

**UNIVERSITA' DEGLI STUDI DI PADOVA**

FACOLTA' DI SCIENZE MATEMATICHE, FISICHE E NATURALI

LAUREA SPECIALISTICA IN BIOLOGIA MOLECOLARE

TESI DI LAUREA

**STUDI STRUTTURALI SULLA PROTEINA CagV  
DELL'ISOLA DI PATOGENICITA' DI *Helicobacter pylori***

RELATORE: CHIAR.MO PROF. GIUSEPPE ZANOTTI  
DIPARTIMENTO DI SCIENZE CHIMICHE

CORRELATORE: DOTT. SSA LAURA CENDRON  
DIPARTIMENTO DI SCIENZE CHIMICHE

LAUREANDO: ZANETTE SANDRO

ANNO ACCADEMICO 2006-2007

# INDICE

ABBREVIAZIONI	3
1. ABSTRACT	5
2. INTRODUZIONE	7
2.1 Caratteristiche generali di <i>Helicobacter pylori</i>	7
2.2 Caratteristiche del genoma di <i>H. pylori</i>	9
2.3 Meccanismi di infezione	10
2.4 I sistemi di secrezione di tipo IV	15
2.5 L'isola di patogenicità	19
2.6 La proteina CagV	21
2.7 Le interazioni tra le proteine CAG	23
3. MATERIALI E METODI	27
3.1 Amplificazione dei frammenti di DNA mediante PCR	27
3.2 Elettroforesi di macromolecole	32
3.3 Purificazione dei prodotti di PCR	33
3.4 Clonaggio dei frammenti di DNA tramite "TOPO cloning"	37
3.5 Trasformazione	39
3.7 Lisi delle cellule batteriche	40
3.8 Purificazione con tecniche cromatografiche	40
3.9 Spettrofotometria	43
3.10 Saggi di interazione tra proteine mediante "Dot-blot"	45
3.11 Dicroismo circolare	46
3.12 Cristallizzazione di proteine	50
4. RISULTATI E DISCUSSIONE	55
4.1 Amplificazione dei frammenti di DNA del gene <i>CagV</i>	56
4.2 Trasformazione dei batteri BI21 con pET151- <i>CagV</i>	58

4.3 Espressione di His <sub>6</sub> CagVn1 e His <sub>6</sub> CagVn2	58
4.4 Purificazione mediante cromatografia di affinità	61
4.5 Purificazione mediante cromatografia a esclusione dimensionale	65
4.6 Dicroismo circolare	68
4.7 Cristallizzazione	70
4.8 Studi di interazione tra le proteine <i>Cag</i>	72
5. RACCOLTA DEI DATI CRISTALLOGRAFICI	75
5.1 Raccolta dei dati cristallografici	75
5.2 Elaborazione dei dati	78
5.3 Calcolo del numero di molecole per cella unitaria (Z)	82
6. CONCLUSIONI	84
APPENDICI	87
BIBLIOGRAFIA	99

## ABBREVIAZIONI e SIMBOLI

### AMMINOACIDI

Ala	A	alanina
Arg	R	arginina
Asp	D	acido aspartico
Asn	N	asparagina
Cys	C	cisteina
Gly	G	glicina
Gln	Q	glutammina
Glu	E	acido glutammico
His	H	istidina
Ile	I	isoleucina
Lys	K	lisina
Leu	L	leucina
Met	M	metionina
Phe	F	fenilalanina
Pro	P	prolina
Ser	S	serina
Thr	T	treonina
Tyr	Y	tirosina
Trp	W	triptofano
Val	V	valina

Å	angstrom
APS	ammonio persolfato
<i>cag</i>	cytotoxin associated genes
Da	Dalton
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	acido etilendiamminotetraacetico
e	coefficiente di estinzione molare
F(hkl)	fattore di struttura
h,k,l	indici di Miller
<i>H. pylori (hp)</i>	<i>Helicobacter pylori</i>
HEPES	acido N-2-idrossietilpiperazina-N'-2-etansolfonico
IPTG	isopropil- $\beta$ -D-tiogalattopiranoside
l	lunghezza d'onda
mA	milliAmpere
M,mM	molare, millimolare

MAD	Multiple Anomalous Dispersion
MES	acido metansolfonico
MIR	metodo di sostituzione isomorfa
MPD	metil pentan diolo
MES	acido 2-(N-morfolino)etan solfonico
MIB	tampone acido malonico-imidazolo-acido borico
MMT	tampone acido malico-MES-tris
µL	microlitri
OD	densità ottica
O.N.	over night (per una notte)
PBS	tampone salino contenente fosfato
PCB	acido propionico-cacodilato-bis tris propano
PEG	polietilenglicole
pellet	precipitato di cellule o membrane batteriche
pI	punto isoelettrico
rpm	round per minute (giri al minuto)
SDS	sodio dodecil solfato
SDS-PAGE	SDS-Polyacrylamide Gel Electrophoresis
TEMED	N,N,N',N'-Tetrametilendiammina
TAE	tampone tris-acido acetico-EDTA
TEV protease	Tobacco Etch virus protease
Tris	tris(idrossimetil)amminometano
V	volt

## **1. ABSTRACT**

The aim of this work is the expression, purification and structural characterization of the CagV protein from *Helicobacter pylori*. This protein belongs to the *cag* (cytotoxin associated genes) pathogenicity island of the bacterium, a region of the genome that codifies for a T4SS (Type 4 Secretion System), an apparatus that confers to it the ability to cause several diseases like gastric ulcers, adenocarcinomas and stomach lymphoma. According to sequence homologies, *CagV* is a functional homologue of the *VirB8* inner membrane protein from *A. tumefaciens*. To obtain soluble domains of CagV, the fragments 59-252 and 90-252 has been amplified by PCR and cloned in pET151 plasmid. The proteins have been expressed in *E.coli* BL21 with a six-histidines tag at the N-terminus, then they have been purified by IMAC and gel filtration chromatography. The purified fragments were used to perform structural studies, such as circular dichroism and crystallization trials. The 59-252 (CagVn1) fragment has been crystallized at a concentration of 5 mg/mL, using as a precipitant a solution containing 0.2 M ammonium sulphate, 0.1 M Mes pH 6.5, 30%w/v PEG monomethyl ether 5000. Diffraction data have been measured at the European Synchrotron Radiation Facility (ESRF, Grenoble); crystals diffracted at 2.7Å resolution. Finally, interaction experiments have been performed using the dot-blot technique to verify possible interactions among proteins of the pathogenicity island.



## **2. INTRODUZIONE**

### **2.1 Caratteristiche generali di *Helicobacter pylori***

*Helicobacter pylori* è un batterio microaerofilo Gram-negativo, lungo 2,5-5 µm, capace di colonizzare lo stomaco umano provocando l'insorgenza di diverse patologie, tra cui gastriti croniche, malattie da reflusso gastroesofageo, ulcere e vari tipi di tumori. Nonostante il pH dello stomaco sia spesso così basso da impedire la crescita della maggior parte delle forme microbiche, *Helicobacter pylori* ha sviluppato diverse strategie per far fronte a questo ambiente ostile, al punto tale da riuscire ad infettare più del 50% della popolazione umana mondiale. Tuttavia nella maggior parte dei casi (80%-85% circa) l'infezione avviene in forma asintomatica o di modesta entità; nel 15%-20% dei casi, invece, provoca patologie gastroduodenali, tra cui varie forme di gastrite, ulcera gastrica e duodenale, adenocarcinoma ed linfoma di MALT (Mucosal Associated Lymphoid Tissue) [Robinson et al., 2007]. La trasmissione può avvenire in diversi modi [Axon, 1995 e Stone, 1999]: oro-orale, gastro-orale ed oro-fecale. Nella trasmissione oro-orale l'evidenza epidemiologica suggerisce che la placca dentaria possa rappresentare un importante serbatoio. In uno studio, molti soggetti hanno presentato una positività in diversi siti orali, quali le tasche periodontali ed il dorso della lingua; *H. pylori* è stato anche ricercato sotto le unghie della mano dominante, dove è stato trovato nel 58% dei casi. I ricercatori hanno dimostrato, quindi, che i portatori orali di *H. pylori* giocano un importante ruolo nella trasmissione dell'infezione e la mano può essere uno strumento di trasmissione. Le secrezioni gastriche possono verosimilmente rappresentare la fonte di *H. pylori* presente nello stomaco. La propagazione oro-fecale potrebbe ricorrere direttamente da persone infette o che vivono in condizioni di sovraffollamento o in condizioni sanitarie scadenti, oppure indirettamente da acqua o cibi contaminati. La trasmissione attraverso l'acqua è riportata in varie parti del mondo ed è probabilmente dovuta alla contaminazione fecale. E' possibile anche la trasmissione sessuale: infatti, poiché l'*Hp* è presente nelle feci, l'infezione può essere trasmessa durante rapporti sessuali oro-anali con un individuo infetto. L'incidenza dell'infezione da *Hp* è andata sempre più riducendosi con lo sviluppo industriale: infatti, nei



paesi sviluppati, soggetti di elevato stato socio-economico hanno percentuali più basse di infezione, mentre in quasi tutti i paesi in via di sviluppo i livelli di infezione rimangono ancora elevati. La prevalenza aumenta con l'età ed è direttamente correlata con un basso livello socio-economico durante l'infanzia.

L'entità dell'infezione dipende da svariati fattori, tra cui la reazione dell'ospite, i polimorfismi del batterio, la fisiologia, il genotipo e le abitudini alimentari dell'organismo che lo ospita.

*Helicobacter pylori* possiede alcune caratteristiche che ne consentono la sopravvivenza; in particolare, presenta una forma elicoidale che, assieme ai flagelli, consente il movimento attraverso lo strato di muco che riveste la superficie dello stomaco; inoltre ha la capacità di produrre degli enzimi specifici che permettono di tamponare il pH. E' capace di provocare danni cellulari legandosi agli antigeni di Lewis presenti sulle cellule gastriche e secernendo diversi fattori chemioattrattori ed immunostimolanti. Inoltre è stato dimostrato che la presenza dell'isola di patogenicità *cag*, che codifica per un sistema di secrezione di tipo IV, è fondamentale per la virulenza del batterio [Montecucco e Rappuoli, 2001].



**Figura 2.1** Fotografia al microscopio elettronico di *Helicobacter pylori* [Montecucco e Rappuoli, 2001].

## 2.2 Caratteristiche del genoma di *H. pylori*

Il genoma del ceppo virulento 26695 di *H. pylori* [Tomb et al., 1997] consiste di un cromosoma circolare lungo 1,667,867 bp ed un contenuto medio G + C del 39%. Ci sono tuttavia cinque regioni specifiche che presentano una significativa differenza nel loro contenuto in G + C. Due di queste contengono almeno una copia di una sequenza di inserzione (IS605) e sono fiancheggiate da una sequenza di RNA ribosomale 5S ad una estremità e da una sequenza di 521 bp ripetuta sette volte nell'altra; queste due regioni sono rilevanti in quanto contengono dei geni implicati nel processamento del DNA; una delle due contiene anche due ortologi del gene *virB4/ptI*, il cui prodotto serve per il trasferimento del T-DNA oncogenico in *Agrobacterium* e per la secrezione della tossina in *Bordetella pertussis*. Un'altra regione è l'isola di patogenicità *cag*, che è fiancheggiata da ripetizioni di 31 bp e sembra derivare da un trasferimento genico orizzontale.

Sono state individuate 36 specie di tRNA, organizzate in 7 clusters e 12 geni singoli, e due serie di geni di rRNA 23S-5S e 16S.

Lungo il cromosoma sono inoltre presenti 8 famiglie di ripetizioni (con identità >97%), che variano in lunghezza da 0.47 a 3.8 Kb. Queste ripetizioni sono presenti sia in regioni codificanti (derivanti per esempio da duplicazioni) che in regioni intergeniche; parecchie sono associate ai geni che codificano per le proteine della membrana esterna (OMP).

Sono state identificate 1590 ORF, aventi una lunghezza media di 945 bp, simili a quelle presenti in altri genomi procariotici. Più del 70% delle proteine codificate da queste ORF hanno, in *H. pylori*, un punto isoelettrico maggiore di 7, diversamente dal 40% di *E. coli* ed *H. influenzae*; ciò riflette ragionevolmente un adattamento da parte del batterio all'acidità dell'ambiente gastrico.

Sono presenti inoltre 95 famiglie di geni paraloghi, comprendenti 266 prodotti genici (16% del totale) identificati. La maggior parte dei paraloghi ha un ruolo assegnato che cade nelle categorie funzionali tipo rivestimento cellulare, trasporto e legame di proteine, replicazione del DNA.

*Helicobacter pylori* presenta una enorme variabilità genetica, determinata da molteplici fattori tra cui l'ambiente, la differente area geografica dell'organismo ospite e le abitudini alimentari delle persone infette. Infatti, un ceppo che si è ben adattato ad un particolare ambiente, dopo il

trasferimento in un altro individuo può evolvere in un genotipo maggiormente adatto al nuovo ambiente (divergenza genetica); inoltre in uno stesso individuo possono convivere diversi ceppi batterici, che possono prendere il sopravvento alternativamente a livello dello stomaco [Blaser e Berg, 2001].

I polimorfismi di *Helicobacter pylori* sono dovuti ai meccanismi genetici classici, cioè mutazioni puntiformi, sostituzioni, inserzioni e delezioni di uno o più geni, oppure riarrangiamenti più ampi che coinvolgono interi cromosomi. Uno speciale gruppo di polimorfismi riguarda le sequenze ripetute mononucleotidiche o dinucleotidiche nei pressi dell'estremità 5' dei geni detti "contingency genes". In queste regioni sono frequenti gli slittamenti della DNA polimerasi, i quali durante la sintesi del DNA determinano spesso alterazioni del registro di lettura che possono accendere o spegnere un determinato gene. Infine, altri sistemi che determinano un aumento della variabilità genetica in modo efficiente sono la trasformazione e la coniugazione [Alm et al., 1999].

### 2.3 Meccanismi di infezione

Molteplici sono i meccanismi attraverso i quali *H. pylori* provoca l'infezione a livello dello stomaco. Il primo comprende la sintesi di un enzima, un'ureasi  $\text{Ni}^{2+}$  dipendente, in grado di idrolizzare l'urea presente nel succo gastrico in ammoniaca ed anidride carbonica, al fine di rendere meno acido il pH nell'intorno del batterio.

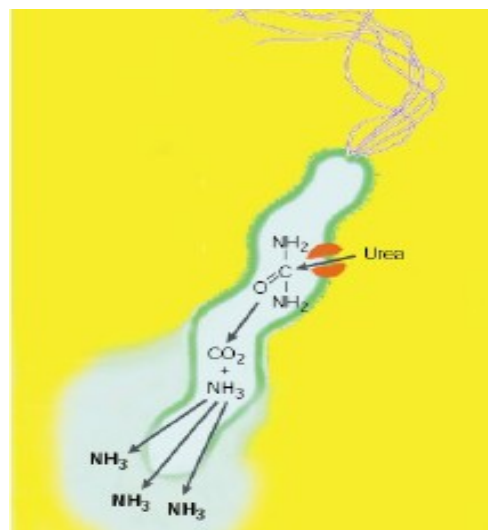


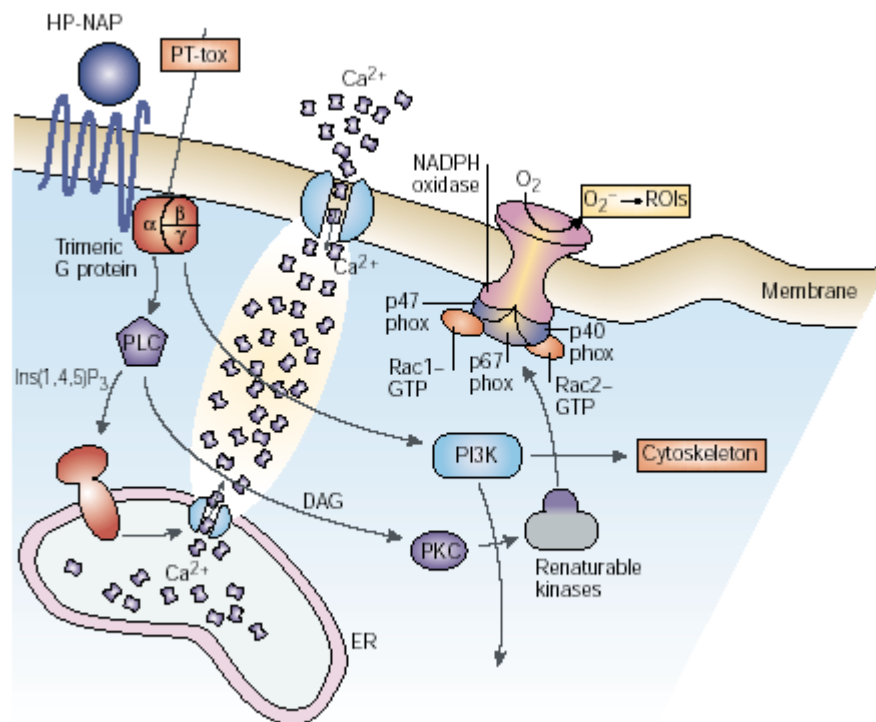
Figura 2.2 Produzione di ammoniaca da parte di *Hp* [Montecucco e Rappuoli, 2001].

Questo enzima è codificato da sette geni, due dei quali, UreA e UreB, codificano per le rispettive subunità dell'enzima, mentre gli altri codificano per le proteine accessorie necessarie per l'inserzione dello ione  $Ni^{2+}$  all'interno dell'enzima. Da un punto di vista strutturale questo enzima è un dodecamero, composto da 6 subunità UreA (26,5 kDa) e 6 subunità UreB (60,4 kDa), formanti nel complesso un anello avente un diametro di 13 nm. Questo enzima ha un ruolo essenziale nell'infezione, infatti batteri mutanti, privi di esso, non sono assolutamente in grado di colonizzare lo stomaco. La produzione di ammoniaca è importante, in quanto le basi deboli permeabili ad alte concentrazioni causano alterazioni cellulari quali il rigonfiamento di compartimenti cellulari, l'alterazione del trasporto vescicolare di membrana, l'inibizione della sintesi proteica, di produzione di ATP e del ciclo cellulare; l'ammoniaca potrebbe inoltre reagire con i metaboliti dell'enzima mieloperossidasi dei granulociti neutrofili per formare sostanze cancerogene, che aumenterebbero la possibilità di trasformazione delle cellule gastriche nella forma neoplastica. Inoltre, l'ureasi potrebbe essere coinvolta nel reclutamento degli stessi neutrofili e dei monociti, nonché nella stimolazione della produzione di citochine proinfiammatorie [Robinson et al., 2007].

L'attacco di *H. pylori* alle cellule del lume gastrico è permesso dall'interazione di diverse proteine batteriche con alcune glicoproteine cellulari. Per esempio BabA, una proteina di membrana esterna del batterio, interagisce con l'antigene tipo B di Lewis (un sistema di antigeni solubili che costituisce uno dei principali sierotipi che distinguono i vari gruppi sanguigni) dell'ospite; una particolare lectina batterica interagisce invece con i residui di acido sialico della laminina. Varie altre proteine di *H. pylori* contribuiscono, in cooperazione con quelle sopra riportate, a stabilire un'interazione molto forte con le cellule dell'ospite. Ciò provoca un profondo riarrangiamento della membrana plasmatica delle cellule sottostanti, la quale via via si estende fino ad unirsi alla superficie del batterio [Segal et al., 1996].

In seguito all'attacco di *H. pylori* alle cellule gastriche si verifica un reclutamento di cellule infiammatorie da parte delle cellule endoteliali, mediante il rilascio di varie chemochine e l'aumento dell'espressione di molecole di adesione. Il batterio è in grado di attivare direttamente le cellule endoteliali e può secernere una proteina oligomerica, detta HP-NAP (*H. pylori* neutrophil-activating protein), che funge da fattore chemotattico per i neutrofili. HP-NAP è un dodecamero di 150 kDa che

viene secreto per autolisi nell'ambiente circostante al batterio e si ritiene che si leghi ad un recettore a sette domini transmembrana presente sui granulociti neutrofili, innescando una cascata di trasduzione del segnale che comprende un aumento di calcio citosolico e la fosforilazione di alcune proteine, portando alla sintesi e al rilascio di specie reattive dell'ossigeno tramite una NADPHossidasi.



**Figura 2.3** Rappresentazione degli eventi coinvolti nell'attivazione dei leucociti da parte di HP-NAP. Il legame di questa proteina al recettore provoca l'apertura dei canali del calcio presenti nella membrana plasmatica e in quella del reticolo citoplasmatico. Inoltre il recettore attivato induce l'attività delle fosfatidilinositolo 3-chinasi (PI3K). L'aumento di calcio citosolico e l'attività di PI3K determinano l'attivazione di chinasi specifiche che fosforilano alcuni siti della NADPH ossidasi, la quale viene quindi attivata e produce anioni superossido e altre specie reattive dell'ossigeno [Montecucco e Rappuoli, 2001].

Inoltre HP-NAP altera l'equilibrio coagulazione-fibrinolisi; questo effetto avviene anche a bassa concentrazione di HP-NAP e potrebbe impedire la rigenerazione tissutale, che richiede la fibrinolisi [Montemurro et al., 2001]. Alcuni studiosi ipotizzano che originariamente fosse una proteina legante il ferro evolutasi come attivatore di leucociti e che quindi sia fondamentale per indurre uno stato di infiammazione tissutale. La sequenza nucleotidica del gene codificante HP-NAP è molto conservata in *Helicobacter pilori* e sono stati trovati anche ortologhi in diverse altre specie; studi strutturali sulla proteina rivelano la presenza di quattro eliche per ogni subunità e

l'oligomerizzazione delle subunità in dodecameri, caratterizzati da una cavità centrale ospitante il ferro (vedi fig. 2.4) [Montecucco e Rappuoli, 2001].

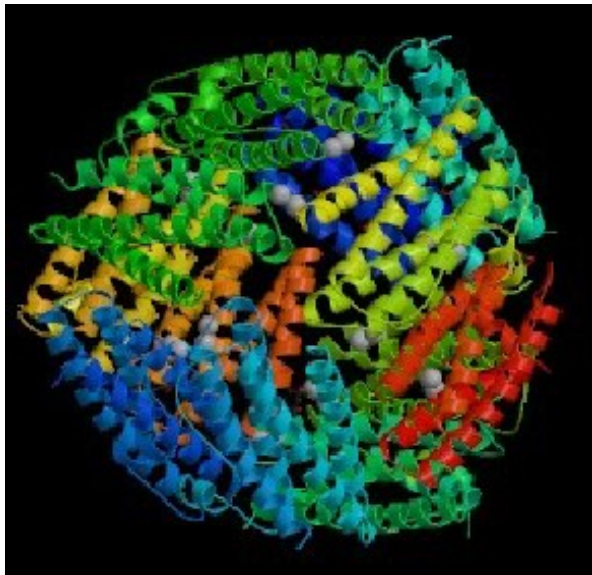
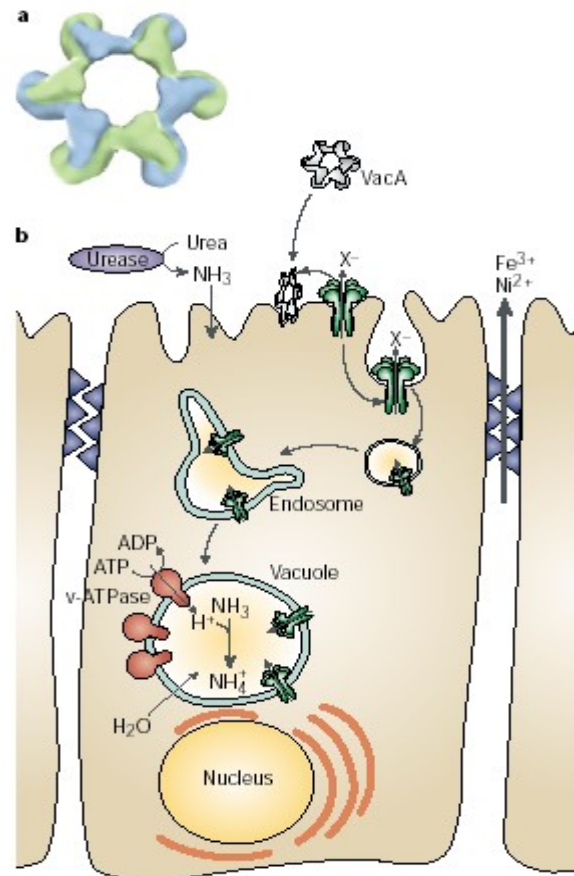


Figura 2.4 Immagine dell'enzima HP-NAP [PDB, 1J14].

Un fattore fondamentale per la crescita del batterio è dato dalla tossina VacA (vacuolating cytotoxin A) [De Bernard et al., 1995], una proteina di circa 90 KDa avente una regione N-terminale (37 KDa) ricca in foglietti  $\beta$  e amminoacidi idrofobici (propensione ad inserirsi in membrana) e la parte C-terminale (58 KDa) data da due domini (uno conservato, l'altro no) separati da un frammento variabile in sequenza e molto flessibile. Vac A funge da un lato da antigene verso cui la risposta immunitaria dell'ospite è diretta, dall'altro da importante fattore di virulenza di *H. pylori*. Essa conferisce un notevole vantaggio competitivo ai ceppi wild-type che la possiedono, rispetto ai mutanti che ne sono privi. Parte della proteina viene rilasciata nel mezzo, ma il 40-60% resta associata alla membrana esterna del batterio. Quella secreta si lega tra la porzione apicale delle cellule gastriche e lo strato di muco che le ricopre, dove oligomerizza in rosette di 6-7 subunità (vedi fig. 2.5a); inoltre forma un canale anionico specifico a bassa conduttanza, tramite il quale vengono rilasciati dal citosol della cellula epiteliale ioni organici e bicarbonato, allo scopo di garantire la crescita del batterio. L'attacco della proteina alla membrana cellulare è indotto dall'ambiente fortemente acido, che permette l'esposizione di siti idrofobici e la sua inserzione nel doppio strato fosfolipidico. I canali costituiti da VacA vengono lentamente internalizzati

per endocitosi negli endosomi, dove instaurano un aumento della permeabilità agli ioni mediato da una pompa protonica con attività ATPasica. Man mano che si produce ammoniaca da parte dell'ureasi, si formano all'interno dell'endosoma degli ioni  $\text{NH}_4^+$ , che a loro volta determinano l'entrata di acqua e la crescita dei vacuoli (vedi fig 2.5b).



**Figura 2.5** In alto è rappresentato l'oligomero di Vac (struttura a rosette), in basso è illustrato un modello del meccanismo di induzione dei danni cellulari da parte di VacA [Montecucco e Rappuoli, 2001].

Un altro passaggio importante dell'infezione di *Helicobacter pylori* è la produzione di CagA, una proteina di 128-145 kDa che, una volta iniettata nel citoplasma della cellula infettata, viene fosforilata a livello degli aminoacidi tirosina. La funzione di questa proteina non è ancora nota, ma si ritiene che possa legarsi a proteine aventi domini SH2 (Src-Homology region 2), determinando un segnale specifico che promuove la riorganizzazione del citoscheletro e di alcuni elementi della membrana apicale. Questa proteina inoltre è altamente immunogenica, come dimostra l'alto titolo di anticorpi contro di essa e l'abbondante localizzazione di linfociti T a livello della mucosa gastrica dei pazienti infetti [Satin et al., 2000].

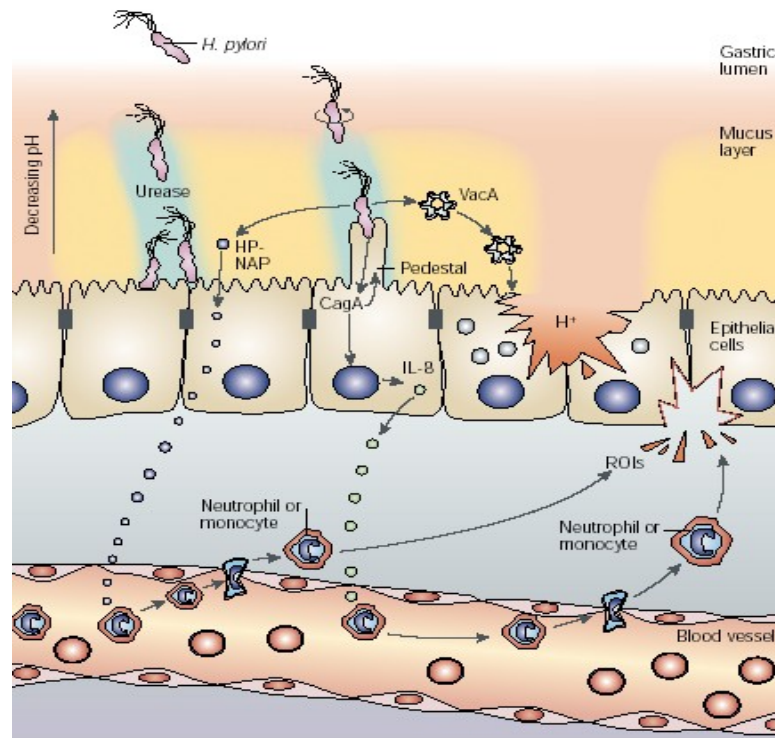


Figura 2.6 L'immagine fornisce un riassunto di tutti i fattori di virulenza di *H. pylori* che intervengono durante la colonizzazione dello stomaco [Montecucco e Rappuoli, 2001].

## 2.4 I sistemi di secrezione di tipo IV

I sistemi di secrezione di tipo IV (T4SS) sono delle strutture, presenti in alcuni batteri patogeni, che hanno la funzione di trasportare molecole (fattori di infettività) attraverso la membrana plasmatica. Essi costituiscono una grande superfamiglia, all'interno della quale si distinguono tre sottocategorie (vedi fig. 2.7). La prima comprende i sistemi di coniugazione che mediano il trasferimento di DNA tra cellule appartenenti a specie anche molto diverse (funghi, cellule animali e vegetali). La seconda comprende sistemi di assunzione e rilascio di DNA indipendenti dall'instaurarsi di un contatto con la cellula bersaglio; questi comprendono sia sistemi per l'acquisizione delle caratteristiche di competenza del batterio, sia sistemi per il rilascio di DNA, quali il plasmide F tipo *Tra* presente nell'isola genetica di *Neisseria Gonorrhoeae*. La terza sottocategoria è costituita da sistemi in grado di translocare macromolecole effettrici (proteine o DNA), dopo che il batterio è riuscito a stabilire un contatto stabile con la cellula ospite [Cascales e Christie, 2003].



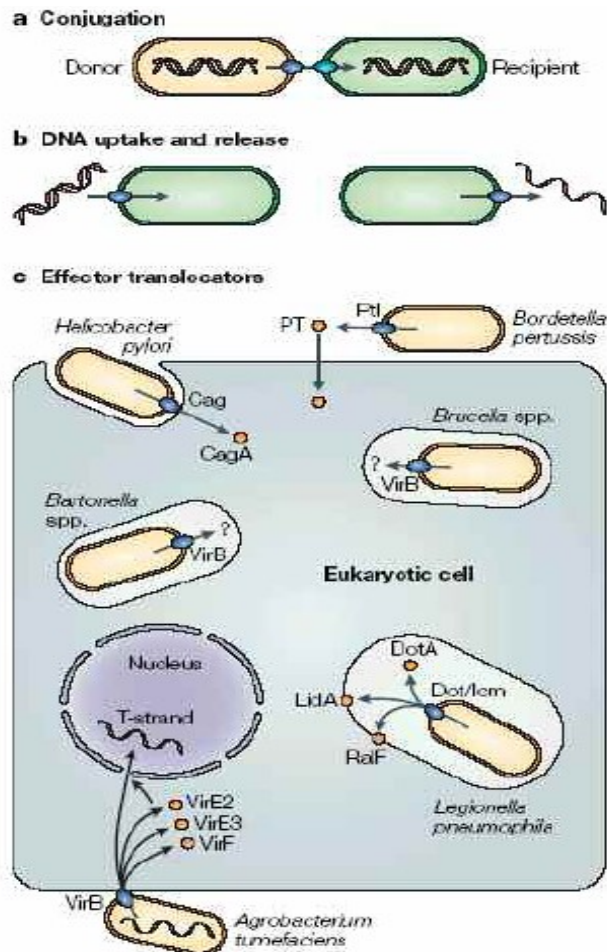


Figura 2.7 Rappresentazione dei vari sistemi di secrezione di tipo IV [Cascales e Christie, 2003].

Tra i batteri in grado di esprimere un sistema di questo tipo, oltre a *Helicobacter pylori*, vi sono *Agrobacterium tumefaciens*, *Legionella pneumophila*, *Brucella* e altri.

I sistemi di secrezione di tipo IV dei batteri Gram-negativi sono caratterizzati dall'assemblaggio di tre tipi diversi di strutture:

- 1) un complesso omoesamerico detto "core", costituito dalle CPs (Coupling Proteins), delle proteine presenti spesso in multipla copia e legate alla membrana attraverso legami acilici con molecole di acido grasso;
- 2) un canale proteico centrale;
- 3) un pilo di coniugazione, denominato pilo T.

Il complesso costituito dalle CPs serve per il reclutamento del DNA tramite l'interazione con diverse proteine, alcune delle quali costituiscono il complesso DTR (DNA Transfer and Replication) che agisce sull'origine di trasferimento (*oriT*) della sequenza di DNA mobile e la processa in modo da ottenerne un singolo filamento (chiamato filamento T). Una delle

proteine del complesso DTR, la relaxasi, rimane legata all'estremità 5' del filamento T, determinando un complesso DNA-proteina che funge da substrato per il trasferimento [Christie, 2004].

Le proteine CP instaurano specifiche interazioni con le proteine del complesso MPF (Mating Pore Formation), le quali determinano l'assemblaggio di una struttura essenziale per la creazione del pilo T e del complesso trans-envelope. La funzione di questo complesso è quella di consentire il passaggio del filamento T o delle proteine effettrici reclutate dalle CP attraverso le membrane batteriche.

Il complesso MPF più caratterizzato e che funge da modello generale anche per altri batteri è il sistema delle proteine VirB di *Agrobacterium tumefaciens*. Queste proteine vengono distinte in tre classi, sulla base della loro funzione:

1. Le proteine strutturali che costituiscono il canale, di membrana interna (VirB6, VirB8 e VirB10) e di membrana esterna (VirB3, VirB7 e VirB9);
2. Due ATPasi, VirB4 e VirB11, localizzate sulla faccia citoplasmatica della membrana interna;
3. La subunità VirB2, che assieme a VirB5 e VirB7, forma il pilo T.

La formazione del pilo è un processo fondamentale per il trasferimento di molecole; tuttavia recenti studi hanno delineato una diversa caratterizzazione di questa struttura. Sono state, infatti, eseguiti degli esperimenti di mutagenesi mirata nelle subunità VirB11 e VirB6 di *Agrobacterium tumefaciens* in cui si è visto che non c'è una correlazione diretta tra biogenesi del pilo ed il trasferimento di DNA. Alcune mutazioni, infatti, consentono la formazione del pilo ma non il trasferimento di substrato, viceversa altre impediscono la genesi del pilo, ma non il trasferimento [Jakuboski *et al.*, 2004].

Sono stati proposti tre modelli che descrivono il trasferimento del DNA T di *Agrobacterium tumefaciens* attraverso il canale trans-envelope.

Il primo modello, quello più semplice, comprende il reclutamento del substrato da parte della Coupling protein VirD4 ed il suo trasferimento attraverso il complesso MPF. Si ritiene che VirD4, assieme a VirB11 e VirB4, inneschino delle variazioni conformazionali essenziali per il trasferimento del substrato attraverso il canale.

Il secondo modello, detto "Shoot and Pump", prevede che il sistema possa funzionare come una pompa e la funzione della VirB11 sarebbe quella di permettere l'unfolding della relaxasi e di coordinarne l'espulsione

attraverso il canale con le altre proteine del complesso MPF, per consentire alla fine il rilascio del T-DNA attraverso il canale stesso.

Il terzo modello, infine, prevede una sorta di “rimbalzo” del complesso relaxasi-DNA tra le CP e la VirB11, per questo motivo è detto modello “Ping pong”. La VirD4 recluta il complesso e trasferisce la relaxasi alla VirB11, la quale, dopo averne permesso l'unfolding, la trasferisce di nuovo alla VirD4; quest'ultima poi, tramite un processo ATP-dipendente, permette la traslocazione del complesso attraverso la membrana interna. Quest'ultimo modello, sulla base dei dati strutturali ottenuti e degli studi biochimici effettuati, è ritenuto essere il più realistico [Christie *et al.*, 2005].

Vi sono inoltre altri due modelli per spiegare il rilascio del materiale all'esterno. Il primo prevede semplicemente il passaggio diretto della molecola attraverso il pilo, il secondo invece prevede che il substrato venga trasportato fino alla base del pilo, il quale, estrudendosi dalla membrana esterna, trasporterebbe il substrato nella cellula bersaglio.

Nella fig. 2.8 sono schematizzati i vari modelli descritti.

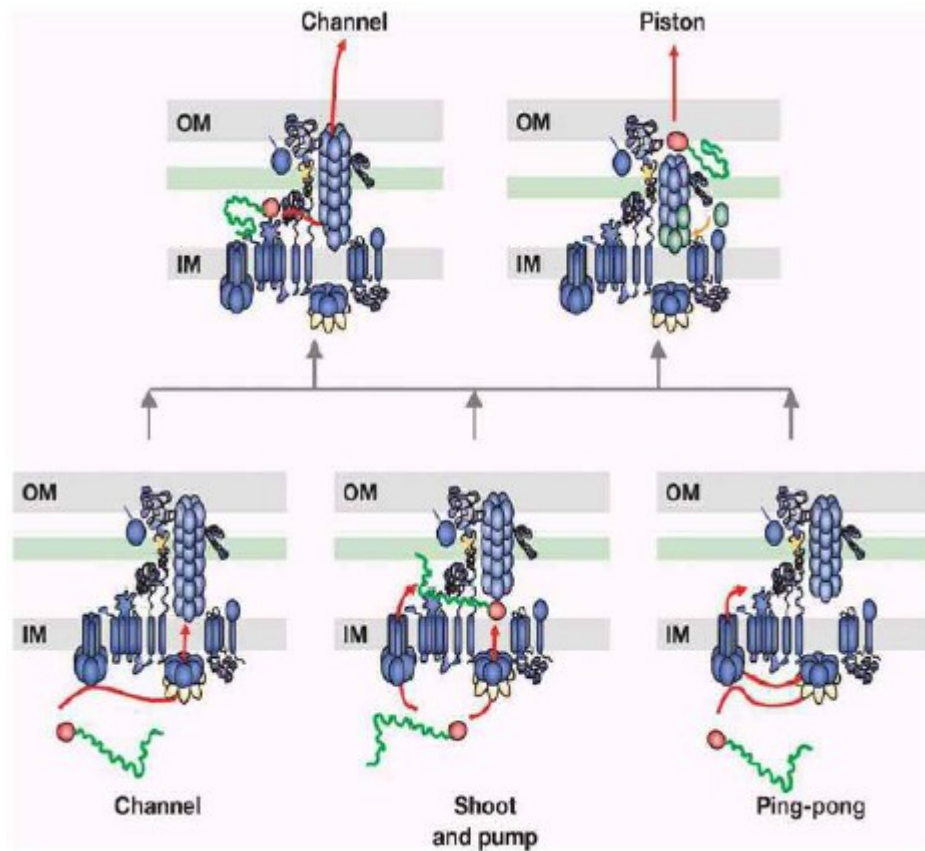


Figura 2.8 Rappresentazione dei modelli di trasferimento e rilascio del T-DNA [Christie *et al.*, 2005].

## 2.5 L'isola di patogenicità

L'isola di patogenicità di *Helicobacter pylori* (cag-PAI) è un frammento di DNA lungo circa 40 Kb, contenente 27 geni putativi. Come accennato precedentemente, tale sequenza presenta un diverso contenuto in C + G rispetto al resto del genoma, pertanto si suppone che essa sia stata acquisita per trasferimento genico orizzontale da un microorganismo sconosciuto (presumibilmente estinto) e integrata a livello delle ultime 31 pb del gene codificante l'enzima glutammato racemasi [Tomb et al., 1997]. I 27 geni della cag-PAI codificano per altrettante proteine putative vedi fig. 2.9), alcune delle quali presentano una omologia di sequenza con alcune proteine di *Agrobacterium tumefaciens* che codificano per il sistema di secrezione di tipo IV descritto nel paragrafo precedente.

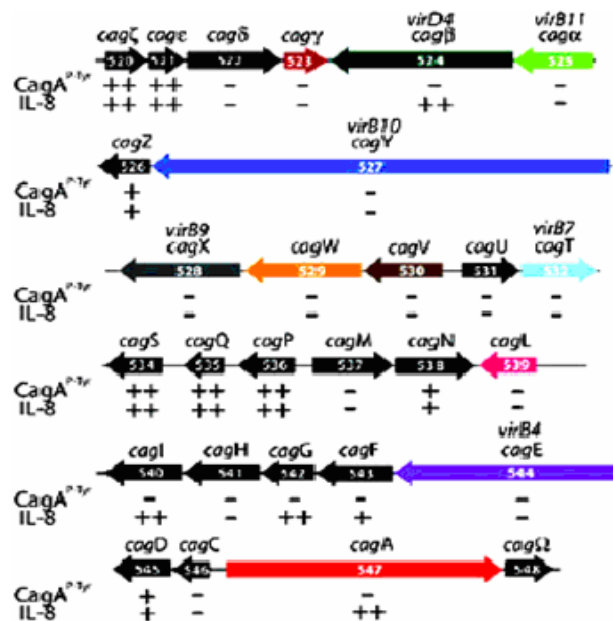


Figura 2.9 Raffigurazione schematica dell'isola di patogenicità cag di *Helicobacter pylori*, che mostra la dimensione dei geni e l'orientazione. Sono rappresentati anche gli effetti (+ o -) delle mutazioni sulla traslocazione di CagA e sull'induzione di IL-8 [Bourzac e Guillemin, 2005].

Le proteine codificate dalla cag-PAI di *H. pylori* vanno a costituire un sistema omologo (vedi fig. 2.10), impiegato in diversi processi, tra cui la traslocazione di CagA nel citoplasma delle cellule epiteliali dell'ospite e l'induzione della secrezione di interleuchina-8 (IL-8) da parte delle cellule dello stomaco. Queste ultime funzioni si riscontrano solamente nei ceppi Hp cag+ (cioè quelli che possiedono la sequenza di 40 KDa codificante per il sistema di secrezione), mentre sono assenti nei ceppi cag - (privi

della sequenza), inoltre, esperimenti di mutagenesi o delezione dei geni dell'isola di patogenicità determinano la perdita di queste funzioni. In particolare, con questi esperimenti è stato dimostrato che 17 geni sono essenziali per la traslocazione di CagA nella cellula ospite, mentre 14 sono fondamentali per indurre la trascrizione di IL-8 [Fischer et al., 2001].

A partire dal modello dell'apparato di secrezione di *Agrobacterium tumefaciens*, è stato possibile individuare le proteine omologhe in struttura e/o funzione nel sistema di secrezione *cag* di *H. pylori*. Tra elementi comuni più importanti vi sono le ATPasi della membrana interna VirB4 (CagE) e VirB11 (CagA), che forniscono energia per la traslocazione del substrato e la proteina accoppiante VirD4 (Cagb). Altre proteine con omologia funzionale sono CagY, che possiede attività di transglicosilasi e CagL, ritenuta essere un chaperone, richiesto per la maturazione delle proteine CagT e CagY [Zahl et al., 2005].

Tra i due sistemi di secrezione si possono tuttavia evidenziare alcune differenze; per esempio CagC, che in *H. pylori* è localizzata in superficie, possiede omologie di sequenza e strutturali con VirB2, che è invece il principale componente del pilo di *Agrobacterium tumefaciens*. Inoltre VirB7 e VirB9 sono proteine localizzate sulla membrana esterna, mentre le rispettive omologhe di *Helicobacter pylori*, CagT e CagX, sono distribuite lungo il corpo del pilo. CagY, la proteina omologa di VirB10, che in *Agrobacterium tumefaciens* è situata nella membrana interna, in *Helicobacter pylori* potrebbe invece costituire il rivestimento del pilo [Andrzejewska et al., 2006].

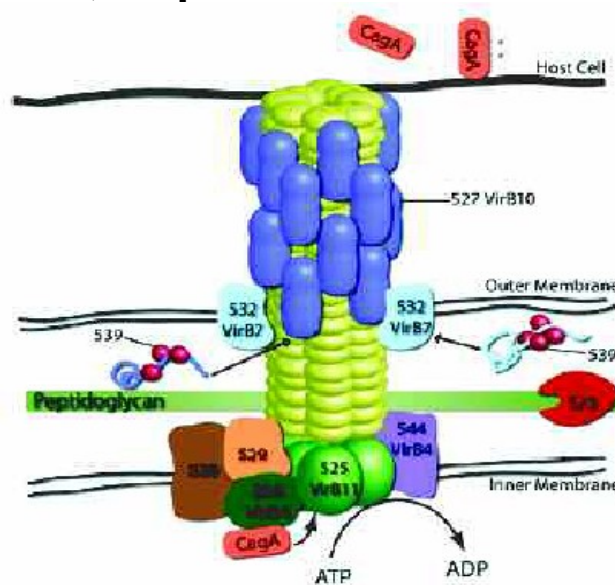


Figura 2.10 Modello ipotetico dell'isola di patogenicità *cag* di *H. pylori* basato sull'omologia con le proteine Vir di *A. tumefaciens* [Bourzac et al., 2005].

Per quanto riguarda *H.pylori*, attualmente sono note le strutture tridimensionali di sole quattro proteine, CagA [Yeo et al., 2000], CagD, CagS [Cendron et al., 2007] e CagZ [Cendron et al., 2004], tutte risolte con il metodo della diffrazione anomala (MAD), dopo aver ottenuto i cristalli contenenti Seleno-metionine delle stesse proteine. Altre proteine, quali CagN, CagL e CagF, sono state espresse e purificate, però ancora la struttura non è stata risolta. Le rimanenti sono state quasi tutte clonate ed espresse, ma non è stato possibile purificarle per varie ragioni, tra cui una bassa solubilità in acqua oppure una certa instabilità che le rende suscettibili di degradazione o di formazione di aggregati molecolari.

## 2.6 La proteina CagV

Lo scopo di questa tesi è la purificazione e caratterizzazione, a livello strutturale, della proteina CagV (HP0530). Poiché in precedenza gli svariati tentativi di cristallizzazione di questa proteina non hanno avuto buon esito, si è pensato di clonare e tentare di cristallizzare due frammenti di essa aventi lunghezza diversa, precisamente gli intervalli 59-252 (CagVn1) e 90-252 (CagVn2), entrambi corrispondenti alla parte periplasmatica e solubile, privi quindi di un tratto N-terminale più o meno lungo.

Utilizzando i programmi disponibili on-line nella banca dati EXPASY, è possibile ricavare alcuni dati riguardanti la proteina a partire dalla sequenza nucleotidica del frammento di DNA che la codifica. Attraverso il programma *ProtParam*, ad esempio, è possibile ottenere alcune caratteristiche essenziali come la composizione atomica, il punto isoelettrico, il coefficiente di estinzione molare ed il peso molecolare.

Di seguito è riportata la sequenza della proteina CagV intera, comprensiva del tratto idrofobico transmembrana:

```
N-MLGKKNEEVLVDENLVGGVIALDRLAKLNKANRTFKRAFYLSMALNVAAVTSIVMMPLKKTDIFVYGID
RYTGEFKIVKRSDARQIVNSEAVVDSATSKFVSLLLFGYSKNSLRDRKDQLMQYCDVSFQTQAMRMFNENI
RQFVDKVRAEAISSNIQREKVKNSPLTRLTFFITIKITPDTMENYEYITKKQVTIYYDFARGNSSQENLIINPFGFKV
FDIQITDLQNEQTVSEILRKIKEVESKNKALKQ-C
```

Nella tabella seguente, invece, sono riportate alcune caratteristiche della proteina intera e delle due forme più brevi della stessa:

<b>Proteina</b>	<b>CagV</b>	<b>CagVn1</b>	<b>CagVn2</b>
Numero Aa	252	227	163
Peso molecolare	29066.6	22574.8	19031.7
Punto isoelettrico	9.52	9.39	9.12
Coeff. estinzione molare [1/(M*cm)]	13410	13410	8940
Coeff. Estinzione [1/cm]	0.461	0.507	0.470

**Tabella 2.1** Sono elencate alcune caratteristiche chimico-fisiche della proteina CagV.

Eseguendo un allineamento di sequenza primaria, è risultato che CagV presenta dei domini omologhi a quelli della proteina VirB8 di *Agrobacterium tumefaciens*, nonché diversi livelli di omologia di sequenza con altre proteine appartenenti a sistemi di secrezione di alcune altre specie. Inoltre è stato dimostrato che CagV è un omologo funzionale della stessa VirB8 [Busler *et al.*, 2006]. Quest'ultima è localizzata a livello della membrana interna, con un corto dominio citoplasmatico ammino-terminale, un'elica transmembrana ed un lungo dominio periplasmatico carbossi-terminale. Da un punto di vista strutturale la proteina si presenta sotto forma di omodimero in cui ogni monomero è costituito da cinque  $\alpha$ -eliche e da quattro foglietti  $\beta$  antiparalleli. La struttura della molecola è stata risolta attraverso la tecnica della sostituzione molecolare partendo da una struttura simile depositata precedentemente, la VirB8 di *Brucella suis*.

Essendo una proteina transmembrana, è stato cristallizzato soltanto il frammento di proteina comprendente l'intervallo aminoacidico 92-231, cioè il tratto periplasmatico, e in questo lavoro di tesi è stata adottata una strategia analoga (le proteine di membrana molto raramente cristallizzano). Nella figura a pagina seguente è riportata la struttura tridimensionale di VirB8 di *Agrobacterium tumefaciens*.

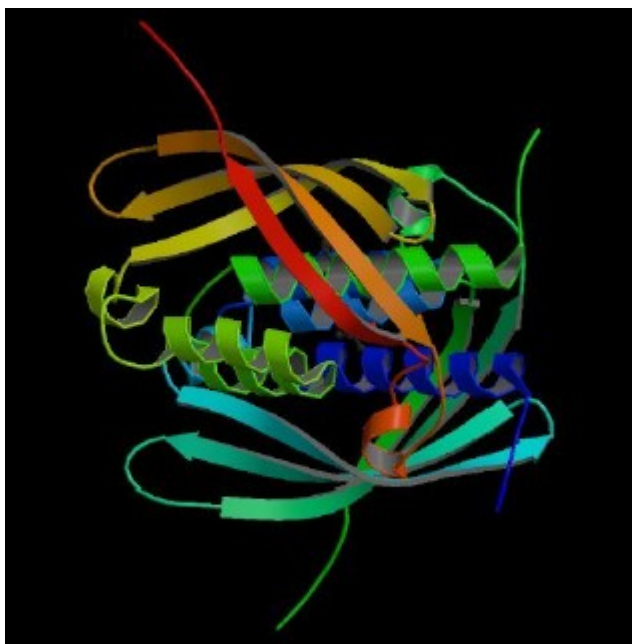


Figura 2.11 Modello di struttura 3D di VirB8 di *Agrobacterium tumefaciens* [PDB ID 2CC3].

## 2.7 Le interazioni tra le proteine CAG

Oltre a cercare di caratterizzare quelle proteine di cag-PAI di cui ancora non si sa la struttura, alcuni gruppi di ricerca stanno intraprendendo degli studi di interazione proteina-proteina per cercare di capire se ci sono proteine che interagiscono tra di loro e se eventualmente si formano dei complessi stabili tra le varie Cag.

In un recente lavoro [Busler V.J., 2006], per effettuare questo tipo di indagini sono state impiegate due metodologie differenti. La prima riguarda il saggio del doppio ibrido applicato in ceppi di *Saccharomyces Cerevisiae*, l'altra invece utilizza saggi biochimici per confermare i risultati ottenuti con la tecnica precedente.

Il saggio del doppio ibrido prevede l'utilizzo di due plasmidi codificanti l'uno per il DNA-binding domain (BD) e l'altro per l'attivatore trascrizionale (AD) del fattore di trascrizione di lievito Gal4. I due domini determinano il funzionamento di Gal4 solo se sono tenuti vicini dall'interazione di due proteine coespresse in lievito, le cui ORF sono clonate appunto nei due plasmidi. Se i due domini sono vicini, Gal4 attiva la trascrizione di vari gene reporter (tra cui HIS3, coinvolto nella biosintesi dell'istidina, o LacZ, che codifica per la galattosidasi), che consentono di confermare l'avvenuta interazione tra le due proteine nonché l'intensità di tale interazione.



Per valutare le potenziali interazioni tra varie Cag, alcuni dei geni che le codificano sono stati clonati nei due plasmidi (pBD e pAD) secondo differenti combinazioni.

Come risultato del saggio, sono state rilevate quattro interazioni omotipiche (cioè due proteine uguali interagenti tra loro) e 36 interazioni eterotipiche (tra due proteine differenti). In tabella 2.2 sono elencate tutte le interazioni riscontrate.

Dal momento che la tecnica del doppio ibrido può generare falsi positivi o falsi negativi, sono stati eseguiti anche degli studi biochimici di controllo. Alcune proteine sono state espresse in ceppi di *Escherichia Coli* competenti fuse o con GST (Glutathione S Transferasi) o con MBP (Maltose Binding Protein). Precisamente, Cag3 e CagZ sono state espresse come proteine di fusione con GST, mentre CagX, CagV, CagS, CagM e CagI sono state espresse con MBP. I lisati batterici sono stati poi mescolati e le proteine con GST sono state quindi purificate. Per evidenziare potenziali interazioni, i campioni derivanti dalla purificazione sono stati analizzati mediante immunoblotting con anticorpi anti-MBP e si è riscontrato che CagV e CagM interagiscono con Cag3, mentre CagX, CagS, CagI, CagV e CagM interagiscono con CagZ.

I risultati di questi studi biochimici indicano altresì che le interazioni rilevate con la tecnica del doppio ibrido non corrispondono a falsi positivi.

Interacting protein partners <sup>b</sup>		β-Galactosidase activity <sup>c</sup>
GAL4AD-	GAL4BD-	
Cag3(B)	Cag3(B)	++
Cag3(B)	CagT/12	++
Cag3(B)	CagM/16	+
Cag3(B)	CagG/21	++
Cag5	Cag5	++
Cag5	CagZ/6	+
Cag5	CagM/16	+
Cag5	CagI/19	+
Cag5	CagG/21	+
VirB11/ATPase	VirB11/ATPase	+++
VirB11/ATPase	CagY/7(B)	++
CagZ/6	CagS/13	+
CagZ/6	CagM/16	+
CagZ/6	CagI/19	++
CagZ/6	CagG/21	++
CagZ/6*	CagF/22*	++
CagY/7(A)	CagF/22	+++
CagY/7(C)	CagZ/6	+
CagY/7(C)*	CagX/8*	+++
CagY/7(C)	CagM/16	+
CagX/8	CagZ/6	+
CagX/8*	CagY/7(C)	+++
CagX/8*	CagM/16*	+
CagV/10	Cag3(A)	++
CagV/10	CagZ/6	+
CagV/10	CagT/12	++
CagV/10	CagM/16	+
CagT/12	CagZ/6	+
CagT/12*	CagM/16*	+
CagM/16*	CagX/8*	++
CagM/16*	CagT/12*	+
CagM/16	CagS/13	+
CagM/16	CagM/16	++
CagI/19	CagT/12	+
CagI/19	CagG/21	++
CagG/21	CagT/12	+
CagG/21	CagM/16	++
CagG/21	CagF/22	++
CagF/22*	CagZ/6*	+
CagF/22	CagT/12	+
CagF/22	CagM/16	+
CagE/23/PicB	CagZ/6	+
CagE/23/PicB	CagY/7(C)	+
CagE/23/PicB	CagM/16	+

**Tabella 2.2** Interazioni tra proteine Cag rilevate mediante la tecnica dei diibridi. L'intensità dell'interazione è proporzionale al numero di "+". Gli asterischi indicano invece le interazioni riscontrate con due reciproche coppie di plasmidi [Busler V.J., 2006].

Dai dati riportati nella tabella precedente è opportuno fare alcune importanti considerazioni: sebbene alcune interazioni trovino la loro controparte in *Agrobacterium tumefaciens*, altre interazioni tra varie cag sono discordanti con il modello ipotetico del T4SS di *Helicobacter pylori*, descritto precedentemente. Ad esempio, CagV (HP0530), l'oggetto di questo lavoro di tesi, è stata vista essere in grado di interagire significativamente con CagA e CagT, mentre nel modello ipotetico di T4SS si troverebbe ancorata alla membrana interna (e CagT alla base del pilo). Questo studio è particolarmente interessante, in quanto consente di formulare nuove ipotesi sulla topologia delle proteine del T4SS di *H.pylori*

e può fungere da solida base di partenza per tutta una nuova serie di indagini.

In questo lavoro di tesi sono stati condotti ulteriori studi di interazione tra le varie proteine Cag utilizzando la tecnica del dot-blot, i dettagli della quale sono descritti nella sezione materiali e metodi.

## **3. MATERIALI E METODI**

### **3.1 Amplificazione dei frammenti di DNA mediante PCR**

La PCR (reazione a catena della polimerasi) è una tecnica rapida e versatile in grado di amplificare specifiche sequenze di DNA *in vitro*. L'amplificazione si ottiene grazie all'utilizzo di un enzima termostabile, la DNA polimerasi di *Thermus Aquaticus*, un batterio in grado di sopravvivere a temperature abbastanza elevate.

La reazione avviene in tre momenti distinti che, reiterati un certo numero di volte (in genere da 25 a 35), costituiscono quello che viene detto "ciclo di PCR". Queste tre fasi sono:

1. denaturazione del DNA, a 90-95°C, allo scopo di separare i due filamenti antiparalleli rompendo i ponti a idrogeno tra basi complementari;
2. annealing, cioè ibridizzazione dei primers agli estremi della sequenza da amplificare; avviene ad una temperatura variabile a seconda della lunghezza e della temperatura di melting della sequenza;
3. elongazione, catalizzata dalla DNA polimerasi, del filamento di DNA di nuova sintesi a partire dai monomeri dNTPs (deossinucleotidi trifosfato).

Al fine di completare la polimerizzazione di tutte le catene, segue una fase finale, detta terminazione, in cui la miscela di reazione viene mantenuta più a lungo a 72°C. L'amplificazione avviene in un termociclatore, un apparecchio in grado di cambiare in modo rapido e programmato la temperatura, affinché possano avvenire nei tempi precisi le 3 fasi citate precedentemente. La soluzione di reazione, oltre al DNA da amplificare, la DNA polimerasi e i primers, contiene un tampone per mantenere il pH costante, una miscela dei 4 dNTPs e sali contenenti  $Mg^{2+}$ , che è il cofattore della DNA polimerasi.

### **3.2 Elettroforesi di macromolecole**

L'elettroforesi è una tecnica analitica che permette la separazione di macromolecole (DNA e proteine) sulla base della loro carica e dimensione. Essa permette di verificare la presenza o meno di una molecola di interesse, l'eventuale degradazione della stessa ed il grado di purezza nelle varie fasi di purificazione.

La separazione delle molecole si basa sulla migrazione differenziale di specie elettricamente cariche sotto l'influenza di un campo elettrico. Una molecola avente carica netta  $q$ , posta in un campo elettrico  $E$ , risente di una forza elettrica proporzionale a  $qE$ . Muovendosi in soluzione, la molecola risente anche di una forza frizionale, avente uguale direzione ma verso opposto, la quale tende ad ostacolarne il moto ed è proporzionale alle dimensioni e alla forma della molecola, nonché alla viscosità del fluido. La molecola si muoverà quindi con una velocità proporzionale al rapporto carica su raggio idrodinamico.

Le proprietà del solvente (come pH, forza ionica) hanno una rilevante importanza, poiché influenzano la carica effettiva della proteina, la sua forma e dimensione idrodinamica.

## ELETTROFORESI SU GEL DI AGAROSIO

Per la separazione di sequenze nucleotidiche si utilizza una matrice a base di agarosio, che presenta una serie di pori di diametro adeguato alla loro dimensione. Prima della solidificazione del gel, sulla matrice si producono, mediante un pettine, dei pozzetti all'interno dei quali vengono caricati i vari frammenti di DNA da far correre. Applicando un campo elettrico, è possibile far migrare i campioni attraverso la matrice grazie alla carica netta assunta. Poiché il DNA è carico negativamente a causa della presenza dei gruppi fosfato, esso migra attraverso il gel dal catodo all'anodo e i frammenti vengono separati a seconda della loro lunghezza. Però solo alcuni frammenti, quelli più piccoli, riescono ad attraversare i pori della matrice e quindi a localizzarsi sulla parte inferiore del gel (a livello dell'anodo), mentre i frammenti più grandi migrano meno e si andranno a localizzare sulla parte superiore.

A seconda del numero di paia di basi dei frammenti da separare, si modula leggermente la concentrazione dell'agarosio (di solito tra 0.8 % e 2%) al fine di variare la dimensione dei pori della matrice. Il quantitativo desiderato di agarosio viene sciolto in 100 mL di tampone TAE, costituito da 40 mM tris base, 0.11% acido acetico glaciale, 2 mM EDTA pH 8.5, e posto a scaldare in forno a microonde fino alla completa solubilizzazione dell'agarosio (che è solubile solo in acqua bollente). Dopo aver fatto raffreddare la soluzione, questa viene versata in un supporto munito di uno o più pettini per formare i pozzetti e viene aggiunta una soluzione di bromuro di etidio, fino ad una concentrazione finale di 0.5 µg/ml. Il

bromuro di etidio è una sostanza in grado di intercalarsi tra le basi azotate del DNA, permettendone la visualizzazione ad un rilevatore a raggi UV (esso emette fluorescenza nel visibile se eccitato da una radiazione ultravioletta). I campioni da caricare su gel vengono preparati mescolando il DNA ad un tampone di caricamento costituito da 40% di saccarosio, 0.25% di xilene cianolo FT e 0.25% blu di bromofenolo, allo scopo di visualizzare la loro migrazione. Assieme ai campioni viene solitamente caricata una soluzione di frammenti di DNA di dimensioni note, detti marcatori di peso molecolare.

Una volta che i campioni sono stati caricati, si impostano, in base alla percentuale di agarosio nel gel, la lunghezza dello stesso e il numero di paia di basi dei frammenti da analizzare, alcuni parametri, quali voltaggio, intensità di corrente e tempo, quindi si avvia la corsa. Una volta che quest'ultima è avvenuta, si visualizza il gel tramite un transilluminatore.

#### ELETTROFORESI SU GEL DI POLIACRILAMMIDE (SDS-PAGE)

Di solito per separare le proteine viene utilizzato come supporto un gel di poliacrilammide, una sostanza ottenuta grazie alla polimerizzazione radicalica di acrilammide (monomero) e N,N'-metilenbisacrilammide (cross-linking agent). La polimerizzazione viene iniziata e mantenuta dalla coppia TEMED (N,N,N',N'-tetrametilendiammina) ed APS (ammonio persolfato). Nella figura riportata di seguito è illustrata la suddetta reazione:

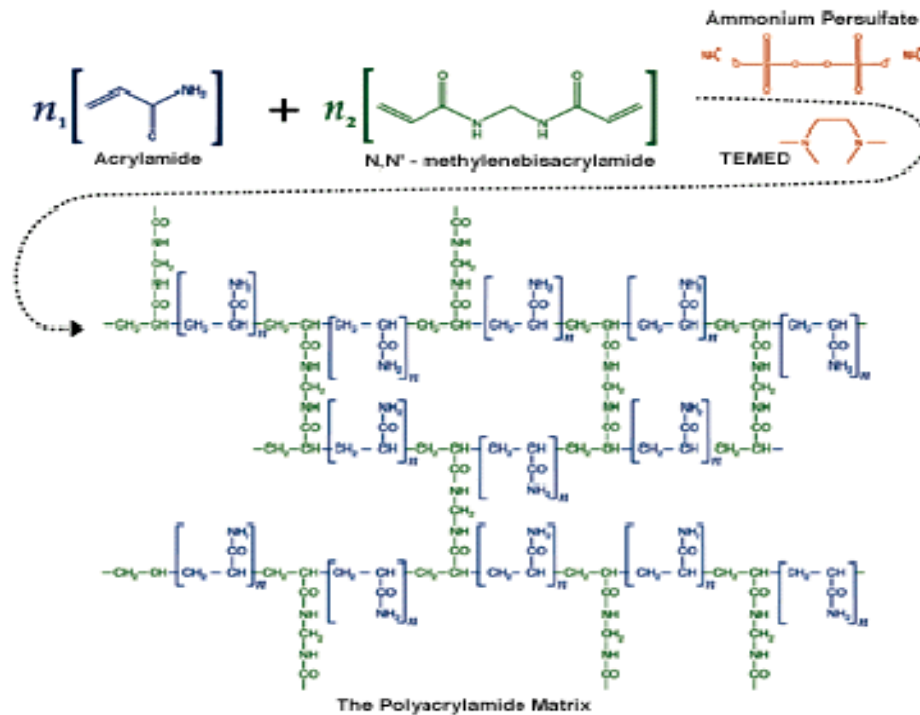


Figura 3.1 Reazione di polimerizzazione dell'acrilammide.

Il gel viene fatto polimerizzare sotto forma di film sottile tra due lastre di vetro, distanziate da opportuni spaziatori ed è costituito da due parti:

1. gel d'impaccamento ("stacking gel"), caratterizzato da una più bassa percentuale di acrilammide, in genere del 5%. In questo caso quindi le maglie sono larghe, in modo da favorire la libera migrazione delle molecole. All'estremità superiore dello stacking gel vengono formati i pozzetti, necessari al caricamento dei campioni; questo serve ad allineare e concentrare le proteine in una sottile banda prima di accedere alla parte inferiore dove avverrà la loro separazione;
2. gel di separazione ("running gel"), caratterizzato da una percentuale di acrilammide maggiore (di solito tra l'8% e il 20%), a livello del quale avviene la separazione delle proteine.

Nella tabella seguente sono elencate tutte le sostanze utilizzate per produrre un gel a diverse percentuali di poliacrilammide (quelle più utilizzate) e le relative quantità:

<b>Componenti (Quantità per 4 gel)</b>	<b>Stacking gel</b>	<b>Running gel</b>	
<b>% di acrilammide</b>	5%	18%	15%
40 % acrilammide	10.35 mL	10.55 mL	13.35 mL
H <sub>2</sub> O ultrapura	4.55 mL	16.9 mL	14.10 mL
1 M Tris (pH 6.8) (mL)	2.70 mL	/	/
1.5 M Tris (pH 8.8) (mL)	/	9.4 mL	9.375 mL
10% SDS (μL)	180 μL	375 μL	375 μL
10% APS (μl)	180 μL	375 μL	375 μL
TEMED (μl)	6 μL	15 μL	15 μL

**Tabella 3.1** Sostanze impiegate per la polimerizzazione di un gel SDS-PAGE.

Generalmente per questo tipo di indagine analitica si fa uso del detergente SDS (sodio dodecilsolfato), una sostanza anionica in grado di denaturare e legare le proteine, conferendo loro una carica netta negativa proporzionale alla massa. In questo modo si rende trascurabile la carica effettiva delle catene polipeptidiche. Siccome tutte le proteine sono denaturate e presentano lo stesso rapporto carica/massa, la mobilità delle molecole dipende unicamente dal peso molecolare e della porosità del gel, che funge da setaccio molecolare.

I campioni, prima di essere caricati su gel, devono essere completamente denaturati, per questo vengono bolliti a 100 °C per almeno 5 minuti nel "sample buffer". Quest'ultimo è dato da una soluzione tamponata (60 mM Tris, pH 6.8), un agente denaturante e complessante (2% SDS), un riducente ( $\beta$ -mercaptoetanololo o DTT) per rompere eventuali ponti disolfuro presenti, un colorante tracciante attraverso cui si possa seguire il fronte della corsa elettroforetica (0.1% blu di bromofenolo) e 10% di glicerolo.

Per poter risalire al peso molecolare delle proteine da analizzare, un pozzetto del gel viene caricato con una miscela di proteine a pesi molecolari noti e viene fatto migrare insieme agli altri campioni.

Il gel viene inserito in posizione verticale su un'apposita struttura, in modo che sia a contatto con una soluzione tampone chiamata "running buffer" (tampone di corsa) contenente: 0.2 M Glicina, 25 mM Tris e 0.1% SDS. La migrazione dei campioni viene ottenuta applicando una corrente elettrica compresa tra i 30 e i 45 mA fino all'uscita del tracciante dal gel.

La mobilità delle catene polipeptidiche è direttamente proporzionale al logaritmo della loro massa. Le proteine più piccole si muoveranno più rapidamente attraverso il gel, mentre quelle più grandi avranno più



difficoltà a migrare, ritrovandosi così a fine corsa più vicino ai pozzetti di caricamento.

A fine corsa, per osservare l'avvenuta separazione delle proteine, si immerge il gel in una soluzione acquosa di colorante: 0.1 % (w/v) Coomassie Brilliant Blue, 40 % (v/v) etanolo, 10% (v/v) acido acetico. La miscela acido-etanolo ha lo scopo di far precipitare le proteine su gel e fissarvi il colorante aggiunto. Infine si prosegue con la decolorazione del gel, al fine di allontanare il colorante in eccesso ed evidenziare solo quello rimasto legato alle proteine. Il limite di rilevabilità del colorante utilizzato è di circa 0.1 µg di proteina.

### **3.3 Purificazione dei prodotti di PCR**

Dopo aver fatto correre il DNA tramite elettroforesi su gel di agarosio, il frammento amplificato mediante PCR deve essere purificato per poter successivamente esser clonato all'interno di un vettore.

La purificazione è una fase importante ed ha lo scopo di eliminare primers, deossinucleotidi residui, sali e prodotti aspecifici.

Nel presente lavoro per la purificazione ci si è avvalsi della metodica denominata "gel extraction", utilizzando il kit "QIAEX II" della Qiagen.

Il metodo consiste nell'effettuare una purificazione del DNA direttamente dalla banda presente sul gel di agarosio, tagliandola con un bisturi dopo aver posto il gel su un transilluminatore a raggi UV.

Il protocollo seguito è il seguente:

- pesare un tubo da 2 mL vuoto, inserirvi la bande e pesare di nuovo il tutto per sapere il peso netto della banda;
- aggiungere 3 volumi di buffer QX1 per ogni frammento di DNA di lunghezza 100bp-4Kb;
- agitare con un vortex per 30 secondi e incubare a 50°C per solubilizzare completamente il gel di agarosio;
- centrifugare il campione per 30 secondi e rimuovere il surnatante con una pipetta;
- lavare il pellet con 500µL di buffer QX1;
- lavare il pellet con 500 µL di buffer PE;
- lasciar seccare all'aria il pellet per 10-15 minuti;
- per eluire il DNA, aggiungere 200 µL di soluzione 10 mM Tris-HCl pH 8.5 e risospendere il pellet tramite il vortex; incubare 5-10 minuti a 50°C;

- centrifugare per 30 secondi, quindi trasferire il supernatante, che contiene il DNA purificato, in un tubo pulito;  
(Gli ultimi due passaggi possono essere ripetuti una o più volte, al fine di aumentare la resa).

### **3.4 Clonaggio dei frammenti di DNA tramite “TOPO cloning”**

Per il clonaggio delle sequenze di DNA purificate dal gel si è adottata una nuova strategia attraverso la quale vengono inseriti, in modo direzionale, frammenti di DNA ad estremità 3' sporgenti all'interno di un vettore. Questa tecnica è denominata “TOPO cloning”, in quanto richiede, per venire attuata, l'enzima topoisomerasi I di *Vaccinia*. A differenza delle tradizionali tecniche di clonaggio, questo metodo non prevede utilizzo di DNA ligasi e di enzimi di restrizione; nonostante ciò ha il vantaggio di essere molto rapido e al contempo di avere una efficienza elevata (spesso maggiore del 90%).

I vettori utilizzati, detti generalmente “pET-TOPO”, sono disponibili in commercio con le seguenti caratteristiche:

- un promotore *T7lac* (del batteriofago T7) per una espressione ad alti livelli, indotta dall'IPTG, del gene di interesse in *Escherichia coli*;
- un sito di clonaggio direzionale dei prodotti di PCR;
- una sequenza detta “fusion tag”, N e/o C terminale, per rilevare e purificare proteine di fusione ricombinanti;
- un sito di riconoscimento per una proteasi capace di rimuovere il frammento tag N terminale;
- il gene *lacI* codificante per il repressore lac, il quale serve a ridurre la trascrizione basale del gene bersaglio, la cui espressione potrebbe condurre ad instabilità plasmidica o a morte del batterio;
- un marcatore di resistenza ad un antibiotico (kanamicina, ampicillina) per poter in seguito selezionare le trasformazioni avvenute correttamente;
- l'origine pBR322 per la replicazione a basse copie ed il mantenimento in *E. coli* del vettore.

In questo lavoro è stato utilizzato il vettore pET151, il quale presenta una sequenza fusion tag di 6 istidine (detta anche “His-tag”) N-terminale, un sito di taglio per la proteasi TEV, ed un marcatore di selezione per l'antibiotico ampicillina.

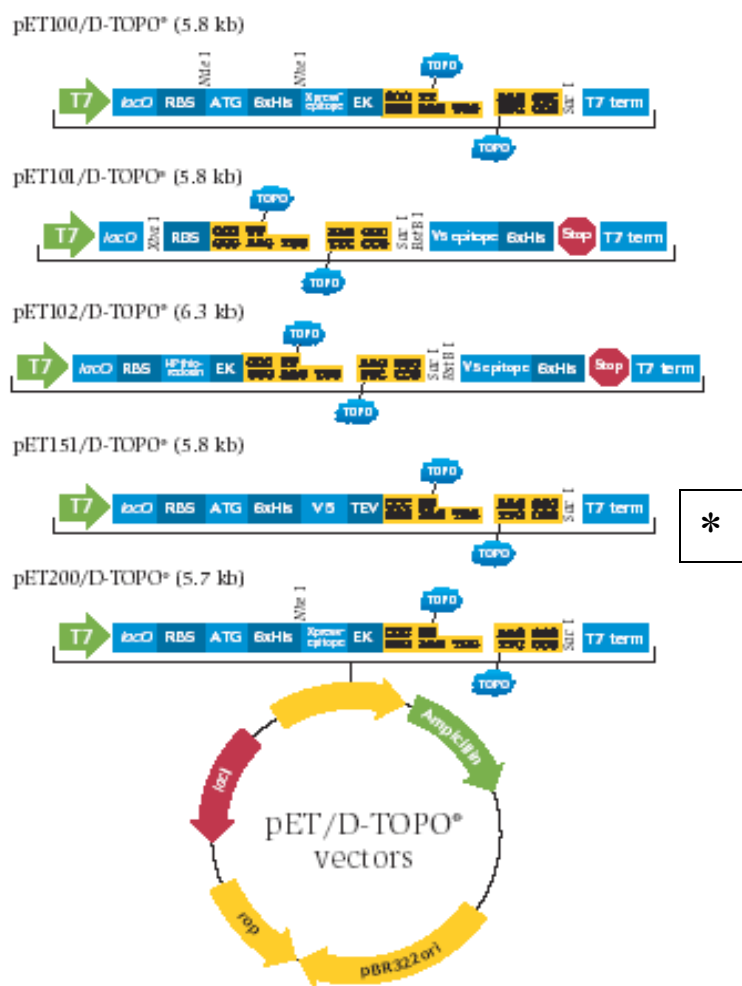


Figura 3.2 Rappresentazione schematica di vettori pET TOPO. L'asterisco indica il vettore utilizzato per clonare CagV.

La tecnica sfrutta il fatto che la topoisomerasi I di *Vaccinia* si lega al DNA in doppia elica del vettore a livello di siti specifici e taglia il legame fosfodiesterico dopo la sequenza 5'-CCCTT (vedi fig. 3.3). L'energia derivante dalla scissione del legame viene conservata attraverso il legame covalente che si instaura tra il fosfato al 3' del filamento interrotto ed il residuo Tyr-274 della topoisomerasi. Il legame fosfotirosinico viene quindi attaccato dal gruppo 5' ossidrilico del frammento prodotto a seguito del taglio, consentendo il rilascio della topoisomerasi e rendendo pertanto la reazione reversibile.

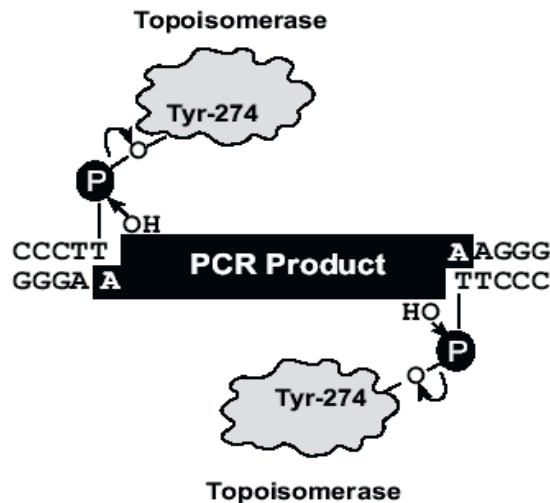
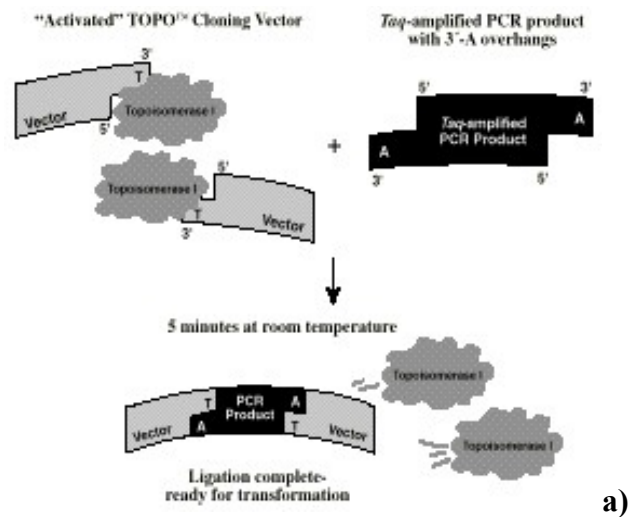


Figura 3.3 a) Reazione di ligazione del prodotto di PCR al vettore;

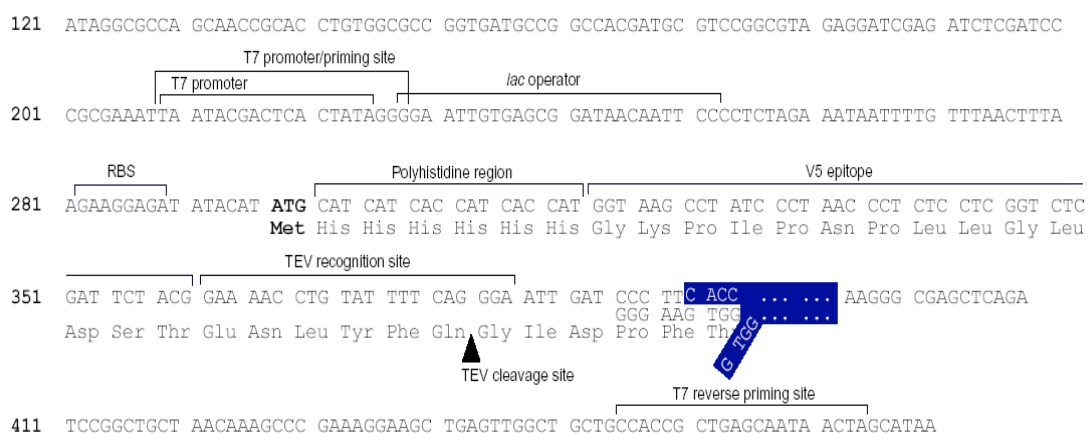
b) Meccanismo con cui avviene il legame della topoisomerasi al DNA.

Il clonaggio direzionale avviene aggiungendo all'estremità 3' del frammento che si vuol clonare un tratto di DNA a singolo filamento che è identico all'estremità 5' del "charged TOPO cloning DNA" (cioè il DNA del vettore attivato, legato alla topoisomerasi). Al DNA del vettore è stata aggiunta una sequenza sporgente di 4 oligonucleotidi (GTGG) e i prodotti di PCR sono clonati direzionalmente aggiungendo al primer forward 4 basi ad essa complementari (CACC). La sequenza GTGG nel vettore di clonaggio invade l'estremità 5' del prodotto di PCR, si appaia alle basi aggiunte, e stabilizza il prodotto di PCR nella corretta orientazione.

Un passaggio critico, quindi, è quello relativo alla costruzione dei primers di PCR. Il primer forward, oltre alla sequenza CACC all'estremità 5', deve avere un codone di stop per la terminazione del peptide all'estremità amminica e due siti di legame al ribosoma (RBS):

1. a monte del codone d'inizio (ATG) nel tag N terminale per assicurare una spaziatura ottimale per consentire una traduzione appropriata;
2. la sequenza consenso AGGAGG 9-10 bp al 5' rispetto al codone d'inizio.

Per quanto riguarda il primer reverse, bisogna viceversa far in modo che non presenti complementarietà con la sequenza GTGG al 5' (altrimenti il clonaggio direzionale non sarebbe più garantito), inoltre deve essere incluso anche in questo caso un codone di stop.



**Figura 3.4** Rappresentazione della mappa del vettore pET151/TOPO; è evidenziato il sito in cui avviene la rottura del legame fosfodiesterico da parte della topoisomerasi.

Dopo aver eseguito l'amplificazione del DNA e verificato la buona riuscita della stessa tramite corsa degli amplificati su gel di agarosio, si procede con la reazione di clonazione. Nella tabella seguente sono riportati i materiali utilizzati e le rispettive quantità:

Materiale	Quantità
Prodotto di PCR	1µL
Soluzione salina	1µL
Acqua sterile	4µL
Vettore TOPO	1µL

**Tabella 3.2** In tabella è elencata la mix di reazione per una reazione di "TOPO cloning".

Una volta preparata la miscela, si incuba il tutto a temperatura ambiente per 5 minuti, poi si trasferisce il tutto in ghiaccio e successivamente si procede alla trasformazione dei vettori contenenti il gene d'interesse in *E.coli*.

### 3.5 Trasformazione

La trasformazione è un processo che avviene spontaneamente in natura, scoperto da Griffith agli inizi del secolo scorso, consistente nell'inserimento di un plasmide contenente uno specifico frammento di DNA all'interno di una cellula (batterio). A livello sperimentale, per aumentare l'efficienza del processo, si trattano i batteri (*E.coli*) da utilizzare per la trasformazione con soluzioni di  $\text{CaCl}_2$  e di  $\text{MgCl}_2$ . Questi sali sono in grado di destabilizzare le membrane batteriche e renderle più permeabili, in modo da facilitare l'entrata del DNA nel citoplasma. Batteri trattati con queste soluzioni saline sono detti "competenti" e possono acquisire DNA con una resa di circa 1%.

Per la preparazione di cellule competenti è necessario innanzitutto coltivarle in terreno liquido fino ad un determinato livello di crescita, che viene monitorato mediante la misura, per via spettrofotometrica, della dispersione della luce da parte dei batteri (densità ottica, "OD") a 600 nm, la quale è direttamente proporzionale al numero di cellule in sospensione. Il protocollo utilizzato per rendere un determinato ceppo batterico competente è riassunto in questi punti:

1. Si inoculano i batteri in una beuta contenente 100 ml di terreno liquido privo di antibiotici;
2. Si incubano a 37°C sotto agitazione (210 rpm) fino al raggiungimento di una densità ottica di 0.94-0.95;
3. Si centrifugano i batteri a 10000 giri per 10 minuti per far precipitare il pellet; si risospende il pellet in 10 mL di una soluzione fredda contenente 80 mM  $\text{CaCl}_2$  e 50 mM  $\text{MgCl}_2$  e si mantiene in ghiaccio per 10 minuti;
4. Si ripete il punto precedente per 2 volte;
5. Si risospende il pellet in una soluzione fredda di 0.1 M  $\text{CaCl}_2$ ;
6. Si mescola la sospensione con un uguale volume di una soluzione di glicerolo 50%.

La sospensione così trattata viene suddivisa in aliquote in tubi da 1.5 mL e le varie frazioni vengono conservate a -80°C. Per ricavare la quantità di

cellule dalla densità ottica misurata si applica la seguente proporzione, valida se il cammino ottico è di 1 cm:

$$1 : 10^9 \text{ cellule} = \text{OD misurata} : X$$

I batteri competenti vengono poi trasformati con il plasmide contenente l'insero di interesse tramite shock termico, processo che sfrutta il fatto che le membrane cellulari, passando da un ambiente freddo (ghiaccio) ad uno molto più caldo, si destabilizzano favorendo così l'incorporazione del plasmide all'interno delle cellule.

Il protocollo che si utilizza è riassunto nei seguenti punti:

- si aggiunge un determinato volume di plasmide (1  $\mu\text{L}$ ) ad una aliquota di batteri competenti (80  $\mu\text{L}$ ) e si incuba in ghiaccio per almeno 20 minuti;
- i batteri vengono trasferiti rapidamente in un bagnetto termostatico e incubati a 42°C per 1 minuto;
- si aggiungono 400  $\mu\text{L}$  di terreno liquido e si incuba il tutto in un agitatore a 200 rpm per circa 45 minuti;
- si piastrano i batteri su piastra petri con terreno solido contenente l'antibiotico (quello specifico della resistenza presente nel plasmide);
- si incuba la piastra a 37°C per una notte.

Sono stati utilizzati esclusivamente ceppi di *Escherichia coli* BL21(DE3) B F<sup>-</sup> ompT hsdS<sub>B</sub> (r<sub>B</sub><sup>-</sup>m<sub>B</sub><sup>-</sup>) gal dcm (DE3). Questo ceppo batterico contiene il fago lisogeno DE3, in cui è presente un gene codificante per la T7 RNA polimerasi sotto controllo del promotore lacUV5. Inoltre possiede pochissime proteasi, per limitare eventuali effetti di degradazione della proteina ricombinante.

Le composizioni dei terreni invece sono riportate di seguito:

- Terreno liquido: brodo Luria-Bertani (LB) composto da 10 g/l Triptone Peptone, 5 g/l Estratto di lievito, 10 g/l NaCl;
- Terreno solido: Luria-Bertani agar (LBA) composto da 10 g/l Triptone Peptone, 5 g/l Estratto di lievito, 10 g/l NaCl, 15 g/l Bacto Agar.

Tutti i terreni, prima di essere utilizzati e prima dell'aggiunta eventuale dell'antibiotico, devono essere sterilizzati in autoclave a 121 °C per 20 minuti.

### 3.6 Espressione proteica

Se la trasformazione è andata a buon fine, si può procedere alla fase di espressione della proteina, che viene fatta utilizzando colture di volumi diversi a seconda dello scopo: in un piccolo volume di terreno per effettuare un controllo dell'espressione o in un maggiore volume per ottenere grandi quantità di proteina.

Per la prova di espressione si utilizzano una o due beute con 25-50 ml di terreno liquido LB con antibiotico e in ognuna si inocula una colonia di batteri trasformati con il vettore di espressione. La coltura viene incubata a 37°C sotto agitazione (180 rpm) e vengono eseguiti a intervalli regolari dei prelievi di 1-2 mL per controllare l'andamento della crescita batterica. Il volume di coltura prelevato viene letto allo spettrofotometro ad una lunghezza d'onda fissa di 600 nm e quando si è raggiunto un valore di densità ottica di circa 0.6 (corrispondente ad una crescita batterica esponenziale) si aggiunge alla stessa coltura una soluzione di IPTG per indurre l'espressione della proteina.

La coltura in grandi volumi di terreno (1 o più litri) viene eseguita per produrre alte quantità di proteina, la quale verrà poi purificata per poter condurre gli studi strutturali successivi. In questa procedura si isola una colonia di batteri trasformati con il vettore di espressione e la si striscia in altre piastre di terreno solido con antibiotico per ottenere, dopo una crescita ON a 37°C, un tappeto di batteri.

I batteri così cresciuti vengono risospesi utilizzando circa 10 mL di terreno liquido per piastra ed un determinato volume viene aggiunto in ogni beuta di terreno. Le colture vengono incubate a 37°C sotto agitazione (180 rpm) e viene monitorata a intervalli periodici la densità ottica. Una volta raggiunto un valore di 0.6, si aggiunge alla coltura un determinato volume di una soluzione di IPTG in ogni beuta per indurre l'espressione della proteina. Infine i batteri, dopo circa 4 ore di induzione, vengono centrifugati per 10 minuti a 4000 rpm e il pellet viene conservato a -80°C.



### 3.7 Lisi delle cellule batteriche

Dopo aver controllato, tramite analisi SDS-PAGE, l'espressione della proteina dai campioni prelevati prima e dopo l'induzione con IPTG, le cellule vengono lisate al fine di liberare la proteina. Per eseguire la lisi, si adopera un metodo basato sull'utilizzo di un enzima, il lisozima, e di un sonicatore, seguendo il seguente protocollo:

- aggiungere al pellet conservato a  $-80^{\circ}\text{C}$  e contenente la proteina un tampone specifico per ogni proteina (circa 20 mL per ogni litro di coltura) e mettere in agitazione finché non si ha la completa sospensione del pellet;
- aggiungere una soluzione 1mg/mL di lisozima e incubare il tutto per un'ora a  $-4^{\circ}\text{C}$ ; il lisozima è una glicosidasi che degrada le pareti di alcuni batteri rompendo il legame glicosidico tra il C-1 dell'acido N-acetilmuramico (NAM) e il C-4 dell'N-acetilglucosammina (NAG), entrambi componenti del peptidoglicano,;
- sonicare il tutto effettuando 4 cicli da 45 secondi, al fine di completare la lisi.

Una volta ottenuto il lisato grezzo, questo viene centrifugato a 25000 rpm per 20 minuti a  $4^{\circ}\text{C}$  e campioni di pellet e surnatante vengono caricati su un gel SDS-PAGE per controllare la localizzazione della proteina. Le proteine idrosolubili, correttamente foldate, si ripartiscono nel surnatante, costituito dal tampone di lisi, mentre proteine di membrana e/o quelle denaturate si ripartiscono nel pellet, costituito soprattutto da membrane cellulari e molecole idrofobiche.

### 3.8 Purificazione con tecniche cromatografiche

#### CROMATOGRAFIA DI AFFINITÀ

La cromatografia di affinità è una tecnica analitica di purificazione, che sfrutta l'affinità delle molecole per specifici ligandi. Il sistema cromatografico, che è una colonna, consiste in una matrice insolubile alla quale viene legata covalentemente, attraverso uno spaziatore opportuno, una molecola verso cui la proteina da purificare ha dimostrato avere un'interazione altamente specifica. Il principio di funzionamento della tecnica si basa sul fatto che la proteina si legherà al ligando, rimanendo in colonna mentre tutte le altre proteine che non interagiscono verranno

allontanate tramite dei lavaggi. Il legame tra il ligando e la molecola di interesse deve però essere reversibile, per far in modo che questa possa essere rimossa senza troppe difficoltà in un successivo passaggio. Si possono utilizzare a tal fine sia metodi specifici, come l'eluizione con ligandi competitivi, che metodi non specifici (variazione del pH o della forza ionica della soluzione).

La metodologia adottata per purificare *CagV* è detta "IMAC" (da Immobilized Metal Affinity Chromatography) e sfrutta l'espressione delle proteine di interesse unita ad una coda di sei istidine (His-tag). L'amminoacido istidina, contenendo nella sua struttura chimica un anello imidazolico, presenta la proprietà di legare fortemente alcuni ioni bivalenti, come il nickel. Gli ioni  $Ni^{2+}$  sono immobilizzati in una resina mediante dei linker costituiti da molecole organiche in grado di legare i metalli tramite legami di coordinazione.

Una volta che la proteina viene fatta passare in colonna, essa si lega fortemente agli ioni nickel immobilizzati tramite la coda di istidine. Per l'eluizione si utilizza lo stesso tampone con cui è stata caricata in colonna, al quale però viene aggiunto imidazolo. Quest'ultimo deve avere una elevata concentrazione, in modo da poter competere con l'istidina per il legame al nickel e favorire così l'eluizione della proteina immobilizzata. E' anche possibile eluire la proteina utilizzando un tampone con un'alta forza ionica, con un pH differente o aggiunto di Guanidinio 6M.

Talvolta per l'eluizione della proteina si effettua un gradiente di imidazolo sfruttando apparati cromatografici collegati ad un software e dotati di uno spettrofotometro capace di misurare l'aumento di assorbanza man mano che avviene l'eluizione e di un rilevatore che invia i dati al computer.

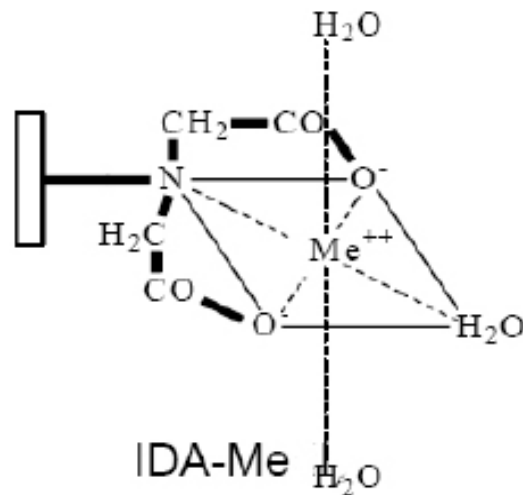


Figura 3.5 Rappresentazione del complesso che si forma mediante legami di coordinazione tra uno ione bivalente e l'acido iminodiacetico legato alla resina.

### CROMATOGRAFIA A GEL PERMEAZIONE

La cromatografia a gel filtrazione, detta anche cromatografia a esclusione dimensionale, permette di separare le molecole in base alla loro dimensione. Le resine utilizzate sono date da polimeri od oligosaccaridi legati da legami crociati (come destrano, agarosio o acrilammide) che definiscono reticoli uniformi nei quali possono diffondere sia il soluto che il solvente. La separazione delle molecole dipende dalla differente abilità, dovuta al diverso peso molecolare e al raggio di Stokes, delle varie molecole di campione di entrare nei pori che costituiscono la fase stazionaria. Le molecole di diametro superiore alla dimensione dei pori sono incapaci di penetrare nei pori della fase stazionaria e si muoveranno attraverso il letto cromatografico molto velocemente (fenomeno dell'esclusione), mentre le molecole di diametro inferiore alla dimensione dei pori trascorreranno molto più tempo nella fase stazionaria, uscendo dalla colonna a volumi di ritenzione più grandi (fenomeno della permeazione).

Il volume di ritenzione  $V_R$  è così definito:

$$V_R = V_0 + kV_i$$

dove:

$V_0$  è il volume della fase mobile al di fuori dei pori;

$V_i$  è il volume della fase mobile contenuta entro i pori della fase stazionari;  $k$  è il coefficiente di distribuzione del soluto tra la fase mobile e quella stazionaria.

Una molecola totalmente trattenuta dal gel avrà coefficiente  $k = 1$  e il volume di ritenzione corrisponderà al valore massimo supportato dalla colonna:  $V_R = V_i + V_0$  (limite di permeazione).

Al contrario, una molecola non trattenuta avrà  $k = 0$  e  $V_R = V_0$  (limite di esclusione).

Le molecole che possiedono volumi di ritenzione compresi tra i due limiti possono essere separate con una certa efficienza. In questo caso, si riesce a determinare dal  $V_R$  della specie eluita il suo peso molecolare per mezzo di una relazione logaritmica, determinata con degli standard di riferimento aventi peso molecolare noto.

Per effettuare la corsa cromatografica della proteina si è utilizzato uno strumento denominato FPLC (Fast Protein Liquid Chromatography) consistente in un sistema di pompe controllato da un software (Unicorn) nel quale si immettono i diversi parametri di corsa, come pressione, flusso di fase mobile e volume delle frazioni da raccogliere. La fase mobile è immessa in colonna tramite due pescanti con filtro collegati alle due pompe, i quali possono essere utilizzati per formare un gradiente di concentrazione tra due tamponi differenti.

L'andamento dell'eluizione è monitorato tramite uno spettrofotometro installato nello strumento e collegato al PC. Selezionando la lunghezza d'onda a cui operare (270-280 nm per le proteine), è possibile rilevare, al passaggio della proteina, un aumento di assorbanza e quindi la formazione di un picco il cui andamento viene visualizzato sulla schermata del computer.

### **3.9 Spettrofotometria**

La spettrofotometria è una tecnica impiegata per la stima quantitativa di diverse molecole, tra cui le proteine. L'assorbimento della luce nel visibile e nell'UV è dovuto all'interazione tra una radiazione elettromagnetica di una determinata lunghezza d'onda e gli elettroni del guscio più esterno degli atomi che compongono la molecola da analizzare. Gli elettroni, infatti, possiedono livelli energetici definiti (quantizzati) e assorbono solo la radiazione avente energia che causa transizioni da un livello all'altro. Se si

considera l'intensità della luce incidente come  $I_0$  e l'intensità della luce trasmessa  $I$ , si definisce trasmittanza (frazione di luce trasmessa) il rapporto tra  $I$  e  $I_0$ . Poiché la trasmittanza diminuisce in modo esponenziale in rapporto alla concentrazione del campione e alla lunghezza del percorso della luce ("cammino ottico"), essa è definita dalla legge di Lambert-Beer come:

$$A = -\log T = \epsilon C l$$

dove:

$A$  è l'assorbanza;

$T$  è la trasmittanza;

$C$  è la concentrazione (in mg/mL);

$\epsilon$  è il coefficiente di estinzione molare per mg/mL di proteina alla lunghezza d'onda corrispondente al massimo di assorbimento;

$l$  è il cammino ottico della cella.

Per misurare l'assorbanza di una soluzione si utilizza uno spettrofotometro, cioè un generatore di radiazioni aventi una lunghezza d'onda specifica, selezionabile grazie alla presenza di un monocromatore. Ogni volta che si vuole effettuare una misura spettrofotometrica di una sostanza in soluzione, è sempre necessario misurare l'assorbanza del tampone in cui questa è sciolta (bianco), poiché questo può assorbire a determinate lunghezze d'onda. Se si esegue una misura di uno spettro, è necessario misurare la linea base (assorbanza del solo tampone in uno spettro di lunghezze d'onda definito) e sottrarre questa alla misura dell'assorbanza del campione. Nel presente lavoro sono state eseguite misure di stima approssimata della concentrazione delle proteine, sfruttando la legge di Lambert-Beer, a partire dall' $\epsilon$  teorico determinato dalla sequenza amminoacidica, e misurando il picco di assorbimento in uno spettro di lunghezze d'onda nell'intervallo 240-350 nm, in cui si ha assorbimento dovuto alla presenza di anelli aromatici nei residui degli aminoacidi Trp, Tyr e, in minor misura, Phe.

### **3.10 Saggi di interazione tra proteine mediante “Dot-blot”**

Il Dot-blot è una tecnica di biologia molecolare utilizzata per studiare l'interazione tra macromolecole; si è sviluppata, parallelamente al Northern e Southern blot, negli anni '70 e inizialmente veniva utilizzata esclusivamente per analisi di interazione tra molecole di DNA e/o RNA, sfruttando il principio dell'appaiamento tra basi complementari. Il Dot-blot è una tecnica semiquantitativa che, nel caso degli acidi nucleici, permette di determinare, per esempio, il numero di copie di un certo gene presenti in un determinato genoma o la quantità di un certo trascritto di mRNA in una cellula. Per lo studio d'interazione tra acidi nucleici, il protocollo prevede sostanzialmente che alcune gocce di DNA bersaglio vengano poggiate su una membrana di nitrocellulosa, denaturate e lasciate asciugare, e ricoperte successivamente con sequenze sonda marcate a singolo filamento per poter individuare l'eventuale ibridizzazione. Negli ultimi anni sono stati messi a punto dei protocolli che consentono di sfruttare alcuni principi di questa tecnica anche per lo studio di interazioni acido nucleico-proteina e proteina-proteina.

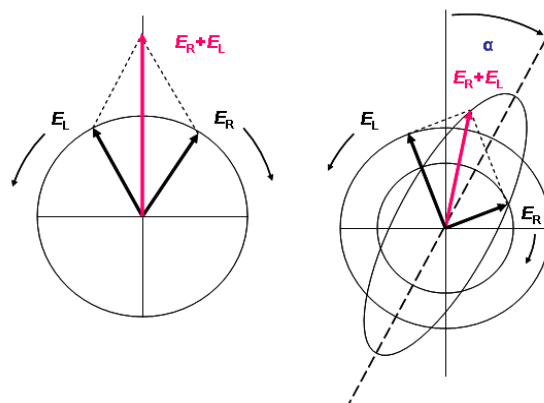
In questo lavoro di tesi sono stati condotti saggi di interazione tra proteine *cag* e il protocollo seguito è riportato nei punti seguenti:

1. si preparano delle lastrine di nitrocellulosa con dei cerchi all'interno dei quali vengono caricate le soluzioni di proteina (proteina primaria) facendole adsorbire lentamente (in modo da farle stare nello spazio delimitato dai cerchi) attraverso passaggi successivi di caricamento; la proteina primaria in genere ha una concentrazione di 5-10 mg/mL;
2. dopo che le soluzioni caricate si sono asciugate, si incubano le membrane di nitrocellulosa con una soluzione data da TBS (Tris buffer saline) + 3-5% BSA (albumina di siero bovino), per almeno un'ora, a temperatura ambiente e sotto blanda agitazione. Il TBS è una soluzione costituita da 20 nM Tris, 150 nM NaCl, pH 7.5;
3. le membrane vengono incubate con la soluzione della seconda proteina (il partner testato per l'interazione), diluita in buffer TBS. Questa incubazione dura 3-4 ore a temperatura ambiente oppure una notte a 4°C, in ogni caso sotto blanda agitazione. La proteina secondaria ha una concentrazione molto bassa, in genere 100-300 µg diluiti in 4 mL di buffer;

4. si effettuano tre lavaggi vigorosi di 5-10 minuti l'uno con TBS addizionato di detergente Tween20, in concentrazione finale dello 0.05% v/v;
5. si inserisce l'anticorpo primario (della ditta CALBIOCHEM, diluito 1:2000) diluito in TBS + 3% BSA; esso deve essere capace di riconoscere la seconda proteina, la quale può per esempio essere dotata di His-tag o di un'altra sequenza riconoscibile in modo specifico. Questa incubazione dura 4 ore a T ambiente o un'intera notte a 4°C, sempre sotto blanda agitazione;
6. si eseguono tre lavaggi di 5 minuti come al punto 4;
7. si incuba per 1-2 ore a T ambiente e sotto blanda agitazione l'anticorpo secondario, cioè quello che riconosce specificamente l'anticorpo primario;
8. si effettuano altri 3 lavaggi come i precedenti e si prosegue con un quarto lavaggio con buffer per fosfatasi alcalina, costituito da 100 mM Tris pH 9.5, 100 mM NaCl, 5 mM MgCl<sub>2</sub>;
9. il blot viene rilevato tramite una miscela **BCIP/NBT** (5-Bromo-4-chloro-3-indolylphosphate, tuluidine salt/Nitroblue tetrazolium chloride), possibilmente al buio e sottoponendo il tutto a blanda agitazione; la rivelazione può essere bloccata attraverso un lavaggio delle membrane con acqua distillata.

### 3.11 Dicroismo circolare

Il dicroismo circolare (CD) è una tecnica molto utile per determinare i contributi di struttura secondaria di macromolecole biologiche in soluzione. Il fenomeno fisico alla base del CD è il diverso assorbimento delle componenti, destrorsa e sinistrorsa, di un fascio di luce linearmente polarizzato da parte di molecole otticamente attive. Il campione, che deve essere chirale, viene colpito da una radiazione linearmente polarizzata, scomponibile in due componenti in fase polarizzate circolarmente, una destrogira e l'altra sinistrogira. A causa dei diversi coefficienti di estinzione molare,  $\epsilon_L$  e  $\epsilon_R$ , si verifica un diverso assorbimento delle due componenti, pertanto la radiazione in uscita risulterà polarizzata ellitticamente. Nel tempo le diverse componenti comporranno nello spazio un'ellisse di asse minore **a** e di asse maggiore **b**.



**Figura 3.6** Sulla sinistra è rappresentato il vettore campo elettrico di un fascio di luce linearmente polarizzato; sulla destra è mostrato l'effetto sulle componenti destrorse e sinistrorse del fascio in seguito all'assorbimento della radiazione da parte di un mezzo chirale.

L'arcotangente del rapporto tra b ed a ( $\arctg(b/a)$  [gradi]) si definisce ellitticità ( $\theta$ ), mentre l'ellitticità molare per residuo (aminoacidico) è data da:

$$[\theta]_R = 100\theta/[Cl(n-1)] \quad [\text{gradi} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}]$$

dove: C è la concentrazione molare del campione (mol/l)

l è la lunghezza della cella (cuvetta) in cui è contenuto il campione

n è il numero di aminoacidi presenti nella molecola

La relazione che lega l'ellitticità all'assorbimento diverso di luce polarizzata circolarmente destra e sinistra è la seguente:

$$\theta = 33 \Delta A = 33 l C \Delta \epsilon \quad [\text{gradi} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}]$$

con  $\Delta A = A_L - A_R$

e  $\Delta \epsilon = \epsilon_L - \epsilon_R$ .

Affinché un composto sia otticamente attivo è necessario che possieda dei cromofori in grado di assorbire radiazioni a determinate lunghezze d'onda e che essi siano intrinsecamente chirali, oppure indotti asimmetricamente da parte dell'intorno. Il legame ammidico e gli anelli aromatici dei residui Trp, Tyr e Phe sono di per sé planari e quindi achirali, ma risultano otticamente attivi per induzione, perché legati a stereocentri come i carboni  $C_\alpha$  della catena polipeptidica oppure perché appartenenti a una conformazione proteica tale da renderli chirali.

I cromofori presenti nelle proteine hanno i seguenti picchi di assorbimento:



- il gruppo peptidico presenta massimi di assorbimento nel lontano UV, attorno a 190 nm;
- le catene laterali dei residui aromatici Phe, Tyr, Trp assorbono sia nel lontano che nel vicino UV, dove hanno massimi di assorbimento attorno a 270 nm;
- il ponte disolfuro, intrinsecamente chirale, presenta un massimo di assorbimento intorno a 250-260 nm.

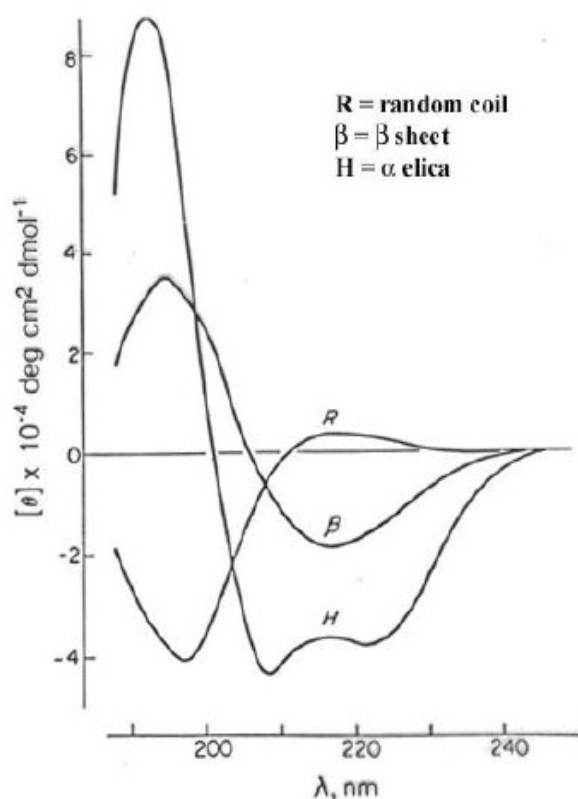
Nel vicino UV il segnale dei residui aromatici è più debole e lo spettro in generale è più difficile da interpretare, ad eccezione della situazione in cui tali residui siano immobilizzati in una struttura terziaria particolare per cui risentano di un intorno asimmetrico.

Gli spettri CD di  $\alpha$ -eliche, strutture  $\beta$  e strutture disordinate (random coil) presentano diverse bande positive e negative, corrispondenti a diverse transizioni elettroniche (vedi tabella sottostante).

Composizione	Transizione	Banda
$\alpha$ -elica	$n \rightarrow \pi^*$ permessa solo magneticamente	222 nm negativa
	$\pi \rightarrow \pi^*$ componente perpendicolare all'asse dell'elica	190 nm positiva
	$\pi \rightarrow \pi^*$ componenti parallela all'asse dell'elica	208 nm negativa
Struttura- $\beta$	$n \rightarrow \pi^*$	216-217 nm negativa
	$\pi \rightarrow \pi^*$	195 nm positiva
Struttura disordinata ( <i>random coil</i> )	$n \rightarrow \pi^*$	215 nm positiva
	$\pi \rightarrow \pi^*$	197 nm negativa

**Tabella 3.3** Sono elencate le varie transizioni elettroniche relative ai diversi elementi di struttura secondaria di una proteina.

Per interpretare gli spettri di dicroismo circolare in termini di struttura secondaria di una proteina si fa riferimento a spettri standard di struttura  $\alpha$ ,  $\beta$  e random coil ottenuti dall'attività ottica di catene polipeptidiche aventi struttura tridimensionale nota.



**Figura 3.7** Spettro generico di dicroismo circolare nelle regioni di assorbimento dei legami peptidici che configurano strutture prevalentemente  $\alpha$ ,  $\beta$  e random coil.

Una volta ottenuto lo spettro dicroico, ci si avvale di programmi computazionali che risolvono, per un insieme definito di lunghezze d'onda, dei sistemi di equazioni del tipo:

$$[\theta(\lambda)] = \chi_{\alpha}[\theta_{\alpha}(\lambda)] + \chi_{\beta}[\theta_{\beta}(\lambda)] + \chi_{rc}[\theta_{rc}(\lambda)] \quad [\text{gradi} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}]$$

dove  $\chi_{\alpha}$ ,  $\chi_{\beta}$ ,  $\chi_{rc}$  sono coefficienti indicanti rispettivamente la frazione di  $\alpha$ -elica, struttura  $\beta$  e random coil presenti nella struttura proteica; pertanto, almeno teoricamente,  $\chi_{\alpha} + \chi_{\beta} + \chi_{rc} = 1$ ;

$[\theta(\lambda)]$  è l'intensità dello spettro dicroico della proteina, alla lunghezza d'onda  $\lambda$ ;

$[\theta_{\alpha}(\lambda)]$ ,  $[\theta_{\beta}(\lambda)]$ ,  $[\theta_{rc}(\lambda)]$  indicano rispettivamente l'intensità di uno spettro CD proprio di catene polipeptidiche con strutture completamente in  $\alpha$ -elica, struttura  $\beta$  e random coil.

Conoscendo l'ellitticità molare per ciascuna lunghezza d'onda selezionata, si può stimare la percentuale media relativa delle varie strutture presenti,

con una precisione massima del 10%. L'incertezza, infatti, è dovuta a molteplici fattori, alcuni dei quali sono i seguenti:

- l'approssimazione nella scelta degli spettri standard per le tre conformazioni. Il valore dell'ellitticità per una stessa lunghezza d'onda e conformazione dipende fortemente dalla lunghezza della catena peptidica e dall'intorno chimico della struttura, i cui effetti sono difficilmente determinabili a priori;
- la presenza, spesso considerevole, di residui aromatici che influenzano la regione del lontano UV;
- l'errore nella stima della concentrazione proteica;
- la presenza di strutture secondarie particolari non riconducibili agli standard proposti.

Sebbene soggetto ad errori di valutazione della struttura, il CD rimane ancora un veloce, valido e non distruttivo metodo d'indagine che richiede piccole quantità di proteina.

Lo strumento utilizzato per la raccolta dei dati sperimentali relativi a *CagV* è il dicrografo a singolo raggio Jasco J-715, che presenta le seguenti caratteristiche:

- temperatura interna della cella di 25 °C;
- intervallo di registrazione dello spettro tra i 190 e i 260 nm;
- larghezza della banda passante di 2 nm;
- risoluzione di 0.2 nm;
- velocità di scansione di 50 nm/min;
- tempo di risposta di 2 sec;
- 10 accumuli di registrazione per spettro;
- cuvetta con cammino ottico di 0.05 mm.

### **3.11 Cristallizzazione di proteine**

#### **PRINCIPI BASE DELLA CRISTALLIZZAZIONE**

La cristallografia di proteine è una tecnica che consiste nello studio della struttura tridimensionale di una proteina mediante analisi della diffrazione dei raggi X di un cristallo, che è una sequenza ordinata e periodica di molecole nello spazio.

La cristallizzazione di una proteina è un processo complesso e delicato, talvolta difficilmente riproducibile, a causa di svariati errori che si compiono, spesso di natura sistematica, quali: la misurazione di piccoli

volumi adoperati per le soluzioni (in genere 0.5-2  $\mu\text{L}$ ), la valutazione del titolo delle soluzioni, lievi variazioni involontarie della temperatura nel corso del processo, piccoli urti meccanici. Tutti questi fattori influenzano gli equilibri a cui è soggetta la proteina in soluzione, dove solo la prevalenza di alcuni di questi portano a cristallizzare la proteina nella conformazione nativa. Anche il solvente è indispensabile al mantenimento della struttura tridimensionale delle macromolecole all'interno del reticolo cristallino; infatti i loro cristalli inglobano, a differenza di quelli di piccole molecole, un contenuto d'acqua elevato, compreso tra il 30% e il 70% del volume totale del reticolo. Pertanto i metodi tradizionali di cristallizzazione per molecole inorganiche che sfruttano l'evaporazione totale del solvente o la variazione drastica della temperatura non possono essere utilizzati, in quanto distruggerebbero la struttura del cristallo.

#### CURVE DI SOLUBILITA'

Per ottenere dei cristalli che forniscano dati di diffrazione soddisfacenti sono fondamentali un'elevata purezza e una lenta precipitazione della proteina. Il concetto di purezza non è solo relativo all'assenza di contaminanti, ma anche alla struttura e conformazione della proteina: la soluzione deve possedere un'omogenea popolazione di conformeri e quindi definite proprietà chimico-fisiche perché si possano formare ordinate strutture cristalline. Invece la lentezza del processo serve ad ottenere cristalli di dimensioni maggiori e un elevato grado di regolarità nella collocazione ed orientazione delle molecole all'interno del reticolo tridimensionale.

E' utile analizzare separatamente le cause della spontanea formazione dei primi nuclei di cristallo e della successiva crescita. Si possono individuare tre tappe, che si realizzano a valori diversi delle concentrazioni di proteina e di agente precipitante (vedi fig. 3.8).

1. Nucleazione: è un fenomeno la cui probabilità di avvenire aumenta con l'aumentare del grado di sovrasaturazione della soluzione proteica. Al crescere del numero di nuovi centri di nucleazione si abbassa il grado di sovrasaturazione, diminuendo la probabilità di formazione di nuovi centri. Inoltre, durante la creazione dei primi aggregati, la formazione di nuclei ordinati può competere con il processo irreversibile di produzione di aggregati amorfi, portando anch'esso ad un abbassamento del grado di sovrasaturazione.

2. Crescita: durante questo processo il cristallo cresce attorno ai primi aggregati proteici.
3. Cessazione della crescita: fase in cui la proteina è in equilibrio tra la sua fase acquosa e quella solida, sia nel caso di un cristallo che di un aggregato irregolare.

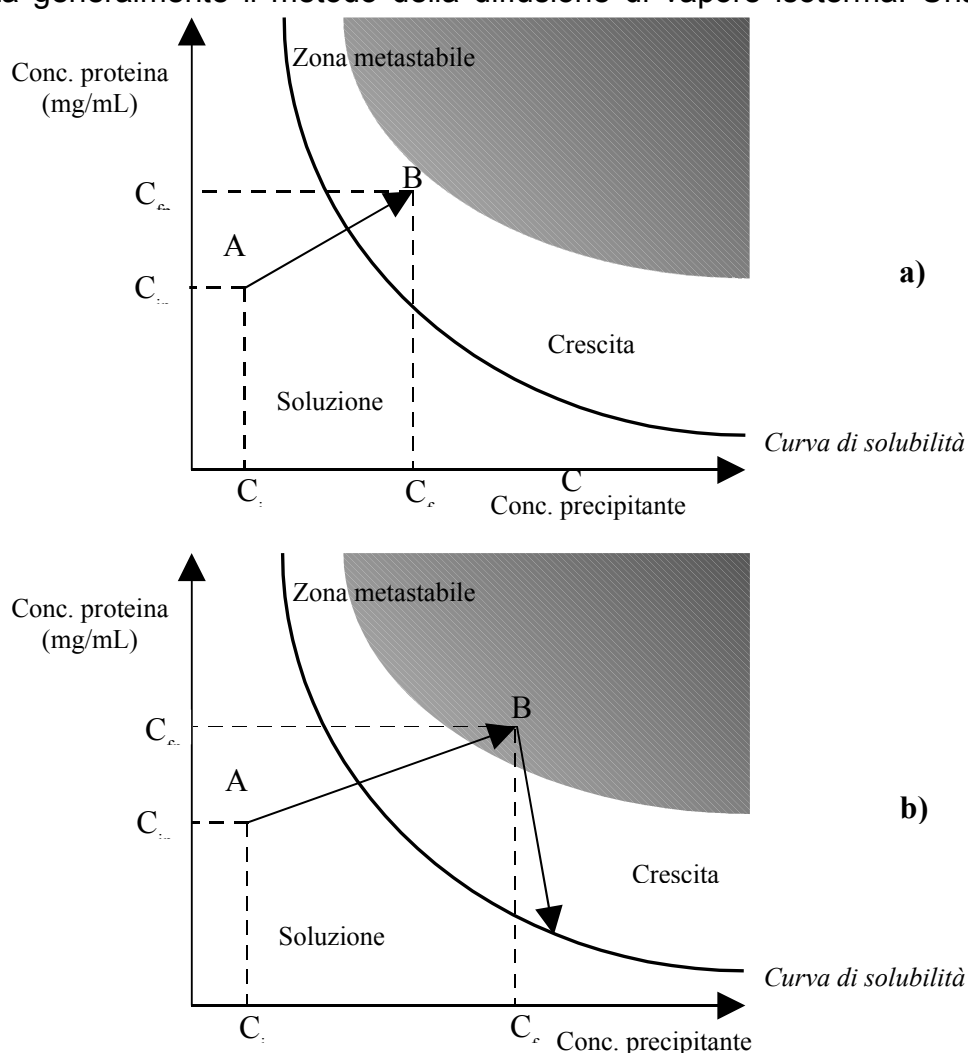
Tutti questi stati avvengono in condizioni di sovrasaturazione della proteina, uno stato termodinamicamente instabile (metastabile) che viene solitamente raggiunto sottoponendo la soluzione proteica a parziale lenta evaporazione. Alcuni dei principali parametri che influenzano questo fenomeno sono:

- a. pH: variando tale parametro mutano le cariche dei residui amminoacidici presenti sulla superficie della proteina, modificandone la carica netta. A valori di pH in cui la macromolecola manifesta una carica formale netta nulla, cioè al punto isoelettrico, essa presenta anche la minima solubilità.
- b. Temperatura: ha un effetto minore sulla precipitazione, anche se comunque presente. Tuttavia si è visto che in genere la solubilità delle proteine decresce al diminuire della temperatura.
- c. Natura del sale: la solubilità di una proteina varia in dipendenza della forza ionica e della natura del sale in soluzione e quindi della sua concentrazione e della carica dei suoi ioni. Generalmente, all'aumentare della concentrazione salina, la solubilità della proteina prima aumenta (salting in), perché gli ioni interagiscono con le sue cariche schermandole, e poi decresce (salting out) a causa dell'aumento di fenomeni di aggregazione della proteina. La capacità degli ioni di influenzare la solubilità delle proteine è rappresentata dalla serie di Hofmeister. Tale serie indica che ad alte concentrazioni alcuni ioni, come per esempio il solfato, sono capaci di indurre la precipitazione in forma nativa delle macromolecole biologiche, mentre altri ioni, come il guanidinio o il tiocianato, tendono a mantenerle in soluzione denaturandole.
- d. Polarità del solvente: questa viene in genere ridotta aggiungendo, ad una soluzione acquosa, una percentuale variabile di solvente organico. La percentuale di solvente organico deve essere comunque contenuta, perché altrimenti c'è il rischio di provocare la denaturazione della proteina.
- e. Agenti precipitanti: si distinguono in 3 categorie principali: sali, solventi organici, e polimeri, il più comune dei quali è il polietilenglicole (PEG).

**Figura 3.8** Curva di solubilità di una soluzione proteica al variare della concentrazione di precipitante e andamento della concentrazione proteica nel caso della tecnica della diffusione di vapore isoterma. a) in questo caso si ha la formazione di cristalli; b) E' rappresentato il caso in cui non si formano cristalli adatti all'analisi ai raggi X.

### TECNICHE DI CRISTALLIZZAZIONE

Per portare la soluzione proteica in uno stato di sovrasaturazione si utilizza generalmente il metodo della diffusione di vapore isoterma. Una



goccia di soluzione proteica, miscelata opportunamente con una soluzione precipitante, viene posta in un sistema chiuso isoterma ad equilibrarsi con la stessa soluzione precipitante, ma a concentrazione maggiore, generalmente doppia. La tensione di vapore della goccia e del precipitante non sono uguali, per cui il sistema tenderà ad evolvere verso l'equilibrio attraverso l'evaporazione delle specie volatili (acqua, solventi organici)

finché le tensioni di vapore della goccia e del precipitante non assumono lo stesso valore. Nella maggior parte dei casi il processo porta ad una perdita di acqua da parte della goccia, favorendo una lenta ma progressiva diminuzione di volume della stessa ed un aumento della concentrazione delle specie in soluzione fino allo stato di sovrasaturazione.

Queste condizioni possono essere ottenute attraverso due sistemi:

- a goccia pendente (“hanging drop”)
- a goccia appoggiata (“sitting drop”)

Nel primo caso una soluzione contenente la proteina e l’agente precipitante viene deposta su un vetrino siliconato, il vetrino viene quindi capovolto e posto sopra un pozzetto contenente il precipitante; quest’ultimo è in concentrazione superiore a quello contenuto nella goccia. Nel caso della “sitting drop” la goccia poggia su un supporto in plastica simile ad un ponticello, posizionato all’interno della soluzione del precipitante; il tutto viene poi chiuso ermeticamente sigillando il pozzetto con un vetrino.

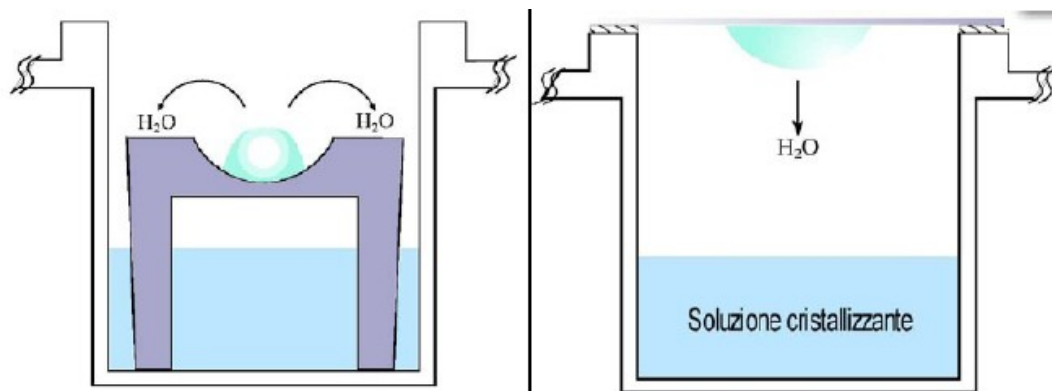


Figura 3.9 Rappresentazione degli apparati di cristallizzazione sitting drop (sinistra) e hanging drop (destra).

## **4. RISULTATI e DISCUSSIONE**

Il lavoro intrapreso si è articolato nei seguenti passaggi:

- Amplificazione e clonazione dei frammenti di DNA codificanti per *CagVn1* e *CagVn2* in opportuni vettori;
- Prove di espressione e solubilità in piccoli volumi, allo scopo di mettere a punto le condizioni ottimali di lavoro;
- Espressione su larga scala per ottenere quantità adeguate di proteina (dell'ordine dei mg) per poter effettuare gli studi successivi;
- Estrazione delle proteine dai batteri tramite lisi cellulare;
- Purificazione delle proteine mediante tecniche cromatografiche;
- Prove di cristallizzazione delle proteine purificate e successiva ottimizzazione delle condizioni per migliorare la qualità dei cristalli;
- Caratterizzazione strutturale tramite misure di dicroismo circolare e di raccolta dei dati di diffrazione ai raggi X dei cristalli ottenuti;
- Studi di interazione tra proteine dell'isola di patogenicità di *H.pylori* mediante dot-blot.

#### **4.1 Amplificazione dei frammenti di DNA del gene *CagV***



L'amplificazione è stata fatta mediante PCR utilizzando il termociclatore "Bio-Rad Gene Cycler". Poiché i due frammenti (*CagVn1* e *CagVn2*) hanno l'estremità 3' in comune, è stato utilizzato lo stesso *primer reverse*, mentre i *primers forward* sono differenti, in quanto la sequenza al 5' è diversa nei due casi. Precisamente i primers utilizzati per l'amplificazione sono:

Primer forward (*CagVn1*): 5' **CACCTTGAAGAAACGGATATATTTG** 3'

Primer forward (*CagVn2*): 5' **CACCTCTGAAGCCGTTGTGGATAG** 3'

Primer reverse: 5' TCATTATTGTTTTAATGCCTTATTTTTTG 3'

La reazione viene eseguita in soluzione a partire da una miscela di reagenti con volume finale di 50 µL. I reagenti e i rispettivi volumi applicati sono riportati in tabella 4.1:

Reagenti	Volume (µL)
DNA	1
Primer forward	0.3
Primer reverse	0.3
Tampone 10X	5
MgSO <sub>4</sub>	2
DNA polimerasi	1.25
dNTPs	1
H <sub>2</sub> O sterile	39.15

**Tabella 4.1** Composizione della mix di PCR

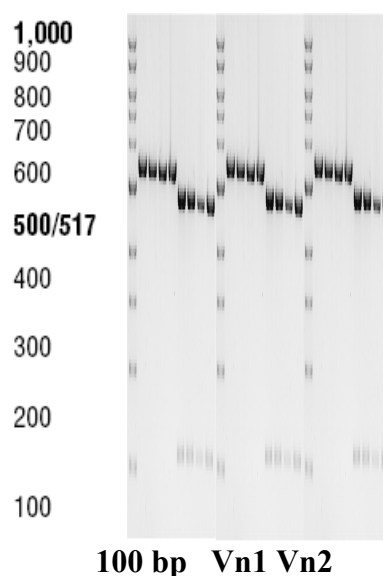
Il ciclo impostato per la PCR è invece riportato nella tabella seguente:

Cicli di PCR	Fasi di PCR	Durata	T (°C)
ciclo n.1 (x 1)	Denaturazione	5 min	94

ciclo n.2 (x 1)	Denaturazione	1 min	94
	Annealing	45 s	55
	Elongation	45 s	72
ciclo n.3 (x 1)	Completamento	10 min	72

**Tabella 4.2 E' riassunto il ciclo di PCR**

Una volta avvenuta la PCR, i prodotti sono stati analizzati mediante corsa su gel di agarosio allo 0.8% e i risultati dell'analisi sono riportati nella figura seguente:



**Figura 4.1 Gel di agarosio di controllo di avvenuta amplificazione.**

I prodotti della PCR sono stati quindi purificati attraverso la procedura della “gel estrazione” e le sequenze sono state clonate usando la strategia del “TOPO cloning”, i cui dettagli sperimentali sono stati descritti nella sezione “Materiali e metodi”.

## **4.2 Trasformazione dei batteri BI21 con pET151-CagV**

Dopo aver clonato i frammenti di DNA codificanti per le due catene polipeptidiche (CagVn1 e CagVn2) nel vettore pET151, è stata fatta la trasformazione dei batteri competenti di *Escherichia coli* ceppo BI21. Quest'ultimi sono stati piastrati su terreno solido contenente ampicillina (50 µg/µL) e fatti crescere incubandoli a 37°C per un'intera notte. Il giorno seguente si sono controllate le piastre e si è visto che i batteri, in entrambi i casi, erano cresciuti bene (dalla presenza delle tipiche colonie batteriche), il che indica che la trasformazione è avvenuta con successo (cioè i batteri cresciuti erano resistenti all'ampicillina, grazie ad un gene specifico contenuto nel vettore, altrimenti non sarebbero potuti crescere su terreno ricco di antibiotico).

### **4.3 Espressione di His<sub>6</sub>CagVn1 e His<sub>6</sub>CagVn2**

Prima di far partire l'espressione delle due catene polipeptidiche, i batteri precedentemente trasformati con il plasmide pET151 sono stati piastrati su terreno solido LB contenente ampicillina e fatti crescere ON a 37°C.

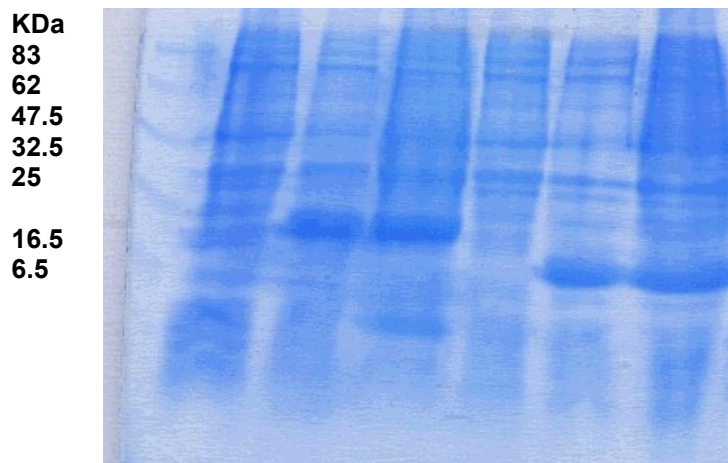
In un primo momento sono state eseguite delle prove di espressione su piccola scala, al fine di verificare la stabilità di ciascuna proteina e le condizioni di espressione; a questo scopo, per ognuna delle due proteine è stata inoculata una colonia batterica, presa dalla piastra Petri, in 40 mL di terreno liquido LB con aggiunta di ampicillina. Le colture sono state incubate a 37°C sotto agitazione (180 rpm) fino ad ottenere una densità ottica di circa 0.6. L'espressione è stata indotta aggiungendo al brodo di coltura una soluzione 0.5 mM di IPTG. A questo punto le colture sono state messe in incubazione ON a 20°C.

Sono stati prelevati, a vari intervalli di tempo, dei campioni di coltura, i quali sono stati lisati attraverso trattamento con sample buffer diluito in tampone e sottoponendoli a bollitura (100°C per 15 minuti).

I campioni sono stati fatti correre in un gel SDS-PAGE al 15% per controllare i livelli di espressione delle due catene polipeptidiche. La corsa è stata fatta ad una intensità di corrente di 40 mA per un tempo di circa 1 ora.

A pagina seguente è riportata l'immagine del gel:

1    2    3    4    5    6    7



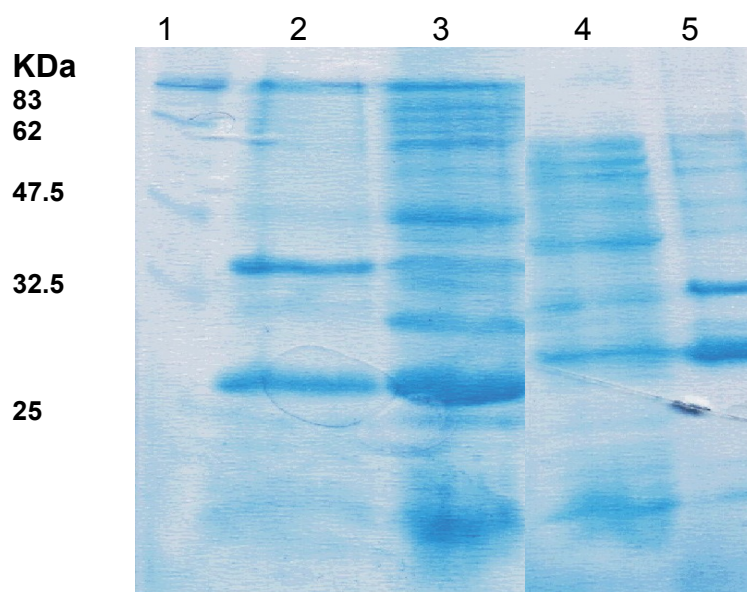
**Figura 4.2 SDS-PAGE relativo alle prove di espressione.**

1. Marker neb
2. CagVn1 N I
3. CagVn1 Indotto 2h
4. CagVn1 Indotto Over night
5. CagVn2 Non Indotto
6. CagVn2 Indotto 2h
7. CagVn2 Indotto Over Night

Dall'analisi del gel si può osservare, nei campioni in cui è stata fatta l'induzione con IPTG, la presenza di una banda più intensa a livello dei pesi molecolari relativi alle due catene polipeptidiche. CagVn1 (22.6 KDa) è corsa di meno rispetto a CagVn2 (19 kDa) in quanto, essendo più lunga, ha più difficoltà a migrare. Si è ripetuta quindi la procedura su ampia scala, inoculando batteri BI21 trasformati con pET151CagVn1 e pET151CagVn2 in 2L di terreno liquido LB addizionato con ampicillina. Dopo aver raggiunto un valore di OD di circa 0.6, i batteri sono stati indotti con 0.5 mM di IPTG e incubati per 4 ore a 37°C e 180 rpm. Una volta che la crescita è terminata, si sono centrifugati i batteri per 15 min a 3900 rpm, dopo di che si è risospeso il pellet in PBS 1X, si è centrifugato per 10 min a 8000 rpm e il nuovo pellet, costituito da cellule e loro derivati, è stato conservato a -80°C, allo scopo di inibire il più possibile l'azione degli enzimi proteolitici che potrebbero degradare le proteine espresse. Il precipitato batterico è stato poi separato dal surnatante e risospeso in circa 20 mL di tampone A (10mL per ogni litro di coltura) che, nel caso di CagVn1 il è costituito da 30 mM Tris pH 7.5 e 200 mM NaCl, mentre per CagVn2 è dato da 30mM Tris pH 7.3, 300 mM NaCl. La composizione dei tamponi è stata scelta in base alle caratteristiche chimiche basate su calcoli teorici (effettuate a partire dalla sequenza aminoacidica) delle catene polipeptidiche, quali solubilità e punto isoelettrico.

Oltre alle prove di espressione, sono stati fatti anche dei saggi di solubilità; per far questo si è proceduto con la lisi batterica, mediante lisozima e sonicazione (vedi materiali e metodi), e il lisato ottenuto è stato centrifugato per 20 min a 25000 rpm per separare le proteine solubili da quelle insolubili.

Il grado di solubilità delle proteine è stato verificato mediante gel SDS-PAGE al 15%, di seguito riportato:



**Figura 4.3 SDS-PAGE relativo alle prove di solubilità**

1. LMW (Neb)
2. CagVn2 pellet
3. CagVn2 surnatante
4. CagVn1 surnatante
5. CagVn1 pellet

Dall'analisi del gel si può notare, in entrambi i casi, la presenza di proteina sia nel surnatante che nel pellet; tuttavia per quanto riguarda CagVn2, l'intensità della banda è nettamente maggiore nel surnatante rispetto al pellet, a differenza di CagVn1 che sembra essere presente maggiormente nel pellet. CagVn2 quindi sembra essere più solubile rispetto a CagVn1, per questo motivo si è deciso di iniziare la parte di purificazione e i successivi tentativi di cristallizzazione da CagVn2.

#### 4.4 Purificazione mediante cromatografia di affinità

Il surnatante relativo ai lisati precedenti è stato caricato su colonnina per cromatografia di affinità "HIS TRAP HIGH PERFORMANCE" (Amersham Biosciences), la quale è costituita da una matrice a base di agarosio con fitti legami crociati, a cui sono legate delle molecole di acido nitrilotriacetico (NTA) che sono in grado di chelare ioni  $Ni^{2+}$ ; per il caricamento dei campioni è stata utilizzata una pompa peristaltica ad un flusso di 4-5 mL al minuto. La colonnina, che era conservata in acqua, è stata equilibrata con tampone A e successivamente si è caricato il campione, costituito da 15 mL di soluzione in 200 mL di tampone A addizionato con 10 mM di imidazolo.

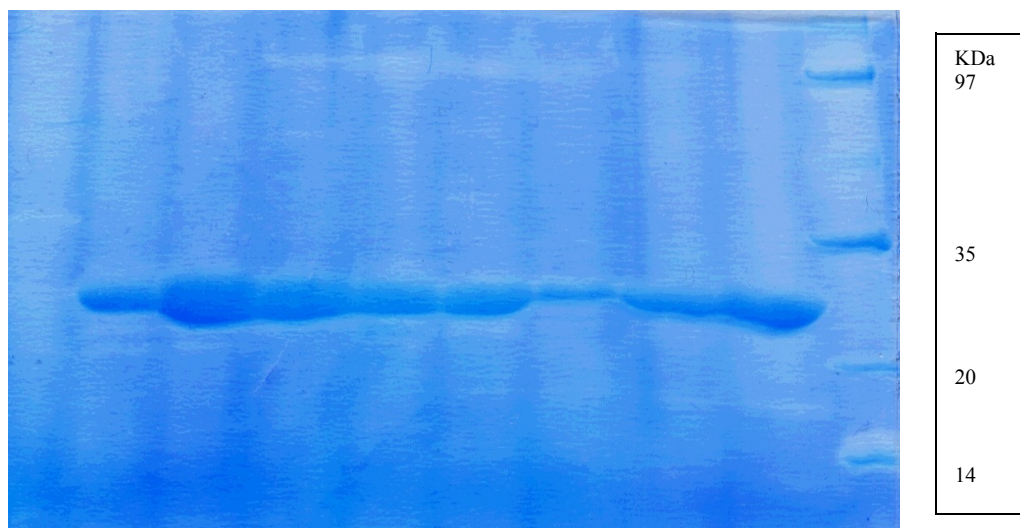
A questo punto si è eseguita l'eluizione, installando la colonnina nell'apparato FPLC, utilizzando due tamponi (A e B). Il tampone B ha la stessa concentrazione e pH del tampone A, con aggiunta di 500 mM di imidazolo. Sono stati impostati i gradienti di B per l'eluizione delle proteine.

Per quanto riguarda His<sub>6</sub>CagVn2 il cromatogramma ottenuto è riportato di seguito:

**Figura 4.4 Cromatografia di affinità per His<sub>6</sub>CagVn2. Il cromatogramma mostra l'andamento dell'assorbanza in funzione del volume di tampone passato in colonna (blu) e il gradiente di tampone B (azzurro). Le frazioni raccolte sono numerate in rosso.**

Alcune delle frazioni delle eluizioni sono state raccolte per essere analizzate mediante SDS-PAGE al 15%. Di seguito è riportata l'immagine di un gel, dal quale si ha la conferma della presenza di proteina nelle varie frazioni:

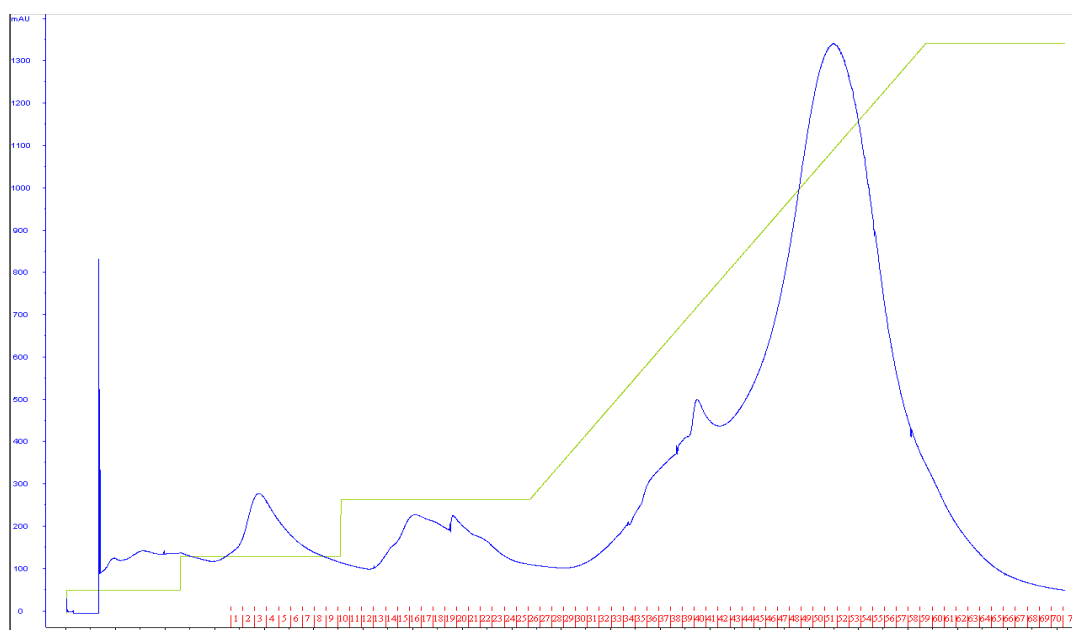
1    2    3    4    5    6    7    8    9    10



**Figura 4.5 SDS-PAGE relativo alla cromatografia di affinità di CagVn2.**

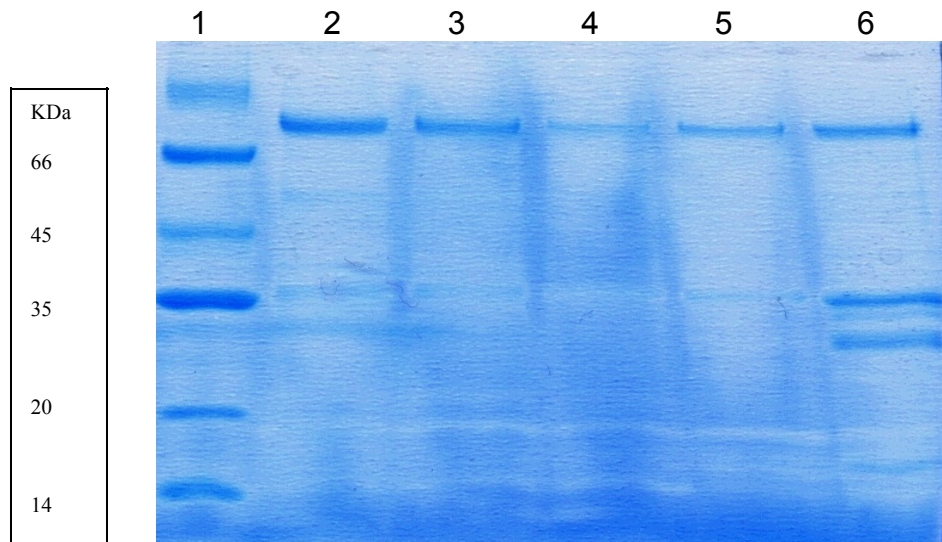
1. Primo picco
2. Secondo picco (prima parte)
3. Secondo picco (seconda parte)
4. Fr.1
5. Fr.4
6. Fr.12
7. Fr.17
8. Fr.23
9. Fr.26
10. LMW

Per quanto riguarda His<sub>6</sub>CagVn1, invece, il cromatogramma è riportato nella pagina seguente:



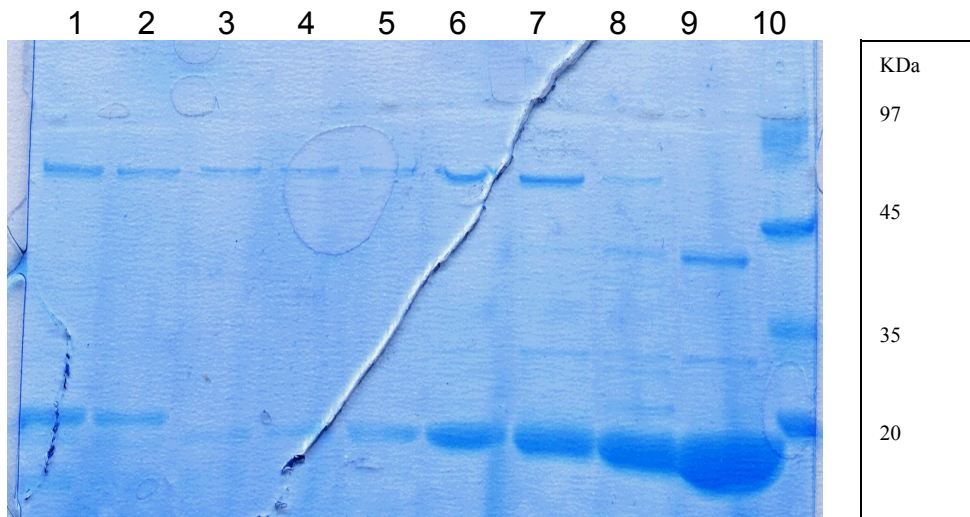
**Figura 4.6 Cromatografia di affinità per His<sub>6</sub>CagVn1. Il cromatogramma mostra l'andamento dell'assorbanza in funzione del volume di tampone passato in colonna (blu) e il gradiente di tampone B (verde). Le frazioni raccolte sono numerate in rosso.**

Anche in questo caso alcune frazioni (quelle relative ai picchi di assorbimento) sono state caricate su gel SDS-PAGE al 15%, le cui immagini sono riportate di seguito:



**Figura 4.7 SDS-PAGE relativo alla cromatografia di affinità di CagVn1 (prima parte).**

1. LMW
2. Fr. 3
3. Fr.5
4. Fr. 9
5. Fr.15
6. Fr.17



**Figura 4.8 SDS-PAGE relativo alla cromatografia di affinità di CagVn1 (seconda parte).**

1. Fr. 20
2. Fr. 23
3. Fr 33
4. Fr. 35
5. Fr. 38
6. Fr. 41
7. Fr. 44



8. Fr. 48
9. Fr. 52
10. LMW

Come si può vedere dalle immagini, la proteina inizia ad essere presente in forma abbastanza pura a partire dalla frazione 40; nelle ultime tre frazioni compare anche una banda, più debole, a livello dei pesi molecolari del dimero.

A questo punto si è deciso di cercare di eliminare la coda di istidine presente all'N terminale di CagVn1. Per far questo si è presa una aliquota di proteina (70 mL alla concentrazione, stimata tramite spettroscopia ultravioletta, di 0.4 mg/mL) e la si è trattata con: 9 mL di TEV (enzima che riconosce il sito di taglio all'N terminale e taglia la proteina subito dopo le 6 His), 70 µL di DTT 1 mM e 400 µL di EDTA 0.5 mM. Si è incubato il tutto a 4°C per un'intera notte sotto blanda agitazione. Il giorno seguente si è notato che la soluzione presentava dei precipitati (flocculi) piuttosto grandi, dovuti forse ad una certa instabilità della proteina. Pertanto si è dovuto continuare gli studi utilizzando la proteina contenente le 6 istidine e si è scelto di modificare la soluzione tampone in cui essa si trovava dopo la cromatografia di affinità per portare la proteina in un ambiente più acido. In particolare, la soluzione finale ottenuta è costituita da 0.05 M BIS TRIS propano pH 6, 0.180 M NaCl, 0.0027 M TRIS pH 7.5, 0.175 M imidazolo. Infine la proteina è stata concentrata mediante ultracentrifugazione fino ad un valore di 10 mg/mL.

#### 4.5 Purificazione mediante cromatografia a esclusione dimensionale

Dopo aver eseguito la cromatografia di affinità, le frazioni ottenute in cui è stata rilevata His<sub>6</sub>CagVn2 tramite SDS-PAGE sono state riunite per poter proseguire con un ulteriore passaggio di purificazione. Per questo si è ricorsi alla cromatografia a gel permeazione, che consente di separare le proteine in base alla loro dimensione. Sono stati caricati 5 mL alla volta su colonna "Superdex G-75" (Pharmacia), le cui proprietà sono riassunte in tabella 4.3, per un totale di 4 passaggi cromatografici.

<b>Composizione della matrice</b>	Composti sferici di agarosio e destrano cross-linked
<b>Diametro particelle</b>	34 µm
<b>Intervallo di separazione</b>	3-70 kDa

<b>Intervallo di pH</b>	1-14
<b>Pressione massima</b>	0.5 MPa
<b>Volume colonna</b>	120.637 mL
<b>Efficienza</b>	>1300 piatti teorici/metro

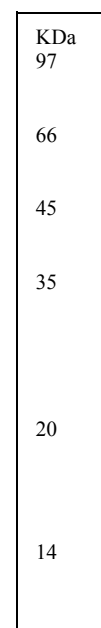
Tabella 4.3 Sono riassunte le caratteristiche della colonna Superdex G-75.

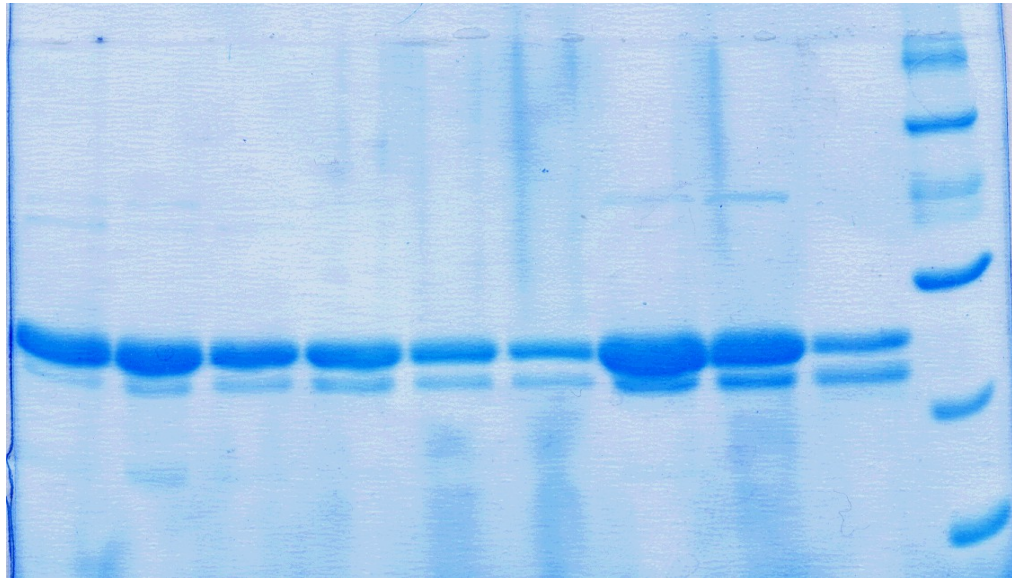
Di seguito sono riportati due dei cromatogrammi registrati:

Figura 4.9 Cromatogrammi a gel permeazione di His6CagVn2. In ascissa è riportato il volume (mL) di eluizione mentre in ordinata l'assorbanza (mAU).

Sono state raccolte, a intervalli regolari, alcune frazioni relative ai picchi cromatografici, le quali sono state fatte correre su gel di poliacrilammide al 15%:

1      2      3      4      5      6      7      8      9      10

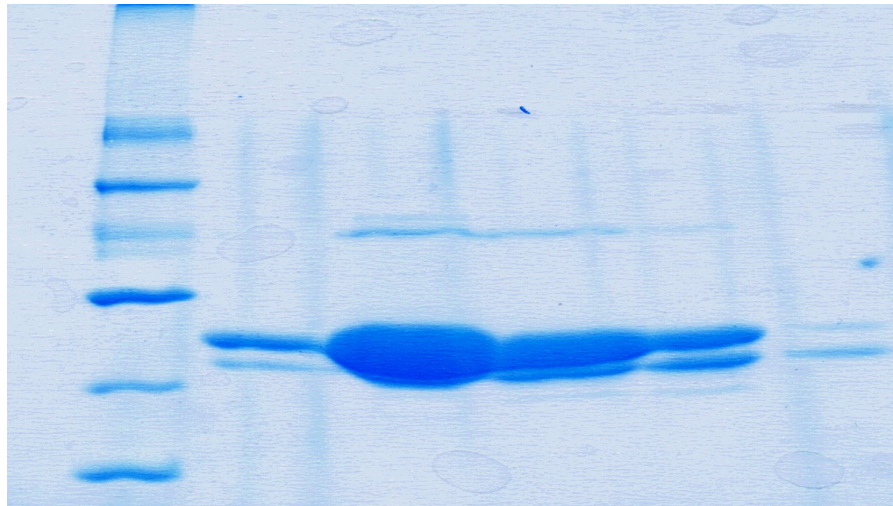




**Figura 4.10 SDS-PAGE di controllo delle frazioni ottenute tramite gel permeazione (I).**

1      2      3      4      5      6

KDa
97
66
45
35
20
14



**Figura 4.11 SDS-PAGE di controllo delle frazioni ottenute tramite gel permeazione (II).**

Si può notare che in tutte le frazioni analizzate è presente la proteina, in forma pura e quasi tutta monomerica. La doppia banda potrebbe indicare una leggera degradazione della proteina, tuttavia il profilo dei gel è stato

ritenuto accettabile per poter considerare His<sub>6</sub>CagVn2 sufficientemente pura e adeguata per portare avanti gli studi successivi.

Sono stati quindi calcolati, mediante un software opportuno, gli integrali dei picchi cromatografici, al fine di stimare la concentrazione di proteina ottenuta mediante la legge di Lambert-Beer, ed stata stimata una quantità totale di proteina pari a 84 mg.

La soluzione contenete la proteina purificata è stata quindi concentrata, tramite ultracentrifugazione, fino ad ottenere una concentrazione di circa 5.6 mg/mL. La soluzione è stata quindi conservata a -80°C, dopo averla suddivisa in aliquote da 300 µL.

Una parte di His<sub>6</sub>CagVn2 è stata inoltre tagliata mediante l'enzima TEV, analogamente a quanto si è fatto per His<sub>6</sub>CagVn1. A tal fine si è incubata la proteina (35 mL) con 1 mL di TEV, 35 µL di DTT 1 mM, 350 µL di EDTA 50 mM; il tutto è stato incubato ON in agitazione a 4°C e, il giorno seguente, la soluzione è stata fatta passare attraverso una colonnina di affinità (la stessa utilizzata per la purificazione) e in questo caso ciò che viene eluito dalla colonna è la proteina tagliata, mentre la coda di istidine si è legata alla fase stazionaria. La presenza di proteina è stata confermata tramite gel SDS-PAGE e misura di uno spettro UV. La concentrazione stimata di proteina tagliata era circa 0.8 mg/mL; questa è stata concentrata, tramite ultracentrifugazione usando un concentratore adeguato, fino ad ottenere una concentrazione di circa 7 mg/mL.

#### **4.6 Dicroismo circolare**

Ultimata la fase di purificazione, His<sub>6</sub>CagVn1 e CagVn2 sono state analizzate mediante dicroismo circolare, in modo da avere una prima idea sulla composizione di struttura secondaria della proteina. Per eseguire questa analisi, i campioni di proteina sono stati diluiti in tampone A fino ad ottenere 30 µL di proteina alla concentrazione di 1 mg/mL. Oltre agli spettri delle proteine, sono stati registrati anche quelli dei soli tamponi, dai quali è stata fatta una sottrazione. Tutti gli spettri sono stati misurati a temperatura ambiente.

La soluzione diluita è stata inserita in una cuvetta di quarzo di cammino ottico 0.02 cm ed è stata quindi condotta la misura. Gli spettri ottenuti, riportati nelle immagini seguenti, sono stati riscalati rispetto alla concentrazione ed al numero di residui aminoacidici della proteina.

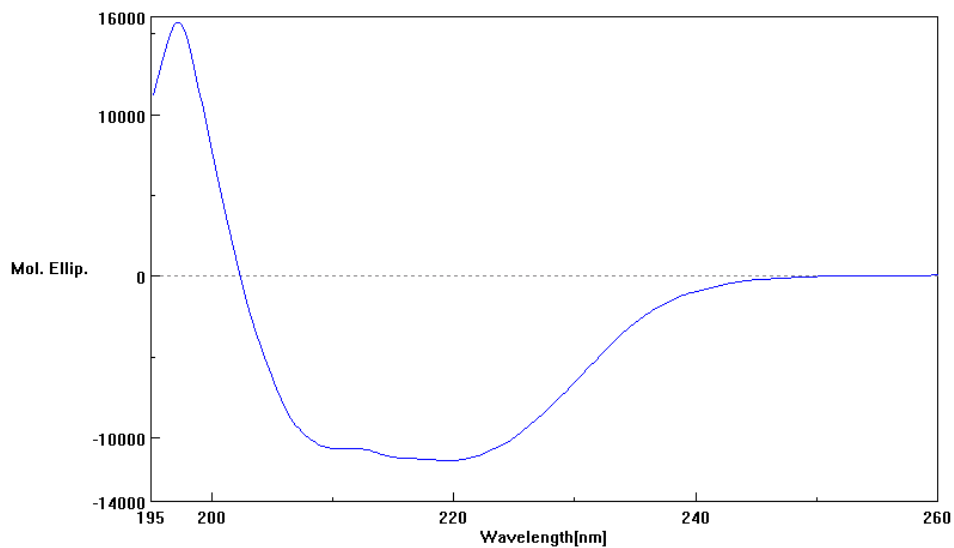


Figura 4.12 Spettro di dicroismo circolare di His6CagVn1.

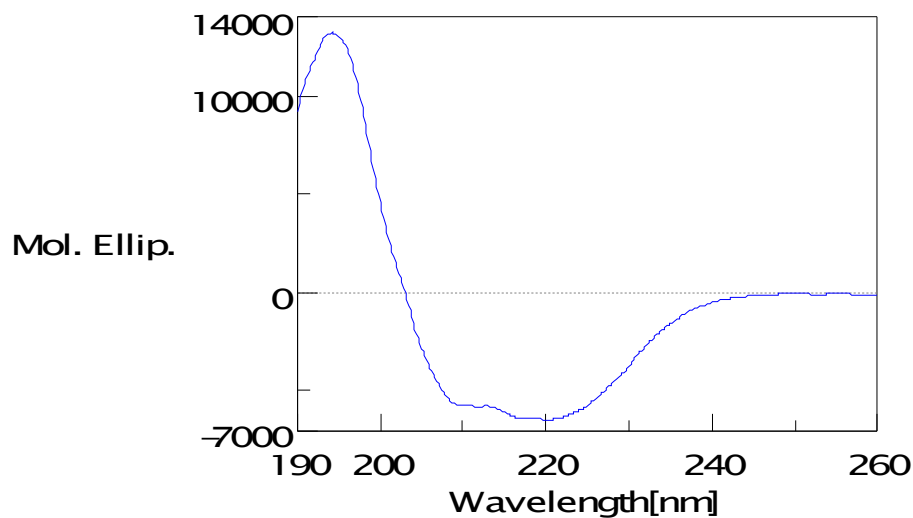


Figura 4.13 Spettro di dicroismo circolare di CagVn2.

I due spettri presentano un andamento molto simile. Nella tabella seguente sono riportate le percentuali di struttura secondaria, calcolate con il programma di interpolazione CDNN, il quale si basa su una serie di spettri base di proteine (almeno 13) aventi struttura nota e utilizza un algoritmo basato sulle reti neurali per la deconvoluzione del segnale ottenuto.

Struttura	% calcolata ( <i>His6CagVn1</i> )	% calcolata ( <i>CagVn2</i> )
-----------	-----------------------------------	-------------------------------

a-elica	35	24
Foglietto $\beta$	17	26
Altro	48	50

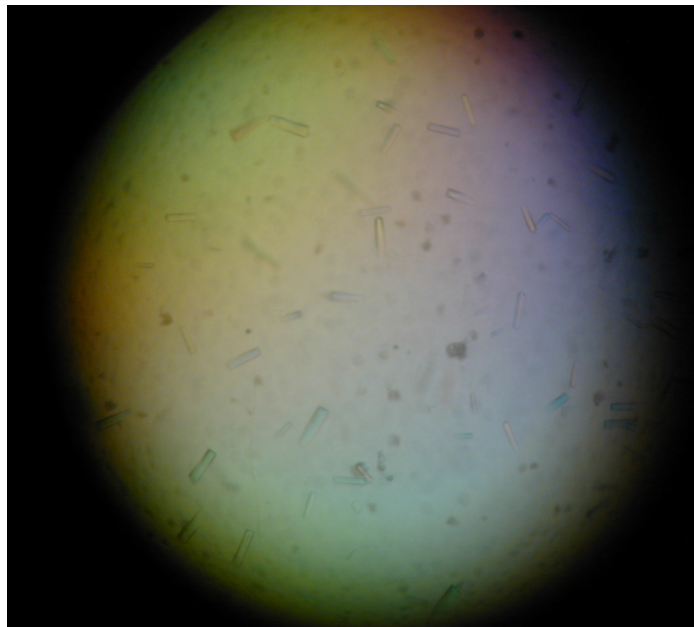
**Tabella 4.4 Riassunto dei contributi di struttura secondaria ottenuti mediante CD.**

In entrambi i casi si ha una percentuale pari al 50% di struttura secondaria ordinata, però la proteina più lunga presenta un contenuto maggiore in  $\alpha$ -elica rispetto ai foglietti  $\beta$ , il che potrebbe indicare una maggiore stabilità in soluzione di questa catena polipeptidica rispetto alla forma più corta. Quest'ultima probabilmente è priva di una parte importante perchè la proteina possa acquisire la struttura terziaria nativa, pertanto potrebbe essere leggermente più flessibile e instabile. Sono noti, infatti, dei casi di proteine instabili, che aggregano o non riescono a ripiegarsi nella conformazione nativa, le quali presentano una tendenza a dare impacchettamenti oligomerici con una struttura analoga a quella dei foglietti  $\beta$ . Tuttavia, tenuto conto della bassa sensibilità della tecnica, i due spettri possono essere ritenuti del tutto comparabili tra loro.

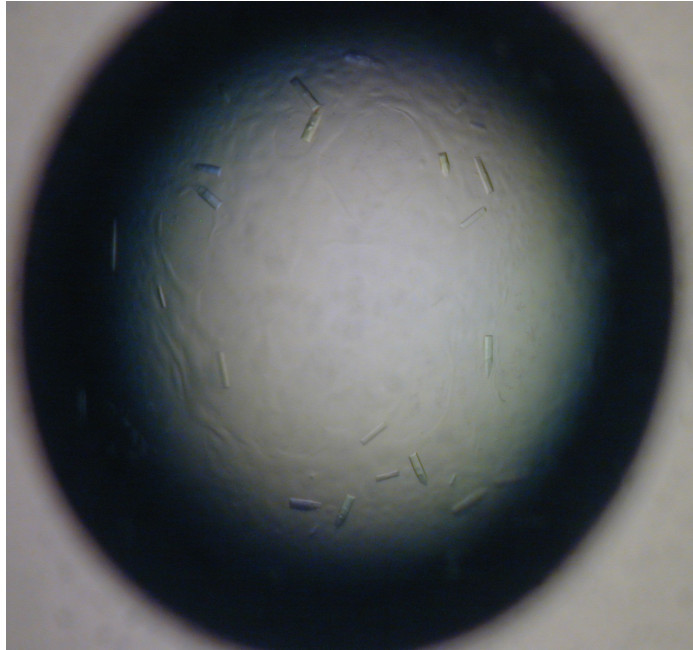
#### **4.7 Cristallizzazione**

Per la ricerca delle condizioni di cristallizzazione è stato sperimentato un approccio a matrice sparsa, utilizzando numerosi precipitanti e variando parametri quali la concentrazione di proteina e di precipitante e la temperatura. In particolare, sono stati impiegati precipitanti forniti dai kit commerciali "Structure Screen 1" (SSI), "Structure Screen 2 (SSII)" (Molecular Dimensions Limited), "Mb Class I", "Mb Class II" e "The PACT suite" (Qiagen). L'intervallo di concentrazioni di proteina utilizzata va dai 20 mg/mL ai 2.5 mg/mL, mentre le temperature testate sono state: 4, 10, 20°C. Per le prime prove di cristallizzazione è stata utilizzata la tecnica hanging drop, depositando 1  $\mu$ L di precipitante sopra 1  $\mu$ L di soluzione proteica. Da questo set di prove (circa 150 per His<sub>6</sub>CagVn2, più di 500 per CagVn2 e circa 400 per His<sub>6</sub>CagVn1), dopo un periodo di incubazione di circa due settimane, sono stati ottenuti dei cristalli di His<sub>6</sub>CagVn1 utilizzando il precipitante numero 26 del kit "Structure Screen II". Quest'ultimo è costituito da 0.2 M ammonio solfato, 0.1 M Mes pH 6.5, 30% w/v PEG monometil etere 5000, mentre la proteina era concentrata a 5 mg/mL. Sono stati quindi eseguiti numerosi tentativi per ottenere cristalli

migliori, sia ripetendo la stessa condizione alle diverse temperature, sia eseguendo una serie di diluizioni della proteina e del precipitante con acqua ultrapura. Si è visto che con questo precipitante i cristalli si formano in modo riproducibile. Questo fatto è importante, perchè un frammento diverso della stessa proteina (l'intervallo aminoacidico 68-252) era già stato cristallizzato in un lavoro precedente, però in quel caso i cristalli non si riproducevano ripetendo le stesse condizioni di cristallizzazione, ma la formazione dei cristalli era un evento occasionale. Inoltre si sono ripetute le prove, con le stesse condizioni ma utilizzando la tecnica sitting drop; in questo caso si sono ottenuti cristalli in un tempo inferiore, tuttavia la dimensione degli stessi è inferiore rispetto a quelli ottenuti tramite hanging drop. Alcuni dei cristalli ottenuti sono riportati nelle immagini seguenti:







**Figura 4.14** Alcuni dei cristalli di His6CagVn1 ottenuti.

Per quanto riguarda invece CagVn2, i numerosi tentativi di cristallizzazione non hanno dato risultati positivi, fatta eccezione per la formazione di microcristalli nel pozzetto relativo alla condizione 44 dello SSII, i quali però dopo pochi giorni si sono sciolti e gli svariati tentativi di riproduzione degli stessi sono stati vani.

#### **4.8 Studi di interazione tra le proteine Cag**

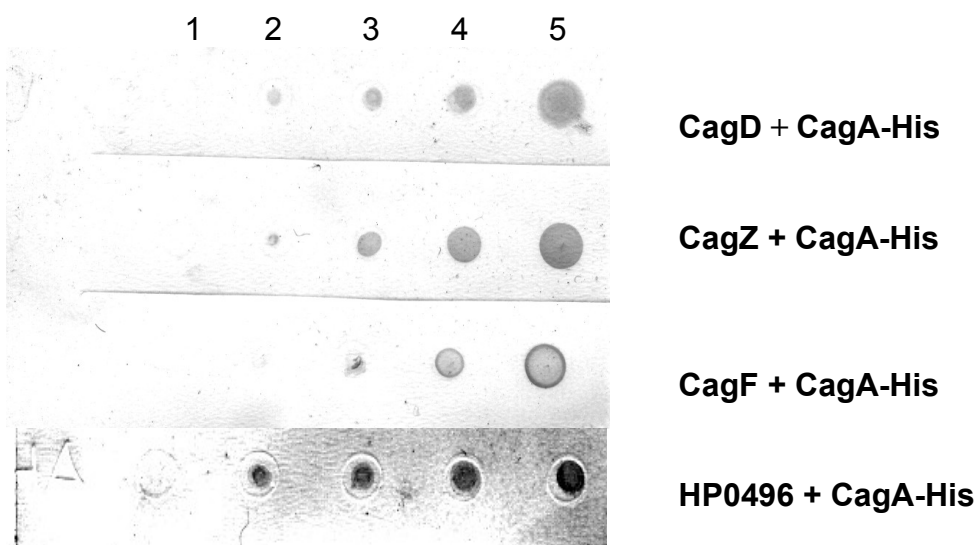
Oltre ad aver effettuato studi riguardanti la struttura di CagV, sono state condotte delle indagini di interazione tra alcune delle proteine dell'isola di patogenicità di *H.pylori* mediante la tecnica del dot-blot. Questi esperimenti danno soltanto un piccolo contributo ad un progetto più ampio che ha lo scopo di stabilire, sulla base di dati provenienti da più tecniche, una mappa delle interazioni tra le varie Cag di cui è nota la struttura tridimensionale. Come descritto nella parte introduttiva, alcuni gruppi di ricerca hanno ottenuto già dei dati, utilizzando metodi biochimici uniti al saggio del doppio ibrido. Poiché il dot-blot è una tecnica semiquantitativa che non presenta una sensibilità molto alta, i risultati ottenuti in questo lavoro servono da un lato per mettere a punto un protocollo efficace di analisi, dall'altro come eventuale conferma di risultati precedenti, oppure

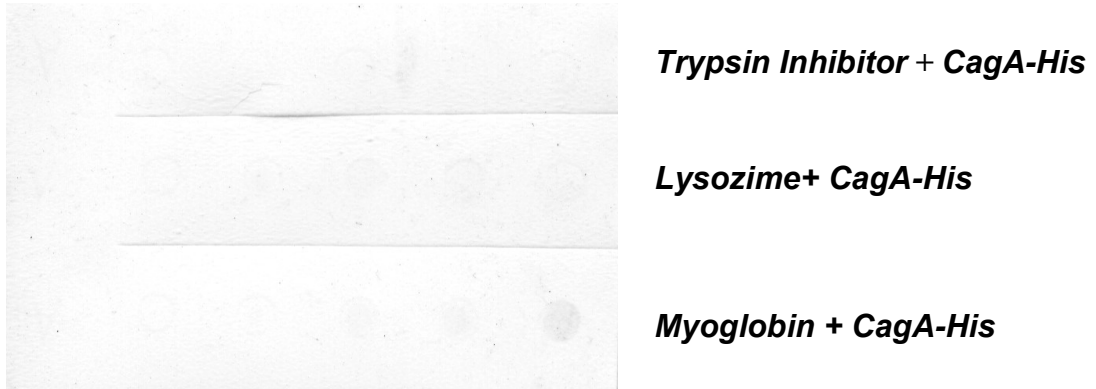
come punto di partenza per affrontare studi più approfonditi condotti anche con altre tecniche. I limiti del dot-blot sono dovuti principalmente al fatto che la proteina si deve legare alla nitrocellulosa, pertanto alcuni siti della stessa potrebbero essere non disponibili per poter interagire con la proteina partner; spesso inoltre ci possono essere delle deboli interazioni aspecifiche dovute alla presenza della carica superficiale della proteina legata al supporto che potrebbero dare dei falsi positivi.

In questo lavoro di tesi sono stati condotti studi di interazione sia tra CagV e altre Cag, che tra proteine Cag diverse da CagV.

CagA, per esempio, avendo un ruolo fondamentale nell'infezione, è stata una delle proteine scelte per questi studi, in particolare è stato utilizzato il dominio centrale della tossina, funzionalizzato con una coda di 6 istidine, gentilmente fornitoci dal Dr. Laurent Terradot.

Di seguito sono riportati alcuni dei risultati che hanno dato esito positivo, con alcuni controlli negativi, eseguiti allo scopo di accertarsi di non ottenere dei falsi positivi:



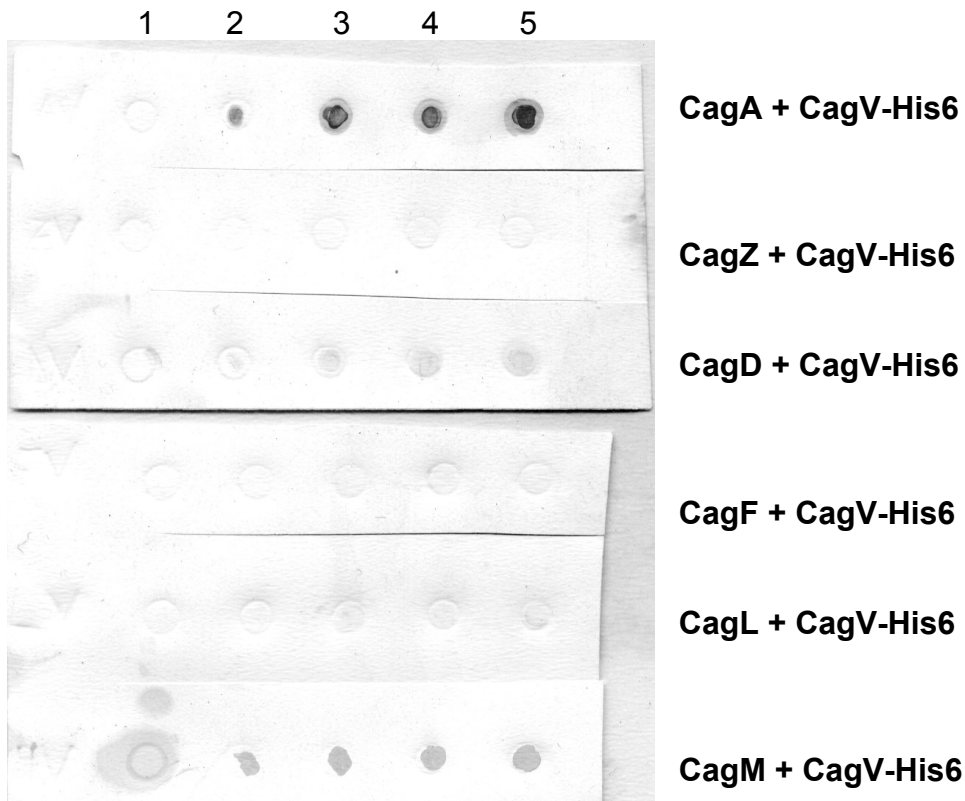


**Figura 4.15** Dot-blot usando come proteina secondaria CagA

1. controllo (nessuna proteina)
2. 0.5 uL proteina
3. 1.0 uL proteina
4. 2.0 uL proteina
5. 4.0 uL proteina

Le concentrazioni delle proteine sono: CagF 3 mg/mL; CagZ 9 mg/mL; CagD 12 mg/mL, HP0496 10 mg/mL, Myoglobin 10 mg/mL; Lysozyme 10 mg/mL; Trypsin Inhibitor 10 mg/mL.

Altri studi di interazione tramite dot-blot sono stati condotti utilizzando His<sub>6</sub>CagV come partner d'interazione; in questo caso si può notare un segnale molto forte usando CagA come prima proteina, mentre negli altri casi il test è risultato negativo, oppure ha dato un segnale più debole, come nel caso di CagM e CagD. Di seguito sono riportate le immagini:



**Figura 4.16 Dot-blot usando come proteina secondaria CagV**

**1 controllo (nessuna proteina)**

**2 0.5 uL proteina**

**3 1.0 uL proteina**

**4 2.0 uL proteina**

**5 4.0 uL proteina**

**CagV aveva una concentrazione di circa 10 mg/mL. Le altre proteine sono state utilizzate ad una concentrazione tale da avere rapporti circa equimolari con CagV.**

L'interazione CagM-CagV era stata rilevata anche tramite il saggio del doppio ibrido [Busler V.J., 2006], mentre quella CagZ-CagV, anch'essa rilevata tramite saggio sopracitato, con il dot-blot non conferma il risultato. Va comunque precisato che in questo esperimento è stata utilizzata solamente una parte della proteina nativa CagV (l'intervallo solubile 68-252) e quindi è probabile che questa fosse priva di qualche amminoacido essenziale per consentire l'interazione.

## **5. MISURA ED ELABORAZIONE DEI DATI**

Per determinare la struttura cristallina tridimensionale di una molecola si utilizzano raggi X aventi una lunghezza d'onda compresa fra 100 e 0.01 nm (1000-0.1Å). Per raggiungere una risoluzione a livello atomico vengono generalmente utilizzate lunghezze d'onda comprese fra 0.7 e 1.9 Å, corrispondenti rispettivamente ad energie di 16.7 e 6.5 KeV.

La diffrazione da parte di un cristallo è un fenomeno che coinvolge tutte le molecole del cristallo; poiché le proteine sono molecole di dimensioni notevolmente maggiori rispetto alla maggior parte delle molecole organiche, all'interno di un cristallo di pari dimensioni sarà presente un numero minore di molecole.

### **5.1 Raccolta dei dati cristallografici**

A causa dell'elevato contenuto di solvente dei cristalli di proteina, e del conseguente loro minor ordine rispetto ai cristalli di molecole più piccole,

essi diffrangono la radiazione X con una certa difficoltà ed in misura minore. Inoltre le proteine sono costituite da elementi leggeri come C, N, O, H, che contengono pochi elettroni per atomo: essendo proprio gli elettroni a generare il fenomeno della diffrazione, gli atomi di questi elementi leggeri diffrangono i raggi X in misura minore rispetto agli atomi di metalli pesanti [Drenth,1994].

A causa di questo basso potere diffrangente, i cristalli proteici richiedono di solito l'uso della radiazione di una sorgente ad alta intensità come il sincrotrone, che ha intensità e brillantezza di svariati ordini di grandezza superiori rispetto a fonti di raggi X convenzionali, quali l'anodo rotante o i tubi sigillati.

Un sincrotrone è una struttura in cui particelle cariche (elettroni o positroni) circolano a velocità prossime a quella della luce. Queste cariche in moto, che variano continuamente la direzione del vettore velocità che le descrive, perdono energia, emettendo radiazioni elettromagnetiche tangenzialmente all'orbita che compiono. La perdita di energia è compensata da un impulso di radiofrequenze applicato ad ogni ciclo. La radiazione emessa viene raccolta dalle *beamlines*, le zone dell'anello in cui sono eseguite le misure.

I dati di diffrazione della nostra proteina sono stati misurati presso la linea ID23-2 del sincrotrone ESRF di Grenoble (Francia).

La raccolta delle intensità della radiazione diffratta è stata effettuata tramite un rivelatore CCD (charge coupled devices) ADSC Q210. I segnali che colpiscono lo schermo del rivelatore, costituito da un *chip* a base di silicio, sono trasmessi tramite fibra ottica al sistema di acquisizione computerizzato.

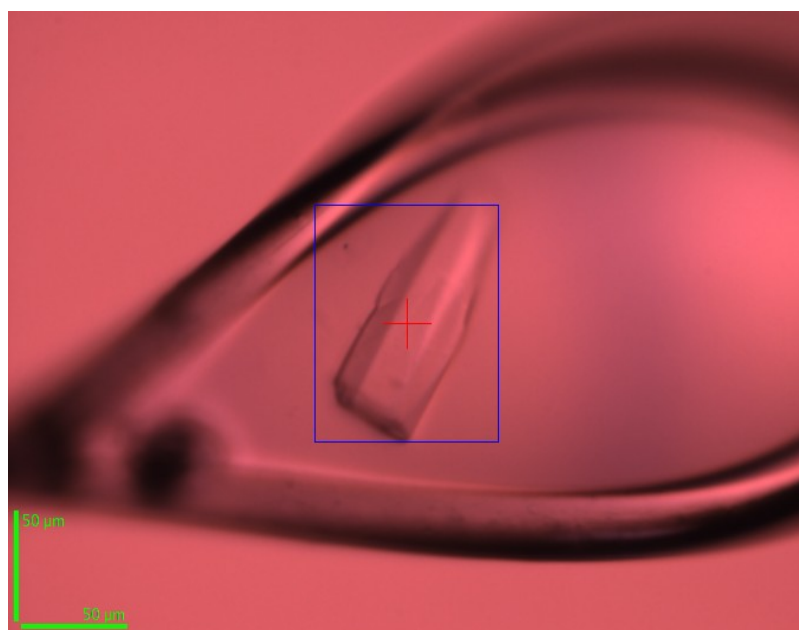
Le misure di diffrazione prevedono la rotazione del cristallo colpito dai raggi X attorno all'asse del perno che lo sostiene. Il rivelatore è posto perpendicolarmente alla direzione di propagazione del fascio e ad una prefissata distanza dal campione. Ogni immagine è il risultato della somma dei segnali raccolti, dopo la rotazione del cristallo di un piccolo angolo (1 grado nel nostro caso). La velocità con la quale avviene questa rotazione dipende dall'intensità della radiazione incidente.

Il cristallo viene prelevato dalla goccia in cui è stato ottenuto mediante un piccolo cappio (*loop*). Quest'ultimo viene inserito su un supporto che permette la rotazione del cristallo durante la raccolta dei dati, ed è sistemato in modo tale che il cristallo sia completamente immerso nel

fascio di raggi X. Il cristallo viene inoltre immerso in un flusso di azoto a 100°K che ha un duplice scopo:

1. ridurre il danneggiamento provocato dal calore per l'elevata intensità del fascio incidente;
2. ridurre la mobilità delle specie radicaliche che si formano per irraggiamento con raggi X e che attraverso successive reazioni chimiche possono distruggere gradualmente l'ordine cristallino.

Nel caso particolare dei cristalli di proteine, l'alto contenuto di solvente può però dare inconvenienti durante il processo di congelamento, portando alla formazione di microcristalli di ghiaccio, che nello spettro di diffrazione, appaiono come anelli, tipici di uno spettro di polveri microcristalline. Inoltre, la formazione di ghiaccio in genere aumenta il disordine a livello del reticolo cristallino. Per questi motivi i cristalli vengono spesso immersi, subito prima della misura, in una soluzione contenente un anti-congelante, detta *cryo-protectant*. Questa soluzione è in genere composta dalla soluzione madre di precipitazione della proteina, alla quale viene aggiunta una sostanza (spesso glicerolo, metilpentandiolo o etilenglicole) con punto di congelamento molto basso. Se l'immersione nel flusso di azoto avviene in modo sufficientemente rapido, si ha la formazione di un solido vetroso che contribuisce solo alla parte amorfa dello spettro. Nel nostro caso i cristalli sono stati immersi per pochi secondi in una soluzione contenente 25% etilenglicole.



**Figura 5.1** Cristallo di His6CagVn1 prelevato mediante “loop”.

Durante la raccolta dei dati il cristallo viene ruotato rispetto a un solo asse, corrispondente all'oscillazione del perno su cui è montato. Per ogni angolo di rotazione viene registrata un'immagine di diffrazione (*frame*), che contiene le “macchie” di diffrazione (riflessioni) raccolte dal rivelatore durante la misura. L'angolo di rotazione per ogni *frame* viene scelto in modo da raccogliere il maggior numero di riflessi possibile, evitando sovrapposizioni tra i segnali e minimizzando la loro registrazione parziale (in modo da avere il minor numero di riflessi “parziali”, cioè che compaiono in parte su un *frame* e in parte su quello successivo). Un angolo piccolo di rotazione permette di separare i punti del reticolo molto vicini tra loro, ma d'altro canto fa aumentare la percentuale di riflessi parziali.

E' necessario infine raccogliere un numero di *frames* tale da permettere di misurare almeno una volta ciascuno i riflessi indipendenti, non legati cioè dalle operazioni di simmetria intrinseche al cristallo analizzato. Questo numero dipende anche dalla capacità del cristallo di sopportare l'intensità dei raggi X a cui è sottoposto.

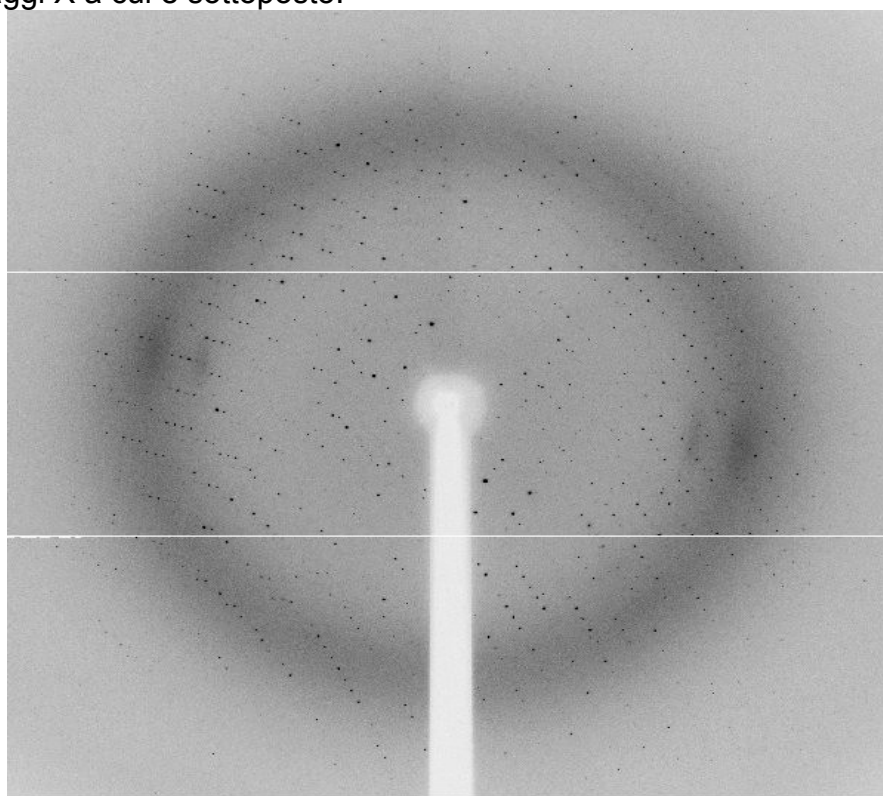


Figura 5.2 Immagine di un frame dello spettro di diffrazione di His<sub>6</sub>CagVn1.

I dati di diffrazione per cristalli della proteina His<sub>6</sub>CagVn1 sono stati raccolti secondo le modalità riportate nella seguente tabella:

Parametri	Descrizione	Valori
-----------	-------------	--------

Ix-ray (Å)	lunghezza d'onda d'irraggiamento	0.97940
Fenditure (nm)	dimensione delle fenditure che regolano la sezione del fascio incidente	20 $\mu$ x 20 $\mu$
$\Delta\phi$	Angolo di rotazione del cristallo in un frame	1°
Risoluzione (Å)	intervallo di risoluzione	103.14-2.85
d	distanza cristallo-rivelatore	389 mm

Tabella 5.1 Modalità di raccolta dei dati di diffrazione.

## 5.2 Elaborazione dei dati

I dati di diffrazione ottenuti devono subire delle elaborazioni prima di poter essere inseriti nel processo di risoluzione strutturale, che porterà all'ottenimento di un modello molecolare tridimensionale. Si è fatto uso di due pacchetti *software*, MOSFLM e SCALA, entrambi contenuti nel pacchetto CCP4 [Potterton et al., 2003].

Ogni *frame* viene registrato sotto forma di un *file*, per essere poi elaborato mediante il programma MOSFLM. Questo *software* è in grado ridefinire la distribuzione dei segnali di diffrazione rispetto ad un sistema di riferimento, cioè di assegnare il valore degli indici cristallografici a ciascun riflesso. In un primo momento, sulla base dei parametri della misura (come la distanza cristallo-rivelatore e la lunghezza d'onda della radiazione incidente), esso determina l'orientazione del cristallo, i parametri di cella ed i probabili gruppi spaziali a cui può appartenere la cella unitaria del cristallo. Ognuno di questi è riportato con un dato di penalità, che si riferisce al grado di corrispondenza tra la posizione effettiva dei riflessi sperimentali e quella prevista sulla base dei diversi gruppi spaziali. Solitamente viene scelto un gruppo spaziale con alta simmetria e bassa penalità, nel nostro caso è stato scelto il gruppo spaziale P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub>, che si è rivelato poi essere corretto.

Dall'elaborazione sono stati ottenuti i seguenti parametri di cella:

a	b	c	a	b	c
59.5623 Å	103.3369Å	130.3304Å	90°	90°	90°

Tabella 5.2 Parametri della cella cristallina.

In un secondo momento il programma MOSFLM determina il valore dell'intensità dei vari riflessi presenti su ogni *frame*, integrando il segnale dell'area corrispondente al riflesso e sottraendo il valore del fondo. I valori delle intensità così ottenute vanno eventualmente sommate tra di loro per i



riflessi “parziali”, cioè presenti su più *frames*. I riflessi vanno poi scalati e mediati per ottenere un insieme di riflessi unici. Alcuni riflessi sono, infatti, presenti più volte in diversi *frames*, a causa della simmetria cristallografica. Questo fatto porta ad una determinazione dell’intensità del riflesso molto più precisa, visto che il suo valore è dato dalla media di più misure. Il numero di volte che uno stesso riflesso si ripete è definito *molteplicità*.

Infine, per ottenere i fattori di struttura dalle intensità integrate si devono introdurre alcuni fattori correttivi (fattore di Lorentz e di polarizzazione, fattore di scala).

Le operazioni sopra descritte vengono eseguite mediante il programma SCALA, appartenente al pacchetto *software* CCP4.

Nella tabella riportata a pagina seguente sono riportate le statistiche finali di uno dei set di dati ottenuti.

<b>Risoluzione massima (Å)</b>	2.7
<b>Riflessi totali misurati</b>	88808
<b>Riflessi indipendenti</b>	21057
<b>Completezza</b>	92.9
<b>Molteplicità</b>	4.2
<b>&lt;I /s(I)&gt;</b>	11.4
<b>Rsym</b>	0.124

Tabella 5.3 Statistiche relative ai dati di diffrazione.

- Rsym = rappresenta l’indice di accordo interno dei dati ed è definito come segue:

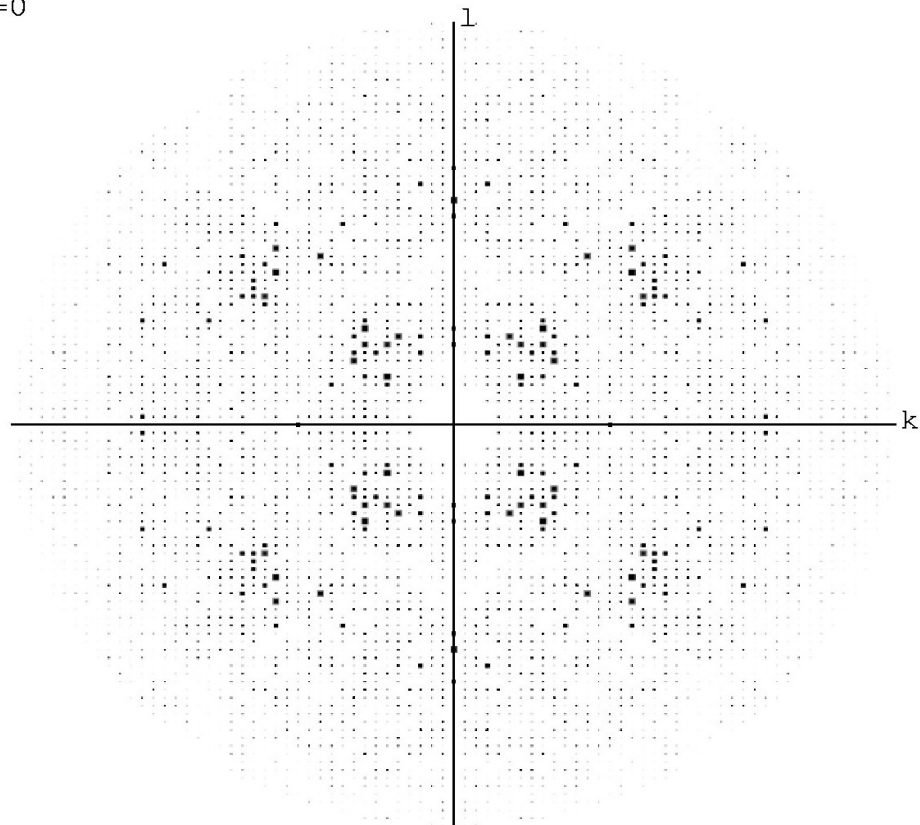
$$R_{sim} = \frac{\sum_{hkl} \sum_i | I_i(hkl) - \bar{I}(hkl) |}{\sum_{hkl} \sum_i I_i(hkl)}$$

dove  $I_i(hkl)$  è l’intensità dell’*i*-esimo segnale di diffrazione e  $\bar{I}(hkl)$  è il valore medio del riflesso (*hkl*);

- *I<sub>av</sub>* = si riferisce all’intensità media dei riflessi;
- *Sigma* = rappresenta la deviazione standard media delle intensità;
- *I<sub>av</sub> / Sigma* = è il rapporto tra l’intensità media dei segnali e la deviazione standard.
- *Rif<sub>tot</sub>* = numero di riflessi totali;
- *Rif<sub>ind</sub>* = numero di riflessi indipendenti;
- *Completezza* = percentuale dei dati di diffrazione raccolti rispetto al numero totale teorico;
- *Molteplicità* = numero di volte in cui mediamente un riflesso equivalente viene misurato.

I dati complessivi ottenuti generano un file di riflessi “unici”, contenenti cioè un solo valore medio per ciascuna misura indipendente. Da questo file è stata ricavata al calcolatore, mediante il programma HKLVIEW, l'immagine sotto riportata, che mostra un piano del reticolo reciproco. Questo piano permette di evidenziare la simmetria del cristallo.

$h=0$



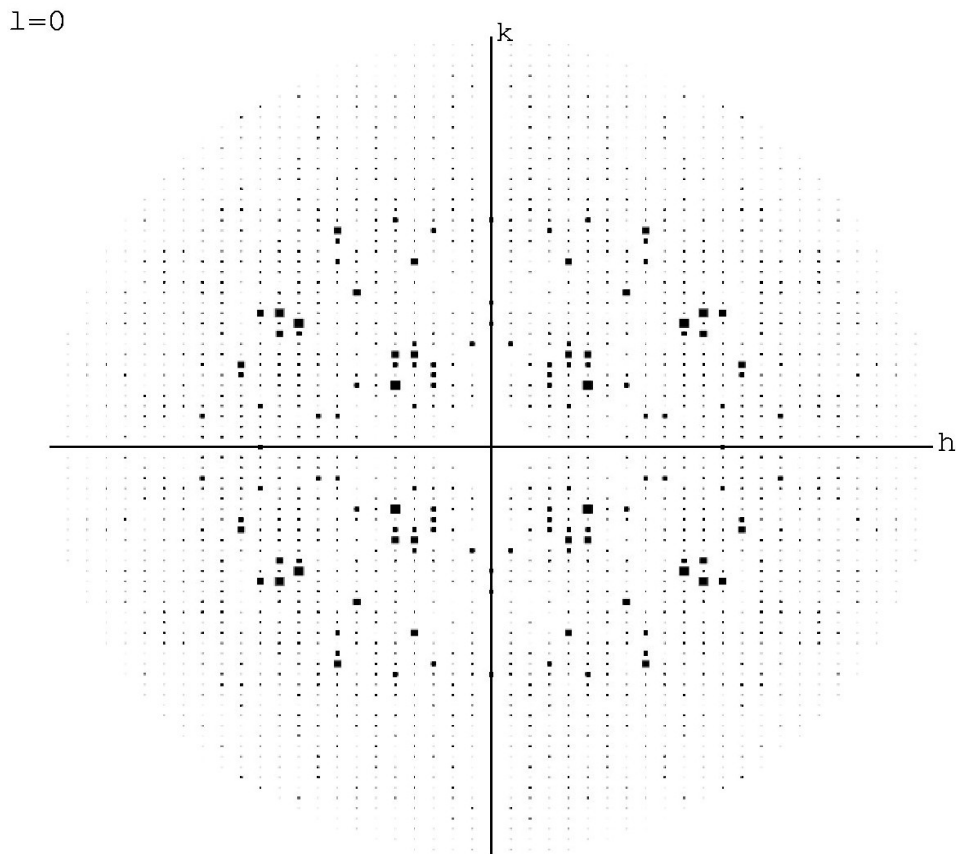


Figura 5.3 Piani reciproci 0hl (in alto) e hk0 (in basso).

### 5.3 Calcolo del numero di molecole per cella unitaria (Z)

Per stimare il numero di molecole contenute nella cella unitaria si è utilizzato il metodo di *Matthews* [Matthews, 1968]. Il parametro  $V_M$  è definito come il rapporto tra il volume della cella unitaria e il peso molecolare totale delle proteine in essa contenute:

$$V_M = \frac{V_{cella}}{Z * PM}$$

dove:

$V_{cella}$  è il volume della cella unitaria;

PM è il peso molecolare della macromolecola;

Z è il prodotto tra il numero di unità asimmetriche della cella unitaria e il numero di molecole di proteina nell'unità asimmetrica, cioè il numero di molecole presenti nella cella unitaria.

$V_M$  è indipendente dal volume dell'unità asimmetrica, ma dipende fortemente dal contenuto di solvente. Poiché da osservazioni empiriche è noto che il termine  $V_M$  è compreso di solito tra 1.6 e 3.5 Å<sup>3</sup>/Da e che la

percentuale di solvente nel cristallo è compresa solitamente tra il 30% e il 70% del volume del cristallo, è possibile determinare il numero di molecole presenti in ogni cella unitaria. Infatti, se indichiamo con  $V_p$  il volume occupato dalla proteina, con  $V_s$  il volume occupato dal solvente e sapendo che:

$$V_p = \frac{1.23}{V_m} \quad \text{e} \quad V_s = 1 - V_p$$

si può calcolare la percentuale di solvente relativa ad un ipotetico numero di proteine contenute nella cella. Il numero di molecole presenti in ogni cella unitaria è così identificato, verificando se la percentuale di solvente cade nell'intervallo sopra citato.

Utilizzando i parametri di cella è possibile calcolare il volume della cella unitaria:

$$V_{\text{cella}} = a b c = 802181.4 \text{ \AA}^3$$

Ora, conoscendo il peso molecolare di un monomero di His<sub>6</sub>CagVn1 (PM = 26471.2 Da) e sapendo che nel gruppo spaziale P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> il numero di unità asimmetriche presenti nella cella unitaria è quattro, si possono stimare le varie percentuali di solvente per un numero crescente di monomeri per unità asimmetrica:

Monomeri per unità asimmetrica	Z	V <sub>M</sub> (Å <sup>3</sup> /Da)	% solvente
1	4	7.576	83.8
2	8	3.788	67.5
3	12	2.525	51.3
4	16	1.894	35.1

**Tabella 5.4** Percentuale di solvente per numero di monomeri nell'unità asimmetrica.

Dai valori sopra riportati, sembra più probabile che nell'unità asimmetrica siano presenti due o tre monomeri, per un contenuto di solvente del 67% o del 51%, rispettivamente. Poiché VirB8 è un dimero, la soluzione corretta sembrerebbe la prima, ma non possiamo escludere che CagV sia presente in forma monomerica.



## **6. CONCLUSIONI**

In questo lavoro di tesi sono stati espressi e purificati due frammenti di lunghezza diversa relativi alla proteina CagV dell'isola di patogenicità di *Helicobacter pylori*, precisamente gli intervalli amminoacidici 59-252 (CagVn1) e 90-252 (CagVn2), entrambi corrispondenti alla regione periplasmatica, più solubile, della proteina. Successivamente sono stati condotti dei tentativi di cristallizzazione delle catene polipeptidiche prodotte e dopo numerose prove sono stati ottenuti i cristalli di CagVn1 unito ad una coda (tag) di 6 istidine. Studi di dicroismo circolare hanno evidenziato che questa catena polipeptidica presenta almeno un 50% di struttura secondaria ordinata, buona parte della quale (circa un 36% del totale) in  $\alpha$ -elica. Allo scopo di costruire un modello di struttura tridimensionale della proteina, sono state effettuate delle misure di diffrazione dei raggi X dei migliori cristalli ottenuti. Queste misure sono state eseguite al sincrotrone di Grenoble (ESRF – European Synchrotron

Radiation Facility) e hanno consentito di ottenere dei dati di diffrazione a 2.7 Å di risoluzione. Questo è un risultato importante, in quanto con questi dati potrebbe essere possibile risolvere la struttura chimica mediante sostituzione molecolare, usando come template le proteine VirB8 di *Agrobacterium tumefaciens* e di *Brucella Suis*, le quali sono presumibilmente omologhe funzionali di CagV. Se in questo modo non sarà possibile risolvere la struttura, si potrà comunque optare per altri metodi, uno dei quali è lo “scattering anomalo”. A tal scopo sarà necessario produrre, purificare e cristallizzare l'isoforma di *CagVn1* contenente, per esempio, Seleno-Metionine, in quanto il Se è un atomo che dà un segnale di diffrazione anomalo, utile per poter calcolare le fasi dei fattori di struttura misurati.

Sono stati eseguiti altresì degli studi di interazione tra varie proteine Cag (compresa CagV) tramite la tecnica del dot-blot, per cominciare ad intraprendere uno studio avente come fine ultimo la realizzazione di una mappa delle interazione tra le varie proteine dell'isola di patogenicità di *H.pylori*.

L'esecuzione di studi strutturali e di interazione delle proteine Cag è di rilevante importanza per comprendere il ruolo e la localizzazione delle stesse nel contesto del sistema di secrezione di tipo IV, consentendo quindi di capire in modo più approfondito quali sono i meccanismi molecolari alla base della patogenesi del batterio. I risultati che si otterranno saranno fondamentali per consentire di effettuare ulteriori studi, di tipo farmacologico, aventi come obiettivo la realizzazione di molecole in grado di inibire l'infezione a livello gastrico del batterio.

## **APPENDICI**

### STRUCTURE SCREEN I COMPOSITION TABLE

1	0.02M Calcium chloride dihydrate	0.1M Na Acetate trihydrate pH	30% v/v 2-methyl-2,4-pentanediol
2	0.2M Ammonium acetate	0.1M Na Acetate trihydrate pH	30% w/v PEG 4000
3	0.2M Ammonium sulphate	0.1M Na acetate trihydrate pH	25% w/v PEG 4000
4	None	0.1M Na acetate trihydrate pH	2.0M Sodium formate
5	None	0.1M Na acetate trihydrate pH	2.0M Ammonium sulphate
6	None	0.1M Na acetate trihydrate pH	8% w/v PEG 4000
7	0.2M Ammonium acetate	0.1M tri-sodium citrate dihydrate	30% w/v PEG 4000
8	0.2M Ammonium acetate	0.1M tri-sodium citrate dihydrate	30% v/v 2-methyl-2,4-pentanediol
9	None	0.1M tri-Sodium citrate dihydrate	20% w/v 2-propanol 20% w/v
10	None	0.1M Na Citrate pH 5.6	1.0M Ammonium dihydrogen
11	0.2M Calcium chloride dihydrate	0.1M Na acetate trihydrate pH	20% v/v 2-propanol



12	None	0.1M Na Cacodylate pH 6.5	1.4M Na acetate trihydrate
13	0.2M tri-sodium citrate dihydrate	0.1M Na Cacodylate pH 6.5	30% v/v 2-propanol
14	0.2M Ammonium sulphate	0.1M Na Cacodylate pH 6.5	30% w/v PFG 8000
15	0.2M Magnesium acetate	0.1M Na Cacodylate pH 6.5	20% PFG 8000
16	0.2M Magnesium acetate	0.1M Na Cacodylate pH 6.5	30% v/v 2-methyl-2,4-pentanediol
17	None	0.1M Imidazole pH 6.5	1.0M Sodium acetate trihydrate
18	0.2M Sodium acetate trihydrate	0.1M Na Cacodylate pH 6.5	30% w/v PFG 8000
19	0.2M Zinc acetate dihydrate	0.1M Na Cacodylate pH 6.5	18% w/v PFG 8000
20	0.2M Calcium acetate hydrate	0.1M Na Cacodylate pH 6.5	18% w/v PFG 8000
21	0.2M tri-sodium citrate dihydrate	0.1M Na HEPES pH 7.5	30% v/v 2-methyl-2,4-pentanediol
22	0.2M Magnesium chloride	0.1M Na HEPES pH 7.5	30% v/v 2-propanol
23	0.2M Calcium chloride dihydrate	0.1M Na HEPES pH 7.5	28% v/v PFG 400
24	0.2M Magnesium chloride	0.1M Na HEPES pH 7.5	30% v/v PFG 400
25	0.2M tri-sodium citrate dihydrate	0.1M Na HEPES pH 7.5	20% v/v 2-propanol
26	None	0.1M Na HEPES pH 7.5	0.8M K Na tartrate tetrahydrate
27	None	0.1M Na HEPES pH 7.5	1.5M Lithium sulphate
28	None	0.1M Na HEPES pH 7.5	0.8M Na dihydrogen phosphate 0.8M K dihydrogen phosphate
29	None	0.1M Na HEPES pH 7.5	1.4M tri-Sodium citrate dihydrate
30	None	0.1M Na HEPES pH 7.5	2% v/v PFG 400 2.0M Amm
31	None	0.1M Na HEPES pH 7.5	10% v/v 2-propanol 20% w/v
32	None	0.1M Tris HCl pH 8.5	2.0M Ammonium sulphate
33	0.2M Magnesium chloride	0.1M Tris HCl pH 8.5	30% w/v PFG 4000
34	0.2M tri-sodium citrate dihydrate	0.1M Tris HCl pH 8.5	30% v/v PFG 400
35	0.2M Lithium sulphate	0.1M Tris HCl pH 8.5	30% w/v PFG 4000
36	0.2M Ammonium acetate	0.1M Tris HCl pH 8.5	30% v/v 2-propanol
37	0.2M Sodium acetate trihydrate	0.1M Tris HCl pH 8.5	30% w/v PFG 4000
38	None	0.1M Tris HCl pH 8.5	8% w/v PFG 8000
39	None	0.1M Tris HCl pH 8.5	2.0M Ammonium dihydrogen
40	None	None	0.4M K Na Tartrate tetrahydrate
41	None	None	0.4M Ammonium dihydrogen
42	0.2M Ammonium sulphate	None	30% w/v PFG 8000
43	0.2M Ammonium sulphate	None	30% w/v PFG 4000
44	None	None	2.0M Ammonium sulphate

45	None	None	4.0M Sodium formate
46	0.05M Potassium dihydrogen	None	20% w/v PEG 8000
47	None	None	30% w/v PEG 1500
48	None	None	0.2M Magnesium formate
49	1.0M Lithium sulphate	None	2% w/v PEG 8000
50	0.5M Lithium sulphate	None	15% w/v PEG 8000

## STRUCTURE SCREEN 2 COMPOSITION TABLE

1	0.5M LiCl	0.5M DMSO	0.5% w/v PEG 8000
2	None	0.5M DMSO	0.5% w/v PEG 8000
3	0.5M LiCl	0.5M DMSO	10% w/v PEG 8000
4	0.5M LiCl	0.5M DMSO	0.5% w/v PEG 1500
5	None	0.5M DMSO	0.5% w/v PEG 1500
6	0.5M LiCl	0.5M DMSO	10% w/v PEG 1500
7	0.5M LiCl	0.5M DMSO	10% w/v PEG 400
8	0.5M LiCl	0.5M DMSO	0.5% w/v PEG 400
9	None	0.5M DMSO	0.5% w/v PEG 400
10	0.5M LiCl	0.5M DMSO	0.5% w/v PEG 8000
11	0.5M LiCl	0.5M DMSO	0.5% w/v PEG 8000
12	None	0.5M DMSO	0.5% w/v PEG 8000
13	0.5M LiCl	0.5M DMSO	10% w/v PEG 8000
14	None	0.5M DMSO	10% w/v PEG 8000
15	0.5M LiCl	0.5M DMSO	10% w/v PEG 1500
16	None	0.5M DMSO	10% w/v PEG 1500
17	0.5M LiCl	0.5M DMSO	10% w/v PEG 1500
18	None	0.5M DMSO	10% w/v PEG 1500
19	0.5M LiCl	0.5M DMSO	10% w/v PEG 400
20	None	0.5M DMSO	10% w/v PEG 400
21	0.5M LiCl	0.5M DMSO	10% w/v PEG 400
22	None	0.5M DMSO	10% w/v PEG 400
23	0.5M LiCl	0.5M DMSO	1.6M Magnesium sulphate
24	None	0.5M DMSO	1.6M Magnesium sulphate
25	0.5M LiCl	0.5M DMSO	0.5% w/v PEG 8000
26	None	0.5M DMSO	0.5% w/v PEG 8000

1	0.1 M tri-Sodium citrate	0.1 M tri-Sodium citrate pH 5.6	10 %(v/v) Isopropanol
2	0.1 M Sodium chloride	0.1 M Sodium acetate pH 4.6	12 %(v/v) Isopropanol
3	0.2 M Calcium chloride	0.1 M Sodium acetate pH 4.6	30 %(w/v) Isopropanol, 20 %(w/v) Glycerol
4		0.1 M TRIS.HCl pH 8.5	0.5 M Ammonium sulfate
5		1 M Na/K phosphate pH 7.5	0.7 M Ammonium sulfate

### Mb Class I COMPOSITION TABLE

1	0.1 M tri-Sodium citrate	0.1 M tri-Sodium citrate pH 5.6	10 %(v/v) Isopropanol
2	0.1 M Sodium chloride	0.1 M Sodium acetate pH 4.6	12 %(v/v) Isopropanol
3	0.2 M Calcium chloride	0.1 M Sodium acetate pH 4.6	30 %(w/v) Isopropanol, 20 %(w/v) Glycerol
4		0.1 M TRIS.HCl pH 8.5	0.5 M Ammonium sulfate
5		1 M Na/K phosphate pH 7.5	0.7 M Ammonium sulfate

6		0.1 M ADA pH 6.5	1.0 M Ammonium sulfate
7		0.1 M TRIS.HCl pH 8.5	1.2 M Ammonium sulfate
8	0.1 M Ammonium acetate		1.4 M Ammonium sulfate, 4 % (w/v) Isopropanol
9		0.1 M tri-Sodium citrate pH 5.6	2.0 M Ammonium sulfate
10	0.25 M Sodium chloride	0.05 M Na/K phosphate pH 7.5	3.5 M Ammonium sulfate
11		0.1 M ADA pH 6.5	1.0 M di-Ammonium phosphate
12	0.1 M Ammonium sulfate	0.1 M HEPES sodium salt pH 7.5	0.5 M di-Sodium phosphate
			0.5 M di-Potassium phosphate
13	0.1 M di-Ammonium phosphate	0.1 M TRIS.HCl pH 8.5	0.5 M di-Sodium phosphate
			0.5 M di-Potassium phosphate
14	0.5 M Lithium chloride		1.0 M tri-Sodium citrate pH 5.6
15		0.1 M TRIS.HCl pH 8.5	0.2 M Lithium sulfate
16	0.1 M K/Na tartrate	0.1 M TRIS.HCl pH 8.5	0.4 M Magnesium sulfate
17	0.1 M Magnesium chloride	0.1 M tri-Sodium citrate pH 5.6	4 % (v/v) MPD
18	0.6 M Magnesium sulfate	0.1 M HEPES sodium salt pH 7.5	4 % (v/v) MPD
19	0.1 M Sodium chloride	0.1 M Sodium acetate pH 4.6	12 % (v/v) MPD
20	0.1 M Sodium chloride	0.1 M tri-Sodium citrate pH 5.6	12 % (v/v) MPD
21		0.1 M ADA pH 6.5	12 % (v/v) MPD
22	0.1 M tri-Sodium citrate	0.1 M HEPES sodium salt pH 7.5	12 % (v/v) MPD
23	0.1 M Lithium sulfate	0.1 M TRIS.HCl pH 8.5	12 % (v/v) MPD
24	0.1 M Sodium chloride	0.1 M TRIS.HCl pH 8.5	12 % (v/v) MPD
25		0.1 M Bis-TRIS Propane pH 7.0	25 % (w/v) MPD
26		0.3 M tri-Sodium citrate pH 5.6	30 % (w/v) MPD
27	0.1 M Lithium sulfate	0.1 M tri-Sodium citrate pH 5.6	4 % (v/v) PEG 400
28	0.3 M Lithium sulfate	0.1 M ADA pH 6.5	4 % (v/v) PEG 400
29	0.6 M Magnesium sulfate	0.1 M HEPES sodium salt pH 7.5	4 % (v/v) PEG 400
30	0.1 M tri-Sodium citrate	0.1 M TRIS.HCl pH 8.5	5 % (v/v) PEG 400
31	0.2 M Calcium chloride	0.1 M HEPES sodium salt pH 7.5	15 % (w/v) PEG 400, 15 % (w/v) Glycerol
32	0.1 M Magnesium chloride	0.1 M Sodium acetate pH 4.6	18 % (v/v) PEG 400
33	0.1 M Sodium chloride	0.1 M tri-Sodium citrate pH 5.6	18 % (v/v) PEG 400
34	0.1 M Magnesium chloride	0.1 M HEPES sodium salt pH 7.5	18 % (v/v) PEG 400
35	0.1 M Ammonium sulfate	0.1 M HEPES sodium salt pH 7.5	18 % (v/v) PEG 400
36	0.05 M Magnesium acetate	0.05 M Sodium acetate pH 4.6	25 % (w/v) PEG 400
37	0.05 M Sodium sulfate	0.05 M TRIS.HCl pH 8.5	30 % (w/v) PEG 400
	0.05 M Lithium sulfate		
38	0.2 M Calcium chloride	0.1 M HEPES sodium salt pH 7.5	48 % (w/v) PEG 400
39		0.01 M TRIS.HCl pH 7.5	20 % (w/v) PEG 550 MME
40	0.1 M Magnesium chloride	0.05 M TRIS.HCl pH 8.5	30 % (w/v) PEG 550 MME
41			35 % (w/v) PEG 600
42	0.35 M Sodium chloride	0.1 M Tricine pH 8.0	28 % (w/v) PEG 1000, 10 % (w/v) Glycerol
43	0.1 M Magnesium chloride, 0.1 M Sodium chloride		10 % (w/v) PEG 1500, 5 % (w/v) Ethanol
44			5 % (w/v) PEG 2000
45	0.5 M Magnesium chloride	0.05 M TRIS.HCl pH 8.5	10 % (w/v) PEG 2000
46	0.02 M tri-Sodium citrate	0.1 M Sodium phosphate pH 6.2	15 % (w/v) PEG 2000
47	0.5 M Sodium chloride	0.1 M Sodium phosphate pH 6.8	15 % (w/v) PEG 2000
48		0.02 M Bis-TRIS propane pH 7.0	15 % (w/v) PEG 2000
49	0.1 M Magnesium chloride	0.05 M HEPES sodium salt pH 7.5	15 % (w/v) PEG 2000
50			15 % (w/v) PEG 2000
51	0.1 M Lithium chloride		15 % (w/v) PEG 2000
52	0.3 M Magnesium nitrate	0.1 M TRIS.HCl pH 8.0	20 % (w/v) PEG 2000, 2 % (w/v) MPD
53	0.3 M Magnesium chloride	0.1 M BICINE pH 9.0	25 % (w/v) PEG 2000, 15 % (w/v) Glycerol
54		0.1 M Sodium acetate pH 4.6	8 % (w/v) PEG 2000 MME
55		0.1 M tri-Sodium citrate pH 5.6	10 % (w/v) PEG 2000 MME, 3 % (w/v) PEG 200
			20 % (w/v) Glycerol
56	0.5 M Sodium chloride	0.05 M TRIS.HCl pH 7.5	12 % (w/v) PEG 2000 MME
57	0.15 M Sodium chloride	0.05 M tri-Sodium citrate pH 5.6	10 % (w/v) PEG 3350
58		0.05 M TRIS.HCl pH 7.5	2 % (w/v) PEG 4000
59	0.1 M Sodium chloride	0.05 M MES sodium salt pH 6.5	5 % (w/v) PEG 4000, 10 % (w/v) Glycerol
60		0.05 M Sodium phosphate pH 6.7	5 % (w/v) PEG 4000
61			5 % (w/v) PEG 4000
62	0.1 M Potassium chloride		5 % (w/v) PEG 4000
63	0.2 M Ammonium sulfate	0.1 M Sodium acetate pH 4.6	10 % (w/v) PEG 4000
64	0.1 M Sodium chloride	0.1 M HEPES sodium salt pH 7.5	10 % (w/v) PEG 4000
65	0.1 M Ammonium sulfate	0.1 M HEPES sodium salt pH 7.5	10 % (w/v) PEG 4000
66	0.5 M Sodium chloride	0.05 M TRIS.HCl pH 8.5	10 % (w/v) PEG 4000
67	0.1 M Lithium sulfate	0.1 M ADA pH 6.5	12 % (w/v) PEG 4000

68	0.1 M Lithium sulfate	0.1 M ADA pH 6.5	12 %(w/v) PEG 4000, 2 %(v/v) Isopropanol
69		0.05 M Sodium phosphate pH 6.8	12 %(w/v) PEG 4000
70	0.5 M Potassium chloride	0.05 M MOPS pH 7.0	12 %(w/v) PEG 4000, 20 %(w/v) Glycerol
71	0.1 M Lithium chloride	0.01 M TRIS.HCl pH 7.5	15 %(w/v) PEG 4000
72	0.5 M Sodium chloride	0.1 M Bis-TRIS propane pH 7.0	20 %(w/v) PEG 4000
73	0.5 M Sodium chloride	0.1 M Sodium phosphate pH 7.0	20 %(w/v) PEG 4000
74	0.15 M Zinc acetate, 0.05 M Zinc chloride	0.05 M TRIS.HCl pH 7.5	20 %(w/v) PEG 4000
75		0.05 M Tricine pH 8.0	22 %(w/v) PEG 4000
76	0.5 M Sodium chloride	0.05 M TRIS.HCl pH 8.5	22 %(w/v) PEG 4000
77			30 %(w/v) PEG 4000
78	0.1 M Magnesium acetate	0.1 M tri-Sodium citrate pH 5.6	10 %(w/v) PEG 5000 MME
79	0.1 M Magnesium sulfate		5 %(w/v) PEG 6000
80	0.15 M Zinc acetate, 0.05 M Zinc chloride	0.05 M TRIS.HCl pH 7.5	10 %(w/v) PEG 6000
81	0.1 M Lithium sulfate	0.1 M tri-Sodium citrate pH 5.6	12 %(w/v) PEG 6000
82	0.15 M Sodium chloride	0.1 M TRIS.HCl pH 8.5	12 %(w/v) PEG 6000
83		0.05 M Sodium succinate pH 6.5	15 %(w/v) PEG 6000
84	0.025 M Potassium phosphate		12 %(w/v) PEG 8000, 10 %(w/v) MPD
85	0.1 M Magnesium acetate	0.1 M tri-Sodium citrate pH 5.6	8 %(w/v) PEG 10000
86			0.05 M Potassium phosphate pH 8.0
87			1.5 M Potassium phosphate pH 7.0
88	0.1 M Lithium sulfate	0.1 M HEPES sodium salt pH 7.5	0.1 M K/Na tartrate
89		0.1 M TRIS.HCl pH 8.5	0.1 M Sodium acetate
90		0.1 M tri-Sodium citrate pH 5.6	0.1 M Sodium chloride
91		0.1 M TRIS.HCl pH 8.5	0.1 M Sodium chloride
92		0.1 M Sodium acetate pH 4.6	1.5 M Sodium chloride
93	0.1 M Sodium formate		2.0 M Sodium chloride
94			0.1 M tri-Sodium citrate pH 4.8
95		0.1 M HEPES sodium salt pH 7.5	1.0 M tri-Sodium citrate
96			1.0 M tri-Sodium citrate

## Mb Class II suite COMPOSITION TABLE

1	0.1 M Lithium sulfate		0.1 M Sodium acetate pH 4.6	1.0 M Ammonium phosphate
2			0.1 M tri-Sodium citrate pH 5.6	1.0 M Ammonium phosphate
3			0.1 M TRIS pH 8.5	2.0 M Ammonium phosphate
4			0.1 M Sodium acetate pH 4.6	2.0 M Ammonium sulfate
5			0.1 M ADA pH 6.5	1.0 M Ammonium sulfate
6				2.0 M Ammonium sulfate
7			0.1 M TRIS pH 8.5	2.0 M Ammonium sulfate

8			0.1 M tri-Sodium citrate pH 5.5	2.5 M Ammonium sulfate
9			0.1 M MES pH 6.5	2.5 M Ammonium sulfate
10			0.1 M HEPES sodium salt pH 7.5	2.5 M Ammonium sulfate
11	2 %(v/v) PEG 400		0.1 M HEPES sodium salt pH 7.5	2.0 M Ammonium sulfate
12				1.0 M di-Ammonium phosphate
13	0.5 M di-Sodium phosphate	0.1 M Ammonium sulfate		0.5 M di-Potassium phosphate
14			0.1 M TRIS pH 8.5	1.5 M Lithium sulfate
15			0.1 M HEPES sodium salt pH 7.5	1.5 M Lithium sulfate
16			0.1 M Sodium acetate pH 4.6	1.0 M Magnesium sulfate
17			0.1 M tri-Sodium citrate pH 5.6	1.0 M Magnesium sulfate
18	0.1 M Lithium sulfate		0.1 M ADA pH 6.5	1.0 M Magnesium sulfate
19		0.1 M Sodium chloride	0.1 M tri-Sodium citrate pH 5.5	30 %(v/v) PEG 400
20			0.1 M MES pH 6.5	30 %(v/v) PEG 400
21		0.1 M Sodium chloride	0.1 M MES pH 6.5	30 %(v/v) PEG 400
22			0.1 M MOPS pH 7.0	30 %(v/v) PEG 400
23		0.1 M Sodium chloride	0.1 M MOPS pH 7.0	30 %(v/v) PEG 400
24			0.1 M HEPES sodium salt pH 7.5	30 %(v/v) PEG 400
25		0.1 M Sodium chloride	0.1 M HEPES sodium salt pH 7.5	30 %(v/v) PEG 400
26		0.1 M Sodium chloride	0.1 M TRIS pH 8.5	30 %(v/v) PEG 400
27	0.1 M Lithium sulphate	0.1 M Sodium chloride	0.1 M tri-Sodium citrate pH 3.5	30 %(v/v) PEG 400
28	0.1 M Lithium sulphate	0.1 M Sodium chloride	0.1 M tri-Sodium citrate pH 5.5	30 %(v/v) PEG 400
29	0.1 M Lithium sulphate	0.1 M Sodium chloride	0.1 M MES pH 6.5	30 %(v/v) PEG 400
30	0.1 M Lithium sulphate	0.1 M Sodium chloride	0.1 M HEPES sodium salt pH 7.5	30 %(v/v) PEG 400
31	0.1 M Lithium sulphate	0.1 M Sodium chloride	0.1 M TRIS pH 8.5	30 %(v/v) PEG 400
32	0.1 M Lithium sulphate	0.1 M Sodium chloride	0.1 M CAPSO pH 9.5	30 %(v/v) PEG 400
33	0.1 M Magnesium chloride	0.1 M Sodium chloride	0.1 M Sodium acetate pH 4.5	30 %(v/v) PEG 400
34	0.1 M Magnesium chloride	0.1 M Sodium chloride	0.1 M tri-Sodium citrate pH 5.5	30 %(v/v) PEG 400
35	0.1 M Magnesium chloride	0.1 M Sodium chloride	0.1 M MES pH 6.5	30 %(v/v) PEG 400
36	0.1 M Magnesium chloride	0.1 M Sodium chloride	0.1 M HEPES sodium salt pH 7.5	30 %(v/v) PEG 400
37	0.1 M Magnesium chloride	0.1 M Sodium chloride	0.1 M TRIS pH 8.5	30 %(v/v) PEG 400
38	0.1 M Magnesium chloride	0.1 M Sodium chloride	0.1 M CAPSO pH 9.5	30 %(v/v) PEG 400
39	0.1 M Magnesium chloride		0.1 M Sodium acetate pH 4.6	30 %(v/v) PEG 400
40	0.1 M Sodium chloride		0.1 M tri-Sodium citrate pH 5.6	30 %(v/v) PEG 400
41	0.1 M Lithium sulfate		0.1 M tri-Sodium citrate pH 5.6	30 %(v/v) PEG 400
42	0.3 M Lithium sulfate		0.1 M ADA pH 6.5	30 %(v/v) PEG 400
43	0.1 M Magnesium chloride		0.1 M HEPES sodium salt pH 7.5	30 %(v/v) PEG 400
44	0.1 M Ammonium sulfate		0.1 M HEPES sodium salt pH 7.5	30 %(v/v) PEG 400
45	0.2 M tri-Sodium citrate		0.1 M TRIS pH 8.5	30 %(v/v) PEG 400
46		0.1 M Sodium chloride	0.1 M tri-Sodium citrate pH 5.5	12 %(w/v) PEG 4000
47			0.1 M MES pH 6.5	12 %(w/v) PEG 4000
48		0.1 M Sodium chloride	0.1 M MES pH 6.5	12 %(w/v) PEG 4000
49			0.1 M MOPS pH 7.0	12 %(w/v) PEG 4000
50		0.1 M Sodium chloride	0.1 M MOPS pH 7.0	12 %(w/v) PEG 4000
51			0.1 M HEPES sodium salt pH 7.5	12 %(w/v) PEG 4000
52		0.1 M Sodium chloride	0.1 M HEPES sodium salt pH 7.5	12 %(w/v) PEG 4000
53		0.1 M Sodium chloride	0.1 M TRIS pH 8.5	12 %(w/v) PEG 4000
54	0.1 M Lithium sulphate	0.1 M Sodium chloride	0.1 M Sodium acetate pH 4.5	12 %(w/v) PEG 4000
55	0.1 M Lithium sulphate	0.1 M Sodium chloride	0.1 M tri-Sodium citrate pH 5.5	12 %(w/v) PEG 4000
56	0.1 M Lithium sulphate	0.1 M Sodium chloride	0.1 M MES pH 6.5	12 %(w/v) PEG 4000
57	0.1 M Lithium sulphate	0.1 M Sodium chloride	0.1 M HEPES sodium salt pH 7.5	12 %(w/v) PEG 4000
58	0.1 M Lithium sulphate	0.1 M Sodium chloride	0.1 M TRIS pH 8.5	12 %(w/v) PEG 4000
59	0.1 M Lithium sulphate	0.1 M Sodium chloride	0.1 M CAPSO pH 9.5	12 %(w/v) PEG 4000
60	0.1 M Magnesium chloride	0.1 M Sodium chloride	0.1 M tri-Sodium citrate pH 3.5	12 %(w/v) PEG 4000
61	0.1 M Magnesium chloride	0.1 M Sodium chloride	0.1 M tri-Sodium citrate pH 5.5	12 %(w/v) PEG 4000
62	0.1 M Magnesium chloride	0.1 M Sodium chloride	0.1 M MES pH 6.5	12 %(w/v) PEG 4000
63	0.1 M Magnesium chloride	0.1 M Sodium chloride	0.1 M HEPES sodium salt pH 7.5	12 %(w/v) PEG 4000
64	0.1 M Magnesium chloride	0.1 M Sodium chloride	0.1 M TRIS pH 8.5	12 %(w/v) PEG 4000
65	0.1 M Magnesium chloride	0.1 M Sodium chloride	0.1 M CAPSO pH 9.5	12 %(w/v) PEG 4000
66	0.1 M Zinc acetate		0.1 M Sodium acetate pH 4.6	12 %(w/v) PEG 4000
67	0.2 M Ammonium sulfate		0.1 M Sodium acetate pH 4.6	12 %(w/v) PEG 4000
68			0.1 M Sodium acetate pH 4.6	12 %(w/v) PEG 4000
69	0.1 M Lithium sulfate		0.1 M tri-Sodium citrate pH 5.6	12 %(w/v) PEG 4000
70	0.1 M Sodium chloride		0.1 M tri-Sodium citrate pH 5.6	12 %(w/v) PEG 4000
71	0.1 M Lithium sulfate		0.1 M ADA pH 6.5	12 %(w/v) PEG 4000
72	0.1 M Sodium chloride		0.1 M HEPES sodium salt pH 7.5	12 %(w/v) PEG 4000
73	0.1 M Ammonium sulfate		0.1 M HEPES sodium salt pH 7.5	12 %(w/v) PEG 4000
74	0.2 M Magnesium chloride		0.1 M TRIS pH 8.5	12 %(w/v) PEG 4000
75	0.2 M Lithium sulfate		0.1 M TRIS pH 8.5	12 %(w/v) PEG 4000
76	0.2 M Ammonium sulfate			12 %(w/v) PEG 4000
77	0.1 M Sodium chloride		0.1 M Sodium acetate pH 4.6	12 %(w/v) PEG 6000

78	0.1 M Magnesium chloride		0.1 M Sodium acetate pH 4.6	12 %(w/v) PEG 6000
79	0.1 M Magnesium chloride		0.1 M ADA pH 6.5	12 %(w/v) PEG 6000
80	0.1 M di-Ammonium phosphate		0.1 M TRIS pH 8.5	12 %(w/v) PEG 6000
81	1.0 M Lithium sulfate			2 %(w/v) PEG 8000
82	0.2 M Sodium acetate		0.1 M MES pH 6.5	10 %(w/v) PEG 8000
83	0.2 M Zinc acetate		0.1 M Imidazole pH 6.5	10 %(w/v) PEG 8000
84	0.2 M Calcium acetate		0.1 M MES pH 6.5	10 %(w/v) PEG 8000
85			0.1 M TRIS pH 8.5	10 %(w/v) PEG 8000
86	0.2 M Ammonium sulfate			10 %(w/v) PEG 8000
87	0.5 M Lithium sulfate			10 %(w/v) PEG 8000
88			0.1 M HEPES sodium salt pH 7.5	1.5 M Potassium phosphate
89			0.1 M TRIS pH 8.5	1.5 M Potassium phosphate
90			0.1 M HEPES sodium salt pH 7.5	1.0 M K/Na tartrate
91			0.1 M MES pH 6.5	1.4 M Sodium acetate
92			0.1 M HEPES sodium salt pH 7.5	1.4 M tri-Sodium citrate
93				2.0 M Sodium formate
94				4.0 M Sodium formate
95			0.1 M tri-Sodium citrate pH 5.5	1.5 M Sodium phosphate
96			0.1 M MES pH 6.5	1.5 M Sodium phosphate

## THE PACT SUITE COMPOSITION TABLE

	<b>SALT</b>	<b>BUFFER</b>	<b>PRECIPITANT</b>
1		0.1M SPG buffer pH 4	25% w/v PEG 1500
2		0.1M SPG buffer pH 5	25% w/v PEG 1500
3		0.1M SPG buffer pH 6	25% w/v PEG 1500
4		0.1M SPG buffer pH 7	25% w/v PEG 1500
5		0.1M SPG buffer pH 8	25% w/v PEG 1500
6		0.1M SPG buffer pH 9	25% w/v PEG 1500

7	0.2M Sodium chloride	0.1M Sodium acetate pH 5	20% w/v PEG 6000
8	0.2M Ammonium chloride	0.1M Sodium acetate pH 5	20% w/v PEG 6000
9	0.2M Lithium chloride	0.1M Sodium acetate pH 5	20% w/v PEG 6000
10	0.2M Magnesium chloride	0.1M Sodium acetate pH 5	20% w/v PEG 6000
11	0.2M Calcium chloride	0.1M Sodium acetate pH 5	20% w/v PEG 6000
12	0.01M Zinc chloride	0.1M Sodium acetate pH 5	20% w/v PEG 6000
13		0.1M MIB buffer pH 4	25% w/v PEG 1500
14		0.1M MIB buffer pH 5	25% w/v PEG 1500
15		0.1M MIB buffer pH 6	25% w/v PEG 1500
16		0.1M MIB buffer pH 7	25% w/v PEG 1500
17		0.1M MIB buffer pH 8	25% w/v PEG 1500
18		0.1M MIB buffer pH 9	25% w/v PEG 1500
19	0.2M Sodium chloride	0.1M MES pH 6	20% w/v PEG 6000
20	0.2M Ammonium chloride	0.1M MES pH 6	20% w/v PEG 6000
21	0.2M Lithium chloride	0.1M MES pH 6	20% w/v PEG 6000
22	0.2M Magnesium chloride	0.1M MES pH 6	20% w/v PEG 6000
23	0.2M Calcium chloride	0.1M MES pH 6	20% w/v PEG 6000
24	0.01M Zinc chloride	0.1M MES pH 6	20% w/v PEG 6000
25		0.1M PCB buffer pH 4	25% w/v PEG 1500
26		0.1M PCB buffer pH 5	25% w/v PEG 1500
27		0.1M PCB buffer pH 6	25% w/v PEG 1500
28		0.1M PCB buffer pH 7	25% w/v PEG 1500
29		0.1M PCB buffer pH 8	25% w/v PEG 1500
30		0.1M PCB buffer pH 9	25% w/v PEG 1500
31	0.2M Sodium chloride	0.1M HEPES pH 7	20% w/v PEG 6000
32	0.2M Ammonium chloride	0.1M HEPES pH 7	20% w/v PEG 6000
33	0.2M Lithium chloride	0.1M HEPES pH 7	20% w/v PEG 6000
34	0.2M Magnesium chloride	0.1M HEPES pH 7	20% w/v PEG 6000
35	0.2M Calcium chloride	0.1M HEPES pH 7	20% w/v PEG 6000
36	0.01M Zinc chloride	0.1M HEPES pH 7	20% w/v PEG 6000
37		0.1M MMT buffer pH 4	25% w/v PEG 1500
38		0.1M MMT buffer pH 5	25% w/v PEG 1500
39		0.1M MMT buffer pH 6	25% w/v PEG 1500
40		0.1M MMT buffer pH 7	25% w/v PEG 1500
41		0.1M MMT buffer pH 8	25% w/v PEG 1500
42		0.1M MMT buffer pH 9	25% w/v PEG 1500
43	0.2M Sodium chloride	0.1M Tris pH 8	20% w/v PEG 6000
44	0.2M Ammonium chloride	0.1M Tris pH 8	20% w/v PEG 6000
45	0.2M Lithium chloride	0.1M Tris pH 8	20% w/v PEG 6000
46	0.2M Magnesium chloride	0.1M Tris pH 8	20% w/v PEG 6000
47	0.2M Calcium chloride	0.1M Tris pH 8	20% w/v PEG 6000
48	0.01M Zinc chloride	0.1M Tris pH 8	20% w/v PEG 6000
49	0.2M Sodium fluoride		20% w/v PEG 3350
50	0.2M Sodium bromide		20% w/v PEG 3350
51	0.2M Sodium iodide		20% w/v PEG 3350
52	0.2M Potassium thiocyanate		20% w/v PEG 3350
53	0.2M Sodium nitrate		20% w/v PEG 3350
54	0.2M Sodium formate		20% w/v PEG 3350
55	0.2M Sodium acetate		20% w/v PEG 3350
56	0.2M Sodium sulphate		20% w/v PEG 3350
57	0.2M Potassium/sodium tartrate		20% w/v PEG 3350
58	0.2M Sodium/potassium phosphate		20% w/v PEG 3350
59	0.2M Sodium citrate		20% w/v PEG 3350
60	0.2M Sodium malonate		20% w/v PEG 3350
61	0.2M Sodium fluoride	0.1M Bis Tris propane pH 6.5	20% w/v PEG 3350
62	0.2M Sodium bromide	0.1M Bis Tris propane pH 6.5	20% w/v PEG 3350
63	0.2M Sodium iodide	0.1M Bis Tris propane pH 6.5	20% w/v PEG 3350
64	0.2M Potassium thiocyanate	0.1M Bis Tris propane pH 6.5	20% w/v PEG 3350
65	0.2M Sodium nitrate	0.1M Bis Tris propane pH 6.5	20% w/v PEG 3350
66	0.2M Sodium formate	0.1M Bis Tris propane pH 6.5	20% w/v PEG 3350
67	0.2M Sodium acetate	0.1M Bis Tris propane pH 6.5	20% w/v PEG 3350
68	0.2M Sodium sulphate	0.1M Bis Tris propane pH 6.5	20% w/v PEG 3350
69	0.2M Potassium/sodium tartrate	0.1M Bis Tris propane pH 6.5	20% w/v PEG 3350
70	0.2M Sodium/potassium phosphate	0.1M Bis Tris propane pH 6.5	20% w/v PEG 3350
71	0.2M Sodium citrate	0.1M Bis Tris propane pH 6.5	20% w/v PEG 3350
72	0.2M Sodium malonate	0.1M Bis Tris propane pH 6.5	20% w/v PEG 3350
73	0.2M Sodium fluoride	0.1M Bis Tris propane pH 7.5	20% w/v PEG 3350
74	0.2M Sodium bromide	0.1M Bis Tris propane pH 7.5	20% w/v PEG 3350
75	0.2M Sodium iodide	0.1M Bis Tris propane pH 7.5	20% w/v PEG 3350
76	0.2M Potassium thiocyanate	0.1M Bis Tris propane pH 7.5	20% w/v PEG 3350
77	0.2M Sodium nitrate	0.1M Bis Tris propane pH 7.5	20% w/v PEG 3350
78	0.2M Sodium formate	0.1M Bis Tris propane pH 7.5	20% w/v PEG 3350
79	0.2M Sodium acetate	0.1M Bis Tris propane pH 7.5	20% w/v PEG 3350
80	0.2M Sodium sulphate	0.1M Bis Tris propane pH 7.5	20% w/v PEG 3350



81	0.2M Potassium/sodium tartarate	0.1M Bis Tris propane pH 7.5	20% w/v PEG 3350
82	0.2M Sodium/potassium phosphate	0.1M Bis Tris propane pH 7.5	20% w/v PEG 3350
83	0.2M Sodium citrate	0.1M Bis Tris propane pH 7.5	20% w/v PEG 3350
84	0.2M Sodium malonate	0.1M Bis Tris propane pH 7.5	20% w/v PEG 3350
85	0.2M Sodium fluoride	0.1M Bis Tris propane pH 8.5	20% w/v PEG 3350
86	0.2M Sodium bromide	0.1M Bis Tris propane pH 8.5	20% w/v PEG 3350
87	0.2M Sodium iodide	0.1M Bis Tris propane pH 8.5	20% w/v PEG 3350
88	0.2M Potassium thiocyanate	0.1M Bis Tris propane pH 8.5	20% w/v PEG 3350
89	0.2M Sodium nitrate	0.1M Bis Tris propane pH 8.5	20% w/v PEG 3350
90	0.2M Sodium formate	0.1M Bis Tris propane pH 8.5	20% w/v PEG 3350
91	0.2M Sodium acetate	0.1M Bis Tris propane pH 8.5	20% w/v PEG 3350
92	0.2M Sodium sulphate	0.1M Bis Tris propane pH 8.5	20% w/v PEG 3350
93	0.2M Potassium/sodium tartrate	0.1M Bis Tris propane pH 8.5	20% w/v PEG 3350
94	0.2M Sodium/potassium phosphate	0.1M Bis Tris propane pH 8.5	20% w/v PEG 3350
95	0.2M Sodium citrate	0.1M Bis Tris propane pH 8.5	20% w/v PEG 3350
96	0.2M Sodium malonate	0.1M Bis Tris propane pH 8.5	20% w/v PEG 3350

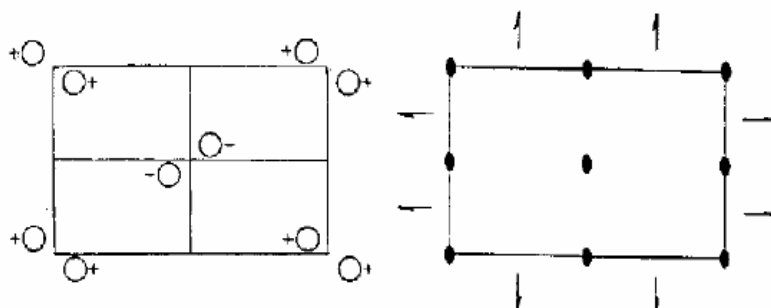
## CARATTERISTICHE DEL GRUPPO P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub>

$P2_12_12$   
 $D_2^3$

No. 18

$P2_12_12$

222 Orthorhombic



Origin at 112 in plane of  $2_12_1$

Number of positions, Wyckoff notation, and point symmetry				Co-ordinates of equivalent positions	Conditions limiting possible reflections
4	c	1		$x, y, z; \bar{x}, \bar{y}, z; \frac{1}{2} + x, \frac{1}{2} - y, z; \frac{1}{2} - x, \frac{1}{2} + y, z.$	General: $hkl:$ } $0kl:$ } No conditions $h0l:$ } $hk0:$ } $h00:$ $h=2n$ $0k0:$ $k=2n$ $00l:$ No conditions
2	b	2		$0, \frac{1}{2}, z; \frac{1}{2}, 0, z.$	Special: as above, plus $hk0: h+k=2n$
2	a	2		$0, 0, z; \frac{1}{2}, \frac{1}{2}, z.$	
Symmetry of special projections					
$(001) pgg; a'=a, b'=b$				$(100) pmg; b'=b, c'=c$	$(010) pgm; c'=c, a'=a$

## ***BIBLIOGRAFIA***

Axon A. T., **Is *Helicobacter pylori* transmitted by the gastro-oral route?**, *Aliment Pharmacol Ther* 1995;9(6):585-8

Amieva MR, Vogelmann R, Covacci A, Tompkins LS, Nelson WJ, Falkow S. **Disruption of the epithelial apical-junctional complex by *Helicobacter pylori* CagA.** *Science* 2003; 300:1430–1434.

Bauerfeind P, Garner R, Dunn BE, Mobley HL. **Synthesis and activity of *Helicobacter pylori* urease and catalase at low pH.** *Gut* 1997; 40:25-30.

Blaser MJ, Berg DE, ***Helicobacter pylori* genetic diversity and risk of human disease.** *The Journal of Clinical Investigation* 2001; 107(7):767-73

Bourzac KM, Guillemin K. ***Helicobacter pylori*-host cell interactions mediated by type IV secretion.** *Cellular Microbiology* 2005; 7(7):911-919.

Buhrdorf R, Forster C, Haas R, Fischer W. **Topological analysis of a putative virB8 homologue essential for the cag type IV secretion system in *Helicobacter pylori*.** *Int J Med Microbiol* 2003; 293: 213–217.

Busler VJ, Torres VJ, McClain MS, Tirado O, Friedman DB, Cover TL. **Protein-protein interactions among *Helicobacter pylori* cag proteins.** *Journal of Bacteriology* 2006; 188(13):4787-4800.

Cascales E, Christie PJ., **The versatile bacterial type IV secretion systems.** *Nat. Rev. Microbiol.* 2003; 1:137–149.

Cendron L., Seydel A., Angelini A., Battistutta R., Zanotti G. **Crystal structure of CagZ, a protein from *Helicobacter pylori* pathogenicity island that encode from a tipe IV secretion system,** *J.Mol Biol* 2004 Jul 16; 340(4):881-9.

Cendron L., Tasca E., Seraglio T, Seydel A., Angelini A., Battistutta R., Montecuccio C., Zanotti G. **The crystal structure of CagS from the *Helicobacter pylori* pathogenicity island,** *InterScience* (in pubblicazione).

Christie PJ., **Bacterial type IV secretion: the *Agrobacterium* VirB/D4 and related conjugation systems**, Biochem. Biophys. Acta 2004; 1694:219–234.

De Bernard M, Papini E, De Filippis V, Gottardi E, Telford J, Manetti R, Fontana A, Rappuoli R, Montecucco C. **Low pH activates the vacuolating toxin of *Helicobacter pylori*, which becomes acid and pepsin resistant**. J. Biol. Chem. 1995; 270:23937-23940.

Drenth J. **Principles of Protein X-Ray Crystallography**. Springer 1994

Fischer W., Puls J., Buhrdorf R., Gebert B., Odenbreit S. and Haas R., **Systematic mutagenesis of the *Helicobacter pylori* cag pathogenicity island: essential genes for CagA translocation in host cells and induction of interleukin-8**, Molecular Microbiology (2001) 42(5), 1337-1348

Harris PR, Mobley HL, Perez-Perez GI, Blaser MJ, Smith PD. ***Helicobacter pylori* urease is a potent stimulus of mononuclear phagocyte activation and inflammatory cytokine production**. Gastroenterol. 1996; 111:419-425

Jakubowski SJ, Krishnamoorthy V, Cascales E, Christie PJ. ***Agrobacterium tumefaciens* VirB6 domains direct the ordered export of a DNA substrate through a type IV secretion system**. J. Mol. Biol. 2004; 341:961–977

Matthews B.W. **Solvent content of protein crystals**, J Mol Biol 1968 Apr 28;33:491-7

Montecucco C, Rappuoli R., **Living dangerously: how *Helicobacter pylori* survives in the human stomach**, Molecular Cell Biology 2001; 2:457-66

Montemurro P, Barbuti G, Dundon WG, Del Giudice G, Rappuoli R, Colucci M, De Rinaldis P, Montecucco C, Semeraro N, Papini E. ***Helicobacter pylori* neutrophil-activating protein stimulates tissue factor and plasminogen activator inhibitor-2 production by human blood mononuclear cells**. J. Infect. Dis. 2001; 183(7):1055-1062

Potterton E, Briggs P, Turkenburg M, Dodson E. **A graphical user interface to the CCP4 program suite.** *Acta Crystallogr. D Biol. Crystallogr.* 2003; 59(7):1131-7.

Robinson K, Argent R. H., Atherton J. C., **The inflammatory and immune response to *Helicobacter pylori* infection,** Best practice and Research clinical Gastroenterology, Vol.21, No.2, 237-259, 2007

Satin B, Del Giudice G, Della Bianca V, Dusi S, Laudanna C, Tonello F, Kelleher D, Rappuoli R, Montecucco C, Rossi F. **The neutrophil activating protein (HP-NAP) of *Helicobacter pylori* is a protective antigen and a major virulence factor.** *J. Exp. Med.* 2000; 191(9):1467-1476.

Saunders NJ, Peden JF, Hood DW, Moxon ER. **Simple sequence repeats in the *Helicobacter pylori* genome.** *Mol. Microbiol.* 1998; 27:1091–1098.

Segal ED, Falkow S, Tompkins LS. ***Helicobacter pylori* attachment to gastric cells induces cytoskeletal rearrangements and tyrosine phosphorylation of host cell proteins.** *Proc. Natl Acad. Sci. USA* 1996; 93:1259-1264

Stone M. A., **Current concepts: Transmission of *Helicobacter pylori*,** *Postgrad. Med J.*, 1999; 75; 198-200

Telford JL, Ghiara P, Dell'Orco M, Comanducci M, Burroni D, Bugnoli M, Tecce MF, Censini S, Covacci A, Xiang Z, Papini E, Montecucco C, Parente L, Rappuoli R. **Gene structure of *Helicobacter pylori* cytotoxin and evidence of its key role in gastric disease.** *J. Exp. Med.* 1994; 179(5):1653-1658.

Tomb JF, White O, Kerlavage AR, Clayton RA, Sutton GG, Fleischmann RD, Ketchum KA, Klenk HP, Gill S, Dougherty BA, Nelson K, Quackenbush J, Zhou L, Kirkness EF, Peterson S, Loftus B, Richardson D, Dodson R, Khalak HG, Glodek A, McKenney K, Fitzgerald LM, Lee N, Adams MD, Hickey EK, Berg DE, Gocayne JD, Utterback TR, Peterson JD, Kelley JM, Cotton MD, Weidman JM, Fujii C, Bowman C, Watthey L, Wallin E, Hayes WS, Borodovsky M, Karp PD, Smith HO, Fraser CM, Venter JC. **The complete genome sequence of the gastric pathogen *Helicobacter pylori*.** *Nature* 1997; 388(6642):539–547.

Yeo HJ, Savvides SN, Herr AB, Lanka E, Waksman G **Crystal structure of the hexameric traffic ATPase of the Helicobacter pylori type IV secretion system.** MolCell.2000Dec; 6(6):1461-72.

Yoshida N, Granger DN, Evans DJ Jr, Evans DG, Graham DY, Anderson DC, Wolf RE, Kviety PR. **Mechanisms involved in Helicobacter pylori-induced inflammation.** Gastroenterology 1993; 105(5):1431-1440

Winn M.D. **An overview of the CCP4 project in protein crystallography: an example of a collaborative project.** J Synchrotron radiat. 2003 Jan;10:23-5. Epub 2002 Dec 24

Zahl D, Wagner M, Bischof K, Bayer M, Zavec B, Beranek A, Ruckstuhl C, Zarfel GE, Koraimann G. **Peptidoglycan degradation by specialized lytic transglycosylases associated with type III and type IV secretion systems.** Microbiology. 2005; 151(11):3455-3467.

*Ringraziamenti*

*Desidero innanzitutto ringraziare il professor Giuseppe Zanotti per avermi dato l'opportunità di svolgere la tesi nel suo laboratorio; sia in questa occasione che in altre precedenti si è sempre dimostrato molto gentile e disponibile nei miei confronti.*

*Ringrazio la dottoressa e correlatrice Laura Cendron per avermi seguito nella fase sperimentale e per i numerosi suggerimenti indispensabili. Un ringraziamento è rivolto anche a tutti i ragazzi che hanno lavorato assieme a me nei laboratori di chimica (Alberto, Alessandro, Anton, Giulia, Luisa, Nicola) e del VIMM (Elena, Elisa, Fabio, Laura Fonso, Marco, Rosa) per il clima sereno e di collaborazione che si è instaurato in tutti questi mesi e per tutte le risate che ci siamo fatti a pranzo e non solo.*

*La lista delle persone che hanno collaborato, sebbene indirettamente, alla realizzazione di questo lavoro e la cui presenza è stata davvero importante sono parecchie, in particolare:*

- i miei famigliari che mi hanno sostenuto economicamente e non solo;*
- i miei amici (Cam Phat, Davide, Giulia, Michele, Migena, Mirko, Giacomo e tutti gli altri) che si sono sempre dimostrati persone eccezionali e che dai quali ho sempre avuto l'appoggio nei momenti critici;*
- i miei compagni di corso (presenti e passati) grazie ai quali ho potuto trascorrere gli anni dell'università nel migliore dei modi;*
- la mia ballerina Alessia che mi ha permesso di coltivare al meglio la mia passione principale, che è il ballo, consentendomi così di essere più rilassato durante la stesura;*
- i miei coinquilini Alberto, Giacomo e Nico, con i quali ho convissuto in un clima sereno e spesso festoso.*

*Tantissime quindi sono le persone che ho incontrato e che mi sento per qualche motivo di ringraziare (la lista dovrebbe continuare....ma sarebbe infinita), quelle che fanno eccezione sono rare (più uniche che rare) ma qui è meglio non citarle!*