



# UNIVERSITA' DEGLI STUDI DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

DIPARTIMENTO di CHIMICA BIOLOGICA

**SCUOLA di DOTTORATO di RICERCA  
in BIOCHIMICA e BIOTECNOLOGIE**

INDIRIZZO BIOTECNOLOGIE

CICLO XX

**ANALISI FUNZIONALE E COMPARATA  
DELLE LONGINE VAMP7 E Ykt6**

**Direttore della Scuola : Ch.mo Prof. Lorenzo Pinna**

**Supervisore : Prof. Francesco Filippini**

**Dottorando : Marco Vedovato**

DATA CONSEGNA TESI  
31 gennaio 2008



# INDICE

	<b>Pagina</b>
<b>SOMMARIO</b> .....	3
<b>ABSTRACT</b> .....	4
<b>INTRODUZIONE</b> .....	5
Origine evolutiva dei compartimenti cellulari.....	5
Ipotesi del trasporto vescicolare e primi studi pilota.....	6
Meccanismi molecolari del trasporto vescicolare.....	8
Struttura e funzione delle SNARE.....	9
I domini N-terminali delle SNARE.....	12
Le famiglie longiniche.....	13
– VAMP7/TI-VAMP.....	13
– Sec22b.....	14
– Ykt6.....	16
Ruoli dei domini longinici.....	16
<i>Longin-like</i> .....	17
<b>MATERIALI E METODI</b> .....	19
<b>ANALISI BIOINFORMATICHE</b> .....	19
– <i>Server</i> principali utilizzati.....	19
– Analisi delle sequenze.....	20
– Analisi filogenetiche.....	20
– <i>Homology modeling</i> .....	21
<b>BIOLOGIA MOLECOLARE</b> .....	22
Mutagenesi sito diretta.....	22
Sequenze degli oligonucleotidi sintetici per la PCR ( <i>primer</i> ).....	22
Reazione a catena della polimerasi (PCR).....	23
Elettroforesi in gel di agarosio.....	24
Plasmidi.....	24
Ceppo batterico e terreni di coltura.....	27
Ceppo di lievito e terreni di coltura.....	27
Clonaggio mediante kit TOPO TA Cloning®.....	28
Estrazione e purificazione del DNA plasmidico da cellule batteriche.....	28
Taglio del DNA con enzimi di restrizione.....	29
Purificazione dei frammenti di DNA dopo PCR o dopo restrizione.....	29
Ligazione.....	29
Induzione della competenza e trasformazione delle cellule batteriche.....	30
Sequenziamento del DNA.....	30
Conservazione dei cloni batterici a – 80°C.....	30
Trasformazione delle cellule di lievito.....	30

<b>SCOPO della TESI</b> .....	32
<b>RISULTATI E DISCUSSIONE</b> .....	34
Distribuzione delle longine negli eucarioti.....	34
Identificazione di una nuova famiglia proteica <i>longin-like</i> : le Fitolongine.	36
Analisi filogenetica delle famiglie longiniche.....	38
Creazione di <i>pattern</i> per le longine.....	43
Modelli delle longine.....	44
Conservazione e variazione del meccanismo di interazione intramolecolare. ....	47
<b>CONCLUSIONI E PROSPETTIVE</b> .....	49
<b>BIBLIOGRAFIA</b> .....	51

## SOMMARIO

Le longine sono proteine caratterizzate da un'estensione N-terminale conservata di 120-140 aminoacidi, definita longin domain (LD). Esse appartengono alle SNARE, proteine fondamentali per il controllo del traffico subcellulare, e sono divise in tre famiglie principali: VAMP7, Sec22 e Ykt6.

L'attività di ricerca svolta nel corso del dottorato è vincolata ad un progetto (finanziato dalla Fondazione CA.RI.PA.RO.) che prevede lo studio della modulazione funzionale e delle interazioni dei LD di VAMP7 e Ykt6. Perciò è stata studiata la distribuzione del LD e delle longine negli eucarioti attraverso uno screening bioinformatico su proteomi e genomi. Ciò ha permesso di evidenziare ruoli diversi per le famiglie longiniche di Ykt6 e VAMP7. Inoltre sono state scoperte le Fitolongine: esse rappresentano il primo esempio di famiglia di longine non-SNARE poiché presentano una regione centrale variabile che ha sostituito lo SNARE *motif*. L'evoluzione delle longine sia tra le diverse famiglie che all'interno di ogni famiglia è stata analizzata attraverso la creazione di alberi filogenetici ed ha rivelato interessanti informazioni soprattutto per quanto riguarda le relazioni tra Ykt6, VAMP7 e Fitolongine. Sono stati creati dei modelli 3D del LD per ottenere suggerimenti funzionali dalla conservazione o divergenza di specifiche regioni sulla superficie, ed individuare anche degli aminoacidi ritenuti "chiave" per eventuali interazioni con altre proteine coinvolte nel traffico vescicolare. La capacità autoinibitoria identificata per Ykt6 di lievito sembra essere basata sulla presenza di una superficie idrofobica nell' $\alpha 1$  del LD, ma dai nostri dati sulla presenza di superfici idrofobiche o polari conservate nell' $\alpha 1$  il meccanismo risulta essere famiglia-specifico e non sempre dovuto a superfici idrofobiche. La mutagenesi di alcuni aminoacidi conservati nelle Ykt6 e nelle VAMP7 e i successivi saggi di doppio ibrido, potranno aiutare a comprenderne l'importanza.

## ABSTRACT

Longins are proteins characterized by a conserved N-terminal extension of 120-140 amino acids, named longin domain (LD). They belong to SNARE that are crucial proteins for subcellular trafficking control, and they are divided into three main families: VAMP7, Sec22 and Ykt6. The research activity performed during this PhD is linked to a project (funded by CA.RI.PA.RO. Fondazione) that aims at studying functional modulation and interactions of VAMP7 and Ykt6 LDs. Therefore the distribution of LD and longin proteins in eukaryotes has been studied, by means of a bioinformatic screening on proteomes and genomes. This allowed for highlighting differential roles of Ykt6 and VAMP7 longin families. In addition, Phytolongins have been discovered: they are representing the first example of non-SNARE longins family, showing a variable central region that replaced the SNARE motif. By means of phylogenetic trees, longins evolution between and inside the families has been analyzed and has shown interesting data, particularly as regards the relationships between Ykt6, VAMP7 and Phytolongins. We performed tridimensional models of LDs in order to infer functional insights from the conservation or divergence of surface specific patches, and also to unravel key amino acids for the interactions with other vesicular proteins. A hydrophobic surface of  $\alpha$  helix 1 seems to mediate the autoinhibitory mechanism in yeast Ykt6 LD, but our data on conserved hydrophobic or polar surfaces in  $\alpha$ 1 indicate that this mechanism is family-specific and not always dependent on hydrophobicity. We have performed site-directed mutagenesis of particular residues that are conserved in Ykt6 and VAMP7, in order to obtain indications about their value in future two-hybrid assays.

## INTRODUZIONE

### **Origine evolutiva dei compartimenti cellulari.**

La più semplice tra le cellule eucariotiche è estremamente più complessa del più avanzato procariote ed è organizzata in compartimenti cellulari ed organelli quali, ad esempio, il nucleo, i mitocondri, il reticolo endoplasmatico (RE) ed il citoscheletro. Il macchinario base di replicazione, trascrizione e traduzione del DNA e i cicli metabolici chiave sono condivisi da tutte le cellule e verosimilmente devono essere stati posseduti da un proto-eucariote (Ranea *et al.*, 2006), mentre l'origine dei nuclei resta tuttora una questione aperta (Lopez-Garcia & Moreira, 2006; Martin, 2005). Strutture analoghe ai nuclei non sono facilmente identificabili in procarioti, anche se ci sono ad esempio, nel phylum dei Planctomycetes, dei nucleoidi limitati da membrane (Fuerst, 2005).

L'origine procariotica dei mitocondri da parte di batteri simbiotici è risaputa e, recentemente, sono stati identificati come mitocondri modificati in eucarioti anaerobi, gli idrogenosomi (che assicurano e generano energia alla cellula) e i mitosomi (trovati in *Entamoeba* e *Giardia*; van der Giezen *et al.*, 2005).

Anche i compartimenti del traffico di membrana presenti nelle cellule eucariotiche non sono chiaramente riconoscibili nei procarioti, ma nell'archeobatterio *Ignicoccus islandicus* è stata osservata tramite microscopia elettronica la presenza di tubuli ed eventi di gemmazione e fusione delle vescicole (Rachel *et al.*, 2002).

Sebbene non ci siano molti omologhi molecolari del traffico di membrana nei procarioti, Vps29, una porzione del rivestimento vescicolare del retromero (unità di recupero dall'endosoma al Golgi), ha un chiaro omologo in molti batteri e archeobatteri e possiede attività fosfoesterasica condivisa da tutti gli omologhi (Damen *et al.*, 2006).

La presenza di queste strutture in alcuni procarioti e la capacità di costruire compartimenti racchiusi da membrane, rende più credibile la loro formazione spontanea, senza dover invocare indefiniti eventi stocastici. Il sistema del traffico di membrana sembra essersi originato anch'esso autonomamente da un aumento di preesistenti blocchi di domini e *motif* di un limitato repertorio di proteine,

presenti nel proto-eucariote o derivate recentemente da trasferimento di geni (Dacks & Field, 2007).

Nella cellula eucariotica primordiale il sistema secretorio ed endocitico doveva essere in grado di gestire due importanti funzioni base: esportazione ed importazione. Nelle cellule eucariotiche esistenti, questi processi si sono complicati e sovrapposti per permettere il trasporto e l'indirizzamento delle proteine e dei lipidi nei corretti compartimenti. L'aumento della grandezza della cellula e del suo complesso di endomembrane è stato accompagnato da un aumento della complessità del macchinario molecolare di queste vie di trasporto, ovvero del numero e dei tipi di proteine di rivestimento, GTPasi regolatorie, fattori di fusione e reclutamento, motori molecolari e segnali di indirizzamento (Sarastre & Goud, 2007).

Una differenza tra gli attuali eucarioti unicellulari e le cellule dei metazoi è la continua richiesta di riciclo della membrana. Ad esempio, il lievito può essere considerato come una macchina biosintetica: è infatti in grado di dividersi molto velocemente grazie ad un apparato secretorio efficiente, ma ha una limitata capacità endocitica, che preclude lo sviluppo di organelli di riciclo sia tra RE ed apparato del Golgi, sia tra membrana plasmatica e vacuolo. Le cellule dei metazoi nel loro percorso biosintetico hanno evoluto invece un compartimento specializzato di membrana, chiamato compartimento intermedio, che permette l'indirizzamento delle molecole tra RE ed apparato del Golgi e favorisce il riciclo della membrana, che è importante per il funzionamento e l'integrità del reticolo (Sarastre & Goud, 2007).

### **Ipotesi del trasporto vescicolare e primi studi pilota.**

La frequente osservazione di proteine secretorie in vescicole ha ispirato l'ipotesi che il trasferimento di molecole tra gli organelli del percorso secretorio sia mediato dal trasporto in vescicole.

Le proteine secretorie appena sintetizzate vengono trasportate attraverso una serie di organelli racchiusi da membrane verso lo spazio extracellulare. Allo stesso modo le proteine destinate alla membrana plasmatica, agli endosomi o ai lisosomi, condividono le prime "stazioni" di questo percorso (cioè RE e Golgi) con le



proteine secretorie, ma quest'ultime si trovano anche dentro piccole vescicole intersperse tra i maggiori organelli del *trafficking*.

Il trasporto può essere suddiviso in quattro fasi principali: la gemmazione della vescicola, la selezione delle proteine da trasportare, l'indirizzamento della vescicola e la fusione con la membrana bersaglio. La vescicola gemma da una membrana definita donatrice, che permette anche l'incorporazione selettiva del carico nella vescicola e trattiene le proteine residenti nel compartimento donatore. Queste vescicole vengono indirizzate verso specifici compartimenti bersaglio dove possono riversare il loro carico dopo la fusione con la membrana (Palade, 1975).

La caratterizzazione di reazioni e componenti proteici coinvolti nel trasporto vescicolare è stata iniziata dai gruppi di Schekman e Rothman.

Il gruppo di Schekman ha avuto la lungimiranza di scegliere come organismo modello *Saccharomyces cerevisiae* quando non era ancora chiaro che avesse un apparato secretorio molto simile a quello dei mammiferi. La facilità di manipolare geneticamente il lievito ha permesso di isolare dei mutanti sensibili alla temperatura (*sec*) e difettivi nella secrezione proteica (Novick *et al.*, 1980).

Questi mutanti contenevano vari tipi di strutture racchiuse da membrana che differivano per il tipo di mutante: potevano essere piccole vescicole da 60-100 nm di diametro che presumibilmente corrispondevano ai trasportatori, una maggiore rete di RE o un complesso Golgi alterato. Tutte queste strutture rappresentano stadi intermedi del percorso secretorio, che mostrano amplificazione a causa di un blocco in una fase precisa del trasporto proteico.

La successiva identificazione dei geni che erano difettivi nei 23 tipi di mutanti identificati, ha rilevato nuove proteine coinvolte nei vari stadi del percorso secretorio (Schekman & Novick, 2004).

Il gruppo di Rothman, invece, ha affrontato il problema misurando, con un ingegnoso saggio *cell free*, il trasporto di proteine tra le cisterne del complesso Golgi di mammifero (Balch *et al.*, 1984).

Questo saggio consisteva nell'incubare una frazione di Golgi da cellule infettate dal virus della stomatite vescicolare (che mancano dell'enzima N-acetilglucosammina transferasi) assieme al Golgi derivante da cellule *wild type*. Aggiungendo il substrato marcato radioattivamente, ATP e citosol, si può

misurare il trasporto tra i compartimenti donatore e accettore attraverso l'immunoprecipitazione delle glicoproteine del Golgi infettato che incorporano il substrato radioattivo.

Questi approcci hanno permesso di scoprire che le proteine Sec identificate geneticamente sono ortologhe a quelle identificate biochimicamente con il saggio *cell free*, rivelando che le componenti chiave del macchinario di trasporto vescicolare sono conservate in eucarioti evolutivamente lontani.

### **Meccanismi molecolari del trasporto vescicolare.**

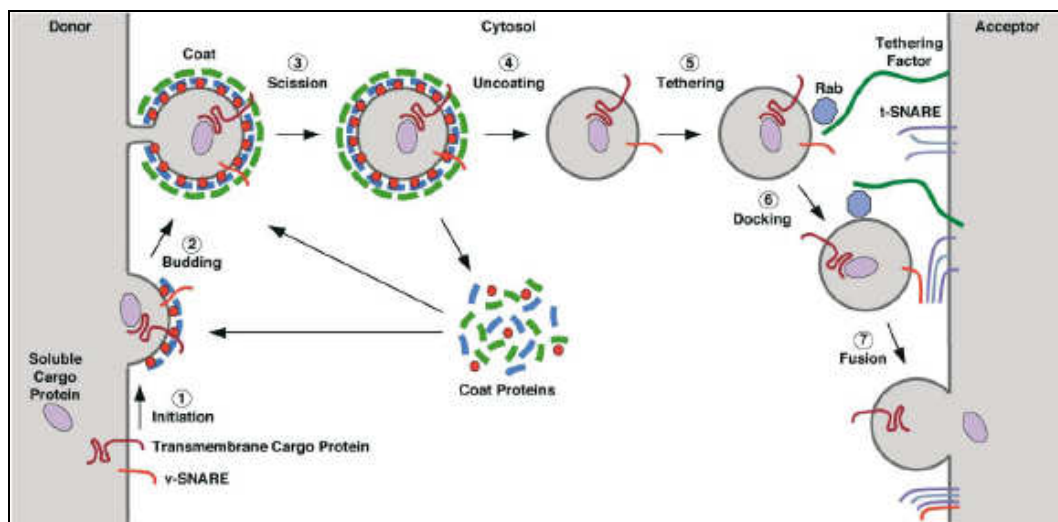
Il primo stadio del trasporto vescicolare consiste nella formazione della vescicola attraverso l'iniziazione del rivestimento (Figura 1), cioè l'assemblaggio di proteine che vengono reclutate dal citosol verso la vescicola nascente. Tali rivestimenti possono essere impalcature di proteine come la clatrina (Pearse, 1975) che è implicata nel trasporto dal trans-Golgi *network* all'endosoma e alla membrana plasmatica; oppure altri due rivestimenti chiamati COPII e COPI che sono formati da complessi proteici: COPII media il trasporto tra RE, compartimenti intermedi RE-Golgi e complesso Golgi (Barlowe *et al.*, 1994); COPI è invece coinvolto nel trasporto tra le cisterne del Golgi e nel trasporto retrogrado verso il RE (Letournerur *et al.*, 1994).

In generale i componenti del rivestimento più vicini alla membrana (in blu in Figura 1) vengono reclutati da GTPasi Arf/Sar associate alla membrana (Serafini *et al.*, 1991) (in rosso in Figura 1); le proteine transmembrana che riconoscono il carico ed altre proteine (dette v-SNARE) vengono inglobate nel rivestimento e comincia la gemmazione. Al complesso si aggiungono altre proteine più distali (ad esempio il pre-complesso COPII formato da Sec23p•Sec24p recluta il subcomplesso Sec13p•Sec31p formato due subunità di Sec13p e di Sec31p; Lederkremer *et al.*, 2001) (in verde in Figura 1) che aumentano la curvatura e la successiva scissione della vescicola. Nel passaggio successivo avviene la dissociazione del rivestimento ed il riciclo delle proteine formanti il complesso. La vescicola deve poi essere indirizzata verso il compartimento accettore: essa si avvicina alla membrana accettrice guidata dal citoscheletro, e viene "incatenata" attraverso un fattore legante (*tether*) e una GTPasi Rab. Viene poi ancorata alla membrana *target* attraverso l'associazione tra proteine v-SNARE (presenti sulla

vescicola) e t-SNARE (sulla membrana *target*) in un fascio di 4  $\alpha$  eliche. Questo complesso *trans*-SNARE promuove la fusione della vescicola nel doppio strato lipidico accettore. La proteina trasportata viene trasferita al compartimento accettore e le SNARE vengono riciclate.

Nonostante la diversità nelle dimensioni degli organelli, la gemmazione e la fusione delle vescicole sono mediate da complessi di famiglie di proteine conservate negli eucarioti (Bonifacino & Glick, 2004).

Il processo necessita di una regolazione nella gemmazione e nella fusione delle vescicole, inoltre ci deve essere anche un trasporto retrogrado che permetta il bilanciamento e il regolamento del trasporto.



**Figura 1:** rappresentazione schematica dei principali eventi nella gemmazione e fusione di vescicole da un compartimento donatore ad un compartimento accettore (tratta da Bonifacino & Glick, 2004).

### Struttura e funzione delle SNARE.

Le proteine SNARE sono recettori di SNAP (SNAP Recceptors) (Clary & Rothman, 1990), ove l'acronimo SNAP sta per Soluble NSF Attachment Protein (Clary *et al.*, 1990) e NSF significa fattore sensibile alla N-etilmaleimide ed indica una ATPasi (Tagaya *et al.*, 1993; Wilson *et al.*, 1989). Queste proteine sono caratterizzate da un dominio ad  $\alpha$  elica conservato (SNARE *motif*) di 60-70 aminoacidi disposti in eptadi ripetute (Weimbs *et al.*, 1997); al C-terminale di molte proteine SNARE un dominio transmembrana si collega con un piccolo

*linker* allo *SNARE motif*. Nella regione N-terminale risiede la differenza principale tra le proteine SNARE, che ne permette la divisione in vari subgruppi per la presenza o assenza di domini che possono variare in lunghezza e mostrare dei *fold* differenti ed indipendenti fra di loro (Fasshauer, 2003; Brunger, 2005).



**Figura 2:** schematizzazione del prototipo di una proteina SNARE. E' omesso il *linker* tra lo *SNARE motif* ed il *DTM* (*DTM* = dominio transmembrana).

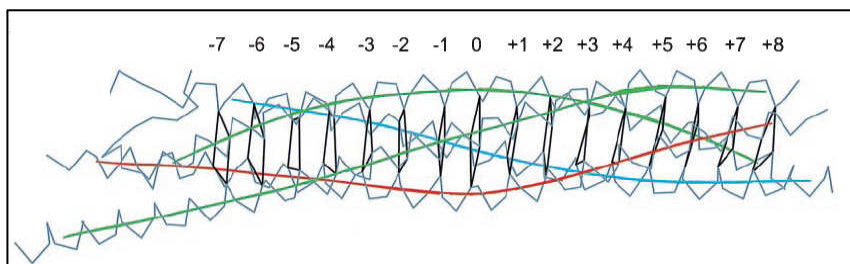
Questa struttura prototipo (N-terminale - *SNARE motif* - dominio transmembrana C-terminale) (Figura 2) presenta delle importanti eccezioni:

- alcune proteine SNARE non possiedono un dominio N-ter strutturato, tra cui le “brevine” (Rossi *et al.*, 2004);
- altre, prive di un dominio transmembrana, possono presentare una modificazione post-trascrizionale per l’ancoraggio in membrana: tra queste è incluso un piccolo gruppo rappresentato dalle SNAP-25 (synaptosomal-associated protein 25 kDa) che contengono due differenti *SNARE motif* uniti da un *linker* flessibile palmitoilato; nel lievito il dominio transmembrana di Ykt6p è sostituito da un box CAAX che viene farnesilato (McNew *et al.*, 1997);
- il dominio transmembrana delle SNARE può anche essere palmitoilato ed è stato dimostrato recentemente che tale modificazione protegge le SNARE dalla ubiquitinazione e conseguente degradazione (Valdez-Taubas & Pelham, 2005).

Le proteine SNARE presenti nelle due opposte membrane sono associate in complessi che vengono disassemblati dall’NSF. Per molti anni si è ritenuto che ci fosse una netta separazione tra le SNARE del compartimento donatore ed accettore, portando ad una classificazione funzionale come v-SNARE e t-SNARE (Söllner *et al.*, 1993). Attraverso il riconoscimento reciproco che porta alla formazione del cosiddetto “complesso SNARE” o “*trans*-SNARE” (Hay, 2001), la cui struttura è stata studiata principalmente a livello di esocitosi di vescicole sinaptiche (Sutton *et al.*, 1998), si ottiene la specificità dell’indirizzamento di una vescicola al compartimento corretto.

Questo sistema di classificazione, tuttavia, risulta inadeguato a descrivere gli eventi di fusione omotipica e alcune funzioni delle SNARE in alcuni passaggi di trasporto con vari partner. Ad esempio, la SNARE Sec22b di *Saccaromices cerevisiae* funziona sia nel traffico retrogrado che anterogrado tra il RE e l'apparato del Golgi. Nel traffico anterogrado Sec22b colocalizza con Bos1 e Bet1 nel trasporto vescicolare, ma solo Bet1 era classificata come v-SNARE, mentre Bos1 e Sec22b come t-SNARE (McNew *et al.*, 2000). Nel trasporto retrogrado, però, Sec22b è l'unica v-SNARE funzionale (Burri *et al.*, 2003).

La conoscenza dei complessi SNARE ha fornito una classificazione più rigorosa ed invariabile. La formazione del complesso viene mediata dallo SNARE *motif* ed è associata a cambiamenti conformazionali e di energia libera. Quando le SNARE sono monomeriche lo SNARE *motif* non è strutturato. Tuttavia quando un appropriato set di SNARE è combinato, si forma un complesso molto stabile (Fasshauer, 2003). Le strutture cristallografiche di due complessi (il primo formato da SNARE neuronali ed il secondo da SNARE endosomali) hanno rilevato un elevato grado di conservazione: i *core complexes* possono essere rappresentati come un *coiled coil* allungato di 4  $\alpha$ -eliche parallele che portano ciascuna lo SNARE *motif* (Sutton *et al.*, 1998, Antonin *et al.*, 2002). Il centro del fascio contiene 16 strati impilati che permettono l'interazione tra le catene: questi strati sono tutti idrofobici, tranne lo strato centrale "0" (chiamato anche strato ionico) che contiene 3 glutammine (Q) molto conservate ed un residuo di arginina (R) conservato (Figura 3). In base al tipo di aminoacido presente nello strato ionico si può avere una nuova classificazione delle SNARE in R e Q-SNARE, queste ultime ulteriormente suddivisibili in Qa-, Qb-, e Qc-SNARE (Fasshauer *et al.*, 1998).



**Figura 3:** visualizzazione degli strati del fascio del complesso di fusione tra le 4  $\alpha$ -eliche parallele delle SNARE (tratta da Fasshauer *et al.*, 1998).

I complessi SNARE funzionali che guidano la fusione di membrana sono etero- ed oligomerici: ogni fascio di 4 eliche parallele è invariante e richiede 3Q-SNARE (Qa, Qb e Qc) e 1R-SNARE. Un'analisi filogenetica delle sequenze SNARE di lievito, *Arabidopsis thaliana* e mammiferi mostra che queste 4 sottofamiglie SNARE sono altamente conservate e divergono presto nell'evoluzione eucariotica (Bock *et al.*, 2001).

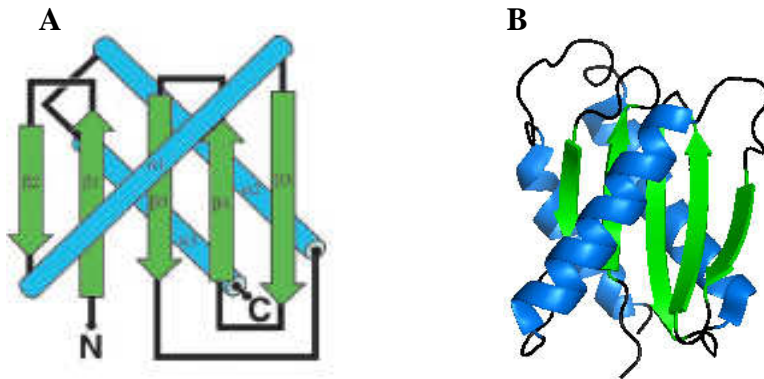
Grazie alla natura anfifilica, lo SNARE *motif* si può associare in altre combinazioni che risultano essere fasci di eliche meno stabili del complesso. Questi complessi, identificati nello studio di SNARE neuronali (Misura *et al.*, 2001a; Misura *et al.*, 2001b; Margittai *et al.*, 2001; Zhang *et al.*, 2002; Weninger *et al.*, 2003; Dennison *et al.*, 2006), non sembrano avere una corretta topologia di membrana e guidare la fusione di membrana, ma piuttosto rappresenterebbero reazioni "off-pathway" (Fasshauer, 2003).

### **I domini N-terminali delle SNARE.**

La presenza di domini N-terminali con *fold* differenti ha reso interessante capirne la funzione nelle SNARE.

Le SNARE Qa, ed alcune Qb e Qc, hanno 3  $\alpha$  eliche antiparallele N-terminali che formano un fascio: questi fasci possono variare in lunghezza ed essere connessi allo SNARE *motif* da un linker flessibile. Questo dominio viene chiamato Habc e interagisce con lo SNARE *motif* formando una configurazione chiusa che viene poi stabilizzata dalla proteina munc18 (Fernandez *et al.*, 1998; Dulubova *et al.*, 1999; Margittai *et al.*, 2003; Toonen & Verhage, 2003). La Qc-SNARE di lievito Vam7p ha un dominio conservato di circa 120 aminoacidi, chiamato *Phox-homology* (PX) perché originariamente identificato nelle subunità p40<sub>phox</sub> e p47<sub>phox</sub> della NADPH ossidasi (Lu *et al.*, 2002). La presenza di ulteriori ripiegamenti nei domini N-terminali di Qb e Qc non è da escludere perché alcuni di questi domini sono divergenti in sequenza e lunghezza (Jahn & Scheller, 2006). Molte R-SNARE hanno invece un dominio N-terminale globulare formato da cinque foglietti  $\beta$  antiparalleli racchiusi da un lato da una  $\alpha$  elica e dall'altro da due  $\alpha$  eliche (Gonzalez *et al.*, 2001; Tochio *et al.*, 2001) (Figura 4). Questo dominio, formato da 120-140 residui aminoacidici, è stato definito dominio

longinico (LD) e le proteine che lo possiedono sono chiamate “longine” per distinguerle dalle R-SNARE che possiedono un dominio N-terminale molto corto e non conservato, e per questo chiamate brevine (Filippini *et al.*, 2001).



**Figura 4:** (A) diagramma topologico (modificato da Gonzalez *et al.*, 2001) e (B) rappresentazione 3D del LD di Sec22b di topo tramite PyMOL.

### Le famiglie longiniche.

#### – VAMP7/TI-VAMP.

Il primo membro studiato di questa famiglia proteica è stata la VAMP7 umana, una proteina di 220 aminoacidi codificata dal gene SYBL1 (D’Esposito *et al.*, 1996), che è stata identificata anche in ratto. Attraverso studi condotti su ratto con anticorpi policlonali anti-VAMP7, è stato dimostrato che è espressa come proteina di circa 25 KDa in vari organi, tra cui polmone, milza, fegato, rene e cervello, mentre è assente in cuore e muscolo (Advani *et al.*, 1999). Questi risultati concordavano con dati di *northern blot* ottenuti precedentemente sugli stessi tessuti (Advani *et al.*, 1998).

Per la sua resistenza alla proteolisi da parte della neurotossina tetanica, VAMP7 è stata di conseguenza nominata anche TI-VAMP (*Tetanus neurotoxin Insensitive-VAMP*); essa risulta insensibile anche alle neurotossine botuliniche B, D, F e G (Galli *et al.*, 1998). VAMP7 è coinvolta nella via endocitica ed è presente negli endosomi precoci e tardivi (Advani *et al.*, 1999); media la fusione delle vescicole dal post-Golgi con la membrana plasmatica in neuroni e cellule epiteliali (Galli *et al.*, 1998; Advani *et al.*, 1998) ed è assente invece nelle vescicole di trasporto dal *trans-Golgi network* agli endosomi (Advani *et al.*, 1999). L’intervento di

VAMP7/TI-VAMP nel trasporto dagli endosomi ai lisosomi è stato dimostrato anche nei macrofagi alveolari (Ward *et al.*, 2000). Dati di microscopia elettronica rivelano che VAMP7/TI-VAMP non è associata ai lisosomi, ma media la fusione, insensibile alle neurotossine clostridiali, di vescicole provenienti dal post-Golgi con la membrana plasmatica nei neuroni (Coco *et al.*, 1999) ed anche nelle cellule epiteliali, definendo in questo modo un nuovo compartimento (Lafont *et al.*, 1999).

La crescita di assoni e dendriti è un evento cruciale nella maturazione e nel differenziamento dei neuroni (Prochiantz *et al.*, 1995) e si basa soprattutto sull'aggiunta di nuovi componenti alla membrana plasmatica, attraverso il trasporto di vescicole lungo la via esocitotica (Futerman & Banker, 1996). L'allungamento dei neuriti è un processo finemente regolato: VAMP7/TI-VAMP (ed in particolare il suo dominio N-terminale) è un componente chiave di questa regolazione (Martinez-Arca *et al.*, 2000). La regolazione avviene tramite l'inibizione, mediata dall'LD, dell'interazione dello SNARE *motif* con il complesso di fusione: misurando il tasso di assemblaggio dell'LD di VAMP7 con SNAP-25 e la syntaxina, si è visto che le interazioni sono circa il doppio in assenza del dominio longinico, portando a formulare l'ipotesi che esso regoli negativamente l'esocitosi (Martinez-Arca *et al.*, 2003). Recentemente è stato dimostrato, attraverso analisi di immunoblot, che VAMP7 è specificamente associata con il RE in cellule dell'intestino di ratto e regola il *trafficking* intracellulare di alcune vescicole specializzate chiamate PCTV (*PreChylomicron Transport Vesicle*) (Siddiqi *et al.*, 2006), che gemmano dal RE e trasportano verso il Golgi l'unica lipoproteina (chilomicrone) prodotta dall'intestino (Siddiqi *et al.*, 2003).

Nei saccaromiceti, la VAMP7 è sostituita da Nyv1, la cui regione N-terminale è stata recentemente risolta per NMR, dimostrandone il *fold* LD (Wen *et al.*, 2006).

#### – **Sec22b.**

L'importanza delle longine Sec22b nel *trafficking* era già stata capita dai primi studi condotti nel lievito per identificare quali geni (e loro prodotti proteici) fossero importanti per il trasporto tra il RE e il complesso del Golgi. SEC22, BET1 e BOS1 furono identificati come geni cruciali per questo trasporto



(Newman *et al.*, 1990). Nel lievito *Saccharomyces cerevisiae* Sec22b (chiamata anche Sec22p, Dascher *et al.*, 1991) interviene nella via di secrezione, in particolare nel trasporto vescicolare, sia anterogrado che retrogrado, tra il reticolo endoplasmatico ed il complesso del Golgi e questa funzione è stata identificata anche in mammiferi (Hay *et al.*, 1996; Paek *et al.*, 1997; Hay *et al.*, 1998; Spang & Schekman, 1998). In lievito, si capì che Sec22b è richiesta assieme a Bos1p, affinché si verifichi un'efficiente interazione tra le SNARE per la formazione del complesso di fusione (Sacher *et al.*, 1997); in ratto è stato dimostrato che Sec22b partecipa al complesso quaternario con altre tre SNARE: rbet1, membrin e syntaxina 5 (Xu *et al.*, 2000). Sec22 è ancorata alla membrana tramite un dominio transmembrana, la cui composizione e lunghezza può determinare la capacità della proteina di partecipare a tappe diverse del trasporto nella via secretoria (Bretscher & Munro, 1993). Può essere sostituita da Ykt6 nel traffico dal reticolo al Golgi in lievito, mentre Sec22 non può sostituire funzionalmente Ykt6 (Burri *et al.*, 2003; Tsui *et al.*, 2001). E' stata risolta la struttura cristallografica del dominio N-terminale (LD) della Sec22b di topo: esso consiste in cinque foglietti  $\beta$  antiparalleli racchiusi da un lato da una  $\alpha$  elica e dall'altro da due  $\alpha$  eliche e presenta, quindi, la stessa struttura del dominio N-terminale di Ykt6p (Gonzalez *et al.*, 2001; Tochio *et al.*, 2001). Poiché la capacità di assemblaggio del complesso SNARE con e senza LD di Sec22b di topo risultava invariato, ciò ha portato a concludere che l'LD non interagisse con lo SNARE *motif* o che l'interazione fosse troppo debole o transiente (Gonzales *et al.*, 2001). Recenti studi in mammifero sulla struttura del complesso che si forma tra Sec23, Sec24 e Sec22b per determinare quali meccanismi molecolari permettessero l'esportazione di Sec22b dal RE, hanno rivelato che il dominio longinico di Sec22b deve interagire con il suo SNARE *motif* e formare una conformazione "chiusa" che possa così interagire con una superficie formata dal subcomplesso Sec23/24. Questo risultato dimostra che il "segnale di trasporto" di Sec22 è un epitopo conformazionale (Mancias & Goldberg, 2007).

## – Ykt6.

Questa famiglia proteica ha la caratteristica principale di non avere un dominio transmembrana, ma di essere legata alla membrana tramite un'ancora lipidica. Quando in lievito è stata identificata la prima Ykt6 (chiamata anche Ykt6p) come componente nel *trafficking* tra il RE ed il Golgi, si è capito che il box C-terminale CAAX (un *motif* consenso per la prenilazione) poteva venire unito con un lipide (Søgaard *et al.*, 1994). Questa proteina, oltre ad essere essenziale in lievito nel trasporto tra RE e Golgi (McNew *et al.*, 1997), è stata successivamente identificata nei complessi SNARE del Golgi in mammiferi (Zhang & Hong, 2001; Xu *et al.*, 2002), come importante componente della fusione omotipica nel vacuolo di lievito (Ungermann *et al.*, 1999; Kweon *et al.*, 2003) ed in due tappe distinte del trasporto verso il vacuolo (Dilcher *et al.*, 2001). Inoltre è stata determinata la sua funzione nel trans-Golgi *network* in *Arabidopsis* (Chen *et al.*, 2005). La localizzazione in membrana di Ykt6p è essenziale per la sua funzione ed è mediata da isoprenilazione a carico dell'estrema regione carbossi-terminale; ciò nonostante la sostituzione della coda lipidica con un'ancora transmembrana non sembra provocare alcuna alterazione funzionale (McNew *et al.*, 1997). L'associazione di Ykt6 a membrane di cellule non neuronali (come le HeLa) avviene attraverso farnesilazione e palmitoilazione (Fukasawa *et al.*, 2004). Ykt6 è in grado di mediare la palmitoilazione sia di altre proteine (Dietrich *et al.*, 2004), sia di se stessa: questa capacità autocatalitica sembra associata al dominio N-terminale (Veit, 2004). Tramite NMR è stato determinato che tale dominio in Ykt6 di lievito possiede la stessa struttura del dominio N-terminale di Sec22b di topo (Tochio *et al.*, 2001), cioè è un dominio longinico (Filippini *et al.*, 2001). Il LD mostra in superficie un motivo idrofobico di quattro fenilalanine, che interagisce con lo SNARE *motif*, regolandolo negativamente (Tochio *et al.*, 2001).

### **Ruoli dei domini longinici.**

Il primo ruolo attribuito al dominio longinico (LD) è stato proprio la capacità autoinibitoria: attraverso l'interazione con lo SNARE *motif*, il LD di Ykt6p di lievito influenza la cinetica della formazione dei complessi SNARE

(Tochio *et al.*, 2001). Questa capacità di interazione intramolecolare ricorda quella del dominio Habc delle syntaxine. Tuttavia, l'LD ha anche la funzione di mediare la localizzazione subcellulare: ad esempio, l'LD delle Ykt6 dei mammiferi ha la capacità di indirizzarle verso compartimenti specifici (Hasegawa *et al.*, 2003).

Il LD di Ykt6 non solo controlla la localizzazione subcellulare, ma agisce anche come chaperone per lo SNARE *motif* e l'ancora lipidica (Hasegawa *et al.*, 2004). Altra funzione identificata per l'LD di Ykt6 di lievito è il suo intervento nella palmitoilazione di Vac8 (Dietrich *et al.*, 2004); inoltre contiene un sito di legame per il palmitoil-CoA e possiede attività autopalmitoilica (Veit, 2004).

Per l'LD della VAMP7 umana è stato dimostrato un ruolo duale: ha la capacità di inibire la formazione dei complessi SNARE e di mediare la localizzazione della VAMP7 nell'endosoma tardivo attraverso l'interazione con la subunità  $\delta$  del complesso AP3 nell'endosoma precoce (Martinez-Arca *et al.*, 2003). L'espressione dell'LD di VAMP7 regola la secrezione lisosomale delle cellule epiteliali ed agisce nel sistema endocitico tardivo nella migrazione cellulare (Proux-Gillardeaux *et al.*, 2007). In *Arabidopsis thaliana*, che possiede più di 10 geni per le VAMP7, i diversi domini longinici delle VAMP7 sono essenziali per permetterne la corretta e diversa localizzazione subcellulare (Uemura *et al.*, 2005). Nell'LD di Sec22b sono presenti alcuni aminoacidi che sembrano particolarmente interessanti nell'interazione con il subcomplesso Sec23/24 del COPII, dimostrando come ci siano ulteriori funzioni del dominio longinico nel *trafficking* vescicolare (Mancias & Goldberg, 2007).

### ***Longin-like.***

Esistono delle proteine che non appartengono alle SNARE, sono coinvolte nel traffico vescicolare e possiedono un dominio avente un ripiegamento *longin-like* che potrebbe svolgere ruoli funzionali nel *trafficking*:

- SEDL (il cui gene è responsabile della *Spondyloepiphyseal dysplasia tarda* - SEDT -), un componente del TRAPP (*transport protein particle*) coinvolto nel trasporto delle vescicole tra il RE ed il Golgi, ha una struttura molto simile al dominio longinico e potrebbe interagire con le proteine SNARE oppure è stato anche ipotizzato che agisca come un adattatore per la costituzione del TRAPP (Jang *et al.*, 2002);

- la subunità  $\sigma 2$  e il dominio N-terminale della subunità  $\mu 2$  del complesso AP2 (formato da quattro subunità) possiedono una struttura simile al LD; sembra che il loro ruolo principale sia nello stabilizzare il centro dell'assemblamento dell'adattatore (Collins *et al.*, 2002);
- il dominio N-terminale della subunità  $\alpha$  di SR (*signal recognition particle receptor*) di mammifero ha un *fold* simile al LD ed è stato identificato come un prototipo per l'interazione con piccole GTPasi (Schlenker *et al.*, 2006);
- la proteina MP1 (*mitogen-activated protein kinase scaffold protein*) viene localizzata nell'endosoma tardivo attraverso l'interazione con una proteina adattatrice chiamata p14. E' stata risolta la struttura del complesso formato da queste due proteine e si pensa possa essere un sito di interazione con le chinasi: entrambe hanno una struttura *longin-like* (Lunin *et al.*, 2004; Kurzbauer *et al.*, 2004);
- una struttura simile al dominio longinico è stata determinata anche nelle famiglie proteiche di ChiPS e DUF254, anch'esse coinvolte nel traffico vescicolare. (Kinch & Grishin, 2006).

Tutte queste proteine che posseggono una struttura *longin-like* e che sono coinvolte nella formazione di complessi del traffico vescicolare, permettono di assegnare al dominio longinico un ruolo nella conservazione del meccanismo tra gli eucarioti.

## MATERIALI E METODI

### ANALISI BIOINFORMATICHE

– *Server principali utilizzati.*

**N.C.B.I.** (<http://www.ncbi.nlm.nih.gov/>). Il *National Center for Biothechnology Information* è una divisione della *National Library of Medicine* (NLM), presso il *National Institute of Health* (NIH), USA. Fondato nel 1988 per decisione governativa, ha lo scopo di fornire un supporto alla ricerca nella biologia molecolare. Gestisce *database* pubblici e sviluppa *software* per l'analisi di dati genomici.

**E.B.I.** (<http://www.ebi.ac.uk/>). L'*European Bioinformatic Institute* è un'organizzazione che fa capo all'*European Molecular Biology Laboratory* (EMBL). Assieme all'NCBI, gestisce i maggiori *database* di sequenze nucleotidiche e proteiche, e rappresenta il tentativo di fornire molteplici programmi per l'analisi delle sequenze.

**EXPASY** (<http://www.expasy.ch/>). L'*Expert Protein Analysis System*, gestito dal *Swiss Institute of Bioinformatics* (SIB), fornisce molteplici programmi per l'analisi di sequenze proteiche.

**J.G.I.** (<http://www.jgi.doe.gov/>). Il DOE (US Department of Energy) *Joint Genome Institute* è stato creato nel 1997 unendo le risorse e le competenze di tre laboratori nazionali: il Lawrence Berkeley National Laboratory (LBNL), il Lawrence Livermore National Laboratory (LLNL), e il Los Alamos National Laboratory (LANL) gestiti dall'Università della California. La caratteristica principale è di avere *database* di sequenze nucleotidiche e proteiche di organismi ritenuti importanti dal Dipartimento dell'Energia e non necessariamente corrispondenti ai "classici" organismi modello.

– **Analisi delle sequenze.**

**Analisi di similarità.** Le analisi di similarità sono state condotte utilizzando BLAST (*Basic Local Alignment Search Tool*; Altschul *et al.*, 1997), un programma per la ricerca di similarità locale tra una sequenza data ed una serie di sequenze depositate nella banca dati scelta. Le ricerche sono state effettuate nei database di sequenze dell'NCBI, EBI e JGI utilizzando i parametri di *default*. Nei casi in cui non si riusciva ad identificare facilmente gli ortologi, avendo a disposizione una sequenza di una specie, si selezionava un database di sequenze derivanti da un singolo organismo evolutivamente vicino (seguendo la classificazione degli eucarioti in sei supergruppi; Adl *et al.*, 2005). Oppure si modificavano altri parametri di ricerca per riuscire a recuperare le sequenze di organismi evolutivamente distanti.

**Allineamenti.** Le sequenze ottenute tramite BLAST sono state inizialmente allineate utilizzando il programma CLUSTAL X (Thompson *et al.*, 1997) in grado di elaborare allineamenti multipli di sequenze sia proteiche che nucleotidiche, evidenziandone le zone di similarità. Successivamente gli allineamenti sono stati corretti manualmente.

**Pattern PROSITE.** Per analizzare i *pattern* creati sulle sequenze è stato utilizzato il programma ScanProsite, che confronta un *pattern* inserito dall'operatore con il database proteico di Swiss-Prot e TrEMBL, dando come *output* le sequenze contenenti frammenti che soddisfano le caratteristiche del *pattern*.

– **Analisi filogenetiche.**

Per determinare il modello che più si adatta ad un insieme di sequenze proteiche e capirne l'evoluzione, è stato utilizzato il programma **Prot-Test** versione 1.4. Dato un allineamento di proteine, questo programma è in grado di calcolare il modello migliore attraverso due criteri principali: il AIC (*Akaike Information Criterion*) ed il BIC (*Bayesian Information Criterion*). Viene calcolata sia la probabilità di ciascun modello candidato attraverso diversi parametri, sia l'importanza di ciascuno di questi parametri; vengono inserite anche le correzioni per il tasso di variazione o di invariabilità di alcuni siti (Abscal *et al.*, 2005).

Gli alberi filogenetici sono stati costruiti utilizzando 3 diversi programmi:

- **MrBayes** versione 3.1.2 (Ronquist & Huelsenbeck, 2003) per analisi Bayesiane: permette di determinare la migliore topologia dell'albero ed i valori di probabilità posteriore (PP) per i nodi attraverso 1000000 di generazioni MCMC (*Markov Chain Monte Carlo*). Un metodo rapido per determinare quando bloccare l'analisi è rimuovere graficamente gli alberi generati prima del *plateau* della probabilità dei dati: questo rappresenta il valore *burn-in*;
- **PhyML** versione 2.4.4 (Guindon & Gascuel, 2003) e **RaxML** versione 2.2.3 (Stamatakis, 2006): sono due programmi che permettono la costruzione di alberi filogenetici e danno valori di *maximum-likelihood* (ML) *bootstrap*. Il *likelihood* è la probabilità di osservare i dati nel caso l'ipotesi sia corretta, e il sistema statistico di ML sceglie l'ipotesi che rende i dati più plausibili e valuta la probabilità che i dati di sequenza siano stati prodotti da un dato modello evolutivo e da un dato percorso filogenetico. La stima statistica dell'affidabilità dei vari raggruppamenti avviene attraverso sottocampionamenti casuali dei dati (*bootstrap*).

Per poter visualizzare graficamente gli alberi filogenetici si è utilizzato il programma **TreeView** (Page, 1996).

– ***Homology modeling.***

La creazione di modelli 3D dei domini N-terminali di alcune longine sono stati effettuati utilizzando due programmi: Geno3D (<http://geno3d-pbil.ibcp.fr>; Combet *et al.*, 2001) e Phyre (<http://www.sbg.bio.ic.ac.uk/phyre>; Bennett-Lovsey *et al.*, 2008) che funzionano come *web server* automatici per generare modelli molecolari di proteine. Fornendo la sequenza da modellizzare, questi programmi cercano, attraverso analisi di similarità (PSI-BLAST), una proteina (o una sua porzione) con struttura depositata che agisce da template per la riproduzione della struttura della sequenza in esame. I programmi estraggono le restrizioni geometriche attraverso la corrispondenza degli atomi tra il template e la sequenza, e la ricostruzione 3D della proteina avviene usando l'approccio di distanza geometrica. I modelli possono essere visualizzati con il programma PyMol (<http://www.pymol.org>; De Lano, 2002): con esso si può non solo effettuare

l'*editing* della struttura, ma anche visualizzare la sovrapposizione del template con il modello generato, in modo da avere la diretta corrispondenza dei singoli aminoacidi.

## **BIOLOGIA MOLECOLARE**

### **Mutagenesi sito diretta.**

Si è seguito il protocollo di “*overlap extension mutagenesis*” (Higuchi et al., 1988; Ho et al., 1989), nel quale per generare la mutazione sono necessari 4 oligonucleotidi (*primer*) utilizzati in tre reazioni di PCR. Per garantire stabilità durante l'appaiamento con lo stampo, gli oligonucleotidi sono lunghi almeno 21 paia di basi (bp) e, quando possibile, recano G o C al 3' e hanno percentuale di GC più vicina possibile al 50%; compatibilmente con i gradi di libertà disponibili, presentano la tripletta codificante l'aminoacido da sostituire al centro della sequenza.

La prima amplificazione è mediata da un *primer forward* senza mutazione al 5' della sequenza da modificare, ed un *primer reverse* che contiene la mutazione. Nella seconda reazione il *primer forward* contiene la mutazione ed il *primer reverse* è al 3' della sequenza da modificare. I due *primer* con la mutazione devono essere sovrapposti e complementari per almeno 15 bp. In tal modo gli amplificati delle due prime reazioni di PCR, essendo sovrapposti nella regione mutagenizzata, in una terza reazione di PCR con i *primer* 5' e 3' saranno in grado di formare eteroduplex che, per estensione, completano la mutagenesi.

Dopo le reazioni di PCR si verifica la presenza delle bande delle dimensioni attese e si esclude la presenza di eventuale aspecifico mediante elettroforesi in gel di agarosio (1.5%).

### **Sequenze degli oligonucleotidi sintetici per la PCR (*primer*).**

Si riportano 6 gruppi di sequenze di *primer* utilizzati in esperimenti diversi:

- oligonucleotidi esterni con le “code” utilizzati sia nella mutagenesi-sito diretta, sia per l'amplificazione della sequenza intera di Ykt6 di lievito e successivo clonaggio nei vettori per i saggi di doppio ibrido:



YKT6\_Sc\_W1\_EcoRI 5' -GCGAATTCATGAGAATCTACTACATCGGTG-3'  
 YKT6\_Sc\_C1\_XhoI 5' -GATCTCGAGCTACATGATGATGCAACAC-3'

- oligonucleotidi utilizzati per la mutagenesi sito diretta del residuo di arginina 50 in alanina, in glutammina e acido glutammico:

YKT6\_Sc\_W\_R50A 5' -GAAACGGTCGCTTCTGCAACTGGTGCAGGAC-3'  
 YKT6\_Sc\_C\_R50A 5' -GTCCTGCACCAGTTGCAGAAGCGACCGTTTC-3'  
 YKT6\_Sc\_W\_R50Q 5' -GAAACGGTCGCTTCTCAAAGTGGTGCAGGAC-3'  
 YKT6\_Sc\_C\_R50Q 5' -GTCCTGCACCAGTTTGAGAAGCGACCGTTTC-3'  
 YKT6\_Sc\_W\_R50E 5' -GAAACGGTCGCTTCTGAAACTGGTGCAGGAC-3'  
 YKT6\_Sc\_C\_R50E 5' -GTCCTGCACCAGTTTCAGAAGCGACCGTTTC-3'

- oligonucleotidi utilizzati per creare la porzione del LD o dello SNARE *motif* di Ykt6 di lievito o umana da clonare nei vettori per i saggi di doppio ibrido:

YKT6\_Sc\_C2\_XhoI 5' -GATCTCGAGCTAATCTTGATATTTGCTAATG-3'  
 YKT6\_Sc\_W2\_EcoRI 5' -GCGAATTCACAGGCTGACGCTATC-3'

- oligonucleotidi esterni con le "code" utilizzati sia nella mutagenesi-sito diretta sia per l'amplificazione della sequenza intera di VAMP7 umana e successivo clonaggio nei vettori per i saggi di doppio ibrido:

PHA10-W1\_EcoRI 5' -GCGAATTCCTTTTTGCTGTTGTTGCCA-3'  
 New\_PHA10-C1\_BamHI 5' -GCGGATCCTCACAGGTTCTTCATACACATG-3'

- oligonucleotidi utilizzati per la mutagenesi sito diretta dei residui leucina 33 e lisina 35 in glutammina e alanina:

VAMP7\_Hs\_W\_L33Q 5' -GTGACAGAGCAGATTTCAGGCTAAGATACCTTCTG-3'  
 VAMP7\_Hs\_C\_L33Q 5' -CAGAAGGTATCTTAGCCTGAATCTGCTCTGTCAC-3'  
 VAMP7\_Hs\_W\_L33A 5' -GTGACAGAGCAGATTGCTGCTAAGATACCTTCTG-3'  
 VAMP7\_Hs\_C\_L33A 5' -CAGAAGGTATCTTAGCAGCAATCTGCTCTGTCAC-3'  
 VAMP7\_Hs\_W\_K35Q 5' -GAGCAGATTCTGGCTCAGATACCTTCTGAAAATAAC-3'  
 VAMP7\_Hs\_C\_K35Q 5' -GTTATTTTTCAGAAGGTATCTGAGCCAGAATCTGCTC-3'  
 VAMP7\_Hs\_W\_K35A 5' -GAGCAGATTCTGGCTGCTATACCTTCTGAAAATAAC-3'  
 VAMP7\_Hs\_C\_K35A 5' -GTTATTTTTCAGAAGGTATAGCAGCCAGAATCTGCTC-3'

- oligonucleotidi utilizzati per creare la porzione del LD o dello SNARE *motif* di VAMP7 umana da clonare nei vettori per i saggi di doppio ibrido:

PHA10-W3\_EcoRI 5' -GGGAATTCATCACTCTGAGAATAAGG-3'  
 New\_PHA10-C2\_BamHI 5' -GCGGATCCTCACTTATTCTCAGAGTGATGC-3'

### Reazione a catena della polimerasi (PCR).

Si conduce ciascuna reazione in 10 µl, utilizzando: 1-50 ng del DNA da amplificare, proveniente da precedenti clonaggi effettuati nel nostro laboratorio; 0.25 µM di *primer forward* e *reverse*; 200 µM di ciascun dNTP; 1.5 mM MgCl<sub>2</sub>; 3.5 unità di Taq polimerasi Expand™ High Fidelity (Roche Diagnostics) nel

tampone 10x dedicato. Le reazioni di PCR sono condotte in un termociclatore Gene Amp PCR System 9700 (Perkin-Elmer). Per ciascuna coppia di *primer* con “code” è stata allestita una reazione di PCR in cui i primi cicli di amplificazione del template sono stati condotti ad una temperatura prossima a quella degli oligonucleotidi sprovvisti delle “code”.

### **Elettroforesi in gel di agarosio.**

La presenza e la corretta lunghezza dei prodotti di amplificazione è verificata mediante separazione elettroforetica su gel di agarosio (ultraPURE agarose electrophoresis grade, Gibco-BRL) ad una concentrazione di 1% (p/v). L'agarosio in polvere viene sciolto nel tampone TAE (40 mM Tris/Acetato, 1mM EDTA), in presenza di 0.25 µg/ml di bromuro di etidio, un agente in grado di intercalarsi nella molecola di DNA ed emettere fluorescenza, se irradiato con luce UV. Ai campioni da analizzare viene aggiunto un tampone di caricamento (*loading buffer* 1X: 0,1% SDS, 5% glicerolo e 0,005% di blu di bromofenolo). Insieme ai campioni da analizzare viene caricato un marcatore di peso molecolare opportuno. Il gel viene successivamente immerso in un tampone di corsa (TAE 1X) e sottoposto ad un campo elettrico di 50-100 V per un tempo variabile. Le concentrazioni di DNA sono determinate per analisi densitometrica tramite il sistema di analisi digitale delle immagini Chemi-Doc (BioRad) ed il *software* dedicato Quantity One (BioRad).

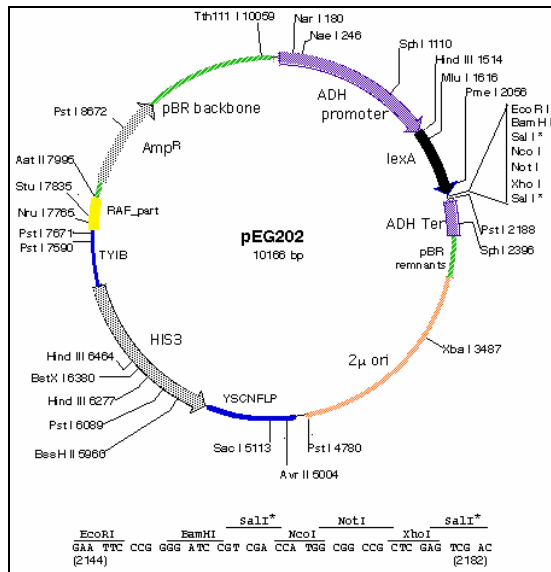
### **Plasmidi.**

**TopoTA** (Invitrogen). Il vettore (pCR<sup>®</sup>2.1-TOPO<sup>®</sup>, 3.9 Kb) è fornito linearizzato, con singole timidine (T) 3'-sporgenti a cui è covalentemente legata la topoisomerasi I di *Vaccinia* virus. Consente di clonare in un passaggio prodotti di PCR. Possiede geni per la resistenza alla kanamicina ed all'ampicillina, e l'origine di replicazione pUC per il mantenimento e la propagazione in batteri. Il sito di policlonaggio è inserito tra il promotore *Plac* ed il gene *lacZα*, per consentire la selezione delle colonie positive mediante saggio colorimetrico su piastra, ed è fiancheggiato dalle sequenze *forward* e *reverse* di *priming* per M13. Il kit, utilizzato seguendone le relative istruzioni, contiene anche tutti i reagenti necessari per il clonaggio (tamponi, dNTP, *primer forward* e *reverse* M13, template e *primer* di controllo).

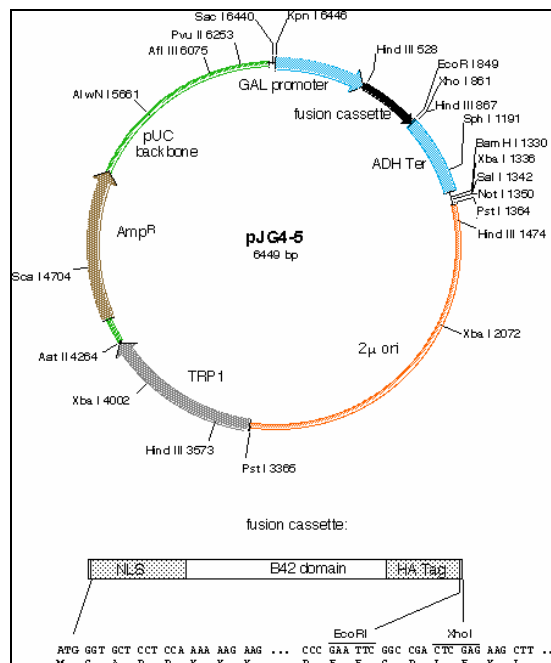
**pEG202** (fornito gentilmente dal Dr D.K. Banfield, Hong Kong University). E' un vettore di circa 10,2 Kb utilizzato per saggi di doppio ibrido; permette la produzione di proteine di fusione formate dal BD della proteina LexA di *E. coli* (residui 1-102) e dalla sequenza di interesse, inserita nel sito di policloneggiamento (MCS) a valle di LexA. La proteina di fusione viene espressa in lievito sotto il controllo del promotore troncato del gene *ADHI* di *S. cerevisiae*; la trascrizione si ferma a livello del terminatore del gene *ADHI* e la proteina tradotta è indirizzata al nucleo dalle sequenze NLS (*Nuclear Localization Signal*) di LexA. Il vettore possiede le origini di replicazione pBR e 2 $\mu$  per la replicazione rispettivamente in batteri e in lievito; per la selezione di trasformanti, porta il gene *bla* per la resistenza all'ampicillina in *E. coli*, ed il marcatore *HIS3* per la sintesi dell'istidina in lievito (Figura 5).

**pJG4-5** (Invitrogen). E' un vettore di 6.5 Kb che esprime proteine fuse all'AD di B42; il MCS comprende solo i siti di restrizione per EcoRI e XhoI. L'espressione della proteina è sotto il controllo del promotore di GAL1 e la trascrizione si ferma a livello del terminatore di GAL1. La proteina è indirizzata al nucleo dalle sequenze NLS a monte dell'AD. Si replica in batteri e lieviti utilizzando le origini di replicazione pUC e 2 $\mu$  rispettivamente. Possiede il gene per la resistenza all'ampicillina per la selezione in batteri e il marcatore nutrizionale *TRP1* per la selezione in lievito (Figura 6).

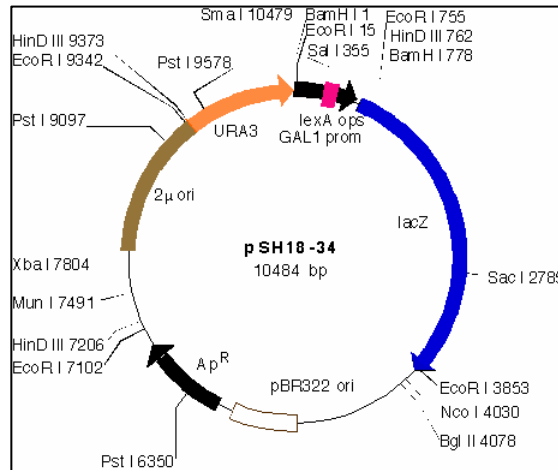
**pSH18-34** (Invitrogen). E' un plasmide *reporter* di circa 10.5 kb, contenente la sequenza del gene *reporter* LacZ. La sua espressione è controllata dal promotore di GAL1 e dall'operatore di *lexA*, presente in 8 copie. Si replica in *E. coli* sfruttando l'origine di replicazione Col E1 derivata da pBR322, e in lievito tramite l'origine 2 $\mu$ . Possiede il gene per la resistenza all'ampicillina per la selezione in batteri e il marcatore nutrizionale URA3 per la selezione in lievito (Figura 7).



**Figura 5:** Mappa e MCS del vettore pEG202 (tratto dalla pagina web: <http://www.fccc.edu/research/labs/golemis/plasmids/pEG202.html>).



**Figura 6:** Mappa e MCS del vettore pJG4-5 (tratto dalla pagina web: <http://www.fccc.edu/research/labs/golemis/plasmids/pJG4-5.html>).



**Figura 7:** Mappa del vettore *pSH18-34* (tratto dalla pagina web: <http://www.fccc.edu/research/labs/golemis/plasmids/pSH18-34.html>).

### Ceppo batterico e terreni di coltura.

Per tutti i clonaggi è stato utilizzato il ceppo di *E. coli* DH10B avente genotipo  $F^-$  *mrcA*  $\Delta$ (*mrr*-*hsdRMS*-*mrcBC*)  $\Phi$ 80 *lacZ* $\Delta$ M15  $\Delta$ *lacX*74 *deoR* *recA1* *endA1* *araA139*  $\Delta$ (*ara*,*leu*)7697 *galU* *galK* $\lambda^-$  *rspL* (*StrR*) *nupG*. Le cellule sono state coltivate in terreno liquido sterile LB (Luria-Bertani, contenente 10g/L di peptone, 5g/L di estratto di lievito, 10g/L di NaCl) o in terreno solido addizionato con agar 1% (SIGMA). Questi terreni sono stati addizionati con 100  $\mu$ g/ml di ampicillina per la selezione delle colonie contenenti i plasmidi che conferiscono resistenza all'antibiotico stesso.

### Ceppo di lievito e terreni di coltura.

Per le trasformazioni di lievito è stato utilizzato il ceppo di *Saccharomyces cerevisiae* EGY48 avente genotipo *MATa*, *ura3*, *his3*, *trp1*, *LexA<sub>op(x6)</sub>-LEU2* e dotato del gene reporter *LEU2*, la cui espressione è sotto il controllo dell'operatore di LexA e del promotore minimo di *LEU2*. Consente la selezione e il mantenimento di plasmidi in grado di sintetizzare uracile, istidina e triptofano, per i quali è auxotrofo.

**Terreni selettivi.** Dal terreno sintetico minimo (SD, costituito da 1,7 g/L Yeast Nitrogen Base (Invitrogen), 5 g/L ammonio solfato, vari aminoacidi aventi concentrazioni comprese tra 20 e 200 mg/L, agar 2% (SIGMA) per il terreno solido), sono stati ottenuti i terreni selettivi privi di uracile, istidina e triptofano

per la selezione e il mantenimento dei lieviti trasformati; è stato utilizzato un terreno privo anche di leucina per lo studio dell'interazione delle proteine clonate.

### **Clonaggio mediante kit TOPO TA Cloning®.**

Dopo la PCR con *Taq* polimerasi “*proofreading*”, l'amplificato è clonato con il kit TOPO TA Cloning® seguendo le istruzioni della ditta: si addiziona il prodotto di PCR con 1 unità di *Taq* polimerasi non “*proofreading*” al fine di aggiungere le A sporgenti al 3' necessarie per il TA *cloning*; lo si incuba a 72°C per 10'; successivamente si prelevano 4 µl di miscela e si aggiungono 1 µl di Salt Solution ed 1 µl di vettore pCR®2.1-TOPO®, incubando a 22-23° C per 5'. 4 µl di questo prodotto sono sufficienti per trasformare 100 µl di cellule batteriche competenti DH10B. Queste ultime sono seminate su piastre addizionate con kanamicina e superficie precedentemente trattata con 40 µl di X-gal (40 mg/ml) per la rilevazione colorimetrica. I cloni positivi appaiono bianchi poiché nei vettori pCR®2.1-TOPO® l'inserto è clonato nel sito di policlonaggio tra promotore *Plac* e gene *lacZα*.

### **Estrazione e purificazione del DNA plasmidico da cellule batteriche.**

Il DNA plasmidico su piccola scala (miniprep; Sambrook & Russell, 2001a) è ottenuto seguendo il metodo della lisi alcalina che prevede due fasi:

estrazione del DNA plasmidico: partendo dalla coltura satura di *E. coli* in terreno liquido selettivo, i batteri sono raccolti per centrifugazione e risospesi in 100 µl/campione di soluzione A (50 mM glucosio, 25 mM Tris-Cl pH 8.0, 10 mM EDTA pH 8.0) contenente 1 mg/ml lisozima e 0.5mg/ml RNasi A (QIAGEN). Dopo 5' a temperatura ambiente si aggiungono 200 µl di soluzione B (0.2 M NaOH, 1% SDS) e, dopo altri 5' di incubazione, 150 µl di soluzione C (3 M K-acetato, 11.5% acido acetico); il tutto è posto in ghiaccio per 5'. Si eliminano i detriti cellulari per doppia centrifugazione (in *microfuge* da banco, alla massima velocità) e si conserva il surnatante;

purificazione del DNA plasmidico: al surnatante, che contiene il DNA plasmidico, è aggiunto 1 volume di isopropanolo per 5' a T° ambiente. Poi si centrifuga per 15' alla massima velocità e il precipitato di DNA è lavato in 500 µl di etanolo al 70% (v/v). Dopo aver eliminato il surnatante, l'etanolo residuo viene fatto evaporare ed infine il *pellet* di DNA plasmidico viene risospeso in H<sub>2</sub>O sterile o in TE pH 8.0 (10 mM Tris-Cl pH 8.0, 1 mM EDTA pH 8.0).

### **Taglio del DNA con enzimi di restrizione.**

Le reazioni di restrizione degli amplificati sono eseguite con gli enzimi opportuni (Promega o Takara) in base alle sequenze di riconoscimento presenti nelle “code” degli amplificati. Gli stessi enzimi sono usati anche per linearizzare i plasmidi da utilizzare per i successivi clonaggi. Ogni reazione di restrizione viene condotta in un tampone adeguato, impiegando per ciascun enzima 1 unità per  $\mu\text{g}$  di DNA. Le reazioni sono condotte per circa 1 ora alla temperatura ottimale per gli enzimi usati.

### **Purificazione dei frammenti di DNA dopo PCR o dopo restrizione.**

I frammenti di DNA (amplificati ottenuti via PCR o plasmidi linearizzati) vengono separati per elettroforesi su gel di agarosio allo 0.8% (p/v) e, dopo verifica dei pesi molecolari per comparazione con i marcatori, le bande sono escisse dal gel. Il DNA è recuperato con il kit di estrazione QIAquick Gel Extraction Kit (QIAGEN), seguendone il protocollo che prevede: aggiunta di tamponi per legare il DNA alla colonna contenente una membrana di gel-silice, lavaggio del DNA ed eluizione dello stesso con tamponi dedicati. I tamponi provvedono a mantenere un'adeguata concentrazione di sali e un opportuno pH per rendere ottimale prima l'adsorbimento del DNA alla membrana e poi il suo recupero. I frammenti di DNA ottenuti dopo restrizione enzimatica di prodotti di PCR vengono purificati direttamente, senza passaggio in gel di agarosio, utilizzando il QIAquick PCR Purification Kit (QIAGEN). Entrambi i kit permettono la purificazione di DNA da circa 100 bp a 10 Kb.

### **Ligazione.**

Le quantità di DNA plasmidico e di inserto necessarie per la ligazione sono calcolate tenendo conto delle dimensioni relative dei due frammenti; in tutti i casi il rapporto molare utilizzato è stato di 1:6 (plasmide:inserto). La ligazione è stata effettuata con 1 unità di T4 DNA ligasi -USB- (diluata nell'apposito tampone: 20 mM Tris-HCl, pH7.6, 1mM EDTA, 5mM DTT, 60mM KCl, 50% glicerolo), nel tampone di ligazione (10X= 660 mM Tris-HCl, pH 7.6, 66 mM  $\text{MgCl}_2$ , 100 mM DTT, 660  $\mu\text{M}$  ATP) in un volume finale di 30  $\mu\text{l}$  in milliQ. La reazione è proseguita per 16 ore a 16°C.

### **Induzione della competenza e trasformazione delle cellule batteriche.**

Per l'induzione della competenza nelle cellule batteriche si è utilizzato il protocollo che si basa su  $\text{CaCl}_2$  come fattore di induzione (Sambrook & Russell, 2001b). La trasformazione è stata ottenuta con il metodo dello shock termico (Sambrook & Russell, 2001b). A 50  $\mu\text{l}$  di cellule competenti è stato aggiunto il prodotto della ligazione; la miscela è stata lasciata 30' in ghiaccio e 90'' a 42°C, per produrre dei pori nella parete batterica. È stato poi aggiunto il terreno LB (900  $\mu\text{l}$ ), e la miscela è stata incubata a 37°C per 1 ora. Infine i batteri sono stati piastrati su terreno LB-ampicillina (utilizzando 1/10 della coltura per una piastra e i 9/10 per una seconda, dopo aver concentrato le cellule per centrifugazione).

### **Sequenziamento del DNA.**

Tutti i prodotti di PCR e i plasmidi con inserto ottenuti dopo ligazione vengono sequenziati per accertare l'assenza di mutazioni puntiformi o l'avvenuta mutagenesi. Per il sequenziamento ci si affida al Sequencing Core presso BMR Genomics s.r.l. (Padova).

### **Conservazione dei cloni batterici a – 80°C.**

I batteri trasformati con il prodotto di una specifica ligazione o con un plasmide vuoto possono essere conservati a –80°C. Si parte da un preinoculo di batteri trasformati in 1.8 ml di terreno liquido selettivo sterile fatto crescere a 37°C per tutta la notte a 180 rpm. Si precipitano i batteri in microcentrifuga alla massima velocità per 5'; dopo aver eliminato il surnatante, il precipitato batterico viene risospeso in 1 ml di LB/antibiotico addizionato con 20% glicerolo e sterilizzato per filtrazione, e si conserva a –80°C.

### **Trasformazione delle cellule di lievito.**

Il primo passaggio per la trasformazione del lievito è la preparazione del DNA carrier (2 mg/ml di DNA di sperma di salmone, SIGMA) in TE (10 mM Tris HCl, 1 mM EDTA pH 8) (Schiestl e Gietz, 1989), la cui qualità viene saggiata con una corsa in gel di agarosio 1%. Prima dell'uso viene denaturato con 3 cicli di bollitura per 5' seguiti da raffreddamento in ghiaccio. Le cellule da trasformare vengono preincolate nel terreno opportuno e fatte crescere tutta la notte a 30°C. Prima della trasformazione, è necessario diluire il preinoculo fino ad un valore di



assorbanza  $O.D._{600} = 0,4$ . Per la trasformazione è stato seguito il metodo del PEG/litio acetato (LiAc) (Agatep *et al.*, 1998). Il lievito viene fatto crescere fino a raggiungere un  $O.D._{600} = 0,8$  in agitazione a  $30^{\circ}C$ ; si prelevano 2,5 ml di coltura e si raccolgono le cellule tramite centrifugazione (5' a 1000g). Si lava con 500  $\mu$ l di TE pH 7,5 e la risospende in 50  $\mu$ l di 100 mM LiAc 1X, pH 7,5. Per ogni trasformazione sono posti in una provetta:

- 25  $\mu$ l di DNA carrier 2mg/ml;
- 500 ng di plasmide in 25  $\mu$ l di TE pH 7,5;
- 50  $\mu$ l di cellule di lievito in 100 mM LiAc;
- 262,5  $\mu$ l PEG 50%:
- 50  $\mu$ l 1M DTT;
- 45  $\mu$ l LiAc 1M, pH 7,5;
- 42,5  $\mu$ l TE 10x, pH 7,5.

La provetta viene incubata per 5' in agitazione a  $28^{\circ}C$  e sottoposta a shock termico ( $42^{\circ}C$  per 30'). Le cellule vengono poi precipitate, risospese nell'appropriato mezzo selettivo e piastrate su mezzo selettivo agarizzato. Le piastre vengono incubate a  $30^{\circ}C$  per la crescita delle colonie.

## SCOPO della TESI

L'attività di ricerca riportata in questa tesi è vincolata ad un progetto finanziato dalla Fondazione CA.RI.PA.RO. che prevede lo studio della modulazione funzionale e delle interazioni dei domini longinici di VAMP7 e Ykt6. Queste due proteine appartengono alla famiglia delle SNARE, i cui membri sono fondamentali per il traffico subcellulare negli eucarioti. Possiedono uno SNARE *motif* coinvolto nella formazione del complesso di fusione, che permette l'ancoraggio della membrana della vescicola con la membrana *target*. Una regolazione di questo meccanismo è stata inizialmente riconosciuta nell'interazione dello SNARE *motif* con il dominio N-terminale Habc delle sintaxine. Successivamente, mentre il nostro gruppo dimostrava che nelle sottofamiglie VAMP7, Sec22b e Ykt6 era presente un'estensione N-terminale che sembrava conservata negli eucarioti e definiva perciò un nuovo dominio (il dominio longinico LD), il gruppo di D.K. Banfield (inizialmente indipendente) rivelava la struttura del dominio N-terminale di Ykt6 di lievito e ne dimostrava l'interazione intramolecolare con lo SNARE *motif*. Queste prime informazioni hanno fatto sì che l'attività del nostro gruppo continuasse con: identificazione della presenza del LD in proteine negli eucarioti; caratterizzazione delle interazioni tra i LD e gli altri domini; studio della regolazione genica e post-trascrizionale; studio del coinvolgimento del LD nel normale sviluppo cerebrale ed in patologie neurologiche e neurodegenerative (anche in collaborazione con altri gruppi).

Gli scopi di questa tesi sono: dare un contributo all'identificazione della presenza del LD nelle proteine eucariotiche e alla comprensione della distribuzione delle famiglie longiniche, attraverso uno *screening* bioinformatico di proteomi e genomi; creare degli alberi filogenetici per chiarire l'evoluzione delle longine sia tra le diverse famiglie che all'interno di ogni famiglia; ottenere suggerimenti funzionali attraverso la creazione di modelli 3D del LD a partire da quelli depositati, osservando conservazione o divergenza di specifiche regioni; verificare se il meccanismo di interazione intramolecolare di Ykt6 risulta essere specie-specifico e se esso presenta delle variazioni nella famiglia delle VAMP7. In particolare, la comparazione tra VAMP7 e Ykt6 permette di capire se il LD di due

longine diverse, che dovrebbero presentare dei meccanismi base conservati dovuti proprio alla presenza del dominio longinico, può avere dei ruoli non solo famiglia-specifici, ma anche proteina-specifici .

## RISULTATI E DISCUSSIONE

### Distribuzione delle longine negli eucarioti.

Lo studio di proteine coinvolte nel trasporto vescicolare effettuato comprendendo anche organismi diversi dagli organismi modello, permette di identificarne la conservazione, presumibilmente essenziale, per il *trafficking* di membrana, così come evidenziarne la diversità di distribuzione tra gli eucarioti. Secondo la nuova classificazione gli eucarioti possono essere suddivisi in sei supergruppi (Chromalveolata, Archeplastida, Opisthokonta, Excavata, Amoebozoa e Rhizaria; Adl *et al.*, 2005) e, finora, studi sui sistemi di trasporto intracellulare sono stati condotti prevalentemente su *S. cerevisiae* ed altri metazoi appartenenti agli Opisthokonta. Non ci sono quindi molte informazioni sul meccanismo di trasporto negli altri supergruppi, ed un primo passo verso la loro comprensione è la ricerca, estesa a rappresentanti dei sei supergruppi, della distribuzione delle proteine coinvolte. In particolare, questa tesi ha tra gli scopi la comprensione di presenza e distribuzione delle famiglie longiniche VAMP7 e Ykt6 (appartenenti alle R-SNARE): tramite l'utilizzo di sequenze sorgenti, con BLAST si sono analizzati *database* di sequenze genomiche, proteiche e di EST di organismi appartenenti a tutti i supergruppi per individuare sequenze omologhe. Sono stati utilizzati i *database* presenti presso NCBI, EBI e JGI, seguendo la metodica di ricerca di similarità delle sequenze riportata nella sezione "Materiali e Metodi". Per aumentare la rappresentanza degli organismi, si sono utilizzati anche *database* di EST, che hanno permesso di includere organismi per i quali il dato genomico è non depositato o parziale.

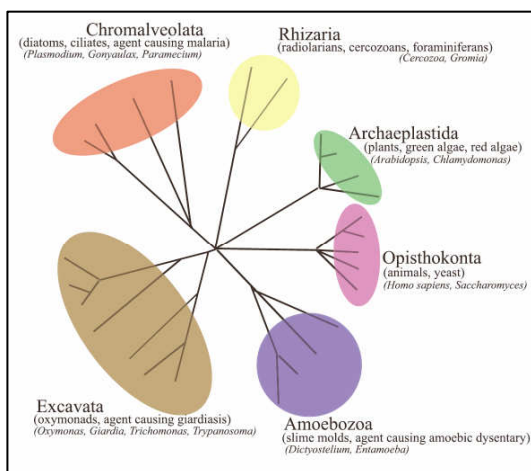
Il *dataset* iniziale così ottenuto ha permesso di identificare più di 200 sequenze omologhe in cinque supergruppi: l'analisi non ha compreso i Rhizaria a causa dell'esigua disponibilità di genomi sequenziati. Infatti, non esiste attualmente nessun rappresentante di questo supergruppo avente il genoma interamente sequenziato.

Sono state incluse nella ricerca anche le sequenze delle famiglie longiniche delle Sec22b, delle Nyv1 e delle brevine per verificare la distribuzione di tutte le famiglie longiniche e confrontarla anche con quella delle brevine.

Come riportato in Figura 8, tutte e tre le famiglie longiniche principali (Ykt6, Sec22b e VAMP7) sono presenti nei 5 supergruppi eucariotici ma con distribuzioni diverse. Ykt6 è presente in tutti gli organismi considerati, e sembra essere quindi la longina “essenziale”; inoltre, in alcuni organismi sono presenti 2 geni. Sec22b, sebbene presente in tutti i supergruppi, può avere una distribuzione particolare: presenza di più di 2 geni in alcuni organismi o assenza come ad esempio in alcuni funghi. VAMP7 è presente in tutti i supergruppi, ma ha una distribuzione più variabile: ci può essere un solo gene, come nei metazoi, fino a più di 10 geni in alcune piante superiori come *A. thaliana*. In *S. cerevisiae* è sostituita da Nyv1p, che è presente solo nei saccaromiceti, e il cui LD è strutturalmente riconoscibile, ma possiede una sequenza molto degenerata (Wen *et al.*, 2006).

La presenza delle brevine, riscontrata solo nei supergruppi degli Opisthokonta e Amoebozoa, rende questa famiglia non conservata come le longine.

Considerando i risultati ottenuti, emerge che le longine costituiscono una famiglia proteica evolutivamente conservata in tutti gli eucarioti, rendendo ipotizzabile un loro ruolo essenziale nel trasporto vescicolare.



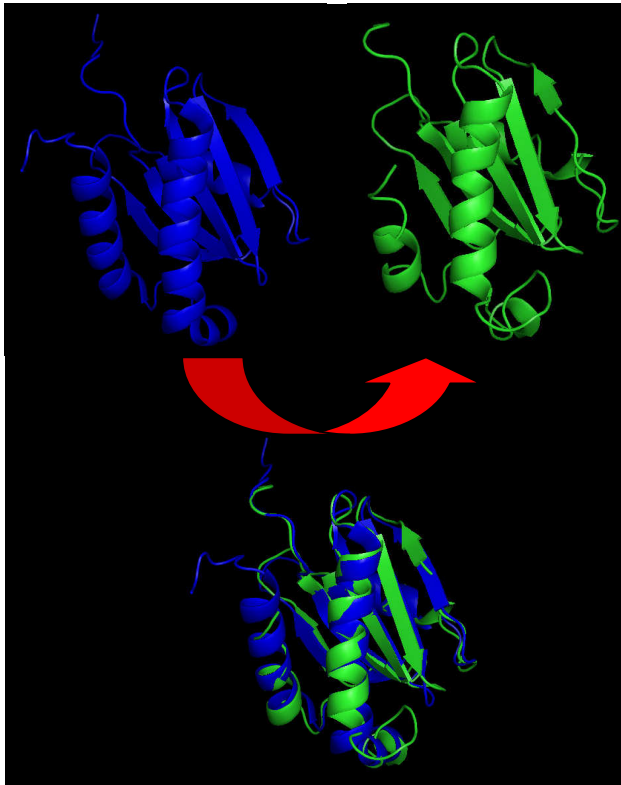
	Ykt6	Sec22b	VAMP7	OSSERVAZIONI
<b>OPISTHOKONTA</b>	1-2	1-2	0-2	Presenza di Nyv1 in Saccaromiceti
<b>AMOEBOZOA</b>	1-2	0-1	2	Assenza di Sec22b in alcuni funghi
<b>ARCHAEPLASTIDA</b>	1-2	1-3	1-11	Massima amplificazione di VAMP7 nelle piante
<b>EXCAVATA</b>	1	1-2	1-8	
<b>CHROMALVEOLATA</b>	1-2	1	1-5	

**Figura 8:** nel diagramma superiore è riportata la classificazione degli eucarioti in 6 supergruppi (da Dacks, in “SNAREs” in press); la distribuzione numerica delle longine nei cinque supergruppi considerati è rappresentata in tabella.

### **Identificazione di una nuova famiglia proteica *longin-like*: le Fitolongine.**

Nella ricerca di rappresentanti delle famiglie longiniche ci siamo imbattuti in alcune sequenze il cui dominio N-terminale presentava un'identità proteica media del 50% con il LD (da analisi di Blastp), ma non sembravano possedere lo SNARE *motif*. Inizialmente, si pensava derivassero da predizioni di sequenze proteiche errate, ma questa ipotesi è stata subito smentita dalla loro presenza anche in *database* di sequenze espresse. Abbiamo perciò ritenuto opportuno verificare la distribuzione di queste proteine negli eucarioti: la loro presenza è specifica delle piante, ed inoltre, analizzando più in dettaglio il supergruppo delle Archaeplastida (che comprende alghe e piante terrestri), si è scoperto che esse sono presenti solo in Embryophyta, a partire da *Physcomitrella patens* (muschio) fino alle piante superiori. In queste ultime, come ad esempio in *A. thaliana*, ci sono 4 geni che codificano per queste proteine.

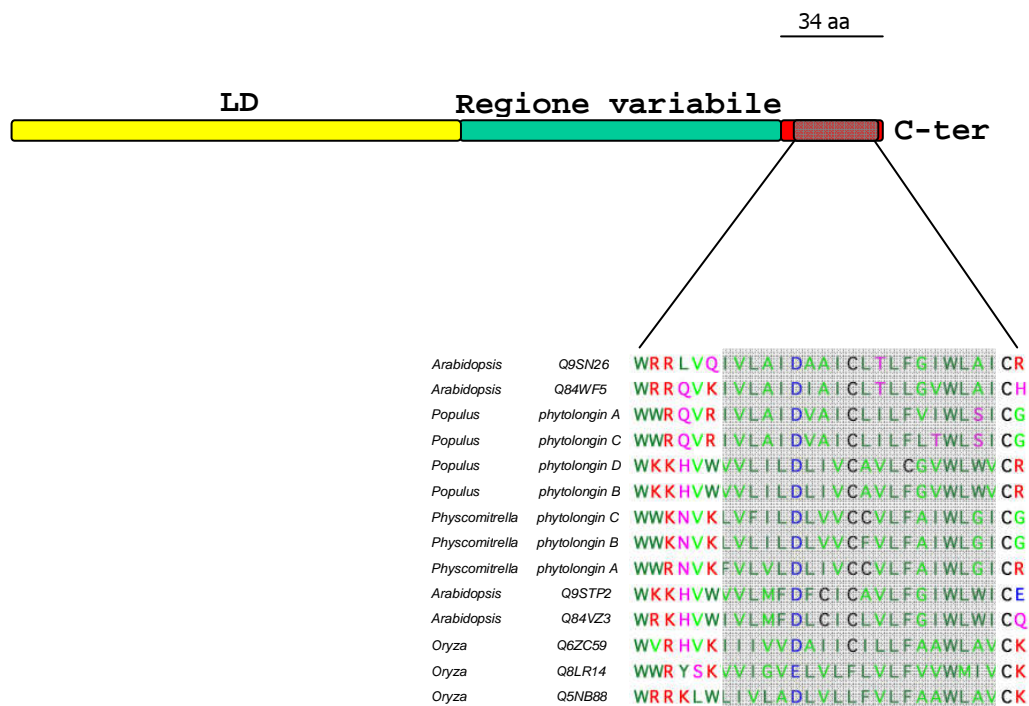
Per ottenere indicazioni sulla struttura del loro dominio N-terminale, abbiamo provato ad utilizzare dei programmi di *homology modeling*, ma inizialmente non siamo stati in grado di avere nessun template utile. I modelli di LD depositati appartenevano ad una Ykt6 di lievito ed una Sec22b di topo (Gonzalez *et al.*, 2001; Tochio *et al.*, 2001), sui quali non era possibile modellizzare né le VAMP7 né le Nyv1. Recentemente sono state depositate le strutture di quest'ultime, ed è stato possibile ottenere il modello del dominio N-terminale di questa nuova famiglia, usando come template la struttura del LD di VAMP7 umana e come sequenza da modellizzare la Q9SN26 di *A. thaliana*. Il modello ottenuto è costituito da cinque foglietti  $\beta$  antiparalleli racchiusi da un lato da una  $\alpha$  elica e dall'altro da due  $\alpha$  eliche: esso quindi rappresenta il dominio longinico (Figura 9). Aver ottenuto il modello a partire da una VAMP7 suggerisce la possibilità che questa famiglia proteica sia più correlata alle VAMP7 rispetto alle altre longine. L'informazione sulla presenza di questa famiglia solo nelle piante, unita alla definizione del loro dominio N-terminale come dominio longinico, ci ha permesso di definire queste proteine col termine Fitolongine.



**Figura 9:** la struttura in blu rappresenta il LD di VAMP7 usato come template per ottenere il modello (in verde) del dominio N-terminale della Fitolongina di *Arabidopsis* (numero di accesso Q9SN26). L'immagine in basso rappresenta la sovrapposizione tra i due modelli ottenuta tramite PyMOL.

La differenza principale tra le Fitolongine e le altre famiglie longiniche risiede nella mancanza dello SNARE *motif*; esse possiedono invece una regione centrale di lunghezza variabile da 60 a 100 aminoacidi con “caratteristiche” conservate, quali la presenza di regioni a bassa complessità ricche in aminoacidi carichi. Per questa regione si è cercato di capire se corrispondesse a qualche altro dominio conosciuto. Nonostante nella ricerca siano stati modificati molti parametri rendendola meno stringente e andando a reiterare le analisi in *database* di sequenze di piante e poi di altri eucarioti, non sono stati identificati domini omologhi, pur avendo analizzato anche gli UTR e gli introni. Inoltre sono state estese le ricerche anche verso i procarioti, ma non si è ottenuta nessuna informazione che permettesse di capire la derivazione di questa regione centrale. Per quanto riguarda la regione C-terminale, è emerso che essa è altamente conservata in tutte le Fitolongine, è molto simile al C-terminale delle VAMP7 e da analisi di predizione topologica, utilizzando programmi quali HMMTOP (Tusnàdy & Simon, 2001), sembra comprendere un dominio transmembrana.

L'architettura delle Fitolongine è quindi descrivibile come una regione N-terminale che possiede un LD (simile a quello delle VAMP7), seguita da una regione centrale variabile e da un putativo dominio transmembrana C-terminale: rappresenta perciò un prototipo di longina non-SNARE (Figura 10) (Vedovato *et al.*, in preparazione).



**Figura 10:** rappresentazione schematica del prototipo di una Fitolongina. E' riportata l'allineamento della porzione C-terminale di alcune Fitolongine dove in grigio è evidenziata la putativa regione transmembrana. Solo per le sequenze depositate è riportato il relativo numero di accesso.

### Analisi filogenetica delle famiglie longiniche.

Avendo a disposizione un ampio *set* di sequenze provenienti da organismi diversi, estratte per le analisi descritte nei precedenti paragrafi, si è pensato di utilizzarlo anche per comprendere l'organizzazione delle longine dal punto di vista evolutivo.

Per acquisire competenza nella creazione di alberi filogenetici e realizzare le analisi in collaborazione con un ricercatore interessato al sistema di *protein trafficking*, mi sono recato a Cambridge (UK) presso il laboratorio del Dr. Joel B. Dacks (attualmente Research Group Leader al Department of Pathology, University of Cambridge) e, sotto la sua supervisione, ho creato alcuni alberi



filogenetici che o contenevano tutti i rappresentanti delle famiglie longiniche, o escludevano alcuni gruppi per comprenderne meglio le rispettive relazioni.

Il procedimento per l'analisi filogenetica di sequenze può essere suddiviso in cinque parti principali:

1. selezione delle sequenze;
2. allineamento multiplo;
3. applicazione di un metodo di ricostruzione filogenetica;
4. visualizzazione dell'albero;
5. validazione della ricostruzione filogenetica.

Le sequenze derivanti dalle ricerche per comprendere la presenza e distribuzione delle famiglie longiniche rappresentano un *set* di dati ampio, ma che non può essere utilizzato nella sua interezza per le analisi filogenetiche. Infatti, per minimizzare eventuali errori derivanti da una sbilanciata rappresentatività delle sequenze, si è cercato di selezionare le sequenze longiniche provenienti da un egual numero di rappresentanti di ciascuno dei cinque supergruppi eucariotici. In alcuni organismi rappresentativi non siamo però riusciti ad ottenere sequenze appartenenti a tutte le famiglie longiniche. La causa potrebbe derivare da una reale assenza di una famiglia longinica, oppure da un'estrema divergenza nella sequenza difficilmente rilevabile tra organismi evolutivamente distanti.

Per la creazione degli alberi filogenetici sono stati inizialmente realizzati degli allineamenti sviluppati sulle strutture secondarie delle sequenze rappresentative di ciascun supergruppo, considerando poi solo le regioni chiaramente omologhe. Con queste regioni allineate si è creato un *file* utilizzabile dal programma Protest che permetteva di scegliere il modello evolutivo migliore. I modelli per l'evoluzione delle proteine sono basati su matrici empiriche, e questo programma permette di individuare la matrice più probabile: dai nostri dati spesso veniva riconosciuta come più probabile la matrice WAG (Whelan & Goldman, 2001). Una volta scelto il modello, per la creazione degli alberi filogenetici si sono utilizzati i programmi PhyML, RaxML e MrBayes (descritti nella sezione Materiali e Metodi). Tutti e tre i programmi sono disponibili presso *bioinformatics and computational biology service (mole)* di Cambridge (<http://www.bio.cam.ac.uk/molbio.html>) previa iscrizione. Questo servizio fornisce non solo i programmi sopra elencati, ma anche più risorse di calcolo grazie all'utilizzo contemporaneo di più computer, permettendo perciò una

maggior rapidità delle analisi computazionali. Sono stati utilizzati tre programmi diversi nella creazione degli alberi filogenetici per avere la possibilità di comparare approcci diversi ed ottenere una maggior affidabilità delle analisi evolutive.

Le analisi filogenetiche iniziali sono state effettuate per identificare eventuali sequenze *long-branch*. Queste possono infatti contribuire ad artefatti, creando il cosiddetto “*long branch attraction*”: due sequenze che possiedono un ramo lungo vengono avvicinate tra loro, anche se nel vero albero non lo sono. Queste sequenze sono state rimosse, si sono ricontrollate le regioni da mantenere e, successivamente, le analisi filogenetiche sono state ripetute per confermare o smentire i dati ottenuti dalle analisi precedenti. Sono state utilizzate le sequenze di tutte le famiglie longiniche (comprese Nyv1 e Fitolongine) non solo per capire la relazione tra le principali famiglie, ma anche per determinare da quale famiglia proteica derivassero le Fitolongine. Dagli alberi creati con i tre programmi, la prima impressione è stata che le tre famiglie principali derivano da un antecore comune, mentre sia Nyv1 che le Fitolongine derivano dalle VAMP7. I valori riportati nell’albero in Figura 11 permettono di capire quali rami siano stati ben risolti; per altri rami, invece, è stato necessario produrre ulteriori alberi filogenetici per ottenere un buon valore che permettesse di sostenere “l’immagine” dell’albero.

Il primo risultato ottenuto è la distinzione della famiglia Ykt6 dalle altre famiglie e la non derivazione delle Fitolongine dalle Ykt6. Questo dato rende corretta l’ipotesi delle Ykt6 come una famiglia longinica altamente specializzata. Si pensa perciò che si sia separata subito dall’antecore delle longine e, una volta acquisite le funzioni che le vengono attribuite, è rimasta altamente conservata ed indispensabile per gli eucarioti.

Per determinare se la famiglia delle Sec22b fosse l’antecore delle Fitolongine, è stato necessario ripetere le analisi escludendo le Ykt6, per permettere ai programmi di “concentrare” i loro calcoli sulle relazioni tra Sec22b, VAMP7 e Fitolongine. I risultati hanno confermato che le Fitolongine non derivano nemmeno dalle Sec22b (Figura 12).



**Figura 11:** *albero filogenetico delle famiglie longiniche. Sono riportati i valori di MrBayes, PhyML e RaxML solo per i rami con valori superiori rispettivamente a 0.80-50-50. I rami con valori > di 0.95-80-80 sono indicati con il simbolo ○, i rami con valori > 1.0-95-95 sono indicati con ● e rappresentano i rami meglio risolti.*



**Figura 12:** albero filogenetico delle famiglie longiniche Sec22b, VAMP7 e Fitolongine. Sono riportati i valori di MrBayes, PhyML e RaxML solo per i rami con valori superiori rispettivamente a 0.80-50-50. I rami con valori > di 0.95-80-80 sono indicati con il simbolo **o**, i rami con valori > 1.0-95-95 sono indicati con **•** e rappresentano i rami meglio risolti.

Nelle piante terrestri sono presenti più geni codificanti le VAMP7 che sono suddivise nelle due sottofamiglie VAMP7.1 e VAMP7.2. La sottofamiglia che presenta la maggiore amplificazione genica è quella delle VAMP7.2, dalle quali le Fitolongine sembrano derivare. Dall'albero riportato in Figura 12, le Fitolongine risultano essere non solo chiaramente evolute dalle VAMP7.2 ma, a loro volta, suddivisibili nelle piante superiori in due sottofamiglie. Questi risultati sono stati confermati dalla costruzione di un ulteriore albero utilizzando solamente le sequenze di VAMP7 di piante e delle Fitolongine, creando così un *set* specifico. I valori ottenuti in quest'ultima analisi sono concordi con i valori riportati nell'albero in Figura 12.

### **Creazione di *pattern* per le longine.**

Dagli allineamenti utilizzati per le analisi evolutive, è emersa la presenza di regioni altamente conservate all'interno del dominio longinico. Per capire la funzionalità di tali regioni però si è utilizzato un *set* di dati in grado di dare la massima rappresentazione della variabilità delle longine negli organismi, perciò lo si è ampliato utilizzando sequenze derivanti da un *clustering* di EST effettuato in collaborazione con il Dr. Rosario Dioguardi. Sono stati definiti dei *consensus* per gli elementi strutturali del LD ed individuate anche le posizioni di elevata conservazione per le tre famiglie longiniche principali. Per le famiglie delle Ykt6 e Sec22b il *consensus* degli elementi strutturali riguarda più di 1/3 delle posizioni e meno di 1/4 per le VAMP7. In particolare le posizioni invariante sono il 12-15% nelle Ykt6 e nelle Sec22, mentre la percentuale scende al 7% nelle VAMP7.

La regione del LD che media l'interazione intramolecolare con lo SNARE *motif* in Ykt6 di lievito è stata dimostrata essere l'alfa elica 1 ( $\alpha 1$ ; Tochio *et al.*, 2001). Analizzando in dettaglio la conservazione di questa regione, essa risulta essere famiglia-specifica ed ha consentito la creazione di tre *signature pattern* che comprendono non solo  $\alpha 1$ , ma anche il foglietto  $\beta 3$  successivo:

### **PATTERN VAMP7 SVILUPPATO SU $\alpha 1$ - $\beta 3$ :**

[ GST ] - [ ND ] - [ FHYAT ] - X - [ EFTNDASG ] - [ VIHFML ] - [ TLAVS ] - X - X -  
 [ IVFLACY ] - [ LIF ] - X - [ KRQDNLS ] - [ ILV ] - [ PGQSTNK ] - X ( 4 , 8 ) -  
 [ KRTNH ] - X - [ TSDIVC ] - [ YLEF ] - X - [ HQYAKDC ]

#### PATTERN Sec22b SVILUPPATO SU $\alpha 1$ - $\beta 3$ :

[DSNQAHY]-[LMIFKASGV]-X-X-[YQHVILA]-[QKS]-[SQNKGLA]-  
[QKRHL]-[AVICF]-K-X-[LVI]-[FVCILTYS]-[RSKCQD]-[KRFPGN]-  
[LMFPSQ]-X(5,7)-[REQL]-[CAMSQ]-[TS]-[LIV]-  
X(0,1)[EGLIT]-[AST]-[GDKNS]

#### PATTERN Ykt6 SVILUPPATO SU $\alpha 1$ - $\beta 3$ :

[FYI]-[QETHKY]-[RK]-[SGPKNQDA]-[STAGN]-[VILYFA]-X-  
[EQD]-[FLYAMH]-[MILFTAV]-X-[FLM]-X-[SATG]-[QAKTER]-  
[LTI]-[ILV]-X-X-R-[STIN]-X-X-[GACTNS]-[TSLQAMV]-[RPK]-  
[ASQE]-[SLDVI]-[VI]-[KRQET]-[EQLH]-[QDKNGSE]

Utilizzando questi tre *pattern* nel *database* di Swiss-Prot si ottiene un indice di precisione quasi del 100% per ciascuno, perciò sembrano essere molto specifici.

Non sono stati creati i *pattern* per le altre due famiglie longiniche Nyv1 e Fitolongine perché le prime sono presenti solo nei Saccaromiceti ed hanno una sequenza molto degenerata, mentre per le Fitolongine, presenti solo nelle piante terrestri, è disponibile solo un numero ridotto di sequenze.

La creazione dei *consensus* e dei *pattern* ha consentito non solo di costruire dei marcatori che potrebbero essere utili per le annotazioni automatiche delle proteine, ma anche di individuare l'importanza di particolari regioni e di alcuni aminoacidi presenti. Ad esempio la fenilalanina in posizione 42 (F42) nell' $\alpha 1$  di Ykt6 di lievito, riconosciuta come residuo importante nell'interazione intramolecolare, è abbastanza conservata nelle Ykt6, mentre non viene mantenuta nelle Sec22b e nelle VAMP7. Nelle VAMP7 risulta invece molto conservata una lisina nell' $\alpha 1$  ed una leucina nella stessa posizione di F42.

#### **Modelli delle longine.**

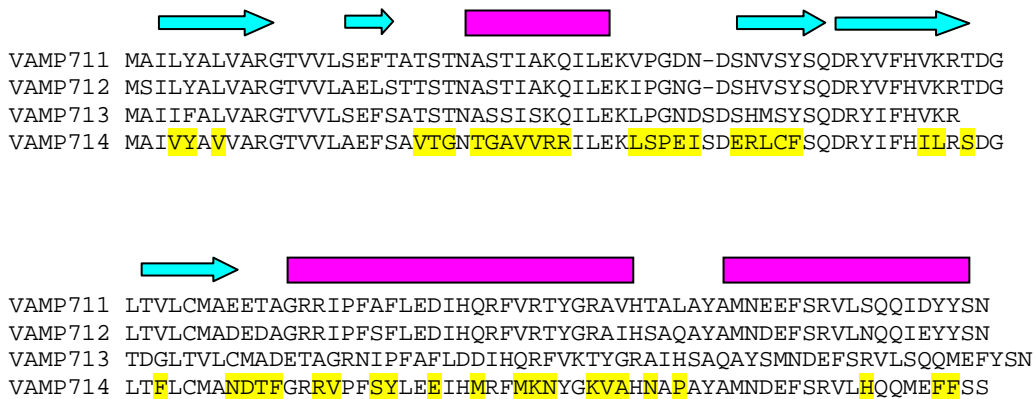
Come descritto nell'introduzione, uno dei ruoli attribuiti al LD è l'intervento nella localizzazione subcellulare delle longine. In particolare la VAMP7 umana localizza nella membrana plasmatica e nell'endosoma (Martinez-Arca *et al.*, 2003); nel *Paramecium tetraurelia*, che possiede 4 VAMP7, tre si localizzano nel RE e una nel vacuolo contrattile (Schilde *et al.*, 2006); in *A. thaliana* sono presenti sette VAMP7 che localizzano in vari compartimenti (Uemura *et al.*, 2004): cinque VAMP7.2 in membrana plasmatica e nell'endosoma tardivo,

VAMP7.2.3 nel RE, VAMP7.2.7 nell'endosoma precoce, tre VAMP7.1 nella membrana vacuolare ed infine VAMP7.1.4 nell'apparato del Golgi.

Coerentemente con la divisione delle VAMP7 nelle piante in due sottofamiglie e con il ruolo del LD, nessuna VAMP7.1 localizza nello stesso compartimento delle VAMP7.2 ma, non essendoci informazioni su quali regioni del LD potessero giustificare queste diverse localizzazioni, si è ritenuto opportuno analizzare in dettaglio non solo le sequenze, ma anche i *patch* di superficie di queste sequenze.

Perciò, seguendo la procedura adottata per le Fitolongine, sono stati creati dei modelli che per omologia si sono basati sulla struttura del LD di VAMP7 umana (essendo questa l'unica struttura di VAMP7 depositata). I modelli, uniti agli allineamenti delle sequenze, hanno permesso di rivelare alcune differenze all'interno dei sottogruppi VAMP7.1 e VAMP7.2.

In particolare, nel confronto tra VAMP7.1.4 (che localizza nel Golgi) ed il gruppo di paraloghi VAMP7.1-2-3 (che condividono la localizzazione nella membrana vacuolare), la maggior parte delle differenze si concentra nella parte iniziale dell' $\alpha 1$ , nella regione tra  $\alpha 1$  e  $\beta 3$  e in buona parte dell' $\alpha 2$  (Figura 13).

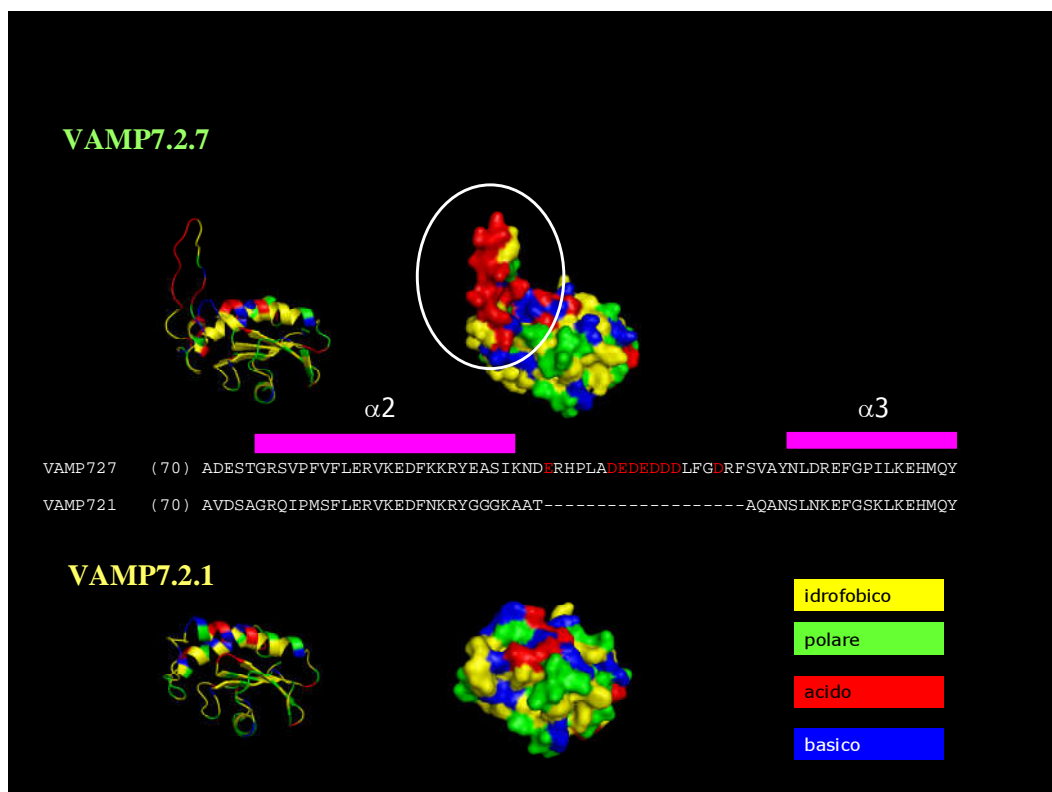


**Figura 13:** allineamento tra le VAMP7.1 di *A. thaliana*. Sono riportate le strutture secondarie: in azzurro i foglietti  $\beta 1-5$ , in viola le  $\alpha$  eliche 1-3. Sono evidenziati in giallo gli aminoacidi differenti in VAMP7.1.4 rispetto alle altre VAMP7.1.

VAMP7.2.7, l'unica presente nell'endosoma precoce, possiede una porzione di sequenza in più che, dall'allineamento con le altre VAMP7.2, risulta compresa tra  $\alpha 2$  e  $\alpha 3$ . Nella creazione del modello, i programmi generavano per quella regione un *loop*. Indagando sulla presenza di questa VAMP7 nelle altre piante, abbiamo

notato che questa porzione di sequenza è conservata: tale conservazione riguarda anche gli aminoacidi, tanto che lo abbiamo definito “loop acido” tra  $\alpha 2$  e  $\alpha 3$  (Figura 14). Non ci sono altre zone rispetto alle altre VAMP7.2 che risultano essere distinte. Queste informazioni ci hanno permesso di ipotizzare che questa zona, che non a caso si distingue dalla zona centrale dell' $\alpha 1$  deputata all'interazione intramolecolare, potrebbe essere fondamentale per l'interazione con altre proteine.

Ad avvalorare queste ipotesi, recentemente è stata riportata la probabile importanza del *loop* tra l' $\alpha 2$  e l' $\alpha 3$  del LD di Sec22b nell'interazione con il subcomplesso Sec23/24 del COPII (un complesso che riveste alcune vescicole), dimostrando come ci siano ulteriori funzioni del dominio longinico nel *trafficking* vescicolare (Mancias & Goldberg, 2007).

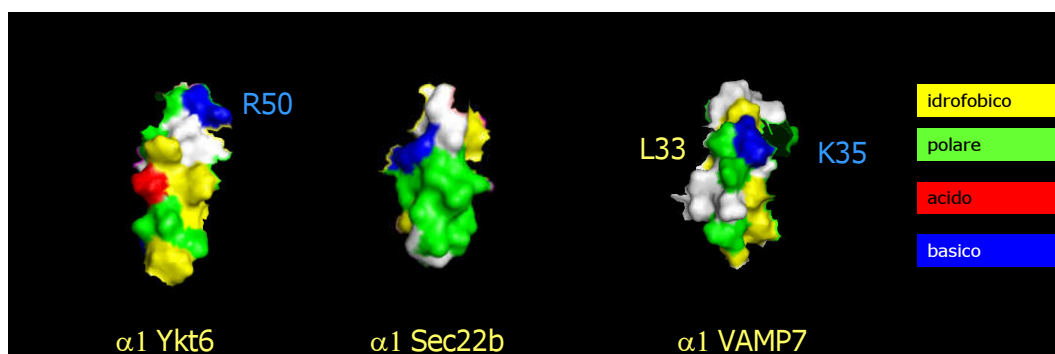


**Figura 14:** Confronto tra le sequenze e le strutture di VAMP7.2.7 e VAMP7.2.1 di *Arabidopsis*. Le superfici sono indicate con il codice a colori riportato, in base alle loro caratteristiche. Il loop acido presente solo in VAMP7.2.7 è cerchiato in bianco.



### Conservazione e variazione del meccanismo di interazione intramolecolare.

La superficie idrofobica presente nell' $\alpha 1$  del LD di Ykt6 di lievito sembra essere fondamentale per la capacità di interazione tra il LD e lo SNARE *motif* (Tochio *et al.*, 2001). Essendo interessati alla comprensione della conservazione di tale meccanismo sia tra le Ykt6 sia esteso alle altre longine, abbiamo utilizzato le strutture depositate del LD delle longine ed analizzato solo la porzione dell' $\alpha 1$ .



**Figura 15:** Visualizzazione tramite PyMOL delle superfici delle  $\alpha$  delle longine principali estrapolate dai modelli degli LD completi con evidenziati gli aminoacidi conservati a più del 95%.

Riportando ed evidenziando sulla struttura gli aminoacidi molto conservati (>95%) e il *consensus* degli altri all'80% (Figura 15), nell' $\alpha 1$  di Ykt6 emerge una regione centrale idrofobica unita alla presenza costante di un residuo di arginina (R50). In Sec22b non ci sono aminoacidi particolarmente conservati e la superficie di  $\alpha 1$  sembra essere fortemente polare. L' $\alpha 1$  di VAMP7 ha una superficie mista, costituita da regioni con aminoacidi idrofobici e aminoacidi polari; inoltre in VAMP7 l'aminoacido corrispondente a F42 di Ykt6 è una leucina molto conservata (L33 riferita alla posizione nella VAMP7 umana) ma sembra essere interna, mentre in superficie è presente una lisina (K35) anch'essa altamente conservata. Queste informazioni sulle diverse caratteristiche della superficie creata da  $\alpha 1$  nelle longine, dimostrano che la presenza di una superficie idrofobica nel LD in Ykt6 di lievito non è alla base del meccanismo generale dell'interazione tra LD e SNARE *motif*. Volendo capire quali potessero essere i determinanti chiave del LD, abbiamo intrapreso la mutagenesi sito-diretta di alcuni residui (la metodica è riportata nella sezione "Materiali e Metodi"), per valutare in analisi di doppio ibrido la variazione specie-specifica del meccanismo di inibizione intramolecolare. Il residuo sul quale abbiamo focalizzato le nostre

analisi in Ykt6 di lievito e umana, è l'arginina 50 (R50). Poiché anche VAMP7 umana mostra conservazione dell'interazione intramolecolare (da nostre evidenze preliminari), abbiamo intrapreso la mutagenesi dei residui di leucina e lisina (L33 e K35) che sono conservati tra le VAMP7, ritenendoli putativi residui chiave nell'interazione.

Per capire il ruolo dell'arginina 50 in Ykt6 abbiamo deciso di creare tre mutanti: il primo che sostituisce l'arginina in acido glutammico (R50E), il secondo mutante con glutammina (R50Q) ed il terzo con alanina (R50A). In base ai risultati dei saggi di doppio ibrido si potrà comprendere innanzitutto se questo aminoacido è in grado di determinare o influenzare l'interazione, ma anche se la sua importanza è dovuta alle caratteristiche di carica o polarità. Con questi primi risultati si potrebbe chiarire se la conservazione del meccanismo di interazione nelle Ykt6 è effettivamente dovuta alla presenza della superficie idrofobica.

Con lo stesso criterio, si è iniziata la creazione dei mutanti di leucina in glutammina o alanina (L33Q e L33A), e lisina in glutammina o alanina (K35Q e K35A) nella VAMP7 umana. Anche in questo caso dai risultati dei saggi di doppio ibrido si potrà elucidare il meccanismo di interazione intramolecolare e determinare se l'importanza di questi aminoacidi così conservati rappresentano la base per l'autoinibizione.

Le analisi di doppio ibrido che verranno prossimamente condotte con i mutanti creati nel corso di questa tesi, si basano sull'*interaction trap* e sfruttano la ricostituzione dell'attivatore trascrizionale LexA. Questo sistema, grazie alla presenza di operatori multipli di lexA a monte di due geni *reporter* (il gene lacZ ed un gene per un'auxotrofia nutrizionale), permette l'identificazione di interazioni anche molto deboli, non facilmente identificabili con i classici sistemi di doppio ibrido.

## CONCLUSIONI E PROSPETTIVE

Le analisi condotte in questa tesi hanno rivelato che le longine, proteine SNARE fondamentali nel traffico vescicolare, sono conservate in tutti gli eucarioti e suddivise in tre famiglie principali (Ykt6, Sec22b e VAMP7), che da analisi filogenetiche sono risultate essere diversamente rappresentate nei vari organismi. L'attenzione è stata rivolta soprattutto a due famiglie longiniche: Ykt6 e VAMP7. La famiglia delle Ykt6 è presente in tutti i supergruppi eucariotici, ha una sequenza molto conservata e si è specializzata diventando essenziale per la vitalità dell'organismo; VAMP7 invece è una longina più "plastica": la sua sequenza risulta meno conservata negli eucarioti rispetto a Ykt6, ha avuto amplificazione genica e possiede ruoli distinti in alcuni eucarioti: in particolare nelle piante (in cui non sono presenti le brevine) tale amplificazione raggiunge livelli molto elevati e le VAMP7 presentano localizzazioni subcellulari specifiche. Rientra nelle plasticità anche il dato che nei Saccaromiceti la longina Nyv1 deriva da e sostituisce VAMP7.

Nell'ambito della ricerca delle tre famiglie longiniche, abbiamo scoperto nelle piante l'esistenza delle cosiddette Fitolongine, ossia proteine che, pur possedendo il LD ed un dominio C-terminale conservato, non possiedono lo SNARE *motif* nella regione centrale: esse quindi rappresentano un prototipo di longina non-SNARE. Da dati di conservazione, amplificazione ed espressione è possibile ipotizzarne un ruolo rilevante nelle piante terrestri. Si cercherà quindi di determinare la localizzazione subcellulare di tali proteine in *A. thaliana* e, attraverso saggi di doppio ibrido, comprendere se esiste un'interazione tra l'LD e la regione non-SNARE.

Poiché in letteratura è riportato un ruolo del LD nel *trafficking*, ed in particolare le VAMP7 di *Arabidopsis* presentano una localizzazione cellulare diversa, sono stati creati dei modelli 3D del loro LD, che hanno permesso di individuare dei *patch* di superficie e putativi aminoacidi chiave. La regione che sembra coinvolta nell'interazione con altre proteine è quella formata da  $\alpha 2$ -loop- $\alpha 3$ . Il quadro che ne emerge è che il LD presenta due superfici opposte con ruoli distinti: quella formata da  $\alpha 1$  è coinvolta nell'interazione intramolecolare, mentre quella formata da  $\alpha 2$ -loop- $\alpha 3$  è coinvolta in interazioni intermolecolari.

Dal recupero di molte sequenze longiniche dai *database*, sono stati creati degli allineamenti che tenevano conto delle strutture secondarie. Su tali allineamenti si è individuata la zona più conservata sulla quale si sono sviluppati i *pattern* specifici per le tre famiglie longiniche principali che sono in grado di riconoscerle e discriminarle. Questa regione, che corrisponde principalmente all' $\alpha 1$ , rende l'interazione intramolecolare famiglia-specifica. Perciò si è intrapresa la creazione di mutanti che verranno utilizzati in saggi di doppio ibrido, per comprendere l'importanza di alcuni residui di  $\alpha 1$  identificati come cruciali per l'interazione tra LD e SNARE *motif* in VAMP7 umana e Ykt6 umana e di lievito.

## BIBLIOGRAFIA

**Abascal, F.; Zardoya, R. & Posada, D.** (2005) ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**: 2104-2105.

**Adl, S. M.; Simpson, A. G. B.; Farmer, M. A.; Andersen, R. A.; Anderson, O. R.; Barta, J. R.; Bowser, S. S.; Brugerolle, G.; Fensome, R. A.; Fredericq, S.; James, T. Y.; Karpov, S.; Kugrens, P.; Krug, J.; Lane, C. E.; Lewis, L. A.; Lodge, J.; Lynn, D. H.; Mann, D. G.; McCourt, R. M.; Mendoza, L.; Moestrup, O.; Mozley-Standridge, S. E.; Nerad, T. A.; Shearer, C. A.; Smirnov, A. V.; Spiegel, F. W. & Taylor, M. F. J. R.** (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J. Eukaryot. Microbiol.* **52**: 399-451.

**Advani, R. J.; Bae, H. R.; Bock, J. B.; Chao, D. S.; Doung, Y. C.; Prekeris, R.; Yoo, J. S. & Scheller, R. H.** (1998) Seven novel mammalian SNARE proteins localize to distinct membrane compartments. *J. Biol. Chem.* **273**: 10317-10324.

**Advani, R. J.; Yang, B.; Prekeris, R.; Lee, K. C.; Klumperman, J. & Scheller, R. H.** (1999) VAMP-7 mediates vesicular transport from endosomes to lysosomes. *J. Cell. Biol.* **146**: 765-776.

**Agatep R., Kirkpatrick R.D., Parchaliuk D.L., Woods R.A. & Gietz R.D.** (1998) Transformation of *Saccharomyces cerevisiae* by the lithium acetate/ single-stranded carrier DNA/polyethylene glycol (LiAc/ssDNA/PEG) protocol. *Technical Tips Online* (<http://www.trends.com>).

**Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller W. & Lipman, D. J.** (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **1997**: 25:3389-3402.

**Antonin, W.; Fasshauer, D.; Becker, S.; Jahn, R. & Schneider, T. R.** (2002) Crystal structure of the endosomal SNARE complex reveals common structural principles of all SNAREs. *Nat. Struct. Bio.* **9**: 107-111.

**Balch, W. E.; Dunphy, W. G.; Braell, W. A. & Rothman, J. E.** (1984) Reconstitution of the transport of protein between successive compartments of the Golgi measured by the coupled incorporation of N-acetylglucosamine. *Cell* **39**: 405-416.

**Barlowe, C.; Orci, L.; Yeung, T.; Hosobuchi, M.; Hamamoto, S.; Salama, N.; Rexach, M.F.; Ravazzola; M., Amherdt, M. & Schekman,R.** (1994) COPII: a membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell* **77**: 895-907.

**Bennett-Lovsey, R.M.; Herbert, A.D.; Sternberg, M.J.E. & Kelley L.A.** (2008) Proteins: Structure, Function. *Bioinformatics* **70**: 611-625.

**Bock, J. B.; Matern, H. T.; Peden, A. A. & Scheller, R. H.** (2001) A genomic perspective on membrane compartment organization. *Nature*. **409**: 839-841.

**Bonifacino, J. S. & Glick, B. S.** (2004) The mechanisms of vesicle budding and fusion. *Cell* **116**: 153-166.

**Bretscher M., Munro S.** (1993) Cholesterol and the Golgi apparatus, *Science* **261**: 1280-1281.

**Brunger, A. T.** Structure and function of SNARE and SNARE-interacting proteins. (2005) *Q. Rev. Biophys.* **38**:1-47.

**Burri, L.; Varlamov, O.; Doege, C. A.; Hofmann, K.; Beilharz, T.; Rothman, J. E.; Söllner, T. H. & Lithgow, T.** (2003) A SNARE required for retrograde transport to the endoplasmic reticulum. *Proc. Natl. Acad. Sci. U S A* **100**: 9873-9877.

**Chen, Y.; Shin, Y. & Bassham, D. C.** (2005) YKT6 is a core constituent of membrane fusion machineries at the *Arabidopsis* trans-Golgi network. *J. Mol. Biol.* **350**: 92-101.

**Clary D.O.; Griff I.C. & Rothman J.E.** (1990) SNAPs, a family of NSF attachment proteins involved in intracellular membrane fusion in animals and yeast. *Cell* **61**: 709-721.

**Clary D.O. & Rothman J.E.** (1990) Purification of three related peripheral membrane proteins needed for vesicular transport. *J. Biol. Chem.* **265**: 10109-10117.

**Coco, S.; Raposo, G.; Martinez, S.; Fontaine, J. J.; Takamori, S.; Zahraoui, A.; Jahn, R.; Matteoli, M.; Louvard, D. & Galli, T.** (1999) Subcellular localization of tetanus neurotoxin-insensitive vesicle-associated membrane protein (VAMP)/VAMP7 in neuronal cells: evidence for a novel membrane compartment. *J. Neurosci.* **19**: 9803-9812.

**Combet, C.; Jambon, M.; Deléage, G. & Geourjon C.** (2002) Geno3D: Automatic comparative molecular modelling of protein. *Bioinformatics.* **18**: 213-214.

**Dacks, J. B. & Field, M. C.** (2007) Evolution of the eukaryotic membrane-trafficking system: origin, tempo and mode. *J. Cell Sci.* **120**: 2977-2985.

**DeLano, W.L.** (2002) The PyMOL Molecular Graphics System <http://www.pymol.org>.

**Damen, E.; Krieger, E.; Nielsen, J. E.; Eygensteyn, J. & van Leeuwen, J. E. M.** (2006) The human Vps29 retromer component is a metallo-phosphoesterase for a cation-independent mannose 6-phosphate receptor substrate peptide. *Biochem. J.* **398**: 399-409.

**Dennison, S. M.; Bowen, M. E.; Brunger, A. T. & Lentz, B. R.** (2006) Neuronal SNAREs do not trigger fusion between synthetic membranes but do promote PEG-mediated membrane fusion. *Biophys. J.* **90**: 1661-1675.

**D'Esposito, M.; Ciccodicola, A.; Gianfrancesco, F.; Esposito, T.; Flagiello, L.; Mazzarella, R.; Schlessinger, D. & D'Urso, M.** (1996) A synaptobrevin-like gene in the Xq28 pseudoautosomal region undergoes X inactivation. *Nat. Genet.* **13**: 227-229.

**Dilcher M., Kohler B. & von Mollard G.F.** (2001) Genetic interactions with the yeast Q-SNARE VTI1 reveal novel functions for the R-SNARE YKT6. *J. Biol. Chem.* **276**: 34537-34544.

**Dietrich, L. E. P.; Gurezka, R.; Veit, M. & Ungermann, C.** (2004) The SNARE Ykt6 mediates protein palmitoylation during an early stage of homotypic vacuole fusion. *EMBO J.* **23**: 45-53.

**Dulubova, I.; Sugita, S.; Hill, S.; Hosaka, M.; Fernandez, I.; Südhof, T. C. & Rizo, J.** (1999) A conformational switch in syntaxin during exocytosis: role of munc18. *EMBO J.* **18**: 4372-4382.

**Fasshauer, D.; Sutton, R. B.; Brunger, A. T. & Jahn, R.** (1998) Conserved structural features of the synaptic fusion complex: SNARE proteins reclassified as Q- and R-SNAREs. *Proc. Natl. Acad. Sci. U S A* **95**: 15781-15786.

**Fasshauer, D.** (2003) Structural insights into the SNARE mechanism. *Biochim Biophys. Acta* **1641**: 87-97.

**Fernandez, I.; Ubach, J.; Dulubova, I.; Zhang, X.; Südhof, T. C. & Rizo, J.** (1998) Three-dimensional structure of an evolutionarily conserved N-terminal domain of syntaxin 1A. *Cell* **94**: 841-849.

**Filippini, F.; Rossi, V.; Galli, T.; Budillon, A.; D'Urso, M. & D'Esposito, M.** (2001) Longins: a new evolutionary conserved VAMP family sharing a novel SNARE domain. *Trends Biochem. Sci.* **26**: 407-409.

**Fuerst, J. A.** (2005) Intracellular compartmentation in planctomycetes. *Annu. Rev. Microbiol.* **59**: 299-328.

**Fukasawa, M.; Varlamov, O.; Eng, W. S.; Söllner, T. H. & Rothman, J. E.** (2004) Localization and activity of the SNARE Ykt6 determined by its regulatory domain and palmitoylation. *Proc. Natl. Acad. Sci. U S A* **101**: 4815-4820.

**Futerman A.H. & Banker G.A.** (1996) The economics of neurite outgrowth: the addition of new membrane to growing axons. *Trends Neurosci.* **19**: 144-149.

**Galli, T.; Zahraoui, A.; Vaidyanathan, V. V.; Raposo, G.; Tian, J. M.; Karin, M.; Niemann, H. & Louvard, D.** (1998) A novel tetanus neurotoxin-insensitive vesicle-associated membrane protein in SNARE complexes of the apical plasma membrane of epithelial cells. *Mol. Biol. Cell.* **9**: 1437-1448.

**Gonzalez, L. C.; Weis, W. I. & Scheller, R. H.** (2001) A novel snare N-terminal domain revealed by the crystal structure of Sec22b. *J. Biol. Chem.* **276**: 24203-24211.

**Guindon, S. & Gascuel, O.** (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**: 696-704.

**Hasegawa, H.; Zinsser, S.; Rhee, Y.; Vik-Mo, E. O.; Davanger, S. & Hay, J. C.** (2003) Mammalian ykt6 is a neuronal SNARE targeted to a specialized compartment by its profilin-like amino terminal domain. *Mol. Biol. Cell.* **14**: 698-720.

**Hasegawa, H.; Yang, Z.; Oltedal, L.; Davanger, S. & Hay, J. C.** (2004) Intramolecular protein-protein and protein-lipid interactions control the conformation and subcellular targeting of neuronal Ykt6. *J. Cell. Sci.* **117**: 4495-4508.

**Hay J.C., Hirling H. & Scheller R.H.** (1996) Mammalian vesicle trafficking proteins of the endoplasmic reticulum and Golgi apparatus. *J. Biol. Chem.* **271**: 5671-5679.

**Hay, J. C.; Chao, D. S.; Kuo, C. S. & Scheller, R. H.** (1997) Protein interactions regulating vesicle transport between the endoplasmic reticulum and Golgi apparatus in mammalian cells. *Cell* **89**: 149-158.

**Hay J.C., Klumperman J., Oorschot V., Steegmaier M., Kuo C.S. & Scheller R.H.** (1998) Localization, dynamics, and protein interactions reveal distinct roles for ER and Golgi SNAREs. *J. Cell. Biol.* **141**: 1489-1502.

**Hay J.C.** (2001) SNARE complex structure and function. *Exp. Cell. Res.* **271**: 10-21.

**Higuchi R.; Krummel B.; & Saiki R.K.** (1988) A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res.* **16**: 7351-7367.

**Ho, S.N.; Hunt H.D.; Horton R.M. & Pullen J.K.** (1989) and Pease LR. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**: 51-59.



**Jahn, R. & Scheller, R. H.** (2006) SNAREs - engines for membrane fusion. *Nat. Rev. Mol. Cell. Biol.* **7**: 631-643.

**Jang, S. B.; Kim, Y.; Cho, Y.; Suh, P.; Kim, K. & Oh, B.** (2002) Crystal structure of SEDL and its implications for a genetic disease spondyloepiphyseal dysplasia tarda. *J. Biol. Chem.* **277**: 49863-49869.

**Kinch, L. N. & Grishin, N. V.** (2006) Longin-like folds identified in CHiPS and DUF254 proteins: Vesicle trafficking complexes conserved in eukaryotic evolution. *Protein Sci.* **15**: 2669-2674.

**Kurzbauer, R.; Teis, D.; de Araujo, M. E. G.; Maurer-Stroh, S.; Eisenhaber, F.; Bourenkov, G. P.; Bartunik, H. D.; Hekman, M.; Rapp, U. R.; Huber, L. A. & Clausen, T.** (2004) Crystal structure of the p14/MP1 scaffolding complex: how a twin couple attaches mitogen-activated protein kinase signaling to late endosomes. *Proc. Natl. Acad. Sci. U S A* **101**: 10984-10989.

**Kweon, Y.; Rothe, A.; Conibear, E. & Stevens, T. H.** (2003) Ykt6p is a multifunctional yeast R-SNARE that is required for multiple membrane transport pathways to the vacuole. *Mol. Biol. Cell.* **14**: 1868-1881.

**Lafont, F.; Verkade, P.; Galli, T.; Wimmer, C.; Louvard, D. & Simons, K.** (1999) Raft association of SNAP receptors acting in apical trafficking in Madin-Darby canine kidney cells. *Proc. Natl. Acad. Sci. U S A* **96**: 3734-3738.

**Lederkremer, G.Z., Cheng, Y., Petre, B.M., Vogan, E., Springer, S., Schekman, R., Walz, T., and Kirchhausen, T.** (2001) Structure of the Sec23p/24p and Sec13p/31p complexes of COPII. *Proc. Natl. Acad. Sci. USA* **98**: 10704-10709.

**Letourneur, F.; Gaynor, E.C.; Hennecke, S.; Demolliere, C.; Duden, R.; Emr, S.D.; Riezman, H. & Cosson, P.** (1994) Coatamer is essential for retrieval of dilysine-tagged proteins to the endoplasmic reticulum. *Cell* **79**: 1199-1207.

**Lippincott-Schwartz, J.** (1993) Bidirectional membrane traffic between the endoplasmic reticulum and Golgi apparatus. *Trends Cell. Biol.* **3**: 81-88.

**López-García, P. & Moreira, D.** (2006) Selective forces for the origin of the eukaryotic nucleus. *Bioessay* **28**: 525-533.

**Lu, J.; Garcia, J.; Dulubova, I.; Südhof, T. C. & Rizo, J.** (2002) Solution structure of the Vam7p PX domain. *Biochemistry* **41**: 5956-5962.

**Lunin, V. V.; Munger, C.; Wagner, J.; Ye, Z.; Cygler, M. & Sacher, M.** (2004) The structure of the MAPK scaffold, MP1, bound to its partner, p14. A complex with a critical role in endosomal map kinase signaling. *J. Biol. Chem.* **279**: 23422-23430.

- Mancias, J. D. & Goldberg, J.** (2007) The transport signal on Sec22 for packaging into COPII-coated vesicles is a conformational epitope. *Mol. Cell.* **26**: 403-414.
- Margittai, M.; Fasshauer, D.; Pabst, S.; Jahn, R. & Langen, R.** (2001) Homo- and heterooligomeric SNARE complexes studied by site-directed spin labeling. *J. Biol. Chem.* **276**: 13169-13177.
- Margittai, M.; Fasshauer, D.; Jahn, R. & Langen, R.** (2003) The Habc domain and the SNARE core complex are connected by a highly flexible linker. *Biochemistry* **42**: 4009-4014.
- Martin, W.** (2005) Archaeobacteria (Archaea) and the origin of the eukaryotic nucleus. *Curr. Opin. Microbiol.* **8**: 630-637.
- Martinez-Arca S., Alberts P., Zahraoui A., Louvard D. & Galli T.** (2000) Role of tetanus neurotoxin insensitive vesicle-associated membrane protein (TI-VAMP) in vesicular transport mediating neurite outgrowth. *J. Cell. Biol.* **149**: 889-900.
- Martinez-Arca, S.; Rudge, R.; Vacca, M.; Raposo, G.; Camonis, J.; Proux Gillardeaux, V.; Daviet, L.; Formstecher, E.; Hamburger, A.; Filippini, F.; D'Esposito, M. & Galli, T.** (2003) A dual mechanism controlling the localization and function of exocytic vSNAREs. *Proc. Natl. Acad. Sci. U S A* **100**: 9011-9016.
- McNew, J. A.; Sogaard, M.; Lampen, N. M.; Machida, S.; Ye, R. R.; Lacomis, L.; Tempst, P.; Rothman, J. E. & Söllner, T. H.** (1997) Ykt6p, a prenylated SNARE essential for endoplasmic reticulum-Golgi transport. *J. Biol. Chem.* **272**: 17776-17783.
- McNew, J. A.; Parlati, F.; Fukuda, R.; Johnston, R. J.; Paz, K.; Paumet, F.; Söllner, T. H. & Rothman, J. E.** (2000) Compartmental specificity of cellular membrane fusion encoded in SNARE proteins. *Nature* **407**: 153-159.
- Misura, K. M.; Scheller, R. H. & Weis, W. I.** (2001a) Self-association of the H3 region of syntaxin 1A. Implications for intermediates in SNARE complex assembly. *J. Biol. Chem.* **276**: 13273-13282.
- Misura, K. M.; Gonzalez, L. C.; May, A. P.; Scheller, R. H. & Weis, W. I.** (2001b) Crystal structure and biophysical properties of a complex between the N-terminal SNARE region of SNAP25 and syntaxin 1a. *J. Biol. Chem.* **276**: 41301-41309.
- Newman, A. P.; Shim, J. & Ferro-Novick, S.** (1990) BET1, BOS1, and SEC22 are members of a group of interacting yeast genes required for transport from the endoplasmic reticulum to the Golgi complex. *Mol. Cell. Biol.* **10**: 3405-3414.
- Novick, P.; Field, C. & Schekman, R.** (1980) Identification of 23 complementation groups required for post-translational events in the yeast secretory pathway. *Cell* **21**: 205-215.

- Paek I., Orci L., Ravazzola M., Erdjument-Bromage H., Amherdt M., Tempst P., Sollner T.H. & Rothman J.E.** (1997) ERS-24, a mammalian v-SNARE implicated in vesicle traffic between the ER and the Golgi. *J. Cell. Biol.* **137**: 1017-1028.
- Page, R. D. M.** (1996). TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12**: 357-358.
- Palade G.** (1975) Intracellular aspects of the process of protein synthesis. *Science* **189**: 347-358.
- Pearse, B.M.** (1975). Coated vesicles from pig brain: purification and biochemical characterization. *J. Mol. Biol.* **97**: 93–98.
- Prochiantz A.** (1995) Neuronal polarity: giving neurons heads and tails. *Neuron* **15**: 743-746.
- Proux-Gillardeaux, V.; Raposo, G.; Irinopoulou, T. & Galli, T.** (2007) Expression of the Longin domain of TI-VAMP impairs lysosomal secretion and epithelial cell migration. *Biol. Cell.* **13**: 177-186.
- Rachel, R.; Wyszchony, I.; Riehl, S. & Huber, H.** (2002) The ultrastructure of Ignicoccus: evidence for a novel outer membrane and for intracellular vesicle budding in an archaeon. *Archaea* **1**: 9-18.
- Ranea, J. A. G.; Sillero, A.; Thornton, J. M. & Orengo, C. A.** (2006) Protein superfamily evolution and the last universal common ancestor (LUCA). *J. Mol. Evol.* **63**: 513-525.
- Ronquist, F. & Huelsenbeck, J. P.** (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572-1574.
- Rossi, V.; Banfield, D. K.; Vacca, M.; Dietrich, L. E. P.; Ungermann, C.; D'Esposito, M.; Galli, T. & Filippini, F.** (2004) Longins and their longin domains: regulated SNAREs and multifunctional SNARE regulators. *Trends Biochem. Sci.* **29**: 682-688.
- Sacher, M.; Stone, S. & Ferro-Novick, S.** (1997) The synaptobrevin-related domains of Bos1p and Sec22p bind to the syntaxin-like region of Sed5p. *J. Biol. Chem.* **272**: 17134-17138.
- Sambrook J., Russel D.W.** (2001a) Preparation of plasmid DNA by alkaline lysis with SDS: Miniprep. In “*Molecular Cloning: a laboratory manual*”, third edition, Irvin, N Editor, Cold Spring Harbor Laboratory Press Publisher, vol 1, section 1.32.

**Sambrook J., Russel D.W.** (2001b). Preparation and transformation of competent *E. coli* using Calcium Chloride. In “*Molecular Cloning: a laboratory manual*”, third edition, Irvin, N Editor, Cold Spring Harbor Laboratory Press Publisher, vol 1, section 1.116.

**Saraste, J. & Kuismanen, E.** (1992) Pathways of protein sorting and membrane traffic between the rough endoplasmic reticulum and the Golgi complex. *Semin. Cell. Biol.* **3**: 343-355.

**Saraste, J. & Goud, B.** (2007) Functional symmetry of endomembranes. *Mol. Biol. Cell.* **18**: 1430-1436.

**Schekman, R & Novick, P.** (2004) 23 genes, 23 years later. *Cell* **116**: S13-5, 1 p following S19.

**Schlenker, O.; Hendricks, A.; Sinning, I. & Wild, K.** (2006) The structure of the mammalian signal recognition particle (SRP) receptor as prototype for the interaction of small GTPases with Longin domains. *J. Biol. Chem.* **281**: 8898-8906.

**Schiestl R.H. & Gietz R.D.** (1989) High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr. Genet.* **16**: 339-346.

**Serafini, T.; Orci, L.; Amherdt, M.; Brunner, M.; Kahn, R.A. & Rothman, J.E.** (1991) ADP-ribosylation factor is a subunit of the coat of Golgi-derived COP-coated vesicles: a novel role for a GTP-binding protein. *Cell*, **67**: 239-253.

**Siddiqi, S. A., Gorelick, F. S., Mahan, J. T. and Mansbach, C. M., II.** (2003). In vitro budding of pre-chylomicron transport vesicles from the endoplasmic reticulum is COPII protein independent. *J. Cell. Sci.* **116**: 415-427.

**Siddiqi, S. A.; Mahan, J.; Siddiqi, S.; Gorelick, F. S. & Mansbach, C. M.** (2006) Vesicle-associated membrane protein 7 is expressed in intestinal ER. *J Cell. Sci.* **119**: 943-950.

**Söllner, T.; Bennett, M. K.; Whiteheart, S. W.; Scheller, R. H. & Rothman, J. E.** (1993) A protein assembly-disassembly pathway in vitro that may correspond to sequential steps of synaptic vesicle docking, activation, and fusion. *Cell* **75**: 409-418.

**Søgaard, M.; Tani, K.; Ye, R. R.; Geromanos, S.; Tempst, P.; Kirchhausen, T.; Rothman, J. E. & Söllner, T.** (1994) A rab protein is required for the assembly of SNARE complexes in the docking of transport vesicles. *Cell* **78**: 937-948.

**Spang A. & Schekman R.** (1998) Reconstitution of retrograde transport from the Golgi to the ER in vitro. *J. Cell. Biol.* **143**: 589-599.

**Stamatakis, A.** (2006) RAxML-VI-HPC: Maximum Likelihood-based Phylogenetic Analyses with Thousands of Taxa and Mixed Models. *Bioinformatics* **22**: 2688-2690.

**Sutton R.B.; Fasshauer D.; Jahn R. & Brunger A.T.** (1998) Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 Å resolution. *Nature* **395**: 347-353.

**Tagaya M.; Wilson D.W.; Brunner M.; Arango N. & Rothman J.E.** (1993) Domain structure of an N-ethylmaleimide-sensitive fusion protein involved in vesicular transport. *J. Biol. Chem.* **268**: 2662-2666.

**Thompson, J. D.; Gibson, T. J.; Plewniak, F.; Jeanmougin, F. & Higgins, D. G.** (1997) The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic. Acids. Res.* **25**: 4876-4882.

**Tochio, H.; Tsui, M. M.; Banfield, D. K. & Zhang, M.** (2001) An autoinhibitory mechanism for nonsyntaxin SNARE proteins revealed by the structure of Ykt6p. *Science* **293**: 698-702.

**Toonen, R. F. G. & Verhage, M.** (2003) Vesicle trafficking: pleasure and pain from SM genes. *Trends. Cell. Biol.* **13**: 177-186.

**Tsui, M. M.; Tai, W. C. & Banfield, D. K.** (2001) Selective formation of Sed5p-containing SNARE complexes is mediated by combinatorial binding interactions. *Mol. Biol. Cell.* **12**: 521-538.

**Tusnàdy, G. E. & Simon, I.** (2001) The HMMTOP transmembrane topology prediction server. *Bioinformatics* **17**: 849-850.

**Uemura, T.; Sato, M. H. & Takeyasu, K.** (2005) The longin domain regulates subcellular targeting of VAMP7 in *Arabidopsis thaliana*. *FEBS Lett.* **579**: 2842-2846.

**Ungermann C., von Mollard G.F., Jensen O.N., Margolis N., Stevens T.H. & Wickner W.** (1999) Three v-SNAREs and two t-SNAREs, present in a pentameric cis-SNARE complex on isolated vacuoles, are essential for homotypic fusion. *J. Cell. Biol.* **145**: 1435-1442.

**Valdez-Taubas, J. & Pelham, H.** (2005) Swf1-dependent palmitoylation of the SNARE Tlg1 prevents its ubiquitination and degradation. *EMBO J.* **24**: 2524-2532.

**van der Giezen, M.; Tovar, J. & Clark, C. G.** (2005) Mitochondrion-derived organelles in protists and fungi. *Int. Rev. Cytol.* **244**: 175-225.

**Veit, M.** (2004) The human SNARE protein Ykt6 mediates its own palmitoylation at C-terminal cysteine residues. *Biochem. J.* **384**: 233-237.

**Ward, D. M.; Pevsner, J.; Scullion, M. A.; Vaughn, M. & Kaplan, J.** (2000) Syntaxin 7 and VAMP-7 are soluble N-ethylmaleimide-sensitive factor attachment protein receptors required for late endosome-lysosome and homotypic lysosome fusion in alveolar macrophages. *Mol. Biol. Cell.* **11**: 2327-2333.

**Weimbs, T.; Low, S. H.; Chapin, S. J.; Mostov, K. E.; Bucher, P. & Hofmann, K.** (1997) A conserved domain is present in different families of vesicular fusion proteins: a new superfamily. *Proc. Natl. Acad. Sci. U S A* **94**: 3046-3051.

**Wen, W.; Chen, L.; Wu, H.; Sun, X.; Zhang, M. & Banfield, D. K.** (2006) Identification of the Yeast R-SNARE Nyv1p as a Novel Longin Domain containing Protein. *Mol. Biol. Cell.* **17**: 4282-4299.

**Weninger, K.; Bowen, M. E.; Chu, S. & Brunger, A. T.** (2003) Single-molecule studies of SNARE complex assembly reveal parallel and antiparallel configurations. *Proc. Natl. Acad. Sci. U S A* **100**: 14800-14805.

**Whelan, S. & Goldman, N.** (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* **18**: 691-699.

**Wilson D.W.; Wilcox C.A., Flynn G.C.; Chen E.; Kuang W.J.; Henzel W.J.; Block M.R.; Ullrich A. & Rothman J.E.** (1989) A fusion protein required for vesicle-mediated transport in both mammalian cells and yeast. *Nature* **339**: 355-359.

**Xu, Y.; Martin, S.; James, D. E. & Hong, W.** (2002) GS15 forms a SNARE complex with syntaxin 5, GS28, and Ykt6 and is implicated in traffic in the early cisternae of the Golgi apparatus. *Mol. Biol. Cell.* **13**: 3493-3507.

**Xu, D., Joglekar, A. P., Williams, A. L., and Hay, J. C.** (2000) Subunit Structure of a Mammalian ER/Golgi SNARE Complex *J. Biol. Chem.* **275**, 39631–39639.

**Zhang, F.; Chen, Y.; Kweon, D.; Kim, C. S. & Shin, Y.** (2002) The four-helix bundle of the neuronal target membrane SNARE complex is neither disordered in the middle nor uncoiled at the C-terminal region. *J. Biol. Chem.* **277**: 24294-24298.

**Zhang T. & Hong W.** (2001) Ykt6 forms a SNARE complex with syntaxin 5, GS28, and Bet1 and participates in a late stage in endoplasmic reticulum-Golgi transport. *J. Biol. Chem.* **276**: 27480-27487.