, con un aumento di selettività, i secondi si legano indiscriminatamente a più geni, con quindi una diminuzione di selettività.

I siRNA presentano dei vantaggi rispetto all'utilizzo di piccole molecole, poiché essi hanno un'interazione migliore con il target, in quanto hanno un appaiamento completo con i nucleotidi del frammento di RNA interessato; le piccole molecole, invece, necessitano di un'interazione specifica con il target, che deve essere studiata e sviluppata affinché la piccola molecola abbia un appaiamento corretto. Potenzialmente, i siRNA possono legarsi a qualsiasi frammento di RNA, in quanto è solo necessario conoscere la sequenza nucleotidica corrispondente. Tuttavia, i siRNA presentano anche significativi svantaggi rispetto alle piccole molecole, in quanto queste molecole, se non protette nel modo corretto, non sono stabili alle RNAsi endogene e possono essere degradate prima del raggiungimento del target. In secondo luogo, essendo grandi molecole, presentano una farmacocinetica difficile da gestire, e possono dare luogo a effetti collaterali. Questi effetti collaterali sono dovuti ad un appaiamento non completo con regioni di RNA di altri geni.<sup>15</sup>

## **2.3 RIBOTACs**

La tecnologia RIBOTACs (Ribonuclease Targeting Chimeras) discende dall'approccio PROTACs (Proteolysis Targeting Chimeras). In questa tecnologia, una piccola molecola, che instaura il legame con un target di origine proteica, viene funzionalizzata con una seconda molecola, che consente il legame con una seconda proteina, ovvero l'ubiquitina. Questa proteina viene utilizzata per segnalare alla cellula che la proteina è difettosa e deve subire un processo di degradazione nei lisosomi; questo meccanismo consente quindi di eliminare la proteina target utilizzando un meccanismo endogeno della cellula.<sup>16</sup>





SARS-CoV-2 AH

Fig. 2.1: Rappresentazione schematica del funzionamento di un RIBOTAC, che consente la degradazione dell'RNA. Questa rappresentazione, in particolare, indica la degradazione dell'RNA di SARS-CoV-2 ad opera di C5-RIBOTAC, una molecola bifunzionale che verrà esposta nel seguente capitolo.

(Haniff H. S., et al. 2020)

L'approccio RIBOTACs (Ribonuclease Targeting Chimeras) utilizza delle piccole molecole, dove gli altri metodi utilizzano degli oligonucleotidi<sup>14</sup> per instaurare l'interazione con l'acido nucleico. I RIBOTACs infatti sono simili ai PROTACs; essi utilizzano una molecola che possiede due siti di legame per instaurare una prima interazione con il target e una seconda interazione con una proteina endogena, che viene utilizzata dal corpo per degradare il target. In particolare, per i PROTACs, si lega l'ubiquitina, mentre per i RIBOTACs, si utilizza l'enzima ribonucleasi, che catalizza l'idrolisi dell'RNA a nucleotidi, degradandolo.

I RIBOTACs hanno dei vantaggi rispetto ai nucleotidi antisense in quanto sono piccole molecole e quindi vengono distribuite in più tessuti. In secondo luogo, i RIBOTACs, una volta che si sono legati all'RNA target e lo hanno degradato, possono legarsi nuovamente, mentre gli ASO vengono degradati insieme al target; questo consente ai RIBOTACs di essere efficaci a concentrazioni catalitiche. Uno svantaggio dei RIBOTACs rispetto ai ASO è che sono più difficili da sviluppare, in quanto è necessario trovare una piccola molecola che legghi in modo selettivo la sequenza di RNA di interesse.<sup>17</sup>



## 3. Utilizzo dell'approccio RIBOTACs contro SARS-CoV-2

Un modo per prevenire il -1 PRF del ribosoma dall'RNA virale è quello di legare una piccola molecola alla FSE, quindi stato effettuato uno studio per capire quale piccola molecola ha delle interazioni compatibili con l'RNA virale su una collezione di 3 271 molecole, delle quali è noto il legame che formano con l'RNA. I dati della collezione di molecole sono ospitati nel database Inforna, che cataloga le interazioni, ottenute sperimentalmente, tra le piccole molecole e l'RNA.<sup>18</sup>

Per vedere che struttura molecolare è necessaria per instaurare un buon legame, si è guardato in letteratura<sup>18,19</sup> ed è stato fatto un iniziale HTS (High Throughput Screening), per capire quali strutture avessero il miglior legame.

### 3.1 AbsorbArray

Per testare l'intensità del legame dei leganti, è stato fatto un secondo screening, chiamato AbsorbArray<sup>20,21</sup>; che è rappresentato schematicamente nell'immagine 3.1.

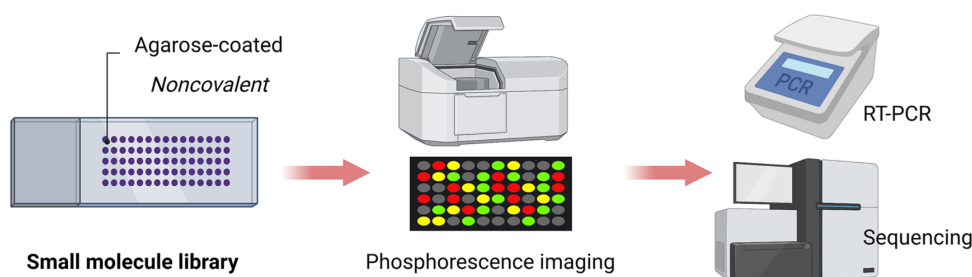


Fig. 3.1: Rappresentazione schematica di AbsorbArray. Questo screening non necessita l'immobilizzazione della molecola in modo covalente, in quanto viene assorbita in uno strato di agarosio e viene lavata. In questo modo, la molecola interagisce con il target, ovvero l'RNA virale, nello stesso modo in cui interagirebbe in soluzione.  
(Velagapudi S. P., et al. 2018)

Questo screening prevede la deposizione di 25 mL di gel di agarosio 1% (w/v) e la successiva deposizione di 100 nL del composto sul quale viene effettuato lo screening. In seguito, il gel viene lavato due volte con acqua NanoPure e viene asciugato.

Una volta preparato il microarray con i composti sul quale viene effettuato lo screening, vengono depositati 200 pmol di RNA wild-type di SARS-CoV-2, nel quale il FSE è stato marcato

con un radioisotopo. L'RNA, inoltre, è stato pre-ibridizzato con un filamento complementare di RNA. In seguito è stata aggiunta una soluzione buffer, per consentire il folding dell'RNA. All'RNA, in seguito, è stato addizionato del BSA (Bovine Serum Albumin), fino a che una concentrazione di BSA pari a  $40\mu/mol$  non è stata raggiunta. L'aggiunta di BSA è servita per migliorare l'accuratezza dei risultati, in quanto impedisce perdite di acido nucleico durante i successivi lavaggi. Il microarray sono stati incubati per 30 minuti e in seguito lavati ed asciugati ed sono stati posti ad imaging. In questo screening sono stati selezionati i composti che hanno mostrato un segnale 1000 volte più elevato rispetto al segnale di background.

Da questo screening sono stati identificati 26 leganti; questi composti sono stati classificati in tre categorie, ovvero i fenil-benzil-idimazoli, le benzil tiouree e i 2,4-diammino-chinazoline; esse sono rappresentate in figura 3.2.

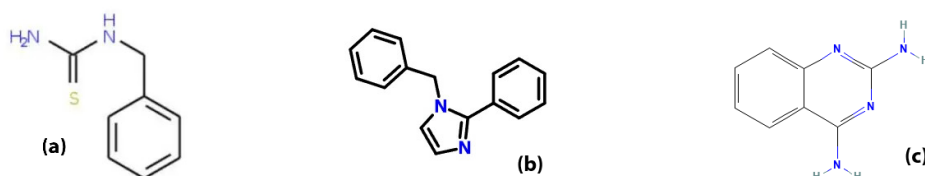


Fig. 3.2: Scaffold delle molecole che formano un buon legame con l'RNA di SARS-CoV-2. In particolare si vede (a) la benzil-tiourea, (b) il fenil-benzil-idimazolo e (c) la 2,4-diammino-chinazolina.

Queste 26 molecole sono state sottoposte ad un test competitivo mediante RT-qPCR.<sup>22</sup> La tecnica RT-qPCR (Reverse Transcription quantitative real-time PCR) permette una determinazione in tempo reale dell'RNA mediante l'utilizzo di una spettroscopia di emissione. Questa tecnica viene resa più sensibile utilizzando due enzimi, ovvero

- una trascrittasi inversa, che permette di trascrivere l'RNA virale in cDNA (complementary DNA).
- una DNA polimerasi, che permette di formare un duplex di DNA a partire da cDNA e di aumentarne la quantità.

In seguito, è stato utilizzato un mix di coloranti fluorescenti, chiamato Power SYBR Green Master Mix (Life Technologies), per misurare la quantità di DNA presente. Questa tecnica permette di ottenere una determinazione quantitativa del DNA presente, in quanto il segnale di luminescenza è proporzionale al segnale di coloranti legati al DNA.

Questo test è stato effettuato con due tipologie di RNA virale, ovvero l'RNA virale wild-type e l'RNA virale dove il loop interno UU è stato sostituito con due nucleotidi appaiati, ovvero

AU. Questo test è stato fatto per per confermare la specificità del legame e si è visto che le 2,4-diammino-chinazoline e le benzil-tiouree non si sono legate all'RNA mutante, e quindi il loro legame con il loop interno UU è specifico. Questo screening ha consentito di ridurre ancora il numero di molecole considerate a cinque. Le molecole sono state chiamate C1-C5 e sono rappresentate in figura 3.3.<sup>23</sup>

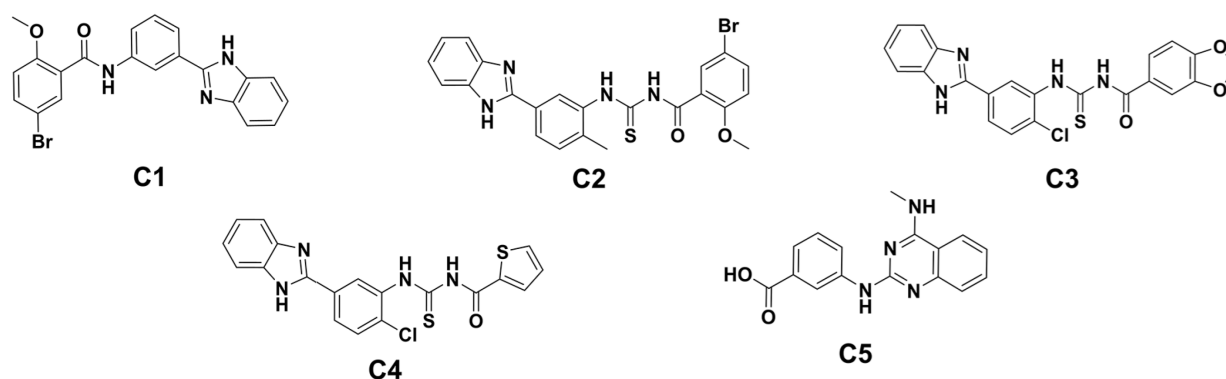


Fig. 3.3: Rappresentazione delle molecole ottenute dopo il test competitivo.

## 3.2 C5

Per comparare la diminuzione di efficienza del frameshift in queste cinque molecole sono state coltivate le cellule HEK293T ed è stato inserito al loro interno un filamento di RNA, in cui sono espresse, in ordine, la luciferasi di Renilla e la luciferasi di lucciola. Queste due sequenze nucleotidiche sono state separate dal frameshift element di SARS-CoV-2, di SARS-CoV. Le cellule HEK293T prendono il nome da (Human Embryonic Kidney 293 T) e sono cellule utilizzate per la crescita virale in laboratorio.<sup>24</sup>

Come controllo, le due sequenze nucleotidiche sono state codificate senza un elemento di frameshift, in un ORF che prende il nome di ORF-CTRL. I campioni sono stati incubati per 24 ore e in seguito sono state misurate le attività delle due luciferasi tramite "Dual-Glo Luciferase Assay Kit" (Promega; catalog # E2980).

La codifica di queste due luciferasi è stata fatta in modo sequenziale; prima è codificata la luciferasi di lucciola e poi, separata dal frameshift element, è stata codificata la luciferasi appartenente a Renilla, come in figura 3.4

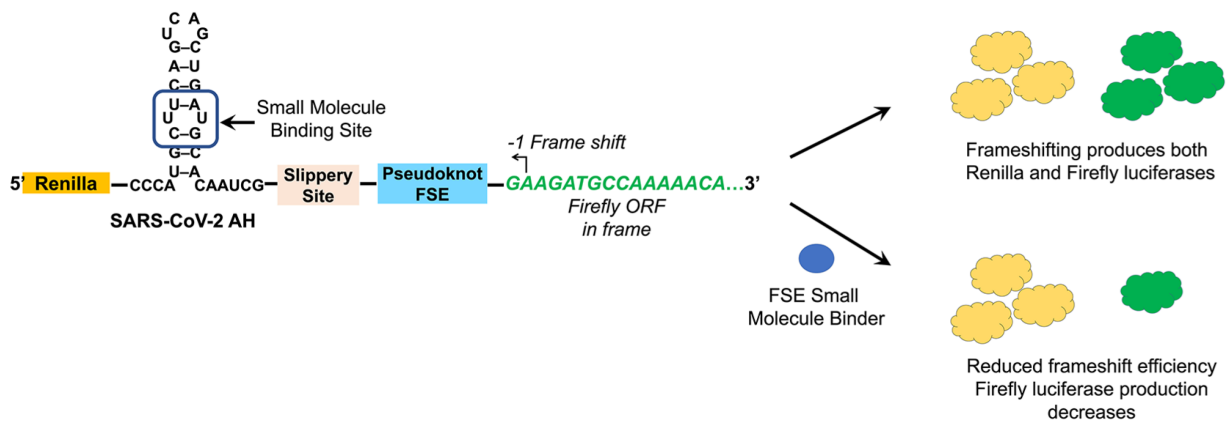


Fig. 3.4: Rappresentazione schematica del sistema utilizzato per verificare l'efficienza di frameshift. Nell'estremità 5' è presente un ORF che codifica la luciferasi di Renilla, mentre nell'estremità 3' è presente un ORF che codifica la luciferasi di lucciola. Il secondo ORF, per essere sintetizzato, necessita di un frameshift out-of-plane, in particolare di  $-1$ -frameshift. Le piccole molecole possono legarsi all'attenuator hairpin, in modo tale da ridurre l'efficienza di frameshift, abbassando quindi i livelli della luciferasi di lucciola. (Haniff H. S., et al. 2020)

L'efficienza di frameshift è stata calcolata determinando il rapporto tra il segnale di luminescenza della luciferasi considerata e il segnale ottenuto dalla codifica di ORF-CTRL, ovvero tramite la formula

$$\text{Efficienza Frameshifting} = \frac{\frac{F_T}{R_T}}{\frac{F_C}{R_C}} \quad (3.1)$$

dove  $F_T$  e  $R_T$  sono i segnali del campione con l'elemento di frameshift, mentre  $F_C$  e  $R_C$  sono i segnali del controllo.

Dai dati contenuti nel grafico 3.5, si vede come il composto C5 abbassa l'efficienza di frameshift del  $25\% \pm 1$ , mentre gli altri composti non lo abbassano in modo significativo. Si vede, inoltre, che l'abbassamento dell'efficienza del frameshift è proporzionale alla dose di C5 somministrata. L'abbassamento di frameshift visto si ottiene con una concentrazione di C5 pari a  $2\mu M$ . Questa molecola, inoltre, presenta una selettiva riduzione del frameshift solo nel caso dell'utilizzo del frameshift element del SARS-CoV-2, mentre non ha nessun effetto se si utilizza il frameshift element del SARS-CoV; questo indica una selettività del C5 verso il loop UU, che è presente esclusivamente nel FSE di SARS-CoV-2. Per capire l'intensità del legame tra C5 e il loop UU, è stata misurata l'affinità al sito di legame tramite la costante di dissociazione  $K_d$ , che è risultata essere pari a  $11 \pm 3.7nM$ , quindi particolarmente elevata. Il C5 appartiene alla classe delle 2,4-diammino-chinazoline.

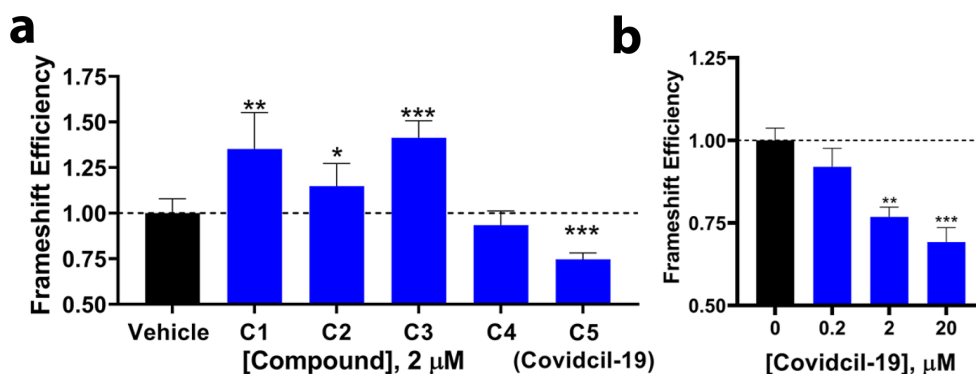


Fig. 3.5: (a) Effetti delle piccole molecole. L'efficienza del frameshifting è stata calcolata come indicato nella formula 3.1. (b) L'effetto di riduzione del frameshifting è dovuto alla dose utilizzata. Il composto in particolare è C5. L'efficienza di frameshifting è stata calcolata come in (a).

### 3.3 C5-PA

Per capire come funzionalizzare C5, sono stati sostituiti dei gruppi funzionali; in particolare, il gruppo carbossilico è stato sostituito con un gruppo propilammidico, ottenendo un composto che prende il nome di C5-PA, che è rappresentato in figura 3.6.

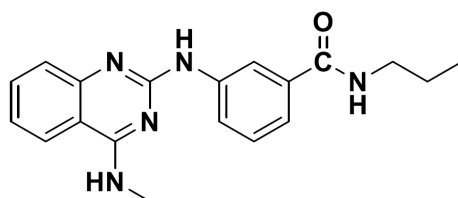


Fig. 3.6: Rappresentazione del C5-PA.

C5-PA presenta una minore riduzione di efficienza del frameshift rispetto a C5, come dimostrato nei grafici in figura 3.7

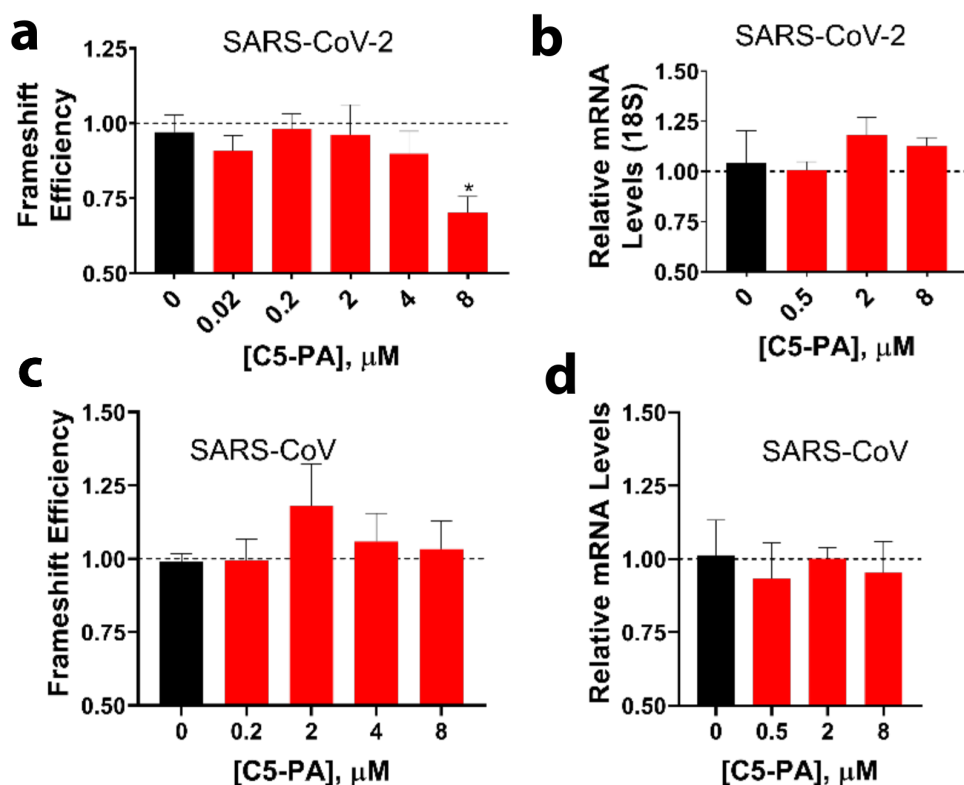


Fig. 3.7: (a) Riduzione dell'efficienza di frameshifting per C5-PA. Anche in questo caso, la riduzione di efficienza è dose-dipendente. (b) C5-PA comporta una riduzione di efficienza del frameshift per SARS-CoV-2, tuttavia non diminuisce la quantità di RNA presente. (c) e (d) C5-PA non presenta effetti né di riduzione del frameshift, né di riduzione dell'RNA di SARS-CoV.

Tuttavia C5-PA presenta una costante di dissociazione  $K_d$  pari a  $4 \pm 1nM$ , che quindi è più elevata rispetto a quella di C5, il che conferma che il gruppo carbossilico può essere sostituito e al suo posto è possibile aggiungere altri gruppi funzionali. In questi grafici si vede anche che C5-PA continua a presentare una selettività per il loop UU del AH di SARS-CoV-2 rispetto all'AH di SARS-CoV. Inoltre, si vede che l'effetto di abbassamento dell'efficienza del frameshift continua a essere dipendente dalla dose di composto utilizzata. Infine, si vede che la presenza di questo composto non riduce la quantità di RNA presente, né per SARS-CoV-2, né per SARS-CoV.

### 3.4 C5-RIBOTAC

Avendo confermato la presenza del legame selettivo tra C5 e l'attenuator hairpin di SARS-CoV-2, si è voluto provare a migliorarlo ulteriormente, utilizzando l'approccio RIBOTAC. L'approccio RIBOTAC è già stato presentato e prevede appunto il reclutamento di un'enzima



endogeno deputato all'idrolisi dell'RNA. Questo approccio permette di utilizzare una quantità stechiometrica di farmaco, in quanto esso non viene immediatamente degradato dalla cellula, ma può agire più volte. Quindi, è stato sintetizzato il C5-RIBOTAC tramite l'aggiunta di un linker, legato al sito dell'acido carbossilico, e di un sistema eterociclico che funge da reclutatore per l'RNAsi, come mostrato in figura 3.8.

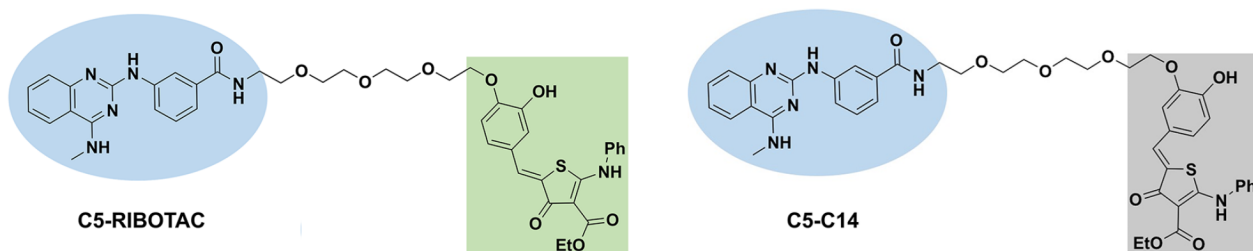


Fig. 3.8: Rappresentazione di C5-RIBOTAC e di C5-C14.

Come controllo, è stato sintetizzato un secondo RIBOTAC, chiamato C5-C14, raffigurato in figura 3.8, che differisce nella tipologia del sistema eterociclico; in questo caso l'eterociclo ha una minore capacità di reclutare l'RNAsi. Si è visto che nessuno dei due composti, nelle concentrazioni biologicamente attive, causa effetti off-target.

Rispetto al C5, il C5-RIBOTAC non si limita all'inibizione del frameshifting, ma causa la degradazione dell'RNA. Infatti, la riduzione di attività di luciferasi di Renilla comporta che nemmeno la prima proteina nel frammento di RNA viene codificata e questo è indice della degradazione dell'RNA causata da RNAsi. La riduzione è dipendente dalla quantità di composto utilizzata, come confermato dai dati espressi in figura 3.9.

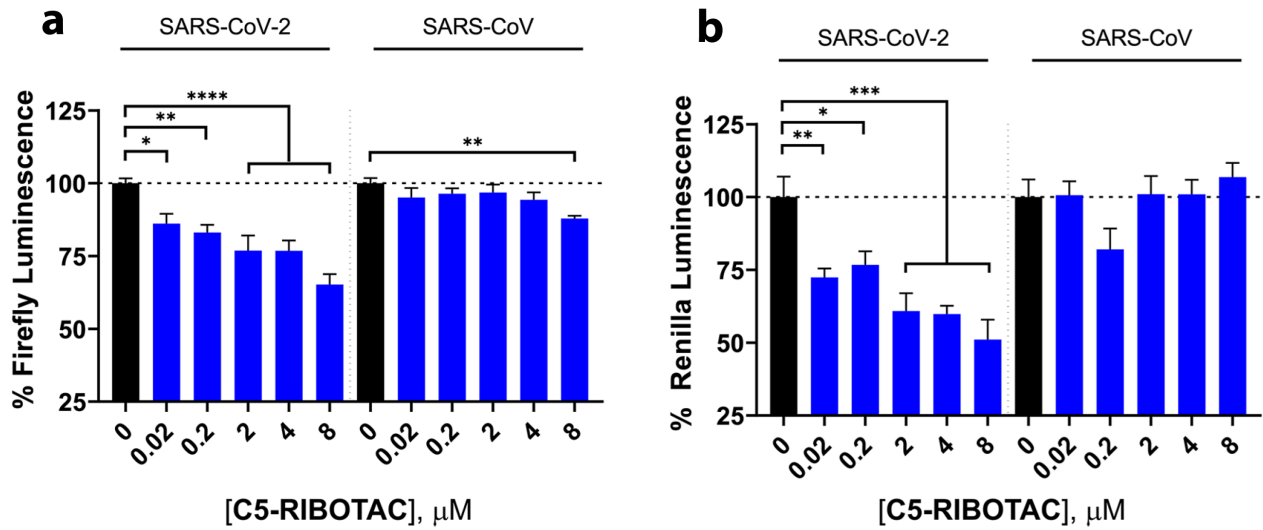


Fig. 3.9: (a) Effetto di C5-RIBOTAC sull'attività di luciferasi di lucciola. (b) Effetto di C5-RIBOTAC sull'attività di luciferasi proveniente da Renilla. (c) e (d) Si vede inoltre che C5-RIBOTAC non ha effetti significativi se viene utilizzato l'attenuator hairpin di SARS-CoV.

Per ottenere un abbassamento di quantità di luciferasi di lucciola pari al 25 %, è stata utilizzata una concentrazione di C5-RIBOTAC 10 volte più piccola, ovvero  $0.2\mu\text{M}$  rispetto alla medesima riduzione di frameshift vista nel caso di C5, ovvero  $2\mu\text{M}$ , quindi si può dire che C5-RIBOTAC è 10 volte più potente di C5. Si vede, inoltre, che C5-RIBOTAC è selettivo e non comporta sostanziali riduzioni quando è stato utilizzato l'AH di SARS-CoV, mentre si è rivelato molto efficiente nel caso dell'AH di SARS-CoV-2.

Il RIBOTAC di controllo, ovvero il C5-C14, ha comunque prodotto una diminuzione di efficienza nel frameshifting, tuttavia questo effetto è da considerarsi una conseguenza del legame tra la molecola di C5 e l'AH del SARS-CoV-2.

## 4. Conclusioni

Concludendo, si è visto che la strategia RIBOTAC comporta dei vantaggi rispetto ad una semplice molecola. In particolare, il composto C5-RIBOTAC ha una azione 10 volte più potente rispetto alla semplice molecola C5. Inoltre, si è visto la possibilità di avere diversi target per un singolo agente patogeno con un possibile aumento delle strategie farmaceutiche impiegate per una possibile cura. Si è visto, inoltre, che è necessaria una costante ricerca in ambito farmaceutico per poter essere preparati a epidemie future.

In ogni caso, l'approccio RIBOTAC offre diversi utilizzi, oltre all'uso illustrato finora, ovvero potrebbe essere efficace contro alcune tipologie di cancro e come trattamento per le malattie genetiche.<sup>17</sup>



# Bibliografia

- (1) V'kovski, P.; Kratzel, A.; Steiner, S.; Stalder, H.; Thiel, V. *Nature Reviews Microbiology* **2021**, *19*, 155–170.
- (2) Hardenbrook, N. J.; Zhang, P. *Current Opinion in Virology* **2022**, *52*, 123–134.
- (3) Wang, M.-Y.; Zhao, R.; Gao, L.-J.; Gao, X.-F.; Wang, D.-P.; Cao, J.-M. *Front Cell Infect Microbiol* **2020**, *10*, 587269.
- (4) Davidson, A. D.; Williamson, M. K.; Lewis, S.; Shoemark, D.; Carroll, M. W.; Heesom, K. J.; Zambon, M.; Ellis, J.; Lewis, P. A.; Hiscox, J. A.; Matthews, D. A. *Genome Medicine* **2020**, *12*, 68.
- (5) Harrison, A. G.; Lin, T.; Wang, P. *Trends Immunol* **2020**, *41*, 1100–1115.
- (6) Hillen, H.; Kokic, G.; Farnung, L.; Dienemann, C.; Tegunov, D.; Cramer, P. *Nature* **2020**, *584*, 154–156.
- (7) Bekheit, M. S.; Panda, S. S.; Girgis, A. S. *European Journal of Medicinal Chemistry* **2023**, *252*, 115292.
- (8) Ullrich, S.; Nitsche, C. *Bioorganic & Medicinal Chemistry Letters* **2020**, *30*, 127377.
- (9) Akinosoglou, K.; Schinas, G.; Gogos, C. *Viruses* **2022**, *14*, 2540.
- (10) Tinoco Jr, I.; Kim, H.-K.; Yan, S. *Biopolymers* **2013**, *99*, 1147–1166.
- (11) Giedroc, D. P.; Cornish, P. V. *Virus Res* **2008**, *139*, 193–208.
- (12) Kelly, J. A.; Woodside, M. T.; Dinman, J. D. *Virology* **2021**, *554*, 75–82.
- (13) Bhatt, P. R.; Scaiola, A.; Loughran, G.; Leibundgut, M.; Kratzel, A.; Meurs, R.; Dreos, R.; O'Connor, K. M.; McMillan, A.; Bode, J. W.; Thiel, V.; Gatfield, D.; Atkins, J. F.; Ban, N. *Science* **2021**, *372*, 1306–1313.
- (14) Dias, N.; Stein, C. A. *Molecular Cancer Therapeutics* **2002**, *1*, 347–355.
- (15) Hu, B.; Zhong, L.; Weng, Y.; Peng, L.; Huang, Y.; Zhao, Y.; Liang, X.-J. *Signal Transduction and Targeted Therapy* **2020**, *5*, 101.
- (16) Zou, Y.; Ma, D.; Wang, Y. *Cell Biochemistry and Function* **2019**, *37*, 21–30.
- (17) Dey, S. K.; Jaffrey, S. R. *Cell Chemical Biology* **2019**, *26*, 1047–1049.
- (18) Velagapudi, S. P.; Gallo, S. M.; Disney, M. D. *Nature Chemical Biology* **2014**, *10*, 291–297.
- (19) Haniff, H. S.; Knerr, L.; Liu, X.; Crynen, G.; Boström, J.; Abegg, D.; Adibekian, A.; Lekah, E.; Wang, K. W.; Cameron, M. D.; Yildirim, I.; Lemurell, M.; Disney, M. D. *Nature Chemistry* **2020**, *12*, 952–961.

- (20) Velagapudi, S. P.; Costales, M. G.; Vummidi, B. R.; Nakai, Y.; Angelbello, A. J.; Tran, T.; Haniff, H. S.; Matsumoto, Y.; Wang, Z. F.; Chatterjee, A. K.; Childs-Disney, J. L.; Disney, M. D. *Cell Chem Biol* **2018**, *25*, 1086–1094.
- (21) Zhao, R.; Fu, J.; Zhu, L.; Chen, Y.; Liu, B. *Journal of Hematology & Oncology* **2022**, *15*, 14.
- (22) Adams, G. *The Biochemist* **2020**, *42*, 48–53.
- (23) Haniff, H. S.; Tong, Y.; Liu, X.; Chen, J. L.; Suresh, B. M.; Andrews, R. J.; Peterson, J. M.; O’Leary, C. A.; Benhamou, R. I.; Moss, W. N.; Disney, M. D. *ACS Central Science* **2020**, *6*, 1713–1721.
- (24) Kelly, J. A.; Olson, A. N.; Neupane, K.; Munshi, S.; San Emeterio, J.; Pollack, L.; Woodside, M. T.; Dinman, J. D. *J Biol Chem* **2020**, *295*, 10741–10748.