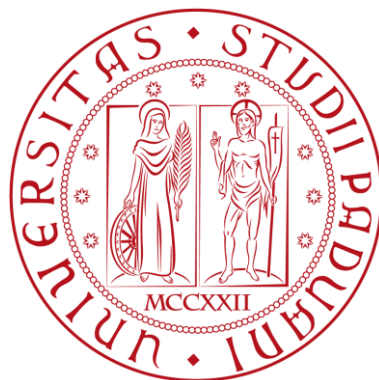


# UNIVERSITA' DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE DEL FARMACO

**CORSO DI LAUREA MAGISTRALE IN FARMACIA**



TESI DI LAUREA MAGISTRALE

MiniFrag: sviluppo di una strategia di screening innovativa  
tramite NMR contro la proteina Bfl-1

RELATORE: Prof. Mattia Sturlese

LAUREANDO: Enrico Penzo

Matricola: 1140250

Anno Accademico: 2021/2022



*A mio nonno Aristide,  
Guida costante nel mio percorso.*





# INDICE

<b>1</b>	<b>SOMMARIO</b>	<b>1</b>
<b>2</b>	<b>INTRODUZIONE</b>	<b>3</b>
<b>2.1</b>	<b>TUMORE</b>	<b>3</b>
<b>2.2</b>	<b>MORTE CELLULARE</b>	<b>6</b>
2.2.1	NECROSI	7
2.2.2	APOPTOSI	7
<b>2.3</b>	<b>MEDIATORI PRO E ANTI APOPTOTICI</b>	<b>12</b>
2.3.1	RUOLO NELLA GENESI TUMORALE	14
<b>2.4</b>	<b>BFL-1</b>	<b>15</b>
2.4.1	GENE BCL2A1	18
2.4.2	INIBITORI DELLE PROTEINE ANTIAPOPTOTICHE BCL-2	19
<b>2.5</b>	<b>FRAGMENT-BASED DRUG DISCOVERY TRAMITE NMR</b>	<b>24</b>
<b>2.6</b>	<b>MINIFRAG</b>	<b>29</b>
<b>3</b>	<b>OBIETTIVI DELLA TESI</b>	<b>31</b>
<b>4</b>	<b>MATERIALE E METODI</b>	<b>33</b>
<b>4.1</b>	<b>TRASFORMAZIONE ATTRAVERSO IL PLASMIDE CHAMPION™ PET SUMO</b>	<b>33</b>
<b>4.2</b>	<b>ESPRESSIONE DI BFL-1 IN TERRENO ARRICCHITO CON <sup>15</sup>N</b>	<b>36</b>
<b>4.3</b>	<b>ESTRAZIONE MEDIANTE SONICAZIONE</b>	<b>40</b>
<b>4.4</b>	<b>PURIFICAZIONE TRAMITE FAST PROTEIN LIQUID CHROMATOGRAPHY (FPLC)</b>	<b>41</b>
4.4.1	IMAC I	42
4.4.2	DESALTING E TAGLIO CON SUMO PROTEASI	43
4.4.3	IMAC II	44
4.4.4	SEC	44
<b>4.5</b>	<b>QUANTIFICAZIONE TRAMITE SPETTROFOTOMETRIA UV</b>	<b>45</b>
<b>4.6</b>	<b>SDS PAGE</b>	<b>47</b>
<b>4.7</b>	<b>SPETTROSCOPIA NMR, PRINCIPI E APPLICAZIONI</b>	<b>48</b>
4.7.1	CHEMICAL SHIFT	51
4.7.2	FID E TRASFORMATA DI FOURIER	52
4.7.3	<sup>1</sup> H- <sup>15</sup> N SOFAST HMQC E CHEMICAL SHIFT PERTURBATION (CSP)	53
4.7.4	COPRESENZA	59
<b>5</b>	<b>RISULTATI SPERIMENTALI</b>	<b>61</b>
<b>5.1</b>	<b>SINTESI DI BFL-1 PER VIA RICOMBINANTE</b>	<b>61</b>

<b>5.2</b>	<b>PURIFICAZIONE</b>	<b>61</b>
5.2.1	IMAC I	62
5.2.2	DESALTING	63
5.2.3	IMAC II	64
5.2.4	GEL FILTRATION	65
<b>5.3</b>	<b>ANALISI QUANTITATIVA TRAMITE SPETTROFOTOMETRIA UV</b>	<b>66</b>
<b>5.4</b>	<b>SDS PAGE</b>	<b>67</b>
<b>5.5</b>	<b>MINIFRAG COSTRUZIONE DELLO STOCK</b>	<b>69</b>
<b>5.6</b>	<b>PREPARAZIONE DEL CAMPIONE PER L'ANALISI NMR</b>	<b>71</b>
<b>5.7</b>	<b>ANALISI DEGLI SPETTRI NMR</b>	<b>73</b>
5.7.1	A1	74
5.7.2	A2	77
5.7.3	B6	79
5.7.4	C2	81
5.7.5	C3	83
5.7.6	C4	85
5.7.7	C5	87
5.7.8	C6	89
5.7.9	C7	91
<b>5.8</b>	<b>COPRESENZE</b>	<b>94</b>
5.8.1	COPRESENZA A2+B6	96
5.8.2	COPRESENZA A2+C5	98
5.8.3	COPRESENZA A2+C7	100
5.8.4	COPRESENZA A2+C7+C2	102
<b>6</b>	<b>CONCLUSIONI</b>	<b>105</b>
<b>7</b>	<b>BIBLIOGRAFIA</b>	<b>109</b>

---

# 1 SOMMARIO

Questo progetto di tesi fonda le sue basi sulla ricerca di molecole in grado di inibire la proteina antiapoptotica Bcl-1. Si tratta di una proteina facente parte della famiglia delle Bcl-2, che si è scoperto venire sovraespressa in tumori sia ematologici che solidi.

Nei tumori la sovraespressione di tale proteina facilita la progressione della malattia in quanto rende possibile l'elusione dei meccanismi fisiologici di morte cellulare che permettono l'eliminazione di cellule alterate, e potenzia anche i meccanismi di farmaco resistenza tumorale, rendendo inefficaci le terapie convenzionali.

La ricerca è stata eseguita mediante la tecnica *Fragment Based Drug Discovery* (FBDD) effettuata tramite spettroscopia NMR. La tecnica prevede lo studio dell'interazione tra piccole molecole (massimo 300 Da), dette 'frammenti', e la proteina di interesse. In questo progetto si è voluto portare al limite minimo la tecnica FBDD, utilizzando i MiniFrag, frammenti ancora più piccoli contenenti un massimo di atomi pesanti tra 5 e 7 (massimo 100 Da). L'utilizzo dei MiniFrag è stato documentato da un solo articolo, che però ha utilizzato come tecnica analitica la cristallografia a raggi X; perciò l'innovazione apportata dal lavoro svolto in questa tesi è data dal verificare la possibilità di svolgere un'analisi similare tramite NMR e di creare una metodologia di analisi dei dati.

Sono stati analizzati dieci MiniFrag, scelti attraverso metodi computazionali dal laboratorio MMS (Prof. Stefano Moro) che, a sua volta, ha voluto portare al limite l'analisi computazionale di molecole così piccole. Perciò i risultati che si sono ottenuti sperimentalmente permetteranno anche di validare e affinare le analisi effettuate tramite software.

Il lavoro che è stato svolto consta di tre parti principali: la produzione, per via ricombinante, e purificazione, tramite *Fast Protein Liquid Chromatography* (FPLC) della proteina di interesse; la preparazione dei campioni da analizzare tramite NMR bidimensionale; l'analisi dei dati, analizzando il *Chemical Shift Perturbation* (CSP), e la loro comparazione con i dati forniti dall'MMS.



## 2 INTRODUZIONE

### 2.1 Tumore

*“I tumori o neoplasie sono l’insieme di una popolazione di cellule somatiche che ha preso origine quasi sempre da una sola cellula dell’organismo che ha subito una serie sequenziale di peculiari danni genomici (mutazioni) trasmissibili alla progenie.”<sup>1</sup>*

Le mutazioni a cui le cellule tumorali vanno incontro sono di vario tipo:

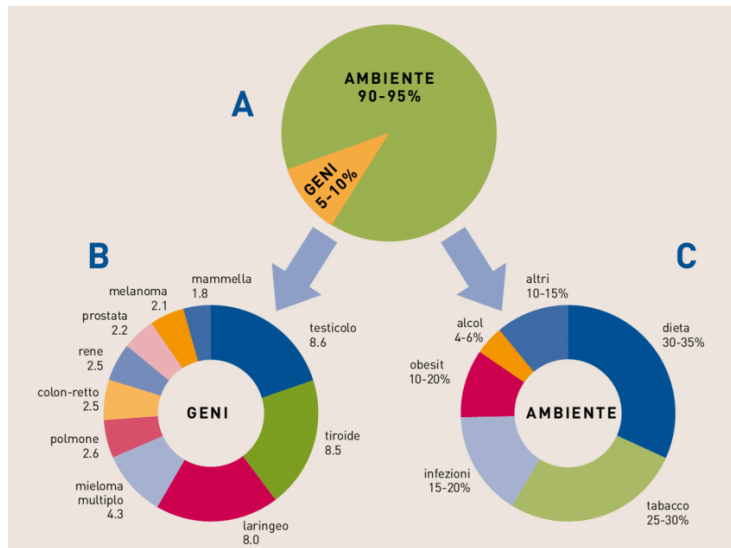
- Perdita dei meccanismi deputati al controllo proliferativo, con conseguente sovra-espressione delle cellule mutate;
- Riduzione o perdita della capacità differenziativa;
- Incapacità di andare incontro a morte cellulare programmata;
- Capacità di esprimere o silenziare geni, codificanti per meccanismi di sopravvivenza, che in condizioni fisiologiche non verrebbero espressi o silenziati, questi permettono alle cellule mutate di sfuggire ai meccanismi di difesa dell’organismo.

Una cellula, per diventare tumorale, richiede che si manifestino più mutazioni a carico di determinati geni. Si può dire che non esista un unico tipo di tumore, ma ne esistono un elevatissimo numero, distinti sulla base dell’intensità e della localizzazione del danno genomico.<sup>1</sup>

I geni maggiormente coinvolti nella neoplasia sono: oncogeni, geni che promuovono la crescita cellulare e che vengono attivati; oncosoppressori, geni che inibiscono la crescita cellulare e che vengono inibiti; sequenze coinvolte nei meccanismi di riparazione del DNA, che vengono silenziate e sequenze coinvolte nella morte cellulare programmata, le quali vengono modulate per inibire l’apoptosi. Con la mutazione di tali geni si ottengono delle cellule aventi una proliferazione illimitata nel tempo, dette ‘cellule immortalizzate’, caratteristica comune a tutti i tipi di tumore.<sup>2</sup>

L’eziologia tumorale è di difficile indagine, i fattori di rischio coinvolti sono numerosi e spesso vi è una concausa tra fattori esogeni e fattori endogeni che concorrono all’insorgenza. Gli agenti in grado di favorire la formazione o la proliferazione tumorale vengono definiti ‘agenti cancerogeni’ e possono essere di diversa natura, come si evidenzia

nella *FIGURA 2.1*. È possibile notare come la componente genetica, abbia un peso inferiore rispetto alla componente ambientale, tuttavia non trascurabile; infatti se i meccanismi deputati alla riparazione del DNA sono alterati nel momento in cui si verificasse una mutazione genetica, questi non saranno in grado di ripristinare il normale corredo genetico. Tra i fattori di rischio ambientali primeggiano una dieta errata; il fumo di tabacco; infezioni date da batteri; obesità e consumo di sostanze alcoliche.<sup>3</sup>



*Figura 2.1 – Fattori di rischio nella genesi tumorale<sup>3</sup>*

I fattori di rischio precedentemente citati non si può dire siano causa diretta dell'insorgenza di una neoplasia, tuttavia ne esistono altri, seppur meno comuni, dei quali si è osservata una stretta correlazione tra fattore di rischio e insorgenza tumorale, tra questi si possono trovare: radiazioni UV, ionizzanti e raggi X; agenti chimici come arsenico e asbesto; agenti biologici come virus (*Papilloma virus*) e batteri (*Helicobacter Pylori*).

Tutte le sostanze cancerogene, al momento conosciute, sono state elencate e classificate dall'*Agenzia Internazionale per la Ricerca sul Cancro (IARC)*; la classificazione che viene proposta (*FIGURA 2.2*) è basata sulla relazione causale tra sostanza e tumore.<sup>4</sup>

<b>Group 1</b>	Carcinogenic to humans	121 agents
<b>Group 2A</b>	Probably carcinogenic to humans	93 agents
<b>Group 2B</b>	Possibly carcinogenic to humans	320 agents
<b>Group 3</b>	Not classifiable as to its carcinogenicity to humans	501 agents

*Figura 2.2 – Classificazione IARC degli agenti cancerogeni in relazione al loro grado di pericolosità<sup>4</sup>*

Lo sviluppo di una neoplasia è un processo complesso e che, come è già stato detto, richiede una concomitanza di più fattori e di molto tempo per avvenire, tuttavia il processo può essere riassunto e suddiviso in tre parti:<sup>2</sup>

- Una prima parte in cui si ha l'*iniziazione*, in cui avviene il danno genetico e/o se ne accumulano di diversi;
- Una seconda parte detta *promozione*, dove entrano in gioco dei fattori esogeni, che di per sé non sono in grado di provocare tumore, ma che sono in grado di promuovere la proliferazione della cellula mutata e/o di potenziare i suoi meccanismi antiapoptotici;
- L'ultimo step del processo è la *progressione*, che porta all'acquisizione di nuove mutazioni e alla comparsa di diversi fenotipi, in grado di rendere la cellula più resistente, più aggressiva o in grado di formare metastasi.

Secondo l'*Organizzazione Mondiale della Sanità* (OMS), il cancro è una delle principali cause di morte nel mondo, con circa 10 milioni di decessi nel 2020, statisticamente quasi una morte su sei è causata dal tumore. I tumori più comuni sono quelli al seno, al polmone, al colon e alla prostata (FIGURA 2.3).<sup>5</sup>

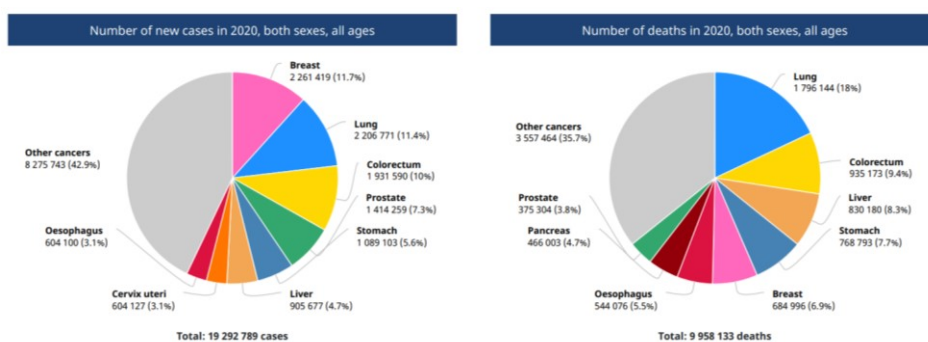


Figura 2.3 – Nuovi casi e morti di tumore nel 2020 nel mondo.<sup>5</sup>

In Italia, in particolare, nel 2020 si è assistito ad una forte incidenza di nuovi casi di tumore, su una popolazione di 60.461.828 cittadini, si sono riscontrati ben 415.269 nuovi casi, in particolare i più incidenti sono stati quelli al seno e alla prostata; inoltre sono stati registrati ben 174.759 decessi dovuti a questa patologia, in particolare la mortalità maggiore è stata riscontrata per tumori al seno e al polmone (FIGURA 2.4).<sup>5</sup>

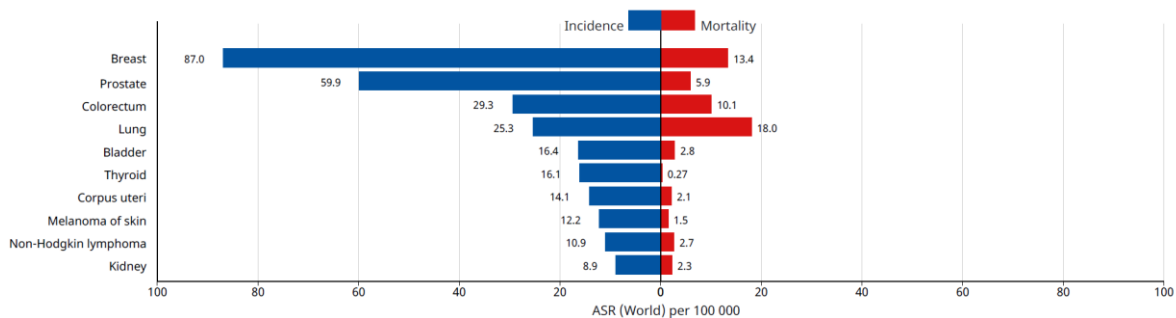


Figura 2.4 – Incidenza e mortalità del tumore in Italia nel 2020<sup>5</sup>

In accordo con la stima fatta dall'OMS, è possibile affermare che circa il 30-50% delle neoplasie può essere prevenuto evitando i fattori di rischio, quindi seguendo uno stile di vita sano e stando attenti a non entrare in contatto con eventuali sostanze mutagene, e facendo controlli periodici o prendendo parte a screening, così da giungere ad una eventuale diagnosi precoce, la quale garantisce maggiori probabilità di sopravvivenza.

Sfortunatamente una cura univoca e che garantisca un'elevata percentuale di guarigione ancora non è stata scoperta, questo perché la cellula neoplastica è in grado, esprimendo o silenziando specifici geni, di aumentare la sua resistenza ai farmaci o alle terapie che impiegano radiazioni ionizzanti. L'elevata resistenza alle terapie, unita alla possibilità di metastatizzare e formare una nuova massa tumorale, con differenti caratteristiche e in un differente sito dell'organismo, rende molto complesso lo studio di una terapia efficace, anche se annualmente vengono pubblicate nuove e promettenti ricerche in merito. In questa tesi l'approccio vuole essere di ricerca di una potenziale molecola, che in futuro potrebbe diventare un farmaco, in grado di favorire l'apoptosi bloccando le proteine antiapoptotiche sovraesprese nella cellula tumorale; così da poter essere usato in associazione a farmaci in grado di indurre l'apoptosi o comunque la morte cellulare.<sup>6</sup>

## 2.2 Morte cellulare

La morte cellulare è un processo necessario per la sopravvivenza e lo sviluppo dell'organismo e per questo può assumere diverse forme; può essere la conseguenza di un processo non fisiologico e incontrollato provocato da una lesione del tessuto, oppure può essere la conseguenza di un complesso processo fisiologico di crescita e differenziazione dell'organismo che risponde a stimoli sia interni che esterni.



Questo processo di morte cellulare viene solitamente distinto in due processi: necrosi, definita come una forma di morte accidentale e passiva dovuta ad un ambiente ostile nel quale la cellula non è stata in grado di adattarsi; e apoptosi, definita al contrario come meccanismo cellulare di suicidio programmato, nel quale la cellula prende parte attivamente e che è dovuto a segnali fisiologici e specifici per permette la preservazione dell'organismo. La più grande differenza tra le due è come il processo si manifesta morfologicamente.<sup>7</sup>

### **2.2.1 Necrosi**

La necrosi è una modalità di morte cellulare passiva e non programmata ed è indotta da agenti non fisiologici; la risposta dell'organismo solitamente è un'inflammatione acuta che però può causare ulteriori danni cellulari.

Il processo è caratterizzato da una progressiva perdita dell'integrità della membrana plasmatica, con conseguente aumento delle dimensioni della cellula; rottura del nucleo e degli organelli cellulari, i quali rilasciano tutto il loro contenuto all'interno della cellula; dilatazione dei mitocondri, che non potranno più provvedere alla respirazione cellulare; infine si avrà rottura della membrana cellulare, evento che porterà l'organismo a riconoscere come estranee le componenti presenti prima all'interno della cellula e a rispondere attivando il processo dell'inflammatione.<sup>1</sup>

Le cause dell'apoptosi possono essere: ipossia, mancato apporto di sangue e nutrienti, ipertermia, esposizione a radiazioni o a composti tossici e infezioni.

La necrosi può verificarsi anche nel tumore se dovesse crescere in maniera eccessiva e non dovessero essere presenti sufficienti vasi sanguigni per l'apporto di ossigeno e nutrienti; inoltre questa possibilità di provocare necrosi è il mezzo sfruttato da chemioterapia, radioterapia e ipertermia, trattamenti antitumorali 'tradizionali'.<sup>8</sup>

### **2.2.2 Apoptosi**

L'apoptosi fa parte di un processo chiamato morte cellulare programmata (PCD, dall'inglese *Programmed Cell Death*); viene attuata nel momento in cui determinate cellule non siano più necessarie alla sopravvivenza dell'organismo, così da mantenere la dimensione e la funzionalità di ogni tessuto, inoltre viene attuata anche su cellule infettate da patogeni o che presentano delle alterazioni genomiche, quindi come meccanismo di autodifesa.<sup>9</sup>

Le alterazioni del DNA, che sono alla base della formazione di una neoplasia, possono essere causate sia da errori nella replicazione, dati dalle polimerasi, sia da stress ambientale dato da radiazioni ultraviolette, radiazioni ionizzanti o molecole in grado di legare il DNA alterandone la replicazione; queste mutazioni provocano una perdita del controllo proliferativo e l'impossibilità di andare incontro ad apoptosi.

La morfologia apoptotica è esattamente contraria rispetto a quella necrotica, infatti, in questo caso, si potrà assistere ad una riduzione della dimensione della cellula, con un aumento della densità del citoplasma e un maggior impacchettamento degli organelli. Nel nucleo la cromatina si presenta estremamente condensata e tende ad aggregare sulla parete della membrana nucleare, la quale tende poi a dividersi in più frammenti. Sulla membrana cellulare tendono poi a formarsi delle estroflessioni, le quali successivamente si staccano formando i cosiddetti corpi apoptotici, formati da piccoli pezzi di membrana cellulare contenenti materiale citoplasmatico. I corpi apoptotici vengono poi riconosciuti dai macrofagi, fagocitati ed in fine distrutti dagli enzimi lisosomiali. Tutto questo processo permette l'eliminazione controllata di cellule senza l'attivazione del sistema infiammatorio.<sup>9</sup>

L'apoptosi viene distinta fondamentalmente in due macro categorie:

- **Via estrinseca:** in cui i recettori presenti sulla membrana plasmatica vengono attivati da specifici ligandi.
- **Via intrinseca:** viene attivata da diversi tipi di stress intracellulare, principalmente legati ad uno stress ossidativo dei mitocondri.

L'attivazione del processo è regolata e reversibile, in quanto controllata da una serie di stimoli contrastanti, che possono essere segnali di morte o di sopravvivenza. Nel momento in cui i segnali di morte, prevalgono sui segnali di vita si va in contro alla morte della cellula; inoltre questi segnali, che inizialmente seguono due vie differenti, convergono poi nell'attivazione di una famiglia di cistein-proteasi chiamate caspasi.<sup>8</sup>

Le caspasi (**cysteinyl aspartate-specific proteinases**), sono enzimi che vengono attivati mediante un taglio proteolitico e a loro volta sono coinvolte nell'idrolisi specifica di legami peptidici dopo un residuo di acido aspartico.

La **via estrinseca** (*Figura 2.5*) ha inizio con il legame tra un ligando extracellulare e il suo recettore sulla membrana cellulare, esistono vari possibili ligandi, i più conosciuti e studiati

sono: *Tumor Necrosis Factor- $\alpha$*  (TNF- $\alpha$ ), che lega il recettore *Tumor Necrosis Factor Receptor* (TNFR), e *Fas Ligand* (FasL), il quale lega il recettore Fas; il primo ligando è una citochina prodotta dal sistema immunitario e dai macrofagi che viene rilasciata solitamente in presenza di tossine batteriche, mentre il secondo è una proteina sempre appartenente alla famiglia dei TNF, ma che viene prodotta dai linfociti T.<sup>10</sup>

I recettori appena visti sono recettori transmembrana aventi una coda amminoacidica nel citosol chiamata *Death Domain* che interagisce, se il recettore viene attivato, con una proteina facente parte delle *Docking Protein*<sup>11</sup>, proteina che lega la procaspasi-8 e ne provoca l'oligodimerizzazione, così da permetterne l'autolisi e l'attivazione della cascata di attivazione delle caspasi.

La procaspasi-8 e -10, vengono così attivate dalla *docking protein* e si trasformano in caspasi -8 e -10, che a loro volta convertono in forma attiva le procaspasi -3, -6 e -7; queste ultime andranno ad interagire con vari enzimi e proteine come: *poly-ADP-ribosylpolymerase* (PARP); lamina (proteina nucleare con funzione strutturale e trascrizionale) e  $\alpha$ -fodrina (proteina citoscheletrica). L'attivazione di queste ultime proteine è lo step definitivo del meccanismo molecolare dell'apoptosi, si tratta di uno step irreversibile che porta alla frammentazione in corpi apoptotici.

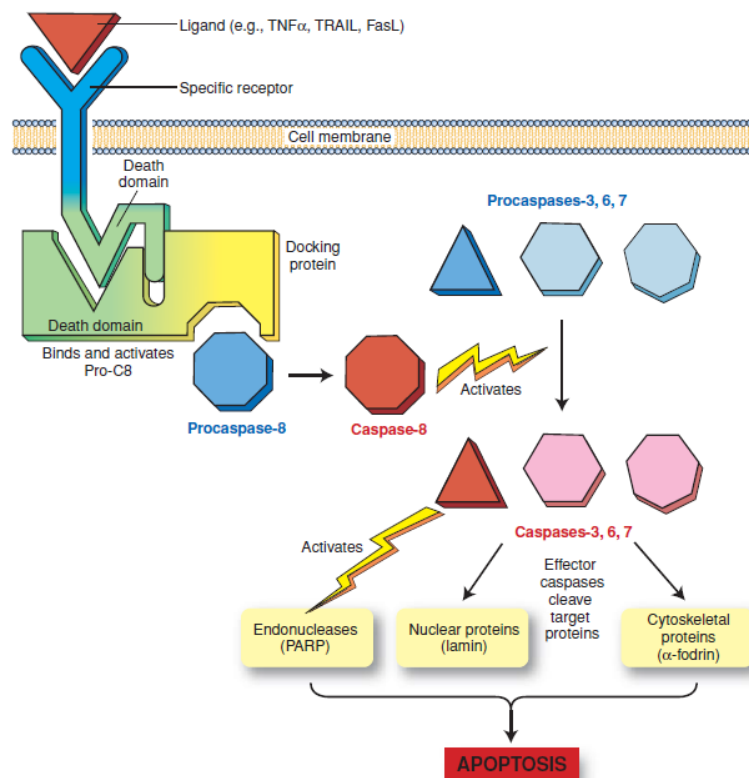


Figura 2.5 – Via estrinseca dell'apoptosi <sup>7</sup>

La via estrinseca dell'apoptosi è un processo finemente regolato, il recettore TNF oltre ad andare ad attivare le caspasi promuovendo la morte cellulare, attiva anche una famiglia di proteine chiamate NFκB; si tratta di fattori di trascrizione nucleare, che codificano per una serie di proteine anti-apoptotiche.

Le vie intrinseca ed estrinseca sono due facce della stessa medaglia, nonostante abbiano due meccanismi di attivazione differenti, ad un certo punto della cascata biochimica si avrà un'intersezione delle due; in particolare nel momento in cui la caspasi-8 viene attivata, oltre ad attivare le altre caspasi, andrà anche ad idrolizzare una proteina plasmatica chiamata Bid (*BH<sub>3</sub>-only protein*). Si tratta di una proteina proapoptotica facente parte della famiglia delle Bcl-2.

Bid una volta attivata (tBid) potrà entrare nel mitocondrio e lì legarsi a Bax/Bak (proteine antiapoptotiche), provocandone l'oligodimerizzazione e la formazione di pori sulla membrana mitocondriale esterna (MOMP) con la conseguente fuoriuscita dei mediatori apoptotici facenti parte della seconda fase della via intrinseca.<sup>12</sup>

La **via intrinseca** dell'apoptosi inizia a livello mitocondriale, gli stimoli che ne permettono l'attivazione sono molteplici: si possono avere un aumento del calcio intracellulare, accumulo di ROS (specie reattive dell'ossigeno), presenza di danni irreparabili al DNA, diminuzione dei fattori di crescita.

Questa via presenta un passaggio chiave che avviene a seguito degli stimoli precedentemente descritti: il rilascio del citocromo c dallo spazio intermembrana del mitocondrio al citosol.

Il citocromo c è una proteina solubile, presente a livello mitocondriale, essenziale per la catena del trasporto degli elettroni e quindi la produzione di ATP.

La membrana mitocondriale interna è una barriera impermeabile che permette di contenere gli organelli facenti parte della matrice mitocondriale; questa membrana è attraversata dalla proteina *Mitochondrial Permeability Transition Pore* (MPTP), una proteina canale che solitamente si trova nella conformazione chiusa, mentre ancorate al lato interno della membrana si trovano una serie di altre molecole coinvolte nell'apoptosi: citocromo c, *Second Mitochondria-derived Activator of Caspases* (Smac/diablo, proteina che promuove l'attivazione delle caspasi), *Apoptosis-Inducing Factor* (AIF, fattore coinvolto nell'induzione

dell'apoptosi). Sempre a livello della membrana mitocondriale interna è importante sottolineare la presenza di un potenziale di membrana ( $\Psi_m$ ) tra l'interno della membrana mitocondriale (potenziale negativo) e il citosol (potenziale positivo); fondamentale per la produzione di ATP.<sup>9</sup>

Il meccanismo apoptotico intrinseco (*FIGURA 2.6*) inizia quando nei mitocondri vi è un accumulo di  $Ca^{2+}$ , un'eccessiva presenza di ROS o uno squilibrio nel potenziale di membrana, tutto ciò provoca l'apertura del canale MPTP.

Un importante meccanismo che viene attivato qualora si inizino ad avere dei considerevoli danni al DNA è l'attivazione di una serie di proteine tra cui la più conosciuta è p53; queste permettono la trascrizione di geni coinvolti nella riparazione del DNA stesso. Tuttavia se il meccanismo di riparazione non dovesse essere in grado di far fronte ai danni già avvenuti, la stessa p53 andrà ad attivare la trascrizione di geni codificanti per proteine proapoptotiche.

L'apertura di MPTP provoca un flusso di protoni ( $H^+$ ) e sali verso la matrice mitocondriale. Il flusso di soluti provocato dall'apertura delle proteine canale provoca un drastico squilibrio del potenziale di membrana ( $\Psi_m$ ), ciò comporta l'impossibilità da parte del mitocondrio di far funzionare la catena di trasporto degli elettroni e con ciò la produzione di ATP. Oltre alla mancata produzione di ATP il passaggio di soluti provoca anche un rigonfiamento dei mitocondri.

La membrana mitocondriale esterna (*MOM, Mitochondrial Outer Membrane*), a causa dello squilibrio che si viene a creare all'interno del mitocondrio, diventa sempre più permeabile, fintanto che si può arrivare o alla rottura della stessa o alla formazione di pori dovuta all'eterodimerizzazione tra le proteine proapoptotiche della famiglia Bcl-2 e Bid.

Nel momento in cui la membrana mitocondriale esterna non funge più da barriera impermeabile si ha avrà il rilascio nel citosol di tutti i componenti della matrice mitocondriale (AIF, Smac/diablo, Citocromo c), con conseguente inizio della seconda fase dell'apoptosi.

Nella seconda fase, si assiste al legame tra il Citocromo c e *Apoptosis Protease Activating Factor-1* (Apaf-1) presente a livello citosolico, il complesso formatosi prende il nome di apoptosoma ed è in grado di legare la procaspasi-9 attivandola e facendo iniziare la cascata caspatica in maniera simile a come accade nella via estrinseca.<sup>7</sup>

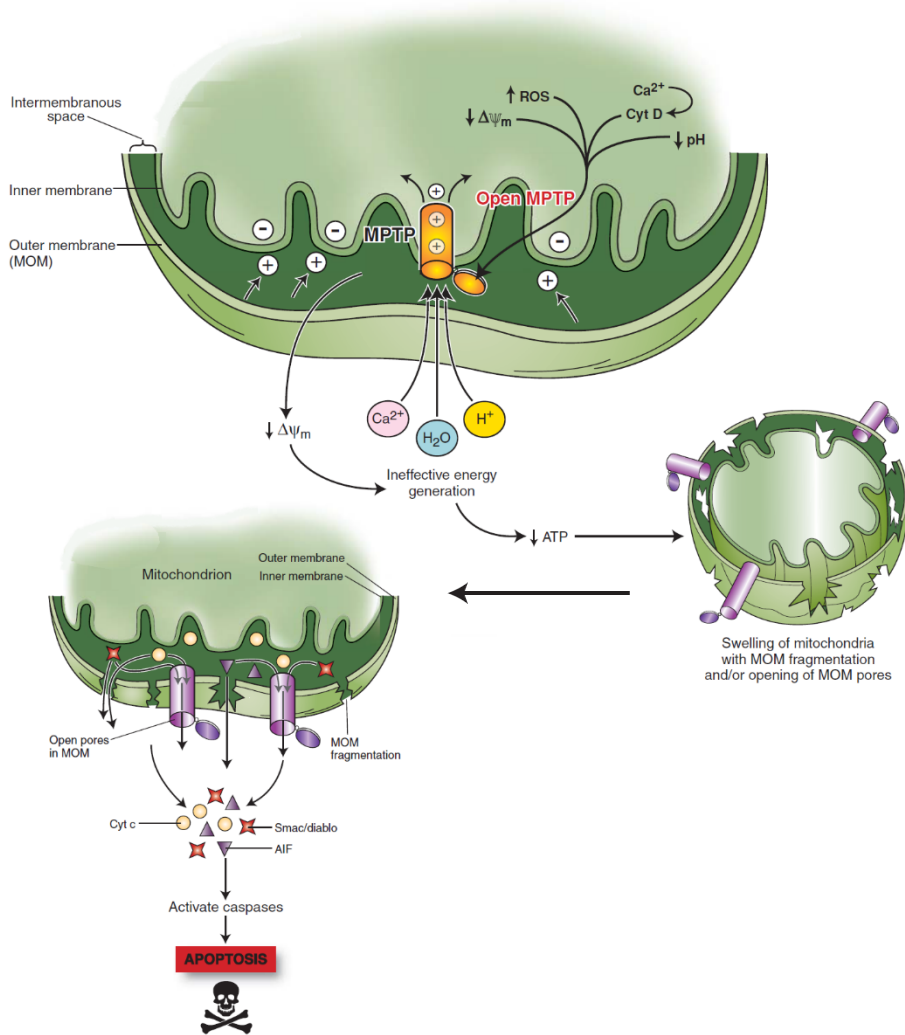


Figura 2.6 – Via intrinseca di attivazione dell'apoptosi <sup>7</sup>

## 2.3 Mediatori pro e anti apoptotici

Precedentemente si è visto come giochino un ruolo fondamentale nel processo apoptotico una serie di proteine che promuovono o inibiscono la morte cellulare, tali mediatori fanno parte della famiglia *B-cell lymphoma 2* (Bcl-2), il cui ruolo principale è il controllo della formazione dei pori sulla membrana mitocondriale esterna (MOMP), con il conseguente rilascio del Citocromo c nel citosol.

Esistono diverse proteine facenti parte della famiglia Bcl-2, ognuna con specifiche peculiarità, ma tutte condividono almeno uno di quattro domini globulari formati da  $\alpha$ -eliche: BH<sub>1</sub>, BH<sub>2</sub>, BH<sub>3</sub>, BH<sub>4</sub> (*Bcl-2 homology domain*).<sup>13</sup>

L'omologia dei domini BH permette di distinguere questa famiglia di proteine in tre fondamentali sottofamiglie (FIGURA 2.7 A).

Il primo sottogruppo è rappresentato dalle proteine multi dominio Bcl-2, Bcl-xL, Bcl-w, Mcl-1, A1, CED-9, Bfl-1 e Bcl-B, hanno funzione antiapoptotica e presentano tutti e quattro i domini BH. La presenza dei quattro domini permette l'eterodimerizzazione con Bax/Bak, al fine di bloccare la loro attività proapoptotica.

Il secondo e il terzo sottogruppo racchiudono tutte proteine proapoptotiche. Nel secondo sottogruppo troviamo Bax, Bak, Diva, Bcl-xs e Bok. Queste proteine possiedono i domini BH<sub>1</sub>, BH<sub>2</sub> e BH<sub>3</sub>.

L'ultima sottofamiglia mediatori è chiamata anche *BH<sub>3</sub>-only* in quanto possiede unicamente il dominio BH<sub>3</sub>, necessario all'interazione tra proteine pro e antiapoptotiche; di questa famiglia fanno parte Bid, Bim, Bad, Puma, Noxa, BMF, HRK e Bik. Come già visto in precedenza, rappresentano l'ultimo passaggio della via estrinseca che permette la formazione dei MOMP e la fuoriuscita del citocromo c. Le *BH<sub>3</sub>-only* che possono legarsi sia alle proteine anti apoptotiche, venendo inibite, sia a quelle proapoptotiche facilitandone l'oligodimerizzazione.<sup>13</sup>

I domini BH delle prime due sottofamiglie hanno una struttura simile, formano un fascio elicoidale che circonda un'elica centrale ( $\alpha_5$ ), questa struttura genera una cavità interna idrofobica, fondamentale per l'interazione tra le stesse proteine attraverso il dominio BH<sub>3</sub> (Figura 2.7 b).

Mentre le proteine BH<sub>3</sub>-only presentano una struttura disordinata, non globulare, fatta eccezione per il dominio BH<sub>3</sub> che invece presenta una struttura ad  $\alpha$ -elica anfipatica.

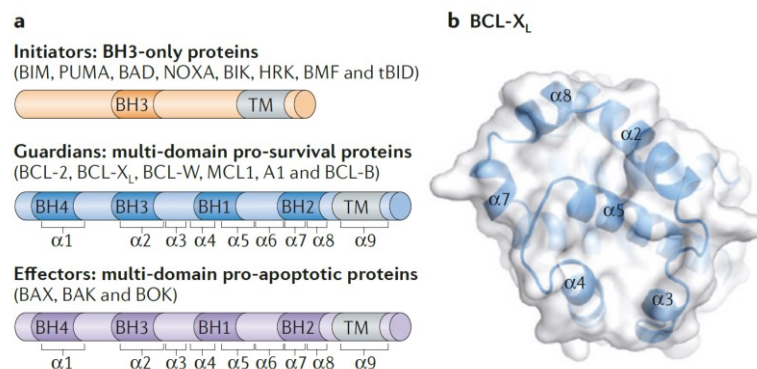


Figura 2.7 a – Rappresentazione dei domini BH nelle tre sottofamiglie di proteine Bcl-2. È stata anche riportata la nomenclatura delle  $\alpha$ -eliche corrispondenti ai vari domini. Si parte dal dominio BH<sub>4</sub> con la catena  $\alpha_1$  sino ad arrivare al dominio BH<sub>2</sub>, con l'elica  $\alpha_8$ . I domini possono essere formati da una o più  $\alpha$ -eliche. Figura 2.7 b – BCL-X<sub>L</sub> esempio di disposizione tridimensionale dei domini BH, la cavità idrofobica interna permette l'interazione con l' $\alpha$ -elica del dominio BH<sub>3</sub> delle BH<sub>3</sub>-only proteins.<sup>13</sup>

### 2.3.1 Ruolo nella genesi tumorale

La formazione di un tumore richiede una combinazione di più errori genetici che rendono possibile alla cellula neoplastica di diventare indipendente nella replicazione e di poter evitare i normali segnali di morte necessari al controllo proliferativo fisiologico.

Come visto in precedenza la morte cellulare programmata è finemente regolata dall'interazione tra i membri pro e antiapoptotici della famiglia di proteine Bcl-2.

Diversi studi eseguiti su topi geneticamente modificati hanno dimostrato che la crescita tumorale può avvenire sia con l'inibizione delle proteine proapoptotiche, sia con la sovraespressione di proteine antiapoptotiche. Queste mutazioni si è inoltre visto che possono trovare il loro scopo sia nella genesi tumorale, sia nella propagazione come mezzo di sviluppo per le cellule malate.<sup>14</sup>

Si è osservato inoltre che tali mutazioni, oltre ad inibire l'apoptosi, rendono anche le cellule neoplastiche più resistenti alle terapie antitumorali, come ad esempio il trattamento con cisplatino.

Dei tre gruppi di proteine della famiglia delle Bcl-2, due di queste sono proapoptotiche, mentre l'altro gruppo è costituito da mediatori antiapoptotici. Del gruppo delle proteine antiapoptotiche ne sono state identificate e caratterizzate sei: bcl-2, Bcl-X<sub>L</sub>, BCL2-related, Mcl-1, Bcl-w e A1/Bfl-1; si tratta di proteine che condividono quattro dei domini BH e hanno tutte una struttura tridimensionale molto simile.

Diversi studi hanno evidenziato che una sovra-espressione di Bcl-2 è caratteristica di specifici tipi di neoplasia: del linfoma follicolare che colpisce i linfociti B (*Human follicular centre B-Cell Lymphoma*), della leucemia linfatica cronica (Chronic Lymphocytic Leukaemia CLL), del linfoma diffuso a grandi cellule B (DLBCL) e del linfoma mantellare (*Mantle Cell Lymphoma*).<sup>15</sup>

Un alto livello di Bcl-2 è tipico dei linfomi sopracitati, che fanno tutti parte della classificazione dei tumori ematologici<sup>16</sup>, tuttavia si riscontra anche una certa sovraespressione in alcuni tumori solidi<sup>17</sup>, in particolare al cervello, al seno, alla prostata e al polmone.<sup>18</sup>



## 2.4 Bfl-1

Allo scopo di questa tesi ci si concentrerà sulla proteina antiapoptotica *Bcl-2-related protein A1* o detta anche Bfl-1, facente parte della famiglia delle Bcl-2.

Come già visto in precedenza si tratta di una proteina composta da 4 domini BH, i quali formano una struttura in grado di legare il dominio BH<sub>3</sub> delle controparti proapoptotiche. Questo sito di legame si è dimostrato molto interessante come target nella ricerca di molecole con struttura simile a BH<sub>3</sub> in grado di inibire o limitare l'azione di tale proteina.<sup>19</sup>

La struttura tridimensionale di una proteina risulta necessaria per lo studio di interazioni con ligandi fisiologici e non; quindi per capire eventuali meccanismi d'azione o per valutare i cambiamenti che avvengono quando la proteina si trova in presenza di ligandi allosterici, covalenti, competitivi o non competitivi.

La struttura di una proteina può essere ottenuta attraverso due principali tecniche: cristallografia a raggi X (*X-ray crystallography*) o attraverso spettroscopia di risonanza magnetica (NMR). Entrambe le tecniche presentano dei limiti, seppur diversi, e ciò comporta che non sempre sia possibile ottenere la struttura precisa di una proteina pura e slegata dal suo ligando (**forma apo**), quindi talvolta risulta più semplice ricavare la struttura della proteina legata ad un ligando (**forma holo**).<sup>20</sup>

Nel caso di Bfl-1 la struttura è stata ricavata con la tecnica della cristallografia a raggi X ed è stata fatta utilizzando la forma apo, quindi della proteina pura non legata ad alcun ligando (*FIGURA 2.8A* PDB ID: 5WHI)

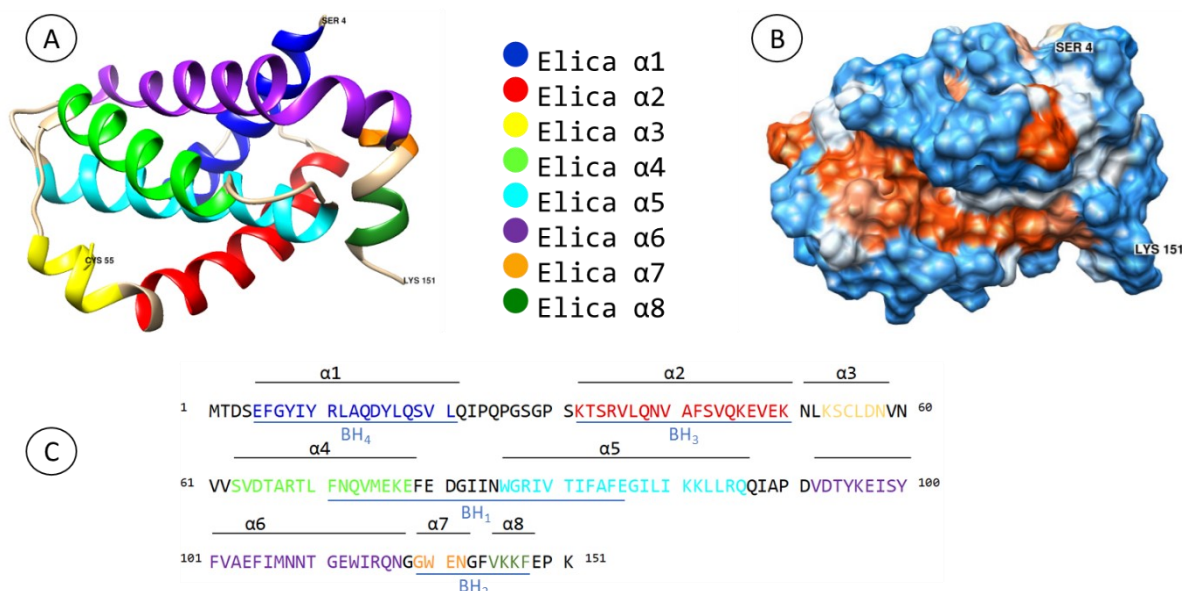


Figura 2.8A – Bfl-1 forma apo (PDB ID: 5WHI) vengono evidenziate le  $\alpha$ -eliche che formano la struttura tridimensionale della proteina, l’N-terminale si trova sulla Serina 4, mentre il C-terminale sulla Lisina 151; inoltre viene evidenziata la Cisteina 55, fondamentale per l’interazione. Figura 2.8B – Bfl-1 rappresentazione della nube elettronica superficiale in cui si evidenzia in rosso la tasca idrofobica ed in blu la parte idrosolubile. Figura 2.8C – sequenza amminoacidica di Bfl-1, in cui vengono evidenziate le porzioni amminoacidiche facenti parte dei domini BH e delle  $\alpha$ -eliche.<sup>21</sup>

Si può notare una struttura globulare composta da una serie di  $\alpha$ -eliche unite da dei loop, formati principalmente da glicine e proline, che permettono un’interruzione dell’elica.

La proteina conserva la struttura a multidominio della famiglia delle Bcl-2, quindi presenta un core formato dalle eliche  $\alpha 5$  e  $\alpha 6$  che, grazie al loop disposto a formare una curva di  $180^\circ$ , si posizionano in modo quasi parallelo formando la tasca idrofobica, tipica di questa famiglia di proteine. Il core è poi circondato dalle altre  $\alpha$ -eliche. L’elica  $\alpha 9$  viene solitamente troncata per facilitare l’espressione e la cristallizzazione e per lo stesso motivo vengono troncati i primi tre amminoacidi, la struttura così inizia dalla serina 4 e termina con lisina 151 (FIGURA 2.8C). Nonostante le modifiche apportate il core rimane intatto e funzionale, permettendo così di studiarne le interazioni.

Un’importante caratteristica che differenzia Bfl-1 dagli altri componenti della stessa famiglia è il fatto che presenta una cisteina, in posizione 55, orientata verso la tasca idrofobica e altamente reattiva, questo permette di formare un legame covalente con il dominio BH3 di altri agenti proapoptotici o di essere un buon target per lo sviluppo di farmaci selettivi.<sup>19</sup>

Ai fini di questa tesi risulta importante chiarire che la tasca idrofobica è a sua volta divisa in più *pockets* o *binding site*, ovvero piccoli gruppi di amminoacidi che interagiscono attraverso

legami deboli con un solo amminoacido del dominio BH3; quindi per definire tali *pockets* si valuta l'interazione con il dominio BH3 che, nel caso della FIGURA 2.9, è stato utilizzato quello di BAK.<sup>22</sup>

Come si può evincere dall'immagine, l'interazione con BAK delinea 4 *pockets*, dalla P1 alla P4; la numerazione parte dall'interazione con l'amminoacido più vicino all'N-terminale di BAK, che in questo caso è valina 74; quindi gli amminoacidi che interagiscono con V74 andranno a formare il P1, quelli che interagiscono con leucina 78 formeranno il P2, isoleucina 81 definirà il P3 ed in fine isoleucina 85 andrà a delineare il P4.

Nella FIGURA 2.9 si nota che è presente anche un P0, il quale però non ha interazioni con BAK, questo perché sperimentalmente si è evidenziato che nonostante non ci sia un legame con gli amminoacidi del dominio BH3 alcune molecole possono interagire con gli aa presenti in questa zona della proteina e provocare un cambiamento conformazionale; per questo viene comunque considerato nell'elenco delle *pockets* e viene studiato per eventuali interazioni.

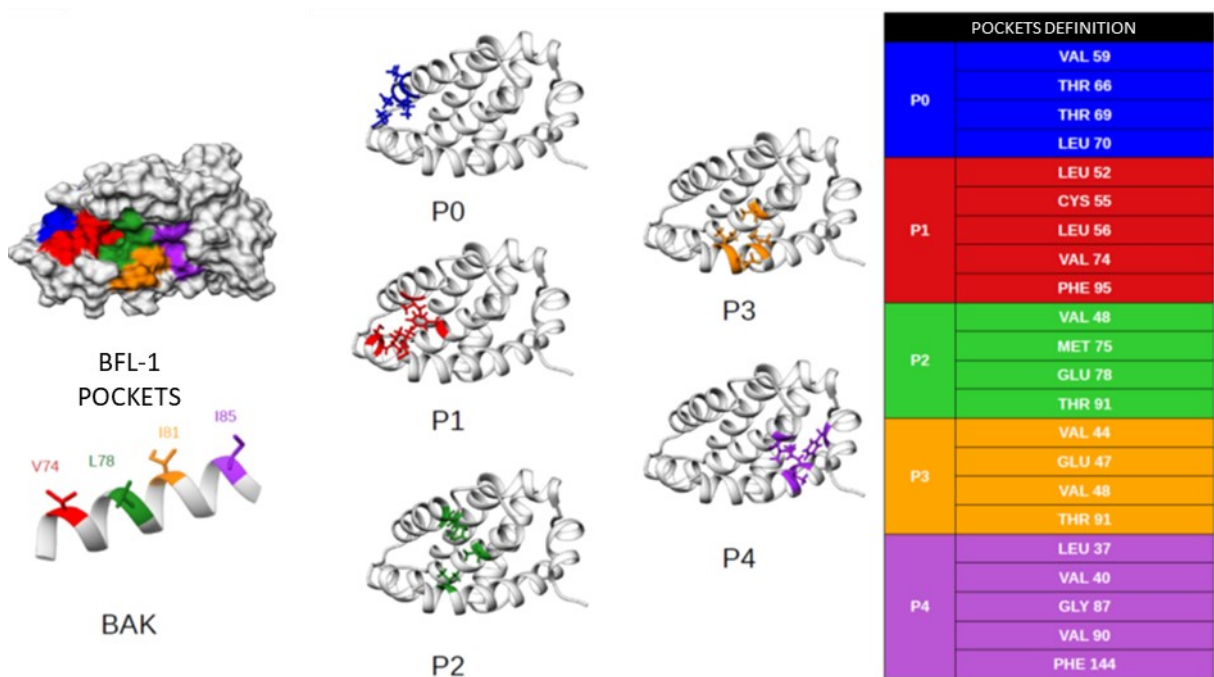


Figura 2.9 – Elenco delle Pockets nella tasca idrofobica di Bfl-1. La numerazione dei pockets viene fatta in relazione al dominio BH3 di una proteina proapoptica, in questo caso BAK. P0 non ha un'interazione diretta, ma è interessante dal punto di vista della ricerca.<sup>21</sup>

### 2.4.1 Gene BCL2A1

Bfl-1 viene codificata dal gene BCL2A1, gene la cui espressione è regolata dal fattore nucleare  $\kappa$ B (NF- $\kappa$ B) e svolge importanti funzioni pro-sopravvivenza della cellula.

In un contesto fisiologico, il gene BCL2A1 si è notato essere espresso principalmente nel sistema ematopoietico, al fine di facilitare la sopravvivenza dei leucociti; tuttavia viene espresso anche in altri tessuti come polmoni, intestino, testicoli e cellule muscolari lisce, solitamente per permettere la sopravvivenza delle cellule dopo un'infezione. Nel caso del tumore invece, l'espressione del gene ha la capacità di facilitare la proliferazione del tumore stesso o di fornire resistenza alla terapia.<sup>23</sup>

BCL2A1 è un gene localizzato nel cromosoma 15q24.2 e contiene tre esoni. L'mRNA solitamente trascrive il gene dall'esone 1 e 3, formando una proteina contenente 175 amminoacidi e nove  $\alpha$ -eliche; ai fini della produzione sperimentale tuttavia viene ridotta a 148 amminoacidi per migliorarne la stabilità e comunque si è visto non avere ricadute sulla capacità di legame della proteina.

Il gene è stato identificato per la prima volta in un fegato fetale e per questo ha assunto il nome di Bfl-1 (*Bcl-2 Related Gene Expressed in Fetal Liver*). Si è poi scoperto che la trascrizione può essere indotta da vari fattori: il fattore di necrosi tumorale  $\alpha$  (TNF $\alpha$ ), *Nuclear Factor Kappa-Light-Chain-Enhancer of Activated B Cells* (NF- $\kappa$ B), acidi retinoici, antagonisti degli acidi retinoici ed in fine una proteina chiamata zinc-finger che viene sovraespressa in particolari forme di leucemia.<sup>23</sup>

I primi due fattori evidenziano la relazione che intercorre tra Bfl-1 e l'infezione, infatti vengono prodotti dalle cellule proinfiammatorie per consentirne la sopravvivenza. Gli altri fattori di trascrizione vengono invece espressi in situazioni di iperossia o di bassi livelli di sostanze reattive dell'ossigeno (ROS).

Patologicamente si è poi scoperto che BCL2A1 subisce un riarrangiamento cromosomico nella leucemia mieloide cronica, che lo collega al fattore di crescita 4 dei fibroblasti.

Al fine di comprendere più approfonditamente il ruolo di Bfl-1 nella patologia tumorale si sono studiati dei topi transgenici in cui veniva sovraespresso il gene BCL2A1, i dati che sono emersi sono stati un aumento della sopravvivenza dei linfociti T, aumentandone

notevolmente la conta nel timo e nella milza. Un altro dato molto interessante che è emerso da questo studio è stato che una sovraespressione di Bfl-1 da sola non risulta sufficiente ad indurre la genesi tumorale nelle cavie studiate.

Da questi studi emerge quindi che nonostante il ruolo marginale nella genesi tumorale, un aumento incontrollato di Bfl-1 può prevenire la morte cellulare permettendo l'avanzamento del tumore e permette, al tumore stesso, di diventare resistente alla chemioterapia, in particolare a etoposide, staurosporina e cisplatino. Mentre un silenziamento del gene BCL2A1, sempre in cavie animali, ha dimostrato una sensibilizzazione delle cellule B tumorali all'apoptosi indotta da agenti chemioterapici, nello specifico è stato utilizzato l'anticorpo monoclonale rituximab.<sup>23</sup>

#### **2.4.2 Inibitori delle proteine antiapoptotiche Bcl-2**

Il trattamento dei tumori ematici negli ultimi dieci anni è molto progredito, grazie alla ricerca sui meccanismi biologici alla base di questa patologia, sono stati sviluppati una serie di farmaci che agiscono sulle proteine della famiglia Bcl-2; in particolare sulla sottofamiglia di oncogeni che impediscono l'apoptosi.

Molte delle molecole in grado di inibire la famiglia Bcl-2 attualmente sono sottoposte a studi clinici per il trattamento del tumore. Attualmente la molecola che ha dato i risultati più promettenti risulta essere *Venetoclax*, una molecola che agisce specificamente su Bcl-2 e che viene somministrata in associazione ad anticorpi monoclonali anti-CD38, migliorandone notevolmente l'efficacia terapeutica.

La sopravvivenza dei pazienti affetti da tumori ematici, negli ultimi anni, sta aumentando grazie allo sviluppo di nuovi farmaci come agenti immunomodulatori (IMiD), inibitori del proteasoma (PI) e anticorpi monoclonali (MoAb). Tuttavia, la capacità di adattamento delle cellule neoplastiche, ha portato alla perdita di efficacia di tali farmaci e ad una sempre più bassa remissione dalla malattia.<sup>24</sup>

I farmaci attualmente in commercio in grado di agire sui meccanismi della via intrinseca dell'apoptosi sono molecole di dimensioni ridotte che inibiscono le proteine Bcl-2, Bcl-xL, Bcl-w, ma non sono selettive per Bfl-1 e Mcl-1, inoltre non sono stati scoperti farmaci che fossero non selettivi per l'intera famiglia di proteine. (FIGURA 2.10)

Gli inibitori che per ora hanno sortito i maggiori effetti sono piccole molecole che legano la tasca idrofobica delle proteine, dove si trova il sito di interazione per il dominio BH3, di fatto spiazzando le proteine proapoptotiche così permettere il proseguimento dell'induzione della morte cellulare con l'esocitosi del citocromo c.

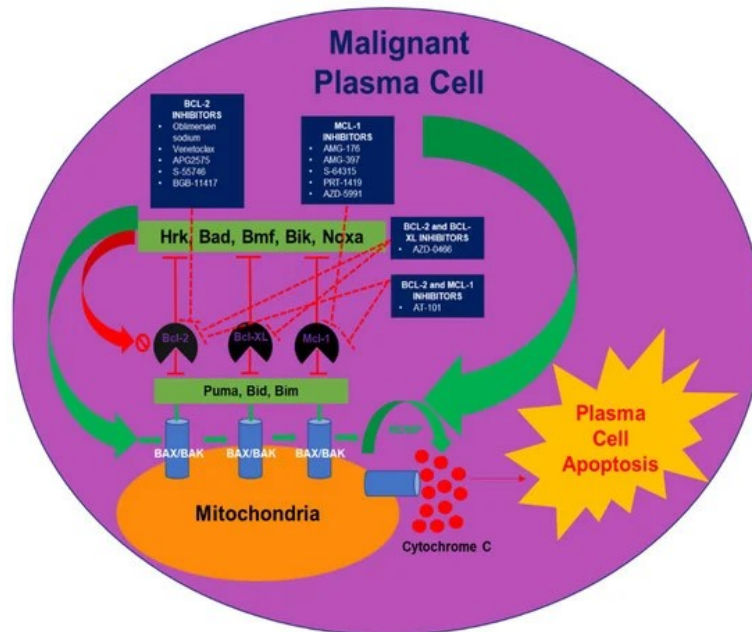


Figura 2.10 – Molecole scoperte in grado di inibire l'interazione tra le proteine antiapoptotiche della famiglia Bcl-2 e le controparti proapoptotiche<sup>24</sup>

**ABT-737:** è una piccola molecola, scoperta tramite uno screening di librerie di composti tramite NMR, che lega Bcl-2, Bcl-xL e Bcl-w. È in grado di legare in maniera specifica la tasca idrofobica antagonizzando efficacemente il dominio BH3 delle proteine proapoptotiche. Tuttavia, nonostante inizialmente abbia mostrato ottimi risultati negli studi preclinici, era richiesta la somministrazione parenterale, in quanto l'effetto di primo passaggio, se assunta oralmente, riduceva troppo la biodisponibilità.<sup>25</sup> (FIGURA 2.11)

**Navitoclax:** precedentemente veniva chiamato ABT-263, si tratta di un analogo strutturale di ABT-737. È stato sviluppato per essere più biodisponibile se assunto oralmente ed infatti ha riscosso buoni risultati negli studi clinici. Si è osservato, in modelli in vitro, che Navitoclax è in grado, di far regredire sia dei modelli di linfoma che di mieloma. Tuttavia le modifiche necessarie ad aumentarne la biodisponibilità, hanno portato ad avere come effetto collaterale una diffusa trombocitopenia. Questo effetto collaterale ha portato a dover rivedere il dosaggio nella sperimentazione; inoltre quando si è passati alla sperimentazione su modelli animali, si è osservato che nei modelli animali l'efficacia risultava ridotta, sia a

causa della bassa dose, sia a causa della bassa biodisponibilità; tuttavia l'associazione con Rituximab sembra essere promettente.<sup>26</sup> (FIGURA 2.11)

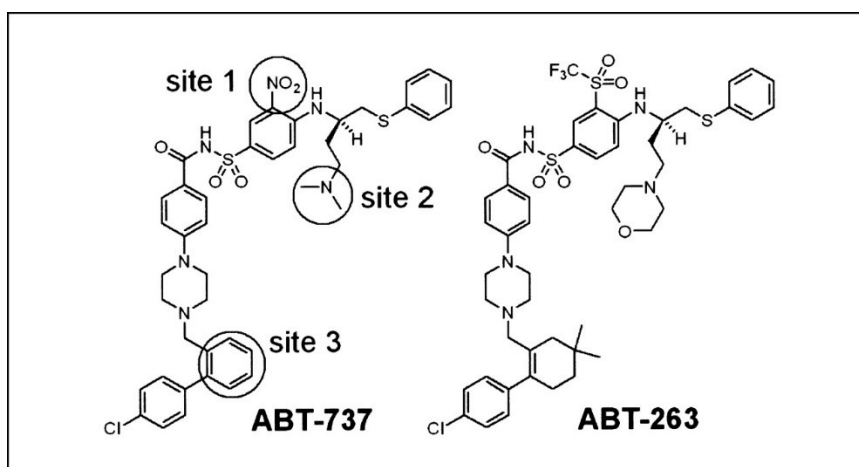


Figura 2.11 – Struttura chimica di ABT-737 (sinistra) e di ABT-263 (destra), nella figura vengono evidenziate le modifiche strutturali necessarie all'aumento della biodisponibilità.<sup>26</sup>

**Venetoclax:** la ricerca su ABT-737 e ABT-263 è proseguita e qualche anno dopo è stato prodotto ABT-199, oggi conosciuto con il nome di Venetoclax. Si tratta di un farmaco autorizzato nel 2016 dall'EMA<sup>27</sup> e dall'FDA<sup>28</sup>, con il nome commerciale di Venclyxto. Si tratta attualmente il primo ed unico farmaco *BH3-mimetics* in commercio.

La molecola instaura un legame specifico con Bcl-2, perdendo quindi di affinità per Bcl-xL e Bcl-w; il suo impiego è stato autorizzato per la terapia contro la leucemia linfatica cronica (CLL) e la leucemia mieloide acuta (AML). Può essere utilizzato in combinazione con Obinutuzumab o con Rituximab solo dopo un precedente trattamento che ha avuto esito negativo. Può essere inoltre utilizzato da solo in pazienti con particolari alterazioni geniche come ad esempio la delezione 17p o la mutazione di TP53, questi pazienti non possono essere trattati con inibitori di della via recettoriale delle cellule-B (es. Ibrutinib and Idelalisib).<sup>29</sup> (FIGURA 2.12)

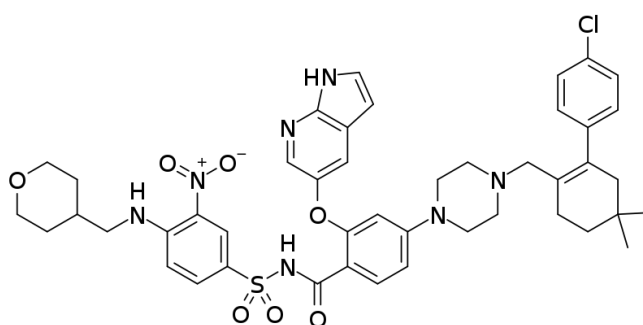


Figura 2.12 – Venetoclax (DrugBank DB11581)

Gli esempi appena illustrati sono tutti di molecole che derivano da screening di librerie di frammenti; queste molecole sono state più o meno migliorate per aumentarne biodisponibilità o selettività verso la tasca idrofobica delle proteine antiapoptotiche, portando alla sintesi di altre molecole come Obatoclox, Navitoclox e ABT-199, attualmente sottoposte a studi clinici. Tuttavia un nuovo approccio, attraverso metodi computazionali, ha portato alla progettazione di inibitori biologici come ad esempio: oligonucleotidi antisenso, piccole sequenze di DNA in grado di legare ed inibire l'mRNA codificante per le Bcl-2 antiapoptotiche; oppure piccoli peptidi che mimano la struttura del dominio BH3, ma che si legano con maggior affinità, così da sostituirsi alle proteine endogene.

Questi ultimi vengono chiamati *staplet peptides* (peptidi pinzati), in quanto sono dei peptidi che si dispongono ad  $\alpha$ -elica, ma si tratta di una conformazione tanto fondamentale quanto difficile da ottenere per un piccolo peptide di questo tipo, perciò si è pensato di aggiungere un ponte idrocarburico tra amminoacidi di catene laterali. I peptidi che si formano vengono anche chiamati SAHB (*Stabilized  $\alpha$ -Helices of BCL-2 Domains*) perché è emerso che questo ponte idrocarburico migliora sia la stabilità, sia la resistenza alle proteasi, sia l'affinità per la tasca della sottofamiglia multidominio delle Bcl-2.<sup>30</sup>

**Oblimersen Sodium:** si tratta di un oligonucleotide antisenso progettato specificamente per legare i sei codoni dell'mRNA umano codificante per Bcl-2. L'mRNA se non espresso verrà degradato e quindi come risultato non si avrà la formazione di Bcl-2.

*Oblimersen* (FIGURA 2.13) è stato il primo oligonucleotide antisenso ad aver dimostrato una certa efficacia nella diminuzione dell'espressione di Bcl-2 nelle cellule tumorali.<sup>31</sup> I dati raccolti negli studi clinici hanno evidenziato un miglioramento della risposta alla terapia antitumorale se combinato con Fludarabina e Ciclofosfamide, tanto che è stata fatta richiesta sia all'FDA<sup>32</sup> che all'EMA<sup>33</sup> per l'autorizzazione all'immissione in commercio, con il nome di Genasense. Attualmente le due agenzie del farmaco non hanno ancora concesso l'autorizzazione in quanto, dopo una valutazione della sopravvivenza in un trattamento quinquennale, purtroppo si è visto che la sopravvivenza non era sufficientemente migliorata rispetto all'utilizzo degli antitumorali in monoterapia. Questo insuccesso tuttavia ha aperto la strada a numerose ricerche su analoghi del Genasense e sulla terapia con oligonucleotidi antisenso.



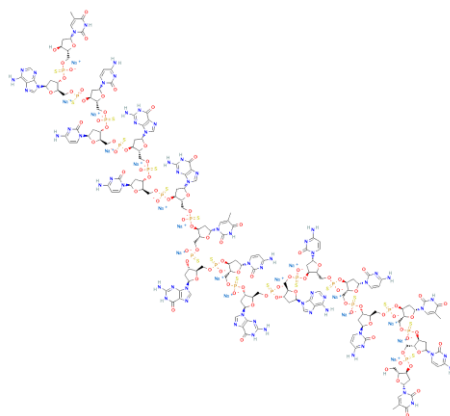


Figura 2.13 - Oblimersen Sodium (PubChem CID 118984457)

La ricerca sui peptidi pinzati, rispetto agli altri inibitori, è relativamente recente (circa una decina di anni), ciò comporta che non siano ancora presenti molecole che siano passate agli studi clinici, tuttavia le ricerche sembrano dare speranzosi risultati per il futuro. La specificità di queste piccole  $\alpha$ -eliche che mimano il dominio BH3 è data dalla composizione e dalla disposizione degli amminoacidi che interagiranno poi con la tasca delle proteine antiapoptotiche. Si è quindi partiti dai ligandi endogeni come BIM, che è in grado di legare la tasca di tutte le proteine antiapoptotiche, oppure BAD, che lega BCL-2, BCL-xL, e BCL-w, o ancora NOXA, che può legare MCL-1 e BFL-1.<sup>34</sup>

Dopo aver ottenuto alcuni risultati con i ligandi endogeni, la ricerca ha portato a modificare alcuni amminoacidi al fine di migliorarne la selettività e/o l'affinità, si sono venuti a definire così i SAHB. Un esempio di SAHB è riportato in FIGURA 2.14, si è partiti dal ligando endogeno BIM, in cui è stato aggiunto il ponte idrocarburico necessario alla stabilità (in verde acqua), e gli amminoacidi Q221 (glutammina) ed E223 (acido glutammico) sono stati sostituiti con fenilalanina; il nuovo peptide formatosi è stato chiamato MCL-1 SAHB<sub>D</sub>. Sperimentalmente è stato riscontrato che le modifiche alla struttura hanno permesso una migliore selettività per la proteina MCL-1 e un'affinità maggiore data dai legami degli amminoacidi D218 (acido aspartico) e R214 (arginina) con la tasca idrofobica della proteina.<sup>34</sup>

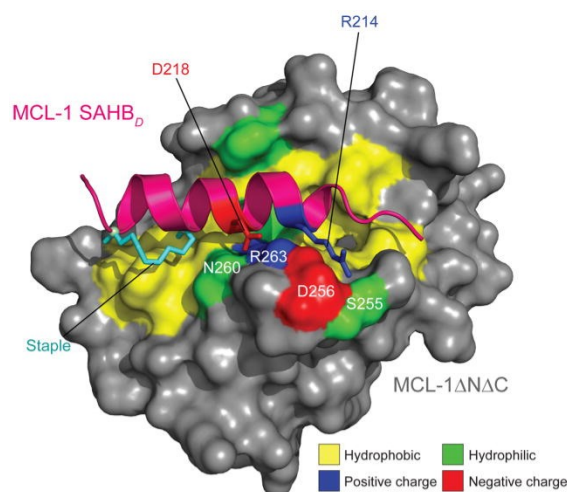


Figura 2.14 – Struttura cristallografica di MCL-1 legato a MCL-1 SAHB<sub>D</sub>, in azzurro è evidenziato il ponte idrocarburico che dona stabilità alla struttura, vengono evidenziati anche gli amminoacidi che permettono l'interazione con la proteina e i punti di interazione sulla superficie della proteina stessa <sup>34</sup>

## 2.5 Fragment-based drug discovery tramite NMR

Negli ultimi vent'anni la spettroscopia di risonanza magnetica nucleare (NMR) si è dimostrata una tecnica analitica molto valida ai fini della scoperta di nuovi farmaci (*drug discovery*). In particolare, il *Fragment-Based Drug Discovery (FBDD)* ha largamente beneficiato dallo sviluppo delle tecniche NMR, trovando ottimi riscontri: nella creazione di librerie di frammenti; nelle tecniche che osservano le interazioni del ligando (*Ligand-Observed*) e in quelle che osservano invece il target (*Target-Observed*); negli screening di possibili ligandi (*Hit Screening*) e nella validazione di ligandi (*Validation*). <sup>35</sup>

La FBDD è un approccio alla ricerca di nuovi farmaci che permette di lavorare su numerosi target come DNA, RNA, chinasi, enzimi, proteine di membrana e proteine intrinsecamente disordinate (*Intrinsically Disordered Protein IDP*). I "frammenti" a cui ci si riferisce sono delle piccole molecole il cui peso molecolare solitamente non supera i 300 Da, inoltre, date le ridotte dimensioni, l'affinità che presentano per il target è dell'ordine del micro ( $\mu\text{M}$ ) o del milli molare (mM).<sup>36</sup> In realtà i criteri di inclusione ai fini della creazione di una libreria di frammenti si basa sulla del 3, derivante dalla regola di Lipinski, ovvero sia:<sup>37</sup>

- La massa molecolare deve essere  $\leq 300$  Da, questo per facilitare la diffusione e la possibilità poi di maturare il composto;
- $\text{ClogP} \leq 3$ , il coefficiente di ripartizione ottanolo/acqua esprime la misura della liposolubilità della molecola;

- Donatori o accettori di legami a idrogeno  $\leq 3$ , questo perché troppi legami ad idrogeno polarizzano la molecola, rendendo più difficile la diffusione nelle parti apolari del target;
- Il numero di legami rotazionali  $\leq 3$ , la regola è necessaria ai fini sperimentali per poter studiare la sua interazione e la specificità di legame.

Solitamente le librerie (*Library*) di frammenti sono composte da centinaia, se non migliaia di composti. Si comprano da ditte specializzate che dopo aver selezionato i frammenti, seguendo le regole di Lipinski, li suddividono ulteriormente in base alle caratteristiche chimiche, inoltre le ditte escludono dalla selezione tutti quei composti che per loro natura si sanno essere chelanti, ligandi aspecifici, ligandi covalenti, aggreganti, etc.... quindi tutti quei composti che non porterebbero a nessun dato rilevante.

L'utilizzo di piccole molecole ha lo svantaggio che l'affinità per il target è molto ridotta rispetto all'utilizzo di molecole più grandi e complesse e che è necessario usare tecniche strumentali in grado di analizzare queste interazioni con sufficiente risoluzione. Tuttavia gli svantaggi sono notevolmente compensati dalla velocità e dalla semplicità di sintetizzare degli analoghi per studi di relazione struttura-attività; oppure dalla possibilità di sondare un ampio spazio chimico date le piccole dimensioni; o ancora dalla possibilità di trovare più facilmente un valido ligando che, sebbene abbia una ridotta affinità rispetto ad uno di dimensioni maggiori, solitamente presenta delle interazioni di miglior qualità con la proteina bersaglio.<sup>38</sup>

Le tecniche utilizzate ai fini della ricerca fragment based sono:

- *Thermal Shift Assay (TSA)*: saggio che misura la variazione di temperatura a seguito di denaturazione termica, variazione che sarà influenzata dalla stabilità della proteina in presenza o meno di un ligando. Ha un costo limitato e permette di avere una grande portata, ma ha anche un tasso, relativamente elevato, di falsi positivi.<sup>39</sup>
- *Surface Plasmon Resonance (SPR)*: tecnica che prevede l'utilizzo di una lamina metallica che interagisce con il ligando e attraverso un fascio di elettroni o di luce si eccitano i plasmoni di superficie della lamina, che adsorbirà e rifletterà in maniera differente la radiazione dipendentemente dal ligando. Come per il TSA è una tecnica che permette di analizzare un gran numero di campioni velocemente ed in tempo reale, ma anche in questo caso è frequente incappare in falsi positivi.<sup>40</sup>

- X-ray Crystallography: la cristallografia a raggi X è una tecnica molto utilizzata e molto efficace, permette di ottenere dei dati molto precisi e di studiare la struttura tridimensionale dei target e dei complessi ligando-target, tuttavia non sempre è possibile ottenere un cristallo o un cristallo di qualità sufficiente all'analisi; inoltre è una tecnica che richiede molto più tempo per essere effettuata e i dati per essere analizzati.<sup>41</sup>
- MicroScale Thermophoresis (MST): è una tecnologia in grado di misurare le interazioni tra biomolecole; viene misurato il cambiamento di fluorescenza del campione sfruttando un gradiente di temperatura, rilevando cambiamenti nel guscio, nella carica o nelle dimensioni di idratazione molecolare. Si tratta di una tecnica che ha un basso costo e una buona efficienza, tuttavia richiede di avere un campione fluorescente.<sup>42</sup>
- Nuclear magnetic resonance (NMR): tecnica basata sull'analisi del comportamento degli spin di nuclei di atomi magneticamente attivi in presenza di un forte campo magnetico. È una tecnica che può essere applicata a campioni in soluzione, e che permette di ottenere dei risultati molto accurati, riducendo la possibilità di falsi positivi, il limite della tecnica è la necessità di inserire nel campione almeno un isotopo che sia magneticamente attivo (<sup>13</sup>C, <sup>1</sup>H, <sup>15</sup>N, <sup>31</sup>P, <sup>23</sup>Na e <sup>19</sup>F).<sup>43</sup>

Rispetto alle altre tecniche, l'NMR presenta numerosi vantaggi, primo tra tutti la possibilità di effettuare l'analisi sia in solvente organico che in *buffer* acquoso, poi permette di stimare un valore di  $K_d$  frammento-proteina (costante di dissociazione rappresentante la concentrazione di composto necessario a saturare il 50% dei siti di legame) nell'ordine del millimolare. L'analisi tramite NMR rende possibile anche valutare l'interazione e la modalità di associazione in siti specifici.

Esistono due approcci alla FBDD tramite NMR: uno che si focalizza sul ligando (*Ligand-Observed NMR*) e l'altro che invece osserva i cambiamenti che avvengono nel target (*target-Observed NMR*).<sup>36</sup>

La prima metodologia, *Ligand-observed NMR*, sfrutta principalmente due tecniche NMR monodimensionali: *Transfer Difference spectroscopy* (STD) e *Water-Ligand Observed via Gradient Spectroscopy* (Water-LOGSY). Si tratta di esperimenti basati sull'effetto *Nuclear Overhauser Effect* (NOE) ovvero nella misurazione della variazione dell'intensità del segnale

di un protone che si trova spazialmente vicino ad un secondo protone che viene irradiato. Tale fenomeno è spiegato dalla formazione di un'interazione nucleare dipolo-dipolo in grado di indurre una variazione dello spin, conseguentemente si ha una variazione della magnetizzazione che induce un'alterazione dell'intensità del segnale NMR. Il vantaggio della tecnica è che non ci sono limitazioni nel peso molecolare del target e la possibilità di rilevare anche dei ligandi deboli, anche se è necessario sfruttare un isotopo come il  $^1\text{H}$  nella soluzione contenente il campione.

Quando si parla di *target-observed* NMR invece ci si focalizza sul segnale del target che viene misurato in termini di *chemical shift perturbation* (CSP)<sup>44</sup>, ovvero si misurano i cambiamenti nella nube elettronica della proteina indotti dal legame con il ligando.

Per poter utilizzare tale tecnica è necessario produrre un ligando che sia costituito da almeno due atomi magneticamente attivi, solitamente si utilizzano  $^1\text{H}$ ,  $^{15}\text{N}$  e  $^{13}\text{C}$ , questo ne costituisce uno dei limiti, oltre alla necessità di avere un target di dimensioni non troppo elevate, però risulta anche una tecnica estremamente precisa, che permette lo studio del sito di interazione e che riduce drasticamente la possibilità di incorrere in falsi positivi. Gli esperimenti che solitamente vengono effettuati quando si parla di *target-observed* sono esperimenti bidimensionali  $^1\text{H}$ - $^{15}\text{N}$  *Hetero-Nuclear Single Quantum Coherence Spectrum* ( $^1\text{H}$ - $^{15}\text{N}$  HSQC) *Selective Optimized-Flip-Angle Short-Transient - Heteronuclear Multiple Quantum Coherence* (SOFAST-HMQC), questi esperimenti vanno a valutare nello specifico il CSP di  $^1\text{H}$  e  $^{15}\text{N}$  nei gruppi ammidici facenti parte del *backbone* della proteina.

Le tecniche sperimentali finora trattate hanno costituito la base della FBB, tuttavia risultano essere tecniche molto costose e che presentano dei peculiari limiti, dovuti ai principi fisici alla base delle tecnologie utilizzate, così negli anni, grazie all'avanzamento della tecnologia computazionale si è pensato di iniziare ad affiancare le tecniche sperimentali ad un approccio computazionale. Grazie alle strutture tridimensionali di ligandi e proteine che sono state fatte tramite mezzi sperimentali, si è potuto creare un database che ha permesso di sviluppare un approccio in silico alla ricerca sperimentale.

Uno dei grandi vantaggi è stato quello di poter prevedere con un buon grado di sicurezza il *binding site* del target e di conseguenza quali ligandi possano essere più affini di altri, facendo risparmiare così molto tempo nello *screening*. Oppure la possibilità di incrociare i dati sperimentali con quelli in silico ai fini di comprendere la disposizione spaziale del ligando

all'interno del *binding site*, ciò con la conseguenza di poter, sempre in silico, maturare il ligando in modo da aumentarne le dimensioni e migliorarne l'interazione con la proteina, velocizzando così sia la sintesi che la validazione con metodi sperimentali.

Il lavoro computazionale inizia, come detto, sfruttando le strutture tridimensionali di target e ligandi ottenute tramite cristallografia a raggi X o NMR, da qui si sono creati dei protocolli in silico in grado di campionare, studiare e mappare i *binding site* e di vederne il possibile comportamento in presenza di ligandi. Inizialmente i metodi computazionali non tenevano conto della flessibilità del target, ovvero del fatto che in soluzione una proteina non è immobile, ma si verificano dei piccoli cambi conformazionali che determinano l'esposizione o meno di *binding site* più interni. Un grande cambio nell'approccio è avvenuto con lo sviluppo di *molecular dynamics simulation* (MD), ovvero algoritmi che tengono conto dei cambi conformazionali che avvengono sia fisiologicamente che in presenza di un possibile ligando. Una delle difficoltà di eseguire degli studi in silico su frammenti è che questi ultimi, data la loro piccola dimensione, si possono posizionare in diversi modi all'interno delle cavità del target, inoltre, essendoci più conformazioni di uno stesso target, il ligando ha più possibilità di inserirsi nel *binding site*. Questo comporta che nelle simulazioni effettuate si possono visualizzare numerose conformazioni target-ligando, che vengono chiamate *poses*; sarà poi compito del ricercatore selezionare le più probabili, che poi comunque saranno sottoposte a validazione sperimentale.<sup>35</sup>

Dopo aver verificato sperimentalmente quali frammenti effettivamente presentino una migliore affinità, il lavoro di *drug discovery* procede con la fase *hit to lead*, ovvero maturando i ligandi più promettenti; ciò può avvenire affidandosi nuovamente ai metodi in silico. La maturazione avviene attraverso tre diverse modalità: *fragment-linking*, *fragment-growing* e *fragment-merging* (FIGURA 2.15).

Nella tecnica di *fragment growing*, si procede migliorando l'affinità di legame del frammento per il target, aggiungendo nuovi gruppi funzionali in grado di interagire con la proteina.

Il metodo di *fragment linking* invece prevede di collegare più frammenti con elevata affinità per punti diversi del *binding site*, attraverso un 'ponte molecolare' (*linker*). Questo permette di mantenere il giusto orientamento dei frammenti nella tasca proteica e conseguentemente di migliorarne l'affinità.

L'ultima strategia di maturazione è il *fragment merging*, ovvero l'unione di due frammenti che legano nello stesso *binding site*, cercando di ottenere un composto con potenza maggiore.<sup>38</sup>

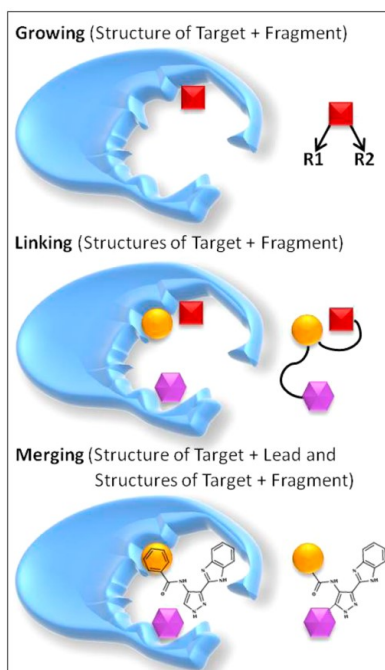


Figura 2.15 – Strategia hit to lead, esemplificazione di *fragment-linking*, *fragment-growing* e *fragment-merging*<sup>38</sup>

## 2.6 MiniFrag

La trattazione effettuata finora sulla *fragment-based drug discovery* è stata necessaria per comprendere le basi di questa tesi, si è visto come vengono definiti i frammenti, come tramite la spettroscopia NMR sia possibile studiare le interazioni deboli target-fragment, gli approcci computazionali e le difficoltà che si possono incontrare nella metodologia.

In questo lavoro si è voluto portare al limite la ricerca tramite NMR di possibili ligandi, sfruttando delle molecole più piccole dei frammenti solitamente utilizzati nelle librerie. Il grande vantaggio nella sperimentazione attraverso tali molecole è quello di poter esplorare un maggior spazio chimico rispetto ai frammenti solitamente utilizzati.

L'azienda farmaceutica Astex, ad un convegno tenutosi a San Diego nel 2018, è stata la prima a proporre l'utilizzo di molecole estremamente piccole nella ricerca di nuovi farmaci.<sup>45</sup> Successivamente, nel 2019, è stato pubblicato un articolo<sup>46</sup> in cui venivano definiti i cosiddetti *ultra-low-molecular-weight fragment* o **MiniFrag**, e la tecnica che era stata utilizzata, ovvero la spettroscopia a raggi X. Successivamente l'azienda Enamine ha reso

disponibile una libreria composta da 80 di queste molecole, rendendo possibile l'avanzamento della ricerca e la possibilità di implementare altre tecniche strumentali.

I MiniFraggs vengono definiti come molecole aventi un numero di atomi pesanti (HAC) compreso tra 5 e 7, inoltre, a differenza dei frammenti che seguono la regola del 3, questi composti seguono la regola dell'1 ovvero:

- Massa molare (MW) < 100 Da
- Numero di donatori di legami a idrogeno  $\leq 1$
- Numero di accettori di legami a idrogeno  $\leq 1$
- ClogP  $\leq 1$

Viste le piccole dimensioni si è osservato che l'affinità per il *binding site* e la rilevabilità allo strumento alle normali concentrazioni con cui si effettuavano gli screening (circa 1-2mM), era troppo bassa, non si avevano interazioni ed era quasi impossibile apprezzarle con la cristallografia. Perciò si sono cominciate ad utilizzare delle concentrazioni molto più elevate, nell'ordine dell'1M, ciò ha permesso sia di poter visualizzare le interazioni con il *binding site* sia di poter vedere qualcosa di inaspettato, ovvero che date le piccole dimensioni dei frammenti questi andavano ad occupare anche *binding site* secondari diversi dal sito attivo primario, che solitamente non venivano indagate in quanto i frammenti non potevano accedervi o non potevano legarsi con affinità sufficiente da poter essere valutati.<sup>46</sup>



### 3 OBIETTIVI DELLA TESI

Il focus di questa tesi si è concentrato sulla proteina antiapoptotica Bfl-1, appartenente alla famiglia delle Bcl-2. Fisiologicamente la sua espressione ha lo scopo di mantenere l'equilibrio nei processi di morte cellulare. Infatti la proteina si lega al dominio BH3 dei membri proapoptotici BAX e BAK, della stessa famiglia, impedendo così la liberazione del citocromo c.

La sovraespressione di Bfl-1 è stata osservata principalmente in tumori ematici, come leucemie e linfomi, e anche in alcuni tumori solidi. Lo squilibrio nell'omeostasi che si viene a creare porta a proliferazione tumorale e farmaco resistenza.

L'obiettivo principale di questo progetto è di ricercare piccole molecole che possano legare in modo selettivo il sito di interazione della proteina con BAX e BAK, spiazzandone il legame e impedendone l'attività. Le molecole che dimostrano tale attività verranno poi maturate per portare alla sintesi di un composto *lead*, a peso molecolare e affinità maggiore. A tale scopo viene utilizzata la tecnica *Fragment Based Drug Discovery* (FBDD) tramite NMR, la quale ha già dimostrato la sua valenza nella ricerca di un inibitore.

Attualmente tramite FBDD sono state già scoperte alcune molecole in grado di competere per il legame alla tasca idrofobica, tra cui *Venetoclax*, l'unica ad essere stata commercializzata. Tuttavia, nessuna di queste molecole risulta essere in grado di legare Bfl-1.

Il secondo scopo di questa tesi riguarda lo sviluppo e la verifica di una strategia di *screening* tramite NMR di molecole più piccole rispetto a quelle utilizzate comunemente nella FBDD, dette MiniFrag.

Il vantaggio nell'utilizzo di queste piccole molecole sta nel poter analizzare nuovi possibili siti di interazione con la proteina al fine di migliorare e potenziare la ricerca di inibitori specifici, sia per Bfl-1, sia, in futuro, per altre proteine.



## 4 MATERIALE E METODI

Il fine di questa tesi è quello di eseguire uno screening di composti a basso peso molecolare, con un numero di atomi pesanti (HAC) compreso tra 5 e 7, sulla proteina anti-apoptotica Bfl-1 utilizzando come tecnica principale l'NMR; per poter applicare tale tecnica è necessario che la proteina venga marcata con isotopi magneticamente attivi quali  $^{15}\text{N}$  e  $^1\text{H}$ .

La produzione di proteina marcata avviene attraverso quattro fasi: la fase di *trasformazione* in cui si inserisce il plasmide necessario all'espressione di Bfl-1 all'interno di cellule di *Escherichia coli*, la fase di *espressione* dove si fanno crescere i batteri in un terreno adeguato a marcare la proteina, la fase di *estrazione* durante la quale le cellule vengono lisate e vengono separati i residui cellulari dalla proteina, fase di *purificazione* effettuata utilizzando l'FPLC in modo da separare completamente la Bfl-1 da tutte le proteine funzionali dei batteri usati per l'espressione.

### 4.1 Trasformazione attraverso il plasmide Champion™ pET sumo

Per la fase di produzione vengono utilizzate le cellule *One Shot™ BL21 (DE3) (ThermoFisher Scientific)*. Si tratta di cellule batteriche chimicamente competenti, che vengono solitamente utilizzate per la produzione di proteine ricombinanti.

La produzione di Bfl-1 inizia con la trasformazione, ovvero un processo mediante il quale dei batteri competenti acquisiscono del DNA esogeno dall'ambiente (in questo lavoro il DNA esogeno è un plasmide codificante per la proteina di interesse).<sup>47</sup>

In questo lavoro la trasformazione batterica è stata effettuata mediante il metodo dello shock termico utilizzando 100 ng di plasmide per 50  $\mu\text{L}$  di cellule. Nello specifico i batteri sono stati lasciati in ghiaccio per 20 minuti, dopodiché sono stati prelevati 50  $\mu\text{L}$  da utilizzare come controllo negativo; al campione da trasformare invece sono stati aggiunti 100 ng di plasmide per 50  $\mu\text{L}$  di cellule (il controllo negativo verrà poi trattato allo stesso modo dei batteri che si vogliono trasformare). Dopo l'aggiunta del plasmide, i batteri sono stati posti in ghiaccio per 15 minuti; al termine dei quindici minuti sono stati spostati in un termoblocco impostato a  $42^\circ\text{C}$  per 1 minuto; subito dopo sono stati rimessi in ghiaccio per 2 minuti. Al termine di questo processo di shock termico sono stati aggiunti 250  $\mu\text{L}$  di terreno di coltura, avendo cura di mescolare il campione delicatamente, e si sono posizionate le vials all'interno

di un incubatore alla temperatura di 37°C per 30 minuti sotto agitazione (200 rpm). Il processo di trasformazione termina prelevando 150 µL di sospensione batterica e distribuendola in una piastra di Petri contenente del terreno di coltura con Kanamicina; la piastra è stata poi incubata a 37°C overnight. Il risultato finale al termine dell'incubazione è stato di trovare due piastre, una contenente le colonie batteriche e l'altra (controllo negativo) in cui non era cresciuto nulla.

Il plasmide ricombinante utilizzato è il plasmide *Champion™ pET SUMO* (ThermoFisher Scientific), contenente, oltre alla sequenza codificante per BFL-1, anche una serie di altre sequenze necessarie sia alla produzione della proteina, come la resistenza alla kanamicina, sia alla sua purificazione, come la presenza di SUMO e dell'His-tag.

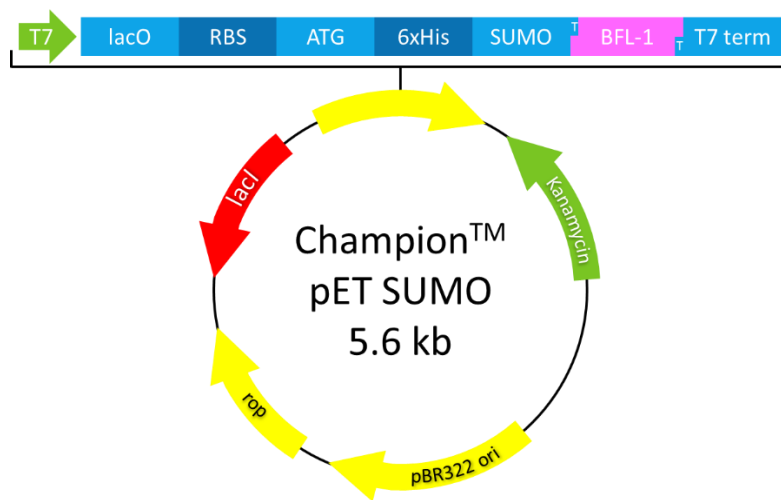


Figura 4.1 - Struttura del plasmide *Champion™ pET SUMO* <sup>48</sup>

Analizzando il vettore (FIGURA 4.1) si possono vedere diverse sequenze con specifiche funzioni:

- *pBR322 ori*, origine della replicazione, permette al plasmide di replicarsi autonomamente rispetto al DNA batterico.
- *rop* (*Repressor of Primer*), si tratta di una sequenza che stabilizza l'interazione RNA-RNA e mantiene basso il numero di copie del plasmide.
- *lacI*, gene codificante per il repressore lac, è una proteina che lega la sequenza lacO e ne impedisce la trascrizione, questo permette di controllare la trascrizione basale del promotore T7lac, ed è possibile inibire la sua azione utilizzando IPTG (isopropil-β-D-1-tiogalattopiranoside), un analogo dell'allolattosio, così da permettere la trascrizione da parte della polimerasi.

- *t7*, si tratta di una sequenza promotrice di trascrizione per la T7 RNA polimerasi, una polimerasi presente nel fago T7, un virus che infetta quasi tutti i ceppi di E. Coli, e che si è visto essere molto affine e specifica per questa polimerasi.

L'RNA polimerasi non viene normalmente espressa nelle cellule di E. Coli, perciò il kit di cellule che è stato acquistato presenta una modifica del genoma in cui è presente una sequenza codificante per la T7 RNA polimerasi.

- *lacO*, sequenza alla quale si lega il repressore lac, fondamentale per il controllo dell'espressione della proteina.
- *RBS*, si tratta del *binding site* del ribosoma.
- *ATG*, sequenza promotrice di trascrizione procariotica, si tratta di una sequenza di basi azotate poste prima della proteina che identificano l'inizio della trascrizione della stessa.
- *6xHis* o *HIS TAG*, è composto da sei istidine legate alla porzione N-terminale di SUMO, ed è un sistema comunemente utilizzato in biologia molecolare per la purificazione delle proteine ricombinanti, in quanto lega in maniera specifica e con grande affinità una specifica colonna per FPLC.
- *SUMO (Small Ubiquitin-Like Modifier)*, si tratta di una piccola proteina posta tra l'HIS TAG e Bfl-1 che favorisce la solubilità della proteina, diminuendo il rischio di degradazione e aumentandone la resa.

Un altro grande vantaggio nell'utilizzare SUMO è la possibilità, in fase di purificazione, di tagliarla all'estremità C-terminale sfruttando una SUMO proteasi altamente specifica così da rimanere con solo la proteina di interesse pura.

- *BFL-1*, la sequenza codificante per la proteina di interesse.
- *T7 term*, sequenza di fine per la polimerasi.
- *Kanamicina*, codifica per la resistenza all'antibiotico kanamicina. Questa resistenza è fondamentale ai fini della produzione, in quanto l'antibiotico viene aggiunto al terreno di coltura e permette di selezionare solamente i batteri contenenti il plasmide ed inoltre si evitano possibili contaminazioni.

La sequenza codificante per Bfl-1 che è stata utilizzata in questa tesi è stata leggermente modificata a fini produttivi. La sequenza presenta 476 paia di basi, codificanti per 176 amminoacidi, tuttavia al termine della traduzione la porzione C-terminale è priva di 25 residui, perciò Bfl-1 è composta da 151 amminoacidi (*FIGURA 4.2*); inoltre due cisteine in

posizione 4 e 19 sono state sostituite da due serine, questo al fine di aumentare la stabilità della proteina ed impedire eventuali oligodimerizzazioni, mentre la cisteina in posizione 55 non è stata modificata in quanto coinvolta nel binding site proteico.

10	20	30	40	50	60
MTDSEFGYIY	RLAQDYLSV	LQIPQPGSGP	SKTSRVLQNV	AFSVQKEVEK	NLKSCLDNVN
70	80	90	100	110	120
VVSVDTARTL	FNQVMEKEFE	DGIINWGRIV	TIFAFEGILI	KKLLRQQIAP	DVDTYKEISY
130	140	150			
FVAEFIMNNT	GEWIRQGGW	ENGFVKKFEP	K		

Figura 4.2 – Sequenza amminoacidica della proteina Bfl-1 utilizzata in questa tesi, in giallo sono evidenziate le serine in posizione 4 e 19 sostituite alle cisteine ed in verde è evidenziata la cisteina 55.

Le modifiche effettuate non comportano alcun cambiamento della struttura, in quanto tra i due amminoacidi varia solamente il gruppo sulfidrilico (-SH) delle cisteine, con il gruppo ossidrilico (-OH) delle serine, mentre per quanto riguarda l'assenza dei 25 residui dal C-terminale non risulta essere un problema perché non si tratta di una porzione coinvolta nel binding site della proteina e non presenta una funzione strutturale per la proteina.

## 4.2 Espressione di Bfl-1 in terreno arricchito con <sup>15</sup>N

I batteri vengono fatti crescere in due terreni di coltura differenti: LB e M9. Il terreno *Lysogen Broth* (LB) è un terreno di coltura ricco, non marcato, usato per far crescere i batteri in condizioni ottimali; contenente (TABELLA 4.1) Triptone come fonte di amminoacidi essenziali, Cloruro di sodio (NaCl) per garantire l'equilibrio osmotico ed estratto di lievito contenente vitamine ed altri elementi essenziali per la crescita batterica. Questo terreno viene utilizzato nella fase pre-inoculo, in cui si fanno crescere i batteri che sono stati trasformati o che sono stati scongelati, in questo modo è possibile lavorare in piccola scala e avere un quantitativo sufficiente di batteri per l'inoculo.

COMPOSIZIONE	QUANTITÀ
Triptone	10 g
NaCl	10 g
Estratto di lievito	5 g
Acqua milliQ	q.b. a 1 L

Tabella 4.1 – Composizione terreno *Lysogen Broth* (LB)

Prima dell'utilizzo il terreno si sterilizza in autoclave e se ne prelevano 20 ml, dopodiché si aggiunge l'antibiotico (Kanamicina) ad una concentrazione di 50 µg/ml, si effettua lo "scratch", ovvero si immerge un puntale della pipetta automatica nell'eppendorf contenente i batteri trasformati, così da prelevarne una piccola aliquota che rimane adesa alle pareti del puntale, poi lo si trasferisce nella provetta contenente l'LB e lo si lascia in incubatore *overnight* a 37°C e sotto agitazione a circa 180 rpm.

Il giorno seguente si effettua il cambio scala, ovvero si passa a far crescere i batteri da piccoli volumi (20 ml di LB) a grandi volumi (2 litri di M9).

Il cambio scala avviene centrifugando la provetta contenente i batteri a 2000 rpm per 5 minuti, si elimina il surnatante e si risospende il pellet in altri 20 ml di M9 (contenente kanamicina), dopodiché si dividono i due litri di M9 in 4 beute, cosicché ogni beuta abbia 500 ml di terreno e poi si dividono i 20 ml di sospensione di batteri in ogni beuta (circa 5 ml per beuta).

Al termine del procedimento si avrà una diluizione dei batteri cresciuti in LB di 1:100 in M9.

Il terreno M9, al contrario del precedente, è un terreno di coltura minimo con una composizione chimica definita (TABELLA 4.2), e che viene arricchito con dell'azoto marcato sottoforma di cloruro d'ammonio ( $^{15}\text{NH}_4\text{Cl}$ ); anche in questo caso viene aggiunto l'antibiotico kanamicina ad una concentrazione di 50 µg/ml.

Un terreno minimo è un terreno che contiene il minimo delle sostanze necessarie alla crescita batterica e questo permette di aggiungere una fonte di nutrimento marcata, che sarà l'unica fonte disponibile per i batteri, cosicché tutte le strutture che esso produrrà saranno marcate e si potranno utilizzare, ad esempio, per eseguire esperimenti bidimensionali all'NMR in cui si andranno ad osservare i segnali dell' $^{15}\text{N}$  e dell' $^1\text{H}$ , obiettivo di questa tesi.

Oltre all'azoto marcato è possibile anche aggiungere del carbonio marcato, sottoforma di glucosio ( $^{13}\text{C}_6\text{H}_{12}\text{O}_6$ ), e questo permette di eseguire degli esperimenti tridimensionali in cui si osservano i segnali di  $^{13}\text{C}$ ,  $^{15}\text{N}$  e  $^1\text{H}$ . Questo tipo di esperimenti permette di studiare la struttura tridimensionale di una proteina o di correlare i picchi che si osservano eseguendo un esperimento bidimensionale agli amminoacidi corrispondenti.

COMPOSIZIONE	QUANTITÀ
Na <sub>2</sub> HPO <sub>4</sub> anidro	6,8 g
KH <sub>2</sub> PO <sub>4</sub> anidro	3,0 g
NaCl	0,5 g
Glucosio anidro	4 g
<sup>15</sup> NH <sub>4</sub> Cl	1 g
MaSO <sub>4</sub> 1M	1 ml
CaCl <sub>2</sub> 2M	50 µl
MEM-Vitamins 100x	10 ml
Trace elements 100x	10 ml
Acqua milliQ	q.b. a 1 L

Tabella 4.2 – Composizione terreno M9

La crescita batterica avviene nell'arco di alcune ore e richiede un costante controllo, in quanto non è un processo costante nel tempo, ma avviene in 4 fasi ed è influenzato da fattori quali temperatura, agitazione e concentrazione di nutrienti presenti.

Solitamente la crescita batterica viene rappresentata mediante l'utilizzo di un grafico che riporta in ascissa il tempo e in ordinata il numero di cellule presenti nella coltura (FIGURA 4.5); si ha una prima fase di latenza (*lag*), in cui le cellule non si riproducono e si adattano al nuovo terreno, una fase di crescita esponenziale (*log*), in cui si ha una proliferazione esponenziale con conseguente esaurimento dei nutrienti, la terza fase è la fase stazionaria, dove si raggiunge un equilibrio tra la quantità di batteri che muoiono e di quelli che si dividono, ed infine si assiste alla fase di morte a causa della mancanza di nutrienti.<sup>49</sup>

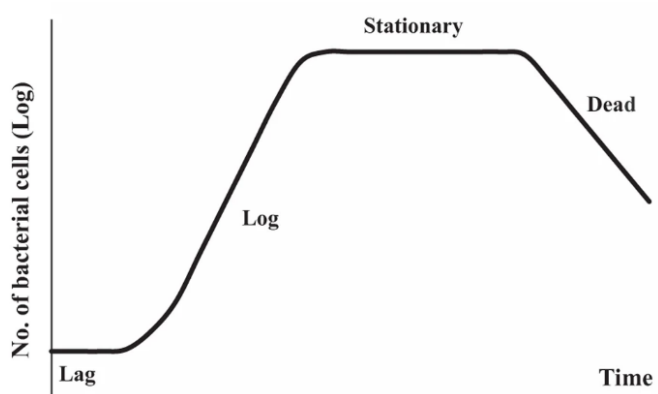


Figura 4.3 – Fasi della crescita batterica <sup>49</sup>

Nel caso di questa tesi sono state necessarie circa 5-6 ore per arrivare alla fase stazionaria. In questa fase si ha il numero massimo di cellule possibili per la quantità di terreno; a questo punto è stato possibile indurre l'espressione di Bfl-1, finora non espressa a causa della presenza del repressore lac. Per l'induzione viene utilizzato IPTG, che essendo un analogo



dell'allolattosio, permette l'inibizione del repressore lac e l'espressione della proteina di interesse. Viene utilizzato IPTG, in quanto permette la formazione di un legame covalente con il repressore lac, inibendolo irreversibilmente e impedendo la metabolizzazione.<sup>50</sup>

La crescita si effettua a temperatura controllata di 37°C e sotto agitazione a 180 rpm, mentre la misura della quantità di cellule presenti si effettua *mediante Fisherbrand™ Cell Density Meter (Biochrom-Ultrspect 10)*, uno spettrofotometro UV che emette nell'UV a 600nm e misura l'assorbanza del campione posto in una cuvetta di plastica trasformandola in densità ottica (OD). Lo strumento permette di ottenere un dato (OD<sub>600</sub>) di concentrazione del campione correlando in maniera proporzionale l'assorbanza, ovvero la quantità di luce assorbita dal campione misurata sfruttando il grado di scattering della luce, con la concentrazione della popolazione cellulare presente nel campione.<sup>51</sup>

La misurazione dell'OD<sub>600</sub> viene effettuata ogni ora sino a raggiungere un valore di circa 0,3, dopodiché la misura viene effettuata ogni mezzora sino a raggiungere un valore compreso tra 0,4 e 0,8 (FIGURA 4.6); a questo punto si induce aggiungendo ad ogni beuta una concentrazione di IPTG 1mM e le si ripone in incubatore ad una temperatura di 18°C con un'agitazione di 180 rpm. La concentrazione di IPTG, la temperatura di incubazione e gli rpm utilizzati per indurre l'espressione non sono standardizzati per ogni tipo di battere o di plasmide, ma bensì sono frutto di osservazioni sperimentali effettuate inizialmente, al fine di trovare la giusta combinazione di parametri in grado di portare alla resa massima possibile dell'espressione della proteina di interesse.

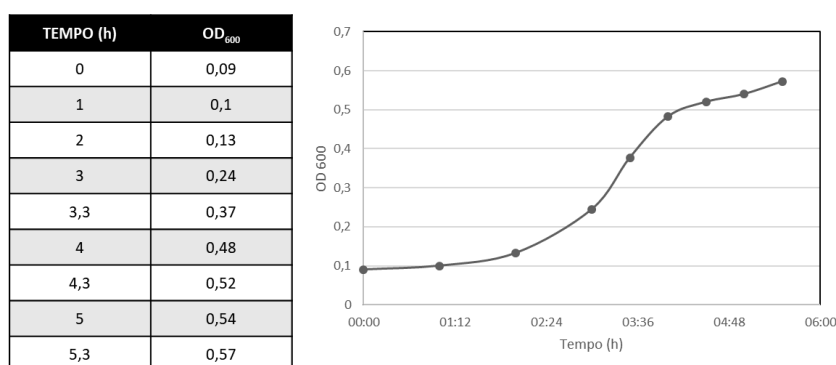


Figura 4.4 - Misura dell'OD<sub>600</sub> nel tempo, si possono vedere le prime due fasi di crescita batterica, sino a circa 1,30h si ha la fase di latenza, poi si assiste alla fase esponenziale e circa a 5h si raggiunge un OD<sub>600</sub> di 0,57, momento in cui si induce con

IPTG

### 4.3 Estrazione mediante sonicazione

Al termine dell'espressione si procede alla lisi delle cellule di *E. coli* mediante sonicazione, così da separare il materiale citosolico delle cellule dalla proteina.

Si procede centrifugando i due litri di coltura a 5000 rpm per 25 minuti. Questo passaggio permette di separare le cellule (che si depositeranno sul fondo) dal terreno. Dopodiché, si eliminerà il surnatante e si risospenderanno le cellule nel *lysis buffer* (TABELLA 4.3), ossia un *buffer* fosfato, contenente NaCl, glicerolo, imidazolo, BME ( $\beta$ -mercaptoetanololo, un agente antiossidante, che impedisce l'ossidazione dei residui di cisteina e PMSF (fenil-metil-sulfonil-fluoruro, un inibitore di serin-proteasi).

COMPOSIZIONE	QUANTITÀ
PO <sub>4</sub> pH 7,4	50 mM
NaCl	300 mM
Glicerolo	10 %
Imidazolo	15 mM
BME	2 mM
PMSF	2 mM
Acqua milliQ	q.b.

Tabella 4.3 – Composizione del *lysis buffer*

Dopo aver risospeso i batteri si procede con la lisi cellulare utilizzando *Vibra-Cell*<sup>®</sup> *VCX 130PB*, uno strumento costituito da una sonda metallica che viene inserita all'interno del campione e genera onde sonore ad alta frequenza. Queste onde sonore generano delle microscopiche bolle di vapore che implodendo causano delle onde d'urto che incontrano i batteri e ne causano la lisi; l'effetto fisico sfruttato viene definito cavitazione sonora.<sup>52</sup>

Lo svantaggio di questa tecnica è che le onde generate causano un aumento di temperatura del campione e questo potrebbe comportare la denaturazione della proteina che si vuole estrarre. Per evitare tale problema si immerge il campione in un bagno di ghiaccio e si effettuano dei cicli on/off, in modo da non far aumentare troppo la temperatura.

Al termine dell'operazione si può notare un cambiamento nel colore del lisato che assume una colorazione più scura a mano a mano che si procede con la sonicazione. Questo comportamento può tornare utile al fine di valutare in maniera empirica l'avvenuta lisi del campione.

## 4.4 Purificazione tramite Fast Protein Liquid Chromatography (FPLC)

Al termine della lisi si procede centrifugando il campione a 14000 rpm per 45 minuti al fine di separare tutte le componenti cellulari, che si depositeranno sul fondo, dalle componenti proteiche che si trovavano all'interno della cellula, che rimarranno in soluzione nel surnatante.

Dopo aver centrifugato, si procede con un primo step di purificazione filtrando il surnatante con un filtro PES con *cut-off* di 0,45  $\mu\text{m}$ , che permette di separare ulteriormente Bfl-1 da eventuali residui di membrana che altrimenti andrebbero ad impaccare la colonna cromatografica impedendo il passaggio della proteina di interesse.

Bfl-1 è stata purificata utilizzando lo strumento ÄKTA<sup>TM</sup> (Cytiva), mediante *Fast Protein Liquid Chromatography* (FPLC), una tecnica cromatografica che permette di lavorare con elevati flussi di liquido a bassa pressione, così da poter analizzare, separare o purificare miscele di proteine. Il vantaggio di questa tecnica, rispetto ad una tecnica simile come l'HPLC, è il fatto che lavorando a bassa pressione si riesce a preservare l'integrità della proteina.<sup>53</sup>

Come tutte le tecniche cromatografiche, l'FPLC si basa sulla separazione delle proteine per mezzo della loro affinità per la fase mobile o per la fase stazionaria della colonna. La fase mobile è rappresentata da una serie di *buffer* (TABELLA 4.4), che possono variare sia in composizione che in pH, in base alla stabilità della proteina; mentre la fase stazionaria varia in base al tipo di colonna utilizzata ed è composta da resine di vario tipo in base alle necessità di purificazione.

IMAC				GEL FILTRATION	
BUFFER A		BUFFER B		BUFFER C	
COMPOSIZIONE	QUANTITÀ	COMPOSIZIONE	QUANTITÀ	COMPOSIZIONE	QUANTITÀ
PO <sub>4</sub> pH 7,4	50 mM	PO <sub>4</sub> pH 7,4	50 mM	PO <sub>4</sub> pH 7,4	50 mM
NaCl	300 mM	NaCl	300 mM	NaCl	150 mM
Glicerolo	10 %	Glicerolo	10 %	DTT	2 mM
BME	2 mM	BME	2 mM	Acqua milliQ	q.b.
Acqua milliQ	q.b.	Imidazolo	500 mM		
		Acqua milliQ	q.b.		

Tabella 4.4 – Composizione dei buffer utilizzati nelle varie corse cromatografiche

Al fine di garantire una corretta purificazione si dovrà pulire lo strumento con un flusso di acqua milliQ dopo ogni step di purificazione. Nel caso di questa tesi saranno quattro passaggi

corrispondenti a quattro colonne cromatografiche con diverse affinità. Durante la corsa cromatografica la fase mobile passerà attraverso due detector, uno che permette di rilevare la concentrazione salina della soluzione andando a misurare la conducibilità, mentre il secondo è un rilevatore UV che permette di valutare l'assorbanza del campione ad una lunghezza d'onda di 280 nm; grazie a quest'ultimo rilevatore è possibile valutare la concentrazione della proteina.

#### 4.4.1 IMAC I

Nel primo step di purificazione si ha Bfl-1 fusa all'N-terminale con la proteina SUMO. Quest'ultima presenta all'estremità N-terminale una coda di 6 istidine, chiamata His-Tag. Quest'ultima sequenza amminoacidica risulta fondamentale ai fini della purificazione, in quanto viene legata dalla colonna *HisTrap™ HP 5 ml* (Cytiva), colonna che sfrutta il principio dell'*Immobilized Metal Ion Adsorption Chromatography* (IMAC), in cui la fase stazionaria è costituita da gel di agarosio cross-linkato e modificato con IDA (chelante imminodiacetato) che trattiene gli ioni  $Ni^{2+}$ ; i quali a loro volta coordinano con buona affinità gli anelli imidazolici delle 6 istidine (FIGURA 4.9), bloccando così nella colonna la proteina avente l'His-Tag e lasciando eluire tutte le altre proteine.<sup>54</sup>

Il grandissimo vantaggio di inserire nel plasmide la sequenza His-Tag sta nel fatto che la colonna IMAC è estremamente specifica per questa sequenza; inoltre in natura l'istidina è poco espressa e comunque l'His-Tag riesce a spiazzare eventuali istidine presenti in altre proteine, perciò si può essere sicuri che nella colonna si legni solamente la proteina di interesse.

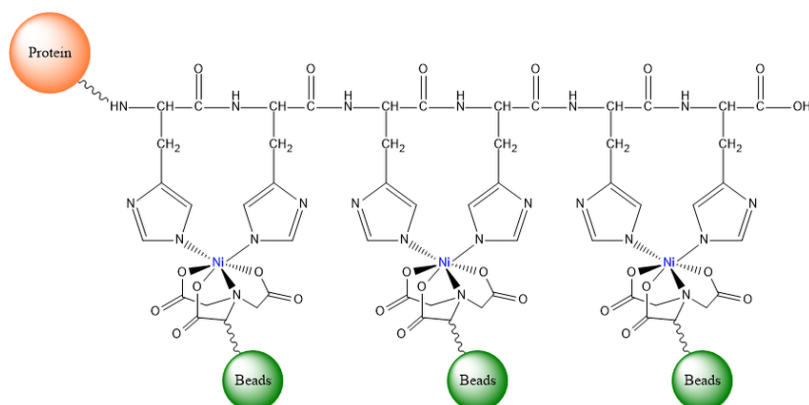


Figura 4.5 – Rappresentazione dell'interazione tra due istidine facenti parte dell'His-Tag che formano un legame di coordinazione con il  $Ni^{2+}$  a sua volta legato ad IDA.<sup>55</sup>

Come fase liquida si utilizzano il *buffer A* e il *buffer B* (TABELLA 4.4). Il primo si utilizza per far passare il campione in colonna e farlo legare al Ni<sup>2+</sup>. Dopo aver fatto passare l'intero campione attraverso la colonna, si procede alla sostituzione del *buffer A* con il *buffer B* mediante un gradiente che va dallo 0% al 100% di B; l'unica differenza tra i due *buffer* è la presenza di imidazolo, che entrando nella colonna compete con le istidine per il legame con il metallo permettendo l'eluizione della proteina con la fase mobile costituita dal *buffer B*.

Il vantaggio di questa tecnica è la possibilità di far passare in colonna elevati volumi di campione, in quanto la separazione continua ad avvenire fin tanto che la colonna non viene saturata.

#### 4.4.2 Desalting e taglio con SUMO proteasi

Il secondo passaggio cromatografico prevede l'utilizzo di una colonna *HiPrep™ 26/10 Desalting (Cytiva)*, che permette di rimuovere l'imidazolo, che essendo un inibitore di proteasi, impedirebbe il successivo taglio proteolitico con la SUMO proteasi.

La fase stazionaria della colonna è costituita da resine polimeriche di destrano, le quali formano dei pori che trattengono molecole e ioni a basso peso molecolare separandoli da macromolecole quali proteine, che non venendo bloccate nei pori avranno un tempo di ritenzione all'interno della colonna inferiore, permettendo di separare le due specie chimiche.<sup>56</sup>

Al termine della *Desalting* si ottiene un campione di proteina puro in soluzione nel *buffer A*. Tuttavia, Bfl-1 presenta ancora legati all'N-terminale: SUMO, che è stata necessaria per aumentare la solubilità della proteina permettendo di estrarla in maggior quantità dai batteri, e His-Tag, coda di sei istidine necessaria al legame con la colonna IMAC. A questo punto si procede aggiungendo SUMO proteasi, così da liberare Bfl-1 dai costrutti ora non più necessari.

Il taglio con SUMO proteasi avviene *overnight* a 20°C con agitazione a 100 rpm. Il vantaggio nell'utilizzare questa cistenil-proteasi risiede nel fatto che riconosce, con elevata specificità, la struttura tridimensionale di SUMO e ne idrolizza il legame peptidico a livello del C-terminale, separandola da Bfl-1. Più precisamente SUMO proteasi riconosce il legame peptidico nella sequenza x-Gly-Gly-x, andando ad effettuare il taglio sul legame successivo alla sequenza delle due glicine che si trova a livello del C-terminale di SUMO.

La proteasi viene prodotta in casa per via ricombinante ingegnerizzandola in modo tale che anch'essa presenti una coda di sei istidine all'estremità N-terminale in modo da facilitare le fasi di purificazione.

#### 4.4.3 IMAC II

Al termine del taglio proteolitico effettuato con SUMO proteasi si procede al quarto step di purificazione riutilizzando la stessa colonna utilizzata in precedenza, ovvero la colonna *HisTrap™ HP 5 ml* (Cytiva). In questo modo sarà possibile chelare nella colonna His-Tag legato a SUMO e la SUMO proteasi, così da avere la fuoriuscita immediata di Bfl-1 che non rimarrà legata alla colonna e verrà raccolta per procedere con l'ultimo step, ovvero la SEC. Come visto in precedenza, la componente rimasta legata nella colonna verrà rimossa utilizzando un gradiente del *buffer B*.

In questo passaggio, andando ad analizzare il picco fornito dal rilevatore UV presente nello strumento, è possibile quantificare la concentrazione di SUMO e di proteina la quale, se il taglio è avvenuto con successo, dovrebbe essere in rapporto 1:1.

#### 4.4.4 SEC

L'ultima colonna utilizzata è la *HiLoad™ 16/60 Superdex™ 75 prep grade* (Cytiva), colonna che sfrutta i principi della *Size Exclusion Chromatography* (SEC). La cromatografia ad esclusione dimensionale, permette la separazione di proteine sulla base del loro peso molecolare, sfruttando la struttura della fase stazionaria. Quest'ultima è formata da destrani cross-linkati a molecole di agarosio, che formano una struttura porosa in grado di rallentare notevolmente il tempo di ritenzione di proteine o peptidi più grandi rispetto alle proteine con dimensione inferiore.<sup>57</sup>

Per questa colonna viene utilizzato il *buffer C* (TABELLA 4.4), che presenta una minor concentrazione di sale, fondamentale per la stabilità in soluzione della proteina, ma che se presente in quantità eccessiva potrebbe inficiare l'analisi all'NMR, e viene tolto anche il glicerolo che, se presente, risulterebbe visibile all'NMR e quindi impedirebbe una corretta analisi dei dati.

Con quest'ultimo passaggio si conclude la purificazione di Bfl-1, che partendo da due litri di coltura porta ad avere un quantitativo di circa 15 mg di proteina pura per litro di coltura. Si

tratta di una quantità molto elevata di proteina marcata, che permette di eseguire uno screening NMR su numerosi campioni con una singola produzione. Visto il costo della proteina (migliaia di euro al grammo) e la quantità necessaria per eseguire i vari esperimenti, si tratta di un risultato decisamente rimarchevole, che ha richiesto un gran lavoro negli anni da parte dei ricercatori del Laboratorio DDbyNMR (Prof. Sturlese) per essere raggiunto.

## 4.5 Quantificazione tramite spettrofotometria UV

Al fine di misurare la concentrazione della proteina si utilizza la spettroscopia UV a 280 nm; si possono utilizzare due strumentazioni il cui principio risulta essere il medesimo, ma la cui applicazione varia a seconda del campione che ci si trova ad analizzare. I due strumenti utilizzati sono: il rilevatore UV all'interno dell'FPLC, utilizzato durante la produzione e che richiede il passaggio del campione attraverso l'FPLC; lo spettrofotometro *NanoDrop TM 1000 Spectrophotometer* (Thermo Scientific), utilizzato quando si ha un campione già purificato e che magari si è dovuto concentrare per ragioni sperimentali. Il vantaggio di utilizzare quest'ultimo strumento è la possibilità di effettuare la misurazione in modo molto preciso e con un quantitativo di campione di pochi microlitri.

Il principio su cui si basa questa tecnica è la misura dell'assorbanza a 280 nm; si andrà quindi a misurare l'assorbimento di questa radiazione da parte di alcuni amminoacidi in grado di produrre delle transizioni energetiche dei loro elettroni esterni.

Per effettuare tale misura nello strumento sono presenti: una sorgente luminosa (lampada a scarica di gas che può essere deuterio o idrogeno) che emette radiazioni nella regione dell'UV (tra i 100 e i 400 nm); un monocromatore, che disperde la luce policromatica in bande monocromatiche, così da poter selezionare la lunghezza d'onda a 280 nm; una cella, che può essere costituita da quarzo oppure, negli strumenti più innovativi (come nel caso di questa tesi), vengono posti 2-5  $\mu\text{l}$  di campione tra il monocromatore e il rilevatore, che sfruttando la tensione superficiale dei liquidi si disporranno a formare una goccia sferica la quale fungerà da cuvetta; un rilevatore, la parte più importante dello strumento in quanto determina sia la sensibilità che la precisione dello strumento.<sup>58</sup>

I dati dello strumento vengono analizzati applicando la legge di Lambert-Beer (*EQUAZIONE 4.1*), la quale mette in relazione l'assorbanza ( $A$ ) che risulta essere proporzionale al

coefficiente di estinzione molare ( $\epsilon$ ), allo spessore dello strato attraversato ( $l$ ) e alla concentrazione del campione ( $C$ ).

$$A = \epsilon l C \quad \text{Equazione 4.1}$$

$\epsilon$  è il coefficiente di estinzione molare (misurato in  $M^{-1}cm^{-1}$ ) e indica il valore di assorbanza del campione in condizioni standard con:  $l = 1$  cm e  $C = 1$  M. Il valore di questo coefficiente dipende dalla lunghezza d'onda della radiazione assorbita e dalla specie chimica che assorbe a tale lunghezza d'onda; quindi nel caso di Bfl-1 e delle proteine in generale, si tende ad utilizzare una radiazione del vicino UV (tra 200 e 380 nm), in quanto gli amminoacidi contenenti un anello aromatico (Phe, His, Trp, Tyr) hanno un coefficiente di estinzione molare ben noto tra i 250 e i 300 nm. Di questi in particolare si vanno ad analizzare Triptofano, Tirosina e Fenilalanina che presentano un massimo di assorbimento a 256 nm, inoltre si analizza anche la Cisteina, nonostante non sia un amminoacido aromatico, in quanto la  $\epsilon$  data dal gruppo tiolico ha una certa rilevanza a questa lunghezza d'onda.<sup>59</sup>

Nel caso della proteina studiata si hanno, nella sequenza amminoacidica: 3 triptofani, cinque tirosine ed una cisteina; al fine di calcolare la  $\epsilon$  totale si andranno a sommare le  $\epsilon$  degli amminoacidi presenti moltiplicate per il numero degli stessi; da questo si ottiene un valore di  $23950 M^{-1}cm^{-1}$ .

Il cammino ottico ( $l$ ) viene misurato in cm e dipende dallo strumento utilizzato. Invertendo l'EQUAZIONE 4.1 e conoscendo i valori di  $A$ ,  $\epsilon$  e  $l$ , è possibile calcolare la concentrazione (in M) del campione.

Gli strumenti utilizzati in questa tesi sfruttano software che valutano la concentrazione in mg/ml questo, in termini di formule matematiche, non modifica ciò che già è stato detto, tuttavia, il coefficiente di estinzione molare verrà diviso per la massa molare corrispondente, così da ottenere un coefficiente di estinzione molare espresso in grammi invece che in moli, che viene chiamato A 0,1%. Nella

Bfl-1		6His-SUMO-Bfl-1	
$\epsilon$	23 950 $M^{-1}cm^{-1}$	$\epsilon$	23950 $M^{-1}cm^{-1}$
MW	17 336,79	MW	25 346,91
A 0,1%	1,38 L/mol cm	A 0,1%	0,945 L/mol cm

Tabella 4.5 – Valori di  $\epsilon$ , Massa molecolare, A 0,1% di Bfl-1 e del costrutto 6His-SUMO-Bfl-1



## 4.6 SDS PAGE

*Sodium Dodecyl Sulphate – PolyAcrylamide Gel Electrophoresis* o SDS PAGE, è una tecnica elettroforetica qualitativa che permette di separare le proteine in base al loro peso molecolare. Tale tecnica viene utilizzata solitamente al termine della purificazione, al fine di controllare la purezza del campione ad ogni step.

La tecnica si basa sulla denaturazione di tutte le proteine presenti nei campioni raccolti, in modo da poterle separare unicamente in base al loro peso molecolare; per fare ciò si utilizzano: SDS (Sodio Dodecilsolfato) che lega gli amminoacidi delle proteine, in un rapporto costante di 1,4 g di SDS per grammo di proteina; quest'interazione in rapporto costante permette oltre alla denaturazione della proteina, anche la presenza di un rapporto massa/carica costante; DTT che riduce eventuali ponti disolfuro ed infine si utilizza un termoblocco a 95°C per completare la denaturazione. In questo modo si otterranno delle proteine tutte con lo stesso rapporto carica/massa e medesima mobilità elettroforetica.

Dopo il trattamento denaturante si caricano le soluzioni nei pozzetti del gel, costituito da acrilamide e metilenbisacrilamide che formano una struttura porosa in grado di separare le proteine in base al loro raggio idrodinamico. Solitamente il primo o l'ultimo pozzetto viene riempito con un marker contenente degli standard a colorati a vari pesi molecolari, così da poter avere un confronto a peso noto.

Per procedere con l'analisi si deve immergere il gel in una soluzione chiamata *Running Buffer*, costituito da MES, Tris base, SDS, EDTA e acqua milliQ; infine si applica una tensione di 180 mV per circa 45 minuti, in questo modo le proteine, cariche negativamente, potranno migrare attraverso il gel verso il polo positivo e lo spazio che percorreranno sarà inversamente proporzionale al loro peso molecolare.

Al termine della corsa elettroforetica si eseguono dei lavaggi del gel in acqua milliQ bollente, in modo da eliminare il *running buffer* in eccesso. Infine si aggiungerà Coomassie G-250, un colorante che lega i residui di Arginina e Lisina, permettendo di distinguere le bande rappresentanti le diverse proteine; dopodiché si procede ad un ulteriore lavaggio con milliQ.

A questo punto del procedimento risulta già possibile visualizzare le bande, ma per fissare il colorante alla proteina, viene lasciato riposare overnight in una soluzione acquosa contenente NaCl 20%.

## 4.7 Spettroscopia NMR, principi e applicazioni

Ai fini di questo progetto di ricerca la principale tecnica di analisi è stata la risonanza magnetica nucleare o spettroscopia NMR, una tecnica che permette di valutare sia la composizione chimica che la struttura molecolare di un campione.

In questa tesi sono stati presi in considerazione dieci molecole che sono state testate sulla proteina anti-apoptotica Bfl-1 precedentemente prodotta. Ai fini dell'analisi si è andati ad effettuare un'analisi NMR dei singoli composti per verificare che il composto comprato corrispondesse a quello arrivato e che non ci fossero modifiche chimiche date da eventuali instabilità molecolari. Dopo questa prima verifica si è effettuata l'analisi vera e propria dell'interazione proteina-ligando, che viene eseguita ponendo a confronto lo spettro NMR bidimensionale della proteina pura ed un secondo spettro della proteina a cui è stata aggiunta una o più molecole; in questo modo è possibile valutare la presenza di eventuali interazioni.

La tecnica si basa sull'applicazione di un campo magnetico che genera specifiche radiofrequenze, le quali vengono assorbite da specifici nuclei ( $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ,  $^{19}\text{F}$ ,  $^{31}\text{P}$ ,  $^{23}\text{Na}$ ); l'energia così assorbita permette la transizione ad un livello energetico maggiore degli spin di tali nuclei. L'energia assorbita viene poi riemessa, sempre sotto forma di radiofrequenza, dagli stessi nuclei, che tornano alla condizione iniziale. L'energia riemessa viene quindi rilevata e registrata sotto forma di grafico NMR.<sup>60</sup>

I nuclei atomici possono comportarsi come se ruotassero attorno al loro stesso asse, generando un momento angolare, il cui vettore ( $P$ ) dipende dal numero quantico di spin ( $I$ ). Lo spin nucleare viene generato dalle particelle costituenti il nucleo, ovvero protoni e neutroni; solitamente gli spin delle particelle costituenti il nucleo sono appaiati e la loro risultante è 0, tuttavia in alcuni atomi la risultante degli spin può essere diversa da 0:

- Se protoni e neutroni sono in entrambi pari nel nucleo, lo spin risultante sarà 0;

- Se protoni e neutroni sono entrambi in numero dispari il nucleo avrà spin Intero (1, 2, 3, 4...)
- Se protoni e neutroni risultano essere gli uni pari e gli altri dispari si avrà uno spin razionale (1/2, 5/2, 7/2...)

Nel caso dell’NMR solamente i nuclei con spin differente da 0 possono essere rilevati e solitamente gli esperimenti vengono condotti sfruttando atomi aventi spin 1/2.

$P$  ed  $I$  non variano mai in modulo, tuttavia il vettore momento di spin ( $P$ ) può variare il suo orientamento che, nel caso di atomi aventi  $I = 1/2$ , può assumere solamente stati energetici:  $+ 1/2$  e  $- 1/2$ , definite dal numero quantico di spin  $m$ . (FIGURA 4.10)

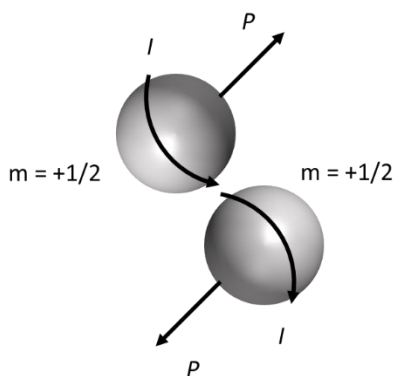


Figura 4.6 – Rappresentazione schematica dei vettori  $P$  ed  $I$  e del numero quantico di spin  $m$

Ogni nucleo dotato di spin essendo carico ed in movimento, genera un campo magnetico, e quindi è dotato di un momento magnetico nucleare di spin ( $\mu$ ). Tale momento magnetico ha modulo proporzionale e verso uguale al momento di spin; perciò si avrà, come nel caso precedente, due possibili versi dati dal numero quantico di spin.

Il rapporto tra il momento magnetico di spin ( $\mu$ ) e il momento di spin ( $P$ ) viene definito rapporto giromagnetico ( $\gamma$ ) (EQUAZIONE 4.2)

$$\gamma = \frac{\mu}{P} \qquad \text{Equazione 4.2}$$

I nuclei dotati di spin generalmente si trovano in egual misura nello stato di spin  $m \pm 1/2$  poiché possiedono la medesima energia; tuttavia se questi atomi vengono posti all’interno di un campo magnetico  $B_0$ , questi si orientano nello stesso verso del campo magnetico o in quello opposto (vengono definite: orientazione parallela o antiparallela), e possiedono

energia differente: l'orientazione parallela sarà a minore energia, mentre l'orientazione antiparallela risulterà essere a maggior energia.<sup>60</sup>

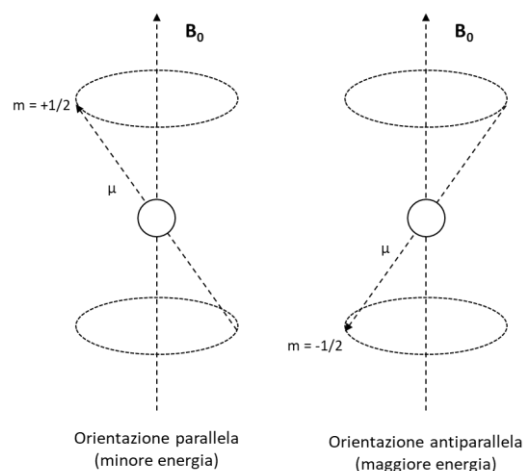


Figura 4.7 – Rappresentazione esemplificativa della variazione di spin in base al campo magnetico

L'interazione tra il campo magnetico ( $B_0$ ) e il momento magnetico ( $\mu$ ) determina l'energia di un nucleo, che può essere misurata tramite NMR.

L'EQUAZIONE 4.3, detta anche frequenza di Larmor, è un rapporto che rappresenta il cambiamento della direzione dell'asse di rotazione dei nuclei atomici attorno alla direzione di un campo magnetico:

$$\nu = \frac{\gamma B_0}{2\pi} \quad \text{Equazione 4.3}$$

La frequenza di Larmor correla il campo magnetico ( $B_0$ ) al rapporto giromagnetico ( $\gamma$ ), viene misurata in Hz (Hertz) e indica che all'aumentare del campo magnetico aumenta anche la frequenza di Larmor ( $\nu$ ); ovvero si avranno più nuclei che assumeranno un'orientazione antiparallela all'aumentare del campo magnetico e ciò comporta una maggior differenza di energia tra i due livelli.

L'equazione indica che all'aumentare del campo magnetico si avrà uno spostamento dei nuclei dallo stato  $+1/2$  (a più bassa energia) verso lo stato  $-1/2$  (a più alta energia); questo aumento energetico risulta necessario in quanto lo strumento necessita di certo quantitativo di energia per poter effettuare una misurazione e più alto sarà l'aumento di energia migliore sarà il dato ottenuto.

Inoltre bisogna anche considerare i nuclei che si vanno ad analizzare, poiché NELL'EQUAZIONE 4.3 si può notare come, nella scelta del campo magnetico, incida anche il rapporto giromagnetico dei nuclei, che più è elevato minore è il campo magnetico necessario per aumentare il divario energetico; oppure, a parità di campo magnetico, all'aumentare del rapporto giromagnetico si avrà maggiore sensibilità dello strumento.<sup>60</sup>

#### 4.7.1 Chemical shift

Se si andasse a valutare la frequenza di risonanza (EQUAZIONE 4.3) di un certo nucleo, preso singolarmente, ad un determinato campo magnetico, risulterebbe identico a tutti gli altri nuclei dello stesso tipo; ad esempio se si andasse a valutare la frequenza di risonanza dell'isotopo  $^1\text{H}$  preso singolarmente, ad un certo campo magnetico, risulterebbe identica a tutti gli altri atomi di  $^1\text{H}$  presi singolarmente e analizzati con lo stesso campo magnetico.<sup>61</sup>

Tuttavia, quando si vanno ad osservare i nuclei magneticamente attivi in una molecola, questi si trovano immersi in una nube elettronica, data dagli atomi circostanti che compongono la molecola. Questa nube elettronica è in grado di schermare il nucleo dal campo magnetico, ovvero il nucleo risentirà in maniera minore del campo magnetico e la sua frequenza di risonanza sarà di conseguenza minore.

Le differenze di frequenza risultano essere costanti in base all'intorno chimico e sebbene minime le une dalle altre è possibile rilevarle e distinguerle, questo permette di correlare la frequenza di risonanza alla struttura chimica della molecola.

L'effetto fisico appena descritto prende il nome di Chemical Shift, tuttavia si presenta un problema dato dal fatto che al variare del campo magnetico applicato, si avrà una variazione della frequenza di risonanza; questo non permetterebbe di paragonare due spettri fatti con campi magnetici anche leggermente differenti. Per ovviare al problema si utilizza una molecola standard di riferimento e si correla la frequenza di assorbimento del campione ( $\nu$ ) alla frequenza di assorbimento del riferimento ( $\nu_{standard}$ ), andando a definire l'EQUAZIONE 4.4 del Chemical Shift:

$$\delta = \frac{\nu - \nu_{standard}}{\nu} \cdot 10^6$$

Equazione 4.4

Il valore  $\delta$  viene comunemente misurato in ppm, ma risulta altrettanto possibile misurarlo in Hz; nella formula viene moltiplicato per  $10^6$  in quanto essendo un numero molto piccolo sarebbe più difficile da maneggiare come dato sperimentale, mentre in questo modo il dato risulta più intuibile e di più facile valutazione.

#### 4.7.2 FID e trasformata di Fourier

Dopo aver compreso che il segnale che si va a registrare è una variazione di energia, dallo stato più energetico (-1/2) allo stato meno energetico (+1/2) dipendente dal campo magnetico applicato, e che tale variazione di energia viene registrata come frequenza, si passa all'analisi dei dati raccolti.

I dati vengono letti sottoforma di grafico (FID), in cui sull'asse delle ascisse troviamo il tempo e sull'asse delle ordinate è posta l'intensità (FIGURA 4.8), il grafico che si presenta è formato da dei segnali sinusoidali che rappresentano la frequenza di risonanza degli atomi marcati analizzati, che andrà a diminuire a mano a mano che i nuclei si rilassano e tornano alla configurazione meno energetica. Il grafico che si andrà a leggere però non sarà formato solamente da un singolo segnale, ma da una serie di frequenze che si sovrappongono, si tratta dell'evidenza pratica del *Chemical Shift* (EQUAZIONE 4.4).<sup>62</sup>

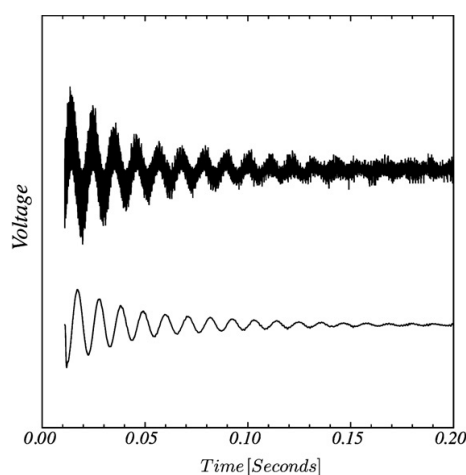


Figura 4.8 – Esempio di FID NMR <sup>62</sup>

La sovrapposizione di frequenze nel FID rende incomprensibili i dati ottenuti all'NMR, perciò si applica la Trasformata di Fourier, ovvero un'operazione matematica in grado di trasformare una funzione in un'altra mediante un'integrazione. La trasformata di Fourier permette quindi di passare dal dominio dei tempi a quello delle frequenze (FIGURA 4.9).

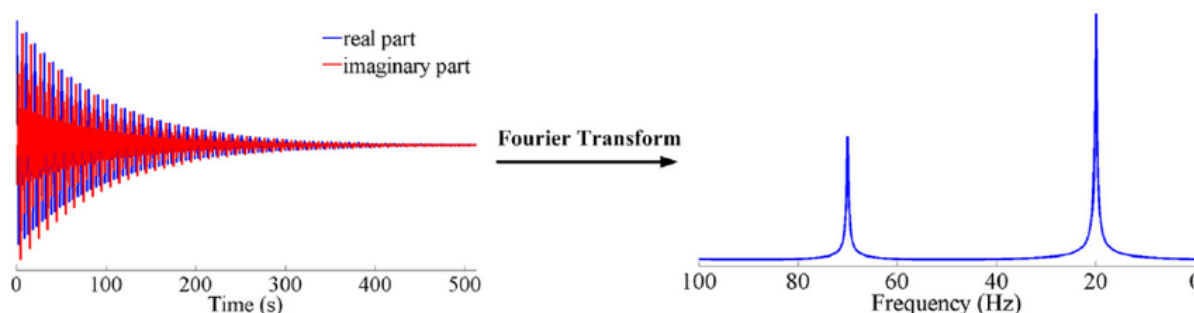


Figura 4.9 – Esempio di FID trasformato in grafico NMR mediante la trasformata di Fourier<sup>63</sup>

I concetti finora visti hanno portato a definire l'ottenimento di uno spettro monodimensionale, tuttavia applicando gli stessi principi è possibile eseguire anche spettri bi o tri dimensionali. Ai fini di questa tesi sono stati eseguiti degli esperimenti bidimensionali sulla proteina Bfl-1, andando ad osservare variazioni nello spettro ottenuto, le quali indicano un possibile legame con le molecole aggiunte.

Uno spettro bidimensionale richiede la presenza di due nuclei magneticamente attivi, solitamente  $^1\text{H}$  e  $^{15}\text{N}$ , a differenza degli esperimenti monodimensionali che ne richiedono solamente uno, solitamente  $^1\text{H}$ .

La misurazione avviene in maniera simile a quello che è stato visto in precedenza, quindi si ottiene un FID, che dovrà essere trasformato per ottenere uno spettro. Tuttavia, in questo caso, avendo due nuclei si osserveranno due differenti frequenze di risonanza ( $\nu$ ), di conseguenza si avranno due FID; al fine di ottenere uno spettro sarà quindi necessario applicare due volte la trasformata di Fourier, una per ogni dimensione.

Lo spettro che si ottiene da un esperimento bidimensionale, si presenta in maniera molto diversa rispetto allo spettro monodimensionale: il grafico presenta in ascissa e in ordinata le frequenze in ppm dei due nuclei magneticamente attivi. Ciò che si può osservare in questo tipo di esperimenti sono una serie di picchi dalla forma più o meno circolare, che rappresentano ognuno un amminoacido. Più in particolare, i picchi rappresentano l'accoppiamento scalare tra i due nuclei magneticamente attivi.<sup>60</sup>

#### 4.7.3 $^1\text{H}$ - $^{15}\text{N}$ SOFAST HMQC e Chemical Shift Perturbation (CSP)

Nel paragrafo precedente si è introdotto l'argomento degli esperimenti NMR bidimensionali, che si rivelano estremamente utili, se non necessari, quando si analizzano proteine; in

quanto, data la complessità strutturale, in uno spettro monodimensionale sarebbe impossibile riuscire ad assegnare i picchi ai relativi gruppi funzionali.

La differenza rispetto ad uno spettro monodimensionale sta nel fatto che il segnale proveniente dal campione è in funzione di un'unica variabile temporale ( $t_1$ ) e di conseguenza lo spettro letto tramite la trasformata di Fourier è dipendente da un'unica variabile di frequenza; mentre in un esperimento bidimensionale il segnale dei nuclei è in funzione di due variabili temporali ( $t_1$  e  $t_2$ ), di conseguenza sarà necessario applicare due volte la trasformata di Fourier in quanto dipendente da due variabili di frequenza.

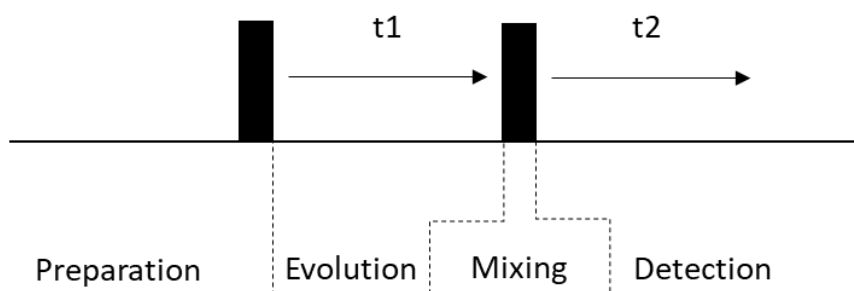


Figura 4.10 – Schema generale di un esperimento bidimensionale<sup>61</sup>

Nella FIGURA 4.10 si può vedere lo schema generale di un esperimento bidimensionale, in cui si ha una prima fase di preparazione (*Preparation*) in cui viene applicato un campo magnetico e, attraverso uno o più impulsi, vengono eccitati i nuclei portandoli ad un livello energetico maggiore ( $-1/2$ ). Il secondo step è l'evoluzione (*Evolution*), in cui durante il periodo di tempo  $t_1$  si lascerà evolvere la magnetizzazione. Al termine del tempo  $t_1$  si procede con il *Mixing time* che consiste nell'applicare un altro impulso o serie di impulsi, per poi passare alla fase di rilevazione (*Detection*), in cui dopo un tempo  $t_2$  si andranno a registrare i segnali (FID) e si applicheranno le due trasformate di Fourier per poter ottenere uno spettro bidimensionale.<sup>61</sup>

I segnali così ottenuti saranno poi trasposti in un grafico e appariranno come dei picchi di forma tondeggiante, definiti da delle coordinate X, Y: rappresentanti la frequenza in Hz o lo spostamento chimico relativo in ppm, dei due nuclei magneticamente attivi presenti nel campione.

La descrizione appena svolta degli aspetti generali alla base della spettroscopia NMR bidimensionale è alla base delle varie tecniche poi sviluppate nei vari campi di applicazione dell'NMR 2D; nel caso di questa tesi ci si sofferma su  $^1\text{H}$ - $^{15}\text{N}$  *Heteronuclear Single Quantum*



*Correlation* (HSQC) e  $^1\text{H}$ - $^{15}\text{N}$  *Heteronuclear Multiple Quantum Correlation* (HMQC). Si tratta di esperimenti NMR bidimensionali etero correlati, in cui la magnetizzazione del protone viene eccitata per poi essere trasferita all'etero nucleo  $^{15}\text{N}$ . Attraverso altri impulsi poi la magnetizzazione viene nuovamente trasferita al protone per poi essere infine rilevata direttamente; al contrario, la magnetizzazione dell'isotopo  $^{15}\text{N}$  viene rilevata per via indiretta durante il tempo di evoluzione  $t_1$ .

Il risultato che si ottiene è una serie di picchi che rappresentano l'accoppiamento scalare tra l'azoto e il protone ammidico (FIGURA 4.11). Ogni picco quindi corrisponde ad un gruppo NH del *backbone* di una proteina e delle catene laterali di quegli aa che contengono un gruppo -NH (ad esempio il triptofano), ovvero ad ogni amminoacido, e fatta eccezione per la prolina e molto spesso per l'N-terminale, il cui azoto tende ad essere un  $\text{NH}_3$ , difficilmente rilevabile in soluzione acquosa.<sup>60</sup>

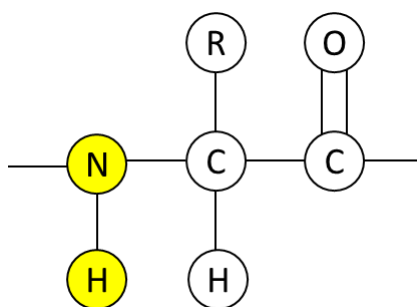


Figura 4.11 – In giallo sono evidenziati gli atomi che vengono correlati dagli esperimenti HSQC e HMQC per ottenere un picco rappresentato in uno spettro 2D

All'atto pratico quando si analizza un'interazione proteina-ligando attraverso uno spettro NMR bidimensionale si confrontano due spettri, un bianco, con sola proteina, ed uno spettro in cui il campione è costituito da proteina e ligando. Solitamente il ligando interagisce con il gruppo funzionale amminoacidico (R), ciò comporta una modificazione spaziale della proteina, che produrrà uno spostamento (*shift*) del picco, corrispondente all'amminoacido che interagisce con il ligando. Inoltre la modifica spaziale della proteina produrrà anche degli spostamenti nei picchi degli amminoacidi che costituiscono l'intorno dell'amminoacido maggiormente coinvolto nell'interazione, andando a definire un pattern di spostamenti che identifica una probabile interazione.

È importante sottolineare che nell'applicazione di questa tecnica sia necessario utilizzare una proteina marcata sia all' $^1\text{H}$  che all' $^{15}\text{N}$ , questo per poter sfruttare l'effetto di trasferimento della magnetizzazione. Oltre a ciò è bene ricordare che vengono generati degli impulsi

selettivi per i protoni amminici al fine di ridurre al minimo l'interferenza dovuta all'acqua nell'acquisizione dello spettro.

Per effettuare questa tesi si è impiegata una variante dell'HMQC, ovvero il SOFAST-HMQC, che permette di acquisire in un tempo ridotto uno spettro 2D. Per fare ciò si è diminuito al massimo il tempo che intercorre tra le scansioni, a questo scopo si sono dovute combinare una serie di accortezze, ovvero utilizzare un piccolo numero di impulsi, si è dovuto eccitare l'angolo di Ernst e si sono dovuti ottimizzare i rilassamenti longitudinali, tutto ciò al fine di migliorare la qualità del segnale rispetto al rumore che si viene a formare durante la veloce ripetizione di questi esperimenti. Tutto ciò permette di poter ottenere uno spettro di una proteina a 20  $\mu\text{M}$  in circa un'ora e mezza.<sup>64</sup>

Come visto in precedenza il *Chemical Shift* permette di distinguere i picchi di diversi nuclei (nel caso degli esperimenti bidimensionali nuclei correlati) in maniera specifica; in quanto, dipendentemente dal loro intorno chimico, avranno una frequenza di risonanza diversa dagli altri nuclei. Questo effetto viene sfruttato per l'analisi degli spettri bidimensionali, infatti viene valutato il cosiddetto *Chemical Shift Perturbation* (CSP), ovvero si monitorano gli spostamenti che avvengono tra i picchi quando viene aggiunto un ligando. Il ligando andando ad interagire con la proteina, in modo specifico nel suo sito di interazione, causerà dei piccoli cambiamenti conformazionali e modificherà l'intorno chimico di quei pochi amminoacidi che farà spostare il picco nel grafico.<sup>44</sup>

Per effettuare questa analisi è necessario acquisire un bianco, ovvero uno spettro della sola proteina, e poi acquisirne un secondo aggiungendo alla proteina un ligando. Questo permetterà di valutare se sono presenti o meno degli *shift* e dove si trovano. L'analisi inoltre è molto sensibile ai cambiamenti di concentrazione del ligando; perciò, è possibile valutare gli spostamenti che avvengono tra i picchi in modo estremamente preciso ed è possibile sfruttare questo spostamento per scoprire nuove *pockets*, per determinare il legame di una o più molecole con le stesse e per calcolare un'eventuale  $K_d$ .

Data l'elevata sensibilità della tecnica è necessario mantenere le condizioni sperimentali il più possibile invariate, perciò si utilizza lo stesso *buffer* in ogni tubo NMR che si vuole comparare, si mantiene la stessa concentrazione di proteina e si controlla sempre che non ci siano variazioni di pH tra i tubi.<sup>44</sup>

Per poter calcolare il CSP, come detto in precedenza, sono necessari due spettri i quali vengono processati prima con il *software* TopSpin.<sup>65</sup> Dopodiché, si utilizza il *software* Sparky<sup>66</sup> per assegnare i picchi nello spettro agli amminoacidi corrispondenti. Il passaggio effettuato mediante Sparky<sup>66</sup> necessita di una *peak list*, ovvero un documento che riporta il nome degli amminoacidi della proteina con le relative coordinate in ppm; quest'ultima si può scaricare da apposite banche dati e viene ricavata da una struttura cristallografica o da uno spettro NMR tridimensionale.

Dopo aver assegnato ad ogni picco l'amminoacido corrispondente si sovrappongono gli spettri del bianco con quello della proteina a cui è stato aggiunto il ligando, si procederà con una calibrazione (nel caso di questa tesi si è utilizzato G139, in quanto si è visto essere un amminoacido molto poco reattivo, quindi che non presenta spostamenti in presenza di ligandi, e semplice da distinguere dagli altri) e si potranno vedere degli spostamenti di alcuni picchi. Lo spostamento dei picchi a questo punto può essere visto come una distanza euclidea, perciò viene applicata l'*EQUAZIONE 4.5*.

$$\Delta\delta_{CSP} = \sqrt{\frac{1}{2}[\Delta\delta_H^2 + (\alpha\Delta\delta_N)^2]}$$

*Equazione 4.5*

Nell'equazione  $\Delta\delta_H$  rappresenta lo spostamento del picco sull'asse del protone (ascisse) in ppm, mentre  $\Delta\delta_N$  rappresenta lo spostamento sull'asse dell'azoto (ordinate).  $\alpha$  invece rappresenta un fattore correttivo, nel caso di questa tesi ha valore 0,14, necessario per poter confrontare i due spostamenti, in quanto nonostante l'utilizzo della stessa unità di misura la scala delle ascisse risulta differente da quella delle ordinate. Senza il fattore correttivo gli spostamenti su uno dei due assi avrebbero un peso diverso rispetto a quelli sull'altro asse, impedendo una corretta valutazione del CSP.<sup>44</sup>

Al fine di velocizzare l'analisi il Dott. Sturlese ha sviluppato uno *script (CSP mapping)* in grado di effettuare il calcolo del CSP per ogni singolo amminoacido, date le due *peak list (apo e holo)*. Inoltre, sempre tramite *script*, vengono applicate delle formule statistiche con diverse funzioni; in particolare in questa tesi se ne sono utilizzate due:

- Il *mean five (mean5)*, si tratta di una media del CSP dei 5 amminoacidi più perturbati. Questo dato permette di valutare più esperimenti con diversi ligandi, confrontandoli tra loro in modo da distinguere numericamente i ligandi più promettenti.

- Il triplo della deviazione standard ( $3*SD$ ), viene impiegato per selezionare, all'interno di uno stesso spettro, gli amminoacidi che possiedono un CSP maggiore del triplo della deviazione standard. Questi amminoacidi definiscono il *binding site*.

Lo *script* si rivela di grande aiuto nell'analisi di numerose miscele, tuttavia richiede una successiva verifica da parte dell'operatore.

Si è detto che per il calcolo del CSP viene valutato lo spostamento dei picchi, ma questo spostamento è strettamente legato all'equilibrio che viene ad instaurarsi nel campione analizzato. La proteina (P) si trova in equilibrio chimico con il ligando (L) (EQUAZIONE 4.6):



L'equilibrio sopra descritto è caratterizzato da due costanti di velocità ( $K_{on}$  e  $K_{off}$ ), che rappresentano le velocità della reazione diretta e di quella inversa; all'equilibrio il rapporto tra queste due costanti corrisponde alla  $K_d$ , ovvero alla costante di dissociazione, che solitamente si calcola secondo l'EQUAZIONE 4.7:

$$K_d = \frac{[PL]}{[P][L]} \quad \text{Equazione 4.7}$$

Quanto appena detto ha una grande rilevanza su ciò che si può visualizzare sullo spettro NMR. Il legame tra proteina e frammento può essere forte, quindi ad alta affinità oppure può essere debole, con una bassa affinità. La differenza di affinità genera due differenti possibilità: la prima viene definita *fast exchange*, ovvero proteina e ligando hanno bassa affinità. In caso di *fast exchange*, la  $K_{off}$  risulta essere più grande del CSP e per ciò si visualizzerà nello spettro uno spostamento lento del picco all'aumentare della concentrazione di ligando. La seconda possibilità al contrario viene definita *slow exchange*; il legame tra proteina e ligando è ad alta affinità e la  $K_{off}$  è minore del CSP. Ciò comporta che all'aumentare della concentrazione di ligando, nello spettro bidimensionale, si vedrà gradualmente sparire il picco e riapparire, sempre gradualmente, in un altro punto. Ciò che è stato appena detto viene esemplificato nella FIGURA 4.12.

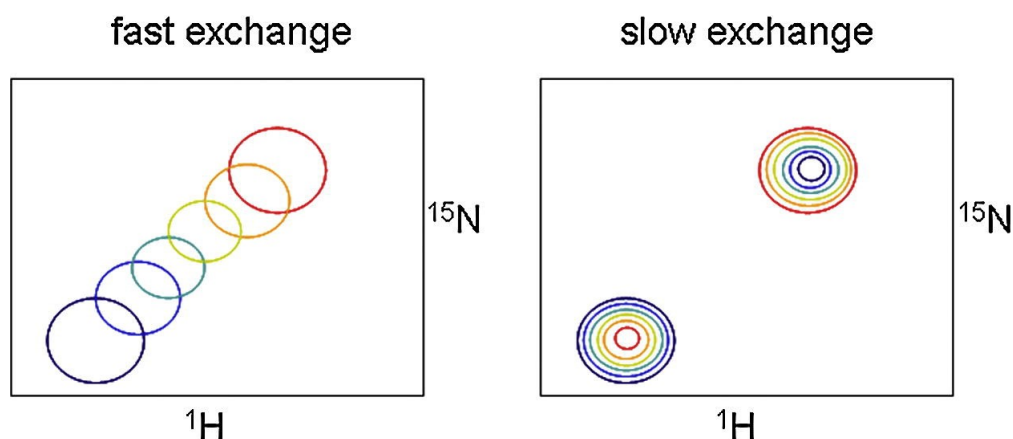


Figura 4.12 – Spettri NMR esemplificativi in cui è stato isolato un unico picco, nel primo viene rappresentata la condizione di fast exchange, mentre nel secondo quella di slow exchange. I picchi vengono rappresentati con colori diversi in quanto ad ogni colore corrisponde uno spettro differente ad una concentrazione di ligando differente <sup>44</sup>

È importante per questo lavoro di tesi sottolineare che lo spostamento che avviene all'aumentare della concentrazione, è sempre rivolto lungo una stessa direzione ed uno stesso verso; se ciò non fosse è possibile che ci si trovi di fronte ad un ligando che presenta più *binding modes* o, più probabilmente, che sia presente qualcosa che interferisce con la proteina alterandone la struttura, come ad esempio un errore nella preparazione del *buffer*, la proteina che inizia a degradarsi o la temperatura non costante.

#### 4.7.4 Copresenza

I dati ottenuti da questo progetto hanno portato alla curiosità scientifica di testare ulteriormente i siti di legame dei MiniFrag; per fare ciò si è deciso di aggiungere alla proteina un mix di frammenti che si è visto legare *pocket* differenti, questa pratica prende il nome di **copresenza**. I motivi per cui è sembrato interessante valutare delle copresenze sono principalmente due:

- Confrontare empiricamente l'affinità di un ligando rispetto ad un altro.
- Analizzare se due frammenti coesistono contemporaneamente all'interno di una stessa tasca, in punti diversi.

Per applicare questa tecnica si procede con la preparazione di un bianco contenente la proteina e uno dei due frammenti, solitamente quello ad affinità maggiore, e poi effettuare un esperimento con entrambe le molecole. In questo modo, quando si effettua l'analisi dello spettro, si potranno apprezzare solo gli *shift* provocati dal secondo frammento.

Ci saranno tre diversi casi che si verranno a configurare quando si applica tale tecnica:

- Nel primo caso si vedrà che gli *shift* presenti nel secondo spettro saranno talmente piccoli da essere considerati irrilevanti. Questa situazione indicherà che il secondo frammento non coesiste con il secondo. Ciò può essere dato dal fatto che la proteina subisce un cambiamento conformazionale, in presenza del primo ligando, che non crea più le condizioni chimiche ideali per l'interazione con il secondo.
- Il secondo caso prevede che gli *shift* che si visualizzano nel secondo spettro sono dovuti agli amminoacidi localizzati nella zona in cui si è vista precedentemente l'interazione con il primo ligando. Questo risultato è indice che il secondo ligando spiazza il secondo nella *pocket*. Perciò si potrà affermare che il secondo frammento sia più affine alla tasca rispetto al primo.
- L'ultimo caso è quello in cui si vedranno degli *shift* rilevanti, ma saranno concentrati in una *pocket* diversa da quella occupata dal primo frammento. Il significato di questo dato è che i due frammenti coesistono all'interno della *binding* site proteico, ma in punti differenti.

Ognuno di questi tre casi permette di acquisire maggiori informazioni sulle modalità di interazione ligando-proteina, nella prospettiva di maturare i ligandi e realizzare un composto *lead*.

## 5 RISULTATI SPERIMENTALI

### 5.1 Sintesi di Bfl-1 per via ricombinante

La fase iniziale di questa tesi è la produzione di Bfl-1 marcata all' $^{15}\text{N}$  per via ricombinante, per fare ciò si sono utilizzati due terreni di coltura LB ed M9, il primo per una prima crescita batterica, mentre il secondo per permettere di ottenere la proteina marcata. Al termine della crescita esponenziale in M9 si è poi indotta l'espressione con IPTG. Nella FIGURA 5.1 è stato riportato il grafico dell'andamento della crescita batterica, in ascissa si trova il tempo, mentre in ordinata il valore di  $\text{OD}_{600}$  corrispondente. Si può notare che in circa 5 ore si è raggiunto un valore di OD di circa 0,57; valore ottimale per l'induzione.

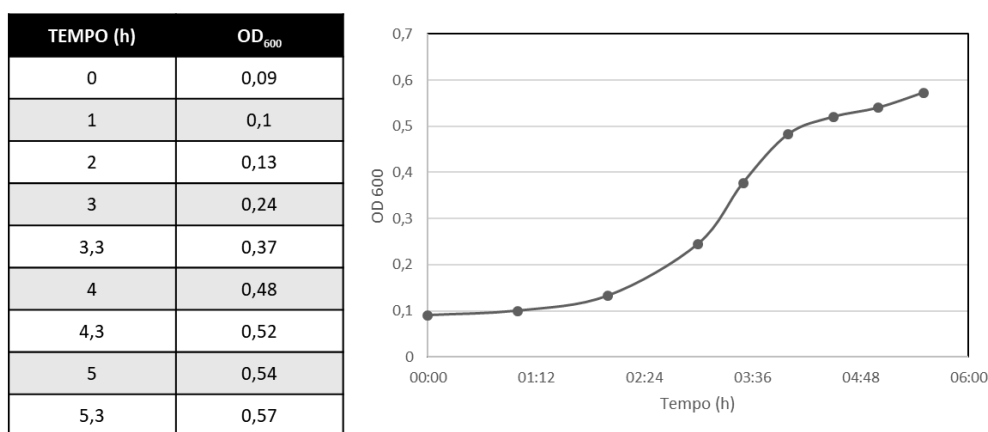


Figura 5.1 - misura dell' $\text{OD}_{600}$  nel tempo, si possono vedere le prime due fasi di crescita batterica, sino a circa 1,30h si ha la fase di latenza, poi si assiste alla fase esponenziale e circa a 5h si raggiunge un  $\text{OD}_{600}$  di 0,57, momento in cui si induce con IPTG

Il giorno seguenti si è misurato nuovamente l'OD per controllare la crescita batterica avvenuta durante la notte e si è ottenuto un valore di 2,1, valore ottimale in quanto ci si aspetta essere compreso tra 2 e 2,5.

### 5.2 Purificazione

Di seguito vengono riportate le varie fasi di purificazione, svolte mediante FPLC. Nella TABELLA 5.1 sono stati riportati i vari *buffer* utilizzati ai fini di questo processo. Dopo la produzione si è svolta una centrifuga; si sono risospesi i batteri nel *lysis buffer*; si sono lisati i batteri tramite sonicazione ed in fine si è eseguita una seconda centrifuga.

LYSIS BUFFER		IMAC				GEL FILTRATION	
COMPOSIZIONE	QUANTITÀ	BUFFER A		BUFFER B		BUFFER C	
		COMPOSIZIONE	QUANTITÀ	COMPOSIZIONE	QUANTITÀ	COMPOSIZIONE	QUANTITÀ
PO <sub>4</sub> ph 7,4	50 mM	PO <sub>4</sub> ph 7,4	50 mM	PO <sub>4</sub> ph 7,4	50 mM	PO <sub>4</sub> ph 7,4	50 mM
NaCl	300 mM	NaCl	300 mM	NaCl	300 mM	NaCl	150 mM
Glicerolo	10%	Glicerolo	10%	Glicerolo	10%	DTT	2 mM
Imidazolo	15 mM	BME	2 mM	BME	2 mM	Acqua MilliQ	q.b.
BME	2 mM	Acqua MilliQ	q.b.	Imidazolo	500 mM		
PMSF	2 mM			Acqua MilliQ	q.b.		
Acqua MilliQ	q.b.						

Tabella 5.1 – Buffer impiegati nella purificazione

### 5.2.1 IMAC I

Al termine della seconda centrifugazione si è filtrato il surnatante e si è proceduti ad effettuare una prima corsa cromatografica IMAC. In questo modo Bfl-1 è stata isolata dalle altre proteine batteriche. Questa separazione è avvenuta grazie all'interazione altamente specifica dell'His-Tag con il nichel della colonna.

Nella FIGURA 5.2 viene mostrato il cromatogramma che si genera in tempo reale nel momento in cui si effettua una purificazione tramite FPLC. Sono stati utilizzati i *buffer A* e *B*, il primo necessario per l'eluizione del campione nella colonna, in cui l'His-Tag si legherà alla colonna stessa; mentre il secondo *buffer* (rappresentato dalla linea verde) sarà necessario allo spiazzamento dell'His-Tag dal nichel.

Prima di caricare il campione è necessario equilibrare la colonna, che viene conservata in etanolo 20%, perciò si inizia facendo passare prima un flusso d'acqua MilliQ filtrata così da eliminare l'etanolo e poi si procede con un flusso di *buffer A* misto ad un 3% di *buffer B*, così da avere nella colonna una soluzione contenente 15 mM di imidazolo, ovvero la composizione del *buffer* nel quale si trova il campione.

A questo punto inizia la fase di *sample application*, in cui lo strumento preleva il campione. Da due litri di terreno M9 si sono ottenuti circa 40 ml di soluzione, infatti dal grafico (in blu) si può vedere il picco dell'UV che va da circa 10 ml a circa 50 ml.

Dopo il passaggio della soluzione in colonna si avrà Bfl-1 legata al nichel, perciò si procede con la fase di *column wash*, in cui si è impostato un flusso di *buffer* misto 97% A e 3% B. Il flusso permette di lavare via tutte le altre proteine solubili che si legano in modo specifico



alla colonna. Nel grafico si può vedere un aumento dell'UV tra i 50 e i gli 80 ml che corrisponde alle proteine non legate alla colonna che escono e vengono eliminate.

Dagli 80 ml sino ai 110 ml circa si assiste alla fase di *elution*, ovvero si imposta un gradiente di *buffer B*, che va dal 3% e arriva al 100%, questo aumento di imidazolo permette di spiazzare l'HIS TAG dal nichel. Lo strumento ha la possibilità di suddividere la soluzione che esce in aliquote da 2 ml.

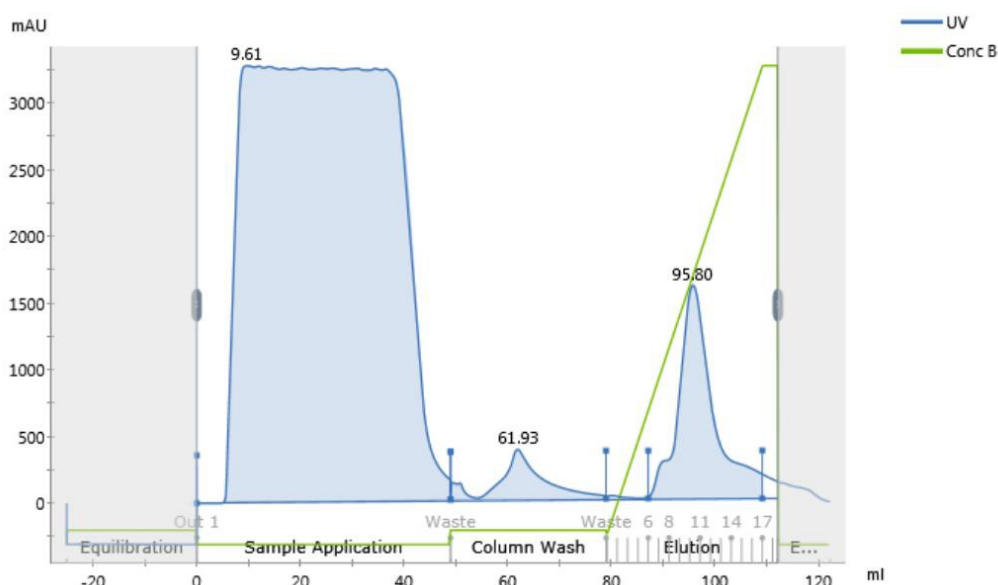


Figura 5.2 – Cromatogramma IMAC I Bfl-1

Osservando il cromatogramma si è deciso di prelevare le aliquote dalla 6 alla 17 esclusa, ovvero un campione di 22 ml e, tramite il software dell'FPLC si è calcolato la concentrazione, che ammonta a 116,94  $\mu\text{M}$ ; corrispondente a circa 65 mg di proteina. È importante ricordare che la proteina purificata finora è composta da 6His-SUMO-Bfl-1, perciò i successivi passaggi saranno necessari per ottenere solamente Bfl-1.

### 5.2.2 Desalting

Prima di procedere con il secondo step di purificazione è stato necessario concentrare il campione a 10 ml (limite massimo della colonna). La Desalting è necessaria per cambiare il *buffer* del campione con il *buffer A*.

Come si può osservare dalla FIGURA 5.3 nella fase di *sample application* il picco (in blu) dell'UV si protrae per circa 20 ml, in cui viene fatto passare il campione misto al *buffer A*, così da eliminare l'imidazolo, che uscirà successivamente (curva arancione nel grafico).

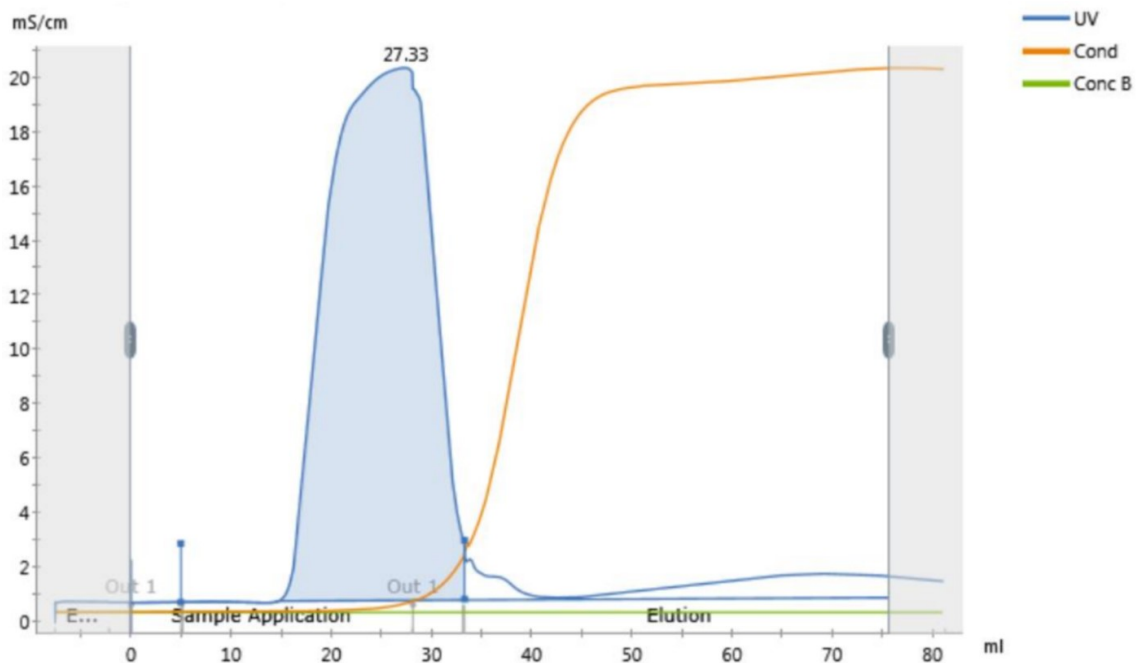


Figura 5.3 – Cromatogramma Desalting Bfl-1

Al termine della desalting si è aggiunto al campione la SUMO proteasi, che separa Bfl-1 da SUMO-6His; il taglio avverrà *over night*, sotto agitazione a 18°C.

### 5.2.3 IMAC II

L'ultimo giorno di purificazione inizia con una seconda cromatografia IMAC, così da poter separare Bfl-1, che non si legherà alla colonna, da SUMO-HIS TAG, che invece si legheranno alla fase stazionaria.

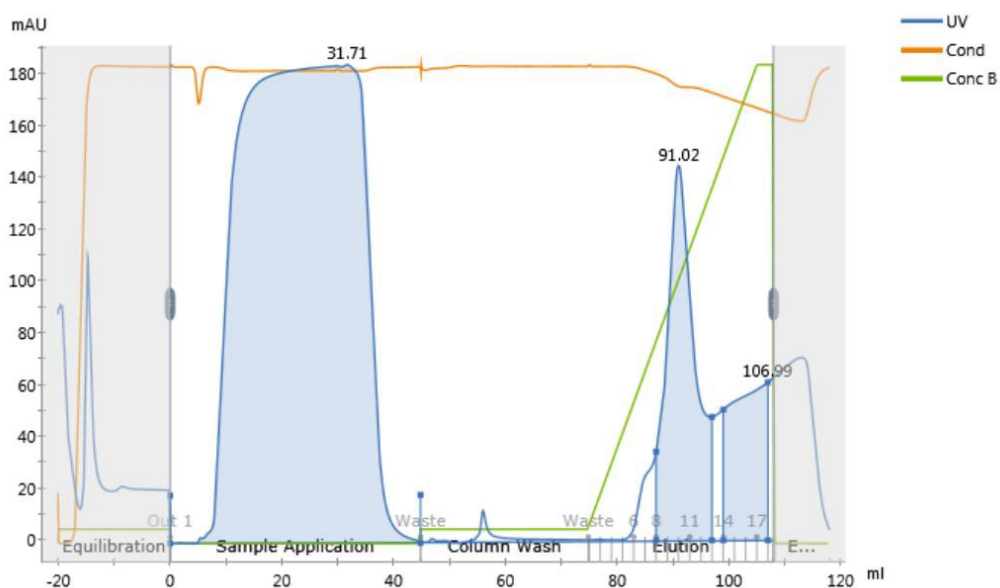


Figura 5.4 – Cromatogramma IMAC II Bfl-1

Come si può vedere dal grafico (*FIGURA 5.4*) si ha una prima fase in cui il campione viene prelevato, passa in colonna ed eluisce la proteina di interesse (*sample application*), successivamente si ha un aumento della concentrazione del *buffer* B (in verde), ovvero un aumento della concentrazione di imidazolo, che permette di spiazzare SUMO-His-tag e di avere il corrispondente picco (in blu) (*elution*).

Una verifica che può essere fatta per valutare l'efficienza del taglio effettuato da SUMO proteasi è quella di quantificare, usando il cromatogramma e inserendo i relativi valori di assorbanza, le moli di Bfl-1 e di 6His-SUMO; se le moli corrispondono o sono quasi le stesse allora il taglio ha avuto un'elevata efficienza.

Al termine della IMAC II si raccoglie la soluzione uscente nella fase di *sample application*, corrispondente a circa 31 ml, ad una concentrazione di 0,387 mg/ml, che se divisa per la massa molare di Bfl-1 (17336,79 g/mol), restituisce una concentrazione di 22,32  $\mu$ M

#### **5.2.4 Gel Filtration**

L'ultimo passaggio che viene effettuato tramite FPLC è la cromatografia *Gel Filtration*. La colonna che viene utilizzata è la *HiLoad™ 16/60 Superdex™ 75 prep grade (GE Healthcare)*, la quale permette di far passare al suo interno un massimo di 5 ml di campione, perciò come nel caso della *desalting*, è necessario concentrare il campione, a 5 ml.

In questo passaggio viene anche effettuato un cambio di *buffer*, verrà utilizzato il *buffer* C (*TABELLA 5.1*), che contiene una percentuale minore di sale e non contiene glicerolo, che altrimenti inficerebbe le analisi tramite NMR.

Nel cromatogramma (*FIGURA 5.5*) si può osservare il picco tra i 75 e i 90 ml, corrispondenti a Bfl-1, anche in questo caso come era stato fatto per l'IMAC I, si è impostato lo strumento in modo che raccogliesse l'eluato in frazioni da 3,5 ml. Come è stato per la corsa *desalting*, anche in questo caso si assiste ad una diluizione del campione, che da 5 ml caricati, ne permette l'eluizione di 10/15 ml; in questo modo però si ottiene la proteina in un *buffer* che non contiene imidazolo o glicerolo ed ha una quantità di NaCl pari alla metà rispetto ai *buffer* utilizzati in precedenza.

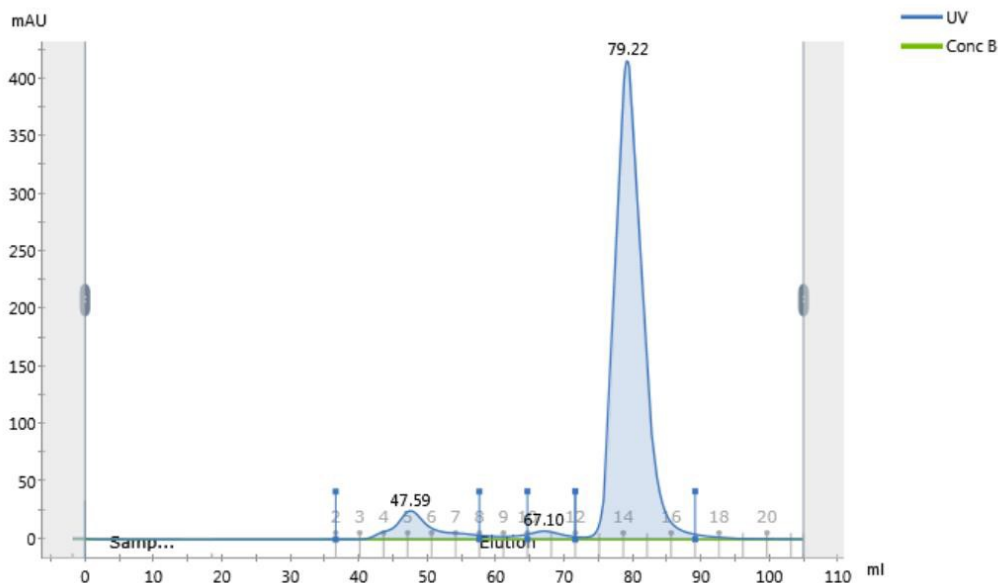


Figura 5.5 - Cromatogramma SEC

In questo caso sono state prese le frazioni dalla 12 alla 16 e riunite in una falcon, il volume così ottenuto è di 17,5 ml. Anche in questo caso si procede con la quantifica della proteina via *software*, inserendo il coefficiente di assorbanza (1,38); si ottiene così un valore di concentrazione pari a 0,393 mg/ml che, divisi per il peso molecolare di Bfl-1 (17336,79 g/mol), danno come risultato una concentrazione di 22,66  $\mu$ M.

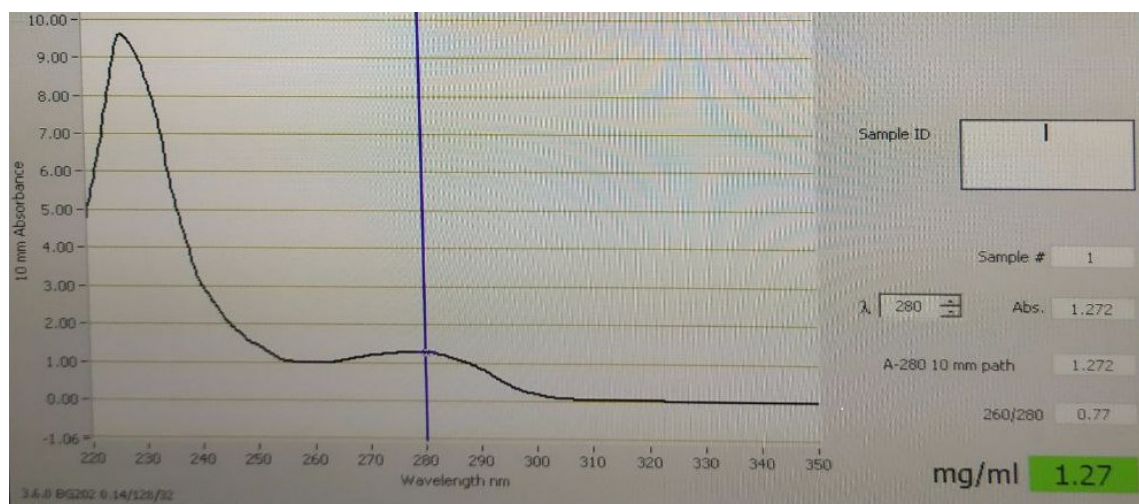
La quantità di proteina così ottenuta è di 6,873 mg partendo da un quantitativo di 2 L di coltura batterica. Questo dato è estremamente significativo, in quanto la resa di questa proteina è estremamente alta e per arrivare ad una resa così elevata sono stati necessari anni di ricerca e miglioramento delle procedure da parte del laboratorio DDbyNMR (Dott. Sturlese).

Il campione puro che si ottiene solitamente non viene utilizzato immediatamente, perciò si conserva la proteina in frigo, aggiungendo sodio azide 0,02%, in modo da migliorarne la stabilità e rallentarne il deterioramento.

### 5.3 Analisi quantitativa tramite spettrofotometria UV

Per avere un'idea dell'andamento della purificazione, solitamente si utilizza lo spettrofotometro UV interno all'FPLC, tuttavia si tratta di uno strumento non molto preciso e che quantifica la proteina nel momento in cui esce dalla colonna, perciò, al fine di avere un dato molto più affidabile, si utilizza lo spettrofotometro ottico *NanoDrop TM 1000 Spectrophotometer*.

Nella *Figura 5.6* si può osservare che il valore di concentrazione risulta essere 1,27 mg/ml, questo dato risulta necessario dividerlo per il valore di massa molare di Bfl-1 (17336,79), ottenendo come valore di concentrazione di 73,25  $\mu$ M. La concentrazione ottenuta risulta essere soddisfacente per poi poter realizzare dei tubi NMR con una concentrazione finale di proteina di 20  $\mu$ M.



*Figura 5.6 – Quantificazione di Bfl-1 effettuata tramite NanoDrop TM 1000 Spectrophotometer*

## 5.4 SDS page

Al termine della purificazione viene sempre effettuata un'analisi qualitativa del processo di produzione e purificazione, per poter avere la certezza che sia stata effettivamente prodotta la proteina di interesse e per valutare la bontà del processo di produzione. Per fare ciò si raccolgono dei campioni prima o al termine di ogni step di purificazione e una volta concluso il processo si effettua un gel tramite SDS page.

Ciò che si ottiene alla fine del processo di colorazione è rappresentato nella *FIGURA 5.7*; in questo modo si possono osservare i vari passaggi di purificazione. Inizialmente si trova il *marker*, ovvero uno standard di pesi molecolari che ha un range da 180 kDa, sino a circa 10 kDa, come detto si tratta di una tecnica qualitativa, perciò la distinzione che viene effettuata non è precisa, ma è sufficiente per poter analizzare il contenuto dei campioni.

Dal momento che la proteina viene sovraespressa nelle cellule batteriche, ci si aspetta che in corrispondenza della stessa si evidenzi una banda blu più intensa e marcata rispetto alle altre. Nella prima *lane* (Pre IMAC I) si è prelevato un campione dopo la seconda centrifuga e prima dell'utilizzo dell'FPLC, ciò che emerge è che effettivamente nel campione ci sia un mix

di proteine costitutive e che tra 40 e 35 kDa ci sia una banda più marcata delle altre, anche se manca di risoluzione permette di avere un'indicazione che probabilmente nel campione sia stato sovraespresso il costrutto di partenza 6His-SUMO-Bfl-1.

Il *flow through* IMAC I è ciò che è eluito dalla corsa, quindi ciò che non si è legato alla colonna; questo campione è stato preso per valutare se la proteina di interesse sia rimasta legata alla colonna o se una parte non si sia legata; è possibile che l'His-Tag non si legni al nichel per vari motivi, i principali sono che la proteina inizi ad aggregare oppure che si sia saturata la colonna o ancora che la colonna non sia più efficiente. Nel caso di questo gel sono presenti numerose bande, dovute a tutte le proteine batteriche che non si sono legate in colonna, e non si vedono bande molto più intense di altre, questo significa che la proteina di interesse si è legata tutta alla colonna e non è presente nell'eluito.

La terza *lane* è il campione preso dopo il taglio con SUMO proteasi e prima della seconda IMAC, in questo caso è stato evidenziato in verde il costrutto 6His-SUMO-Bfl-1 (25346,91 Da), ciò indica che SUMO proteasi non ha tagliato completamente il campione, questo dato è importante per eventualmente cercare di migliorare le condizioni di temperatura-agitazione a cui avviene il taglio o la concentrazione della proteasi; tuttavia in questo caso specifico la quantità di proteina che è stata tagliata è più che sufficiente. In giallo è stata evidenziata la banda corrispondente alla SUMO proteasi, che si trova ancora nel campione. In fine in rosso si evidenzia Bfl-1 tagliata che ha un peso molecolare che si aggira sui 17 kDa, perciò la si può vedere vicino alla banda segnata dal marker dei 15 kDa.

Nella quarta *lane* è stato prelevato un campione nella fase di *Elution* dell'IMAC II, ovvero la fase in cui aumenta l'imidazolo e ciò che era legato in colonna viene spiazzato e può eluire. Come nella *lane* precedente si possono vedere in verde il costrutto intero, che conserva ancora l'HIS TAG, e in giallo SUMO proteasi, che viene prodotta anch'essa con l'HIS TAG appositamente per poterla far legare alla colonna e separarla dalla proteina di interesse.

L'ultimo campione è stato prelevato al termine della SEC, per poter constatare che nel campione sia presente solo ed esclusivamente Bfl-1, ed infatti la si può vedere in rosso come unica banda in tutta la *lane*.

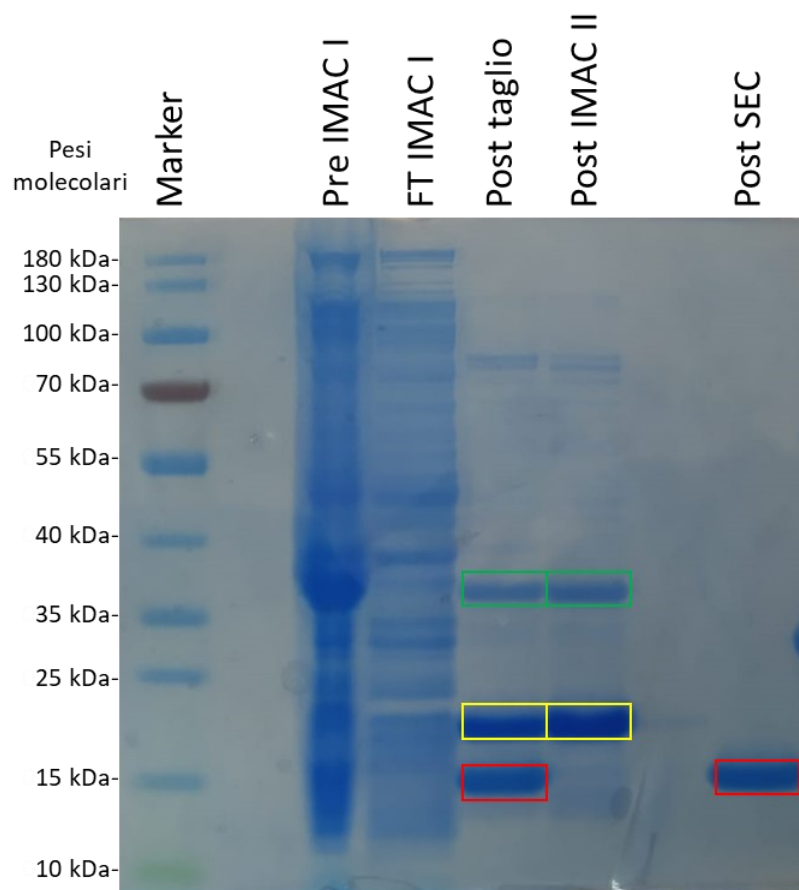


Figura 5.7 – Gel SDS Page, sono evidenziati vari passaggi di purificazione, vengono evidenziati in verde il costrutto 6His-SUMO-Bfl-1, in giallo 6His-SUMO, mentre in rosso la proteina di interesse Bfl-1

## 5.5 MiniFrag costruzione dello stock

In questa tesi si è partiti da una piccola libreria di 10 MiniFrag, che sono stati selezionati tramite screening computazionale di 100 composti, in cui si è esplorato uno spazio chimico potenziale di  $10^4$ , e in cui sono state scelte delle molecole che potessero interagire con delle specifiche *pocket* della tasca idrofobica di Bfl-1. La selezione effettuata con questo criterio è stata fatta al fine di poter poi legare assieme, tramite ponte idrocarburico, le molecole più reattive nei diversi *pocket* e ottenere così un composto *lead*.

Nella FIGURA 5.8 vengono mostrate le formule e la nomenclatura interna utilizzata per distinguerli in modo semplice negli esperimenti.



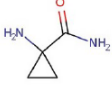
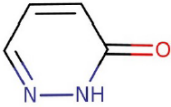
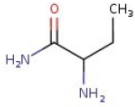
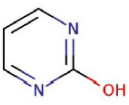
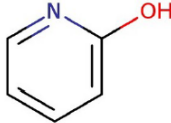
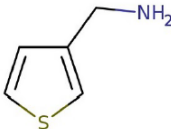

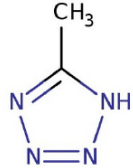
				
A1	A2	B2	B6	C2
				
C3	C4	C5	C6	C7

Figura 5.8 – Struttura MiniFrag con corrispondente codice interno

Come spiegato precedentemente le simulazioni che sono state effettuate tengono conto della dinamica molecolare (*molecular dynamics simulation*), quindi sia dei piccoli cambiamenti conformazionali della proteina, sia delle diverse disposizioni spaziali del frammento, questo ha comportato che per ogni frammento siano state fornite le 5 migliori *poses* per ogni *pocket* della tasca, quindi in totale per ogni MiniFrag sono state fornite 25 immagini rappresentanti una possibile interazione tra proteina e ligando. Le *poses* fornite sono state poi valutate se coerenti o meno e confrontate con i dati sperimentali, ottenendo al massimo quattro o cinque *poses* per frammento.

L'utilizzo di MiniFrag ai fini della *fragment-based drug discovery* (FBDD), da letteratura, per ora è stato effettuato solo tramite cristallografia a raggi X,<sup>46</sup> perciò ci si è basati su quell'articolo per iniziare la sperimentazione tramite NMR.

Il primo passo è stato quello di preparare gli *stock* concentrati di frammento per poterli poi aggiungere nei tubi NMR. Nell'articolo si parla di concentrazioni in *stock* acquoso estremamente elevate (1M), per poi diluirle nel campione da analizzare a circa 50-100 mM. È stato necessario effettuare alcune prove di solubilità, inizialmente si è provato ad utilizzare come solvente il dimetilsolfossido (DMSO), un solvente organico che permette la solubilizzazione dei composti ad alte concentrazioni, ma che, se in concentrazione troppo elevata, nel campione da analizzare all'NMR ne inficia l'analisi andando a sovrastare il segnale del frammento. Dopo aver provato con il DMSO si è passati all'acqua milliQ, ma effettuando qualche prova di solubilità, i frammenti non risultavano sufficientemente solubili. Infine si è optato per provare ad utilizzare come solvente un tampone fosfato



concentrato (500 mM), così da permettere di avere un solvente acquoso, di aumentare la solubilità grazie alla presenza del tampone e di non aggiungere nulla, al momento della preparazione del campione per l'analisi NMR, che possa far precipitare la proteina o che ne infici l'analisi. (TABELLA 5.2)

Solubilità in PO <sub>4</sub> 500 mM pH 7,4										
Posizione	A1	A2	B2	B6	C2	C3	C4	C5	C6	C7
Concentrazione	500 mM	500 mM	1 M	1 M	1 M	1 M	1 M	250 mM	1 M	1 M

Tabella 5.2 – Concentrazione delle soluzioni madre di MiniFrag

Nell'effettuare le solubilizzazioni per lo stock ci si è resi conto che a concentrazioni così alte di frammento il pH tendeva ad esasperarsi nonostante la presenza di un tampone molto concentrato, perciò è stato necessario aggiustare il pH il più vicino possibile a 7,4; questo perché poi nel momento in cui si andava a realizzare il campione contenente la proteina questo teneva a precipitare per aggiunta di una soluzione a pH estremamente acido o basico.

Oltre alla problematica del pH ci si è resi conto anche che la quantità di frammento puro, necessario per realizzare la soluzione di *stock*, era veramente alta rispetto alle quantità di frammento ordinate; per questo motivo il frammento B2 non è stato possibile analizzarlo, in quanto non è risultato sufficiente per effettuare un'analisi soddisfacente.

## 5.6 Preparazione del campione per l'analisi NMR

Dopo aver realizzato le soluzioni madre dei frammenti si è passati alla valutazione della preparazione dei campioni. Si è tenuto conto sia delle condizioni sperimentali descritte nell'articolo di *O'Reilly et al*<sup>46</sup> sia degli esperimenti solitamente effettuati nella FBDD tradizionale. Perciò si è deciso di utilizzare una concentrazione di proteina di 20 µM, che risulta essere relativamente bassa, ma che permette di ottenere uno spettro sufficientemente risoluto.

La decisione a questo punto è stata sulla concentrazione di frammento da aggiungere per poter effettivamente apprezzare un'eventuale interazione; ci si è sempre basati sull'articolo di O'Reilly, e si sono effettuate alcuni esperimenti di prova, alla fine si è deciso di procedere effettuando gli esperimenti con due concentrazioni differenti di frammento: a 20 mM e a

100 mM. Il primo è un rapporto proteina/frammento di 1:1000, mentre il secondo è un rapporto di 1:5000, questi due rapporti sono stati scelti perché si sono visti dei cambiamenti nel CSP già a 20 mM, ma spesso risultavano molto esigui. Perciò si è deciso di utilizzare un rapporto 5000 volte maggiore per poter apprezzare uno spostamento maggiore.

Il campione risulta quindi costituito da Bfl-1 marcata all' $^{15}\text{N}$ , acqua deuterata che fornisce  $^1\text{H}$ , il MiniFrag ed eventualmente il *buffer* in cui si trova la proteina per portare a volume di 550  $\mu\text{L}$ . Nella *TABELLA 5.3* vengono riportate le concentrazioni dei componenti del campione NMR; una cosa fondamentale dopo aver miscelato le varie soluzioni è stata quella di aggiustare il pH che fosse il più vicino possibile a 7,4, questo per evitare eventuali errori di valutazione dello spettro dovuto al pH. Si è deciso di accettare variazioni di pH nell'ordine di  $\pm 0,02$  punti pH.

COMPOSIZIONE CAMPIONE NMR	
Composizione	Concentrazione
Proteina	20 $\mu\text{M}$
D <sub>2</sub> O	5%
MiniFrag	20 mM o 100 mM
Buffer PO <sub>4</sub> 50 mM NaCl 150 mM DTT 2 mM pH 7,4	q.b. a 550 $\mu\text{L}$

*Tabella 5.3 – Composizione del campione NMR*

La tecnica applicata per l'analisi è il SOFAST-HMQC, che permette di ottenere in circa un'ora e mezza lo spettro bidimensionale. Sono stati effettuati più esperimenti in giorni differenti, perciò è sempre stato realizzato un bianco, da utilizzare poi per il calcolo del CSP, mettendolo a confronto con gli esperimenti effettuati lo stesso giorno.

Si è osservato empiricamente che per ottenere risultati migliori fosse necessario porre il campione in incubazione prima di procedere all'analisi. La procedura di incubazione è stata effettuata preparando il campione il giorno prima dell'esperimento, lasciando il campione in frigorifero *over night*, questo ha permesso al MiniFrag, che ricordiamo essere un ligando molto debole, di instaurare un equilibrio di legame. In questo modo la quantità di molecole legate alla proteina risulta essere maggiore e più facilmente apprezzabile all'NMR.

## 5.7 Analisi degli spettri NMR

Gli spettri ottenuti vengono prima processati mediante TopSpin<sup>65</sup> e poi vengono aperti su Sparky<sup>66</sup>, da questo software è possibile aggiustare la *peak list* contenente l'assegnazione dei picchi sugli spettri da analizzare.

L'esperimento SOFAST-HMQC permette di avere uno spettro con una buona risoluzione, ma non la migliore possibile. Questo comporta che molti picchi non siano facilmente distinguibili. Per questo motivo dalla *peak list* di partenza vengono eliminati tutti quegli amminoacidi che cadono sullo stesso sistema di spin.

Dopo aver creato le nuove *peak list* è possibile calcolare il CSP di ogni spettro contenente il frammento confrontandolo con lo spettro del bianco. Per fare ciò è stato utilizzato lo *script 'CSP mapping'*, che permette anche di calcolare il *mean5* e il triplo della deviazione standard automaticamente.

Gli spostamenti dei picchi, all'aumentare della concentrazione, avvengono solitamente lungo una stessa direzione ed in uno stesso verso. In questa tesi vista la dimensione dei frammenti e gli eventuali problemi dovuti al pH si è deciso di aggiungere un criterio di esclusione all'analisi che prevede, a fronte di un alto valore di CSP, di non prendere in considerazione quei picchi che nei due spettri, effettuati a diverse concentrazioni di ligando, non hanno medesima direzione e verso.

Dopo aver selezionato gli amminoacidi con CSP maggiore, si sono esclusi quelli che non si spostavano nella stessa direzione e/o verso e si è poi passati all'analisi su Chimera<sup>21</sup>, altro software che permette di visualizzare la struttura tridimensionale delle proteine. Una volta che si sono evidenziati gli amminoacidi perturbati si è confrontato il risultato sperimentale con le 25 *poses* fornite dal laboratorio MMS, cercando un riscontro tra dato sperimentale e computazionale.

## 5.7.1 A1

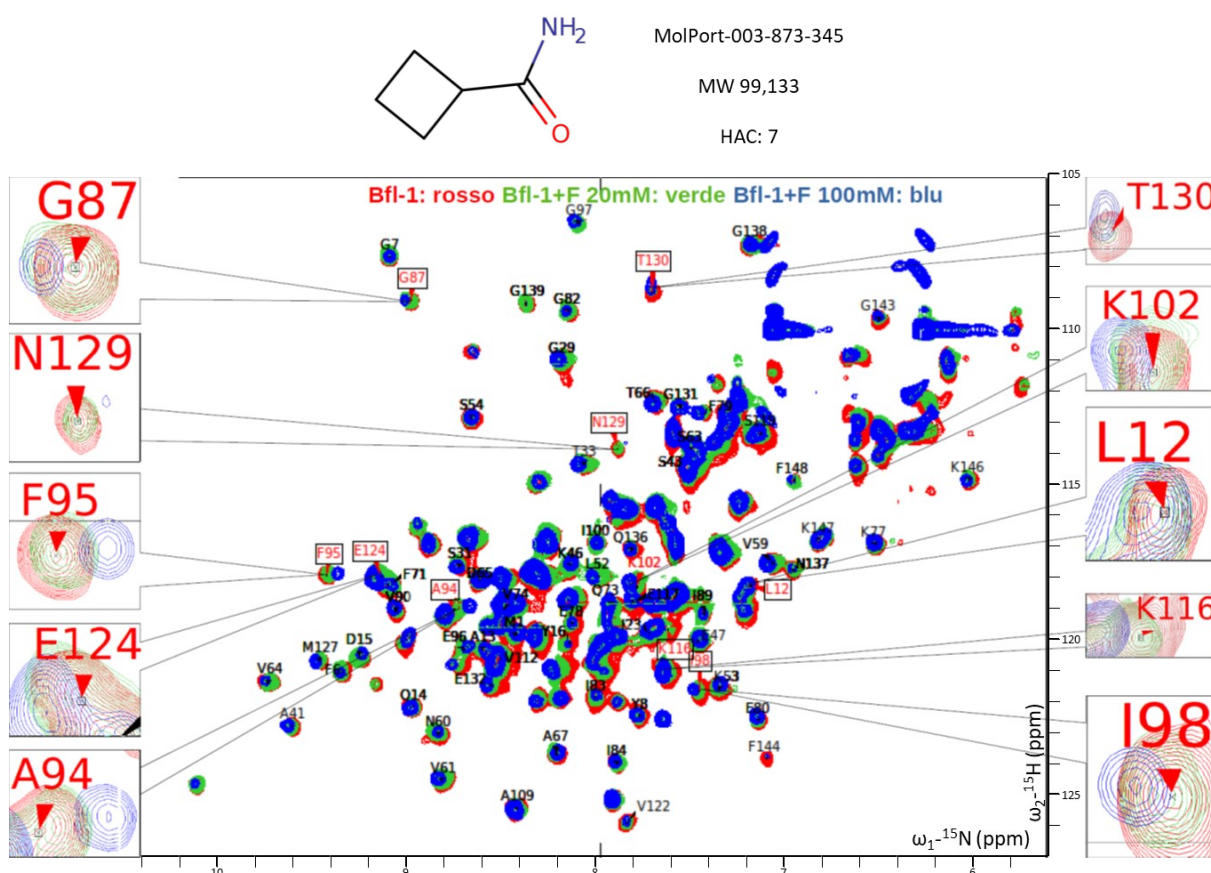


Figura 5.9 – Sovrapposizione degli spettri del bianco con solo Bfl-1, A1 a 20 mM e A1 a 100 mM, con ingrandimento degli aminoacidi che presentano maggiori shift<sup>66</sup>

Al fine di facilitare la comprensione dell'analisi effettuata e di poter apprezzare visivamente gli spostamenti che si sono riscontrati, è stato deciso di riportare lo spettro NMR (FIGURA 5.9) come sovrapposizione di tre spettri, il primo, in rosso, che rappresenta lo spettro effettuato su un campione contenente solo Bfl-1 20  $\mu$ M; il secondo (in verde) che contiene anche il frammento, in questo caso A1, ad una concentrazione di 20 mM (1:1000); ed in fine il terzo, in blu, che contiene il frammento ad una concentrazione di 100 mM (1: 5000).

È stata effettuata l'analisi tramite lo script *CSP mapping*, e gli aminoacidi che superavano il triplo della deviazione standard sono stati poi ricontrollati visivamente nello spettro e sono stati esclusi quelli che nella sovrapposizione dei tre spettri non si spostavano nella stessa direzione e/o verso.

Attraverso lo script si è ricavato anche il valore di *mean5* di entrambi gli spettri contenuti il frammento, così da poterli poi confrontare con i valori di tutta la libreria (TABELLA 5.4).

MEAN5	
A1 20 mM	0,012 ppm
A1 100 mM	0,032 ppm

Tabella 5.4 – Mean 5

Dopo aver selezionato gli amminoacidi con CSP più elevato (nella FIGURA 5.9 sono stati ingranditi ed evidenziati ai lati) si sono evidenziati sulla struttura della proteina. (FIGURA 5.10)

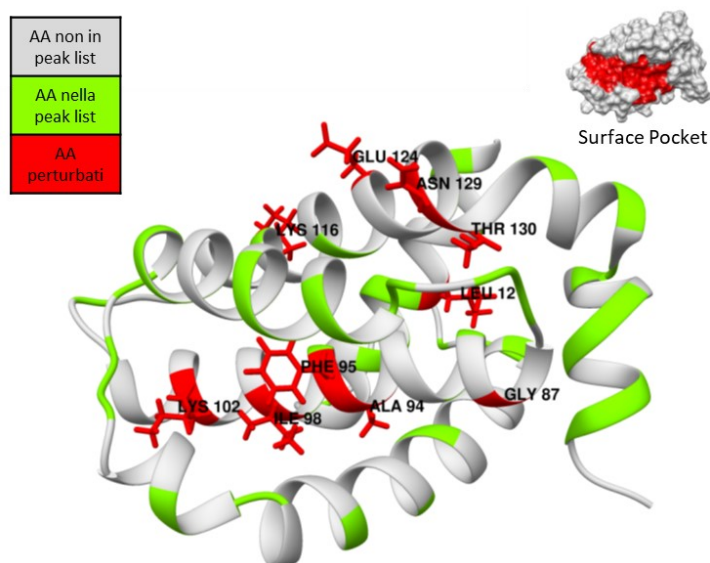


Figura 5.10 – Amminoacidi più perturbati in rosso, in verde gli aa presenti nella peak list e in grigio quelli non presenti nella peak list. In alto a destra si evidenzia la tasca idrofobica di Bfl-1<sup>21</sup>

Gli amminoacidi presenti nella *peak list* vengono rappresentati in verde, mentre in rosso si hanno quelli più perturbati. Si può notare come siano presenti degli aa perturbati anche all'esterno della tasca idrofobica, questo fenomeno è dovuto alla ridotta dimensione dei frammenti, che hanno la possibilità di interagire anche con amminoacidi più esterni alla tasca, e anche all'elevata concentrazione del ligando nella soluzione.

L'ultimo passaggio dell'analisi prevede il confronto con le *poses* computazionali, per vedere se esistono delle similitudini, tenendo comunque presente che non si stanno valutando tutti gli amminoacidi, ma solo quelli rimasti nella *peak list*.

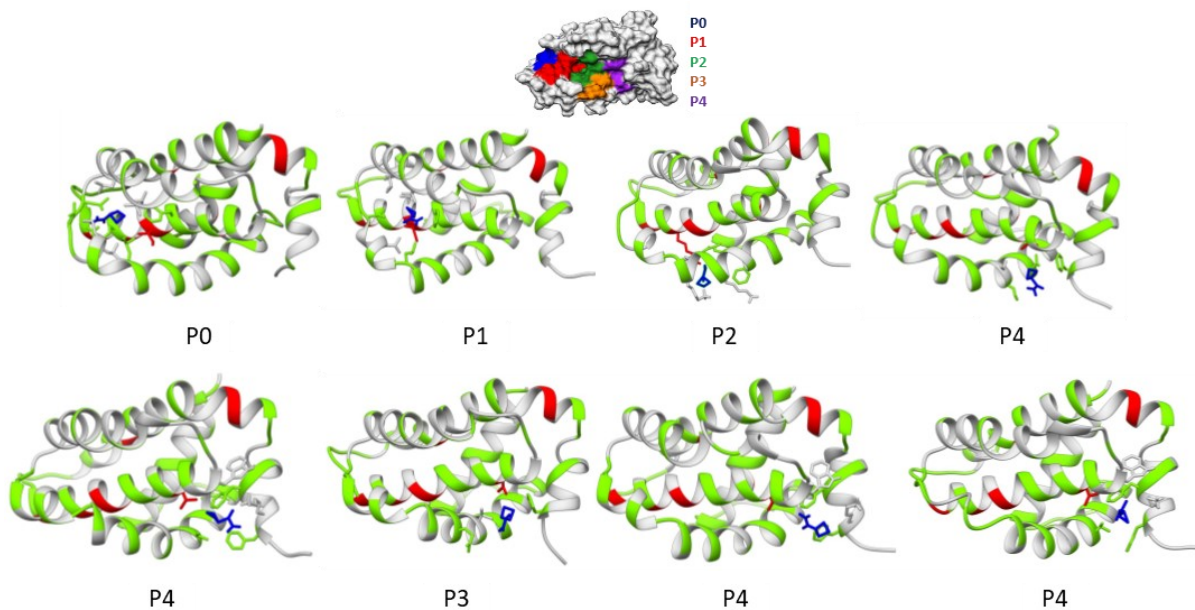


Figura 5.11 – Poses del frammento, in rosso sono evidenziati gli amminoacidi maggiormente perturbati, in verde quelli presenti nella peak list, mentre in blu viene evidenziato il frammento <sup>21</sup>

Nella FIGURA 5.11 in rosso sono stati segnati gli amminoacidi maggiormente perturbati, Si può notare che gli amminoacidi che interagiscono con il frammento (in blu), vengono rappresentati con anche la catena laterale. Delle 25 poses sono state selezionate le 8 riportate in FIGURA 5.11, sfruttando come criterio di inclusione il fatto che almeno uno degli amminoacidi coinvolti nel legame fosse anche uno degli amminoacidi perturbati. In alto è stata aggiunta come riferimento la tasca idrofobica con le relative pockets colorate in modo da poterle facilmente distinguere.

Dai dati sperimentali (FIGURA 5.10) e dal confronto con le poses (FIGURA 5.11) emerge che il sito più probabile di interazione sia a livello della P2 (Val 48, Met 75, Glu 78, Thr 91).

## 5.7.2 A2

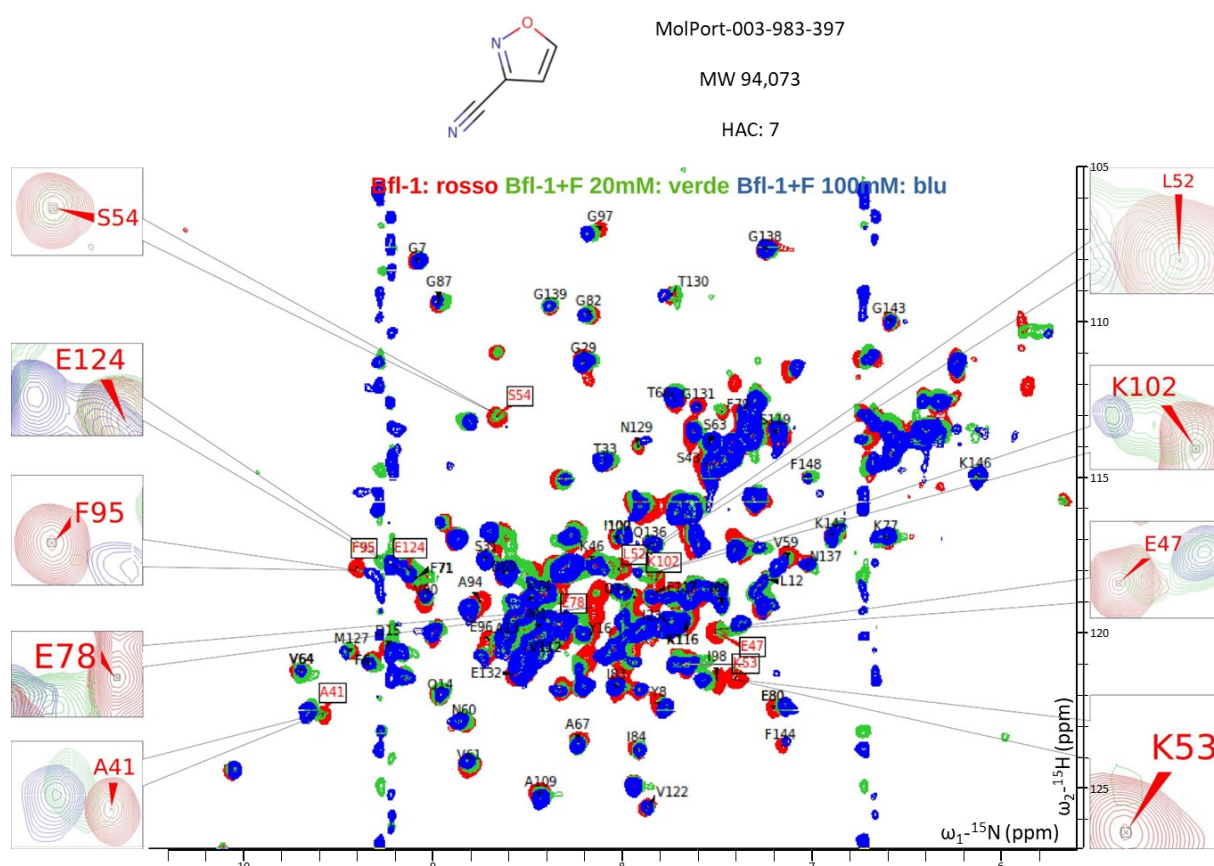


Figura 5.12- Sovrapposizione degli spettri del bianco con solo Bfl-1, A2 a 20 mM e A2 a 100 mM, con ingrandimento degli aminoacidi che presentano maggiori shift <sup>66</sup>

Lo spettro del frammento A2 è stato uno degli spettri più sorprendenti, presenta un valore di CSP (TABELLA 5.5), ad entrambe le concentrazioni, più alto di tutti gli altri frammenti e paragonabile a quello di un buon ligando preso da una libreria di frammenti tradizionale.

Il MiniFrag presenta un gruppo nitrile potenzialmente molto reattivo. Inoltre sono presenti ben due atomi di azoto all'interno della molecola. Questo, a causa dell'elevata concentrazione nel campione, provoca nello spettro (FIGURA 5.12) quell'effetto di rumore rappresentato dalle due bande di picchi molto vicini visibili a circa 6,8 e 9,2 ppm.

MEAN5	
A2 20 mM	0,063 ppm
A2 100 mM	0,072 ppm

Tabella 5.5 – Mean 5

Gli aminoacidi più perturbati sono stati riportati nella FIGURA 5.13, si può vedere che si trovano quasi tutti nella tasca idrofobica, indice di un legame piuttosto specifico, inoltre



sono molto perturbati gli aa nell'intorno della cisteina 55 (amminoacido molto reattivo ma non presente nella *peak list*).

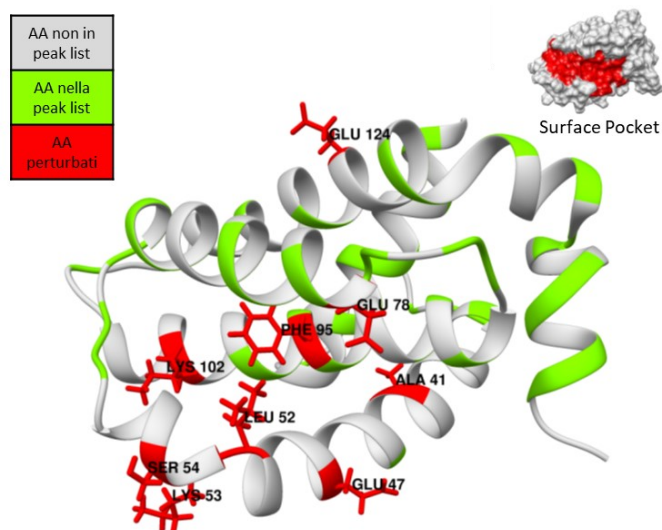


Figura 5.13 - Amminoacidi più perturbati in rosso, in verde gli aa presenti nella *peak list* e in grigio quelli non presenti nella *peak list*. In alto a destra si evidenzia la tasca idrofobica di Bfl-1 <sup>21</sup>

Andando poi ad analizzare le *poses* si può vedere che sono presenti sei possibili riscontri. In particolare, la *pose* che prevede l'occupazione di P2 sembra essere la più probabile. Tuttavia, osservando gli amminoacidi perturbati e confrontandoli con quelli costituenti le *pocket* (FIGURA 2.9), si può supporre che il frammento si posiziona più probabilmente all'interno di P1.

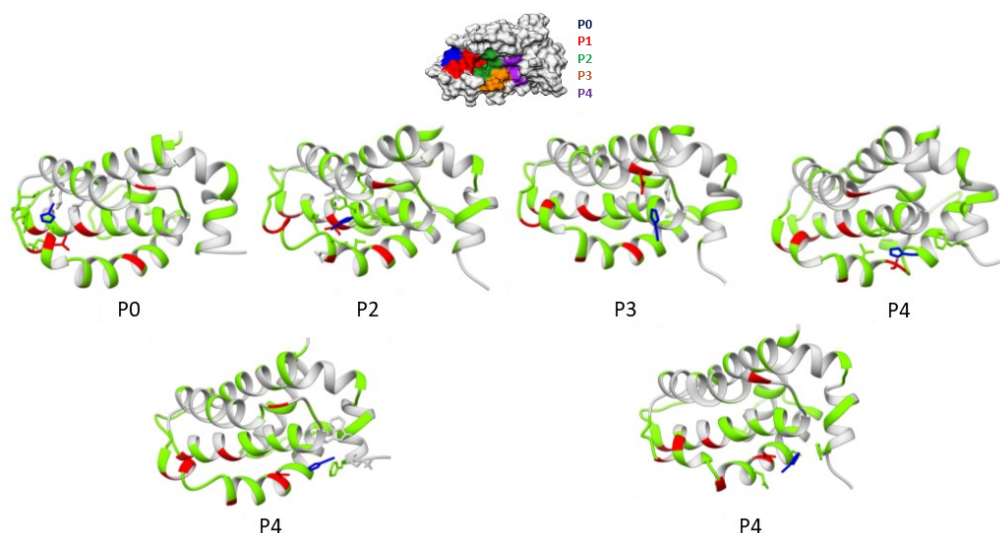


Figura 5.14 - *Poses* del frammento, in rosso sono evidenziati gli amminoacidi maggiormente perturbati, in verde quelli presenti nella *peak list*, mentre in blu viene evidenziato il frammento <sup>21</sup>



### 5.7.3 B6

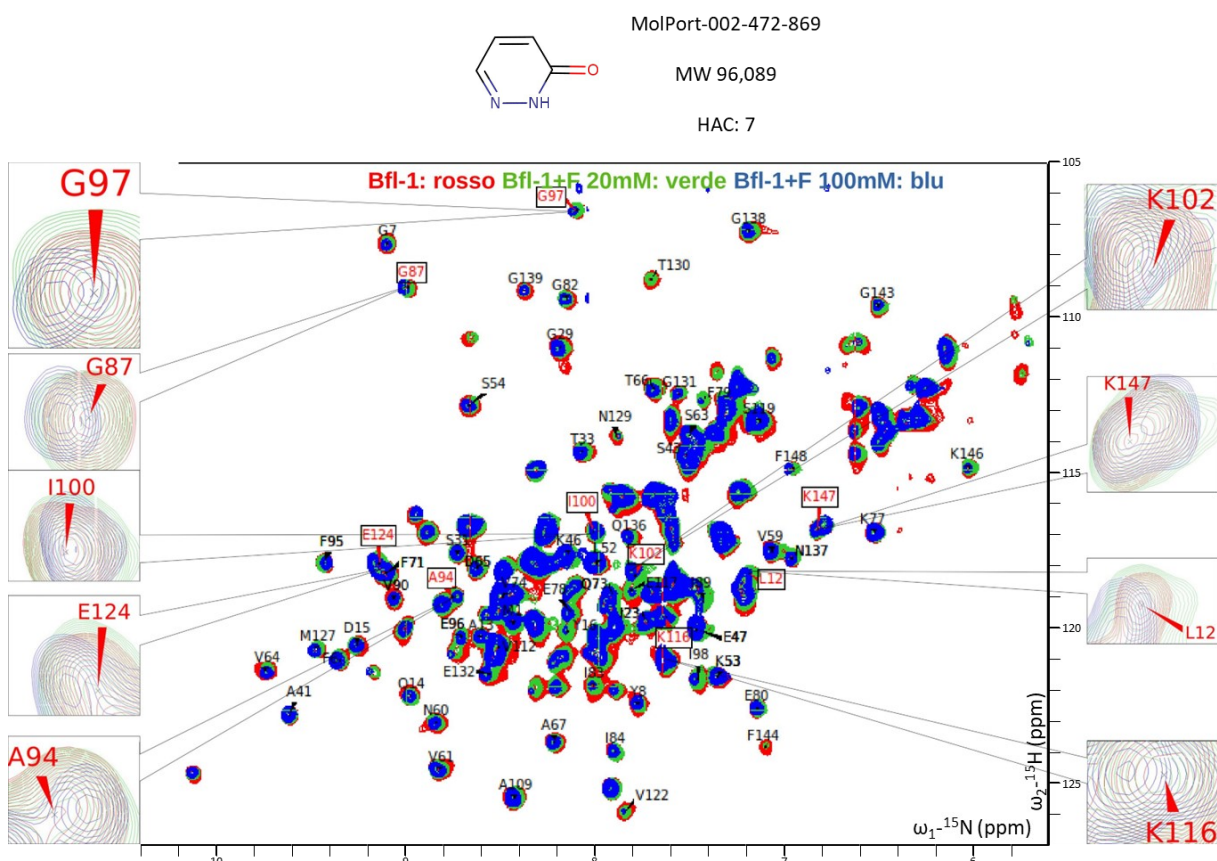


Figura 5.15 - Sovrapposizione degli spettri del bianco con solo Bfl-1, B6 a 20 mM e B6 a 100 mM, con ingrandimento degli aminoacidi che presentano maggiori shift <sup>66</sup>

Nella FIGURA 5.15 è presentato lo spettro del frammento B6, come si può vedere dagli ingrandimenti degli aminoacidi più perturbati, che non presentano un grande spostamento, e dal *mean5* (TABELLA 5.6), che rispetto ai precedenti due casi risulta essere più esiguo, si tratta di un frammento che non sembra avere un legame molto forte con la proteina.

MEAN5	
B6 20 mM	0,007 ppm
B6 100 mM	0,023 ppm

Tabella 5.6 – Mean 5

Gli aminoacidi più perturbati (FIGURA 5.16) sono molto dispersi nella molecola, sintomo di interazione debole e apparentemente aspecifica.

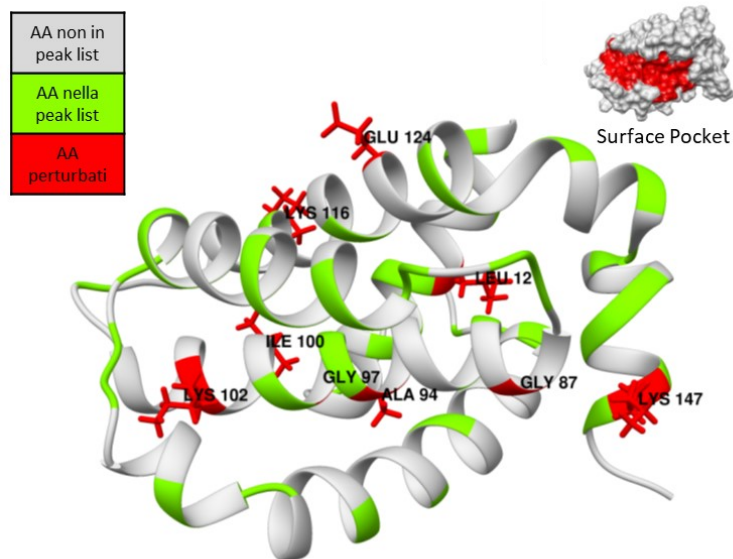


Figura 5.16 - Amminoacidi più perturbati in rosso, in verde gli aa presenti nella peak list e in grigio quelli non presenti nella peak list. In alto a destra si evidenzia la tasca idrofobica di Bfl-1<sup>21</sup>

Nel momento in cui si sono confrontati i dati con le *poses*, si sono trovati due riscontri piuttosto interessanti (FIGURA 5.17), ovvero sembra poterci essere un'interazione con la *pocket 4*.

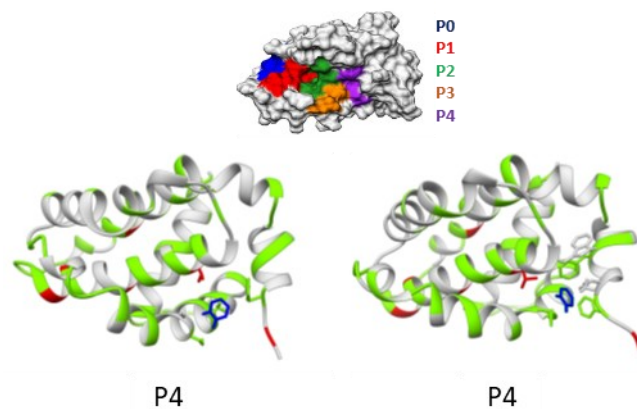


Figura 5.17 - Poses del frammento, in rosso sono evidenziati gli amminoacidi maggiormente perturbati, in verde quelli presenti nella peak list, mentre in blu viene evidenziato il frammento<sup>21</sup>

## 5.7.4 C2

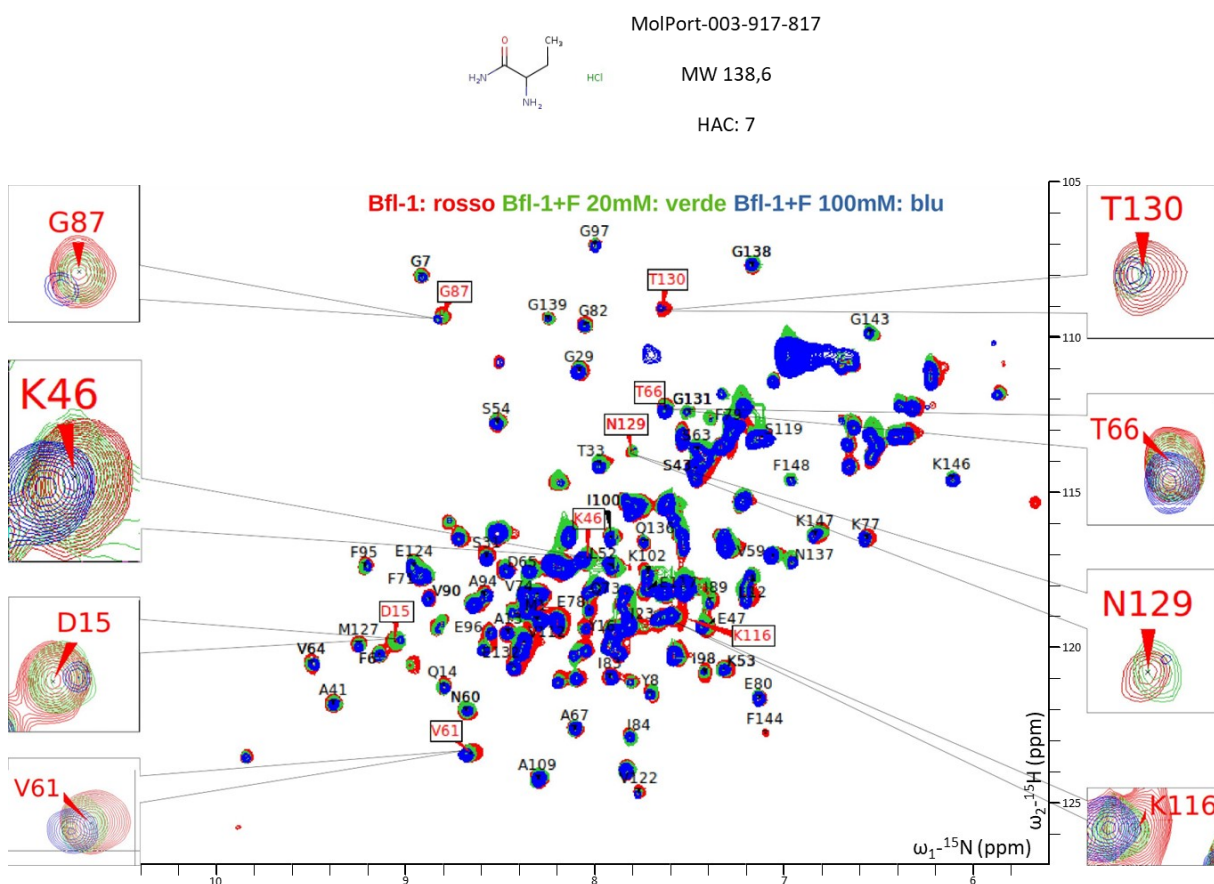


Figura 5.18 - Sovrapposizione degli spettri del bianco con solo Bfl-1, C2 a 20 mM e C2 a 100 mM, con ingrandimento degli aminoacidi che presentano maggiori shift <sup>66</sup>

Dall'analisi dello spettro (FIGURA 5.18) e dai valori di Mean5 (TABELLA 5.7), si può notare che questo frammento non presenta grandi valori di CSP, gli shift dei picchi sono molto esigui e apparentemente non sembrano esserci forti *binding*.

MEAN5	
C2 20 mM	0,011 ppm
C2 100 mM	0,020 ppm

Tabella 5.7 – Mean 5

Dall'analisi su Chimera emerge che la distribuzione degli aminoacidi perturbati è molto sparpagliata sulla struttura della proteina. Una nota interessante potrebbe essere costituita dagli aminoacidi tirosina 66 e valina 61 che si trovano nella tasca P0.

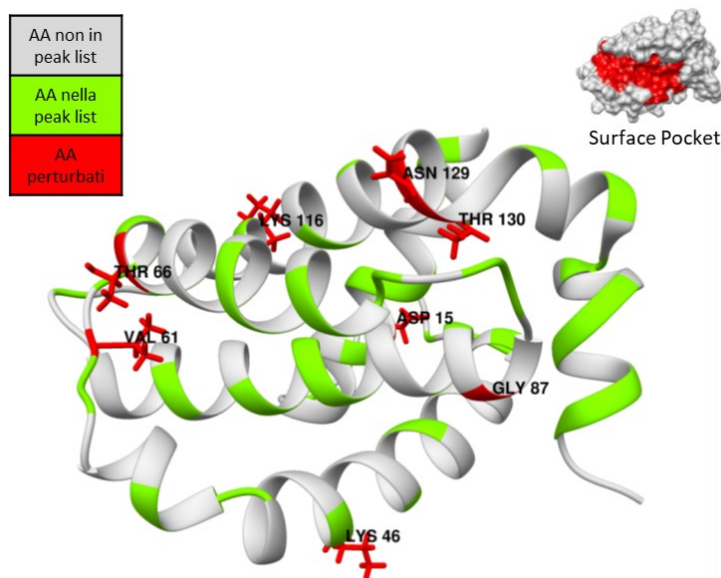


Figura 5.19 - Amminoacidi più perturbati in rosso, in verde gli aa presenti nella peak list e in grigio quelli non presenti nella peak list. In alto a destra si evidenzia la tasca idrofobica di Bfl-1 <sup>21</sup>

Dalla comparazione con le poses (FIGURA 5.20) si può notare che effettivamente tre delle quattro poses mostrano il frammento legarsi a P0; in particolare la terza sembra essere la più convincente, in quanto il MiniFrag si posiziona all'interno della *pocket*. La quarta pose è stata inserita perché presentava un amminoacido perturbato nella zona di predizione del legame. Tuttavia come si può vedere dall'immagine il frammento si trova in una zona molto esterna alla proteina, e quindi alla tasca idrofobica. Questo fa presupporre più ad un legame aspecifico che ad un legame interessante ai fini della ricerca di un inibitore proteico.

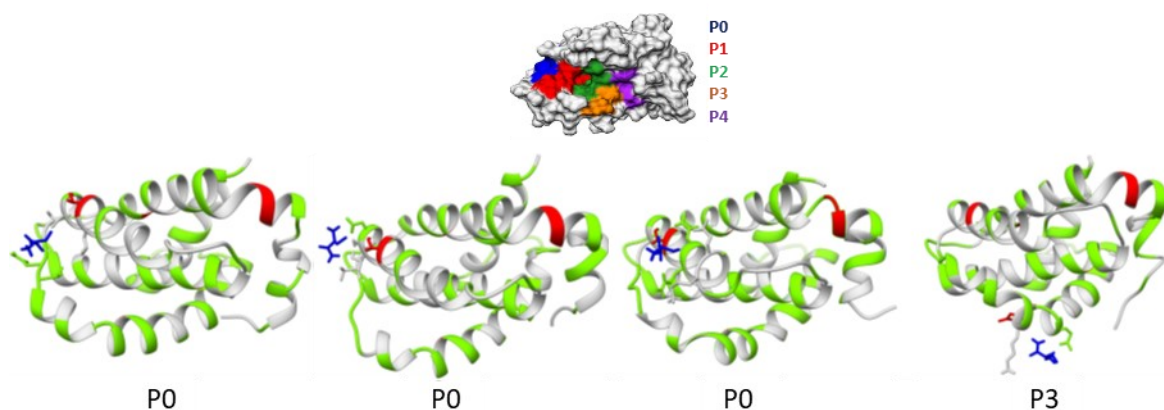


Figura 5.20 - Poses del frammento, in rosso sono evidenziati gli amminoacidi maggiormente perturbati, in verde quelli presenti nella peak list, mentre in blu viene evidenziato il frammento <sup>21</sup>

## 5.7.5 C3

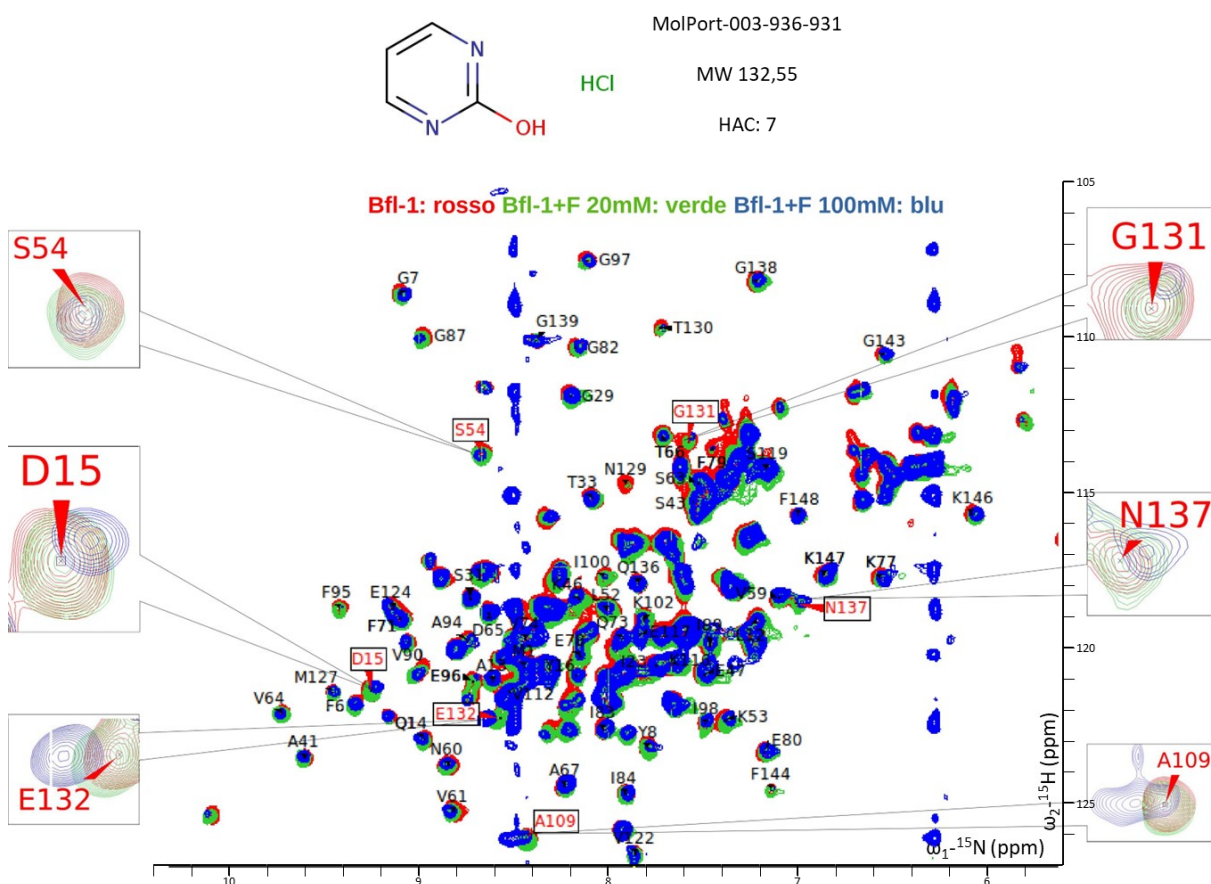


Figura 5.21 - Sovrapposizione degli spettri del bianco con solo Bfl-1, C3 a 20 mM e C3 a 100 mM, con ingrandimento degli aminoacidi che presentano maggiori shift<sup>66</sup>

Dall'analisi dello spettro (FIGURA 5.21) e dall'applicazione dello script, si sono selezionati gli aminoacidi che presentano *shift* maggiori rispetto al triplo della deviazione standard. Come è possibile notare dagli ingrandimenti gli aminoacidi maggiormente perturbati e che non rientrano nei criteri di esclusione imposti per l'analisi, sono molto pochi.

MEANS	
C3 20 mM	0,013 ppm
C3 100 mM	0,031 ppm

Tabella 5.8 – Mean 5

Nonostante il valore di *mean5* (TABELLA 5.8) per lo spettro contenente il frammento a 100 mM sia relativamente elevato (vicino a quello di A1), gli aminoacidi perturbati, visualizzati nella FIGURA 5.22, sono in punti molto lontani della proteina. Questa particolarità fa presumere che non sia effettivamente presente alcun legame specifico, ma una serie di legami aspecifici non molto interessanti per ulteriori ricerche.



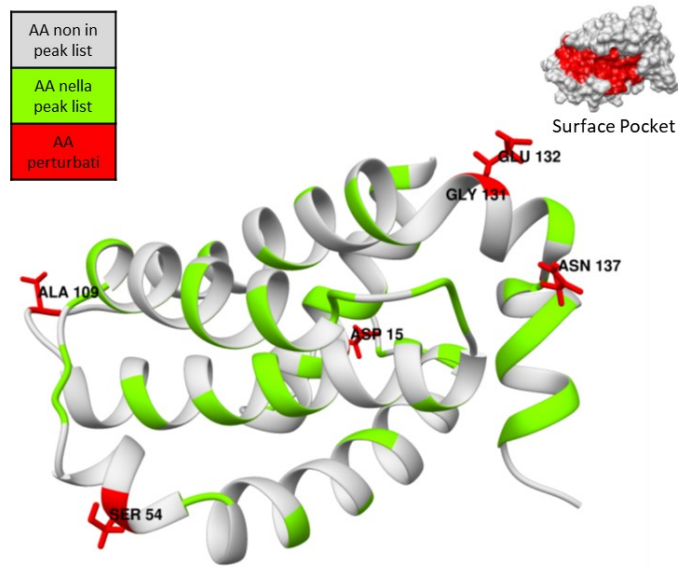


Figura 5.22 - Amminoacidi più perturbati in rosso, in verde gli aa presenti nella peak list e in grigio quelli non presenti nella peak list. In alto a destra si evidenzia la tasca idrofobica di Bfl-1<sup>21</sup>

Nonostante i dati sperimentali suggeriscano la mancanza di un legame specifico, si è trovato comunque un riscontro con una *pose* (FIGURA 5.23). Questo riscontro confrontato con i dati sperimentali però non sembra essere particolarmente rilevante.

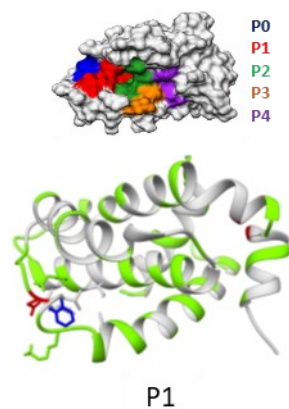


Figura 5.23 - Poses del frammento, in rosso sono evidenziati gli amminoacidi maggiormente perturbati, in verde quelli presenti nella peak list, mentre in blu viene evidenziato il frammento<sup>21</sup>

## 5.7.6 C4

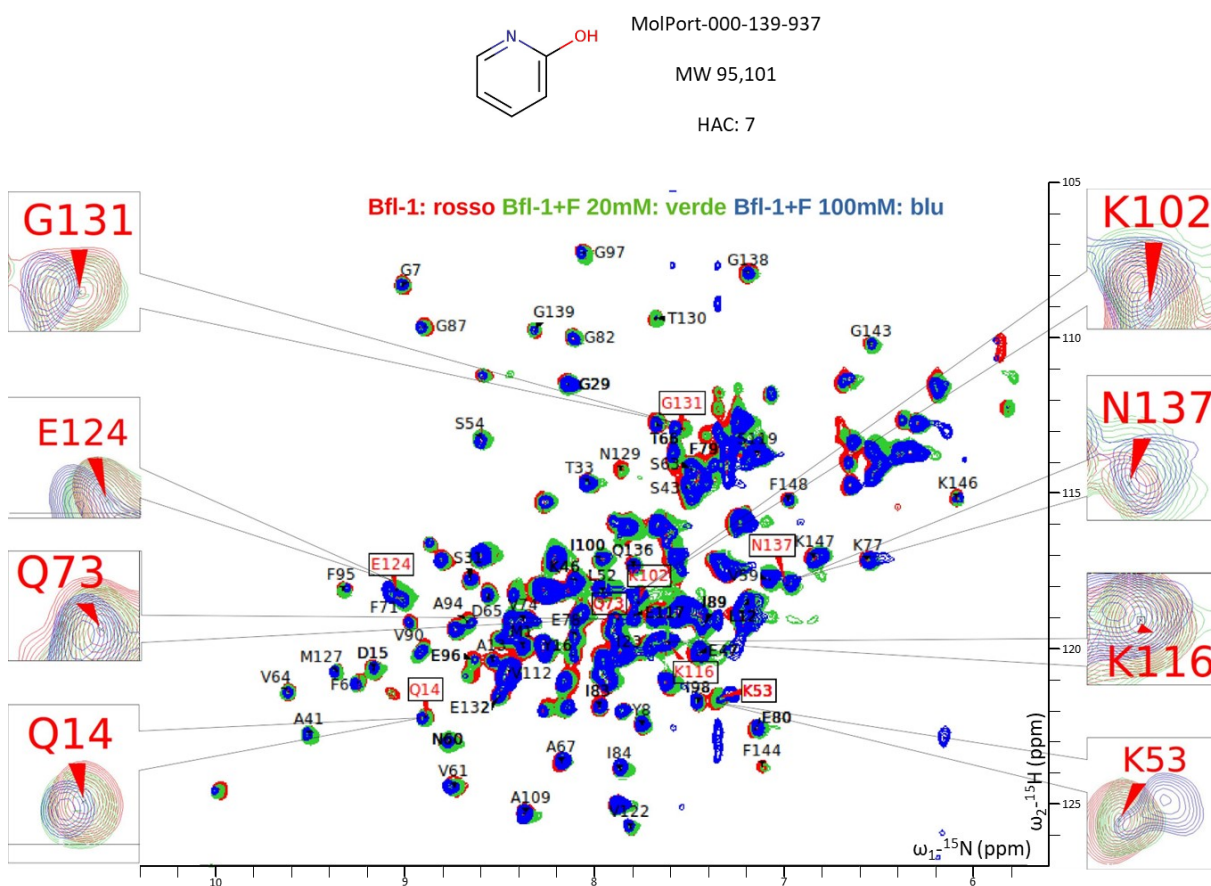


Figura 5.24 - Sovrapposizione degli spettri del bianco con solo Bfl-1, C4 a 20 mM e C4 a 100 mM, con ingrandimento degli aminoacidi che presentano maggiori shift <sup>66</sup>

I valori di *mean5* (TABELLA 5.9) per questi due spettri sono molto simili a quelli del MiniFrag precedente (C3), infatti come è possibile notare dagli ingrandimenti dello spettro (FIGURA 5.24), gli *shift* sono piuttosto esigui anche a grandi concentrazioni. Una differenza che si ha in questo caso è la presenza di più picchi con un valore di CSP superiore al triplo della deviazione standard.

MEAN5	
C4 20 mM	0,012 ppm
C4 100 mM	0,023 ppm

Tabella 5.9 – Mean 5

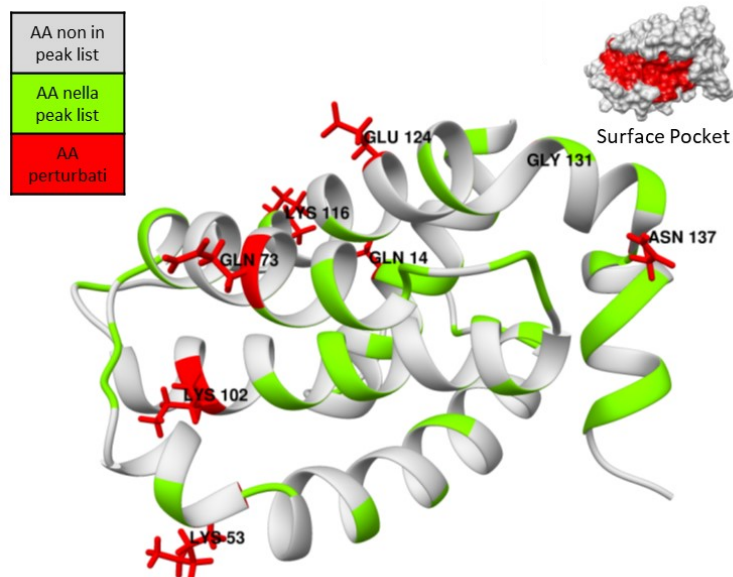


Figura 5.25 - Amminoacidi più perturbati in rosso, in verde gli aa presenti nella peak list e in grigio quelli non presenti nella peak list. In alto a destra si evidenzia la tasca idrofobica di Bfl-1<sup>21</sup>

Dalla FIGURA 5.25 si può notare che gli amminoacidi più perturbati sono abbastanza distanziati gli uni dagli altri, rispetto alla tasca idrofobica i più interessanti sono glicina 73 e lisina 102, che indicano un potenziale legame in P1. Una particolarità rilevante di questo MiniFrag risulta essere la zona perturbata al di sopra della tasca idrofobica, rappresentata da glutammina 14, lisina 116, acido glutammico 124 e glutammina 73, questi amminoacidi sembrano quasi definire una nuova *pocket*.

Nel confronto con le poses computazionali non si sono tuttavia avuti riscontri.



## 5.7.7 C5

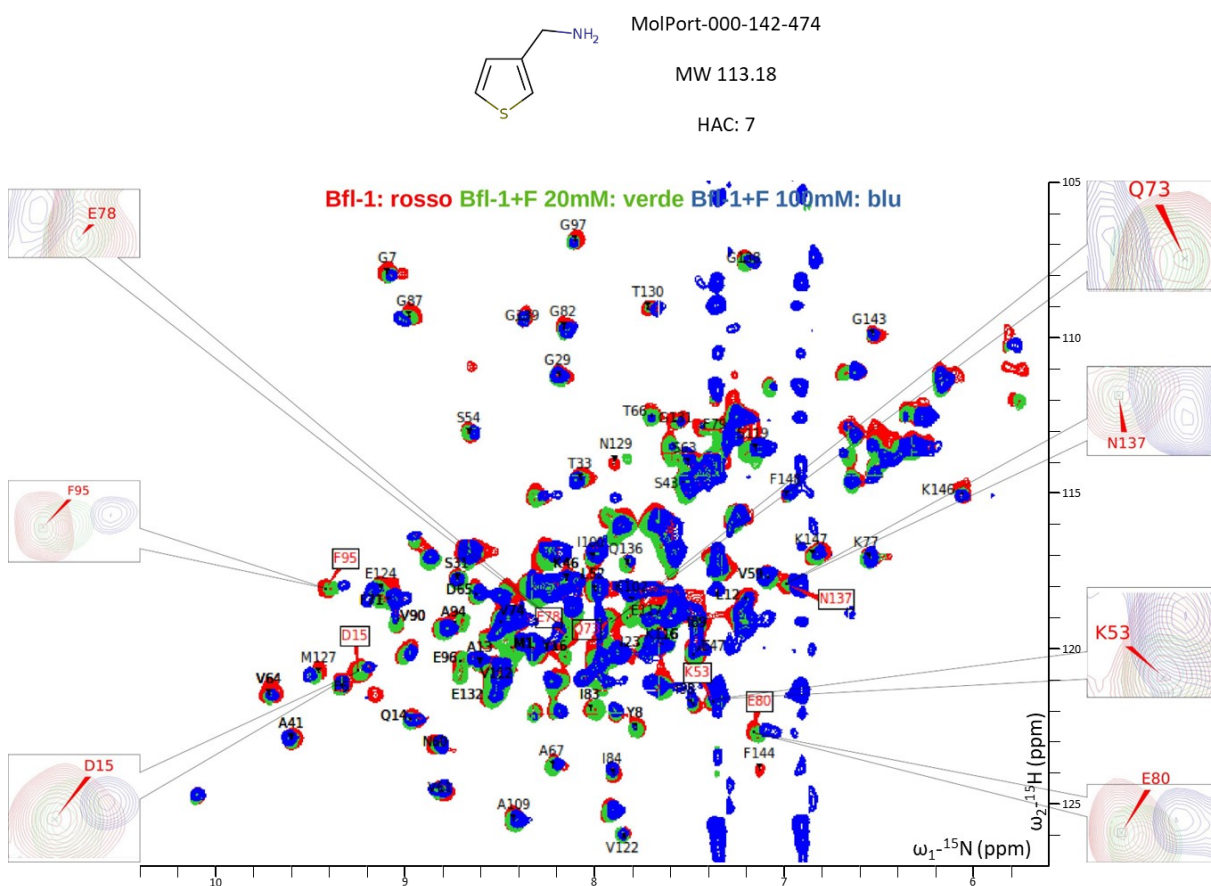


Figura 5.26 - Sovrapposizione degli spettri del bianco con solo Bfl-1, C5 a 20 mM e C5 a 100 mM, con ingrandimento degli aminoacidi che presentano maggiori shift <sup>66</sup>

Il MiniFrag C5 è uno dei più promettenti, con un valore di *mean5* (TABELLA 5.10) molto simile ad A2, seppur minore. Dallo spettro (FIGURA 5.26) si può notare che gli *shift* alla concentrazione di 20 mM sono relativamente esigui, rispetto ad A2. Questo dato è molto più aspettato rispetto all'altro frammento, ed in linea con ciò che ci si aspetterebbe da frammenti con affinità così bassa. Quando si va a valutare lo spettro a 100 mM invece gli *shift* si fanno più evidenti e ampi, sempre in linea con ciò che ci si aspetterebbe.

MEAN5	
C5 20 mM	0,012 ppm
C5 100 mM	0,047 ppm

Tabella 5.10 – Mean 5

Quando si vanno ad evidenziare gli aa maggiormente interessati sulla struttura (FIGURA 5.27), si delinea un sito di interazione tra la tasca P2 e P3, formato dagli aa: glutamina 80, 78 e fenilalanina 95.

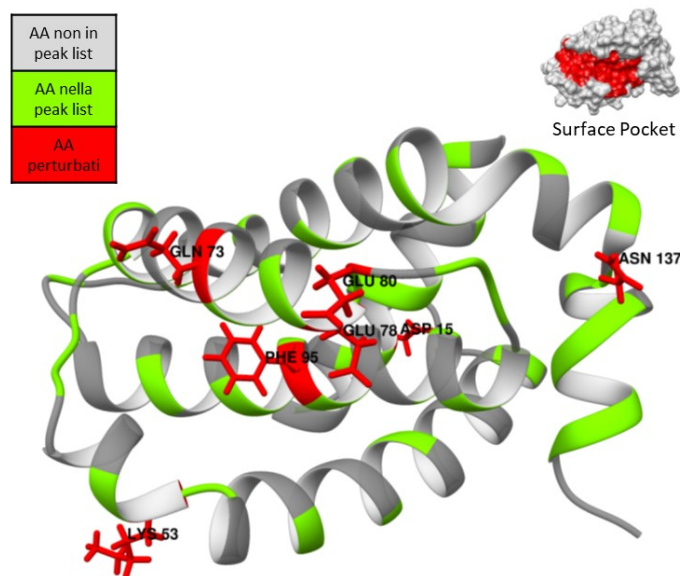


Figura 5.27 - Amminoacidi più perturbati in rosso, in verde gli aa presenti nella peak list e in grigio quelli non presenti nella peak list. In alto a destra si evidenzia la tasca idrofobica di Bfl-1 <sup>21</sup>

Nella FIGURA 5.28 vengono rappresentate le *poses* compatibili con i dati sperimentali. In questo caso le prime due sono abbastanza in accordo con i dati ottenuti; mentre le seconde due non prevedono un'interazione con la tasca idrofobica, ma piuttosto con una parte superficiale della proteina. Le seconde due poses presentano un'interazione all'infuori dei parametri dati al *software*, ovvero di cercare delle possibili conformazioni ligando-proteina all'interno delle 5 principali *pocket*; questo dato è dovuto al fatto che i programmi computazionali simulano ciò che accade nella soluzione in cui si trovano la proteina e il ligando: quindi con la proteina che subisce alcune piccole variazioni conformazionali e il frammento che si avvicina e interagisce con la stessa. Questo comporta che si possano trovare, tra le 25 *poses* consegnate, alcune che evidenziano un'interazione non convenzionale.

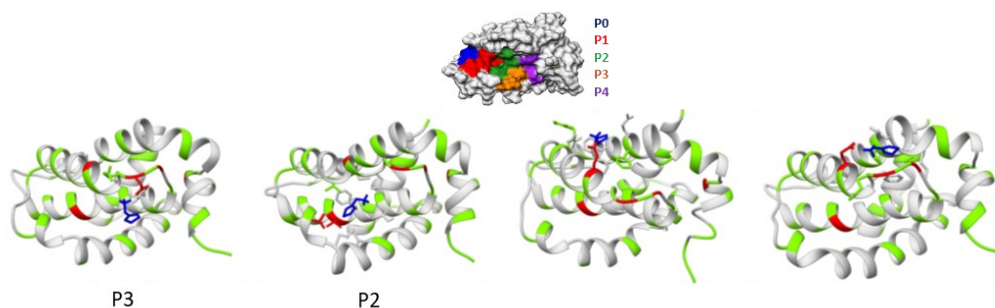


Figura 5.28 - Poses del frammento, in rosso sono evidenziati gli amminoacidi maggiormente perturbati, in verde quelli presenti nella peak list, mentre in blu viene evidenziato il frammento <sup>21</sup>

## 5.7.8 C6

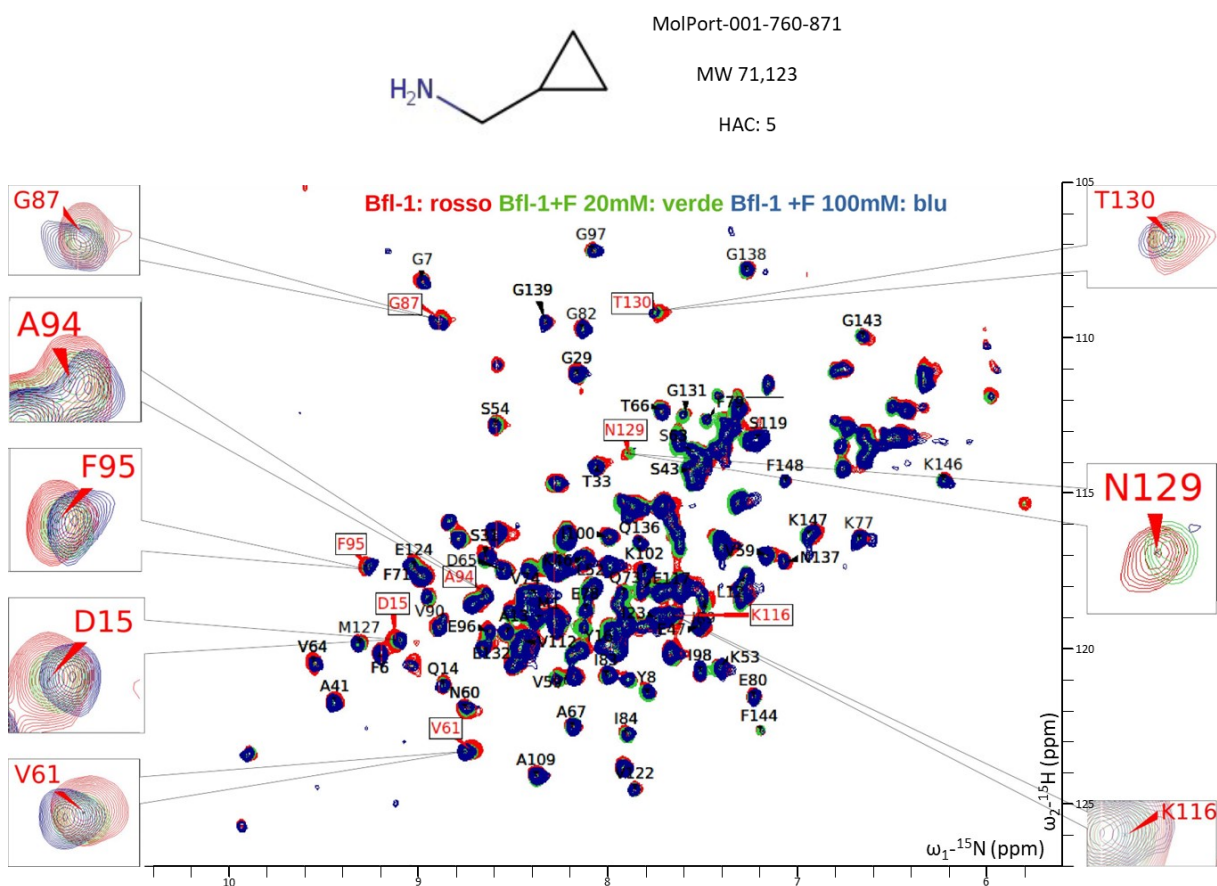


Figura 5.29 - Sovrapposizione degli spettri del bianco con solo Bfl-1, C6 a 20 mM e C6 a 100 mM, con ingrandimento degli aminoacidi che presentano maggiori shift <sup>66</sup>

Analogamente al MiniFrag B6, il CSP per il MiniFrag C6 è estremamente basso (TABELLA 5.11), il che può essere notato anche dallo spettro, in cui gli *shift* ingranditi ai lati risultano essere minimi anche ad alte concentrazioni (FIGURA 5.29).

MEAN5	
C6 20 mM	0,009 ppm
C6 100 mM	0,018 ppm

Tabella 5.11 – Mean 5

Nonostante il basso valore di CSP è stata effettuata l'analisi tramite script del triplo della deviazione standard, e si sono selezionati gli aminoacidi più perturbati, riportati in FIGURA 5.30.

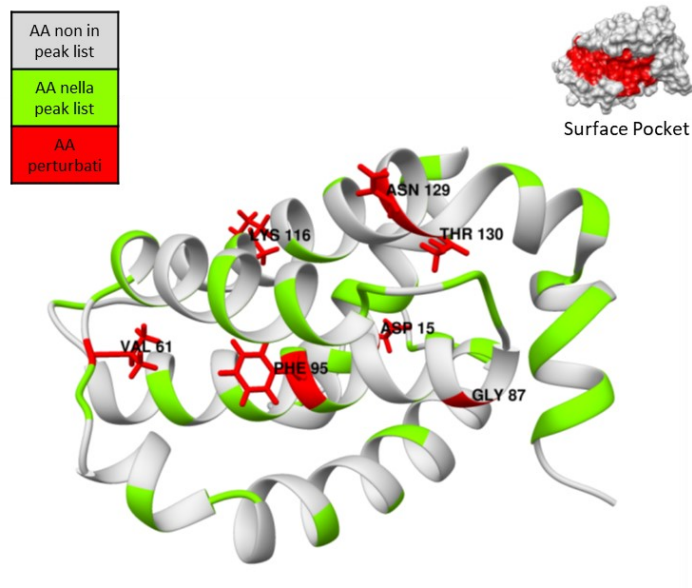


Figura 5.30 - Amminoacidi più perturbati in rosso, in verde gli aa presenti nella peak list e in grigio quelli non presenti nella peak list. In alto a destra si evidenzia la tasca idrofobica di Bfl-1 <sup>21</sup>

Nonostante tra gli amminoacidi perturbati ci fossero valina 61 e fenilalanina 95, che fanno parte del *pocket 1*, non ci sono stati riscontri con nessuna delle 25 poses.

## 5.7.9 C7

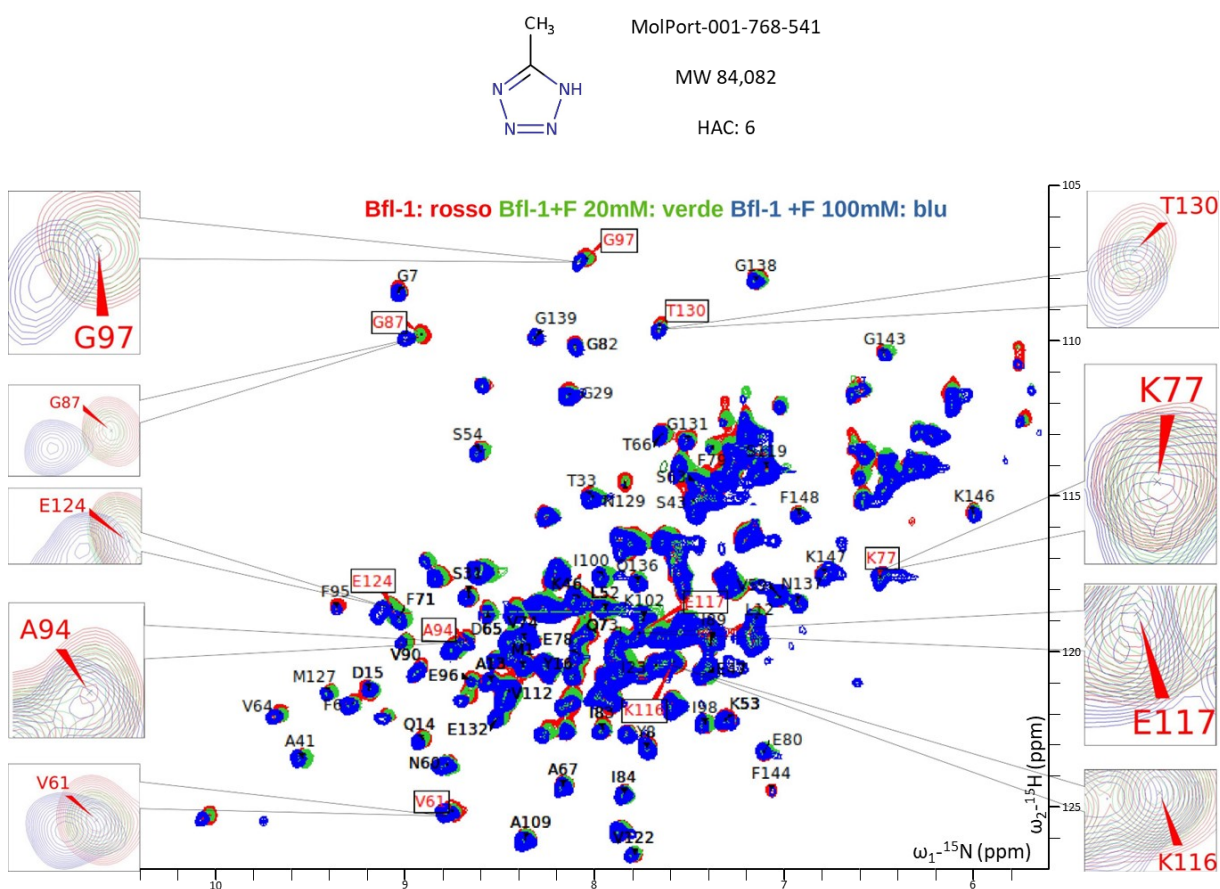


Figura 5.31 - Sovrapposizione degli spettri del bianco con solo Bfl-1, C7 a 20 mM e C7 a 100 mM, con ingrandimento degli amminoacidi che presentano maggiori shift <sup>66</sup>

Il frammento C7, come si può vedere dagli ingrandimenti dello spettro (FIGURA 5.31) non presenta quasi alcuno shift a basse concentrazioni (20 mM), mentre ad alte concentrazioni (100 mM) si possono apprezzare degli spostamenti piuttosto interessanti. Questo aspetto lo si può notare anche dai valori di CSP (TABELLA 5.12). Questo è indice di un possibile legame a bassissima affinità, apprezzabile solamente ad alte concentrazioni.

MEANS	
C7 20 mM	0,009 ppm
C7 100 mM	0,031 ppm

Tabella 5.12 – Mean 5

Nella FIGURA 5.32 si può visualizzare la presenza di tre amminoacidi vicini (lisina 77, glicina 97 e alanina 94), facenti parte delle *pocket* centrali P3/P4, che lasciano presumere un legame in quella zona della tasca idrofobica.



È presente anche un'altra zona relativamente perturbata (glutammato 117, 124 e lisina 116), che però si trovano in una zona superficiale della proteina e lontana dalla tasca idrofobica, questo rende la zona relativamente interessante ai fini della ricerca di un possibile binding. È possibile che siano amminoacidi facenti parte di una nuova *pocket*, lontana da quella principale, la quale richiede uno studio più approfondito nella definizione.

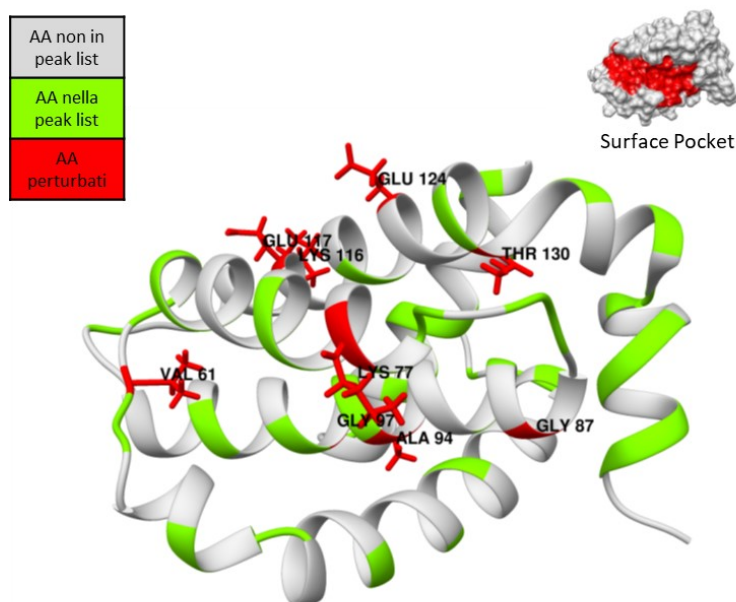


Figura 5.32 - Amminoacidi più perturbati in rosso, in verde gli aa presenti nella peak list e in grigio quelli non presenti nella peak list. In alto a destra si evidenzia la tasca idrofobica di Bfl-1 <sup>21</sup>

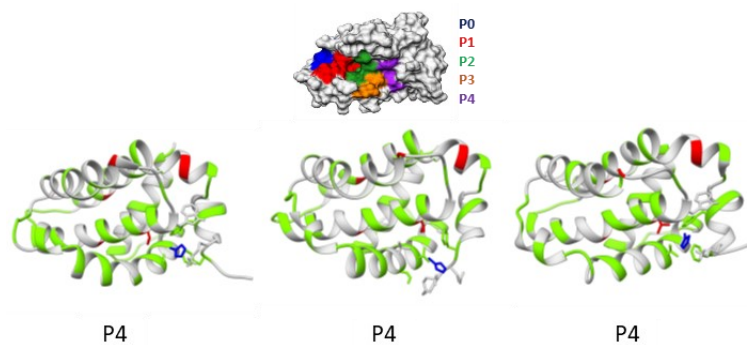
Nel confronto con le *poses* (FIGURA 5.33) si nota una corrispondenza con i dati sperimentali solo in P4, tuttavia i dati sperimentali mostrano che il legame più probabile non sia in quella specifica *pocket*, ma piuttosto in P3.

La discrepanza tra dati sperimentali e computazionali risulta estremamente interessante, è un dato che può essere integrato ai protocolli computazionali per migliorarne la precisione nella predizione.

Questo dato fa emergere uno degli aspetti più interessanti nell'utilizzo dei MiniFrag ai fini della FBDD, ossia la possibilità di scoprire nuovi siti di interazione difficilmente sondabili utilizzando i tradizionali frammenti.

La nuova possibile *pocket* che è stata evidenziata non era mai stata vista o presa in considerazione nella ricerca in quanto, frammenti di dimensioni maggiori e con affinità maggiore per la tasca idrofobica, non sono mai andati a posizionarsi in quel punto più

esterno alla proteina; mentre utilizzando i MiniFrag, aventi dimensioni minori e minor affinità, si è potuta apprezzare anche questa interazione.



*Figura 5.33 - Poses del frammento, in rosso sono evidenziati gli amminoacidi maggiormente perturbati, in verde quelli presenti nella peak list, mentre in blu viene evidenziato il frammento <sup>21</sup>*

## 5.8 Copresenze

Dai risultati ottenuti dal primo set di esperimenti emergono due MiniFrag in particolare che presentano un alto valore di CSP (TABELLA 5.13): A2 e C5. Inoltre, è emerso anche che sono presenti delle corrispondenze tra dati sperimentali e *poses*. Oltre a ciò, è anche emerso che, come pensato nella scelta delle molecole, i MiniFrag tendono ad avere una certa selettività per *pocket* diverse.

MEAN5			
MiniFrag	ppm	MiniFrag	ppm
A1 20 mM	0,012	A1 100 mM	0,032
A2 20 mM	0,063	A2 100 mM	0,072
B6 20 mM	0,007	B6 100 mM	0,023
C2 20 mM	0,011	C2 100 mM	0,020
C3 20 mM	0,013	C3 100 mM	0,021
C4 20 mM	0,012	C4 100 mM	0,023
C5 20 mM	0,022	C5 100 mM	0,047
C6 20 mM	0,009	C6 100 mM	0,018
C7 20 mM	0,009	C7 100 mM	0,031

Tabella 5.13 – Mean 5 di tutti i MiniFrag a confronto

Si è deciso, visti dei risultati così promettenti, di procedere con l'analisi. Per fare ciò si sono aggiunti due frammenti diversi nello stesso campione contenente la proteina, ovvero si è deciso di andare a valutare una copresenza di due o più frammenti nello stesso campione.

Per applicare la tecnica è stato necessario effettuare un bianco contenente la proteina con uno o due frammenti, in questo modo non si andavano a considerare i CSP dovuti ai primi frammenti; dopodiché si è avanzati effettuando un esperimento con la proteina e i due o tre frammenti contemporaneamente.

Per questo tipo di analisi si è scelto di utilizzare i MiniFrag A2 e C5, come bianco, questo perché presentano un'affinità estremamente maggiore rispetto agli altri, e ciò permette, in linea teorica, che non vengano spiazzati dal loro sito di legame se nel campione è presente un altro frammento. Un altro motivo per il quale si sono scelti questi due frammenti come bianco è che, provocando degli *shift* maggiori rispetto agli altri MiniFrag, si rischiava di incorrere in falsi positivi.

In questo caso si è deciso di porre A2 ad una concentrazione di 10 mM all'interno del campione, questo perché dagli esperimenti precedenti, si è visto che gli *shift* provocati da questo frammento erano già elevati ad una concentrazione di 20 mM, addirittura maggiori



rispetto al *mean5* degli altri MiniFrag a 100 mM. Gli altri frammenti invece si sono posti a 100 mM.

L'analisi dei dati si è svolta in maniera leggermente differente rispetto agli esperimenti precedenti, ovvero dopo aver creato le *peak list*, del bianco e del frammento da analizzare, si è calcolato il triplo della deviazione standard per selezionare i picchi con i migliori spostamenti, ma non avendo un terzo spettro di confronto non si è applicata la regola che lo *shift* dovesse essere nella stessa direzione e verso. Il valore di *mean5* è stato calcolato, ma non presenta una valenza pragmatica nel confronto con altri MiniFrag, questo perché dato uno stesso bianco (contenente uno o due frammenti) si è analizzato un solo frammento. Un'altra differenza rispetto all'analisi precedente è che non sono presenti *poses* della nuova situazione analizzata; perciò, non è stato fatto un confronto.

L'interesse di analizzare le copresenze è stato quello di andare a comprendere se, aggiungendo un ligando, che si è visto interagire con una *pocket* diversa, si vedesse ancora l'iterazione del secondo per la stessa *pocket*, o se i due non possano interagire con la proteina contemporaneamente.

## 5.8.1 Copresenza A2+B6

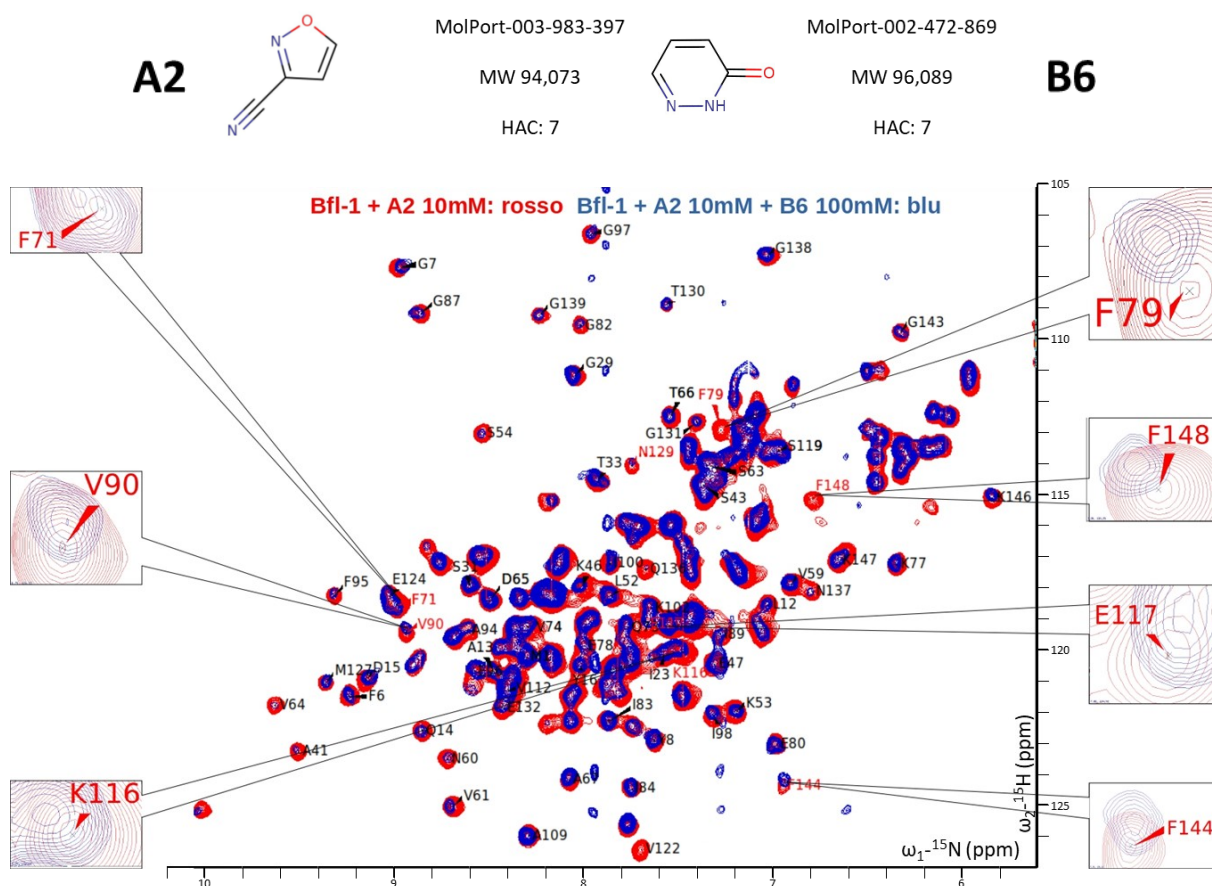


Figura 5.34 - Sovrapposizione degli spettri del bianco (Bfl-1 + A2 10 mM), e di B6 a 100 mM, con ingrandimento degli aminoacidi che presentano maggiori shift <sup>66</sup>

La copresenza di questi due MiniFrag è stata sperimentata perché A2 è un ligando molto affine per P1, mentre B6 interagisce con P4; perciò, essendo *pockets* lontane si è pensato essere una valida combinazione per la valutazione della copresenza.

Dallo spettro (FIGURA 5.34) emergono degli *shift* di buona intensità, infatti se si vanno poi a riportare nella struttura proteica (FIGURA 5.35), si nota che ricompare il sito di interazione su P4. Questo fa pensare che queste due molecole si leghino contemporaneamente all'interno della tasca idrofobica.

Inoltre, è molto interessante notare che si ripresenta, come successo per C4 e C7, un sito di interazione posto al di fuori della tasca idrofobica.

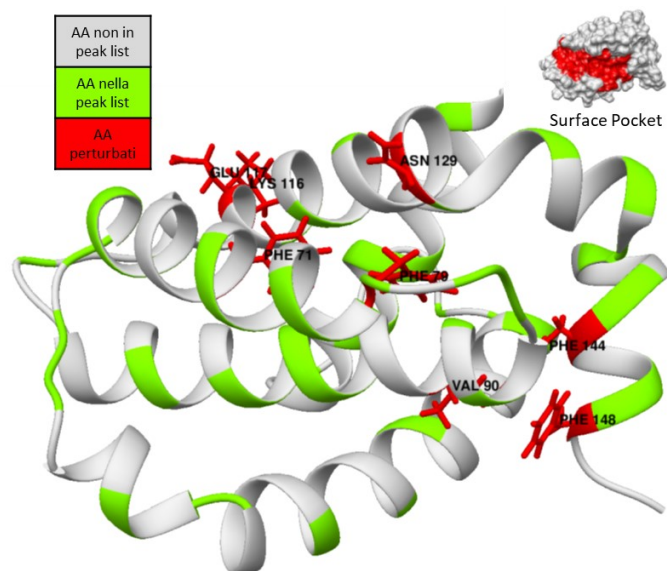


Figura 5.35 - Amminoacidi più perturbati in rosso, in verde gli aa presenti nella peak list e in grigio quelli non presenti nella peak list. In alto a destra si evidenzia la tasca idrofobica di Bfl-1 <sup>21</sup>

Come già detto il valore di *mean5* non troverà riscontro con le altre copresenze, in quanto non si sta andando a cercare un confronto con altre molecole, ma solamente la presenza o meno di un'interazione; tuttavia, per completezza il dato viene riportato.

MEAN5	
A2 10 mM + B6 100 mM	0,021 ppm

Tabella 5.14 – Mean 5

## 5.8.2 Copresenza A2+C5

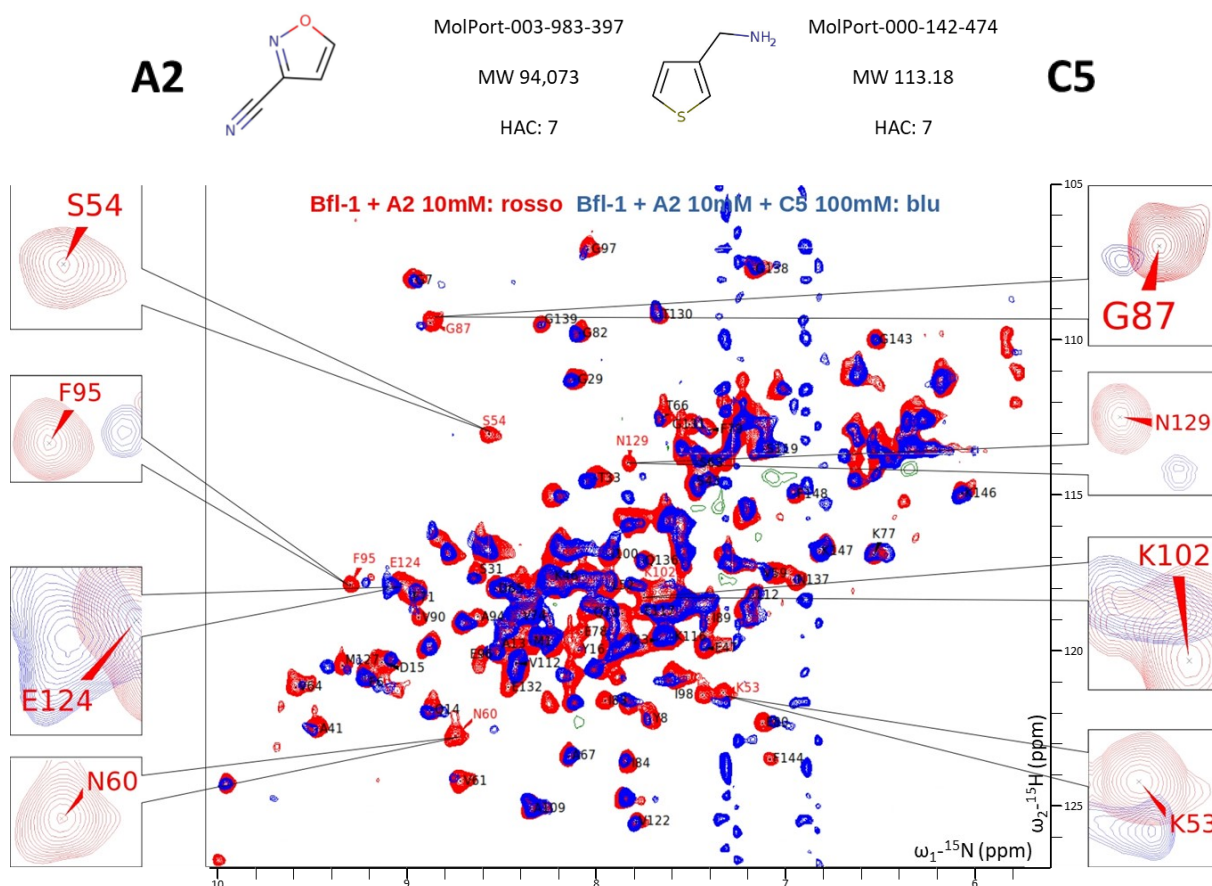


Figura 5.36 - Sovrapposizione degli spettri del bianco (Bfl-1 + A2 10 mM), e di C5 a 100 mM, con ingrandimento degli aminoacidi che presentano maggiori shift<sup>66</sup>

Questa copresenza è stata decisa di svolgerla per dipanare un dubbio che si era venuto a formare durante l'analisi degli spettri del primo set di esperimenti; ovvero si era cominciato a pensare che il frammento A2, visti gli *shift* molto maggiori rispetto agli altri aminoacidi, instaurasse un legame covalente con la proteina.

Per avere un primo riscontro è stato scelto di aggiungere anche C5, ovvero il secondo frammento con valore di *mean5* più alto rispetto agli altri e che lega la proteina tra P2 e P3. Essendo due MiniFrag con elevata affinità, in più essendo A2 ad una concentrazione di 10 mM e C5 ad una di 100 mM, ci si è messi nelle condizioni di avere affinità di legame paragonabile (valore di *mean5*) e quindi di favorire lo spiazzamento di A2 nel caso non fosse un ligando covalente.

MEAN5	
A2 10 mM + C5 100 mM	0,05 ppm

Tabella 5.15 – Mean 5

Ciò che emerge dallo spettro (FIGURA 5.36) è che sono presenti degli *shift* molto ampi, infatti se si valuta il valore di *mean5* (TABELLA 5.15) ci si rende conto che si tratta di spostamenti in ppm elevatissimi e molto vicini ai valori di *mean5* di A2. Se si va poi ad osservare la posizione degli amminoacidi perturbati nella struttura proteica (FIGURA 5.37) ci si rende conto che si tratta degli stessi amminoacidi perturbati da A2 negli esperimenti del primo set. Il dato che è così emerso indicherebbe che A2 venga di fatto spiazzato da C5, e che quindi non sia covalente. Tuttavia, per fugare ogni dubbio si è valutato anche di effettuare in futuro una LC Massa, tecnica distruttiva in grado di fornire un dato certo sulla covalenza o meno del frammento.

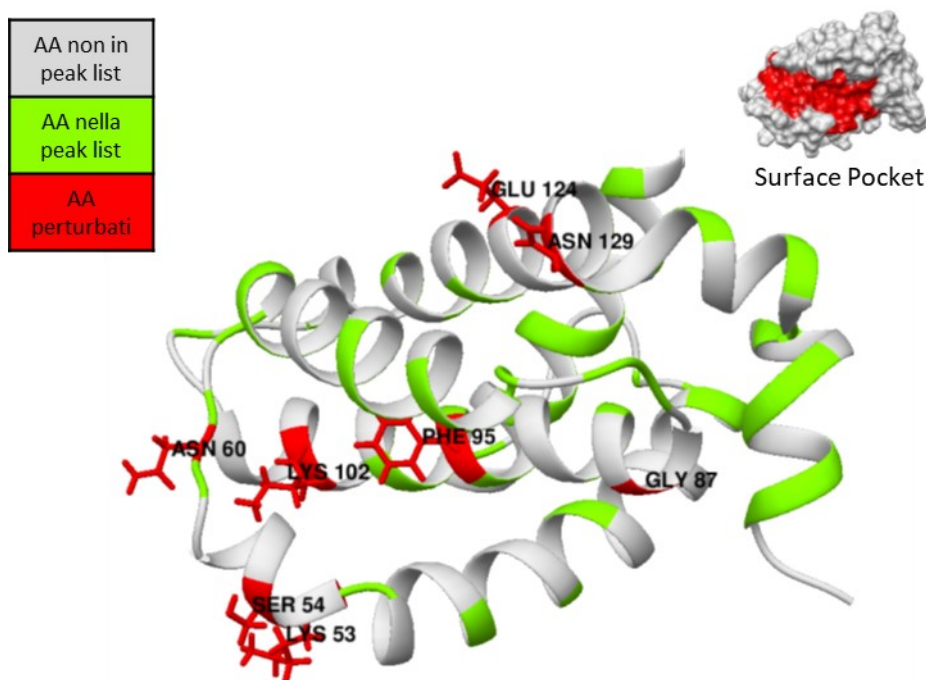


Figura 5.37 - Amminoacidi più perturbati in rosso, in verde gli aa presenti nella peak list e in grigio quelli non presenti nella peak list. In alto a destra si evidenzia la tasca idrofobica di Bfl-1<sup>21</sup>

### 5.8.3 Copresenza A2+C7

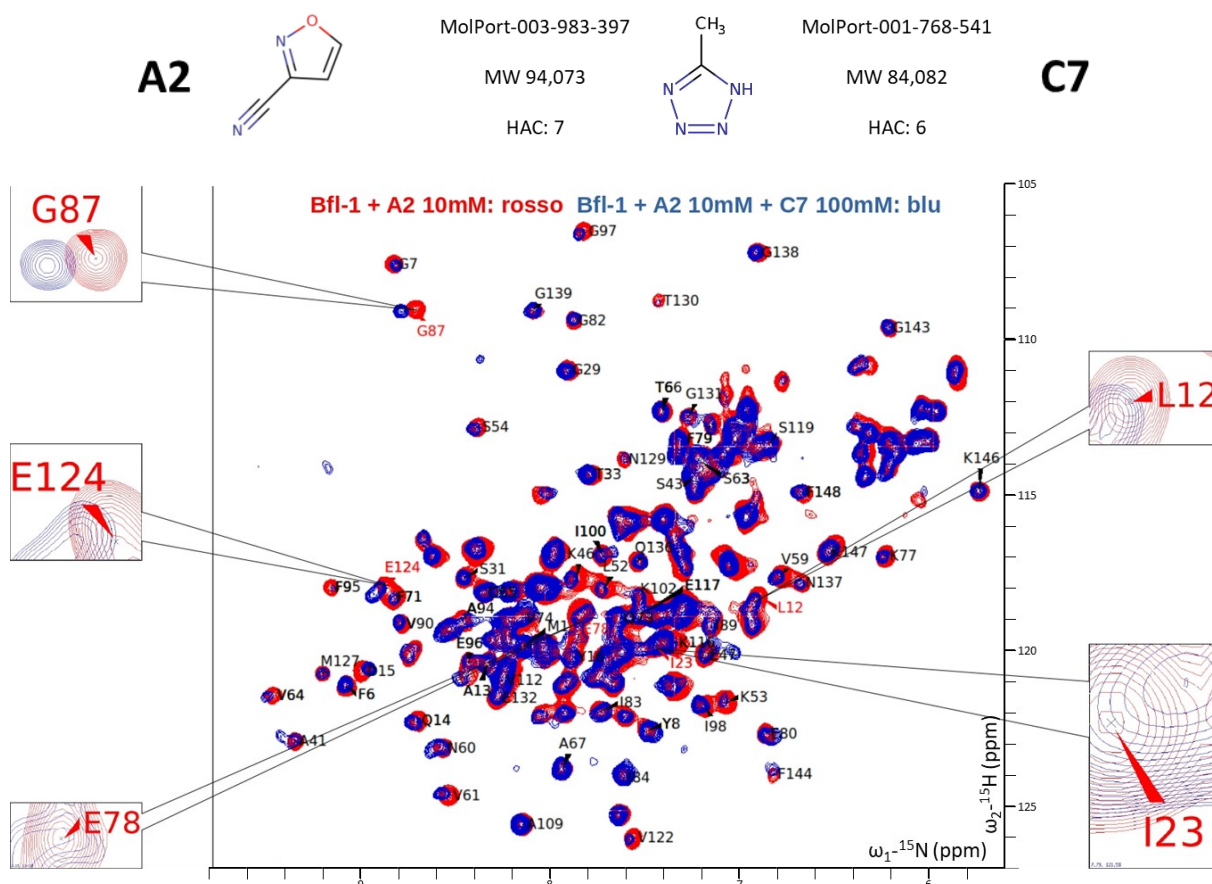


Figura 5.38 - Sovrapposizione degli spettri del bianco (Bfl-1 + A2 10 mM), e di C7 a 100 mM, con ingrandimento degli amminoacidi che presentano maggiori shift<sup>66</sup>

In questa analisi si è andati a valutare la copresenza di A2 e C7, avendo riscontrato che il secondo MiniFrag si legasse con buona probabilità a P3 e interagisse anche con la nuova *pocket* osservata sperimentalmente, si è pensato che una copresenza dei due forzasse il secondo ligando preferibilmente in una delle due tasche.

Dai dati NMR (FIGURA 5.38) si possono vedere degli *shift* di bassa intensità, fatta eccezione per G87. Andando poi ad analizzare la struttura della proteina (FIGURA 5.39) si osserva che la zona più perturbata sia in corrispondenza del terzo *pocket*, zona in cui si trova anche G8, questi due dati permettono di affermare che con buona probabilità il MiniFrag C7 si leghi in quella zona, confermando il dato del primo set di esperimenti.

Si osserva inoltre che la zona della seconda *pocket*, esterna alla tasca idrofobica, non viene più perturbata, indice sia che l'affinità per il P3 sia maggiore, oppure che i cambiamenti conformazionali, a seguito del legame con A2, modifichino la conformazione della tasca impedendo a C7 di entrarvi.



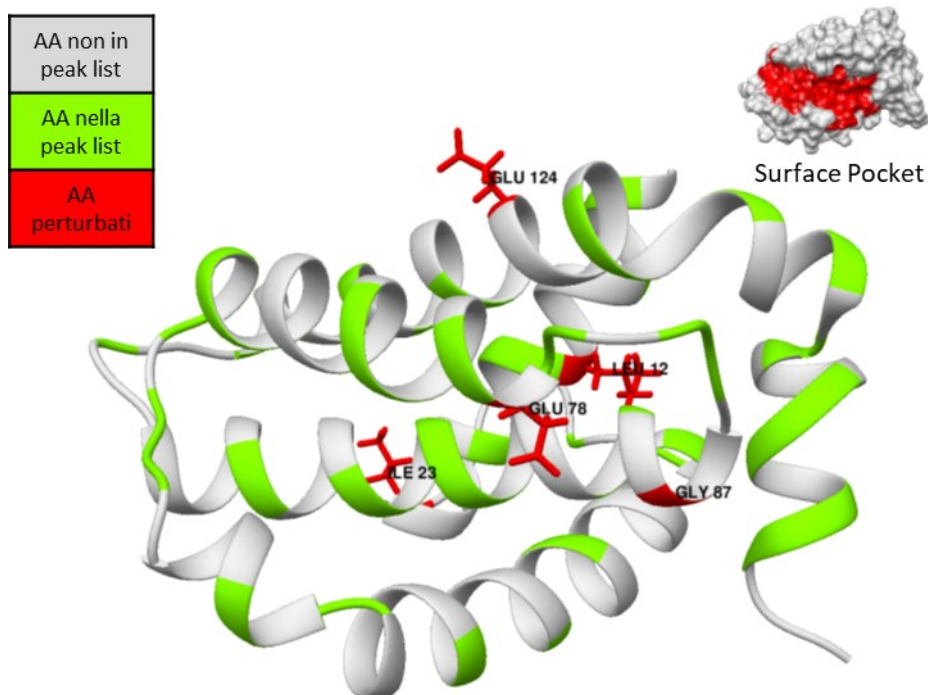


Figura 5.39 - Amminoacidi più perturbati in rosso, in verde gli aa presenti nella peak list e in grigio quelli non presenti nella peak list. In alto a destra si evidenzia la tasca idrofobica di Bfl-1<sup>21</sup>

Il valore di *mean5* (TABELLA 5.16) risulta molto simile al valore ottenuto nell'esperimento con solo C7, questo è indice che effettivamente l'interazione avvenga con la stessa affinità del primo esperimento. Questo dato fa di questa copresenza un buon candidato per un futuro *linking* dei due MiniFrag.

MEAN5	
A2 10 mM + C7 100 mM	0,027 ppm

Tabella 5.16 – Mean 5

## 5.8.4 Copresenza A2+C7+C2

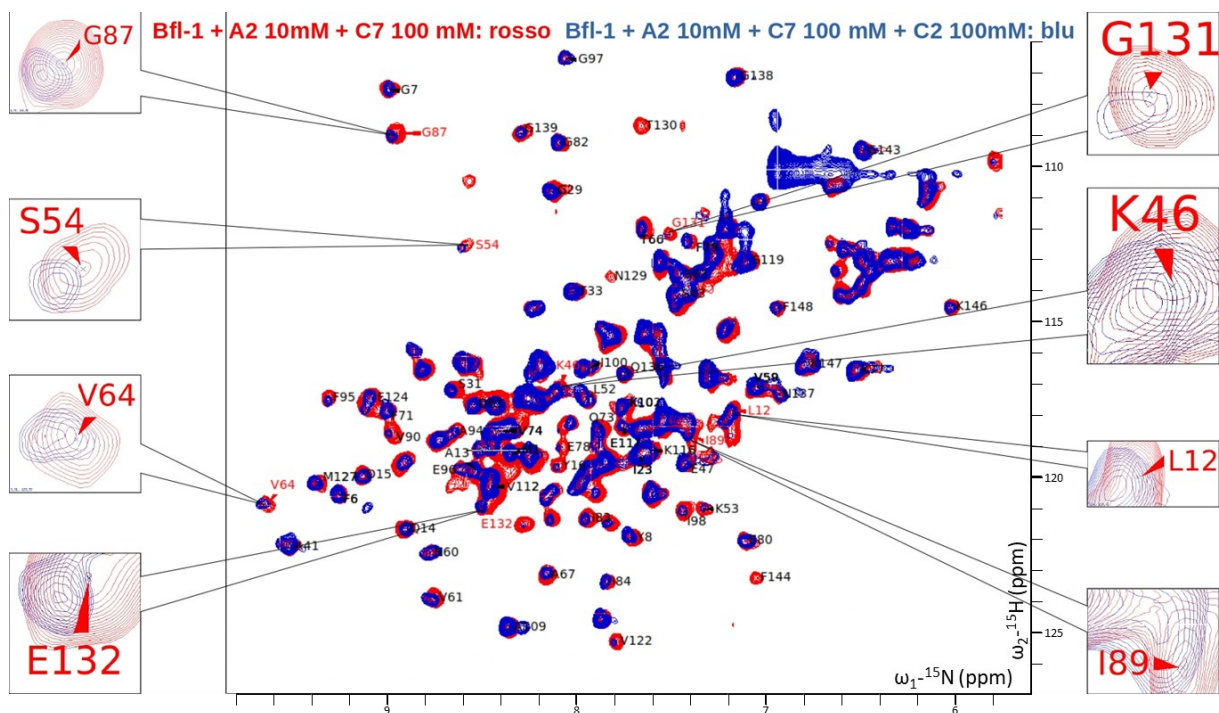
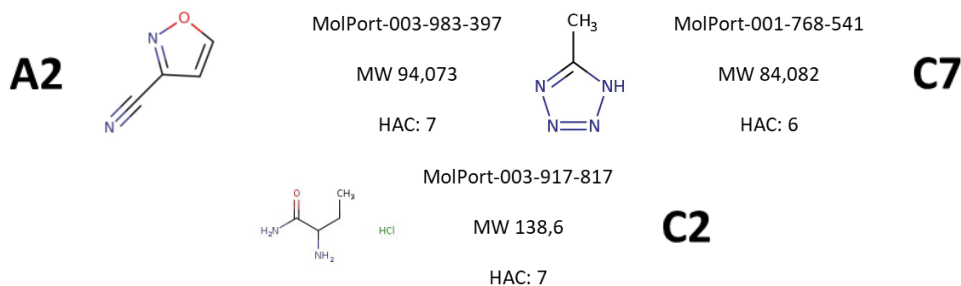


Figura 5.40 - Sovrapposizione degli spettri del bianco (Bfl-1 + A2 10 mM + C7 100 mM), e di C2 a 100 mM, con ingrandimento degli aminoacidi che presentano maggiori shift <sup>66</sup>

Visti i risultati ottenuti con A2+C7 si è voluto provare spingersi ancora oltre con l'analisi provando ad eseguire una triplice copresenza; si è utilizzato anche il MiniFrag C2, che si era visto posizionarsi in corrispondenza del P0.

Ciò che è emerso dallo spettro (FIGURA 5.40) è che si assiste a degli *shift* di molto bassa intensità, ciò lo si può vedere anche dal valore di *mean5* (TABELLA 5.17)

MEAN5	
A2 10 mM + C7 100 mM + C2 100 mM	0,020 ppm

Tabella 5.17 – Mean 5



Se si vanno poi ad osservare gli amminoacidi perturbati nella FIGURA 5.41 si può notare che nella zona del P0 non ci sono perturbazioni rilevanti, perciò questo MiniFrag non andrà a legarsi in quel punto.

Una cosa che emerge invece è che a livello di P4 vi è uno *shift*, che coinvolge alcuni degli amminoacidi (glicina 87) evidenziati nel precedente esperimento con C7 (FIGURA 5.31). Questo dato potrebbe essere spiegato come nella copresenza tra A2+C5, ovvero che in questo caso C2 si leghi con maggior affinità al sito di legame di C7; un'altra spiegazione al dato potrebbe essere data dal fatto che C2 leghi un sito di interazione diverso che però provoca un cambiamento conformazionale, impedendo a C7 di entrare nella tasca.

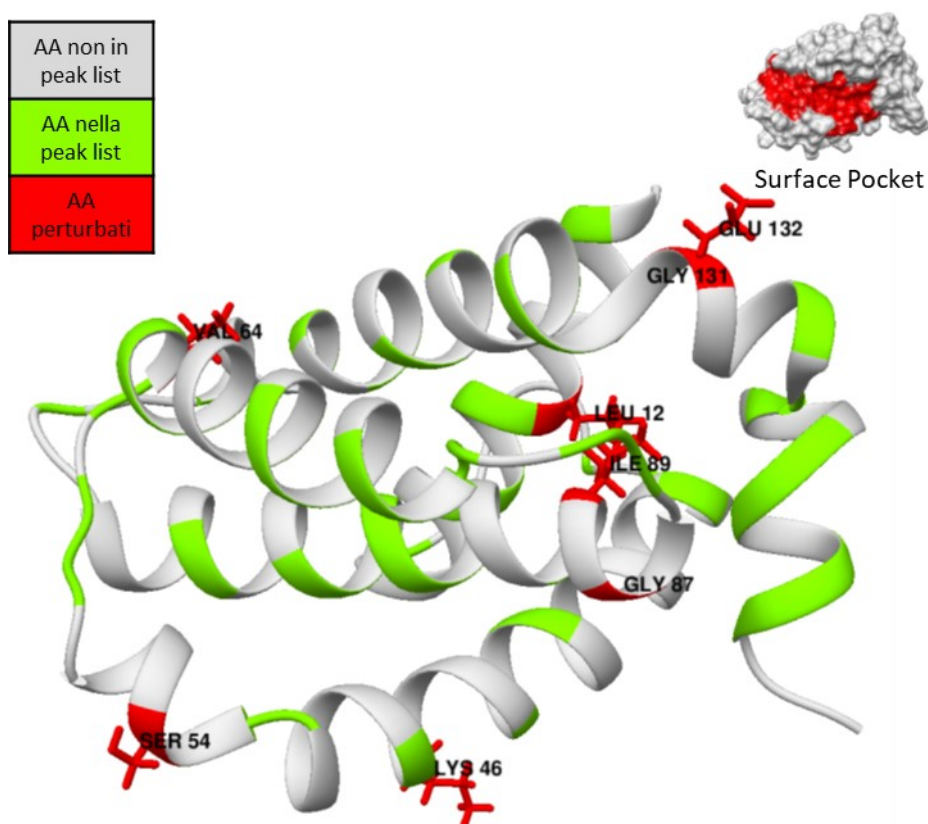


Figura 5.41 - Amminoacidi più perturbati in rosso, in verde gli aa presenti nella peak list e in grigio quelli non presenti nella peak list. In alto a destra si evidenzia la tasca idrofobica di Bfl-1<sup>21</sup>



## 6 CONCLUSIONI

Il progetto di tesi si è focalizzato su Bfl-1, una proteina antiapoptotica della famiglia delle Bcl-2. Tale proteina, interagisce con il dominio BH3 dei membri proapoptotici della stessa famiglia, inibendoli e quindi bloccando l'inizio del processo apoptotico.

Da letteratura è emerso che un gran numero di pazienti affetti da tumori ematologici (leucemie), presentassero all'interno delle cellule neoplastiche una sovraespressione di Bfl-1. Inoltre, sebbene in maniera meno frequente, è stato osservato che anche in alcuni tumori solidi, come il carcinoma al seno, al colon e gastrico, fosse presente un aumento considerevole dell'espressione di Bfl-1.

Da alcuni studi è risultato che la sovraespressione della proteina di interesse risulta essere coinvolta in meccanismi di aumento proliferativo incontrollato della neoplasia, rendendola in grado di sfuggire alla morte cellulare programmata indotta dall'organismo. Si è osservato inoltre che un aumento nell'espressione di Bfl-1 nella cellula tumorale causasse farmacoresistenza, perciò trattamenti farmacologici tradizionali con il cis-platino o etoposide diventavano inefficaci.

Le scoperte riguardanti la sovraespressione delle proteine della famiglia delle Bcl-2 nel contesto tumorale hanno portato alla ricerca di nuovi farmaci specifici per questa famiglia di proteine. Sono stati provati vari approcci, sia tramite *screening* di librerie *Fragment Based Drug Discovery* (FBDD), sia attraverso la creazione di oligonucleotidi antisenso, sia attraverso un innovativo approccio che cerca di riprodurre il sito BH3, così da generare un falso substrato per la proteina antiapoptotica. Attualmente l'unica molecola ad aver superato tutti i test necessari per l'autorizzazione all'immissione in commercio è *Venetoclax*<sup>28</sup>, un farmaco scoperto tramite FBDD, che lega selettivamente la proteina Bcl-2. Questo eccellente risultato apre la strada alla ricerca di nuovi inibitori selettivi per tutte le proteine della famiglia. Infatti la ricerca effettuata su Bfl-1 è stata intrapresa in quanto non sono ancora stati identificati inibitori selettivi in grado di legarla selettivamente ed inibirla.

Lo scopo di questa tesi è appunto quello di trovare un inibitore selettivo per Bfl-1 e lo si è perseguito attraverso la tecnica di FBDD tramite spettroscopia a risonanza magnetica nucleare (NMR), ma lo si è voluto fare utilizzando i MiniFrag, molecole molto più piccole dei frammenti solitamente utilizzati nella FBDD; questo perché grazie alle loro piccolissime

dimensioni permettono di sondare uno spazio chimico esponenzialmente più grande e permettono di interagire con punti più interni o inusuali della proteina, altrimenti insondabili con molecole di maggiori dimensioni. Per effettuare questa analisi si è utilizzata una piccola libreria di 10 composti detti MiniFrag, ovvero dei piccoli frammenti contenenti un numero di atomi pesanti (HAC) compreso tra 5 e 7, mai impiegati per la ricerca di possibili ligandi tramite NMR. La metodologia attraverso cui poter ottenere uno spettro NMR analizzabile e, di conseguenza il metodo necessario ai fini dell'analisi dei dati di questa particolare libreria, è stata un altro obiettivo di questa tesi.

Prima di poter procedere con l'analisi della libreria è stato necessario produrre per via ricombinante la proteina Bfl-1 marcata all' $^{15}\text{N}$ , utilizzando dei batteri *E. Coli*. Nei batteri è stato inserito un plasmide codificante per la proteina e per altre due strutture (6His e SUMO), necessarie alla purificazione. Successivamente è stato necessario separare la proteina di interesse dalle altre proteine solubili presenti all'interno dei batteri, utilizzando come tecnica la *Fast Protein Liquid Chromatography* (FPLC).

Le strutture prodotte assieme alla proteina sono state eliminate attraverso SUMO proteasi durante gli step di purificazione, permettendo così di ottenere circa 7 mg di proteina pura da 2 L di coltura.

Successivamente ai fini della realizzazione degli esperimenti tramite NMR si è partiti dall'unico articolo presente in letteratura in cui sono stati definiti ed utilizzati i MiniFrag per la FBDD tramite cristallografia a raggi X. Tramite tecniche computazionali svolte dal laboratorio MMS (Prof. Stefano Moro) sono stati campionati 100 MiniFrag andando a valutare un potenziale spazio chimico di  $10^4$ : questo ha portato alla selezione di 10 molecole, le quali si sono ritenute essere le più affini o quelle che con maggior probabilità avrebbero legato punti diversi della tasca idrofobica.

Si è utilizzata come tecnica analitica l' $^1\text{H}$ - $^{15}\text{N}$  SOFAST-HMQC, che partendo da un campione di proteina marcato all' $^1\text{H}$  e  $^{15}\text{N}$ , permette di ottenere uno spettro bidimensionale della proteina di interesse. Data la bassa affinità di legame dei MiniFrag si sono dovute effettuare varie prove di solubilità dei frammenti, sia per poterli portare ad una concentrazione massima di 100 mM nel campione da analizzare, sia per poter creare uno *stock* di soluzione madre ad un massimo di 1 M da utilizzare poi per la diluizione dei frammenti nel campione.

Dopo aver stabilito la formulazione dei campioni si è passati all'analisi dei dati, che è avvenuta tramite calcolo del *Chemical Shift Perturbation* (CSP), utilizzando un apposito *script*. Questo passaggio ha permesso l'identificazione degli amminoacidi più perturbati, attraverso il calcolo del triplo della deviazione standard, ed ha anche permesso di ottenere un valore utile al confronto con altri spettri, ovvero il *mean5*, calcolato come la media dei CSP dei 5 amminoacidi più perturbati. Si sono eseguiti tre esperimenti consecutivi, uno costituente il bianco (con sola proteina), mentre gli altri due effettuati a concentrazioni crescenti di frammento (20 mM e 100 mM). Da questi esperimenti è stato deciso di inserire un criterio di esclusione, ovvero di non accettare amminoacidi i cui spostamenti non fossero nella stessa direzione e nello stesso verso rispetto agli altri due spettri.

Dopo aver analizzato gli spettri con i criteri descritti, si è passati a rappresentare gli amminoacidi perturbati sulla struttura tridimensionale proteica, così da poter definire in che modo i MiniFrag si posizionassero all'interno della tasca idrofobica. Una volta eseguito questo passaggio, si sono confrontati i dati con quelli computazionali, in modo da rilevare eventuali riscontri e comprendere in che modo la molecola potesse disporsi spazialmente all'interno della tasca.

Dalle osservazioni si può evidenziare che il metodo sopra descritto ha comportato due conseguenze: innanzitutto la possibilità di vedere in maniera sufficientemente chiara e definita le interazioni proteina-MiniFrag, ed inoltre si sono potute vedere interazioni molto interessanti in 4 dei composti analizzati, e di questi 4 composti se ne sono visti 3 interagire in punti diversi della tasca idrofobica. Inoltre, dagli spettri emerge la presenza di una nuova possibile *pocket* di interazione con la proteina al di fuori della tasca idrofobica.

Dati così promettenti sono da prendere in considerazione in modo molto cauto, in quanto essendo i primi ad essere raccolti utilizzando questo tipo di approccio saranno passibili di ulteriori verifiche future. Tuttavia, dimostrano la possibilità di effettuare l'analisi di composti di tali dimensioni tramite NMR, aprendo la strada ad un nuovo metodo analitico che potrà essere successivamente migliorato.

La presenza di interazioni di tale rilevanza è stata un motore per spingere la sperimentazione ad un passo successivo, ovvero valutare se più frammenti aggiunti contemporaneamente allo stesso campione si legassero contemporaneamente alla tasca idrofobica in punti differenti. Il risultato di questa seconda serie di esperimenti è stato che alcuni frammenti, di

fatto, coesistano all'interno della tasca, mentre altri vengono spiazzati dal sito di interazione dal secondo frammento. Questa ulteriore scoperta potrà essere utilizzata in futuro per far maturare i frammenti, creando un *linker* (una catena idrocarburica), tra i frammenti che legano in posizioni differenti della tasca; questo ai fini di creare un composto *lead*, con caratteristiche di migliore affinità e maggior peso molecolare.

Per concludere, il lavoro svolto rappresenta l'inizio di un nuovo possibile approccio al FBDD tramite NMR. Attraverso l'impiego dei MiniFrag è possibile sondare uno spazio chimico elevatissimo, andando ad esplorare interazioni specifiche in più punti, anche contemporaneamente, del *binding site* proteico. Inoltre, la collaborazione con ricercatori che sfruttano tecniche computazionali ha permesso di rendere più efficiente la ricerca, provando una piccola quantità di molecole e avendo un tasso di successi estremamente elevato rispetto ad un normale *screening* di librerie FBDD.

In futuro le molecole che sono state utilizzate in questo progetto, probabilmente verranno rianalizzate alla luce di nuove metodologie analitiche che verranno implementate, e se manterranno la solidità dei dati osservati allora potranno venire maturate nella ricerca di una nuova molecola antitumorale che permetta di inibire la proteina Bfl-1 e di permettere ai farmaci antitumorali di agire eliminando le cellule malate.

La prospettiva futura di impiego delle molecole analizzate sarà quella di continuare con la validazione dei dati ottenuti, apportando delle migliorie al protocollo di analisi degli esperimenti e alla luce dei nuovi risultati ottenuti, procedere ad una seconda analisi computazionale. Se i risultati continueranno ad essere convincenti allora si potrà procedere con la maturazione dei MiniFrag sino ad ottenere una nuova molecola antitumorale, in grado di inibire la proteina Bfl-1 e così di coadiuvare i farmaci antitumorali nell'eliminazione delle cellule neoplastiche.

## 7 BIBLIOGRAFIA

- (1) Pontieri, G. M. *Elementi di patologia generale & fisiopatologia generale*; Piccin: Padova, 2018.
- (2) Stewart, B. W. *Mechanisms of Carcinogenesis: From Initiation and Promotion to the Hallmarks*; International Agency for Research on Cancer, 2019.
- (3) Tarantini, L.; Gori, S.; Inno, A.; Maurea, N.; Del Sindaco, D.; Faggiano, P.; Barni, S.; Falci, C.; Venturini, E.; Foglietta, J.; Barbieri, A.; Pulignano, G.; Tinti, M.; Tolone, S.; Spallarossa, P.; Ciccarese, M.; Belli, L.; Zustovich, F.; Tuccia, F.; Lecis, P. *Cardio-Oncologia 2017*; 2017.
- (4) *Agents Classified by the IARC Monographs, Volumes 1–131 – IARC Monographs on the Identification of Carcinogenic Hazards to Humans*. <https://monographs.iarc.who.int/agents-classified-by-the-iarc/> (accessed 2022-06-10).
- (5) *Cancer today*. <http://gco.iarc.fr/today/home> (accessed 2022-06-10).
- (6) *Cancer*. <https://www.who.int/news-room/fact-sheets/detail/cancer> (accessed 2022-06-10).
- (7) *Rubin's Pathology: Clinicopathologic Foundations of Medicine*; Lippincott, Williams, & Wilkins: Philadelphia, 2005.
- (8) Lehninger, A. L.; Nelson, D. L.; Cox, M. M.; Melloni, E. *I principi di biochimica di Lehninger*; Zanichelli: Bologna, 2018.
- (9) *Robbins & Cotran Pathologic Basis of Disease*, Tenth edition.; Kumar, V., Abbas, A. K., Aster, J. C., Turner, J. R., Perkins, J. A., Robbins, S. L., Cotran, R. S., Eds.; Elsevier: Philadelphia, PA, 2021.
- (10) Nagata, S. Apoptosis by Death Factor. *Cell* **1997**, *88* (3), 355–365. [https://doi.org/10.1016/S0092-8674\(00\)81874-7](https://doi.org/10.1016/S0092-8674(00)81874-7).
- (11) Brummer, T.; Schmitz-Peiffer, C.; Daly, R. J. Docking Proteins. *FEBS J.* **2010**, *277* (21), 4356–4369. <https://doi.org/10.1111/j.1742-4658.2010.07865.x>.
- (12) Wang, Y.; Tjandra, N. Structural Insights of TBid, the Caspase-8-Activated Bid, and Its BH3 Domain. *J. Biol. Chem.* **2013**, *288* (50), 35840–35851. <https://doi.org/10.1074/jbc.M113.503680>.
- (13) Czabotar, P. E.; Lessene, G.; Strasser, A.; Adams, J. M. Control of Apoptosis by the BCL-2 Protein Family: Implications for Physiology and Therapy. *Nat. Rev. Mol. Cell Biol.* **2014**, *15* (1), 49–63. <https://doi.org/10.1038/nrm3722>.
- (14) Kelly, P. N.; Strasser, A. The Role of Bcl-2 and Its pro-Survival Relatives in Tumorigenesis and Cancer Therapy. *Cell Death Differ.* **2011**, *18* (9), 1414–1424. <https://doi.org/10.1038/cdd.2011.17>.
- (15) Kitada, S.; Andersen, J.; Akar, S.; Zapata, J. M.; Takayama, S.; Krajewski, S.; Wang, H. G.; Zhang, X.; Bullrich, F.; Croce, C. M.; Rai, K.; Hines, J.; Reed, J. C. Expression of Apoptosis-Regulating Proteins in Chronic Lymphocytic Leukemia: Correlations with In Vitro and In Vivo Chemoresponses. *Blood* **1998**, *91* (9), 3379–3389.
- (16) *Tumori ematologici - Istituto Europeo di Oncologia*. <https://www.ieo.it/it/PER-I-PAZIENTI/Cure/Tumori-ematologici/> (accessed 2022-08-21).
- (17) *Definition of solid tumor - NCI Dictionary of Cancer Terms - NCI*. <https://www.cancer.gov/publications/dictionaries/cancer-terms/def/solid-tumor> (accessed 2022-08-21).
- (18) Ikegaki, N.; Katsumata, M.; Minna, J.; Tsujimoto, Y. Expression of Bcl-2 in Small Cell Lung Carcinoma Cells. *Cancer Res.* **1994**, *54* (1), 6–8.
- (19) Harvey, E. P.; Seo, H.-S.; Guerra, R. M.; Bird, G. H.; Dhe-Paganon, S.; Walensky, L. D. Crystal Structures of Anti-Apoptotic BFL-1 and Its Complex with a Covalent Stapled

- Peptide Inhibitor. *Struct. Lond. Engl.* 1993 **2018**, 26 (1), 153-160.e4. <https://doi.org/10.1016/j.str.2017.11.016>.
- (20) Schulz, G. E.; Schirmer, R. H. *Principles of Protein Structure.*; Springer: New York, NY, 2013.
- (21) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera--a Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, 25 (13), 1605–1612. <https://doi.org/10.1002/jcc.20084>.
- (22) Birkinshaw, R.; Gong, J.; Luo, C.; Lio, D.; White, C.; Anderson, M.; Blombery, P.; Lessene, G.; Majewski, I.; Thijssen, R.; Roberts, A.; Huang, D.; Colman, P.; Czabotar, P. Structures of BCL-2 in Complex with Venetoclax Reveal the Molecular Basis of Resistance Mutations. *Nat. Commun.* **2019**, 10. <https://doi.org/10.1038/s41467-019-10363-1>.
- (23) Vogler, M. BCL2A1: The Underdog in the BCL2 Family. *Cell Death Differ.* **2012**, 19 (1), 67–74. <https://doi.org/10.1038/cdd.2011.158>.
- (24) Parrondo, R. D.; Paulus, A.; Ailawadhi, S. Updates in the Use of BCL-2-Family Small Molecule Inhibitors for the Treatment of Relapsed/Refractory Multiple Myeloma. *Cancers* **2022**, 14 (14), 3330. <https://doi.org/10.3390/cancers14143330>.
- (25) Oltersdorf, T.; Elmore, S. W.; Shoemaker, A. R.; Armstrong, R. C.; Augeri, D. J.; Belli, B. A.; Bruncko, M.; Deckwerth, T. L.; Dinges, J.; Hajduk, P. J.; Joseph, M. K.; Kitada, S.; Korsmeyer, S. J.; Kunzer, A. R.; Letai, A.; Li, C.; Mitten, M. J.; Nettekheim, D. G.; Ng, S.; Nimmer, P. M.; O'Connor, J. M.; Oleksijew, A.; Petros, A. M.; Reed, J. C.; Shen, W.; Tahir, S. K.; Thompson, C. B.; Tomaselli, K. J.; Wang, B.; Wendt, M. D.; Zhang, H.; Fesik, S. W.; Rosenberg, S. H. An Inhibitor of Bcl-2 Family Proteins Induces Regression of Solid Tumours. *Nature* **2005**, 435 (7042), 677–681. <https://doi.org/10.1038/nature03579>.
- (26) Tse, C.; Shoemaker, A. R.; Adickes, J.; Anderson, M. G.; Chen, J.; Jin, S.; Johnson, E. F.; Marsh, K. C.; Mitten, M. J.; Nimmer, P.; Roberts, L.; Tahir, S. K.; Xiao, Y.; Yang, X.; Zhang, H.; Fesik, S.; Rosenberg, S. H.; Elmore, S. W. ABT-263: A Potent and Orally Bioavailable Bcl-2 Family Inhibitor. *Cancer Res.* **2008**, 68 (9), 3421–3428. <https://doi.org/10.1158/0008-5472.CAN-07-5836>.
- (27) EMA. *Venclyxto*. European Medicines Agency. <https://www.ema.europa.eu/en/medicines/human/EPAR/venclyxto> (accessed 2022-08-06).
- (28) Research, C. for D. E. and. FDA Grants Regular Approval to Venetoclax in Combination for Untreated Acute Myeloid Leukemia. *FDA* **2021**.
- (29) *BCL-2 mutations and venetoclax binding. a location of G101V and F104L...* ResearchGate. [https://www.researchgate.net/figure/BCL-2-mutations-and-venetoclax-binding-a-location-of-G101V-and-F104L-mutants-relative-to\\_fig2\\_333581164](https://www.researchgate.net/figure/BCL-2-mutations-and-venetoclax-binding-a-location-of-G101V-and-F104L-mutants-relative-to_fig2_333581164) (accessed 2022-08-21).
- (30) Bird, G. H.; Bernal, F.; Pitter, K.; Walensky, L. D. Synthesis and Biophysical Characterization of Stabilized  $\alpha$ -Helices of BCL-2 Domains. *Methods Enzymol.* **2008**, 446, 369–386. [https://doi.org/10.1016/S0076-6879\(08\)01622-4](https://doi.org/10.1016/S0076-6879(08)01622-4).
- (31) Scarfò, L.; Ghia, P. Reprogramming Cell Death: BCL2 Family Inhibition in Hematological Malignancies. *Immunol. Lett.* **2013**, 155 (1), 36–39. <https://doi.org/10.1016/j.imlet.2013.09.015>.
- (32) *Genasense (oblimersen sodium) FDA Approval Status*. Drugs.com. <https://www.drugs.com/history/genasense.html> (accessed 2022-08-07).
- (33) EMA. *Genasense*. European Medicines Agency. <https://www.ema.europa.eu/en/medicines/human/EPAR/genasense> (accessed 2022-08-07).



- (34) Stewart, M. L.; Fire, E.; Keating, A. E.; Walensky, L. D. The MCL-1 BH3 Helix Is an Exclusive MCL-1 Inhibitor and Apoptosis Sensitizer. *Nat. Chem. Biol.* **2010**, *6* (8), 595–601. <https://doi.org/10.1038/nchembio.391>.
- (35) Bissaro, M.; Sturlese, M.; Moro, S. The Rise of Molecular Simulations in Fragment-Based Drug Design (FBDD): An Overview. *Drug Discov. Today* **2020**, *25* (9), 1693–1701. <https://doi.org/10.1016/j.drudis.2020.06.023>.
- (36) Shi, L.; Zhang, N. Applications of Solution NMR in Drug Discovery. *Molecules* **2021**, *26* (3), 576. <https://doi.org/10.3390/molecules26030576>.
- (37) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings1PII of Original Article: S0169-409X(96)00423-1. The Article Was Originally Published in *Advanced Drug Delivery Reviews* 23 (1997) 3–25.1. *Adv. Drug Deliv. Rev.* **2001**, *46* (1), 3–26. [https://doi.org/10.1016/S0169-409X\(00\)00129-0](https://doi.org/10.1016/S0169-409X(00)00129-0).
- (38) Joseph-McCarthy, D.; Campbell, A. J.; Kern, G.; Moustakas, D. Fragment-Based Lead Discovery and Design. *J. Chem. Inf. Model.* **2014**, *54* (3), 693–704. <https://doi.org/10.1021/ci400731w>.
- (39) Jafari, R.; Almqvist, H.; Axelsson, H.; Ignatushchenko, M.; Lundbäck, T.; Nordlund, P.; Molina, D. M. The Cellular Thermal Shift Assay for Evaluating Drug Target Interactions in Cells. *Nat. Protoc.* **2014**, *9* (9), 2100–2122. <https://doi.org/10.1038/nprot.2014.138>.
- (40) Nguyen, H. H.; Park, J.; Kang, S.; Kim, M. Surface Plasmon Resonance: A Versatile Technique for Biosensor Applications. *Sensors* **2015**, *15* (5), 10481–10510. <https://doi.org/10.3390/s150510481>.
- (41) Chilingaryan, Z.; Yin, Z.; Oakley, A. J. Fragment-Based Screening by Protein Crystallography: Successes and Pitfalls. *Int. J. Mol. Sci.* **2012**, *13* (10), 12857–12879. <https://doi.org/10.3390/ijms131012857>.
- (42) *Microscale Thermophoresis (MST) as a novel, alternative biophysical method for fragment screening and hit profiling.* Sygnature Discovery. <https://www.sygnaturediscovery.com/publications/posters/microscale-thermophoresis-mst-as-a-novel-alternative-biophysical-method-for-fragment-screening-and-hit-profiling/> (accessed 2022-08-21).
- (43) Harner, M. J.; Frank, A. O.; Fesik, S. W. Fragment-Based Drug Discovery Using NMR Spectroscopy. *J. Biomol. NMR* **2013**, *56* (2), 65–75. <https://doi.org/10.1007/s10858-013-9740-z>.
- (44) Williamson, M. P. Using Chemical Shift Perturbation to Characterise Ligand Binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16. <https://doi.org/10.1016/j.pnmrs.2013.02.001>.
- (45) Erlanson, D. Practical Fragments: Review of 2018 Reviews. *Practical Fragments*, 2018.
- (46) O'Reilly, M.; Cleasby, A.; Davies, T. G.; Hall, R. J.; Ludlow, R. F.; Murray, C. W.; Tisi, D.; Jhoti, H. Crystallographic Screening Using Ultra-Low-Molecular-Weight Ligands to Guide Drug Design. *Drug Discov. Today* **2019**, *24* (5), 1081–1086. <https://doi.org/10.1016/j.drudis.2019.03.009>.
- (47) Hanahan, D.; Jessee, J.; Bloom, F. R. [4] Plasmid Transformation of Escherichia Coli and Other Bacteria. In *Methods in Enzymology*; Bacterial Genetic Systems; Academic Press, 1991; Vol. 204, pp 63–113. [https://doi.org/10.1016/0076-6879\(91\)04006-A](https://doi.org/10.1016/0076-6879(91)04006-A).
- (48) *Champion™ pET SUMO Expression System.* <https://www.thermofisher.com/order/catalog/product/K30001> (accessed 2022-08-21).
- (49) Wang, L.; Fan, D.; Chen, W.; Terentjev, E. M. Bacterial Growth, Detachment and Cell Size Control on Polyethylene Terephthalate Surfaces. *Sci. Rep.* **2015**, *5* (1), 15159. <https://doi.org/10.1038/srep15159>.

- (50) Hausjell, J.; Kutscha, R.; Gesson, J. D.; Reinisch, D.; Spadiut, O. The Effects of Lactose Induction on a Plasmid-Free E. Coli T7 Expression System. *Bioeng. Basel Switz.* **2020**, *7* (1), E8. <https://doi.org/10.3390/bioengineering7010008>.
- (51) Mohammadi, S.; Nadaraja, A. V.; Luckasavitch, K.; Jain, M. C.; Roberts, D. J.; Zarifi, M. H. A Label-Free, Non-Intrusive, and Rapid Monitoring of Bacterial Growth on Solid Medium Using Microwave Biosensor. *IEEE Trans. Biomed. Circuits Syst.* **2020**, *14* (1), 2–11. <https://doi.org/10.1109/TBCAS.2019.2952841>.
- (52) Zhang, P.; Zhang, G.; Wang, W. Ultrasonic Treatment of Biological Sludge: Floc Disintegration, Cell Lysis and Inactivation. *Bioresour. Technol.* **2007**, *98* (1), 207–210. <https://doi.org/10.1016/j.biortech.2005.12.002>.
- (53) Madadlou, A.; O'Sullivan, S.; Sheehan, D. Fast Protein Liquid Chromatography. In *Protein Chromatography*; Walls, D., Loughran, S. T., Eds.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2011; Vol. 681, pp 439–447. [https://doi.org/10.1007/978-1-60761-913-0\\_25](https://doi.org/10.1007/978-1-60761-913-0_25).
- (54) Block, H.; Maertens, B.; Spriestersbach, A.; Brinker, N.; Kubicek, J.; Fabis, R.; Labahn, J.; Schäfer, F. Chapter 27 Immobilized-Metal Affinity Chromatography (IMAC): A Review. In *Methods in Enzymology*; Burgess, R. R., Deutscher, M. P., Eds.; Guide to Protein Purification, 2nd Edition; Academic Press, 2009; Vol. 463, pp 439–473. [https://doi.org/10.1016/S0076-6879\(09\)63027-5](https://doi.org/10.1016/S0076-6879(09)63027-5).
- (55) Pacesa, M. Purification of Recombinant Adenoviral Hexon Proteins for Generation of Virus-Specific Antibodies & Next-Generation Sequencing of Adenoviral Genomes, 2016. <https://doi.org/10.13140/RG.2.2.20211.53282>.
- (56) *Overview of dialysis, desalting, buffer exchange and protein concentration - IT.* <https://www.thermofisher.com/uk/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/overview-buffer-exchange.html> (accessed 2022-08-21).
- (57) *A Theoretical Study of the Separation Principle in Size Exclusion Chromatography / Macromolecules.* <https://pubs.acs.org/doi/full/10.1021/ma902377g> (accessed 2022-08-21).
- (58) Shrestha, U. T. NANODROP. *My Scientific Blog - Research and Articles*, 2018.
- (59) Grimsley, G. R.; Pace, C. N. Spectrophotometric Determination of Protein Concentration. *Curr. Protoc. Protein Sci.* **2004**, Chapter 3, Unit 3.1. <https://doi.org/10.1002/0471140864.ps0301s33>.
- (60) Jacobsen, N. E. *NMR Spectroscopy Explained Simplified Theory, Applications and Examples for Organic Chemistry and Structural Biology*; 2008.
- (61) Keeler, J. *Understanding NMR Spectroscopy*, 2nd ed.; John Wiley and Sons: Chichester, U.K, 2010.
- (62) Parnell, S.; Woolley, E.; Boag, S.; Frost, C. Digital Pulsed NMR Spectrometer for Nuclear Spin-Polarized <sup>3</sup>He and Other Hyperpolarized Gases. *Meas. Sci. Technol.* **2008**, *19*, 045601. <https://doi.org/10.1088/0957-0233/19/4/045601>.
- (63) Qu, X.; Cao, X.; Cai, S.; Chen, Z. Reconstruction of Self-Sparse 2D NMR Spectra from Undersampled Data in the Indirect Dimension. *Sensors* **2011**, *11*, 8888–8909. <https://doi.org/10.3390/s110908888>.
- (64) Schanda, P.; Kupče, Ě.; Brutscher, B. SOFAST-HMQC Experiments for Recording Two-Dimensional Deteronuclear Correlation Spectra of Proteins within a Few Seconds. *J. Biomol. NMR* **2005**, *33* (4), 199–211. <https://doi.org/10.1007/s10858-005-4425-x>.
- (65) Donaldson, A. F.; Miller, A. A Computational Group Theoretic Symmetry Reduction Package for the Spin Model Checker. In *Algebraic Methodology and Software*

- Technology*; Johnson, M., Vene, V., Eds.; Lecture Notes in Computer Science; Springer: Berlin, Heidelberg, 2006; pp 374–380. [https://doi.org/10.1007/11784180\\_29](https://doi.org/10.1007/11784180_29).
- (66) Lee, W.; Tonelli, M.; Markley, J. L. NMRFAM-SPARKY: Enhanced Software for Biomolecular NMR Spectroscopy. *Bioinformatics* **2015**, *31* (8), 1325–1327. <https://doi.org/10.1093/bioinformatics/btu830>.