Università degli Studi della Tuscia di Viterbo

Dipartimento di Agrobiologia e Agrochimica

Sezione di Genetica

Dottorato di ricerca in Genetica Agraria

XIX ciclo

Caratterizzazione strutturale e funzionale di geni codificanti per proteine della famiglia PDI in frumento tenero

AGR07

Coordinatore:

Prof. Enrico Porceddu

Tutori:

Prof. Mario Ciaffi

Prof. Oronzo Antonio Tanzarella

Dottoranda:

Elisa Laura d'Aloisio

INDICE

1. INTRODUZIONE 1
1.1. Il frumento 1
1.1.1. Le proteine di riserva del frumento 2
1.1.2. Il glutine 6
1.1.3. Sintesi e deposizione delle proteine di riserva 8
1.2. La PDI 13
1.3. La struttura della PDI 14
1.4. La famiglia della PDI nell'uomo 18
1.5. Funzioni della PDI e delle proteine PDI-like 24
1.5.1. Attività redox isomerasica 25
1.5.1.1.La superfamiglia della tioredossina 25
1.5.1.2.Catalisi dell'attività redox 27
1.5.1.3.Domini e funzioni della PDI 29
1.5.1.4.L'origine degli equivalenti ossidativi 30
1.5.1.5.Il ruolo del glutatione 32
1.5.1.6.Il ruolo della PDI in vivo 33
1.5.1.7.Gli omologhi della PDI e le vie biosintetiche alternative 36
1.5.2. Attività di chaperone molecolare 38
1.5.3. Il ruolo della PDI come subunità di complessi enzimatici 40
1.5.4. PDI e proteine sue omologhe in compartimenti cellulari diversi dal RE 41
1.5.4.1.Localizzazione extracellulare 41
1.5.4.2.Localizzazione e funzione delle proteine PDI-like nel citoplasma 43
1.5.4.3.Localizzazione e funzione delle proteine PDI-like nel nucleo 44
1.5.4.4.Localizzazione di una proteina omologa della PDI nel cloroplasto 45
1.6. La famiglia delle PDI nelle piante 46
1.7. Il ruolo della PDI nelle piante 52
1.8. La PDI in frumento 55
2. SCOPO DELLA TESI 64
3. ARTICOLAZIONE DELLE ATTIVITA' 66
4. MATERIALI E METODI 67
4.1. Materiale vegetale 67

4.2. Preparazione degli acidi nucleici 68
4.2.1. Estrazione di DNA nucleare 68
4.2.2. Estrazione di DNA con CTAB 69
4.2.3. Estrazione di RNA totale 70
4.3. Identificazione delle sequenze in banche dati 72
4.4. Analisi RT-PCR 73
4.5. Isolamento di sequenze cDNA "Full Length" mediante RACE 74
4.5.1. Defosforilazione 75
4.5.2. Rimozione del cap dall'estremità 5' degli mRNA completi 76
4.5.3. Reazione di ligasi del GeneRacer™ Oligo agli mRNA completi privati del cap all'estremità 5' 76
4.5.4. Trascrittasi inversa 76
4.5.5. Amplificazione mediante PCR delle estremità 5' e 3' delle sequenze cDNA 77
4.6. Amplificazione PCR del DNA genomico e del cDNA 78
4.7. Clonaggio dei prodotti di amplificazione 81
4.8. Analisi delle sequenze nucleotidiche 82
4.9. Analisi filogenetica 82
4.10. Analisi Southern 84
4.10.1. Preparazione dei filtri di ibridazione (Southern blot) 84
4.10.2. Marcatura delle sonde 84
4.10.3. Ibridazione molecolare 86
4.11. Analisi Northern 87
4.12. Ottenimento dei costrutti molecolari 88
4.12.1. Preparazione ed analisi dei plasmidi donatori delle sequenze di interesse 88
4.12.2. Amplificazione degli inserti con primer adattatori 88
4.12.3. Ottenimento dei frammenti di restrizione dai plasmidi e dagli amplificati 90
4.12.4. Ligasi e trasformazione 90
4.13. Trasformazione biolistica 92
4.13.1. Raccolta e sterilizzazione delle cariossidi immature 92
4.13.2. Isolamento e trattamento degli scutelli estratti da cariossidi immature 92
4.13.3. Bombardamento 93
4.13.4. Induzione dei calli 94
4.13.5. Rigenerazione e selezione in vitro dei tessuti 94

4.14. Identificazione delle piante contenenti il costrutto per la sovra-espressione mediante PCR 95
5. RISULTATI E DISCUSSIONE 97
5.1. Identificazione e clonazione di sequenze cDNA codificanti per proteine PDI-like in frumento 97
5.2. Analisi filogenetica 98
5.3. Struttura delle sequenze cDNA codificanti per proteine PDI-like in frumento ... 103
5.4. Isolamento delle sequenze genomiche mediante PCR 110
5.5. Struttura dei geni PDI-like 113
5.6. Localizzazione cromosomica delle sequenze dei geni *TaPDIL4-1* e *TaPDIL5-1* . 120
5.7. Analisi Northern dei trascritti relativi ai geni *TaPDIL4-1*, *TaPDIL5-1* e *TaPDIL2-1* 123
5.8. Ottenimento del costrutto per la sovra-espressione della PDI classica 127
5.9. Ottenimento dei costrutti per lo studio funzionale del promotore del gene della PDI classica 132
5.9.1. Ruolo funzionale dei motivi regolatori individuati nel promotore di TaPDIL1-1 133
5.9.2. Analisi funzionale dei frammenti di delezione del promotore di TaPDIL1-1.. 136
5.10. Co-trasformazione della cultivar MPB-Bobwhite 26 per la sovra-espressione di TaPDIL1-1 e per l'analisi funzionale del suo promotore 140
5.11. Identificazione mediante PCR delle piante contenenti i costrutti per la sovra-espressione della PDI 144
6. CONCLUSIONI 147
7. PROSPETTIVE FUTURE 151
8. BIBLIOGRAFIA 153
9. RINGRAZIAMENTI 169

1. INTRODUZIONE

1.1. Il frumento

Sono numerose le specie coltivate di cereali, tuttavia tre di esse, frumento, riso e mais, forniscono più del 70% della produzione cerealicola totale. Il contenuto proteico dei cereali è minore di quello dei legumi, tuttavia essi producono più di 200 milioni di tonnellate di proteine per la nutrizione umana ed animale. Il frumento, in particolare, è alla base dell'alimentazione di due miliardi di persone, più di un terzo della popolazione mondiale, ed è coltivato su una superficie che supera 217 milioni di ettari e va dalla Scandinavia all'Argentina, con una produzione annuale di più di 620 milioni di tonnellate di granella (Ciaffi et al., 2005). L'importanza alimentare del frumento deriva dalla sua adattabilità ambientale, dal valore nutrizionale e dalla facilità di conservazione delle cariossidi e dei prodotti derivati, ma anche dalla possibilità, mediante trasformazione, di dare origine ad un'ampia varietà di prodotti di largo consumo. Le farine prodotte dalle cariossidi di frumento, infatti, vengono consumate prevalentemente previa trasformazione in pane, pasta ed altri prodotti ricchi in fibre, carboidrati, lipidi, proteine, vitamine ed elementi minerali quali Fe, Ca, K, Zn, P e Mg. Il frumento è il cereale più ricco in proteine, anche se non è il migliore per il bilanciamento tra gli aminoacidi, essendo carente in quelli basici, quali lisina, treonina e triptofano, essenziali per l'alimentazione umana.

In aggiunta alla loro importanza nutrizionale, le proteine di riserva del frumento determinano le proprietà caratteristiche ed uniche delle sue farine nei processi di pastificazione e panificazione e, quindi, la qualità del prodotto finito ottenuto. È nella fase di idratazione della farina di frumento che queste proteine interagiscono e vanno a formare il glutine, una matrice visco-elastica in grado di inglobare granuli di amido e lipidi, da cui dipenderanno le proprietà funzionali degli impasti.

Una migliore comprensione della struttura di queste proteine, delle loro proprietà biofisiche e funzionali, nonché dei meccanismi biologici che regolano la sintesi, il trasporto e la deposizione nelle cariossidi, è alla base di qualsiasi tentativo di migliorare la qualità tecnologica del frumento.

1.1.1. Le proteine di riserva del frumento

Le proteine della cariosside si possono classificare in tre gruppi: proteine metaboliche e strutturali, di riserva e di difesa. Il 10-15% dell'endosperma delle cariossidi mature di frumento è costituito da proteine. Una parte di queste proteine svolge funzioni metaboliche e strutturali durante lo sviluppo del seme, un'altra, costituita dalle proteine di riserva, funge da fonte di aminoacidi e di energia per l'embrione durante la germinazione. Le proteine di riserva vengono sintetizzate durante lo sviluppo dei semi e rappresentano l'80-90% delle proteine totali nelle cariossidi mature. Le principali proteine di riserva del frumento appartengono alla classe delle prolammine, mentre solo una piccola parte è rappresentata dalla classe delle globuline. Le globuline di riserva contenute nell'endosperma di frumento sono denominate "triticine" e rappresentano circa il 5% delle proteine totali del seme (Singh et al., 1988). Ricchi in globuline di riserva sono l'embrione e lo strato aleuronico esterno dell'endosperma, ma l'insieme di questi tessuti, che vengono comunemente rimossi con la macinazione, costituisce solo il 10% circa del peso secco delle cariossidi (Shewry e Halford, 2002).

La divisione delle prolammine di frumento in gliadine, solubili in soluzioni idroalcoliche, e glutenine, solubili in soluzioni diluite di acidi ed alcali, basata sulla classificazione di Osborne (1924), è stata spesso criticata, ma viene tuttora comunemente usata in quanto risulta molto utile da un punto di vista funzionale e tecnologico. Questa classificazione si basa, infatti, sulla diversa solubilità delle componenti proteiche, che, a sua volta, è funzione della loro capacità di polimerizzazione mediante la formazione di ponti disolfuro intermolecolari, caratteristica che, come sarà spiegato in seguito, è alla base delle proprietà reologiche del glutine. Le glutenine, in presenza di agenti riducenti, sono estraibili in soluzioni idroalcoliche.

Le gliadine sono molecole monomeriche capaci di associarsi solamente mediante legami idrogeno ed interazioni idrofobiche, i legami disolfuro tra cisteine, infatti, sono assenti o sono saturati dalle interazioni intramolecolari. Le gliadine possono essere separate mediante elettroforesi su gel di poliacrilammide in tampone acido (A-PAGE) e comprendono quattro gruppi: α, β, γ ed ω-gliadine, in ordine decrescente di mobilità (Woychik et al., 1961). In base alla loro sequenza N-terminale esse possono essere distinte in tre tipi, α, γ ed ω-gliadine. Le ω-gliadine hanno peso molecolare compreso tra 44 e 78 kDa, non contengono aminoacidi solforati, mentre sono ricche di acido glutammico, glutammina, prolina e fenilalanina (circa 80% dei residui totali). Le α- e γ-gliadine contengono invece aminoacidi solforati (cisteina e

metionina), coinvolti solamente nella formazione di legami disolfuro intramolecolari, e minor quantità di acido glutammico, prolina e fenilalanina (60-70% dei residui totali). Queste due classi proteiche hanno un peso molecolare compreso tra 32 e 42 kDa.

Le glutenine formano, invece, aggregati proteici ad elevato peso molecolare costituiti da diverse catene polipeptidiche tenute insieme da legami disolfuro intermolecolari, le quali, dopo riduzione e separazione in SDS-PAGE, vengono distinte in subunità ad elevato (HMW-GS, 80-120 kDa) ed a basso peso molecolare (LMW-GS, 40-55 kDa). Le subunità ad elevato peso molecolare, oltre che per il peso molecolare, differiscono dalle gliadine per un maggior contenuto di glicina ed un più basso contenuto in prolina. Le subunità a basso peso molecolare sono simili alle α– e γ–gliadine per la composizione aminoacidica ed il peso molecolare, ma sono, come si è detto, in grado di formare ponti disolfuro intermolecolari e di legarsi alle HMW-GS, formando aggregati insolubili in soluzioni idroalcoliche. Le glutenine sono state classificate, in base alla loro mobilità elettroforetica in SDS-PAGE e al loro punto isoelettrico, in quattro gruppi denominati A, B, C e D (Payne e Corfield, 1979; Jackson et al., 1985). Il primo gruppo comprende tutte le HMW-GS, mentre le subunità B costituiscono il gruppo principale delle LMW-GS, sia per numero sia per quantità, e sono le componenti più basiche tra le proteine di riserva. Studi sulle sequenze N-terminali hanno indicato che il gruppo C è costituito prevalentemente da proteine simili alla α- e γ-gliadine (Lew et al., 1992) mentre le subunità appartenenti al gruppo D sono strettamente correlate alle ω-gliadine (Masci et al., 1993). Al contrario il gruppo B sembra essere strutturalmente distinto, sulla base delle sequenze N-terminali, rispetto alle altre proteine che costituiscono il glutine.

Questa classificazione tradizionale, come già accennato, è molto utile, da un punto di vista funzionale e tecnologico, in quanto le glutenine e le gliadine appaiono responsabili delle diverse proprietà funzionali degli impasti. Le glutenine e i relativi legami disolfuro sono considerati, infatti, responsabili dell'elasticità dell'impasto e del glutine, mentre le gliadine, dotate solo di interazioni non covalenti, principalmente legami idrogeno e interazioni idrofobiche, sono apparentemente responsabili della viscosità ed estensibilità.

Una classificazione più accurata si basa invece sull'allineamento, in base all'omologia di sequenza, di singole proteine del glutine con proteine simili isolate in orzo e segale (Shewry et al., 1986). Questa nuova classificazione assegna tutte le prolammine delle Triticeae (frumento, orzo e segale) a tre grandi gruppi: prolammine ricche in zolfo (S-rich), prolammine povere in zolfo (S-poor) e prolammine ad alto peso molecolare (HMW), con diverse sottoclassi nel primo gruppo (Tab. 1.1). Come si può vedere nella tabella 1.1, non c'è una corrispondenza diretta tra questa classificazione e le frazioni monomeriche e polimeriche

di frumento (rispettivamente gliadine e glutenine), in quanto sia alle prolammine ricche in zolfo che a quelle povere in zolfo appartengono forme monomeriche e polimeriche. In Fig. 1.1 sono riassunte le strutture di tre tipiche prolammine di frumento. Tutte le prolammine, sia le HMW che le S-rich e le S-poor, contengono sequenze ripetute basate su motivi ricchi in prolina e glutammina, tali motivi sono chiaramente correlati tra le S-rich e le S-poor. Anche nei domini non ripetitivi delle S-rich e delle HMW si riscontra similitudine di sequenza, in particolare nelle posizioni delle cisteine conservate e dei residui aminoacidici adiacenti. L'allineamento di sequenza chiaramente dimostra che i tre tipi di prolammine, S-rich, S-poor e HMW, hanno un'origine evolutiva comune (Shewry e Halford, 2002).

Tab. 1.1 - Classificazione delle prolammine in frumento e loro caratteristiche.

Componenti	M_r (kDa) (% total)	Polimeriche o monomeriche	Composizione aminoacidica parziale (mol%)
Prolammine HMW			
Glutenine HMW[a]	65–90 (6–10%)	Polimeriche	30–35% Gly, 10–16% Pro, 15–20% Gly, 0.5–1.5% Cys, 0.7–1.4% Lys
Prolammine S-rich			
γ-gliadine	30–45 (70–80%)	Monomeriche	30–40% Gln, 15–20% Pro, 2–3% Cys, <1.0% Lys
α-gliadine		Monomeriche	
Glutenine LMW (tipo B e C)[a, b]		Polimeriche	
Prolammine S-poor			
ω-gliadins	30–75 (10–20%)	Monomeriche	40–50% Gln, 20–30% Pro, 8–9% Phe, 0 –0.5% Lys, 0–<0.5% Cys[c]
Glutenine LMW (tipo D)[a, b]		Polimeriche	

[a] = Subunità
[b] = Le subunità LMW di tipo C sono sostanzialmente forme polimeriche di α- e γ-gliadine, mentre le subunità LMW di tipo D sono forme polimeriche delle ω-gliadine. Le subunità LMW di tipo B costituiscono un gruppo distinto di prolammine ricche in zolfo.
[c] = Cys è presente nelle subunità LMW di tipo D, ma non nelle ω-gliadine (da Shewry e Halford, 2002).

Le HMW-GS, come si può vedere in tabella 1.1, formano un gruppo autonomo, definito prolammine ad alto peso molecolare. Nei frumenti esaploidi ci sono 6 geni codificanti per subunità ad alto peso molecolare. Su ogni cromosoma omologo del gruppo 1 (1A, 1B e 1D) ci sono due geni codificanti per due differenti tipi di subunità chiamate x e y. Alcuni di questi geni possono essere silenti e le diverse cultivar di frumento tenero possono presentare da 3 a 5 subunità ad alto peso molecolare. Le subunità 1Dx, 1Dy e 1Bx sono presenti in tutte le cultivar, mentre 1By e/o 1Ax sono presenti solo in alcune (Shewry e Halford, 2002). Il gene codificante per la subunità 1Ay, infine, è silente in tutte le varietà di frumento tenero. Generalmente le subunità di tipo y contengono 6 cisteine, con un elevato livello di conservazione, tutte ubicate nelle zone non ripetitive, 5 nella regione N terminale e 1 nella C terminale e le subunità x conservano 4 di queste cisteine, ma esistono varianti alleliche in cui sono presenti singoli residui addizionali (Shewry e Tatham, 1997).

Alle prolammine ricche in zolfo appartengono invece sia le α-, β- e γ-gliadine sia le LMW-GS di tipo B e C. Tutte le prolammine ricche in zolfo hanno una struttura simile, in cui

è possibile identificare due domini principali: il dominio N terminale, che copre da un terzo a metà della proteina e consiste per lo più di sequenze peptidiche ripetute ricche in prolina e glutammina, e il dominio C terminale, non ripetitivo (Shewry e Tatham, 1997). Questo ultimo dominio è caratterizzato da 6 (α- e β- gliadine) e 8 (γ-gliadine) residui cisteinici conservati, probabilmente coinvolti nella formazione di legami disolfuro intramolecolari che stabilizzano la struttura proteica (Shewry e Tatham, 1997). Le LMW appartenenti a questo gruppo presentano un numero variabile di questi residui cisteinici conservati e contengono residui addizionali responsabili dei legami disolfuro intermolecolari.

Alle prolammine povere in zolfo appartengono, infine, le ω-gliadine, che non presentano residui cisteinici, e le LMW-GS di tipo D. Queste ultime vengono considerate forme mutate delle ω-gliadine, in cui un singolo residuo cisteinico permette l'interazione di queste proteine con il polimero gluteninico.

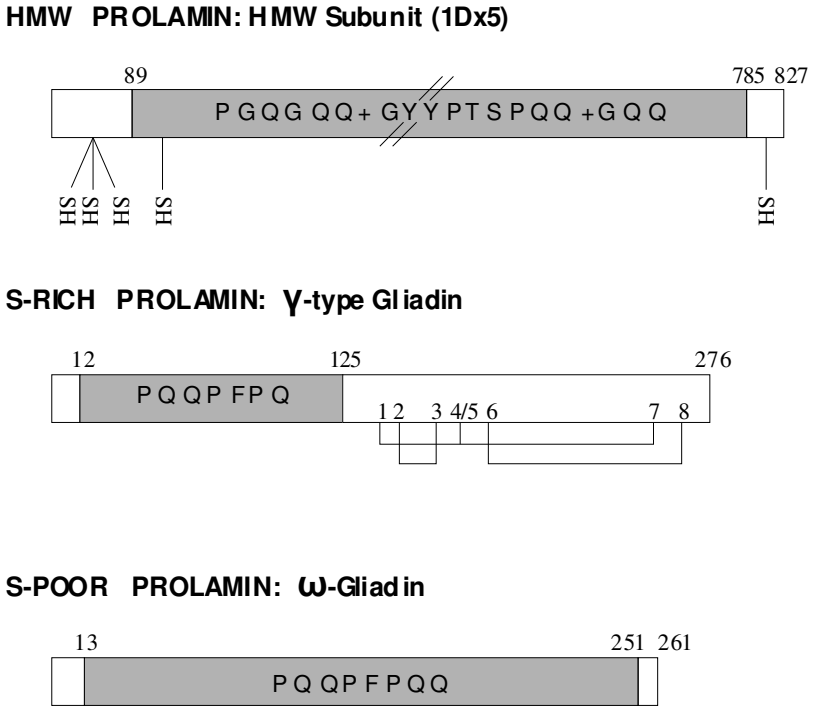

Fig. 1.1 - Rappresentazione schematica della struttura delle tipiche prolammine ad alto peso molecolare (HMW), ricche in zolfo (S-rich) e povere in zolfo (S-poor). Le sequenze ripetute sono evidenziate in grigio, i legami disolfuro tra i residui cisteinici conservati (1-8) nelle γ-gliadine indicati da linee. Con SH vengono indicate le posizioni dei residui cisteinici nelle prolammine ad alto peso molecolare (da Shewry e Halford, 2002).

1.1.2. Il glutine

La maggior parte del frumento destinato all'alimentazione umana viene consumato come prodotti alimentari derivati dalla farina. Durante la fase di macinazione vengono rimossi sia il germe (embrione) che la crusca (pericarpo, testa, strato nucleare ed aleuronico), lasciando solo le cellule dell'endosperma ricche in amido e glutine. Le proteine del glutine formano una matrice continua nelle cellule dell'endosperma mature. Quando la farina viene mescolata ad acqua per formare l'impasto le matrici proteiche delle singole cellule vengono portate in contatto a formare un network continuo capace di conferire all'impasto quelle proprietà reologiche che lo rendono adatto alla panificazione e alla pastificazione.

Le subunità gluteniniche ad alto peso molecolare in particolare formerebbero una struttura polimerica complessa, stabilizzata da legami disolfuro (Shewry e Tatham, 1997) e legami idrogeno intermolecolari (Belton, 1999), e responsabile dello scheletro elastico di base, con il quale le subunità gluteniniche a basso peso molecolare e le gliadine possono interagire con ponti disolfuro intermolecolari e tramite interazioni non covalenti.

Le quantità di proteine monomeriche e polimeriche, in particolare la distribuzione dei pesi molecolari dei polimeri del glutine, sono i fattori che maggiormente influiscono sul grado di tenacità e elasticità degli impasti (Orth e Buishuk,1972; Huebner e Bietz, 1972; Field et al., 1983b; Gupta et al., 1993; Weegels et al,1996). La distribuzione del peso molecolare MWD (molecular weight distribution) delle proteine del glutine è un ottimo parametro per descrivere la correlazione tra proprietà della farina e composizione proteica. Spostando questo parametro a favore di più alti pesi molecolari si migliora la tenacia dell'impasto, fino ad un punto in cui, superati certi valori, si ha un effetto deleterio sulla sua estensibilità. Le proprietà elastiche del glutine, che determinano la forza e la stabilità degli impasti, sono attribuite alla frazione polimerica gluteninica e alla presenza di ponti disolfuro e tra il contenuto di polimeri gluteninici insolubili e la tenacità dell'impasto esiste una correlazione positiva (Field et al.,1983b; MacRitchie, 1987). Le forme monomeriche, interagenti con i polimeri del glutine solo mediante legami non covalenti, contribuiscono invece alla viscosità e all'estensibilità dell'impasto.

Il grado di polimerizzazione delle proteine del glutine va da semplici dimeri, con pesi molecolari di 40-150 kDa, fino a grandi polimeri, le cui dimensioni possono arrivare fino a migliaia di kDa (Shewry et al., 1992, Wrigley, 1996, Wahlund et al., 1996, Stevenson e Preston, 1996). La maggior parte delle subunità gluteniniche hanno almeno due residui

cisteinici nelle regioni terminali della proteina, disponibili per legami disolfuro inter-molecolari, e partecipano quindi alla formazione dei polimeri gluteninici.

La qualità delle farine è in particolare influenzata dal numero di subunità gluteniniche ad alto peso molecolare espresse. Come già accennato in precedenza, questo numero può variare, a seconda della cultivar, da un minimo di 3 ad un massimo di 5 subunità. Le cultivar che presentano 5 subunità gluteniniche HMW espresse hanno una maggiore quantità di aggregati ad elevato peso molecolare rispetto alle cultivar con tre o quattro subunità (Halford et al., 1992), evidenziando, quindi, qualità panificatorie superiori. Oltre a questo effetto quantitativo esiste, inoltre, un effetto qualitativo, funzione delle particolari varianti alleliche per le subunità ad elevato peso molecolare presenti nelle diverse cultivar. Le varianti alleliche dei geni che codificano per le proteine ad elevato peso molecolare influenzano, infatti, i valori della distribuzione del peso molecolare delle proteine del glutine. Payne e Lawrence (1983) hanno analizzato circa 300 cultivar di frumento tenero e hanno individuato diverse forme alleliche ai loci dei geni codificanti per le subunità ad elevato peso molecolare e hanno assegnato diversi punteggi sulla base del loro effetto sulle proprietà reologiche degli impasti (Payne et al., 1979 e 1981).

Per il locus *Glu-D1*, ad esempio, sono state identificate due varianti alleliche designate 5+10 e 2 +12. Dallo studio delle sequenze aminoacidiche dedotte è emerso che ci sono differenze strutturali tra i geni codificanti per le subunità 2 e 5 e tra quelli codificanti per le subunità 12 e 10, in particolare, che le subunità 5 e 2 differiscono per una cisteina addizionale presente, nella subunità 5, in prossimità del dominio N-terminale. La presenza di questa cisteina potrebbe avere ripercussioni sulla quantità e dimensione dei polimeri gluteninici (Lafiandra et al., 1993), giustificando la migliore qualità panificatoria della varietà contenente la variante allelica 5.

Proteine quali le prolammine a baso peso molecolare di tipo D, che possiedono un solo residuo cisteinico, hanno invece un effetto opposto sulla distribuzione dei pesi molecolari, in quanto fungono da terminatori, avendo a disposizione un solo residuo cisteinico disponibile per legami disolfuro intermolecolari.

Sulle proprietà funzionali del glutine influiscono anche fattori ambientali, ad esempio l'innalzamento della temperatura atmosferica ha un effetto negativo sulla qualità delle farine, apparentemente dovuto all'alterata capacità di formare polimeri di maggiori dimensioni (Ciaffi et al., 1996). Anche le pratiche agronomiche che modificano la disponibilità di zolfo nel terreno si ripercuotono sulla capacità della pianta di sintetizzare o meno aminoacidi

solforati (Fullington et al., 1987; MacRitchie e Gupta, 1993) e sull'ottenimento di farina di elevata o bassa qualità.

Tutti i fattori che favoriscono la formazione di aggregati di maggiori dimensioni hanno, pertanto un'influenza positiva sulla qualità delle farine, non ultima la capacità di formazione dei legami disolfuro intermolecolari durante il processo di deposizione delle proteine del glutine nelle cariossidi. L'importanza dei legami disolfuro nel determinare le proprietà funzionali e strutturali del glutine è stata confermata anche da numerosi esperimenti condotti *in vitro,* attraverso l'aggiunta di agenti riducenti e ossidanti al glutine. La riduzione dei ponti disolfuro compromette la tenacità degli impasti, come conseguenza della diminuzione del valore medio dei pesi molecolari. Gli agenti ossidanti, al contrario, portando alla formazione di un maggior numero di legami disolfuro, contribuiscono ad aumentare il valore medio dei pesi molecolari e della tenacità.

1.1.3. Sintesi e deposizione delle proteine di riserva

La sintesi di tutte le proteine di riserva, che nel frumento sono principalmente prolammine, ha luogo nelle cellule dell'endosperma delle cariossidi, iniziando pochi giorni dopo la fioritura (Greene et al., 1985) e protraendosi per tutto il periodo di granigione. Il massimo di espressione dei trascritti si raggiunge durante la fase intermedia della formazione della cariosside, in seguito declina progressivamente fino alla maturità dei semi, circa 35 giorni dopo l'antesi. L'espressione delle diverse classi proteiche risulta quindi altamente coordinata (Greene et al., 1985; Reeves et al., 1986; Grinwade et al., 1996), anche se l'accumulo delle differenti frazioni avviene in maniera sfalsata. La sintesi delle gliadine monomeriche prevale, infatti, su quello delle glutenine polimeriche fino a circa 30 giorni dopo la fioritura, mentre per le restanti due settimane si verifica il contrario (Gupta et al., 1996; Ciaffi et al., 1998). Nell'ambito delle proteine polimeriche totali è stata notata una sostanziale differenza per le quantità relative delle frazioni solubili (aggregati di piccole dimensioni) ed insolubili (aggregati a più elevato peso molecolare) che le compongono. Infatti, mentre le polimeriche solubili compaiono in quantità relativamente elevata già dopo 15 giorni dall'antesi ed hanno un incremento continuo fino al 30°-35° giorno, la formazione degli aggregati a più elevato peso molecolare risulta particolarmente evidente solamente dopo 25-30 giorni dalla fioritura ed il loro accumulo continua anche durante le ultime fasi della granigione, con un massimo tra il 30° e il 40° giorno dopo l'antesi (Gupta et al., 1996; Ciaffi et al., 1998).

La sintesi delle proteine di riserva avviene su polisomi aderenti al lato citosolico della membrana del reticolo endoplasmatico rugoso (Shotwell e Larkins, 1989). Il reticolo endoplasmatico è un comparto subcellulare costituito da un sistema di tubuli e cisterne delimitati da membrana, funzionalmente collegato, attraverso una rete di vescicole, con il complesso del Golgi, la membrana plasmatica ed i vacuoli, a formare il cosiddetto sistema di endomembrane. La catena polipeptidica nascente viene traslocata attraverso la membrana del RE, nell'ambiente luminale. Questo processo di smistamento delle proteine dal citoplasma verso l'interno del lume del RE è determinato dalla presenza di un particolare segnale peptidico all'estremità N-terminale della proteina neosintetizzata. Questo segnale consiste di una sequenza di lunghezza variabile da 20 a 50 aminoacidi, caratterizzata dalla presenza di una regione carica positivamente, seguita da una sequenza costituita da almeno 6 residui idrofobici consecutivi. Il segnale peptidico viene rimosso proteoliticamente, rimanendo avvolto nella membrana, mentre il resto del polipeptide si sposta nel lume del reticolo endoplasmatico. La rimozione del segnale peptidico è essenziale affinché avvenga la traslocazione delle proteine all'interno del lume ed il successivo ripiegamento "folding" della catena polipeptidica (Coleman et al., 1995; Gillikin et al., 1997). Per la traslocazione è inoltre necessario che la catena polipeptidica si trovi in una conformazione non ripiegata, dovuta all'attività di proteine dispieganti, quali HSP70 e BiP, che, legandosi alla catena polipeptidica nascente, ne impediscono il ripiegamento (Miernykk et al., 1992; Panzer et al., 1995).

Nel lume del RE la struttura delle proteine viene modificata in maniera cotraduzionale e post-traduzionale attraverso reazioni di tipo covalente (N-glicosilazione, formazione di legami disolfuro, ecc.) e attraverso la modificazione della struttura tridimensionale, che consiste nel ripiegamento della catena polipeptidica, seguito dall'assemblaggio di aggregati oligomerici nelle proteine multimeriche. La catena polipeptidica contiene nella sua struttura primaria tutte le informazioni necessarie per il suo corretto ripiegamento o folding. L'ottenimento della corretta conformazione ripiegata, tuttavia, pur essendo favorita dal punto di vista termodinamico, non lo è necessariamente anche da un punto di vista cinetico e, in esperimenti condotti *in vitro*, il processo è sempre accompagnato da un elevato numero di proteine con errori di ripiegamento. Le cellule eucariotiche hanno ovviato a questo problema con un ambiente redox altamente specializzato, il reticolo endoplasmatico (Frand et al., 2000), dotato di diversi enzimi luminali, tra cui le chaperonine molecolari della famiglia delle HSP70/BiP (Boston et al., 1991) e la proteina disolfuro isomerasi o PDI (Roden et al., 1982; Bulleid e Freedman, 1988a; Shimoni et al., 1995a,b), responsabile della formazione ed isomerizzazione dei ponti disolfuro. La formazione dei legami disolfuro intramolecolari, tra

residui cisteinici di una stessa molecola, serve a stabilizzarne la struttura secondaria. Questi legami possono essere cruciali per la stabilità strutturale della proteina (Tu e Weissmann, 2004). La formazione dei legami disolfuro intermolecolari serve a legare covalentemente tra di loro differenti polipeptidi.

Il destino finale delle diverse proteine di riserva varia in funzione della specie vegetale, della classe di proteine, globuline o prolammine, della tipologia di prolammina, monomerica o polimerica, e forse anche dello stadio di sviluppo dell'endosperma.

In mais e riso le prolammine appaiono tutte trattenute nel lume e depositate in corpi proteici che si originano dal reticolo endoplasmatico (Larkins e Hurkman, 1978; Krishnan et al., 1986). A differenza del mais, le cui principali proteine di riserva sono, come in frumento, le prolammine, in riso queste ultime sono solo una componente minore, in quanto le principali proteine di riserva in questa specie sono le globuline. Queste vengono trasportate attraverso l'apparato del Golgi fino al vacuolo, dove vanno a formare i corpi proteici del riso di origine vacuolare, chiamati PB-II, in contrapposizione con quelli originatisi nel reticolo endoplasmatico, chiamati PB-I, e contenti prolammine (Krishnan et al., 1986).

In frumento la compartimentazione varia a seconda del tipo di prolammina. Le glutenine, in particolare le HMW, si accumulano direttamente nel lume del reticolo endoplasmatico, dal quale si origina un primo tipo di corpi proteici definiti pesanti (Rubin et al., 1992). Le gliadine passano, invece, attraverso il complesso del Golgi e quindi vengono depositate in piccoli vacuoli PSV (protein storage vacuoles), che si fondono a formare un unico grande vacuolo centrale. I corpi proteici contenenti le gliadine vengono chiamati leggeri (Rubin et al., 1992). Mentre i corpi proteici leggeri solo a volte contengono, oltre alle proteine di riserva, anche la BiP, la più nota chaperonina molecolare localizzata nel RE, i corpi proteici pesanti contengono sempre la BiP. I corpi proteici originatisi direttamente dal reticolo endoplasmatico vengono successivamente assorbiti dal vacuolo mediante un processo simile alla autofagìa (Levanony et al., 1992). I corpi proteici in cariossidi in via di sviluppo di frumento contengono anche inclusioni di globuline di riserva (Bechtel et al., 1991), le triticine, presumibilmente trasportate nel vacuolo passando attraverso il complesso del Golgi.

I corpi proteici elettron-densi sono distinguibili al microscopio elettronico già a partire dal decimo giorni dall'antesi e aumentano di dimensioni fino alle ultime fasi della granigione. Field et al. (1983a) hanno dimostrato che corpi proteici isolati durante le fasi di sviluppo intermedie delle cariossidi di frumento contengono polimeri gluteninici uniti da legami disolfuro. Non è ancora noto il meccanismo esatto di fusione dei diversi corpi proteici, ma a maturazione completa, all'interno delle cellule dell'endosperma di frumento ormai secche,

viene persa la struttura in corpi proteici e le proteine si ritrovano disperse a formare proprio quella matrice proteica continua tra i granuli di amido responsabile delle proprietà reologiche del glutine.

Rimane ancora da chiarire, tuttavia, come alcune prolammine, pur non contenendo un segnale di ritenzione per il reticolo endoplasmatico, ad esempio la sequenza tetrapeptidica KDEL o uno dei suoi derivati funzionali, possano essere accumulate nel reticolo endoplasmatico. Una delle spiegazioni proposte è l'interazione delle prolammine con gli enzimi residenti nel lume del RE che assistono il folding proteico. A sostegno di questa ipotesi ci sono numerosi studi in riso, mais e frumento (Galili et al., 1998). Gli studi di Li et al., (1993) dimostrano, ad esempio, l'interazione della BiP con le prolammine in riso. Evidenze biochimiche e molecolari hanno dimostrato che la BiP risulta associata, infatti, ai polisomi impegnati nella sintesi delle prolammine, alle subunità prolamminiche mature e rimane localizzata sulla superficie dei corpi proteici di questi cereali. Anche la PDI è stata localizzata sulla superficie dei corpi proteici di frumento (Levanony et al., 1992; Rubin et al., 1992; Altschuler e Galili, 1994) e osservazioni in mais, orzo e frumento hanno evidenziato che le prolammine all'interno dei corpi proteici sono organizzate in una struttura ben precisa, stabilizzata dalla formazione di legami disolfuro intermolecolari (Rechinger et al., 1993; Coleman et al., 1996; Bagga et al., 1997). In realtà la presenza sia della BiP che della PDI nei corpi proteici può semplicemente riflettere un intrappolamento passivo. È infatti molto improbabile che la BiP si leghi in maniera stabile ad una catena nascente di prolammine, se non nel caso di forme difettose o mal ripiegate (D'Amico et al., 1992; Zhang e Boston, 1992; Pedrazzini et al., 1994).

Una spiegazione molto più semplice ed immediata potrebbe invece essere trovata nel fatto che le prolammine sono sintetizzate a livelli sufficientemente elevati per formare aggregati nel lume del reticolo endoplasmatico, che vengono stabilizzati da legami idrogeno, interazioni idrofobiche e legami disolfuro intermolecolari (Argos et al., 1982; Shewry et al., 1993). Questi aggregati, aumentando di dimensione, rendono impossibile il loro trasporto lungo la via di secrezione, portando alla formazioni di corpi proteici. In conclusione, la formazione di corpi proteici derivati dal reticolo endoplasmatico potrebbe dipendere dalle caratteristiche fisiche di alcune prolammine e dalla loro concentrazione all'interno del lume, sufficientemente alta perché avvenga l'aggregazione. È possibile anche che la BiP si associ a questi aggregati, ma questa associazione sarebbe funzione del controllo di qualità del RE e non conseguenza del suo ruolo nell'assemblaggio proteico (Hammond e Helenius, 1995).

Alcuni studi di espressione eterologa hanno messo in evidenza che le sequenze ripetute ricche in prolina delle γ-gliadine sono necessarie per la ritenzione all'interno del reticolo endoplasmatico (Altschuler et al., 1993; Altschuler e Galili, 1994). Probabilmente queste regioni partecipano a interazioni proteina-proteina che portano alla formazione di aggregati insolubili che, di conseguenza, si accumulano direttamente nel reticolo endoplasmatico (Coleman e Larkins, 1999; Shewry, 1999). Nel caso specifico del frumento, la formazione di aggregati di grandi dimensioni, insolubili, di glutenine le renderebbe meno facilmente trasportabili rispetto alle gliadine monomeriche, solubili nell'ambiente luminale (Shewry, 1999).

Nel lume del RE, un ambiente considerato altamente ossidante, visto il suo più elevato rapporto tra glutatione ossidato (G-S-S-G) e ridotto (G-SH) rispetto al citoplasma (Hwang et al., 1992), avviene quindi il folding proteico e la formazione dei legami intra ed intermolecolari. A differenza dei legami disolfuro intramolecolari delle proteine di riserva, che si formano molto rapidamente, in tempi dell'ordine di secondi a partire dal momento della sintesi, la formazione di polimeri costituiti da subunità gluteniniche, soprattutto quelle ad elevato peso molecolare, avviene attraverso un processo piuttosto lento, che si protrae per molti giorni dopo la sintesi (Shani et al., 1994).

L'enzima che sembra maggiormente coinvolto nella formazione dei legami disolfuro delle proteine di riserva è la disolfuro isomerasi (PDI). La PDI è un enzima che catalizza la formazione e l'isomerizzazione dei legami disolfuro in un ampio numero di substrati proteici. È localizzata nel lume del reticolo endoplasmatico ed è stato dimostrato il suo ruolo nella formazione di legami disolfuro in numerose proteine secretorie in vari sistemi biologici (Freedman, 1984; Freedman, 1989). I dettagli dei meccanismi coinvolti, nonché il ruolo specifico della PDI nel catalizzare i legami disolfuro tra le proteine del glutine nel RE, non sono però ancora stati chiariti. I paragrafi successivi tratteranno proprio della PDI, allo scopo specifico di esporre lo stato attuale delle conoscenze riguardo a questo enzima, che ad oggi rappresenta il principale candidato per la catalisi della formazione ed isomerizzazione dei legami disolfuro nelle prolammine dei cereali e, quindi, potenzialmente importante per determinare le caratteristiche funzionali del glutine, con tutte le conseguenze riguardanti le successive fasi di panificazione e pastificazione.

1.2. La PDI

La disolfuro isomerasi (PDI) è una proteina residente nel lume del RE delle cellule eucariotiche, isolata per la prima volta nel 1963 da fegato di ratto (Goldberger et al., 1963). La sua particolare abbondanza nel lume del reticolo endoplasmatico di cellule che sintetizzano attivamente proteine di secrezione e l'osservazione che le variazioni dei suoi livelli quantitativi erano correlate con quelle della sintesi proteica (Freedman, 1984) hanno suggerito sin dall'inizio un suo possibile coinvolgimento nel processo di folding delle proteine contenenti residui cisteinici. In seguito la PDI è stata rinvenuta in un'ampia varietà di organi e tessuti di specie animali e vegetali (Bjelland, 1987), risultando altamente conservata tra le diverse specie.

Negli anni numerosi nuovi geni codificanti per proteine PDI-like sono stati isolati e studiati in un ampio numero di organismi. E' interessante notare che il nome PDI indica non solo la famiglia genica, ma anche quel particolare membro della famiglia che è stato per primo caratterizzato e studiato in maniera approfondita.

Dopo più di 40 anni dalla prima scoperta di questa proteina, sorgono sempre nuovi quesiti estesi a un gran numero di membri della famiglia di proteine PDI-like, che da un lato mostrano un'enorme diversificazione funzionale e strutturale, dall'altro hanno in comune un'importantissima caratteristica di base, la presenza di almeno un dominio, attivo o inattivo, con il tipico ripiegamento in α-eliche e foglietti β della PDI, evolutosi dal dominio della tioredossina.

Il folding proteico è un passaggio critico per l'attività cellulare, proteine ripiegate male, infatti, non solo perdono la loro capacità di funzionare, ma si possono accumulare come grandi aggregati insolubili che, a loro volta, possono interferire con il corretto funzionamento dell'intera cellula, fino a compromettere la funzionalità di tutto l'organismo. Numerose malattie umane sono legate all'accumulo di proteine non correttamente ripiegate (Dobson, 2001; Thomas et al., 1995; Koo et al., 1999), tra queste l'Alzheimer (Harper et al., 1997), l'enfisema (Cabral et al., 2001) e le malattie dovute a prioni (DeBurman et al., 1997).

Le strutture terziaria e quaternaria di molte proteine extracellulari sono spesso stabilizzate da legami disolfuro (Freedman, 1987 e 1989), la cui presenza a volte risulta essenziale (Tu e Weissmann, 2004).

Nelle fasi iniziali del folding la formazione dei legami disolfuro è soggetta ad errori, dovuti non solo alla formazione di legami disolfuro tra cisteine errate (Creighton, 1979; Rothwarf et al., 1998), ma anche ad una successione temporale imprecisa, in cui un legame tra

una coppia di cisteine, nonostante sia corretto, essendosi formato troppo presto impedisce altre fasi del ripiegamento, che di conseguenza rimane incompleto (Creighton, 1991; Weissman et al., 1991; Chang et al., 2000).

Inoltre il folding spontaneo di proteine contenenti legami disolfuro, che può impiegare da ore a giorni, o addirittura non avvenire mai, è incompatibile con i tempi impiegati dalla secrezione in vivo, che vanno da 30 a 60 minuti (Wilkinson e Gilbert, 2004). L'efficienza e l'esattezza del folding proteico nelle cellule eucariotiche deriva dall'ambiente redox altamente specializzato (Frand et al., 2000) nel quale avviene, il reticolo endoplasmatico. Nel lume del RE sono infatti presenti numerosi enzimi, quali chaperoni molecolari, che inibiscono l'aggregazione, e catalizzatori del folding, che assistono il ripiegamento tramite la formazione e l'isomerizzazione dei ponti disolfuro (Denecke, 1996; Fassio e Sitia 2002), insieme a sistemi di degradazione delle proteine mal ripiegate (Brodsky et al., 1999). La PDI è uno di questi enzimi, capace non solo di catalizzare la formazione netta, la rottura (riduzione) e l'isomerizzazione (riarrangiamento) dei ponti disolfuro, ma anche dotata di una funzione di chaperone molecolare prevenendo le interazioni non favorevoli durante il folding ed inibendo l'aggregazione tra gli intermedi.

1.3. La struttura della PDI

Le principali caratteristiche strutturali e funzionali della PDI sono quindi conservate tra le diverse specie e sono state studiate approfonditamente nei mammiferi, in particolare nell'uomo (Ferrari e Soling, 1999; Wilkinson e Gilbert, 2004; Ellgaard e Ruddock, 2005).

Sebbene siano state documentate anche forme monomeriche e omotetrameriche (Gilbert, 1998), con ciascuna subunità di peso molecolare compreso tra i 55 e i 60 kDa, questo enzima per lo più è stato isolato da diverse specie animali sotto forma di omodimero, costituito da due subunità identiche (Freedman et al., 1994). Sono state clonate e sequenziate diverse sequenze cDNA e genomiche per alcune specie di vertebrati (Edman et al., 1985; Pihlajaniemi et al., 1987; Parkonnen et al., 1988).

Negli anni sono stati proposti diversi modelli strutturali per la PDI (Fig. 1.2), primo fra tutti c'è il modello di Edman (1985), basato sulla struttura primaria dedotta dalle sequenze nucleotidiche dei cloni cDNA della PDI ottenuti da mRNA di fegato e pancreas di ratto (Edman et al., 1985). Secondo questo modello la PDI è un polipeptide di 508 residui aminoacidici, alla cui estremità N-terminale si trova un segnale peptidico di 19 aminoacidi e

costituito di 4 regioni distinte, suddivisibili in due coppie che presentano omologia interna, definite domini ***a***, ***b***, ***b'***, ***a'***.

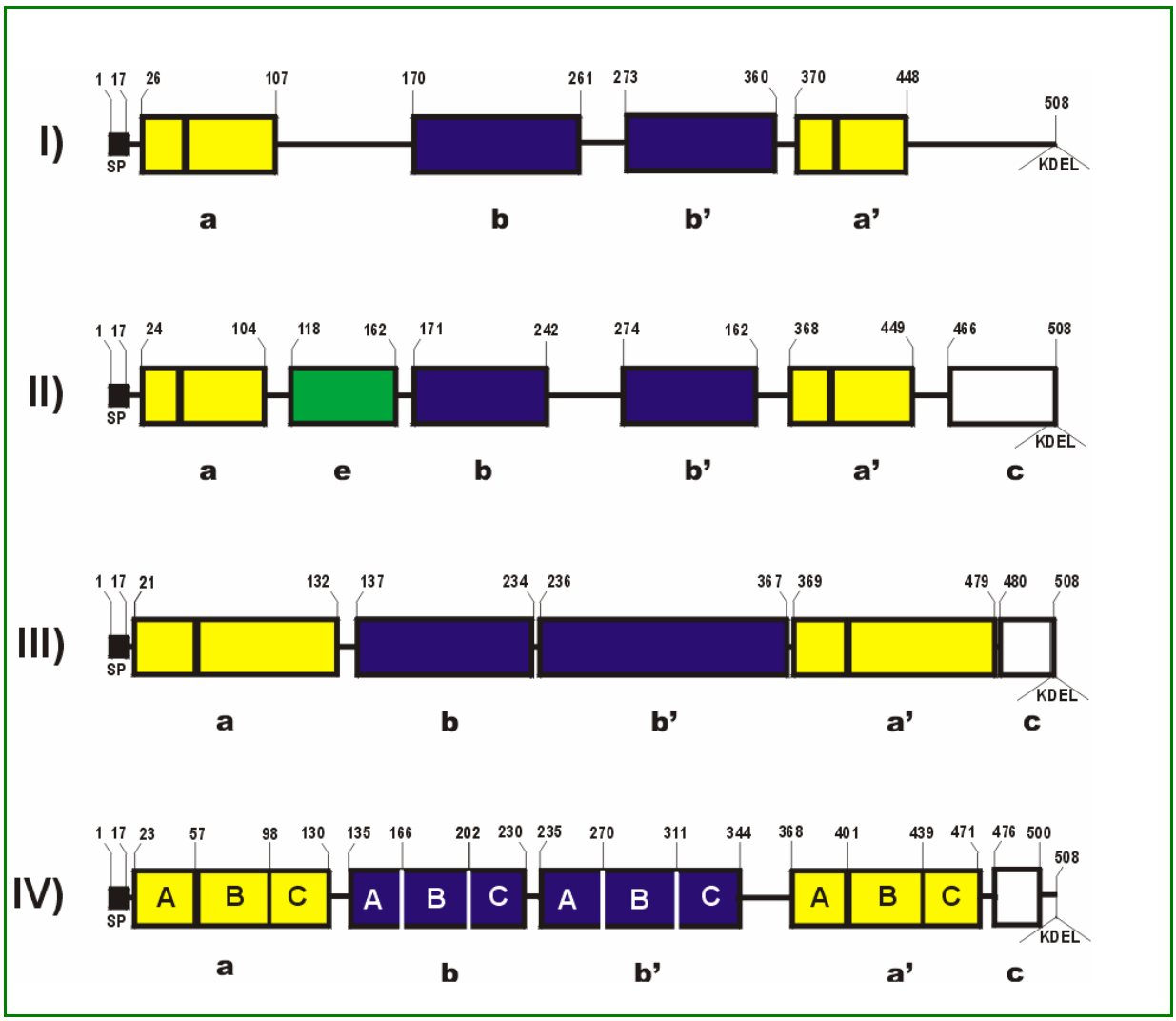

Fig. 1.2 - Modelli strutturali proposti per la PDI:
I) Modello di Edman (1985)
II) Modello di Freedman (1994)
III) Modello di Darby (1995)
IV) Modello di Ferrari (1998)

Nel 1994 Freedman e collaboratori, basandosi sulle giunzioni introne-esone, sulle omologie di sequenza esistenti tra le PDI dei vertebrati e tra queste e le altre proteine correlate e su studi condotti mediante proteolisi parziale delle catene peptidiche, che avviene preferenzialmente alle giunzioni tra i moduli di sequenza, identificarono per la PDI due ulteriori domini, un dominio ***e***, compreso tra il dominio ***a*** e quello ***b***, e un dominio ***c***, nella regione C-terminale della proteina. Il dominio ***c*** contiene una sequenza di residui prevalentemente acidi, caratteristica comune alle proteine che legano il calcio, come le reticoloplasmine (Lucero et al., 1994; Macer e Koch, 1988), e una sequenza KDEL (HDEL

nel lievito), che controlla la ritenzione delle proteine nel RE (Haugejorden et al., 1991; Denecke et al., 1992). Il dominio ***e*** fu chiamato così perché simile al sito di legame degli estrogeni nell'uomo (Tsibris et al., 1989). In seguito, grazie alla manipolazione dei domini mediante tecniche di ingegneria genetica e all'espressione dei domini correttamente ripiegati, i domini ***a*** e ***a'*** (Darby e Creighton 1995a e 1995b) e quelli ***b*** (Darby et al., 1996) e ***b'*** (Darby et al., 1999) sono stati confermati e le giunzioni tra i vari domini identificati con maggiore precisione.

Nel modello corrente, basato su studi NMR sui domini isolati, preparati con tecniche di ingegneria genetica (Kemmink et al., 1995, 1996, 1997), la struttura multidominio della PDI è composta da 5 domini: ***a***, ***b***, ***b'***, ***a'*** e ***c***. I due domini di tipo ***a*** hanno un'elevata omologia di sequenza con la tioredossina, una piccola proteina citoplasmatica coinvolta in numerose reazioni redox (Freedman et al., 1994). Entrambi i domini contengono, inoltre, una sequenza aminoacidica consenso: Cys-Gly-His-Cys (Noiva e Lennarz, 1992), che rappresenta il sito catalitico per l'attività redox-isomerasica (LaMantia e Lennarz, 1993), attribuita alla reattività dei due residui cisteinici. I due domini centrali di tipo ***b*** presentano una certo grado di omologia tra loro e non presentano alcuna identità di sequenza con la tioredossina. La struttura secondaria della tioredossina (Fig 1.3), determinata sperimentalmente mediante misure di risonanza magnetica nucleare (NMR), è caratterizzata da un ripiegamento di tipo α/β (thioredoxin fold), nel quale è contenuta la sequenza del sito attivo, costituita dal tetrapeptide -Cys-Gly-Pro-Cys- chiamato "thioredoxin box".

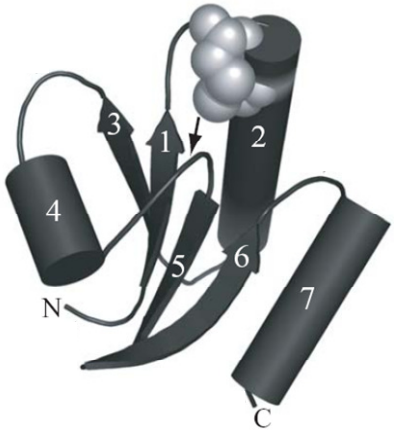

Fig. 1.3 – Struttura secondaria del ripiegamento di tipo α/β (thioredoxin fold) della tioredossina di *E. coli* (PDB file 2TRX) (Katti et al., 1990). I residui da 1-20, che formano un β foglietto ed un'α-elica supplementari al thioredoxin fold generico, non sono mostrati in figura. Le α-eliche sono rappresentate da cilindri mentre i foglietti β da freccie. Il sito attivo è rappresentato da sfere di colore grigio. I numeri da 1 a 7 seguono l'ordine degli elementi nella struttura secondaria (da Sevier e Kaiser, 2006).

I diversi elementi di questo ripiegamento, β-α-β-α-β-α-β-β-α, sono organizzati in modo che i 5 foglietti β, tutti paralleli tra loro eccetto il β4, si trovino al centro circondati dalle 4 α-eliche (Fig. 1.3). Il sito attivo si trova in una sequenza esposta che collega il secondo foglietto β alla seconda α-elica, con l'atomo di zolfo del residuo cisteinico N-terminale in posizione superficiale, tale da essere esposto al solvente, e quello del residuo C-terminale nascosto (Freedman et al., 2002). Studi di spettroscopia NMR, condotti sui singoli domini della PDI isolati e preparati con tecniche di ingegneria genetica, utilizzando la sequenza cDNA di uomo (Kemmink et al., 1996 e 1997), hanno dimostrato che sia il dominio ***a*** sia il dominio ***b***, nonostante la mancanza di omologia di quest'ultimo con la tioredossina, mostrano il caratteristico ripiegamento della tioredossina. Gli elementi della struttura secondaria dei domini ***a*** e ***b*** presentano, infatti, arrangiamenti in 5 foglietti β e 4 α-eliche, del tipo: β-α-β-α-β-α-β-β-α. Ciascun dominio ***a*** e ***b*** può essere suddiviso in tre elementi distinti: A, B e C. La struttura secondaria degli elementi A comprende un motivo β-α-β (Kemmink, 1997), tali elementi A rappresentano le regioni meno conservate tra i domini ***a*** e ***b*** appartenenti alle proteine della famiglia delle PDI, mentre gli elementi C, caratterizzati dal motivo ripetuto β-β-α, costituiscono le regioni più altamente conservate (Ferrari e Söling, 1999). Anche i domini ***a'*** e ***b'*** hanno una probabilità maggiore del 95% di avere una struttura secondaria riconducibile alla tioredossina (Kelley et al., 2000). La PDI contiene, quindi, quattro domini simili alla tioredossina, di cui solo due, ***a*** e ***a'***, conservano il sito attivo, mentre i domini ***b*** e ***b'*** vengono considerati domini tioredossinici inattivi, nei quali è stata persa la sequenza aminoacidica caratteristica del sito attivo per effetto di mutazioni puntiformi (Ferrari et al., 1998).

Sempre in base a studi di risonanza magnetica, il dominio putativo ***e***, individuato da Freedman (1994), non viene considerato un'unità autonoma, ma viene attribuito in parte al dominio ***a*** ed in parte al dominio ***b*** (Darby et al., 1996). Nella parte C-terminale, infine, si trova l'ultimo dominio, il dominio ***c***. Questo dominio, come abbiamo visto in precedenza, contiene una sequenza di residui prevalentemente acidi. Rimane ancora da chiarire come questi domini siano orientati gli uni rispetto agli altri. Recentemente è stata identificata una regione "linker" di 19 aminoacidi tra i domini ***b'*** e ***a'*** (Pirneskoski et al., 2004), che potrebbe determinare una maggiore flessibilità tra questi due domini.

1.4. La famiglia della PDI nell'uomo

Studi, condotti prevalentemente nei mammiferi, hanno evidenziato la presenza di un discreto numero di geni codificanti per proteine con un certo grado di omologia per la PDI classica, indicando l'esistenza di una famiglia di proteine simili alla PDI. Il termine PDI indica, come già detto, sia il primo membro isolato e caratterizzato, definito anche PDI classica, sia l'intera famiglia. Nell'uomo la famiglia delle proteine PDI-like, comprende, ad oggi 17 membri (Fig. 1.5 e Tab. 1.2), la maggior parte dei quali sono stati isolati negli ultimi anni (Ellgaard e Ruddock, 2005). Inizialmente nel RE sono stati identificati i seguenti membri della famiglia della PDI umana: **PDI**, **PDIp**, **ERp57**, **ERp72**, **P5**, **PDIr** e **Erp28** (Ferrari e Söling, 1999; Freedman et al., 2002; Fig. 1.4), a cui in seguito si sono aggiunti **TMX** (Matsuo et al., 2001), **ERp18** (Alanen et al., 2003; Knoblach et al., 2003), **ERp44** (Anelli et al., 2002, 2003), **ERp46** (Knoblach et al., 2003; Sullivan et al., 2003), **ERdj5** (Cunnea et al., 2003; Hosoda et al., 2003), **TMX2** (thioredoxin related transmembrane protein 2, Meng et al., 2003) e **PDILT** (van Lith et al., 2005). In aggiunta a queste 14 proteine ci sono 3 altri membri, **ERp27**, **TMX3** e **TMX4**, per i quali è stata confermata la presenza nel RE (Ellgaard e Ruddock, 2005; Fig. 1.5).

Tutte le proteine della famiglia sono presenti nel reticolo endoplasmatico, in alcuni casi (**PDI** classica e **ERp57**) in quantità relativamente elevate, e sono spesso considerate svolgere la loro funzione esclusivamente in questo compartimento cellulare. Recentemente, come si vedrà meglio in seguito, è stata tuttavia evidenziata la presenza di alcune proteine PDI-like anche in altri compartimenti cellulari, nei quali le funzioni svolte da queste proteine potrebbero essere diverse da quelle esercitate nel RE. Come verrà chiarito nei prossimi paragrafi, l'ambiente altamente ossidante del RE, rispetto agli altri compartimenti cellulari, indirizza l'attività catalitica delle proteine PDI-like verso la formazione di ponti disolfuro, mentre quello maggiormente riducente degli altri compartimenti permette che le stesse proteine PDI-like catalizzino la riduzione dei legami disolfuro.

Al contrario dei domini ***a***, contenenti i siti catalitici attivi, i domini ***b*** non sono facilmente identificabili da programmi di allineamento automatico come BlastP. I domini ***b***, nonostante la loro sequenza aminoacidica risulti altamente divergente, conservano inalterati gli elementi della struttura secondaria, che ne determinano il ripiegamento nel tipico thioredoxin fold.

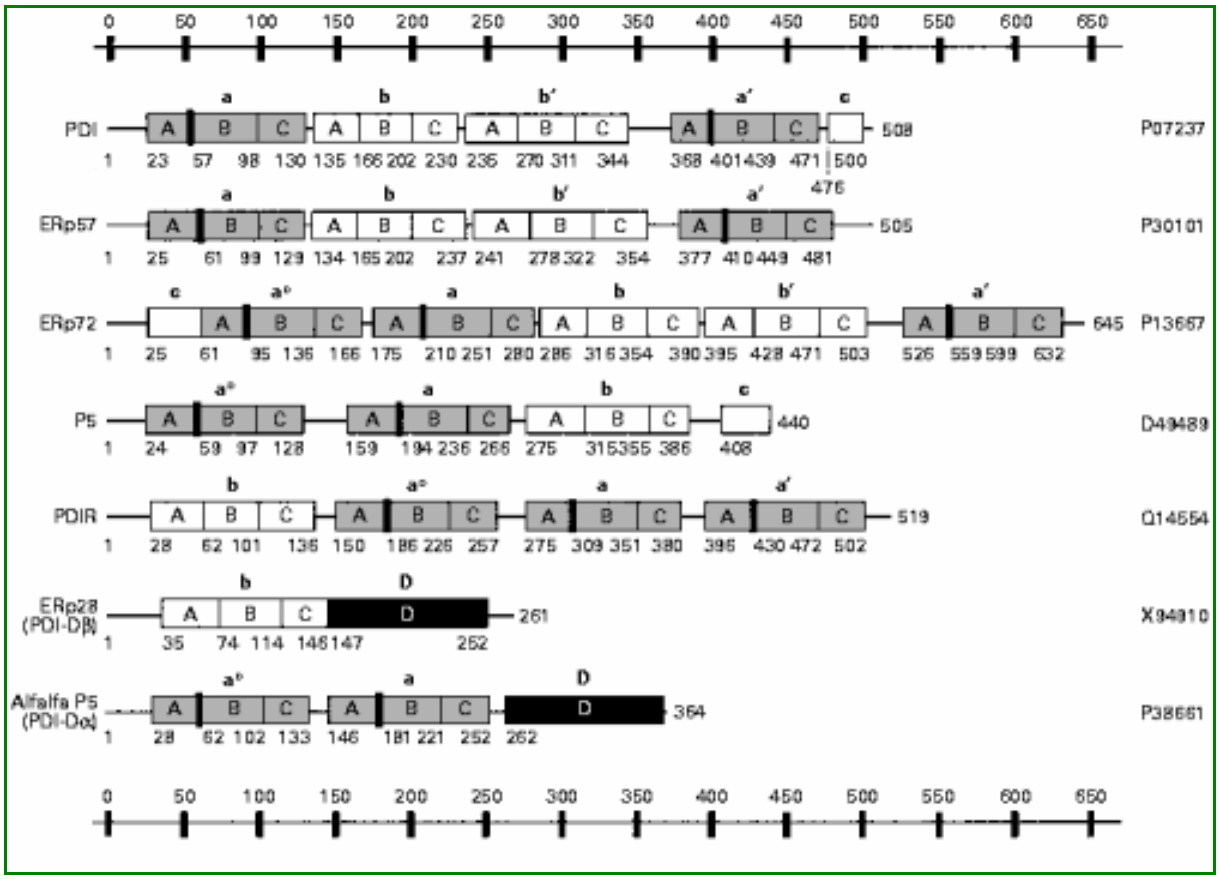

Fig. 1.4 - Strutture multidominio di alcune proteine PDI-like.
I domini tioredossinici attivi, di tipo ***a***°, ***a*** e ***a'*** sono rappresentati da rettangoli grigi, quelli inattivi di tipo ***b*** e ***b'*** da rettangoli bianchi. Un tratto verticale nero indica il sito attivo -Cys-Xaa-Xaa-Cys-. Il dominio ***D***, in nero, è un dominio ad α-elica tipico delle proteine PDI-D, mentre il dominio ***c***, in bianco, è un dominio ricco in residui acidi. A, B e C sono le regioni omologhe con gli elementi del ripiegamento tioredossinico (strutture secondarie β-α-β, α-β-α e β-β-α). I numeri sotto i riquadri indicano gli aminoacidi che delimitano le regioni. I numeri di accessione delle sequenze sono indicati a destra (da Ferrari e Söling, 1999).

Questi domini rivestono una notevole importanza funzionale, in quanto sono responsabili del legame con altre proteine. Vista la difficoltà di identificare questi domini, in letteratura c'è disaccordo sul loro numero e sulla loro presenza in alcune proteine PDI-like. Un caso particolare, che rappresenta un'eccezione, è quello di **ERp28** (Fig. 1.4)**,** che è stata proposta come membro della famiglia della PDI in base alla presenza di una sequenza di circa 100 aminoacidi riconducibile al dominio ***b*** della PDI classica, nonostante l'assenza del sito catalitico tioredossinico (Ferrari et al., 1998). Solo successivamente i dati sulla struttura della proteina ortologa in ratto, ottenuti mediante NMR, hanno indicato la presenza di un dominio riconducile alla tioredossina (Liepinsh et al., 2001), oltre ad un dominio C-terminale composto di 5 α-eliche simili al dominio C-terminale delle proteine PDI-D. Recentemente, grazie ad un nuovo approccio bioinformatico, basato sull'utilizzo del programma VAST, che confronta le coordinate tridimensionali degli atomi di una struttura proteica determinata precedentemente mediante cristallografia o NMR con quelle presenti in banche dati proteiche, è stato possibile identificare domini di tipo ***b*** in altre proteine (Clissold e Bicknell, 2003).

Il polipeptide **ERp57** (Fig. 1.4 e 1.5) presenta un'identità, a livello aminoacidico, solo del 33% con la **PDI,** ma ne conserva tuttavia parte della struttura modulare. Infatti la posizione dei suoi domini catalitici attivi ed inattivi corrisponde a quella dei domini ***a*** , ***a'***, ***b*** e ***b'*** della PDI (Urade e Kito, 1992; Koivunen et al, 1997), mentre manca il dominio acidico C-terminale (Ferrari e Söling, 1999). **ERp57** agisce, sia *in vitro* che *in vivo*, attraverso l'interazione con la calnexina e la calreticulina (Zapun et al., 1998*)*, due lectine residenti nel RE, su proteine nascenti monoglicosilate in maniera indipendente dalla presenza dei legami disolfuro (Elliott et al., 1997; Oliver et al., 1997). Le regioni importanti per l'interazione di **ERp57** sono in primo luogo il dominio ***b'*** ed in secondo la regione C-terminale carica positivamente (Pollock et al., 2004; Russell et al., 2004; Urade et al., 2004). **ERp57,** al contrario della **PDI** che interagisce con le proteine indipendentemente dal loro stato di glicosilazione, non agisce su proteine non glicolsilate. Il sito di legame per la calreticulina nel dominio ***b'*** di **ERp57** consiste in residui equivalenti a quelli che formano il sito di legame per il substrato nella **PDI** (Russell et al., 2004).

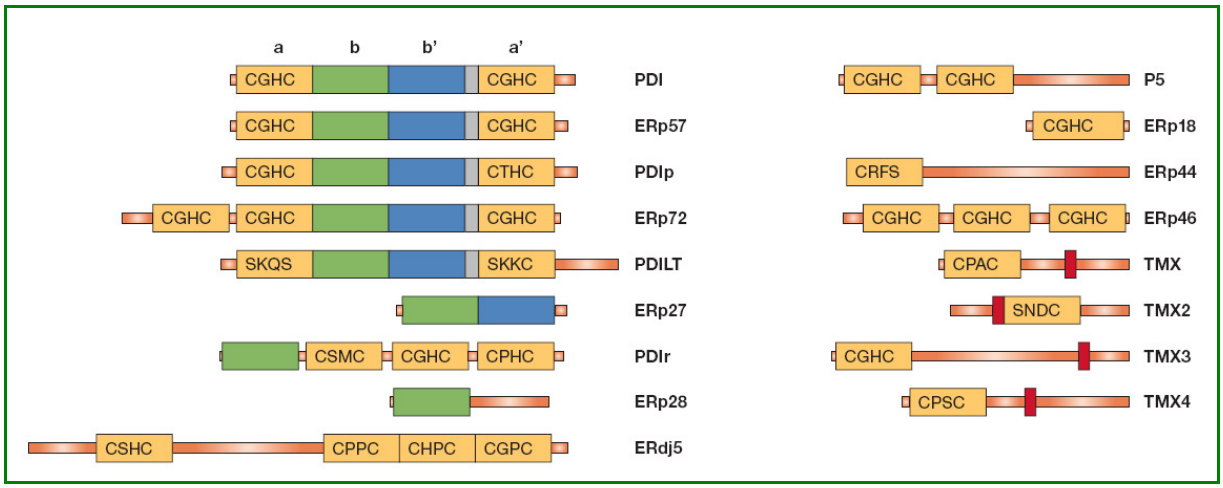

Fig. 1.5 - La famiglia delle PDI nell'uomo. I domini simili alla tioredossina attivi sono rappresentati da rettangoli gialli, quelli inattivi di tipo ***b'*** in blu, altri domini catalitici inattivi in verde. La regione di giunzione tra ***b'*** e ***a'*** è rappresentata in grigio, mentre le regioni transmembrana sono di colore rosso. Non sono messi in evidenza né i peptidi segnale, né i domini ***c*** terminali (da Ellgaard e Ruddock, 2005).

Apparentemente i membri della famiglia della PDI hanno un sito di legame conservato, localizzato in un dominio tioredossina simile non catalitico, responsabile per l'elevata affinità di legame con il substrato o con cofattori interagenti con il substrato (Ellgaard e Ruddock, 2005). Nella maggior parte dei casi questo sito di legame è localizzato in domini di tipo ***b'***. Domini di questo tipo sono presenti in: **PDI**, **PDIp**, **ERp57**, **ERp72, Erp27** e **PDILT** (Fig. 1.5). L'allineamento di questi domini ha mostrato che quelli di **PDIp, Erp27** e **PDILT** sono più simili a quelli della PDI e, quindi, probabilmente coinvolti nel

legare proteine non native. Il sito di legame di **ERp72** è invece più simile a **ERp57** e molti dei residui coinvolti nell'interazione ERp57-calreticulina risultano conservati (Ellgaard e Ruddock, 2005). Non ci sono comunque evidenze sperimentali che suggeriscano una interazione tra **ERp72** e calreticulina. È interessante notare che l'interazione con calnexina o calreticulina aumenta l'attività isomerasica di **ERp57** su proteine monoglicosilate, mentre nella **PDI** lo stesso tipo di interazione è responsabile della riduzione della sua attività e della capacità di legarsi ai peptidi (Baksh et al., 1995; Zapun et al., 1997; Zapun et al., 1998).

Tab. 1.2 – Caratteristiche dei membri della famiglia della PDI nell'uomo (da Ellgaard e Ruddock, 2005).

Name*	SwissProt accession	Length	ER retention	*N*-glycosylation sites (putative)	a-like domains	Charge pair sequence	Conserved arginine
PDI	P07237	508	KDEL	0	2	E47–K81, E391–K424	Yes, yes
ERp57	P30101	505	QEDL	0	2	E51–K82, E400–K433	Yes, yes
PDIp	Q13087	525	KEEL	3	2	E65–K99, K412–E445	Yes, yes
ERp72	P13667	645	KEEL	0	3	E85–K119, E200–K234, E549–K582	Yes, yes, yes
ERp65	Q8N807	584	KEEL	8	2	L66–K100, M411–K444	No, no
ERp27	Q96DN0	273	KVEL	1	0	–	–
PDIr	Q14554	519	KEEL	0	3	M176–N209, M299–A334, M420–A453	No, yes, no
ERp28	P30040	261	KEEL	1	0	–	–
ERdj5	Q8IXB1	793	KDEL	1	4	N152–A183, D474–T505, D582–S613, D694–K725	Yes, no, yes, yes
P5	Q15084	440	KDEL	0	2	E49–A80, E184–A219	Yes, yes
ERp18	O95881	172	EDEL	1	1	I60–N93	No
ERp44	Q9BS26	406	RDEL	0	1	N52–R89	Yes
ERp46	Q8NBS9	432	KDEL	0	3	M83–K118, K211–K244, K344–E378	Yes, yes, yes
TMX	Q9H3N1	280	Unknown	0	1	E50–K82	Yes
TMX2†	Q9Y320	296	KKDK	2	1	E161–K193	?‡
TMX3	Q96JJ7+	454	KKKD	2	1	D47–K81	Yes
TMX4	Q9H1E5	349	Unknown	1	1	K58–K90	No

ERp72 (Fig. 1.4 e 1.5) è un'abbondante proteina ubiquitaria, indotta da stress, con la capacità di legare il calcio (Ferrari e Söling, 1999). Essa differisce dalle altre proteine PDI-like per la presenza di tre domini attivi tioredossina simili (Mazzarella et al., 1990), in particolare la distribuzione dei domini è ***c-a°-a-b-b'-a'*** (Fig. 1.4). La proteina possiede un'attività redox e disolfuro isomerasica significativa (Van Nguyen et al., 1993; Rupp et al., 1994) e può complementare l'assenza della **PDI** in lievito (Günter et al., 1993). *In vitro,* **ERp72** interagisce con proteine denaturate, in associazione con chaperoni molecolari quali la PDI, la BiP e la GRP94 (Kuznetsov, 1994; Nigam et al., 1994).

Assieme a **ERp57**, anche **PDIp** (Fig. 1.5) conserva la struttura nei 4 domini, ***a-b-b'-a'*** della **PDI** classica, con la quale ha il 45% di identità, e manca della regione C-terminale acida. Ci sono però differenze sostanziali nella sequenza del dominio ***b'***, ciò suggerisce che le tre proteine possano avere differenti e specifiche proprietà di legame al substrato. **PDIp,** infatti, si lega in modo specifico a peptidi contenenti un singolo residuo di Tyr o di Trp nelle

cui vicinanze non ci siano cariche negative (Ruddock et al., 2000). Questa proteina tessuto specifica è espressa esclusivamente nelle cellule del pancreas dell'uomo (Desilva et al., 1997), nel quale è responsabile del folding di un gruppo particolare di proteine, i zimogeni pancreatici (Freedman et al., 2002).

La **P5** (Fig. 1.4) ha un'organizzazione in domini di tipo ***a***°-***a***-***b***-***c*** (Chaudhuri et al., 1992). La sua omologa in ratto CaBP1, dalla struttura ***a***°-***a***-***b***-***c***, è la meglio caratterizzata delle proteine di tipo P5. CaPB1 è in grado, come la PDI classica, di catalizzare la riduzione dell'insulina (Van Nguyen et al., 1993; Füllekrug et al., 1994). Analisi biochimiche dettagliate hanno dimostrato che CaBP1 può complementare la mutazione PDI1, letale in lievito, ma non è in grado di catalizzare il ripiegamento redox indipendente di GAPDH (Kramer et al., 2001).

Oltre alle proteine P5 nei mammiferi, ora sono note proteine omologhe in *A. thaliana* che presentano una struttura di tipo ***a***°-***a***-***b***-***c*** (Lemaire e Miginiac-Maslow, 2004). Alcune proteine originariamente attribuite al gruppo delle proteine P5, isolate in erba medica (Shorrosh e Dixon, 1992) ed in *Dictyostelium discoideum* (Dd-PDI Monnat et al., 1997), sono state ascritte alla sub-famiglia della PDI-D (Ferrari e Söling, 1999), i cui membri sono tutti caratterizzati dalla presenza di un dominio, chiamato ***D***, C-terminale, con struttura ad α-elica di circa 110 aminoacidi, la cui funzione è sconosciuta. A questo gruppo appartengono proteine redox attive (PDI-Dα), come la **PDI-D** di erba medica, con i suoi domini ***a***°-***a***-***D***, o inattive (PDI-Dβ), come **ERp28** umana, con i suoi domini ***b***-***D*** (Ferrari et al., 1998), o la **ERp29** di ratto (Demmer et al., 1997; Mkrtchian et al., 1998).

La **PDIr** (Fig. 1.4 e 1.5) con la sua organizzazione in domini ***b***-***a***°-***a***-***a'*** (Hayano e Kikuchi, 1995), risulta molto interessante in quanto, nonostante non si abbiano informazioni sulla sua funzione, ciascun dominio di tipo ***a*** presenta una sequenza differente nel sito attivo: -Cys-Ser-Met-Cys- (***a***°),-Cys-Gly-His-Cys- (***a***), -Cys-Pro-His-Cys- (***a'***). Ciò ha fatto ipotizzare che ogni dominio possa esplicare l'attività catalitica con velocità differenti a seconda del substrato (Ferrari e Söling, 1999).

La sequenza più comune del sito attivo in membri della famiglia della PDI umana è –CXHC-, caratteristica di proteine del RE dotate di un'efficiente attività ossidante. Questo motivo è presente in **PDI**, **PDIp**, **ERp57**, **ERp72**, **P5, ERp46** e **TMX3**, oltre che in uno dei siti catalitici attivi di **ERdj5** e, come abbiamo appena visto, in due domini della **PDIr** (Ellgaard e Ruddock, 2005; Fig. 1.5). Tre dei siti attivi di **ERdj5** contengono, invece, il motivo -CXPC-, trovato nella tioredossina caratterizzata da attività riducente (Fig. 1.5). **ERp27** e **ERp28** (Fig. 1.5), in quanto non presentano il sito attivo, così come **PDILT** e

TMX2 (Fig. 1.5), per mutazione delle cisteine al sito attivo, non hanno la cisteina N-terminale del motivo -**C**XXC- che, come si vedrà meglio in seguito, essendo esposta alla superficie è essenziale per l'attività catalitica di ossido-riduzione. In **ERp44** (Fig. 1.5) manca, invece, la cisteina C-terminale, necessaria affinché le reazioni redox procedano in maniera efficiente e che assicura la liberazione della PDI dal legame con il substrato, in modo che non si accumulino legami disolfuro misti intermedi. **ERp44**, quindi, formando legami disolfuro misti stabili, partecipa al meccanismo di ritenzione nel RE di varie proteine (Anelli et al., 2002, 2003), tra cui Ero1 e le varie proteine nascenti. La funzione dei 2/3 C-terminali, non tioredossina simili, di questa proteina non è nota, come non sono note le caratteristiche che ne determinano la specificità di legame al substrato (Ellgaard e Ruddock, 2005).

Oltre alla presenza di un sito attivo -CXHC- e ad un dominio di tipo ***b***, responsabile dell'interazione con il substrato, altre due caratteristiche sono importanti nel determinare l'attività enzimatica delle proteine PDI-like (Ellgaard e Ruddock, 2005). La presenza di un'arginina conservata (Lappi et al., 2004) (Tab. 1.2), che influenza il $pK_{a,}$ e la conservazione dell'acido glutammico che, probabilmente, funge da accettore di protoni di una coppia carica acido glutammico-lisina localizzata sotto il sito attivo e risultata importante per l'attività catalitica della tioredossina (Dyson et al., 1997). Si suppone che **PDI**, **PDIp**, **ERp57**, e **ERp72,** in quanto dotate di queste caratteristiche, siano coinvolte nella formazione e isomerizzazione dei legami disolfuro, il che è stato confermato *in vitro* per **PDI**, **ERp57** e **ERp72** (Rupp et al., 1994; Darby et al., 1998; Frickel et al., 2004). **ERp27, ERp28, PDILT** e **TMX2,** al contrario, probabilmente non sono direttamente coinvolte nella formazione di ponti disolfuro, visto che non presentano il sito attivo -CXXC-. **P5, ERp46** e **TMX3**, a cui manca un sito tioredossinico di tipo ***b'*** ma presentano le altre caratteristiche e un sito attivo -CXHC-, potrebbero essere delle efficienti ossidasi, mentre **ERdj5,** contenente un sito attivo -CXPC-, potrebbe essere un'efficiente riduttasi. Il sito inusuale di **TMX** rende difficile predire la sua funzione. La proteina presenta comunque le altre caratteristiche necessarie per un efficiente catalisi di reazioni di scambio tiolo-disolfuro. Ai restanti membri, **PDIr**, **ERp18** e **TMX4,** manca l'acido glutammico, perciò si può prevedere che siano inefficienti nella catalisi di questo tipo di reazione. Ciò è stato confermato *in vitro* sia per **ERp18**, che mostra di possedere solo il 15% di attività ossidativa del dominio ***a*** della PDI (Alanen et al., 2003), che per **PDIr**, che conserva solo il 2% di attività isomerasica della PDI (Horibe et al., 2004) e 6% della sua attività ossidativa (Ellgaard e Ruddock, 2005).

A parte i domini di tipo ***a*** e ***b*,** le nuove PDI hanno sequenze diverse, che non possono essere allineate con quelle già identificate nelle PDI-like, e che possono contribuire alla

specificità di legame con il substrato. **ERdj5**, ad esempio**,** ha un dominio DnaJ seguito da 4 domini tioredossina simili (Cunnea et al., 2003). Questa proteina si lega alla BiP, un'ATPasi, attraverso il suo dominio DnaJ. I domini DnaJ sono ubiquitari e prendono il loro nome dalla proteina DnaJ descritta nei procarioti, un membro della famiglia delle Hsp40 (Heat-shock protein 40). Nei batteri DnaJ si lega alla proteina DnaK, membro della famiglia delle Hsp70. Assieme queste due proteine assistono il folding proteico attraverso l'energia rilasciata dall'idrolisi dell'ATP (Suh et al., 1999). In *E. coli* è stato visto che le chaperonine di tipo DnaK/DnaJ assistono la riparazione del DNA mantenendo le proteine coinvolte in questo processo correttamente ripiegate (Zou et al., 1998). Anche alcuni membri della famiglia PDI possono fungere da chaperoni molecolari interagendo con il substrato, un dominio DnaJ potrebbe conferire loro maggiore specificità (Clissold e Bicknell, 2003).

Altri domini particolari sono quelli transmembrana presenti in un gruppo particolare di PDI: **TMX, TMX2, TMX3** e **TMX4** (Fig. 1.5). Per alcune di queste proteine è stato dimostrato il loro ruolo nella regolazione dello sviluppo, ad esempio in *C. elegans* sia il dominio transmembrana che il dominio attivo della proteina DPY-11 sono essenziali per la determinazione della forma (lunghezza) del corpo e della morfologia dei raggi (Ko e Chow, 2002).

La presenza di un numero rilevante di membri della famiglia delle PDI nell'uomo evidenzia la complessità di un sistema che siamo lontani dal comprendere. Sarà essenziale verificare se i diversi omologhi presentano affinità per il substrato ridondanti e sovrapponibili o separate e distinte e definire i tipi di interazione dei diversi membri di questa famiglia tra di loro e con altri enzimi e chaperoni residenti nel RE.

1.5. Funzioni della PDI e delle proteine PDI-like

Dalla sua scoperta alla PDI classica fu attribuita la funzione di formare ponti disolfuro (Freedman et al., 1994) e di catalizzare l'isomerizzazione dei legami disolfuro non nativi (Laboissière et al., 1995). La sua capacità di legare proteine parzialmente o non ripiegate, prevenendone così l'aggregazione, ha suggerito anche un suo ruolo come chaperone molecolare (Hayano et al., 1996; Yao et al., 1997). Sono state identificate, inoltre, diverse funzioni addizionali della PDI (Ferrari e Söling, 1999; Pihlajanieni et a., 1987; Wetterau et al., 1990; Lucero e Kaminer, 1999; Cheng et al., 1987; Bennet et al., 1988; Tsai et al., 2001; Tanaka et al., 2000).

Successivamente sono stati isolati un gran numero di omologhi della PDI, precedentemente enumerati, tutti caratterizzati dalla presenza di almeno un dominio che ha il tipico thioredoxin fold, attivo o inattivo, e dotati di domini aggiuntivi diversi tra loro, che probabilmente sono responsabili delle loro funzioni specifiche e differenziate.

Tutti gli omologhi della PDI contenenti un sito attivo CXXC o CXXS, valutati fino ad oggi per la loro attività redox, ad eccezione della **ERdj5**, hanno dimostrato *in vitro* di avere la capacità di ossidare tioli liberi e/o catalizzare l'isomerizzazione di legami disolfuro (Sevier e Kaiser, 2006). *In vivo* rimane però difficile determinare quale sia la loro attività redox preferenziale, a causa delle funzioni dei diversi membri della famiglia che si sovrappongono e complementano (Norgaard et al., 2001), ma anche la funzione della PDI stessa rimane poco chiara (Laboissière et al., 1995; Solovyov et al., 2004).

1.5.1. Attività redox isomerasica

La formazione dei ponti disolfuro nativi è un meccanismo complesso, in cui non solo i legami disolfuro devono essere formati (ossidazione), ma i legami non corretti devono esseri rotti (riduzione) e/o riarrangiati (isomerizzazione). Specialmente per proteine ricche in legami disolfuro nativi, l'isomerizzazione di questi legami implica un sostanziale cambiamento conformazionale, rappresentando una tra le fasi più critiche del folding. Ad oggi si ritiene che negli eucarioti solo i membri della famiglia della PDI siano capaci di catalizzare questo passaggio.

1.5.1.1. La superfamiglia della tioredossina

Le reazioni di ossido-riduzione sono catalizzate dai membri di una grande famiglia di ossidoreduttasi, la super-famiglia della tioredossina, a cui appartiene anche la PDI (Ferrari e Söling, 1999). I membri di questa famiglia sono dotati di uno o più domini che presentano il tipico ripiegamento in α eliche e foglietti β (thioredoxin fold). Questi domini vengono considerati attivi se contengono un sito attivo caratterizzato da un motivo -CXXC-. La PDI possiede due di questi domini attivi, ***a*** e ***a'***, i cui siti attivi presentano il motivo -CGHC-. Altri membri di questa grande famiglia possono invece avere un solo dominio (DsbA), due domini (ERp57), tre domini (ERp72) o quattro domini (ERdj5).

Nel trasferimento di equivalenti redox al substrato, il sito attivo passa dalla forma ossidata a quella ridotta e viceversa. Infatti il legame disolfuro tra i due residui cisteinici del

sito attivo può essere ridotto formando due tioli, che a loro volta possono essere ossidati formando di nuovo il ponte disolfuro intramolecolare. La stabilità relativa di queste due forme determina la facilità con cui il legame disolfuro del sito attivo può essere ridotto e, quindi, se un enzima è un buon ossidante (forma ridotta più stabile) o un buon riducente (forma ossidata più stabile). Nel caso di un buon ossidante il legame disolfuro del sito attivo è meno stabile, più difficile da formare e trasferito con facilità al substrato. Nel caso contrario, quando un enzima è un buon riducente, il legame disolfuro del sito attivo si forma facilmente ed è più difficile da trasferire al substrato. Il potenziale redox standard della PDI è di circa -180 mV, molto più ossidante rispetto a quello della tioredossina (-260 mV) ma meno di quello di DsbA (-100 mV) (Hawkins et al., 1991; Lundström e Holmgren, 1993; Wunderlich e Glockshuber, 1993). Parte della variazione del potenziale redox è causata dalla stabilizzazione dei tioli cisteinici (Chivers et al., 1997) dovuta, in parte alle interazioni elettrostatiche in prossimità del sito attivo (Jacobi et al., 1997). È stato dimostrato che per l'enzima DsbA (Grauschopf et al., 1995), l'ossidante più forte di questa super-famiglia, e per la tioredossina (Chivers et al., 1997; Lundström et al., 1992), uno dei migliori riducenti, i due aminoacidi compresi fra le due cisteine sono importanti nel determinarne il potenziale redox. Mutando il sito attivo della tioredossina da -CGPC- a quello della PDI -CGHC-, un ottimo ossidante, rende la prima un ossidante 10 volte più forte (Lundström et al., 1992), mentre mutando il sito attivo della DsbA da -CPHC- a quello della tioredossina la rende 1200 volte più riducente (Chivers et al., 1997). I due domini di tipo ***a*** della PDI hanno proprietà redox intermedie e possono agire sia come riduttasi che come ossidasi, in funzione del substrato e dell'ambiente redox (Zapun et al., 1993; Darby e Creighton, 1995 a,b).

Nessuna di queste proteine contenenti un singolo dominio attivo ha mostrato *in vitro* una significativa attività isomerasica (Hawkins et al., 1991). A parte gli omologhi della PDI nelle cellule eucariotiche, infatti, esiste solo un altro membro della famiglia che mostra una significativa attività isomerasica ed è un enzima batterico, la DsbC, che dimerizza creando un sito di legame fiancheggiato da due siti attivi (Zapun et al., 1995).

La PDI è il membro più versatile di questa super-famiglia, se si considera che *in vitro* è capace di catalizzare la formazione netta, la rottura e il riarrangiamento dei legami disolfuro in una varietà di substrati (Freedman, 1989). Il suo ruolo preferenziale *in vivo* è invece rimasto ad oggi poco chiaro. A sostegno di questo ruolo attribuito alla PDI come catalizzatore del folding delle proteine, esistono numerosi studi cinetici condotti utilizzando diversi substrati proteici, in particolare la ribonucleasi e l'inibitore della tripsina pancreatica bovina (BIPT), e cambiando le condizioni redox dell'ambiente della reazione. Da questi studi è

emerso che la disolfuro isomerasi può catalizzare tre tipi differenti di reazione (Creighton et al., 1980; Zapun et al., 1992; Weissman et al., 1993; Huth et al., 1994). Con substrati proteici ridotti, in ambiente ossidante dovuto alla presenza di glutatione ossidato GSSG, la PDI catalizza la formazione di legami disolfuro, sia in piccole proteine che in complessi multimerici, accelerando gli stadi che coinvolgono gli scambi tiolo/disolfuro e i cambiamenti conformazionali, senza modificare la natura chimica degli intermedi. Con substrati proteici ossidati, che contengono legami disolfuro non nativi, l'enzima catalizza i riarrangiamenti dei legami. In condizioni moderatamente riducenti, in presenza di glutatione ridotto GSH (10-3 M di GSH o 10-5 di DTT), la reazione conduce alla proteina con i legami disolfuro corretti compatibili con la conformazione nativa, come nel caso della reazione di conversione della RNase scrambled in RNase nativa, che rappresenta il saggio classico per l'analisi dell'attività isomerasica della PDI. In condizioni marcatamente riducenti (10-2 M di GHS), tuttavia l'enzima funziona come una ossidoreduttasi, catalizzando la riduzione netta dei legami disolfuro della proteina.

1.5.1.2. Catalisi dell'attività redox

La catalisi della formazione e dell'isomerizzazione dei legami disolfuro, come già detto, è funzione dei due residui cisteinici conservati nei due siti attivi (**C**GH**C**) della PDI. Nella PDI ossidata questi due residui cisteinici sono legati da un ponte disolfuro intramolecolare e possono catalizzare la formazione netta di legami disolfuro (ossidazione) nel substrato ridotto, mentre nella PDI ridotta si trovano sotto forma di ditiolo e possono catalizzare la rottura (riduzione) o il riarrangiamento (isomerizzazione) dei legami disolfuro del substrato.

Durante il riarrangiamento dei legami disolfuro, la cisteina N-terminale di ciascun sito attivo (**C**GHC) della PDI ridotta, particolarmente reattiva in quanto esposta al solvente (Walker et al., 1996), inizia l'isomerizzazione attaccando un legame disolfuro del substrato per formare un disolfuro misto e liberando un gruppo Cys-SH (tiolo) del polipeptide. A questo punto potrebbero verificarsi tre eventi (Wilkinson e Gilbert, 2004), illustrati in seguito.

1) Si riforma il legame disolfuro intramolecolare iniziale liberando la PDI nella sua forma ridotta, senza che ci sia isomerizzazione del substrato. Questo meccanismo potrebbe permettere alla PDI di riconoscere i substrati già correttamente ripiegati o, viceversa, che semplicemente non risultano un buon substrato per l'isomerizzazione.

2) Nel substrato si forma un legame disolfuro intramolecolare diverso, che a cascata provoca il riarrangiamento intramolecolare dei legami disolfuro non nativi ed il rilascio del substrato correttamente ripiegato e della PDI nella sua forma ridotta; in questo secondo caso è necessario che nel substrato sia presente una cisteina in forma tiolica, capace di reagire con un disolfuro del substrato per innescare il riarrangiamento intramolecolare.

3) La cisteina C-terminale della PDI reagisce con quella N-terminale legata al substrato formando un legame disolfuro con il conseguente rilascio del substrato ridotto e della PDI nella sua forma ossidata. Il substrato ridotto può quindi essere nuovamente ossidato dalla PDI. In quest'ultimo caso l'isomerizzazione avviene tramite cicli di riduzione ed ossidazione ripetuti fino a che tutti i legami disolfuro risultano corretti e la proteina nativa non risulta più un buon substrato per la PDI.

Il residuo cisteinico N-terminale è quindi responsabile della interazione diretta con il substrato, mentre il secondo residuo cisteinico (CGH**C**), più interno e nascosto, reagisce solo con la cisteina N-terminale (Walker et al., 1996) ed il suo ruolo principale *in vivo* potrebbe essere quello di facilitare la rottura rapida dei legami disolfuro intermolecolari, permettendo l'efficiente scansione dei diversi isomeri disolfuro e prevenendo l'intrappolamento della PDI nei disolfuri misti intermedi (Laboissière et al., 1995; Chivers et al., 1996; Walker et al., 1996; 1997).

Sono stati quindi proposti due meccanismi differenti per l'isomerizzazione: nel primo caso sarebbe il risultato di una reazione a cascata intramolecolare innescata all'interno del substrato, mentre nel secondo deriverebbe da cicli ripetuti di riduzione ed ossidazione (Fig. 1.6). Nel primo caso è necessaria solo la PDI ridotta, nel secondo servono sia la PDI ridotta che quella ossidata. Diverse evidenze suggeriscono che l'isomerizzazione della RNasi "scrambled" richieda cicli di riduzione/ossidazione (Wilkinson e Gilbert, 2004). In assenza di un tampone redox, siti attivi con la sola cisteina C-terminale non hanno attività isomerasica (<1%). Con mutanti –CGH**S**- e non con wtPDI vengono trovati intermedi PDI-substrato legati covalentemente (Walker e Gilbert, 1997), indicando che la seconda cisteina probabilmente ne limita l'accumulo. In queste condizioni la seconda cisteina CGH**C** è essenziale per l'isomerizzazione (Walker et al., 1996). La mancanza di PDI ossidata come di quella ridotta inibisce l'isomerizzazione catalizzata dalla stessa PDI (Schwaller et al., 2002). Mentre la sRNasi richiede cicli di riduzione e ossidazione è probabile che altri substrati utilizzino la via dei riarrangiamenti intramolecolari (Wilkinson e Gilbert, 2004).

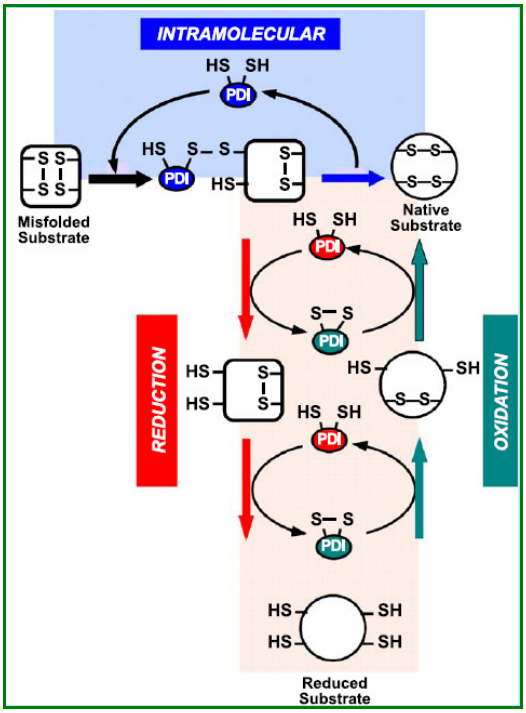

Fig. 1.6 – I due meccanismi proposti per l'isomerizzazione dei ponti disolfuro da parte della PDI. In alto è raffigurato il meccanismo intramolecolare in cui l'isomerizzazione è il risultato di una reazione a cascata intramolecolare mentre in basso sono raffigurate le reazioni di riduzione ed ossidazione del substrato che rappresentano un meccanismo di isomerizzazione alternativo (da Wilkinson e Gilbert, 2004).

Analogamente nell'attività ossidativa la PDI, nella sua forma ossidata, trasferisce direttamente al substrato ridotto il legame disolfuro mediante la formazione di un disolfuro misto dal quale la PDI viene rilasciata nella sua forma ridotta e il substrato nella forma ossidata. L'attività ossidativa richiede, come vedremo, solo uno dei due domini di tipo ***a*** e gran parte della catalisi dell'ossidazione può essere attribuita alla reattività del legame disolfuro al sito attivo, che risulta circa 500 volte più reattivo di quello del glutatione (Walker e Gilbert, 1995). Il secondo residuo cisteinico è richiesto anche per la completa ossidazione dei substrati proteici *in vitro.*

1.5.1.3. Domini e funzioni della PDI

Il contributo dei singoli domini e di combinazioni lineari di questi all'attività ossidoreduttasica della PDI è stato studiato usando saggi basati sul folding *in vitro* di proteine contenenti legami disolfuro, quali la BPTI (inibitore della tripsina pancreatica bovina), per cui il preciso meccanismo di folding è noto, e costrutti per quasi tutte le combinazioni lineari dei diversi domini (Darby e Creighton, 1995c; Darby et al., 1998).

L'espressione dei singoli domini ***a*** e ***a'*** è sufficiente a svolgere la funzione ossidativa, ma presenta una ridotta attività isomerasica (Darby e Creighton, 1995c; Darby et al., 1998). Per reazioni di isomerizzazione semplici sono necessarie combinazioni lineari di un dominio catalitico ***a*** o ***a'*** con il dominio ***b'***, e per reazioni di isomerizzazioni che implicano cambiamenti sostanziali nella struttura del substrato sono necessari costrutti che contengano tutti i domini, con il caratteristico thioredoxin fold (***a***, ***a'***, ***b*** e ***b'***) (Darby et al., 1998). Appare chiaro che c'è un azione sinergica tra i domini catalitici ***a*** e ***a'*** e con il dominio ***b'***, nonostante questo sia inattivo. La catalisi di reazioni di isomerizzazione complesse implica che la PDI nella sua interezza sia in grado di stabilizzare la conformazione parzialmente dispiegata del substrato legandosi a quest'ultimo. Il dominio ***b'*** è risultato essenziale e sufficiente per legare peptidi di dimensioni ridotte (Klappa et al., 1998). Peptidi di dimensioni maggiori e substrati proteici non nativi, per legarsi richiedono la presenza congiunta dei domini ***b'***, ***a'*** e ***c*** (Klappa et al., 1998)**.** Il dominio ***b*** probabilmente ha un ruolo strutturale (Ellgaard e Ruddock, 2005) e non catalitico, in quanto non sembra essere implicato né nel legare il substrato (Klappa et al., 1998), né nel contribuire all'attività catalitica dei costrutti lineari della PDI (Darby et al., 1998).

1.5.1.4. L'origine degli equivalenti ossidativi

Le proteine nascenti sono introdotte nel reticolo endoplasmatico (RE) con i residui cisteinici ridotti. Nel RE vengono formati i legami disolfuro nativi grazie al trasferimento dei legami disolfuro dalla PDI al substrato e al conseguente rilascio del substrato ossidato e della PDI ridotta. Per poter continuare la sua attività catalitica la PDI dopo ogni ciclo di ossidazione deve essere a sua volta riossidata. Viene naturale domandarsi da dove vengano gli equivalenti ossidativi necessari per la formazione dei legami disolfuro nativi e con quale meccanismo lo stato redox del RE sia mantenuto tale da permettere la formazione dei legami disolfuro (ossidazione), mantenendo contemporaneamente un potere riducente sufficiente per rompere i legami non corretti.

A lungo si è pensato che il glutatione ossidato fosse la fonte degli equivalenti ossidativi nel RE (Hwang et al., 1992). Nel RE, infatti, il glutatione è presente in un ambiente relativamente riducente (GSH/GSSG=3-5) (Hwang et al., 1992; Taniyama et al., 1990; Bass et al., 2004), ma altamente ossidante se paragonato al citoplasma (GSH/GSSG=100). Per questo motivo i saggi dell'attività catalitica della PDI *in vitro* sono stati effettuati usando il glutatione come tampone redox. Nuove evidenze hanno dimostrato non solo che la

formazione dei legami disolfuro procede indipendentemente dalla presenza di glutatione, ma anche che quest'ultimo nel RE funziona come riducente (Cuozzo e Kaiser, 1999).

Una combinazione di studi genetici e biochimici, inizialmente in *Saccharomyces cerevisiae* ed in seguito in mammiferi e piante, ha fatto luce sull'origine degli equivalenti ossidativi nel RE. Nel RE la PDI nella sua forma ridotta, dovuta all'ossidazione delle proteine nascenti, viene riossidata dall'azione di Ero1, una proteina di membrana FAD dipendente, che usa il potere ossidativo dell'ossigeno molecolare per rigenerare i propri ponti disolfuro intramolecolari (Sevier e Kaiser, 2006). Parallelamente alla via biosintetica maggiore, coinvolgente Ero1, esiste, specifica per i funghi, una via biosintetica minore coinvolgente una seconda proteina di membrana FAD dipendente, Erv2, capace di assicurare, se sovraespressa, la sopravvivenza di *S. cerevisiae* quando la funzione di Ero1 è compromessa (Sevier et al., 2001). In lievito è stato trovato Erv1 omologo di Erv2 localizzato nello spazio intermembrana dei mitocondri (Lange et al., 2001). Il fatto che Erv2, a differenza di Ero1, non sia un enzima essenziale e la sua assenza non abbia effetti visibili sull'ossidazione delle proteine suggerisce che si tratti di una via biosintetica minore o specializzata per la formazione di legami disolfuro in un gruppo ancora non identificato di proteine non essenziali (Gerber et al., 2001; Sevier et al., 2001; Tu e Weissman, 2002).

Il trasferimento degli equivalenti ossidativi, tra i substrati proteici e la PDI e tra questa ed Ero1, avviene attraverso una serie di scambi diretti, che comportano la formazione di legami disolfuro misti (Tu et al., 2000). La PDI, grazie alla sua solubilità, funziona da trasportatore di equivalenti ossidativi fra le proteine nascenti e le ossidoreduttasi di membrana Ero1 ed Erv2 (Sevier e Kaiser, 2006; Fig. 1.7).

Nei batteri, similmente, la proteina di membrana DsbB trasferisce alla proteina solubile DsbA (tioredossina simile) i legami disolfuro e quest'ultima è responsabile dell'ossidazione dei substrati proteici (Kadokura et al., 2003; Nakamoto e Bardwell, 2004; Tan e Bardwell, 2004). Anche nei batteri l'accettore finale di elettroni è l'ossigeno molecolare e, grazie all'ubiquinone, viene coinvolta la catena respiratoria (Bader et al., 1999 e 2000), che efficientemente riduce l'ossigeno molecolare formando acqua. Negli eucarioti invece, probabilmente anche per la separazione fisica dei due eventi, il RE per il folding ed i mitocondri per la respirazione, non è così ed Ero1 associato a FAD passa direttamente gli elettroni all'ossigeno molecolare. Ancora non è stato chiarito quale sia il destino dell'ossigeno molecolare consumato da Ero1, ma si ipotizza che l'attività di Ero1 potrebbe causare la formazione di forme reattive di ossigeno (ROS, reactive oxygen species), contribuendo allo stress ossidativo all'interno della cellula (Harding et al., 2003).

1.5.1.5. Il ruolo del glutatione

Si pone il problema di comprendere il ruolo del glutatione e la sua origine nel RE. Nei batteri la via biosintetica che porta all'ossidazione dei substrati proteici è distinta e separata da quella responsabile della riduzione ed isomerizzazione. Infatti nei batteri la proteina solubile DsbC, tioredossina simile come DsbA, viene ridotta da una proteina di membrana DsbD, che trae gli equivalenti riducenti dalla tioredossina riduttasi e dal NADPH (Sevier e Kaiser, 2006). DsbC si trova sotto forma di omodimero, condizione che ne impedisce l'ossidazione da parte di DsbB (Bader et al., 2001), ed è responsabile dell'isomerizzazione dei substrati proteici non nativi. Recentemente anche negli eucarioti sono emerse alcune evidenze di una via riducente separata, in cui il glutatione ridotto funge da agente riducente.

Le prime prove che il glutatione agisce da agente riducente all'interno del RE derivano da studi in lievito. Infatti abbassando il livello intracellulare di glutatione, intervenendo sul gene *GSH1* codificante per il primo enzima coinvolto nella sintesi del glutatione, si è visto che, in cellule di lievito contenenti un allele mutato di *ERO1*, ciò ripristina la formazione di legami disolfuro (Cuozzo e Kaiser, 1999). La formazione di glutatione ossidato è inoltre associata *in vivo* all'attività di Ero1 (Cuozzo e Kaiser, 1999), forse come risultato della riduzione da parte del glutatione ridotto dei ponti disolfuro della PDI e delle proteine ossidate (Tu et al., 2000). Ero1 infatti non può direttamente ossidare il glutatione (Tu et al., 2000). Il rapporto GSH/GSSG nel RE è quindi determinato dall'equilibrio tra l'attività ossidativa di Ero1 e quella riducente del glutatione (Fig. 1.7).

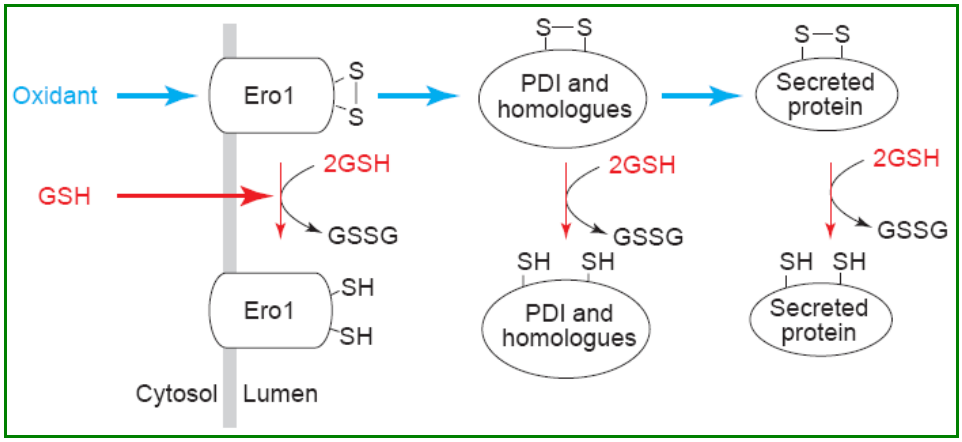

Fig. 1.7 – Equilibrio redox nel reticolo endoplasmatico (ER). L'ossidazione del glutatione compete con l'ossidazione delle proteine. Il glutatione ridotto (GSH), importato nel RE dal citosol, è coinvolto in reazioni di ossidoriduzione che portano alla netta riduzione delle proteine secrete ed alla netta ossidazione del glutatione. Il glutatione potrebbe interagire con le ossidoreduttasi del RE e con le proteine ripiegate. Il glutatione ridotto (GSH, in rosso) contribuisce con equivalenti riducenti che potrebbero essere importanti per mantenere le condizioni redox ottimali per il folding proteico controbilanciando il flusso di equivalenti ossidativi proveniente da Ero1 (frecce blu). GSSG, glutatione ossidato; (da Frand et al., 2000)

Inoltre cellule di mammifero private di glutatione, come il mutante di lievito *gsh1*, non mostrano difetti nel folding proteico, ma se queste cellule sono trattate con DTT (dithiothreitol) esse accumulano proteine con legami disolfuro incompleti (Chakravarthi e Bulleid, 2004) e aggregati proteici insolubili (Molteni et al., 2004), dimostrando che il glutatione protegge la cellula dall'accumulo di proteine mal ripiegate in condizione di stress cellulare dovuto alla presenza di un agente riducente. Non è ancora stato chiarito completamente se il glutatione ridotto influenzi la formazione di legami disolfuro nativi riducendo direttamente i substrati proteici o se la sua azione sia mediata da una o più proteine tioredossina- o PDI-like. Per ERp57, un membro della famiglia della PDI nell'uomo, è stato dimostrato il ruolo del glutatione nella sua riduzione, in quanto ERp57 non può ritornare nel suo stato ridotto in seguito a stress ossidativo in cellule private di glutatione (Jessop e Bulleid, 2004). D'altro canto il glutatione sembra essere in grado di ridurre direttamente i substrati proteici, in quanto gran parte della frazione di glutatione nel RE forma legami disolfuro misti con proteine (Bass et al., 2004).

La fonte di glutatione ridotto sembrerebbe essere il citoplasma (Fig. 1.7), dove la glutatione riduttasi deriva gli equivalenti riducenti dal NADPH (Sevier e Kaiser, 2006). È stata ipotizzata, quindi, la presenza di un sistema di trasporto preferenziale di glutatione ridotto, responsabile del flusso continuo di glutatione ridotto dal citosol al RE. L'elevato livello di glutatione ossidato all'interno del RE sarebbe invece conseguenza dell'accumulo di glutatione ridotto speso ed intrappolato nel RE (Sevier e Kaiser, 2006).

1.5.1.6. Il ruolo della PDI in vivo

Come già descritto in precedenza, *in vitro* la PDI è in grado di catalizzare, in funzione del substrato e dell'ambiente redox, sia la formazione che la rottura e il riarrangiamento dei legami disolfuro. Il suo ruolo *in vivo* è rimasto, invece, poco chiaro e quindi discusso e controverso.

Il ruolo *in vivo* della PDI nella formazione di legami disolfuro nativi nel RE fu indicato per la prima volta da analisi mutazionali in lievito, nel quale il gene *PDI1* risultò essenziale per la sopravvivenza della cellula e il folding proteico ossidativo (LaMantia et al., 1993; Scherens et al., 1991). Un mutante di Pdi1p, in cui sono assenti le cisteine dei siti attivi e privo di attività redox, non può complementare la delezione letale del gene *PDI1* (Laboissière et al, 1995). Appare chiaro che o l'attività di ossidasi o quella di isomerasi della PDI sono essenziali per la sopravvivenza e la normale crescita delle cellule di lievito.

In lievito era stato osservato che mutanti in cui la cisteina C-terminale dei siti attivi della PDI risultava mutata in una serina (CGH**S:**CGH**S**), condizione in cui la proteina, *in vitro,* perde la sua attività di ossidasi, conservando solo quella di isomerasi, erano in grado di garantire la sopravvivenza a cellule prive di Pdi1p. Era stato ipotizzato, quindi, che il ruolo principale della PDI fosse quello di riorganizzare i legami disolfuro non nativi (Laboissière et al., 1995). Mutazioni al sito attivo (CXXC) di Pdi1p, causavano una maggiore sensibilità dei mutanti al DTT, confermando il ruolo della PDI nella formazione dei ponti disolfuro *in vivo* (Holst et al., 1997). In realtà mutanti con una sola cisteina attiva per sito catalitico non solo mancano dell'attività ossidativa *in vitro*, ma anche la loro attività isomerasica risulta molto ridotta (Xiao et al., 2004). Questi mutanti, dotati solo di una certa attività isomerasica, garantiscono la sopravvivenza del lievito, ma solo se l'omologo della PDI in lievito Mpd2p è presente. Quindi Mpd2p, uno dei quattro omologhi non essenziali della PDI1 nel RE di lievito (Cherry et al,., 1997), diviene essenziale (Norgaard et al., 2001) nei mutanti della PDI in cui è presente solo l'attività isomerasica. Mpd2p è inoltre un substrato di Ero1 (Frand e Kaiser, 1999) che si trova prevalentemente in forma ossidata nel RE. Sulla base di questi elementi si può dedurre che Mpd2p sia dotato di attività ossidativa e sia in grado, quindi, di sopperire alla mancanza di questa funzione (Xiao et al., 2004). Inoltre è stato visto che l'espressione del solo dominio ***a'*** della *PDI1*, dotato della sola attività ossidativa, in ceppi di lievito *pdi1* garantisce la loro sopravvivenza e la normale crescita (Xiao et al., 2004), suggerendo che il ruolo essenziale della PDI, almeno in lievito, sia proprio la formazione dei legami disolfuro. Anche in ceppi di lievito *pdi1* in cui sono assenti tutti gli altri omologhi, l'espressione del dominio ***a'*** garantisce la sopravvivenza e la normale crescita del ceppo. In quest'ultimo caso, però, la maturazione di CPY (carboxypeptidase), un enzima vacuolare non essenziale, risulta compromessa, mentre l'espressione della full lenght *PDI1* in questi stessi ceppi sopperisce a questo difetto, dimostrando che la PDI, l'isomerasi più attiva *in vitro*, nella sua interezza anche *in vivo* può catalizzare l'isomerizzazione dei legami disolfuro (Xiao et al., 2004). Comunque anche uno o alcuni degli omologhi della PDI in lievito sembrano contribuire all'attività isomerasica, visto che l'espressione del dominio ***a'*** della PDI1, in presenza degli altri omologhi, non mostra alterazioni nella maturazione della CPY (Xiao et al., 2004).

La PDI *in vivo* può svolgere, quindi, sia attività di isomerasi che di ossidasi, rimane da stabilire quale sia la sua attività principale. Mentre per svolgere la sua funzione di isomerasi la PDI ha bisogno che i suoi siti attivi siano in forma ridotta, la formazione dei legami disolfuro implica che le cisteine nei siti attivi siano presenti in forma ossidata di disolfuro. In lievito la PDI, se fosse in equilibrio con il glutatione, dovrebbe essere ripartita equamente tra la sua

forma ossidata e ridotta (Frand et al., 2000). Misurazioni dirette dello stato redox *in vivo* della PDI nel RE hanno dimostrato che essa si trova invece principalmente nella sua forma ossidata (Frand e Kaiser, 1999). La spiegazione di questa apparente incongruenza è stata attribuita al fatto che la via ossidativa coinvolgente Ero1 è isolata dal resto dell'ambiente redox, in quanto il flusso di equivalenti ossidativi è controllato principalmente dalle cinetiche delle interazioni proteina-proteina, permettendo che lo stato redox di una proteina nel RE sia determinato in primo luogo dalla sua reattività relativa con altre proteine redox e possa quindi discostarsi significativamente da quello predetto sulla base della misurazione del potenziale redox in equilibrio con il glutatione (Frand et al., 2000).

La misurazione diretta dello stato redox dei diversi omologhi è stato quindi usata per ipotizzare la loro funzione preferenziale. Ad esempio la Pdi1p e la Mpd2 di lievito si trovano principalmente in forma ossidata e si pensa che la loro funzione primaria sia quella di catalizzare la formazione di legami disolfuro agendo da ossidasi, mentre alla maggior parte degli omologhi umani, che per lo più si trovano in forma ridotta, viene attribuita una funzione di riduttasi e isomerasi (Sevier e Kaiser, 2006). Anche la PDI umana si trova principalmente nella sua forma ridotta, ma è generalmente accettato il suo ruolo come ossidasi e il suo stato redox potrebbe riflettere semplicemente un trasferimento rapido dei legami ai substrati e/o una più lenta reazione di riossidazione della PDI da parte di Ero1 (Sevier e Kaiser, 2006). In realtà, da studi effettuati sulla PDI in lievito, la situazione risulta essere ben più complessa. Xiao e collaboratori (2004) hanno infatti ripetuto la misurazione diretta del potenziale redox della Pdi1p in lievito, trovando una prevalenza della sua forma ossidata, ma una parte (≈32%) dei suoi siti attivi si trova in condizione ridotta, in accordo con il suo duplice ruolo di isomerasi e ossidasi.

Kulp e collaboratori (2005) hanno usato un altro approccio per determinare le funzioni redox *in vivo* della PDI in lievito. Considerando che la funzione della PDI dipende dal suo potenziale redox, che a sua volta è il risultato della sua interazione con la ossidasi di membrana Ero1, è proprio studiando come quest'ultima e la PDI interagiscano che si può determinare il ruolo della PDI *in vivo.* L'uso di mutanti in cui le cisteine del sito attivo del dominio ***a*** o di quello ***a'*** erano state mutate in alanine, assieme a mutanti esprimenti uno di questi domini attivi solo o in combinazione lineare con uno o entrambi i domini di tipo ***b***, ha permesso di determinare che *in vivo* il dominio attivo C-terminale è un'ossidasi efficiente, ma manca dell'attività isomerasica a differenza del dominio N-terminale, che è una debole ossidasi, agendo principalmente come isomerasi (Kulp et al., 2005). La spiegazione di questo fatto si trova analizzando l'interazione della Ero1 con i due domini attivi. Infatti i due domini

catalitici della PDI full lenght, ossia in presenza degli altri domini tioredossina simili, sono ossidati da Ero1 in maniera indipendente e la velocità di ossidazione iniziale del dominio ***a'*** è doppia di quella del dominio ***a*** (Kulp et al., 2005). Questa disparità è ulteriormente accentuata dal fatto che in presenza di substrato l'interazione di questo con la PDI influisce negativamente sull'ossidazione del dominio ***a***, abbassandone il tasso di ossidazione al 40%. Entrambi questi effetti sono legati alla struttura in domini della PDI full lenght. In assenza del dominio ***a***, infatti, la velocità iniziale di ossidazione del dominio ***a'*** si dimezza, risultando paragonabile a quella del dominio ***a*** (Kulp et al., 2005) e anche l'effetto inibitorio del substrato sull'ossidazione del dominio ***a*** è presente solo se sono presenti anche i domini ***b'*** e ***a'*** (Kulp et al., 2005), i quali sono i principali responsabili del legame con il substrato (Klappa et al., 1998). Entrambi i domini di tipo ***a***, se espressi singolarmente, sono ossidati da Ero1 alla stessa maniera, in mancanza della struttura multidominio della PDI si perde, quindi, l'asimmetria nell'ossidazione, responsabile del fatto che il dominio ***a'*** è efficacemente ossidato da Ero1 e può quindi svolgere la funzione di ossidasi, mentre il dominio ***a*** si trova prevalentemente nella forma ridotta e quindi può fungere da isomerasi (Kulp et al., 2005).

La cinetica delle interazioni Ero1-PDI e l'asimmetria nelle interazioni cinetiche tra Ero1 e i domini attivi della PDI permettono che il potenziale redox dei diversi siti attivi risulti indipendente dall'ambiente redox e dal tampone glutatione. Ciò spiega come lo stato redox del RE venga mantenuto tale da permettere la formazione dei legami disolfuro (ossidazione), mantenendo contemporaneamente un potere riducente sufficiente per rompere i legami non corretti. Infatti siti attivi differenti all'interno di una stessa proteina e siti attivi di proteine diverse, ad esempio quelli dei diversi omologhi della PDI, possono quindi svolgere funzioni differenti e opposte all'interno di uno stesso ambiente redox.

È quindi essenziale analizzare a fondo l'interazione di Ero1, Erv2 e dei loro omologhi con la PDI e le proteine PDI-like per comprendere a fondo quale ruolo ognuno di questi enzimi svolga nella biosintesi proteica.

1.5.1.7. Gli omologhi della PDI e le vie biosintetiche alternative

Sembrerebbe quindi che, come nei batteri, negli eucarioti la formazione, la rottura e l'isomerizzazione dei ponti disolfuro possano avvenire attraverso vie biosintetiche distinte, coinvolgenti non solo la PDI ma anche i suoi omologhi. Ad oggi, come abbiamo visto, sono stati identificati nell'uomo 17 omologhi PDI (Ellgaard e Ruddock, 2005;), 2 paraloghi Ero1 (Cabibbo et al., 2000; Pagani et al., 2000) e tre proteine Erv2-simili (Hoober et al., 1999;

Lisowsky et al., 1995; Wittke et al., 2003) e in lievito 5 omologhi PDI (Scherens et al., 1991; Tachibana e Stevens, 1992; Tachikawa et al., 1995 e 1997; Wang e Chang, 1999), una proteina Ero1 (Frand e Kaiser, 1998; Pollard et al., 1998) e due proteine Erv2 simili (Gerber et al., 2001; Lee et al., 2000; Sevier et al., 2001; Stein e Lisowsky, 1998).

I diversi membri della famiglia di proteine PDI-like, come abbiamo visto, sono strutturalmente variabili, contengono infatti da zero a quattro domini di tipo ***a*** e comprendono, ad oggi, tre proteine umane ed una di lievito con un sito attivo modificato, in cui una delle cisteine del sito attivo risulta mutata in serina, e proteine contenenti solo domini di tipo ***b*** inattivi (Sevier e Kaiser, 2006). I diversi omologhi della PDI isolati in uomo hanno una capacità variabile nell'assicurare la sopravvivenza di *S. cerevisiae* quando la funzione della PDI è compromessa, suggerendo che tali proteine abbiano funzioni distinte (Sevier e Kaiser, 2006). È possibile che alle diverse strutture delle proteine PDI-like corrispondano funzioni diverse e che i diversi omologhi possano partecipare a reazioni redox preferenziali (di ossidazione, di riduzione o di isomerizzazione) o avere funzione di chaperoni molecolari. È stato proposto che in *S. cerevisiae* sia Ero1 che Erv2 abbiano una diversa affinità per i diversi omologhi della PDI e che questi, di conseguenza, potrebbero rappresentare vie ossidative indipendenti le una dalle altre. In lievito sono stati isolati solo i disolfuri misti tra Ero1 e PDI ed Ero1 e l'omologo Mpd2, mentre non è stato possibile individuare interazioni tra Ero1 e Mpd1 (Frand e Kaiser, 1999). Anche in uomo sono stati isolati solo alcuni disolfuri misti tra cui quello tra la Ero1-L umana e la PDI umana (Benham et al., 2000; Mezghrani et al., 2001) nonché tra Ero1-L e gli omologhi ERp44 (Anelli et al., 2002) e PDILT (van Lith et al., 2005), mentre non è stato possibile individuare interazioni con gli omologhi ERp57, P5 e ERp72 (Mezghrani et al., 2001).

L'incapacità di Ero1 di interagire con alcune delle proteine PDI-like suggerisce che essa sia capace di distinguere tra queste e la PDI, permettendo che le diverse proteine PDI-like, a seconda che siano riconosciute e quindi ossidate o meno, possano avere funzioni distinte tra loro partecipando a reazioni di ossidazione o di riduzione ed isomerizzazione (Tu e Weissman, 2004). Ad esempio, per ERp57 non solo non è stato individuato un legame disolfuro misto con Ero1-L, ma una sovra-espressione dell'ossidasi non altera lo stato redox ridotto di questo omologo (Mezghrani et al., 2001), per il quale invece è stata dimostrata la dipendenza dal glutatione ridotto (Jessop e Bulleid, 2004).

Alcuni geni omologhi di quello della PDI sono espressi preferenzialmente in alcuni tessuti, tra questi nell'uomo i geni codificanti per la PDIp, EndoPDI e la PDILT, ciò suggerisce che queste proteine contribuiscano al folding di substrati proteici tessuto specifici

(Desilva et al., 1997; Sullivan et al., 2003; van Lith et al., 2005). Alcune proteine omologhe interagiscono, inoltre, con altre proteine per mediare interazioni substrato specifiche. Ad esempio ERp57 interagisce con calnexina e calreticulina per promuovere la formazione di legami disolfuro in una classe di glicoproteine (Ellgaard e Frickel, 2003). È interessante notare che le regioni di legame per i peptidi nella PDI e i siti di legame per la calnexina e calreticulina nell'omologo ERp57 sono stati identificati e localizzati in regioni distinte dal dominio di tipo ***a*** (Ellgaard e Ruddock, 2005), suggerendo che le interazioni tra gli omologhi della PDI e i diversi substrati siano determinate, almeno in parte, da regioni distinte dal dominio tioredossinico attivo. Di conseguenza la fusione di un dominio di tipo ***a*** con una regione conferente un'interazione specifica con il substrato determinerebbe un numero illimitato di interazioni substrato specifiche (Sevier e Kaiser, 2006). Appare chiaro quanto complesso sia il meccanismo di formazione, riduzione e isomerizzazione dei legami disolfuro nel RE. L'idea che i diversi membri della famiglia PDI possono essere dedicati a vie biosintetiche separate e specifiche appare sempre più evidente.

1.5.2. Attività di chaperone molecolare

Da vari studi è emerso che la PDI presenta alcune caratteristiche tipiche di un chaperone molecolare (Puig e Gilbert, 1994a; Primm et al., 1996; Cai et al., 1994; Song e Wang, 1995). La PDI infatti è capace, quando è presente in concentrazioni stechiometriche, di impedire l'aggregazione dei substrati proteici nascenti o mal ripiegati durante il folding proteico. Questa attività non è solo evidente in substrati proteici contenenti residui cisteinici, dove potrebbe essere spiegata semplicemente dall'aumentata velocità di formazione dei ponti disolfuro corretti, che di per se limiterebbe la possibilità di aggregazione delle proteine (Goldberg ed al., 1991), ma è stata osservata anche nel refolding *in vitro* di proteine non contenenti ponti disolfuro, quali la fosfolipasi A_2 e la gliceraldeide 3–fosfato deidrogenasi (GAPDH) (Cai et al., 1994; Song et al., 1995; Primm et al., 1996; Yao et al., 1997). L'attività come chaperone molecolare della PDI è coerente con l'osservazione che l'accumulo di proteine mal ripiegate nel lume del RE di cellule di lievito e mammiferi può indurre l'espressione della PDI, come nel caso di chaperonine solubili nel RE, quali la BiP.

La funzione di chaperone molecolare viene spiegata dalla capacità della PDI di interagire con i peptidi, prevenendone così l'aggregazione attraverso la formazione di complessi solubili non-covalenti, che mascherano i possibili punti di aggregazione idrofobici (Quan et al., 1995). I peptidi possono legarsi alla PDI attraverso interazioni idrofobiche

(Klappa et al., 1997). L'azione di chaperone molecolare non richiede nessuna delle cisteine dei siti catalitici (Quan et al., 1995), mentre la perdita dei 51 aminoacidi C-terminali della PDI causa la perdita di questa funzione, suggerendo che un sito di legame per il substrato necessario all'attività di chaperone sia localizzato tra i domini ***a'*** e ***c*** (Dai e Wang, 1997). Come per l'attività redox isomerasica della PDI, anche l'attività di chaperone molecolare si basa sulla capacità della PDI di interagire con peptidi. Il grado con cui nel folding si esplica l'attività di chaperone molecolare o quella di isomerasi, e quindi il coinvolgimento dei diversi domini, può essere substrato dipendente, come osservato in studi condotti in vitro con diversi substrati proteici (Puig et al., 1994; Hayano et al., 1996; Yao et al,, 1997; Song et al., 1997). Una forma mutata di PDI, in cui è stato eliminato il sito che presiede all'attività isomerasica è ancora in grado di favorire il folding delle proteine (Hayano et al., 1996). In ogni caso l'interazione con i peptidi sembra competere con l'attività catalitica, dimostrando che i siti di legame per i peptidi sono uguali o vicini a quelli in cui risiede questa attività (Morjana e Gilbert, 1991; Quan et al., 1995). L'affinità di legame sembra aumentare con la lunghezza delle catene peptidiche, con la presenza di residui cisteinici e di residui idrofobici (Morjana e Gilbert, 1991; Klappa et al., 1997).

A concentrazioni di 1/5-1/10 di quelle del substrato la PDI ha un comportamento inusuale, definito anti-chaperone (Puig e Gilbert, 1994a), precipitando varie proteine mal-ripiegate, normalmente solubili, inducendone l'aggregazione (Sideraki e Gilbert, 2000). La PDI risulta integrata in questi aggregati assieme alle proteine non-ripiegate, mentre proteine correttamente ripiegate di controllo non sono integrate (Puig e Gilbert, 1994b). Anche la BiP mostra lo stesso comportamento anti-chaperone (Puig e Gilbert, 1994b). Questa attività della PDI può essere spiegata dalle interazioni multiple tra la PDI e i substrati che provocano la formazione di grandi aggregati, causandone la precipitazione (Sideraki e Gilbert, 2000).

Proteine non completamente ripiegate o mal ripiegate sono trattenute nel RE attraverso due meccanismi. Le proteine con attività di chaperone molecolare, grazie alla loro sequenza di ritenzione KDEL, sono trattenute nel RE, assieme al loro carico di proteine non ripiegate, attraverso un sistema che dal Golgi le rimanda al RE (Pelham, 1988). Le proteine, una volta che sono state correttamente formate, non sono più trattenute nel RE, in quanto non interagiscono con i chaperoni molecolari (Hammond e Helenius, 1994). Il secondo meccanismo che impedisce la secrezione di proteine non native si basa sulla formazione di grandi aggregati di proteine mal ripiegate associate ai chaperoni molecolari. Generalmente le interazioni idrofobiche che causano l'aggregazione e la precipitazione sono tra i fattori principali che riducono la produttività del folding (Goldberg et al.,1991), esistono tuttavia

meccanismi che recuperano le proteine dagli aggregati, almeno nelle fasi iniziali di aggregazione, rendendoli disponibili per il folding e la secrezione (Lodish e Kong, 1993; Verde et al., 1995). L'attività di anti-chaperone della PDI (Whiteley et al., 1997) e di altre proteine (Smith et al., 1995) potrebbe contribuire alla formazione di aggregati ed in seguito alla loro utilizzazione come fonte di proteine secrete.

1.5.3. Il ruolo della PDI come subunità di complessi enzimatici

In aggiunta al suo ruolo primario nella formazione dei legami disolfuro ed alla sua attività come chaperone molecolare, la PDI nei mammiferi possiede due funzioni ben note come subunità di sistemi enzimatici complessi presenti nel RE.

Studi condotti sulle cellule di mammifero hanno dimostrato che ciascuna delle subunità che costituiscono la struttura omodimerica della proteina PDI può fungere da subunità β della prolil-4 idrossilasi (P4H) (Pihlajaniemi, et al., 1987), un enzima solubile residente nel RE. La P4H catalizza l'idrossilazione dei residui di prolina in alcune proteine, come nelle catene pro-α del procollagene e in molte proteine della membrana cellulare. L' enzima è una proteina eterotetramerica (α2β2), composta da polipeptidi α e β, dei quali la subunità α possiede l'attività catalitica. La funzione della PDI come subunità β è quella di stabilizzare e trattenere nel RE il complesso enzimatico, senza partecipare all'attività catalitica (Wilkinson e Gilbert, 2004); da studi condotti *in vitro* è infatti emerso che potrebbe prevenire il cattivo ripiegamento e l'aggregazione delle subunità α, stabilizzando la conformazione ripiegata delle catene polipeptidiche (John et al., 1993).

La PDI costituisce inoltre la subunità β del complesso microsomico di trasporto dei trigliceridi (MTP) (Wetterau et al., 1990), che risulta costituito da un dimero αβ . Questo complesso catalizza l'incorporazione dei triacilgliceroli nelle lipoproteine nascenti nel fegato e nelle cellule intestinali. La subunità β sembra necessaria per mantenere il complesso in soluzione, dal momento che la sua rimozione conduce alla formazione di aggregati insolubili, costituiti dalle subunità α, impedendo l'attività di trasferimento dei triacilgliceroli (Wetterau et al., 1991). Probabilmente la subunità β è importante per mantenere la subunità α in una conformazione non aggregata, necessaria per l'esplicazione della sua attività.

1.5.4. PDI e proteine sue omologhe in compartimenti cellulari diversi dal RE

La PDI e tutte le proteine omologhe sono presenti nel RE, ma almeno tre proteine della famiglia, PDI, Erp57 e ERp72, sono state trovate anche in compartimenti cellulari diversi dal RE, nonostante tutte abbiano il segnale di ritenzione per il RE. Il meccanismo con il quale la PDI ed i suoi omologhi, come anche altri tipi di proteine localizzate prevalentemente nel RE, come la calreticulina, sfuggano al meccanismo di ritenzione nel RE è tuttora sconosciuto. Mentre per la calreticulina è stata non solo dimostrata la sua secrezione, ma anche la sua localizzazione sulla superficie cellulare, nel citoplasma e nel nucleo (Roderick et al., 1997; Holaska et al., 2001; Johnson et al., 2001), la presenza di proteine appartenenti alla famiglia della PDI in compartimenti extra-RE viene spesso messa in discussione. Infatti se la localizzazione è basata su esperimenti di separazione fisica di compartimenti sub-cellulari è molto difficile escludere eventuali eventi di contaminazione.

1.5.4.1. Localizzazione extracellulare

Proteine della famiglia della PDI sono state trovate fuori dalla cellula sia come proteine secrete che come proteine localizzate sulla superficie cellulare.

La PDI classica, nonostante il suo segnale di ritenzione C-terminale KDEL, è secreta a bassi livelli da diversi tipi cellulari, quali le cellule epatiche (Terada et al., 1995), le cellule esocrine del pancreas (Yoshimori et al., 1990), le cellule dell'endotelio (Hotchkiss et al., 1998) e le piastrine (Chen et al., 1992). Anche per Erp57 (Hirano et al., 1995a) e ERp72 (Dorner et al., 1990), in condizioni di sovra-espressione, è stata osservata la secrezione da cellule in coltura.

Per la maggior parte di queste proteine secrete il ruolo biologico è rimasto sconosciuto, ma Delom e collaboratori (2001) hanno chiaramente dimostrato la funzione della PDI secreta dalle cellule dell'endotelio nel lume dei follicoli della tiroide. La PDI, in cooperazione con la BiP, è coinvolta nel controllo del folding e della multimerizzazione della tiroglobulina, probabilmente riducendo i legami disolfuro intermolecolari e quindi limitando il livello della formazione dei complessi multimerici.

La PDI è stata trovata anche localizzata sulla superficie di vari tipi cellulari. È stato suggerito che la PDI, dopo essere stata secreta, si leghi alla superficie cellulare attraverso interazioni elettrostatiche (Terada et al., 1995), anche perchè la PDI è una proteina solubile ed una sua inserzione nella membrana sarebbe poco probabile. Alla PDI localizzata sulla

superficie esterna della membrana plasmatica sono stati assegnati diversi ruoli. I siti tioredossinici della PDI sembrano coinvolti nell'attività riducente dell'ambiente esterno alla cellula (Mandel et al., 1993), catalizzando la riduzione o l'isomerizzazione dei legami disolfuro presenti nelle macromolecole che vengono in contatto con la superficie esterna della membrana plasmatica o che sono in essa presenti. Il livello dei gruppi tiolici (SH) presenti sulla superficie esterna della membrana plasmatica nei linfociti e nelle cellule del fibrosarcoma è positivamente correlato con la presenza della PDI (Lawrence et al., 1996; Jiang et al., 1999).

Differenti studi hanno dimostrato che i gruppi tiolici delle proteine localizzate sulla superficie esterna della membrana plasmatica sono coinvolti nei processi di adesione cellulare. Nella maggior parte dei casi il ruolo della PDI è indiretto, avvenendo infatti attraverso la riduzione e/o riorganizzazione dei legami disolfuro di proteine direttamente coinvolte, e spesso incerto. In alcuni casi, invece, è stato dimostrato un ruolo diretto della PDI. Un chiaro ed inequivocabile esempio dell'intervento diretto della PDI nell'adesione cellulare è riportato per le cellula della retina dell'embrione di pollo (Hausman e Moscona, 1975), dove una proteina localizzata sulla superficie esterna della membrana plasmatica, denominata R-cognin, in seguito risultò essere identica alla PDI (Krishna Rao e Hausman, 1993; Pariser et al, 2000) e i suoi siti tioredossina simili attivi necessari per le proprietà di adesione.

Un ruolo indiretto della PDI è stato recentemente dimostrato nella regolazione del processo di adesione dei leucociti (Bennett et al., 2000), controllato da una proteina presente sulla superficie della membrana plasmatica chiamata L-selectina. L'attività riducente della PDI mantiene questa proteina in una particolare conformazione, che la rende inaccessibile all'azione degli enzimi proteolitici. In particolari condizioni (attivazione dei leucociti) avviene un cambiamento conformazionale, con il conseguente taglio del dominio extra-cellulare della proteina e la perdita delle proprietà adesive dei leucociti. Lo stesso effetto viene provocato da agenti chimici che inibiscono l'attività ossido-riducente della PDI o di anticorpi anti-PDI (Bennett et al., 2000).

La presenza della PDI è stata anche dimostrata sulla superficie esterna delle piastrine, la cui attivazione è accompagnata da un aumento di gruppi tiolici sulla superficie della membrana plasmatica, in particolare da un aumento di siti attivi della PDI in forma ridotta (Burgess et al., 2000). I tioli sulla superficie delle piastrine sono coinvolti nell'adesione cellulare mediata dalla proteina integrina, la cui forma attiva richiede la riduzione e/o il riarrangiamento dei legami disolfuro.

Nonostante molte altre funzioni della PDI localizzata sulla superficie esterna della membrana cellulare debbono essere ancora chiarite, in letteratura sono riportati alcuni interessanti esempi del coinvolgimento della PDI in fenomeni patologici. La PDI localizzata sulla superficie esterna della cellula è coinvolta, ad esempio, nella riduzione dei ponti disolfuro della tossina difterica, che in questo modo acquisisce la capacità di entrare nella cellula ed esercitare la sua attività citotossica (Mandel et al., 1993). Anche il potenziale infettivo di numerosi virus sembra dipendere dall'attività riducente presente sulla superficie esterna della cellula. In particolare, nel caso del virus HIV (Ryser et al., 1994), la cui infettività è strettamente legata alla riduzione dei legami disolfuro della glicoproteina di rivestimento virale gp120 (Ryser et al., 1994; Fenouillet et al., 2001), è stato accertato che l'agente riducente è proprio la PDI. L'utilizzo di agenti inibenti l'attività specifica ossido-riducente della PDI o di anticorpi anti-PDI impedisce l'entrata del virus HIV in cellule linfoidi umane.

La PDI non è l'unico membro della famiglia ad essere stato trovati sulla superficie cellulare, sono stati trovati infatti anche ERp72, sulla membrana di leucociti neutrofili (Weisbart, 1992), ed ERp57, tra le proteine della membrana spermatica (Bohring et al., 2001).

1.5.4.2. Localizzazione e funzione delle proteine PDI-like nel citoplasma

La PDI è stata trovata nel citoplasma ed in particolare nella frazione citosolica delle cellule epatiche umane e di scimmia. È stato proposto che la PDI agisca in cooperazione con l'enzima IDE, responsabile della degradazione dell'insulina, catalizzando la rottura dei legami disolfuro e favorendo quindi la proteolisi (Wroblewski et al., 1992).

Poco dopo la sua caratterizzazione anche la ERp57 fu trovata essere presente in notevoli quantità non solo nel RE, ma anche nel citosol (Lewis et al., 1986). La sua presenza nel citosol, inizialmente basata su esperimenti di separazione fisica dei compartimenti subcellulari, è stata dimostrata definitivamente dalla sua associazione con proteine citosoliche. Ad esempio con la STAT3, un membro della famiglia delle proteine STAT coinvolte nella trasduzione del segnale e nella regolazione della trascrizione. Questa proteina, in particolare, è stata trovata complessata nel citoplasma delle cellule epatiche insieme ad altre proteine, tra cui la ERp57 (Ndubuisi et al., 1999). Anche se il ruolo di questa associazione non è stato ancora chiarito, si è ipotizzato che possa essere importante per la penetrazione della proteina STAT3 nel nucleo. Tutte le proteine del citoplasma possono essere importate nel nucleo se in possesso di una sequenza aminoacidica di riconoscimento

definita segnale di localizzazione nucleare. La proteina ERp57, a differenza della STAT3 è dotata di questo segnale (Turano et al., 2002).

1.5.4.3. Localizzazione e funzione delle proteine PDI-like nel nucleo

Sulla base di due tipi di evidenze sperimentali almeno due membri della famiglia della PDI, PDI e ERp57, sono presenti nel nucleo. Tali evidenze consistono nella determinazione diretta della presenza di queste proteine nel compartimento nucleare e nella dimostrazione *in vitro* che esse possono influenzare processi che avvengono a livello di DNA, ad esempio la capacità dei fattori di trascrizione e delle proteine strutturali del nucleo di legarsi al DNA. Questi risultati vanno comunque presi con precauzione. In primo luogo perchè entrambe le proteine sono risultate essere presenti ad un livello molto basso nel nucleo, e ciò rende concreto il rischio che si possa trattare di contaminazione di PDI proveniente dal RE, ed in secondo luogo il fatto che *in vitro* queste proteine siano coinvolte in processi a livello di DNA non dimostra che ciò avvenga anche *in vivo*.

ERp57 è stata isolata dalla matrice nucleare interna di nuclei purificati da cellule di fegato di gallina (Altieri et al., 1993). PDI ed ERp57 sono state trovate, inoltre, tra le proteine della matrice nucleare di linfociti e monociti umani (Gerner et al., 1999). La prima prova sicura della localizzazione nucleare di ERp57 è stata data da esperimenti di immunoistochimica su cellule intere. In questo modo ERp57 è stata localizzata nei nuclei di alcuni tipi di cellule, ad esempio nei nuclei di cellule HeLa (Coppari et al., 2002). Un'altra prova sicura è il ritrovamento di queste proteine in complessi con il DNA. L'uso della radiazione UV e di *cis*-diamminodicloroplatino ha dimostrato non solo la presenza di ERp57 nel nucleo, ma anche la sua interazione con il DNA. Recentemente VanderWaal et al. (2002) hanno dimostrato che anche la PDI è presente in complessi con il DNA in nuclei isolati da cellule HeLa.

ERp57 inoltre è capace di legarsi al DNA, riconoscendo preferenzialmente sequenze nucleotidiche ricche in A/T e MAR-like (Ferraro et al., 1999). Queste ultime sono regioni associate alla matrice nucleare, che sono alla base dei domini ad anello della cromatina in cui sono organizzati i genomi degli eucarioti e forniscono la base di ancoraggio del DNA alla matrice (Berezney, 1991). Queste MAR sono attive nell'organizzazione strutturale del genoma e possono facilitare, grazie all'interazione con specifiche proteine, la trascrizione di un gene o di un gruppo di geni, promuovendo la formazione di strutture meno condensate della cromatina. Una interessante caratteristica del legame DNA-ERp57 è che questo dipende

dallo stato redox delle proteine, in quanto solo la forma ossidata è capace di legarsi al DNA mentre la sua forma ridotta perde qualsiasi affinità per il DNA. Nonostante non sia stato possibile ad oggi assegnare uno specifico ruolo alle PDI nucleari sono state fatte alcune ipotesi sulla loro potenziale funzione. Considerando infatti che l'attività di legame al DNA di alcuni fattori di trascrizione dipende dallo stato redox in cui si trovano, essendo nella maggior parte dei casi la forma ridotta quella attiva, e che *in vitro* sia la PDI che la ERp57 possono influenzare processi che avvengono a livello di DNA, un loro intervento *in vivo* è plausibile (Turano et al., 2002). Inoltre è stato suggerito che PDI e ERp57 contribuiscano all'ancoraggio degli anelli del DNA alla matrice nucleare. In questo caso la capacità di ERp57 di legarsi al DNA solo nella sua forma ossidata potrebbe rappresentare un meccanismo per l'interazione del DNA con la matrice nucleare modulato dallo stato redox di queste proteine, il quale potrebbe avere importanti conseguenze nella regolazione della trascrizione (Stein et al., 1995). A differenza di ERp57, la PDI non possiede un segnale di localizzazione nucleare. Anche se non è stato ancora chiarito come la PDI possa venire quindi importata nel nucleo è stato ipotizzato che questo avvenga grazie alla sua tendenza ad associarsi con altre proteine, che quindi potrebbero fungere da trasportatori (Turano et al., 2002).

1.5.4.4. Localizzazione di una proteina omologa della PDI nel cloroplasto

RB60 è una proteina PDI-like localizzata nei cloroplasti di *Chlamydomonas reinhardtii*, dove è coinvolta, come parte di un complesso proteico, nel controllo della traduzione indotta dalla luce degli mRNA di cloroplasto attraverso la sua attività redox. La struttura multidominio di questa proteina, ***c*-*a*-*b*-*b'*-*a***, ricorda quella della proteina hERp72, in quanto entrambe contengono un gruppo N-terminale acidico, ad RB60 manca però il dominio ***a***° addizionale. Inizialmente si riteneva che RB60 fosse una PDI cloroplasto specifica delle piante (Kim e Mayfield, 1997), in seguito è stato dimostrato che RB60 è presente sia nel cloroplasto che nel RE (Levitan et al., 2005). Questi risultati sono in accordo con la presenza del segnale di ritenzione nel RE (KDEL) all'estremità C-terminale della proteina. RB60 presenta una sequenza N-terminale lunga 50 aminoacidi, sufficiente a garantire la localizzazione di RB60 sia nel cloroplasto che nel RE. Questa sequenza di 50 amminoacidi contiene un potenziale sito di taglio dopo il 28simo residuo aminoacidico, seguito da un'ulteriore regione di 22 aminoacidi ricca di residui carichi negativamente. È stato ipotizzato che si possa formare un dominio transmembrana ad α-elica nella zona N-terminale che precede il dominio acidico. Ci sono evidenze che tale dominio transmembrana venga tagliato

in seguito all'importazione cotraduzionale della proteina nel lume del reticolo endoplasmatico e che invece rimanga intatto in seguito all'importazione post-traduzionale della proteina nel cloroplasto. In particolare i primi 28 aa sembrano sufficienti affinché la proteina venga importata nel cloroplasto, ma per la sua associazione con le membrane dei tilacoidi sembra necessario anche il dominio N-terminale acidico. Nel cloroplasto RB60 ha la funzione di sensore redox in un complesso proteico che lega l'mRNA. Questo complesso, che si lega alla regione 5' non tradotta dell'mRNA di *psbA,* si pensa sia coinvolto nella regolazione della traduzione dell'ORF a valle, in risposta a segnali fotosintetici e luminosi. In particolare, RB60 modula la traduzione di mRNA di *psbA* in risposta alla luce o al buio. La riduzione dei legami disolfuro di RB60, che è proporzionale all'intensità di luce a cui i cloroplasti sono esposti, stimola la formazione del complesso proteico, che invece è ostacolata dalla fosforilazione ADP dipendente di RB60, che si verifica a concentrazioni di ADP nei cloroplasti determinate dal buio (Levitan et al., 2005).

1.6. La famiglia delle PDI nelle piante

A differenza dei numerosi studi condotti nei vertebrati, ed in particolare nell'uomo, sulla complessità e diversità strutturale dei membri della famiglia della PDI e sulle loro funzioni, si conosce meno sulle caratteristiche molecolari e strutturali delle PDI delle piante, sulla regolazione della loro espressione durante lo sviluppo della pianta e sul loro ruolo fisiologico.

Inizialmente nelle piante le conoscenze riguardanti le proteine appartenenti alla famiglia della PDI erano limitate a qualche sequenza cDNA isolata e caratterizzata in alcune specie come erba medica (Shorrosh e Dixon, 1991; 1992), orzo (Chen e Hayes, 1994), frumento (Shimoni et al, 1995b, Ciaffi et al., 2001), mais (Li and Larkins, 1996;), ricino (Coughlan et al., 1996) e carota (Xu et al., 2002). Per lo più queste sequenze riguardavano la forma classica della PDI, mentre solo alcune codificavano per proteine PDI-like (Shorrosh e Dixon, 1992; Xu et al., 2002).

Tutti i membri della super-famiglia della tioredossina venivano attribuiti ai differenti gruppi funzionali in base a dimensione, numero e organizzazione dei domini tioredossinici, sequenza del tetrapeptide al sito attivo, localizzazione sub-cellulare e attività enzimatica. La crescente disponibilità di sequenze genomiche ed EST ha determinato la possibilità di cercare

e identificare un gran numero di sequenze omologhe tramite BLAST, rendendo sempre più aleatori questi raggruppamenti funzionali.

Houston et al. (2005), nello sforzo di creare una classificazione unica ed esauriente di proteine contenenti domini catalitici tioredossina-simili in *A. thaliana*, da usare per individuare proteine PDI-like in altre specie vegetali, hanno estratto dalle banche dati contenenti sequenze genomiche ed EST di *A. thaliana* tutte le sequenze codificanti per domini tioredossina simili. Da 104 sequenze di *A. thaliana* sono stati identificati 117 singoli domini tioredossina simili, la cui analisi filogenetica ha permesso di classificare la grande super-famiglia della tioredossina nelle sue varie sotto-famiglie, glutaredossine, ecc., e di individuare quindi tutti i potenziali membri codificanti per proteine PDI-like.

Nel loro lavoro si sono basati sulla suddivisione in 4 classi strutturali, basate sul numero e sulla posizione dei siti tioredossinici attivi, dei membri della famiglia della PDI dei mammiferi, la quale è attualmente la meglio caratterizzata (Kanai et al., 1998).

Classe 1: proteine dotate di due siti attivi, il primo nella regione N-terminale ed il secondo nella regione C-terminale.
Classe 2: proteine con due domini disposti in sequenza nella regione N-terminale.
Classe 3: proteine che hanno, oltre ai domini della classe 1, un dominio attivo aggiuntivo nella regione N-terminale.
Classe 4: proteine con tre domini attivi disposti in sequenza nella regione N o C terminale.

Gli organismi unicellulari sembrano avere una famiglia PDI-like meno complessa di quella presente nei mammiferi. In *S. cerevisiae* ci sono solo 5 membri: l'omologo della PDI classica (Pdi1p), essenziale per la sopravvivenza, e 4 proteine PDI-like (Eug1p, Mpdi1p, Mpdi2p e Eps1p) (Norgaard et al., 2001; Freedman et al., 2002). Come *S. cerevisiae,* anche *Giardia lamblia* contiene 5 proteine PDI-like (McArthur et al., 2001). Tutte le proteine PDI-like in *G. lamblia* ed alcune in *S. cerevisiae* contengono un solo sito tioredossinico attivo. Tutte le proteine contenenti un singolo sito attivo vengono raggruppate in una quinta classe strutturale da Houston et al. (2005).

Classe 5: proteine con un solo sito tioredossinico.

Sono stati così individuati 13 geni PDI-like in *A. thaliana* distribuiti sui suoi 5 cromosomi (Houston et al., 2005; Tab. 1.3). Una ricerca in banche dati pubbliche e private

usando le sequenze individuate in *A. thaliana* ha permesso l'individuazione di ulteriori sequenze PDI-like in mais ed in riso (Tab. 1.3). L'analisi filogenetica, basata sulle sequenze aminoacidiche e nucleotidiche dei soli domini tioredossina simili, di questi geni ha rilevato la presenza di VIII classi filogenetiche distinte. Tutti i geni presentano uno o due siti tioredossina simili e ricadono, quindi, in 3 distinte classi strutturali (1, 2 e 5). In particolare alla classe strutturale 1, che è la più estesa, appartengono i gruppi filogenetici I-III, alla classe 2 quelli IV-V, mentre i restanti gruppi filogenetici VI-VIII appartengono alla classe strutturale 5, caratterizzata dalla presenza di un solo sito tioredossinico.

Le proteine dei gruppi filogenetici I, II e III sono simili in lunghezza (500-560 aa) e diverse analisi *in silico* suggeriscono che si tratti di proteine secretorie contenenti sia il peptide segnale che la sequenza KDEL di ritenzione nel RE (Houston et al., 2005). Le proteine del gruppo IV hanno una lunghezza approssimativa di 360 aa e sono prive del segnale di ritenzione KDEL, a differenza dei membri del gruppo V, lunghi circa 440 aa, che lo presentano nella regione C-terminale.

Le restanti proteine, appartenenti tutte al gruppo strutturale 5 in quanto dotate di un solo dominio tioredossina simile, hanno poche caratteristiche comuni, confermando la loro distinta origine filogenetica. Le proteine del gruppo filogenetico VI, approssimativamente lunghe 150 aa, sono risultate le più piccole tra le proteine isolate, mentre i membri dei gruppi VII e VIII avevano una dimensione maggiore (418-485 aa). Le proteine PDI-like dei gruppi VI e VII probabilmente sono proteine secretorie, dotate di un peptide segnale. Nessuna delle proteine con un singolo dominio tioredossina simile presenta un segnale di ritenzione KDEL-simile, solo una proteina del gruppo VII contiene una sequenza C-terminale KKED di ancoraggio alla membrana. Le proteine umane TMX2 e TMX3 hanno una sequenza C-terminale KKDK e KKKD, riconducibile alla sequenza consenso KKXX per il richiamo nel RE di proteine transmembrana di tipo I. Da analisi con i programmi TMpred e TMHMM, le proteine dei gruppi VII e VIII sono risultate potenziali proteine di membrana (Houston et al., 2005). I membri di questi gruppi filogenetici hanno sequenze C-terminali la cui potenziale funzione è di domini transmembrana, che si potrebbero ancorare al RE (Wang e Chang, 1999; Matsuo et al., 2001).

L'analisi filogenetica ha inoltre indicato che i diversi gruppi filogenetici potrebbero essere emersi per duplicazione di geni (I-III), per duplicazione di domini (V) e forse per delezione di domini (VI-VIII) in una fase iniziale dell'evoluzione eucariotica. Inoltre ci sono evidenze di duplicazioni più recenti avvenute in mais, riso e *A. thaliana,* vista la presenza di più sequenze appartenenti allo stesso gruppo filogenetico in una stessa specie.

Tab. 1.3 – Caratteristiche delle proteine PDI-like di *Arabidopsis*, riso e mais (modificato da Houston et al., 2005).

Gruppo Filogenetico	PDIL	Dimensione del polipeptide dedotto	N. di accessione del gene	No. di accessione del cDNA	Classe strutturale	Dominio TRX	Tetrapeptide C-terminale
I	AtPDIL1-1	501	**At1g21750**	NM_102024	1	2	KDEL
I	AtPDIL1-2	508	**At1g77510**	NM_106400	1	2	KDEL
I	OsPDIL1-1	512		AK068268	1	2	KDEL
I	OsPDIL1-2	517		AY739308	1	2	KDEL
I	OsPDIL1-3	492		AK073161	1	2	KDEL
I	ZmPDIL1-1	514		AY739284	1	2	KDEL
I	ZmPDIL1-2	512		AY739285	1	2	KDEL
II	AtPDIL1-3	579	**At3g54960**	NM_115353	1	2	KDEL
II	AtPDIL1-4	597	**At5g60640**	NM_180903	1	2	KDEL
II	OsPDIL1-4	563		AK071514	1	2	KDEL
II	ZmPDIL1-3	568		AY739286	1	2	KDEL
II	ZmPDIL1-4	561		AY739287	1	2	KDEL
III	AtPDIL1-5	537	**At1g52260**	NM_104105	1	2	KDEL
III	AtPDIL1-6	534	**At3g16110**	NM_112481	1	2	KDEL
III	OsPDIL1-5	533		AK073970	1	2	KDEL
III	ZmPDIL1-5	529		AY739295	1	2	KDEL
IV	AtPDIL2-1	361	**At2g47470**	NM_130315	2	2	VASS*,^{o}
IV	OsPDIL2-1	366		NM_185280	2	2	TFSS*,^{o}
IV	OsPDIL2-2	371		NM_183927	2	2	TFSS*,^{o}
IV	ZmPDIL2-1	367		AY739288	2	2	TFSS*,^{o}
IV	ZmPDIL2-2	366		AY739289	2	2	IFSS*,^{o}
V	AtPDIL2-2	443	**At1g04980**	NM_100376	2	2	KDDL
V	AtPDIL2-3	440	**At2g32920**	NM_128852	2	2	KDEL
V	OsPDIL2-3	441		AK062254	2	2	NDEL
V	ZmPDIL2-3	439		AY739290	2	2	NDEL
VI	AtPDIL5-1	146	**At1g07960**	NM_202059	5	1	DKEL
VI	OsPDIL5-1	147		AK063663	5	1	LQDS*
VI	ZmPDIL5-1	150		AY739291	5	1	LEAD*
VII	AtPDIL5-2	440	**At1g35620**	NM_103262	5	1	KKED
VII	OsPDIL5-2	423		AK069367	5	1	AHEE*
VII	OsPDIL5-3	425		ND	5	1	AHED*
VII	ZmPDIL5-2	418		AY739292	5	1	IHEE*
VII	ZmPDIL5-3	420		AY739293	5	1	IHEE*
VIII	AtPDIL5-3	483	**At3g20560**	NM_112948	5	1	GKNI*
VIII	AtPDIL5-4	480	**At4g27080**	NM_118842	5	1	GKNF*
VIII	OsPDIL5-4	485		AK099660	5	1	GKNI*
VIII	ZmPDIL5-4	483		AY739294	5	1	GKNI*

ND: non è disponibile la sequenza cDNA; una ricerca con tBLASTn di Gramene ha permesso di identificare il trascritto dedotto GRMT00000163510 (OsPDIL5-3). *: non ci sono evidenze per il coinvolgimento del tetrapeptide nella ritenzione nel RE. ^{o}: il polipeptide dedotto ha una dominio COOH-terminale coinvolto nella ritenzione nel RE (Monnat et al., 2000).

La presenza nelle diverse specie vegetali dei principali gruppi di ortologhi avvalora l'ipotesi che la duplicazione dei domini tioredossinici sia avvenuta prima della divergenza dei principali gruppi tassonomici vegetali, contribuendo all'evoluzione di questa complessa famiglia genica a partire da un comune eucariote ancestrale (McArthur et al., 2001). Il fatto che le sequenze dei geni ortologhi si siano conservate nel tempo suggerisce che i differenti gruppi di proteine PDI-like presentino funzioni distinte ed importanti nella cellula e che il loro mantenimento sia stato necessario durante l'evoluzione (Houston et la., 2005).

Lemaire e Miginiac-Maslow (2004) avevano identificato 9 dei 13 geni di *A. thaliana* ed avevano identificato gli ortologhi di otto di essi nei mammiferi, includendo nelle loro analisi anche sequenze di uomo e di *Dictyostelium discoideum.* I due geni di *A. thaliana* assegnati al gruppo filogenetico I (numeri di accessione **At1g21750** e **At1g77510**), corrispondono in struttura ad hERp57 (h=umano). Essi hanno una struttura di tipo ***a-b-b'-a'*** a cui manca il dominio ***c*** terminale. Anche la PDI classica in uomo, che presenta una struttura multidominio del tipo ***a-b-b'-a'-c***, fa parte del gruppo filogenetico I. I due geni di *A. thaliana* del gruppo filogenetico II (numeri di accessione **At5g60640** e **At3g54960**), vengono raggruppati assieme ad hERp72 e Rb60 di *C. reinhardtii*. hERp72 lo ricordiamo ha una struttura del tipo ***c-a°-a-b-b'-a'***, mentre Rb60 e i due geni di *A. thaliana*, pur contenendo un gruppo N-terminale acidico, mancano del dominio ***a°*** addizionale, presentando, quindi, la seguente struttura multidominio: ***c-a-b-b'-a'***. La presenza di ERp72 in questo gruppo potrebbe essere dovuta alla sola presenza di un dominio di tipo ***c*** nella regione N-terminale (Lemaire e Miginiac-Maslow, 2004).

Le due sequenze di *A. thaliana* (numeri di accessione **At1g04980** e **At2g32920**) appartenenti alla classe strutturale 2 e al gruppo filogenetico V vengono attribuite da Lemaire e Miginiac-Maslow (2004) assieme a CrPDI4, al gruppo delle P5, caratterizzato dall'organizzazione strutturale: ***a°-a-b-c***. Le sequenze omologhe nei mammiferi (hP5 e CaBP1 di ratto; Füllekrug et al., 1994; Ferrari e Söling, 1999) sono entrambe proteine residenti nel RE. Queste proteine del gruppo P5 di *A. thaliana* e di *C. reinhardtii* conservano tutte il segnale C-terminale di ritenzione nel RE e sono probabilmente residenti nel RE come i loro omologhi in mammiferi.

Come già visto, alcune proteine inizialmente definite P5 sono state attribuite ad un nuovo gruppo chiamato PDI-D (Ferrari e Söling, 1999). La **PDI-D** di erba medica (Shorrosh e Dixon, 1992) e la **Dd-PDI** di *D. discoideum* (Monnat et al., 1997) hanno entrambe una struttura multidominio di tipo: ***a°-a-D,*** conservata anche nel gene di *A. thaliana* con numero di accessione **At2g47470** (Lemaire e Miginiac-Maslow, 2004), appartenente alla classe

strutturale 2 gruppo filogenetico IV. Nell'alga *C. reinhardtii* è presente una PDI di tipo D, la **CrPDI2**, la cui organizzazione ***a*-*D*** è più corta di circa 100 aa di quella di *A. thaliana*, grandezza che corrisponde a quella di un dominio di tipo ***a*** (Lemaire e Miginiac-Maslow, 2004). La proteina di *A. thaliana* ha un tetrapeptide C-terminale (VASS) che non funziona come segnale di ritenzione nel RE. Anche nelle PDI-Dα di erba medica e di *D. discoideum* non c'è un segnale C-terminale di ritenzione per il RE, nonostante sia noto che queste proteine sono presenti nel RE (Shorrosh e Dixon, 1992; Monnat et al., 1997). Recentemente è stato dimostrato che la regione C-terminale del dominio ***D*** è responsabile per la localizzazione nel RE della PDI-D in *D. discoideum* (Monnat et al., 2000), ciò rende plausibile che anche le proteine di *A. thaliana* e di *C. reinhardtii*, in quanto dotate di un dominio ***D***, siano presenti nel RE (Lemair e Miginiac-Maslow, 2004). Queste proteine sono redox attive, in quanto presentano almeno un dominio di tipo ***a*** attivo, e vengono chiamate PDI-Dα in contrapposizione alle proteine PDI-Dβ inattive, alle quali appartengono ERp28 e ERp29 nei mammiferi. Per le proteine PDI-Dβ non sono stati trovati omologhi né in *A. thaliana* né in *C. reinhardtii,* probabilmente per la difficoltà di individuare i domini di tipo ***b***.

Alle proteine con un solo dominio tioredossinico identificate in *A. thaliana* da Houston et al. (2005) e da Lemaire e Miginiac-Maslow (2004) appartiene quella codificata dal gene con numero di accessione **At1g07960** (classe strutturale 5, gruppo filogenetico VI) e dal suo omologo CrPDI3 in *C. reinhardtii*, attribuite da Lemaire e Miginiac-Maslow ad un nuovo gruppo di PDI definito PDI-A, caratterizzato dalla presenza del solo dominio catalitico attivo di tipo ***a***. Un'altra proteina appartenente alla classe strutturale 5, codificata dal gene con numero di accessione **At1g35620** (gruppo filogenetico VII) ha un'organizzazione dei domini di tipo: ***a*-*b*-*b'***. Per questo gene, di cui apparentemente non c'è un omologo in *C. reinhardtii,* analisi nelle banche dati EST hanno mostrato la presenza di omologhi in altre piante superiori. I geni con numeri di accessione **At1g52260** e **At3g16110 (**classe strutturale 1, gruppo filogenetico III) e **At3g20560** e **At4g27080** (classe strutturale 5, gruppo filogenetico VIII) non erano stati identificati da Lemaire e Miginiac-Maslow.

Non per tutti i geni PDI-like identificati nei mammiferi, in particolare nell'uomo, sono stati trovati omologhi nelle piante. Per alcuni, ad esempio *ERp28*, *ERp27* ed *ERp29*, questo potrebbe dipendere solo dalla difficoltà di identificare i domini tioredossinici inattivi di tipo ***b***. Per altri come *Erdj5*, che codifica per una proteina ibrida tra DnaJ/PDI, solo recentemente caratterizzata, ci sono già evidenze, non pubblicate, che ci sia un omologo in *A. thaliana* ed in *C. reinhardtii* (Lemaire e Miginiac-Maslow , 2004). Solo ulteriori ricerche potranno definire quali delle proteine PDI-like presenti nei mammiferi siano presenti nel regno vegetale.

Proteine come PDILT ed ERp44, ad esempio, potrebbero essere sfuggite a quest'analisi, perchè non presentano un sito attivo classico contenente due cisteine attive. Per membri come ERp46 e PDIr, invece, è più difficile fare ipotesi, e solo una ricerca accurata ed esaustiva e una migliore conoscenza dei membri già identificati permetterà di definire questa grande famiglia genica nelle sue differenze e nelle sue similitudini.

1.7. Il ruolo della PDI nelle piante

Nonostante le recenti analisi sulla complessità e diversità della famiglia genica delle PDI nelle piante ci sono ancora numerosi quesiti irrisolti riguardanti la localizzazione cellulare e la funzione fisiologica delle singole proteine. La maggior parte degli studi è stata rivolta alla caratterizzazione molecolare, all'analisi dell'espressione e alla localizzazione intracellulare della PDI classica di alcuni cereali (Chen e Hayes, 1994; Shimoni et al., 1995b; Li e Larkins, 1996). Da questi studi è emerso che la PDI probabilmente ha un ruolo importante nel folding delle proteine secretorie delle piante, in particolare nella formazione dei corpi proteici nell'endosperma.

L'analisi dell'espressione dei geni codificanti per alcune proteine appartenenti alla famiglia delle PDI nelle piante ha indicato che, sebbene i relativi trascritti siano presenti costitutivamente in tutti i tessuti ed organi analizzati, talvolta il loro livello di espressione può variare sensibilmente. Gli mRNA relativi alla PDI classica, isolati in mais, orzo, frumento e ricino, sono espressi preferenzialmente nei semi in formazione, rispetto agli altri tessuti analizzati (radici, foglie, cotiledoni, embrioni, semenzali, ecc.) (Li e Larkins, 1996; Chen e Hayes, 1994; Shimoni et al, 1995b, Ciaffi et al., 2001; Coughlan et al., 1996). In erba medica, il livello di trascritti della PDI classica (alfPDI-1) risulta nettamente più elevato nei fiori maturi che nelle foglie e nelle radici (Shorrosh e Dixon, 1991), mentre la quantità degli mRNA rilevati per la PDI appartenente alla gruppo delle PDI-D (alfPDI-2) era approssimativamente la stessa, indipendentemente dal tipo di tessuto e dallo stadio di sviluppo della pianta (Shorrosh e Dixon, 1992). Il livello di trascrizione del gene codificante per la PDIL1 isolato in carota risulta particolarmente elevato negli embrioni somatici, nei fiori immaturi e nelle foglie giovani, rispetto a radici e foglie di piante adulte (Xu et al., 2002). La PDIL1 di carota è caratterizzata dalla presenza, nella regione N-terminale, di un dominio di circa 70 aa che contiene più del 40% di residui acidi, oltre alla presenza dei due siti catalitici, di un potenziale segnale peptidico e del segnale KDEL, importante per la ritenzione nel RE.

La PDIL1 di carota ha un'organizzazione strutturale simile a quella della Rb60 e a quella codificata da due geni di *A. thaliana* (**At5g60640** e **At3g54960**) appartenenti al gruppo filogenetico II e classe strutturale 1.

Studi condotti sull'espressione e localizzazione della PDI classica in mais (Li e Larkins, 1996) e soprattutto in frumento, come si vedrà in dettaglio nei paragrafi successivi, suggeriscono che, come nei vertebrati, questo enzima possa giocare un ruolo importante nel folding delle proteine di secrezione delle piante ed in particolare nella formazione dei corpi proteici nell'endosperma dei semi, intervenendo nella formazione dei legami disolfuro e nell'assemblaggio delle proteine di riserva.

Lo studio di alcuni mutanti di mais e riso, che producono semi i cui corpi proteici nell'endosperma risultano alterati, ha confermato il ruolo della PDI classica nella deposizione delle proteine di riserva nei cereali. È stato osservato, infatti, che nell'endosperma del mutante di mais *floury-2,* che esprime una α-zeina non correttamente processata, l'espressione della PDI aumenta in maniera significativa. Tale incremento di espressione della PDI a livello proteico e di mRNA è correlato positivamente con il dosaggio di alleli *fl2*. Oltre ad un aumento della PDI, che viene principalmente accumulata nelle cisterne del RE, c'è un incremento dei livelli di BiP. L'aumento della Bip avviene principalmente all'interno dei corpi proteici del RE (Li et Larkins; 1996). L'induzione della PDI nei mutanti *fl2* potrebbe riflettere un suo ruolo come chaperone molecolare ed indicare che la PDI funzioni in concerto con la la BiP a differenti stadi del processing della zeina e nell'assemblaggio dei corpi proteici (Li et Larkins; 1996).

Due ulteriori mutanti di mais, *mucronate* (*mc*) e *defective endosperm B30* (*de*-B30*), esibiscono un aumento endosperma specifico della PDI in seguito a cambiamenti strutturali nelle proteine di riserva (Wrobel, 1996; Kim et al., 2004).

Le principali proteine di riserva del riso sono le gluteline, proteine omologhe alla globulina 11S di soia e pisello, e le prolammine (Zhao et al., 1983; Takaiwa et al., 1987; Shotwell e Larkins, 1989). Nonostante entrambi i tipi di proteine di riserva vengano sintetizzati sulla membrana del RE e traslocati all'interno del lume del RE, una volta maturi vanno a formare corpi proteici (PB) distinti (Tanaka et al., 1980; Krishnan et al., 1986; Yamagata e Tanaka, 1986).

Le prolammine vengono assemblate all'interno del RE e accumulate in forma di granuli di inclusione intra-cisterne. I corpi proteici delle prolammine vengono chiamati PB-I e hanno forma sferica, con diametro di circa 1-2 μm. Le gluteline invece si accumulano sotto forma di cristalloidi nei corpi proteici di tipo II (PB-II) di origine vacuolare, la cui forma è

irregolare e il cui diametro è di circa 3-4 μm. Una proteina in particolare, la α-globulina, che è codificata da un solo gene (Shorrosh et al., 1992) e da sola costituisce fino al 12% di tutte le proteine del seme (Krishnan e White, 1995), risulta sequestrata nella matrice che circonda i cristalloidi nei corpi proteici di tipo II (Bechtel e Juliano, 1980; Krishnan et al., 1992). Il precursore delle gluteline di 57 kD viene sintetizzato sulla membrana del RE e quindi trasportato via Golgi al vacuolo di riserva, dove viene processato in subunità acide e basiche (Yamagata et al., 1982).

Nei mutanti naturali di riso *esp2* (*endosperm storage protein2*), in cui è stata verificata la mancanza di espressione della PDI, sia a livello mRNA che di proteina, sono presenti e risultano normali i corpi proteici di tipo II, ma mancano quelli di tipo I (Takemoto et al., 2002). Questi mutanti sono invece dotati di numerosi corpi proteici più piccoli (0,5 μm), che hanno origine dal RE (Takemoto et al., 2002) e al cui interno si accumulano sia le prolammine ricche in Cys che il propeptide di 57 kD delle gluteline, probabilmente legati da ponti disolfuro intermolecolari, interazione che impedisce al propeptide di raggiungere il complesso del Golgi e quindi il vacuolo. La PDI risulta quindi essenziale per la normale segregazione delle prolammine e del propeptide delle gluteline nel RE (Takemoto et al., 2002), probabilmente assistendo la formazione dei ponti disolfuro intramolecolari ed intermolecolari di queste proteine.

Molte proteine di riserva, tra cui alcune prolammine, le 2S albumine e la α-globulina, contengono sequenze conservate in tre distinte regioni chiamate A, B e C, i cui motivi consenso sono, rispettivamente, Lxx**C**, **CC**xQL e Pxx**C**. La struttura tridimensionale delle 2S albumine in girasole (Pandya et al., 2000; Pantoja-Uceda et al., 2004) ed in ricino (Pantoja-Uceda et al., 2003) consiste in 4 eliche, H1, H2, H3 e H4, è stabilizzata da legami disolfuro intramolecolari ed è conservata anche nella α-globulina. Nella α-globulina monomerica i quattro residui cisteinici formano ponti disolfuro intramolecolari, in particolare la cisteina del motivo A (C48) si lega con la prima cisteina del motivo B (C78), mentre la seconda cisteina del motivo B (C79) si lega con la cisteina del motivo C (C171). Recentemente Kawagoe et al. (2005) hanno dimostrato, usando la proteina GFP (green fluorescent protein) fusa a differenti segmenti della α-globulina monomerica, che proprio i ponti disolfuro intramolecolari sono necessari per il suo trasporto nella matrice dei PB-II. La proteina di fusione GFP-ABC viene normalmente accumulata nei PB-II, ma una mutazione della C79 del motivo conservato B determina l'accumulo nei PB-I di questa stessa proteina di fusione GFP-ABC(C79F). Inoltre la proteina di fusione GFP-AB viene accumulata nei corpi PB-I, mentre la GFP-AB(C79S) risulta accumularsi nei corpi proteici PB-II. Apparentemente la presenza di residui cisteinici

disponibili, la C79 nel motivo conservato CCxxQL nella proteina di fusione GFP-AB e la C171 nel motivo conservato PxxC nella proteina di fusione GFP-ABC(C79F), è responsabile dell'accumulo della proteina corrispondente nei corpi proteici di tipo I. Nella proteina di fusione GFP-ABC sia il legame tra la C43 e la C78 che quello tra C79 e C171 sono intramolecolari, mentre si è ipotizzato che nella proteina di fusione GFP-AB la C79 formi un legame intermolecolare con la Cys del motivo conservato PxxC delle prolammine. È stato dimostrato che questo residuo cisteinico (PxxC) nelle prolammine partecipa alla polimerizzazione (Kawagoe et al., 2005). Il fatto che la cisteina C79, in mancanza della C171 del motivo C nella proteina GFP-AB, formi un legame disolfuro intermolecolare con il motivo conservato C delle prolammine spiegherebbe perché la proteina rimanga bloccata nel RE, invece di essere trasportata attraverso il Golgi nel vacuolo di riserva. Probabilmente anche la C171 può interagire con il motivo CCxxQL delle prolammine. Comunque l'intera regione conservata CCxxQL è necessaria per l'efficiente accumulo nei corpi PB-I della proteina di fusione GFP-AB. Mutando infatti la Leu-82, GFP-AB(L82F) si riduce l'accumulo della proteina di fusione nei PB-I e la mutazione della Cys-78, GFP-AB(C78S), porta al suo accumulo nel lume del RE.

Questi tre motivi conservati sono presenti non solo nelle proteine monomeriche 2S albumine e α-globuline, ma anche nelle prolammine polimeriche. Kreis et al. (1985) hanno discusso l'evoluzione delle prolammine polimeriche e delle albumine monomeriche 2S da un gene ancestrale comune. Si ipotizza che i legami disolfuro intermolecolari delle prolammine si siano evoluti dai legami intramolecolari indispensabili nelle albumine monomeriche 2S. Una mutazione ipotetica nell'"hypervariable loop", tra le regioni B e C delle 2S albumine (Pantoja-Uceda et al., 2004), sarebbe responsabile dell'alterata struttura del loop che ha dato origine ad una proteina polimerica da cui si sarebbero evolute le prolammine.

Nella regione A della α-globulina è inoltre presente un segnale VTS (Vacuolar Targeting Signal) responsabile della localizzazione nel vacuolo di riserva della α-globulina correttamente ripiegata, ma non sufficientemente forte da indirizzare l'accumulo di proteine covalentemente legate a prolammine.

1.8. La PDI in frumento

Inizialmente la presenza della PDI nell'endosperma di frumento è stata rilevata da Roden e collaboratori (1982), i quali hanno dimostrato che l'attività della PDI è associata a

frazioni del RE isolate attraverso ultracentrifugazione di omogenati di endosperma in via di sviluppo. La PDI risulta associata anche con frazioni microsomiche derivanti da embrione e strati aleuronici di cariossidi di frumento in germinazione (Livesley et al., 1992). La PDI è stata purificata da cariossidi in via di sviluppo di frumento (*T. aestivum*) ed è risultata una glicoproteina di 60 kDa (Shimoni et al., 1995b). Analizzando i livelli della PDI mediante western blotting durante la granigione è risultato che, come la BIP, risulta particolarmente abbondante nell'endosperma e nelle radici rispetto a foglie ed embrioni (Shimoni et al., 1995). Analisi della localizzazione subcellulare della PDI e fotografie con la tecnica immunogold di microscopia elettronica hanno dimostrato che la PDI non solo è presente nel lume del RE, ma è anche co-localizzata con le proteine di riserva nei corpi proteici. La presenza della PDI nel RE dell'endosperma di frumento non prova necessariamente che essa sia coinvolta nel folding delle proteine di riserva anche *in vivo*, benché Bulleid e Freedman (1988a) abbiano osservato che la PDI (PDI bovina purificata), almeno *in vitro*, possa catalizzare la formazione di ponti disolfuro intramolecolari in una γ-gliadina, precedentemente sintetizzata da cDNA e traslocata in vescicole microsomiche di pancreas di cane private delle proteine solubili intraluminali. Infatti, prima dell'aggiunta della PDI bovina ai microsomi privati di proteine intraluminali, nella γ-gliadina si aveva una bassa formazione di legami disolfuro, mentre in seguito è stato registrato un aumento significativo nella formazione dei legami disolfuro intramolecolari.

Studi simili sono stati condotti anche per le subunità di glutenine HMW 1Dy10 e 1Dy12 e per una glutenina LMW (Bulleid e Freedman, 1988b; 1992), dimostrando l'intervento della PDI nella formazione dei legami disolfuro intramolecolari, ma non nella formazione di oligomeri stabilizzati da ponti disolfuro intermolecolari. La formazione di ponti disolfuro in una proteina veniva visualizzata mediante separazione elettroforetica, basata sul fatto che proteine contenenti legami disolfuro intramolecolari avevano una mobilità maggiore rispetto a proteine completamente ridotte.

Un supporto indiretto per il ruolo della PDI nella maturazione delle gliadine e nel loro assemblaggio per formare i corpi proteici è stato proposto da Shimoni e Galili (1996). Studi condotti mediante pulse chase *in vivo,* ad uno stadio precoce di sviluppo delle cariossidi, mostrano che le gliadine vengono dapprima sintetizzate come monomeri e che la loro deposizione nei corpi proteici avviene tramite un processo di assemblaggio ordinato. La somministrazione *in vivo* di un agente riducente come il DTT (ditiotreitolo) risulta in una aggregazione prematura di gliadine nell'endosperma. Inoltre *in vitro* il trattamento riducente con DTT di monomeri di gliadine ossidati porta all'aggregazione di complessi difettivi. Questo indica una partecipazione *in vivo* delle proteine residenti nel reticolo endoplasmatico,

che regolano in qualche modo la formazione e l'isomerizzazione dei legami disolfuro intramolecolari.

La localizzazione cromosomica in frumento dei geni codificanti per la PDI appartenente al gruppo filogenetico I e classe strutturale 1 è stata eseguita mediante la tecnica del Southern blotting usando una sonda ottenuta mediante amplificazione PCR con primer disegnati sulla base della sequenza cDNA precedentemente pubblicata da Shimoni e collaboratori (1995a). La sonda è stata ibridata con DNA estratto dalla cv di frumento tenero Chinese Spring e dalle sue linee aneuploidi e digerito con diversi enzimi di restrizione. Sono stati così individuati 4 frammenti di diversa dimensione localizzati sui bracci 4AL, 4BS, 4DS e 1BS delle linee aneuploidi ditelosomiche di CS (Ciaffi et al., 1999). Recenti ricerche hanno evidenziato la possibilità che il gene della PDI localizzato sul cromosoma 1BS possa corrispondere ad uno pseudogene privo della regione codificante in 3' (Johnson e Bhave, 2004). La particolarità che il gene della PDI sia localizzato sul braccio lungo e non su quello corto del cromosoma 4A può essere spiegata dalla presenza di una inversione pericentrica in questo cromosoma (Devos e al., 1995).

È stata quindi studiata la variabilità delle sequenze codificanti per la PDI in specie coltivate di frumento esaploidi e tetraploidi mediante analisi RFLP (Restriction Fragment Length Polymorphism) utilizzando DNA digeriti con *EcoRI*. Da queste analisi è risultato che i frammenti di restrizione localizzati sui cromosomi del gruppo di omeologia 4 sono altamente conservati sia nei frumenti esaploidi che tetraploidi, questi ultimi ovviamente non presentano il frammento di restrizione localizzato sul cromosoma 4D, e che solo il frammento localizzato sul cromosoma 1B mostrava un certa variabilità. In particolare sono stati individuati tre differenti profili. Alcuni genotipi presentavano un frammento con la medesima mobilità elettroforetica del frammento 1B osservato nella cv CS, altri un frammento con una maggiore mobilità elettroforetica, mentre in altri il frammento corrispondente al cromosoma 1B era assente (Ciaffi et al., 1999). Una caratteristica distintiva dei frumenti tetraploidi rispetto a quelli esaploidi sembra la netta prevalenza nei primi (>70% dei genotipi analizzati) del profilo caratterizzato dal frammento 1B con maggiore mobilità elettroforetica, mentre nei frumenti esaploidi analizzati i tre profili individuati erano equamente distribuiti (Ciaffi et al., 1999). Le analisi RFLP sono state estese a 23 specie selvatiche diploidi appartenenti ai generi *Triticum* e *Aegilops* (Ciaffi et al., 2000). Queste analisi hanno indicato che i frammenti di restrizione della PDI sono altamente conservati entro ciascuna specie e hanno altresì confermato che la PDI nelle piante è codificata da una o poche copie per genoma, rispettivamente in specie diploidi e poliploidi. Delle 15 accessioni di *T. urartu* (AA) analizzate, 14 hanno evidenziato

un frammento di restrizione con la stessa mobilità di quello localizzato sul cromosoma 4A di CS e presente sia in *T. dicoccoides* (AABB) che nei frumenti poliploidi coltivati (AABB ed AABBDD). Una sola accessione di *T. urartu* di origine siriana ha mostrato un frammento di mobilità inferiore, che è risultato conservato in due accessioni di *T. dicoccoides* provenienti da Israele. Il frammento localizzato sul braccio cromosomico 4BS è risultato conservato in tutte le 15 accessioni di *T. dicoccoides* analizzate ed è risultato identico a quello presente in CS, inoltre tre delle accessioni presentavano anche il frammento 1BS, in particolare due presentavano quello di maggiori dimensioni ed uno quello di minore dimensione. *T. boeoticum* (A^mA^m) e *T. monococcum* (A^mA^m) hanno mostrato invece un identico frammento di restrizione, la cui mobilità è nettamente maggiore di quello presente nel genoma A dei frumenti coltivati. Questi risultati rafforzano l'ipotesi che *T. urartu* sia la specie donatrice del genoma A di *T. dicoccoides* e dei frumenti poliploidi coltivati e che *T. monococcum* sia la forma domesticata di *T. boeoticum*. Infine tutte le accessioni di *Ae. squarrosa* analizzate hanno evidenziato un unico frammento identico a quello presente nel genoma D dei frumenti poliploidi coltivati (Ciaffi et al., 2000). Le cinque specie di *Aegilops* appartenenti alla sezione Sitopsis, ad eccezione di *A. searsii* che possiede un singolo frammento di restrizione che ibrida con la sonda per la PDI, hanno mostrato invece un profilo di ibridazione complesso, con un elevato livello di variabilità intra-specifica soprattutto per l'unica specie allogama *Ae. speltoides*. Nessuna delle cinque specie ha mostrato un profilo di ibridazione riconducibile a quello del genoma B dei frumenti coltivati e di *T. dicoccoides.* Non è stato quindi possibile trarre nessuna conclusione sull'origine del genoma B.

Il profilo di ibridazione di *T. zhukovsky* (AAA^mA^mGG) era identico a quello di *T. timopheevi* (AAGG), eccetto per la presenza di un frammento di ibridazione addizionale con la medesima mobilità di quello di *T. boeoticum* e *T. monococcum*. Queste analisi confermano che uno dei genomi A di *T. zhukovsky*, una specie isolata da una coltura mista di *T. monococcum* e *T. timopheevi* e considerata un anfiploide delle due specie (Upadahya e Swaminthan, 1963), potrebbe derivare da *T. monococcum* (A^mA^m).

Le sequenze nucleotidiche dei tre geni localizzati sui genomi A, B e D (GPDI-4A, GPDI-4B e GPDI-4D) sono lunghe, rispettivamente, 3561, 3527 e 3466 pb (paia di basi). Il confronto e l'allineamento delle sequenze genomiche con le rispettive sequenze cDNA (CPDI-4A, CPDI-4B e CPDI-4D) hanno dimostrato che i tre geni possiedono una struttura conservata, costituita da 10 esoni (Ciaffi et al., 2006; Fig. 1.8).

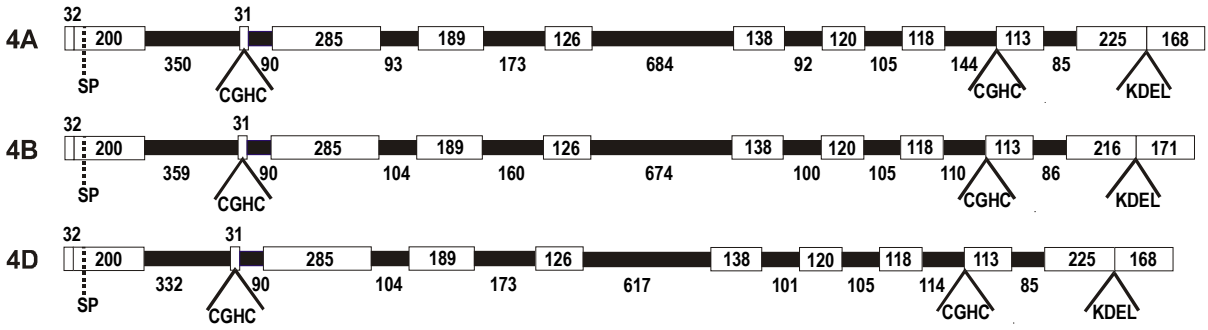

Fig. 1.8 - Struttura esoni/introni dei tre geni della WHEPDI1 localizzati sul gruppo di omeologia 4. In bianco sono indicati gli esoni ed in nero gli introni, i numeri rappresentano le dimensioni in paia di basi (pb) dei relativi elementi. Sono indicati il potenziale peptide segnale (SP), i due siti attivi tioredossina simili (CGHC) e la sequenza C-terminale KDEL di ritenzione nel reticolo endoplasmatico (da Ciaffi et al., 2006).

Gli introni hanno mostrato dimensioni molto variabili, con il più piccolo di 85 PB (introne 9 nel gene 4A e 4D) e 86 pb (gene 4B) ed il più grande, situato in posizione centrale, di 684, 674 e 617 pb (introne 5 nei geni 4A, 4B e 4D). Anche il gene codificante per la PDI in frumento duro, isolato dalla varietà Langdon, è composto di 10 esoni e la sua dimensione è di circa 3.5 Kb (Ciaffi et al., 2001).

Il confronto delle tre sequenze genomiche ha evidenziato, inoltre, una notevole conservazione nell'organizzazione introni/esoni dei tre geni omeologhi isolati in frumento tenero. Il primo esone in 5' è costituito da una sequenza non tradotta (5'UTR) di 32 nucleotidi e da 200 pb della regione codificante, che inizia con una sequenza nucleotidica di 75 pb codificante per un potenziale segnale peptidico di 25 aminoacidi. I codoni dei due siti catalitici, caratterizzati dalla sequenza aminoacidica –CGHC-, sono localizzati, in tutti i tre geni, sul secondo esone, a partire dal primo nucleotide, e sul nono esone, a partire dal 26° nucleotide. Le triplette di un potenziale sito di N-glicosilazione (NFS) si trovano nel sesto esone a partire dal 15° nucleotide. Il decimo esone contiene una regione non tradotta (3'UTR) lunga 171 pb in GPDI-4B e 168 pb in GPDI-4A e GPDI-4D ed una regione tradotta lunga 216 pb in GPDI-4B e 225 pb in GPDI-4A e GPDI-4D. Nella regione codificante del 10° esone in GPDI-4B è presente una delezione di tre triplette codificanti per alanina. I tre geni contengono, inoltre, all'estremità 3' tradotta, una sequenza consenso codificante per il segnale di ritenzione nel RE, il tetrapeptide KDEL (Ciaffi et al., 2006).

Gli Open Reading Frames (ORFs) dei geni PDI localizzati sui cromosomi 4A e 4D consistono di 1545 pb, corrispondenti ad una proteina di 515 aminoacidi con un peso molecolare stimato di 56,6 kDa, mentre l'ORF del gene localizzato sul cromosoma 4B ha una lunghezza di 1536 pb, corrispondente ad una proteina di 512 aminoacidi con un peso molecolare stimato di 56,3 kDa. Le tre sequenze proteiche dedotte sono ricche, in residui acidi e hanno un punto isoelettrico pI di 4,7. Le tre sequenze nucleotidiche dei tre ORF hanno

mostrato un identità compresa tra 96% e 97.5%. Solo 14 delle 66 sostituzioni nucleotidiche nella regione codificante determinano modificazioni nella sequenza aminoacidica, solo sei di esse potrebbero modificare le proprietà chimico-fisiche delle proteina matura (Ciaffi et al., 2006).

Le sequenze genomiche GPDI-4B e GPDI-4D hanno mostrato tra loro un'identità maggiore, del 94,5%, rispetto a GPDI-4A che ha un'identità del 92% con entrambe le sequenze genomiche localizzate sui cromosomi 4B e 4D. Il confronto delle sequenze nucleotidiche dei tre geni ha indicato, come atteso, un minor livello di identità tra gli introni, compresa tra 89-91%, che tra gli esoni, compresa tra 94-100%. Inoltre, mentre le differenze riscontrate tra gli esoni, ad eccezione della delezione di nove nucleotidi nella regione codificante del 10° esone, consistono di semplici sostituzioni nucleotidiche, quelle tra gli introni sono dovute, oltre alle singole sostituzioni, anche ad inserzioni e/o delezioni, come evidenziato dalla differente dimensione di sette dei nove introni dei tre geni clonati.

La sequenza genomica del gene della PDI localizzato sul cromosoma 4D di CS ha mostrato un'identità del 99,7% con quella di *Ae. tauschii* (Johnson e Bhave, 2004). Le differenze più rilevanti sono una delezione di 34 pb alla fine del primo introne e una delezione di due pb nell'ottavo introne di GPDI-4D. Le sequenze del gene per la PDI del genoma A di specie esaploidi e tetraploidi sono altamente conservate e mostrano un'identità del 99,6% (Ciaffi et al., 2001). È inoltre interessante notare che la struttura introni/esoni, nonché la dimensione degli esoni della PDI, sono quasi perfettamente conservate tra i geni di frumento e quelli di *Arabidopsis thaliana* e riso (Ciaffi et al., 2006). Ciò è dovuto verosimilmente alla relazione tra organizzazione in domini della proteina codificata e sequenza genomica dei geni corrispondenti.

Il confronto e l'analisi delle sequenze aminoacidiche dedotte hanno evidenziato che, nonostante tra i geni della PDI di frumento e dell'uomo sia stata rilevata un'omologia ridotta con 31% di identità, le loro strutture modulari, in termini di numero, dimensione, posizione e disposizione dei motivi strutturali dei singoli domini costituenti, sono risultate molto simili (Ciaffi et al., 2005).

Omologie di sequenze, sia interne che con quella della PDI umana, indicano che la proteina di frumento è composta da 4 regioni principali (Fig. 1.9), che corrispondono quasi esattamente ai domini ***a, b***, ***b'*** e ***a'*** della PDI umana, i cui limiti, come già detto, sono stati identificati sperimentalmente. Le regioni ***a*** e ***a'***, che contengono i due siti catalitici, hanno evidenziato una notevole omologia tra loro (43% di identità), con i domini ***a*** e ***a'*** della PDI umana e con la tioredossina. Non è stata riscontrata, invece, nessuna omologia tra i due

ipotetici domini ***b*** e ***b'*** di frumento, né tra questi ed altri domini tioredossina simili. Nonostante la divergenza riscontrata a livello di sequenza aminoacidica, tutti i quattro potenziali domini della PDI di frumento, inclusi ***b*** e ***b'***, presentano una struttura secondaria molto simile tra loro e a quella della tioredossina (Ciaffi et al., 2005). Nonostante siano necessari ulteriori studi per identificare sperimentalmente i limiti dei domini, la struttura multidominio, proposta per la PDI di frumento solo sulla base della sequenza aminoacidica dedotta, suggerisce che nelle piante, come nei vertebrati (Doolittle, 1995), i quattro domini della PDI potrebbero derivare da una duplicazione parziale o dallo slittamento (slippage) della replicazione di un comune gene ancestrale della tioredossina, seguiti da modificazione della sequenza. Probabilmente questi domini si sono formati prima della diversificazione della maggior parte delle specie eucariotiche in quanto sequenze omologhe della PDI sono presenti in organismi eucariotici tassonomicamente molto lontani quali funghi, insetti, mammiferi e piante (Freedman et al., 1994; Sahrawy et al, 1996; Ferrari e Söling, 1999).

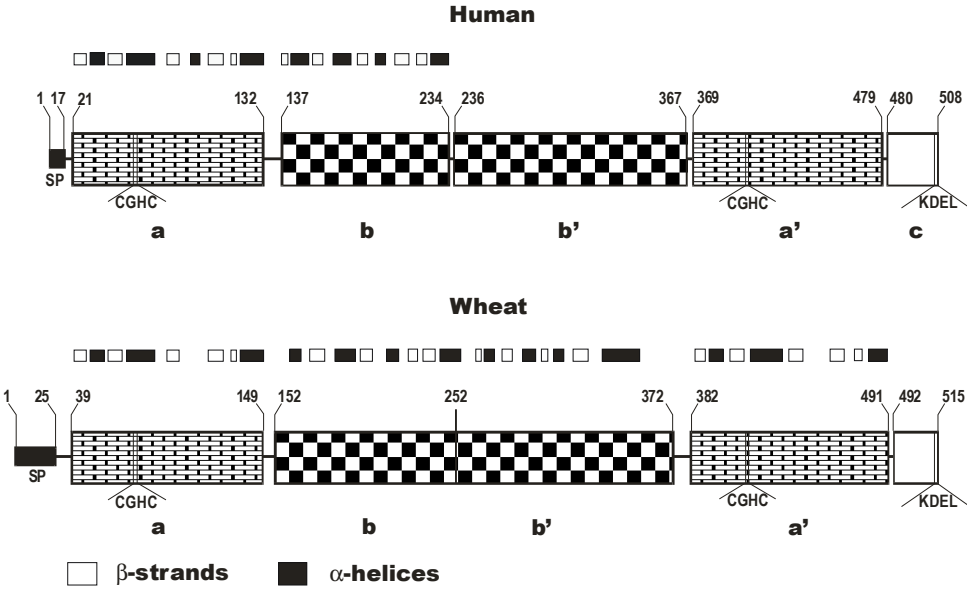

Fig. 1.9 - Confronto delle strutture proteiche modulari della PDI umana e di frumento. La prima è stata determinata sperimentalmente mediante studi di proteolisi, analisi della struttura e funzionalità dei singoli domini, la seconda è stata determinata sulla base dell'analisi della sequenza aminoacidica della proteina e del suo confronto con quella della PDI umana. La struttura secondaria dei domini ***a*** e ***b*** nella PDI umana è stata determinata mediante NMR (risonanza magnetica), mentre quella dei potenziali domini ***a***, ***a'***, ***b*** e ***b'*** della CPDI-4A è stata predetta attraverso il procedimento di Deleage et al. (1997). Sono indicati gli elementi della struttura secondaria: in bianco le α-eliche ed in nero i foglietti β (da Ciaffi et al., 2005).

Ciaffi et al. (2006) hanno clonato le tre sequenze della regione 5' a monte del codone di inizio dei tre geni omeologhi della PDI. Queste tre sequenze, la cui lunghezza è di 1352 pb per il gene GPDI-4A, 1370 pb per il gene GPDI-4B e 1292 pb per il gene GPDI-4D, hanno mostrato un'identità di 89%, un elevato grado di conservazione nei 700 nucleotidi prossimali,

con identità superiore a 93%, e un minor livello di conservazione nella regione distale, con identità di 80%. Le differenze erano causate sia da sostituzioni nucleotidiche che da brevi inserzioni e delezioni. I tre promotori presentano un TATA-box a -79 nucleotidi dal codone di inizio della traduzione ATG e numerosi CAAT-box, che aumentano la forze e l'efficienza del promotore ma non la sua specificità. Ciaffi et al. (2006) hanno inoltre identificato, per tutte e tre le sequenze regolatrici, una serie di elementi regolatori conservati. In particolare la ricerca nel database dei promotori vegetali PlantCARE (Lescot et al., 2002; http://sphinx.rug.ac.be:8080/PlantCARE/) ha permesso di identificare numerosi motivi coinvolti nella regolazione di geni espressi specificamente nei tessuti endospermatici del seme (Guilfoyle, 1997; Albani et al., 1997; Wu et al., 1998). Gli elementi regolatori identificati, elencati nella tabella 1.4 sono: 4 elementi GCN4-like, un elemento Skn-1, un Ry-element, una motivo AACA e due prolamin box (P-box).

L'analisi dell'espressione dei tre geni della PDI, effettuata in CS mediante RT-PCR su mRNA estratti da radici, piantine, foglie, spighette e cariossidi in formazione raccolte tra 6 e 34 giorni dopo l'antesi (GDA), indica che i tre geni sono espressi costitutivamente in tutti i tessuti analizzati, con un'espressione molto intensa nelle cariossidi immature. I livelli di espressione erano simili per i tre geni. I livelli di trascrizione dei tre geni omeologhi sono molto elevati nelle prime fasi di sviluppo delle cariossidi (6-14 GDA) e decrescono in seguito nella fase medio tardiva del riempimento delle cariossidi (18-34 GDA). Il livello di espressione inferiore è stato osservato, per tutti i tre geni, nelle piantine, mentre differenze nei livelli di espressione sono state osservate in spighe, radici e foglie. La trascrizione di CPDI-4A è più elevata nelle spighe, quella di CPDI-4B nelle radici e quella di CPDI-4D nelle foglie. Non è stata però rilevata alcuna differenza nell'espressione tra i tre geni durante le diverse fasi di sviluppo delle cariossidi (Ciaffi et al., 2006).

Tab. 1.4 - Principali elementi regolatori identificati nella sequenza regolatrice dei tre geni omeologhi codificanti per la PDI classica in frumento (da Ciaffi et al., 2006). Le sequenze consenso sono indicate in grassetto. Le sostituzioni di basi sono sottolineate.

Motivo	Funzione	Cromosoma	Filamento	pb dall'ATG	Sequenza
Prolamin box	motivo comunemente presente in sequenze	4A	+	-226	**TGAAAAGT**
	regolatrici di proteine di riserva. Strettamente	4B	+	-227	CGAAAAGT
	associato al motivo GCN4 nelle prolammine.	4D	+	-226	CGAAAAGT
	Coinvolto nell'espressione endosperma	4A	-	-1117	**TGAAAAGT**
	specifica.	4B	-	-1135	**TGAAAAGT**
		4D	-	-1105	**TGAAAAGT**
AACA	motivo conservato nelle sequenze promotrici di	4A	+	-1111	CAACCAATTTCG
	geni codificanti per gluteline in riso e coinvolto	4B	+	-1129	**CAACAAACTTCG**
	nell'espressione endosperma specifica.	4D	+	-1099	CAACCATTTTCG
RY-element	motivo coinvolto nell'espressione seme	4A	+	-815	**CATGCATT**
	specifica di molte proteine nelle leguminose	4B	+	-810	**CATGCATT**
	nei cereali.	4D	+	-804	**CATGCATT**
GCN4	motivo centrale coinvolto nell'espressione	4A	+	-356	CGACTCA
	endosperma specifica.	4B	+	-358	**CGAGTCA**
		4D	+	-357	TGAGTCA
		4A	+	-492	CATGTCA
		4B	+	-493	CATGTCA
		4D	+	-491	**CGTGTCA**
		4A	-	-484	**CGTGTCA**
		4B	-	-485	**CGTGTCA**
		4D	-	-483	CATGTGA
		4A	-	-780	**CGAGCCA**
		4B	-	-775	**CGAGCCA**
		4D	-	-769	**CGAGCCA**
Skn-1	motivo coinvolto nell'espressione	4A	-	-480	ACGAC
	endosperma specifica.	4B	-	-481	ACGAC
		4D	-	-479	**ATGAC**
TATA-BOX		4A	+	-79	TATTAAA
		4B	+	-79	TATTAAA
		4D	+	-79	TATTAAA
		Cromosoma		N° filamento +	N° filamento -
CAAT-box	elemento comune nelle regioni regolatrici	4A		12	12
		4B		14	17
		4D		12	15
GC - motif	motivo comune nelle regioni regolatrici	4A		6	11
		4B		7	11
		4D		10	10

2. SCOPO DELLA TESI

La PDI è una delle proteine più abbondanti nel reticolo endoplasmatico delle cellule eucariotiche, dove interviene nella formazione ed isomerizzazione dei legami disolfuro delle proteine in esso secrete. Nei vertebrati, in particolare nell'uomo, sono state studiate in maniera dettagliata sia la struttura che le funzioni di questo enzima. La PDI umana possiede una struttura multidominio ***a-b-b'-a'-c***, con due domini attivi, ***a*** e ***a'***, e due domini inattivi, ***b*** *e* ***b'***, simili a quello della tioredossina, e un quinto dominio C-terminale, ***c***, ricco di residui acidi. In aggiunta alla sua funzione principale redox-isomerasica, la PDI interviene anche nell'adesione cellulare ed agisce come chaperone molecolare. La PDI fa parte di una famiglia multigenica, che da essa prende il nome, i cui membri possiedono un numero variabile di domini tioredossina simili attivi o inattivi. Nei vertebrati sono state identificate numerose proteine PDI-like, caratterizzate da differenze nella struttura primaria, nel profilo di espressione genica e nelle funzioni metaboliche.

Nelle piante le informazioni riguardanti la famiglia PDI sono molto meno approfondite e riguardano prevalentemente l'isolamento di sequenze cDNA, per lo più codificanti per proteine omologhe alla forma classica della PDI, da varie specie vegetali. Recentemente Houston e collaboratori (2005) hanno condotto uno studio approfondito in *Arabidopsis* per individuare e classificare tutte le proteine PDI-like presenti in questa pianta e confrontarle con le informazioni disponibili in riso e mais, individuando nel regno vegetale 8 distinti gruppi filogenetici. Cinque di questi gruppi filogenetici comprendono proteine PDI-like caratterizzate dalla presenza di due domini catalitici attivi simili a quello della tioredossina, mentre agli altri tre gruppi appartengono proteine che presentano uno solo dominio attivo. I numerosi geni codificanti per proteine PDI-like presenti nelle piante differiscono non solo nel numero e posizione dei siti attivi tioredossina simili, ma anche per la presenza/assenza del segnale KDEL di ritenzione nel reticolo endoplasmatico e per la presenza di domini particolari. Nelle piante si conosce relativamente poco sul ruolo fisiologico svolto dai singoli membri di questa numerosa famiglia genica e sulla struttura ed organizzazione delle sequenze genomiche codificanti per proteine PDI-like. Nella cultivar di frumento tenero Chinese Spring (CS) recentemente sono stati isolati e caratterizzati i tre geni omeologhi codificanti per la forma classica della PDI e le loro sequenze promotrici (Ciaffi et al., 2006). I tre geni, che possiedono una struttura esoni/introni conservata costituita da 10 esoni, sono stati localizzati sui cromosomi del gruppo di omeologia 4 di CS. L'analisi di espressione ha evidenziato che la PDI classica è espressa costitutivamente in tutti i tessuti

analizzati, ma è particolarmente abbondante nelle cariossidi in via di sviluppo, specialmente durante le prime fasi di granigione. L'identificazione nella sequenza promotrice del gene della PDI di diversi elementi regolatori coinvolti nell'espressione endosperma specifica è in accordo con l'abbondante presenza di trascritti nelle cariossidi in via di sviluppo (Ciaffi et al., 2006).

L'intento generale di questa tesi è stato di proseguire nella caratterizzazione strutturale e funzionale di questa grande famiglia genica codificante per le proteine PDI e PDI-like in frumento.

Un primo obiettivo è stato l'identificazione e la caratterizzazione di nuovi geni codificanti per proteine della famiglia PDI in frumento tenero, in particolare dei geni contenenti due domini tioredossina simili attivi, che, grazie alla loro struttura, sono tutti potenzialmente coinvolti nell'isomerizzazione e nella formazione dei legami disolfuro.

Un secondo obiettivo è stato l'ottenimento di costrutti molecolari e di linee di frumento transgeniche nelle quali un gene codificante per la PDI classica di frumento viene sovra-espresso sotto il controllo di un promotore forte endosperma specifico. Tali linee saranno utilizzate in futuro per la caratterizzazione funzionale del gene.

Terzo ed ultimo obiettivo è stato l'ottenimento di costrutti molecolari e di linee di frumento transgeniche in cui il gene reporter *UidA* (GUS) viene posto sotto il controllo di 4 diversi frammenti di delezione della regione promotrice di uno dei geni codificanti per la PDI classica di frumento. Ciò permetterà di chiarire il ruolo funzionale delle diverse regioni del promotore e dei motivi regolatori in esso contenuti nella regolazione dell'espressione del gene codificante per la PDI.

3. ARTICOLAZIONE DELLE ATTIVITA'

L'attività di ricerca svolta durante il dottorato è stata articolata per raggiungere i seguenti obiettivi:

1) Ricerca in banche dati EST per l'identificazione ed il clonaggio di sequenze cDNA codificanti per proteine PDI-like in frumento.
2) Ottenimento mediante RACE di cloni cDNA relativi all'intera sequenza trascritta (full-length) di geni codificanti per proteine PDI-like.
3) Analisi delle relazioni filogenetiche tra sequenze della famiglia genica PDI di frumento e di altre specie.
4) Determinazione della sequenza aminoacidica dedotta delle proteine PDI-like codificate dalle sequenze cDNA clonate.
5) Determinazione delle sequenze genomiche relative ai cloni cDNA dei geni *TaPDIL**2**-1, TaPDIL**4**-1 e TaPDIL**5**-1* e della loro organizzazione strutturale.
6) Localizzazione cromosomica, mediante Southern blotting, delle sequenze genomiche codificanti per *TaPDIL**4**-1 e TaPDIL**5**-1*.
7) Analisi di espressione, mediante tecniche di northern blotting, dei trascritti dei geni *TaPDIL**2**-1, TaPDIL**4**-1 e TaPDIL**5**-1*.
8) Ottenimento del costrutto per la sovra-espressione del gene *TaPDIL**1**-1* codificante per la PDI classica.
9) Ottenimento dei costrutti per l'analisi funzionale del promotore del gene *TaPDIL**1**-1*.
10) Co-trasformazione biolistica di embrioni immaturi di frumento tenero della cultivar MPB-Bobwhite 26. Induzione di calli embriogenetici, selezione, rigenerazione *in vitro* ed allevamento in serra delle piantine di frumento potenzialmente transgeniche per i costrutti ai punti 8 e 9.
11) Identificazione mediante analisi PCR delle piante contenenti i costrutti per la sovra-espressione del gene *TaPDIL**1**-1*.

4. MATERIALI E METODI

4.1. Materiale vegetale

Per le analisi di espressione genica sono stati utilizzati differenti tessuti della varietà di frumento tenero Chinese Spring (CS). Le piante di questa varietà sono state allevate, durante la stagione 2003-2004, in parcelle di 5 m^2 presso i campi sperimentali dell'Azienda agraria dell'Università della Tuscia. Le spighe in via di sviluppo sono state raccolte a partire dallo stadio di primo nodo visibile, quando l'infiorescenza è già formata ed ha una dimensione di circa 1-1,5 cm. I prelievi sono stati eseguiti ad intervalli di circa 5-7 giorni fino alla fase di spigatura completa. Sono state prelevate spighe a 7 differenti stadi di sviluppo. Le cariossidi in via di sviluppo sono state raccolte dal 4° al 36° giorno dopo l'antesi (GDA). In particolare le cariossidi sono state raccolte a 9 differenti stadi di sviluppo, ossia a 4, 8, 12, 16, 20, 24, 28, 32 e 36 giorni dopo l'antesi. Sono state inoltre raccolte, a spigatura completa, le foglie a bandiera.

I semi delle varietà di frumento tenero Chinese Spring (CS) sono stati messi a germinare in cella climatica al buio alla temperatura di 10°C per 7 giorni, avvolti in carta da filtro inumidita. Dopo la germinazione le piantine sono state allevate in laboratorio a temperatura ambiente e fotoperiodo naturale in vaschette, sopra un supporto di rete, contenenti 2 L di soluzione nutritiva (Hoagland e Arnon, 1950). La soluzione nelle vaschette è stata tenuta areata in continuo tramite un gorgogliatore ed è stata rinnovata ogni 3 giorni. 20 giorni dopo la germinazione le giovani piantine, composte dallo stelo e dalle primissime foglie vere, e le relative radici sono state raccolte separatamente. Tutto il materiale raccolto è stato immediatamente immerso in azoto liquido e conservato a –80°C, fino al momento dell'estrazione dell'RNA totale.

La localizzazione cromosomica è stata condotta utilizzando la cultivar Chinese Spring assieme alle sue linee aneuploidi nulli-tetrasomiche (NT) e ditelosomiche (DT), costituite da Sears (1966). Dopo la germinazione dei semi, effettuata alle medesime condizioni descritte in precedenza, questi sono stati trasferiti in vaso e le piante allevate in serra per la raccolta di giovani foglie. Dopo la raccolta, il materiale è stato immediatamente immerso in azoto liquido e conservato a –80°C in attesa dell'estrazione del DNA.

Per l'ottenimento di linee di frumento transgeniche, la linea di grano tenero MPB-Bobwhite 26 è stata allevata in serra con un livello di radiazione solare di 1000-1200 $Jm^{-2}h^{-1}$ ed una temperatura variabile tra 18 e 25°C. L'irrigazione è stata praticata ogni 2-3 giorni in condizioni di poco sole, o quotidianamente nei periodi più caldi. Sono state praticate due fertilizzazioni per ciclo, una in coincidenza della levata ed una in concomitanza dell'antesi, con 2,5 g di urea (46-00-00)/fosfato triplo (00-60-00) nel rapporto 2:1 (v/v). Le spighe sono state raccolte 12-16 giorni dopo l'antesi e le cariossidi immature sono state immediatamente estratte e sterilizzate.

Anche le giovani piante provenienti dalla coltura *in vitro* sono state allevate in serra alle medesime condizioni di luce, temperatura etc. fino alla completa maturazione delle cariossidi. Da questo materiale sono state raccolte foglie per l'estrazione del DNA genomico. Il materiale potenzialmente transgenico è stato allevato in una serre approvata dalla Biosafety Committee del CIMMYT (Centro Internazionale per il Miglioramento del Mais e del Frumento, Messico).

4.2. Preparazione degli acidi nucleici

4.2.1. Estrazione di DNA nucleare

Il DNA è stato estratto da giovani foglie della cultivar Chinese Spring e delle sue linee aneuploidi nulli-tetrasomiche e ditelosomiche con il metodo descritto da Dvorak et al. (1988), adottando alcune modificazioni. Qui di seguito sono riportate le varie fasi previste nel protocollo di estrazione utilizzato.

- Da 3 a 5 g di materiale vegetale sono stati omogeneizzati in Waring Blender insieme a 45 mL di tampone di estrazione antiossidante (H1) pH 9,5 (Saccarosio 0,5 M; Spermidina 1 mM; Spermina 4 mM; EDTA 10 mM; KCl 80 mM; Tris 10 mM; PMSF (fenil metil sulfonil fluoruro) 1 mM; 0,1% DTT).
- L'omogenato è stato filtrato mediante garza e filtro Miracloth e l'estratto successivamente centrifugato a 3000 rpm per 20 minuti a 4°C per precipitare le cellule.
- Dopo la centrifugazione il supernatante è stato eliminato ed il pellet è stato risospeso in 30 mL del tampone H1, contenente 0,5% di Triton X-100, un detergente che lisa le membrane dei cloroplasti e mitocondri, lasciando intatte quelle dei nuclei.

- Dopo un'ulteriore centrifugazione a 3000 rpm per 20 minuti a 4°C il pellet è stato risospeso in 20 mL del tampone precedente.
- Dopo avere ricentrifugato a 2800 rpm, il pellet è stato risospeso in 1 mL di tampone PDB (NaCl 0,1 M; EDTA 0,1 M; Tris 0,05 M pH8), quindi sono stati aggiunti 300 µL dello stesso tampone contenente lo 0,2% di Proteinasi K ed il campione è stato incubato a 60°C per 30 minuti. In questa fase il detergente (SDS) ha determinato la lisi della membrana nucleare e la Proteinasi K ha degradato le proteine.
- Dopo aver lasciato raffreddare le provette a temperatura ambiente, è stato aggiunto 1/2 volume di fenolo (2,5 mL) per estrarre la frazione proteica. Dopo aver agitato delicatamente, l'emulsione è stata lasciata per tutta la notte a 4°C.
- Il giorno successivo sono stati aggiunti 2 mL di tampone TE (Tris 10 mM, EDTA 1 mM) e 2,5 mL di cloroformio; dopo agitazione delicata per evitare la rottura delle molecole di DNA, il campione è stato centrifugato a 5000 rpm per 15 minuti.
- Mediante pipette Pasteur (tagliate in punta in modo da evitare rotture meccaniche del DNA) è stata asportata dal tubo da centrifuga la fase superiore, facendo attenzione a non asportare la fase inferiore (fenolo/cloroformio) e le proteine degradate presenti nell' interfaccia tra le due fasi.
- Il DNA è stato quindi precipitato, aggiungendo alla fase acquosa acetato di sodio 2 M pH 5,5 (1/10 del volume) ed un volume doppio di etanolo assoluto (100%).
- Mediante pipetta Pasteur con punta uncinata è stato prelevato il DNA precipitato ed è stato immerso 4-5 volte in etanolo al 70% per eliminare l'eccesso di sali.
- Il DNA recuperato è stato lasciato asciugare all'aria e quindi risospeso in tampone TE sterile (0,5-1 mL) e conservato a 4°C.

4.2.2. Estrazione di DNA con CTAB

Il materiale vegetale conservato a -80°C è stato liofilizzato e macinato finemente per l'estrazione del DNA. Qui di seguito sono riportate le varie fasi del protocollo di estrazione utilizzato.

- Il tampone di estrazione, composto da 100 mM di Tris a pH 7,5, 700 mM di NaCl, 50 mM di EDTA a pH 8,0, 1% CTAB (Mixed alkyltrimethyl-ammonium bromide; SIGMA M-7635) e 140 mM di 14 M BME (β-mercaptoethanol), è stato preparato fresco ed è stato riscaldato a 60-65°C poco prima di ogni estrazione.

- Ad ogni provetta contenente il materiale vegetale in polvere sono stati aggiunti 400 µL del tampone di estrazione.
- I campioni sono stati agitati quindi per 90 minuti a 60-65°C in un termostato.
- Dopo 10 minuti di raffreddamento, lasciando i campioni a temperatura ambiente, sono stati aggiunti 400 µL di cloroformio-ottanolo (24:1) e il tutto è stato agitato per inversione per circa 10 minuti.
- I campioni sono stati quindi centrifugati per 10 minuti a 3500 rpm e 300 µL di supernatante sono stati trasferiti in nuovi tubi.
- Ai campioni sono stati quindi aggiunti 15 µL di RNasi A (10 mg/mL) lasciando il tutto a 37°C per 30 minuti.
- Ad ogni tubo sono stati aggiunti 35 µL di acetato di sodio 2,5 M e 600 µL di alcol etilico assoluto per precipitare il DNA. Dopo alcune inversioni i campioni sono stati posti a -80°C per 30 minuti.
- I campioni sono stati quindi centrifugati per 30 minuti a 3500 rpm dopo di che il supernatante è stato buttato.
- Il pellet è stato lavato con etanolo al 70% per 5 minuti. L'etanolo è stato quindi eliminato invertendo i tubi su carta assorbente, avendo cura di non disturbare il pellet.
- I campioni sono stati fatti asciugare sotto cappa chimica ed il DNA è stato infine risospeso in 100 µL di acqua sterile (Sigma).

4.2.3. Estrazione di RNA totale

L'RNA totale è stato estratto da radici, piantine, foglie a bandiera, spighe e cariossidi a diversi stadi di sviluppo, seguendo il metodo descritto da Colot et al. (1989). Qui di seguito sono riportate le fasi previste dal protocollo utilizzato.

- Il materiale vegetale (0,5-1 g), dopo l'aggiunta di azoto liquido, è stato polverizzato finemente con pestello e mortaio.
- Il tessuto polverizzato è stato trasferito in un tubo da centrifuga e, dopo completa evaporazione dell'azoto, risospeso nel tampone di estrazione (150 mM di LiCl; 50 mM di Tris-HCl pH 9,0; 5 mM di EDTA; 5% di SDS), nel rapporto 5 mL/g di campione.
- Alla sospensione è stato aggiunto un uguale volume di fenolo-cloroformio pH 8,0 emulsionando il tutto mediante agitazione.

- Il materiale così preparato è stato mantenuto in ghiaccio per 5 minuti e successivamente centrifugato a 4000 rpm per 10 minuti alla temperatura di 4°C.
- Al supernatante, trasferito in un nuovo tubo da centrifuga, veniva quindi aggiunto un uguale volume della miscela fenolo-cloroformio. Dopo averlo agitato, il campione è stato centrifugato a 4000 rpm per 10 minuti alla temperatura di 4°C. Questo passaggio viene ripetuto tre volte, usando, l'ultima volta, cloroformio al posto della miscela fenolo-cloroformio.
- Alla fase acquosa, nuovamente trasferita in un tubo da centrifuga, veniva aggiunta una quantità di LiCl 8M, in modo da ottenere una concentrazione finale 2M (333 µL/mL).
- Il campione veniva quindi lasciato a –20°C per almeno due ore.
- Dopo scongelamento, il campione è stato centrifugato a 10.000 rpm per 30 minuti alla temperatura di 4°C.
- Il pellet veniva quindi risospeso in 1 mL di acqua, alla quale era stato aggiunto precedentemente DEPC (dietil pirocarbonato 1 mL/L di acqua) e, successivamente, veniva aggiunto LiCl 8M, in modo da avere una soluzione ad una concentrazione 2M (333 µL/mL).
- Il campione veniva quindi lasciato in ghiaccio per due ore.
- Dopo centrifugazione a 15.000 rpm per 30 minuti a 4°C, il pellet veniva risospeso in 400 µL di acqua DEPC ed utilizzato per la purificazione di RNA messaggero (mRNA), o conservato a –20°C dopo l'aggiunta di 2,5 volumi di etanolo assoluto e acetato di sodio (NaAc) a pH 5,2.

4.3. Identificazione delle sequenze in banche dati

Per l'identificazione di sequenze codificanti per proteine PDI-like è stata effettuata una ricerca in due differenti banche dati EST di frumento utilizzando le 11 sequenze di riso assegnate alle otto classi filogenetiche della famiglia delle PDI identificate nelle piante (Houston et al., 2005): "TIGR wheat gene index" (TaGI, versione 10, Gennaio 2005) e HarvEST Wheat (versione 1.13, Marzo 2005), disponibili rispettivamente presso i siti internet http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=wheat e http://harvest.ucr.edu.

La prima banca dati contiene 580.155 EST (Expressed Sequence Tag) di cui 502.968 assemblate in contig (TC, Tentative consensus sequence) mentre le rimanenti 77.187 sono risultate sequenze uniche (singletons). L'analisi dei differenti contig ottenuti ha evidenziato la presenza di 44.954 sequenze consenso uniche (Tentative consensus sequences). Attraverso l'opzione "Sequence homology Search" utilizzando il programma BLAST è possibile verificare l'esistenza di contig o singleton che mostrano significative omologie con sequenze isolate e precedentemente caratterizzate in altre specie vegetali.

HarvEST Triticeae è un programma che gestisce una banca dati contenente circa 200.000 sequenze EST isolate da orzo e da diverse specie appartenenti al genere *Triticum*, di cui più di 70.000 da frumento tenero (*Triticum aestivum*). Questo programma viene continuamente aggiornato in relazione all'immissione di nuove sequenze EST ed è in diretta connessione con la divisione EST delle banche dati di GeneBank, EMBL, DDBJ e NCBI.

Il programma HarvEST possiede 4 opzioni per la ricerca: 1) "Search ESTs by Expression Patterns" che, tramite la selezione della libreria cDNA, permette di isolare sequenze relative a particolari tessuti (es. endosperma, sacco embrionale, radici, ecc), o a determinati stadi di sviluppo della pianta (es. spighe in pre-antesi, in antesi o post antesi), o a particolari condizioni di stress (es. piante a diversi stadi di sviluppo sottoposte ad elevate o basse temperature, allevate in condizioni di carenza idrica o in presenza di metalli tossici, quali l'alluminio); 2) "Search by GeneBank #, EST name, or Contig #" che permette di identificare una particolare sequenza sulla base del numero di accessione presente in banca dati o del nome con cui è stato inizialmente contrassegnato il clone EST; 3) "Search ESTs by best BLAST hit keywords" che permette di identificare sequenze presenti nella banca dati sulla base della loro funzione (nel nostro caso è stata utilizzata come parola chiave PDI-like protein); 4) "Find a Triticeae EST for Your Favorite Gene" che permette di verificare la presenza di sequenze EST delle Triticeae che mostrano significative omologie con sequenze

precedentemente isolate e caratterizzate in altre specie vegetali. In particolare, in quest'ultimo caso sono state utilizzate le sequenze degli 11 geni PDI-like precedentemente isolati in riso.

Le sequenze identificate mediante una ricerca incrociata utilizzando le opzioni 3 e 4 sono state ulteriormente analizzate mediante i programmi di "Multiple Alignment" e "Sequence Assembly" del software per l'analisi di sequenza DNAMAN (Lynnon BioSoft), allo scopo di eliminare sequenze ridondanti molto simili tra loro nell'intera lunghezza od assemblare sequenze che si allineavano perfettamente solo alle estremità.

In base alle sequenze EST che sono risultate differenti sono stati disegnati primer specifici con il programma "PCR primer design" del software DNAMAN. I trascritti corrispondenti sono stati quindi isolati mediante RT-PCR, utilizzando mRNA estratto da cariossidi immature raccolte 8 giorni dopo l'antesi.

Qui di seguito verrà descritto il metodo utilizzato per la reazione di trascrittasi inversa. Per la successiva amplificazione mediante PCR dei trascritti si veda il paragrafo 4.6.. Dopo separazione elettroforetica su gel d'agarosio 1,5%, i diversi prodotti di amplificazione sono stati purificati, clonati e caratterizzati per la determinazione della sequenza nucleotidica come verrà descritto in seguito. Mediante analisi di sequenza utilizzando il software DNAMAN, i diversi cloni sono stati classificati in distinti gruppi e quindi assemblati.

4.4. Analisi RT-PCR

Dall'RNA totale, inizialmente estratto da cariossidi immature raccolte 8 giorni dopo l'antesi, è stato purificato l'mRNA (Poly (A) $^{+}$ mRNA) usando il sistema magnetico PolyATract System IV (Promega), secondo la procedura suggerita dai produttori.

5 mg di RNA totale, sospesi in 500 µL di acqua DEPC, sono stati denaturati a 65°C per 10 minuti. Al campione sono stati aggiunti 10 µL di sonda biotinilata-oligo (dT), che si è ibridata con la regione poli(A) in 3' dell'mRNA. Gli ibridi formati sono stati catturati aggiungendo al tubo streptavidina accoppiata a particelle paramagnetiche e poi lavati ad alta stringenza. Mediante un supporto magnetico, la fase solida che si è formata (MagneSphere) è stata ancorata sul fondo del tubo. L'mRNA è stato eluito dalla fase solida attraverso l'aggiunta di acqua deionizzata priva di RNasi.

La reazione di trascrittasi inversa per ottenere il cDNA è stata condotta in un volume di 20 µL, contenente 50 unità di enzima (Expand^{TM} Reverse Transcriptase, Roche), il tampone di reazione 1X, 10 mM di ditiotreitolo (DTT), 20 unità di inibitore della RNasi, una

concentrazione pari a 1 mM di ciascun dNTP, 2,5 μM di primer oligo$(dT)_{16}$ e 50 ng di mRNA.

Inizialmente l'mRNA e gli oligo dT sono stati denaturati incubando le provette da 0,5 ml a 65°C per 10 minuti. Le provette sono state quindi poste in ghiaccio per aggiungere gli altri reagenti. Le provette contenenti tutti i reagenti sono state incubate per 60 minuti a 42°C e poi per 2 minuti a 95°C, per disattivare la trascrittasi inversa, ed infine sono state trasferite in ghiaccio per 5 minuti. La reazione è stata conservata a -20°C. 1 μL della reazione è stato in seguito usato per amplificare le sequenze cDNA dei nuovi geni PDI-like identificati.

4.5. Isolamento di sequenze cDNA "Full Length" mediante RACE

Per l'isolamento dei trascritti completi relativi alle diverse sequenze consenso parziali (tentative consensus sequences) ottenute dalla ricerca in banca dati "TIGR wheat gene index", mediante RACE (Rapid Amplificacion of cDNA ends), è stato usato il GeneRacer™ Kit (Invitrogen). Il sistema permette di ottenere solamente trascritti "full length" mediante l'eliminazione selettiva di mRNA parziali dal processo di amplificazione. In una prima fase l'RNA è trattato con fosfatasi intestinale di vitello (CIP) per rimuovere il fosfato libero all'estremità 5' degli mRNA parziali o di altri acidi nucleici, eliminando la possibilità di aggiungere, alle estremità di questi acidi nucleici, una breve sequenza di nucleotidi (GeneRacer™ Oligo), che serve da stampo per l'amplificazione dell'estremità 5' della sequenza parziale cDNA. Naturalmente la reazione di defosforilazione effettuata mediante CIP non ha avuto alcun effetto su mRNA dotati del cap (7-metilguansina) all'estremità 5'. All'RNA defosforilato è stata aggiunta la pirofosfatasi acida di tabacco (TAP) per rimuovere dagli mRNA completi il cap all'estremità 5', rendendo disponibile il gruppo fosfato legato al carbonio 5' del ribosio, necessario per la reazione di ligasi con il GeneRacer™ Oligo. L'mRNA è stato quindi retrotrascritto utilizzando l'enzima Superscript™ II Reverse Transcriptase (RT) e il GeneRacer™ Oligo dT Primer, che contiene all'estremità 3' una sequenza di 36 bp che serve da stampo per l'amplificazione dell'estremità 3' non conosciuta della sequenza cDNA parziale. La reazione di trascrittasi inversa ha permesso di ottenere singoli filamenti di cDNA alle cui estremità 5' e 3' erano presenti siti di "annealing" conosciuti. Per ottenere l'estremità 5' non conosciuta della sequenza cDNA, il primo filamento cDNA è stato amplificato mediante PCR utilizzando un primer specifico (Reverse GSP), disegnato sul filamento 3'-5' della sequenza parziale, e il "GeneRacer™ 5' Primer"

(omologo al GeneRacer™ RNA Oligo). In questo caso solo gli mRNA che avevano all'estremità 5' la sequenza oligonucleotidica GeneRacer™ RNA Oligo e che erano stati completamente retrotrascritti sono stati amplificati mediante PCR. Analogamente, per ottenere l'estremità 3'sconosciuta della sequenza parziale cDNA, il primo filamento cDNA è stato amplificato utilizzando un primer specifico disegnato sul filamento 5'-3' (Forward GSP) della sequenza parziale ed il primer omologo al "GeneRacer™ Oligo dT Primer" (GeneRacer™ 3' Primer). Solo gli mRNA che possedevano una coda di poli (A) e che sono stati retrotrascritti venivano amplificati mediante PCR. Per aumentare la specificità e la sensibilità della RACE sono state condotte ulteriori amplificazioni utilizzando dei primer "nested" specifici sia per la sequenza cDNA parziale che per i siti di annealing alle estremità 5' e 3'. L'autenticità dei prodotti RACE è stata successivamente verificata mediante analisi di sequenza ed i trascritti completi dei geni PDI-like sono stati successivamente ottenuti, mediante RT-PCR su mRNA estratto da cariossidi immature, utilizzando primer specifici disegnati basandosi sulle sequenze delle regioni non tradotte alle estremità 5' e 3'. Nei prossimi paragrafi verranno descritte in maggior dettaglio le diverse fasi del sistema RACE utilizzato.

4.5.1. Defosforilazione

La reazione di defosforilazione è stata condotta in un volume di 10 µL contenente 200 ng di mRNA, 1 µL del tampone di reazione 10X (CIP Buffer), 40 unità di inibitore della RNasi (RNaseOut™) e 10 unità dell'enzima CIP ed incubata a 50°C per 1 ora. Per la precipitazione dell'RNA defosforilato sono stati aggiunti 90 µL di acqua DEPC e 100 µL di fenolo-cloroformio, agitando il tutto vigorosamente per 30 secondi mediante vortex. Quindi è stato centrifugato alla massima velocità a temperatura ambiente per 5 minuti. Dopo aver trasferito la fase acquosa (circa 100 µL) in una nuova provetta, sono stati aggiunti 2 µL di "mussel glycogen" (glicogeno di militi) (10mg/mL), 10 µL di sodio acetato 3 M a pH 5,2 e 220 µL di etanolo al 95%, agitando il tutto velocemente mediante vortex. La provetta è stata quindi lasciata in ghiaccio secco per 10 minuti e successivamente centrifugata alla massima velocità per 20 minuti a 4 °C e, dopo aver eliminato il supernatante, sono stati aggiunti 500 µL di etanolo al 70%. Infine, dopo un'ultima centrifugazione di 2 minuti a 4°C, è stato rimosso l'etanolo ed il pellet è stato asciugato all'aria e quindi risospeso in 7 µL di acqua DEPC.

4.5.2. Rimozione del cap dall'estremità 5' degli mRNA completi.

La rimozione del cap è stata effettuata mediante l'enzima pirofosfatasi acida di tabacco (TAP) in una reazione con volume finale di 10 µL. All'RNA defosforilato e risospeso in 7 µL di acqua DEPC sono stati aggiunti 1 µL del tampone di reazione 10x (TAP Buffer), 40 unità di inibitore della RNasi (RNaseOut™) e 0,5 unità dell'enzima TAP. Il campione è stato incubato a 37°C per 1 ora e successivamente l'RNA è stato precipitato come descritto nel paragrafo precedente.

4.5.3. Reazione di ligasi del GeneRacer™ Oligo agli mRNA completi privati del cap all'estremità 5'

Alle provette contenenti l'oligo RNA GeneRacer™, liofilizzato e prealiquotato (0,25 µg), sono stati aggiunti 7 µL di RNA defosforilato e trattato con l'enzima TAP per l'eliminazione del cap. Una volta risospeso l'oligo RNA, il tutto è stato incubato a 65°C per eliminare eventuali strutture secondarie dell'RNA. Dopo raffreddamento in ghiaccio per circa 2 minuti, per la reazione di ligasi, in un volume finale di 10 µL, sono stati aggiunti: il tampone di reazione 10X (Ligase Buffer), 10mM ATP, 40 unità di inibitore della RNasi (RNaseOut™) e 5 unità dell'enzima T4 RNA ligasi. Dopo una breve centrifugazione, le provette sono state incubate a 37°C per 1 ora, quindi di nuovo brevemente centrifugate e poste in ghiaccio. L'RNA è stato precipitato come già descritto nel paragrafo 4.5.1.

4.5.4. Trascrittasi inversa

Per la reazione di trascrittasi inversa è stato utilizzato l'enzima Superscript™ II RT e come primer il "GeneRacer™ Oligo dT Primer". Questo primer, lungo 54 bp, contiene una coda di poliT di 18 nucleotidi, che serve da innesco per la reazione di trascrittasi inversa, ed una sequenza all'estremità 5' di 36 bp, che contiene i siti di legame per i primer GeneRacer™ 3' e GeneRacer™ 3' Nested, che vengono utilizzati in seguito per l'amplificazione PCR. La reazione di trascrittasi inversa è stata condotta in un volume di 20 µL contenente l'RNA ligato con l'oligo RNA GeneRacer™, 1mM dATP, dTTP, dCTP e dGTP, 2,5 µM GeneRacer™ Oligo dT Primer, il tampone di reazione 5X (First Strand Buffer), 40 unità di inibitore della RNasi (RNaseOut™) e 200 unità dell'enzima Superscript™ II RT. Dopo un pre-trattamento dell'RNA a 65°C per 5 minuti, le provette contenenti tutti i reagenti sono state incubate a

42°C per 50 minuti, prima di essere poste a 70°C per 15 minuti per inattivare l'enzima. Dopo raffreddamento in ghiaccio per 2 minuti ed una breve centrifugazione sono state aggiunte 2 unità di RNasi H e la miscela di reazione è stata incubata a 37°C per 20 minuti per la degradazione dell'mRNA. La miscela di reazione è stata usata immediatamente per l'amplificazione del primo filamento cDNA o conservata a -20°C.

4.5.5. Amplificazione mediante PCR delle estremità 5' e 3' delle sequenze cDNA

Le estremità 5' e 3' delle sequenze cDNA parziali sono state ottenute amplificando il primo filamento di cDNA utilizzando due coppie di primer, ognuna costituita da una sequenza oligonucleotidica gene-specifica (Reverse o Forward GSP, rispettivamente per le estremità 5' e 3') e dal primer GeneRacer™ 5'o GeneRacer™ 3', rispettivamente omologhi alle sequenze aggiunte alle estremità 5' e 3' dei trascritti completi GeneRacer™ RNA Oligo e GeneRacer™ Oligo dT.

Nel disegnare i primer gene-specifici sono state seguite le seguenti condizioni:

- il contenuto in GC doveva essere compreso tra 50-70% in modo da ottenere una elevata temperatura di annealing (>72°C);
- la lunghezza dei primer doveva essere compresa tra 23-28 nucleotidi per aumentare la specificità di legame e quindi della reazione;
- i primer dovevano comunque avere un basso contenuto in GC all'estremità 3', per impedire che la DNA polimerasi iniziasse la sua attività partendo da siti non-specifici;
- doveva essere evitata la presenza di sequenze complementari all'interno dei GSP o tra questi e i primer dati in dotazione con siti di annealing alle estremità 5' e 3':
- i primer dovevano avere una temperatura di annealing maggiore di 72°C, per aumentare la specificità di legame, riducendo quindi eventuali amplificazioni non specifiche.

La reazione di amplificazione è stata condotta in un volume di 50 µL contenente 0,6 µM GeneRacer™ 5' o 3' Primer, 0,2 µM primer gene specifici (GSP Reverse o Forward), il tampone di reazione 10X (High Fidelity PCR Buffer; Invitrogen), una concentrazione pari a 300 µM di ciascun dNTP, 1 mM di $MgSO_4$, 0,25 unità di enzima (Taq DNA polymerase High Fidelity, Invitrogen) ed 1 µL della reazione di trascrittasi inversa come templato. Per aumentare la specificità e l'efficienza delle reazioni PCR è stato utilizzato un sistema di amplificazione definito "Touchdown PCR", che consiste nell'utilizzare, durante i primi cicli di amplificazione, temperature di annealing più elevate di quelle specifiche dei primer utilizzati e, nei cicli successivi, temperature di annealing leggermente più basse. Dopo una

iniziale denaturazione a 94°C per 2 minuti, sono seguiti una prima serie di 5 cicli di amplificazione così composti: 94°C per 30 secondi, 72°C per 2 minuto e 30 secondi; una seconda serie di altri 5 cicli così composti: 94°C per 30 secondi, 70°C per 2 minuti e 30 secondi; ed infine una terza serie di 25 cicli ciascuno cosi costituito: 94°C per 30 secondi, 65°C per 30 secondi e 68°C per 2 minuti, a cui seguiva un passaggio finale di estensione di 68°C per 10 minuti.

Successivamente sono state condotte ulteriori amplificazioni utilizzando come stampo 1 μL della prima reazione di amplificazione e dei primer "nested" specifici sia per la sequenza parziale cDNA che per i siti di annealing alle estremità 5' e 3', adottando le stesse condizioni descritte precedentemente.

Dopo separazione elettroforetica su gel di agarosio 1,5%, i diversi prodotti RACE sono stati purificati, clonati e sequenziati.

4.6. Amplificazione PCR del DNA genomico e del cDNA

Per l'isolamento delle sequenze genomiche dei geni codificanti per le proteine PDI-like e delle corrispondenti sequenze cDNA l'analisi PCR è stata condotta utilizzando il sistema Expand Long Template PCR System (Roche) con il tampone di reazione 3 in dotazione (10X PCR buffer con 27,5 mM di $MgCl_2$). Le reazioni PCR sono state eseguite in un volume finale di 50 μL, contenente 0,75 μL della miscela di enzimi fornita (*Taq* DNA polimerasi e *Tgo* DNA polimerasi con attività "proofreading"; 5U/μL), 1X del tampone di reazione, 500 μM di ciascun dNTP, 0,25 μM di ciascun primer specifico per le diverse sequenze codificanti per proteine PDI-like e 100-300 ng di DNA genomico o 1 μL di reazione di trascrittasi inversa come templato.

Nel caso di DNA genomico, questo è stato inizialmente denaturato a 94°C per 2 minuti, sono quindi seguiti 30 cicli così composti: 94°C per 2 minuti, 57-65°C per 1 minuto (a seconda della temperatura ottimale di annealing di ogni coppia di primer), e 68°C per un numero variabile di minuti (a seconda della lunghezza dell'amplicone). Dall'undicesimo ciclo in poi il tempo di estensione scelto è stato allungato di ulteriori 20 secondi per ogni ciclo successivo. La reazione di PCR è terminata con uno step di estensione finale a 68°C per 7 minuti.

Per l'isolamento delle corrispondenti sequenze cDNA è stato adottato il seguente programma di amplificazione: dopo 2 minuti di denaturazione iniziale a 94°C sono seguiti 30

cicli di amplificazione, ciascuno così composto: 94°C per 30 secondi, 57-65°C per 1 minuto e 68°C per 3-5 minuti e quindi uno step finale di estensione a 68°C per 7 minuti.

I primer usati per amplificare l'intera regione codificante dei geni PDI-like identificati durante lo svolgimento di questa tesi sono stati disegnati nelle relative zone 5' e 3' trascritte ma non tradotte. In tabella 4.1 sono elencate le coppie di primer usate per amplificare ed isolare i 9 geni PDI-like.

Tab. 4.1 - Elenco dei primer forward e reverse, relativi alle regioni 5' e 3' non tradotte, usati per l'amplificazione dell'intera regione codificante dei geni PDI-like identificati.

Gene	Primer	sequenza nucleotidica
*TaPDIL**2**-1*	**2**-1F1	5'-CGTCCGCGCTCCCACTTTCC-3'
	2-1R1	5'-TGTTCTCAGCATACACCTCCAAC-3'
*TaPDIL**3**-1*	**3**-1F1	5'-TTCTAGATCGCGAGCAGCA-3'
	3-1R1	5'-CCCAACCTGCTTGAGCTTT-3'
*TaPDIL**4**-1*	**4**-1F1	5'-GACTGGAGCACGAGGACACT-3'
	4-1R1	5'-GCCGCTAAACTTTCACTGC-3'
*TaPDIL**5**-1*	**5**-1F1	5'-TCGATCGCCACCATCTAGC-3'
	5-1R1	5'-CCACCTTGCACATCAGAGCTT-3'
*TaPDIL**6**-1*	**6**-1F1	5'-CTCCGATCGAGCCATGGATC-3'
	6-1R1	5'-TAAGCCCAAGCCCGCAAA-3'
*TaPDIL**6**-2*	**6**-2F1	5'-GTCCGTGCTAGCCATGGAT-3'
	6-2R1	5'-AGGTTGAGACAATGTTGCAAGC-3'
*TaPDIL**7**-1*	**7**-1F1	5'-CAGACGCGGCGGAAGATCGAA-3'
	7-1R1	5'-TGATACACCATATAACCCTCGCA-3'
*TaPDIL**7**-2*	**7**-2F1	5'-CCACTTGTACAAAGAAGCTGGGT-3'
	7-2R1	5'-GACATGAAATCGCTAAACAGAGG-3'
*TaPDIL**8**-1*	**8**-1F1	5'-GGCACACCTTTCGAAACG-3'
	8-1R1	5'-GACTGGTGACAAGATCAGGATGA-3'

Per tre geni PDI-like, *TaPDIL**2**-1*, *TaPDIL**4**-1* e *TaPDIL**5**-1*, sono state amplificate anche le corrispondenti sequenze genomiche. Inizialmente per l'amplificazione del DNA genomico dei tre geni sono state utilizzate le stesse coppie di primer F1-R1 usate precedentemente per l'amplificazione dell'intera regione codificante dei trascritti (Tab. 4.1). Successivamente sono stati usati primer interni (Tab. 4.2), alcuni dei quali disegnati in base alla sequenza cDNA, mentre altri sono stati disegnati in seguito in base alle nuove sequenze genomiche parziali ottenute dall'amplificazione con le coppie di primer interni precedentemente disegnati. È stato così possibile amplificare l'intera sequenza genomica per i tre geni. Come templato per le successive amplificazioni interne è stata usata la sequenza amplificata e clonata con la coppia di primer più esterna (F1-R1). Tutti i prodotti di amplificazione, sia quelli derivanti dall'amplificazione del cDNA che quelli derivanti dal DNA genomico, sono stati separati in gel di agarosio 1,5%, purificati con "GFX PCR DNA and Gel Band Purification Kit" (Amersham Biosciences), clonati e sequenziati.

Tab.4.2 – Elenco dei primer forward e reverse usati per l'amplificazione delle sequenze genomiche intere relative ai geni *TaPDIL**2**-1*, *TaPDIL**4**-1* e *TaPDIL**5**-1*. I primer disegnati in base a sequenze genomiche sono sottolineati.

Primer/gene	Sequenza nucleotidica dei primer
*TaPDIL**2**-1*	
2-1F1	5'-CGTCCGCCGCTCCCACTTTCC-3'
2-1F2	5'-CGAGTACCTCATCAAGAACGC-3'
2-1F3	5'-ATGAGCTTGCTGCTGCTT-3'
2-1F4	5'-CGTCAAAGTTGTTGTTGGCAA-3'
2-1F5	5'-CCCACCCTCCTCTTCTTCAT-3'
2-1F6	5'-TGATCACCACCCTCACACA-3'
2-1F7	5'-GTGGCATCGACTCCCTTGTAA-3'
2-1R1	5'-TGTTCTCAGCATACACCTCCAAC-3'
2-1R2	5'-CCTACAACTCGTCCTTCGGG-3'
2-1R3	5'-GAGATACTTGGCCAGCTTGTT-3'
2-1R4	5'-GCTCTGAACTCGCCATCTG-3'
2-1R5*	5'-TACATCTCCACCACCGTCCG-3'
*TaPDIL**4**-1*	
4-1F1	5'-GACTGGAGCACGAGGACACT-3'
4-1F2	5'-ATGGCGACCCCTCAGATC-3'
4-1F3	5'-TCTGACCGAGGAAACCTTTGA-3'
4-1F4	5'-AGCGTCCACTTTAAGGGTCAT-3'
4-1F5	5'-AGGTGTGGCCACTGCAAGA-3'
4-1R1	5'-GCCGCTAAACTTTCACTGC-3'
4-1R2	5'-CCACTAAGAGGAGAAGGCTGA-3'
4-1R3	5'-AGCTTTGTTACCCTTCGGGAAGAA-3'
4-1R4	5'-TTATGTCCATTAGAGCGGTGC-3'
4-1R5	5'-CAGTGCGTTGACCCTCATACC-3'
*TaPDIL**5**-1*	
5-1F1	5'-TCGATCGCCACCATCTAGC-3'
5-1F2	5'-GGCACTGTAAGCAGCTGACG-3'
5-1F3	5'-TTGTGGAGTTCTTTGCACCG-3'
5-1F4	5'-GGCTGCAAAGAACTTGAAGGG-3'
5-1F5	5'-CCCTCCAAACCATCTCGTGTT-3'
5-1F6	5'-CACCAGTGGATTATGAAGGCG-3'
5-1R1	5'-CCACCTTGCACATCAGAGCTT-3'
5-1R2	5'-GCTCTCCTTATCGGCACCAA-3'
5-1R3	5'-CAAAGCAACGCTATCCTTGGC-3'
5-1R4	5'-AGCTACATACCACGGTGCAA-3'

Le combinazioni di primer usate per l'amplificazione del DNA genomico relativo ai tre geni sono elencate qui sotto.

*TaPDIL**2**-1*: **2**-1(F2-R4), **2**-1(F4-R2), **2**-1(F3-R3),**2**-1(F5-R4), **2**-1(F7-R5) e **2**-1(F6-R3).

*TaPDIL**4**-1*: **3**-1(F2-R3), **3**-1(F3-R2), **3**-1(F4-R4), **3**-1(F5-R4) e **3**-1(F4-R5).

*TaPDIL**5**-1*: **4**-1(F2-R2), **4**-1(F3-R1), **4**-1(F3-R2), **4**-1(F5-R4), **4**-1(F6-R4) e **4**-1(F4-R3).

4.7. Clonaggio dei prodotti di amplificazione

I prodotti di amplificazione purificati sono stati inseriti mediante ligasi nel sito di restrizione modificato *Eco*RV del vettore plasmidico pGEM-T Easy Vector (Promega).
Per ottimizzare il rapporto molare tra inserto e vettore è stata applicata la seguente formula:

$$\frac{\text{ng del vettore x kb dell'inserto}}{\text{kb del vettore}} \text{ x rapporto molare (I/V) = ng dell'inserto}$$

Nel caso del pGEM-T Easy Vector, si usano circa 50 ng del vettore (V) che ha lunghezza di 3,0 kb e si provano rapporti molari (I/V = inserto/vettore) da 1:3 a 3:1.

In una provetta da 0,5 ml in ghiaccio si uniscono 5 μL di 2x Rapid Ligation Buffer, 1 μL di pGEM-T Easy Vector (50 ng), 1 μL di T4 DNA Ligase (3 Weiss units/μL) e i μL corrispondenti alle quantità necessarie del prodotto di amplificazione. Si porta ad un volume finale di 10 μL con acqua deionizzata. La ligasi viene incubata per una notte a 16°C.

Per la trasformazione sono state usate cellule competenti di *Escherichia coli*, ceppo DH5α. Le cellule trasformate sono quindi state piastrate su terreno di coltura LB contenente ampicillina (100 μg/ml), necessaria per selezionare le cellule contenenti il plasmide, dato che quest'ultimo reca un gene Amp^r per la resistenza a questo antibiotico. Al terreno di coltura di è stato aggiunto IPTG ed X-gal, un cromogeno analogo al galattosio, che ha permesso la selezione delle cellule contenenti plasmidi ricombinanti attraverso lo screening per colorazione bianco/blu.

Dopo l'estrazione del DNA plasmidico a partire da cellule di cloni positivi (Maniatis et al., 1982) i plasmidi ricombinanti sono stati digeriti con gli enzimi *Eco*RI e *Not*I, i cui siti di restrizione sono presenti alle due estremità del polylinker del vettore plasmidico pGEM-T Easy Vector, e separati su gel di agarosio per valutare le dimensioni dei frammenti inseriti. Per le analisi delle sequenze è stata effettuata una purificazione mediante il "Plasmid Mini Prep Kit" della Qiagen seguendo le indicazioni del produttore. Il sequenziamento è stato eseguito dalla MWG-Biotech AG, Germania. Come primer sono stati usati quelli vettore-specifici, rispettivamente T7 (5'-AATACGACTCACTATAG-3') e SP6 (5'-TACAGGAAACAGCTATGACC-3').

4.8. Analisi delle sequenze nucleotidiche

I dati relativi alle sequenze ottenute sono state analizzati mediante il software DNAMAN (Sequence Analysis Software for Windows) (Lynnon BioSoft, 1999). In particolare, sono stati utilizzati i programmi "Two sequence alignment", "Multiple Alignment", "Sequence Assembly" e "Restriction Analysis" del suindicato software, per rimuovere le sequenze nucleotidiche del vettore, correggere eventuali errori, assemblare sequenze dei diversi cloni che si allineavano perfettamente solo alle estremità, costruire le mappe di restrizione e determinare la struttura esoni/introni delle sequenze genomiche. Questi ultimi risultati sono stati confermati anche utilizzando il software per lo studio dei siti di splicing sviluppato da Hebsgaard et al. (1996) disponibile sul sito Internet http://www.cbs.dtu.dk/services/NetGene2 traduzione delle sequenze cDNA nelle corrispondenti proteine è stata effettuata mediante il programma "Translating DNA sequence to protein" ed il confronto delle sequenze nucleotidiche ed amminoacidiche dedotte mediante il programma "Multiple Alignment", mentre i valori dei pesi molecolari e dei punti isolelettrici delle proteine sono stati stimati mediante i programmi "Prediction of molecular weight" e "Prediction of charge and pI".

L'analisi delle sequenze aminoacidiche dedotte dei cloni cDNA codificanti per proteine PDI-like è stata condotta ricercando motivi conservati nelle banche dati Pfam HMMs (Bateman et al., 2000), InterPro (Apweiler et al., 2001) e SMART (Simple Motif Architecture Research Tool) versione 4 (Schultz et al., 2000; Letunic et al, 2002) disponibili rispettivamente presso i siti Internet: http://pfam.wustl.edu/hmmsearch.shtml/, http://www.ebi.ac.uk/interpro/ e http://smart.embl-heidelberg.de.

4.9. Analisi filogenetica

Allo scopo di determinare le relazioni evolutive tra le sequenze cDNA clonate da frumento ed altri geni codificanti per proteine PDI-like, precedentemente isolati da diverse specie vegetali, è stata condotta un'analisi filogenetica utilizzando le sequenze aminoacidiche di geni codificanti per proteine PDI-like, 13 isolati da *Arabidopsis*, 11 da riso, 12 da mais, 2 dal muschio *Physcomitrella patens* ed uno ciascuno da erba medica, ricino, orzo ed alga unicellulare *Chlamydomonas reinhardtii* (Tab. 4.3) insieme con le proteine dedotte di 10 sequenze cDNA di frumento, di cui nove isolate durante lo svolgimento della tesi ed una

relativa alla PDI classica (I gruppo filogenetico), localizzata nel cromosoma 4A e precedentemente clonata e caratterizzata (Ciaffi et al., 2006). Le sequenze aminoacidiche dedotte dei cloni di frumento isolati sono state ottenute mediante il programma "Translating DNA sequence to protein" del software DNAMAN. L'allineamento delle sequenze proteiche è stato ottenuto mediante il programma ClustalX versione 1.83 (Thompson et al., 1997) utilizzando i seguenti parametri: gap penalty 8, gap extension penalty 2, PAM protein weight matrix. L'albero filogenetico è stato costruito tramite il metodo neighbor-joining (NJ) (Saitou e Nei, 1987) utilizzando il programma NEIGHBOR del software PHYLIP versione 3.6 (Felstein, 1989) e le matrici delle distanze genetiche sono state stimate con il programma PRODIST dello stesso software, utilizzando il modello PAM per la transizione degli aminoacidi. Per valutare la significatività statistica dell'albero ottenuto sono state ottenute 1000 repliche mediante analisi bootstrap utilizzando il programma SEQBOOT del software PHYLIP.

Tab. 4.3 – Sequenze geniche PDI-like di diverse specie vegetali utilizzate per l'analisi filogenetica.

Designazione PDIL	N° accessione cDNA	Specie	Designazione PDIL	N° accessione cDNA	Specie
AtPDIL1-1	NM_102024	*A. thaliana*	OsPDIL4-1	NM_185280	*O. sativa*
AtPDIL1-2	NM_106400	*A. thaliana*	OsPDIL4-2	NM_183927	*O. sativa*
AtPDIL2-1	NM_115353	*A. thaliana*	OsPDIL5-1	AK062254	*O. sativa*
AtPDIL2-2	NM_180903	*A. thaliana*	OsPDIL6-1	AK063663	*O. sativa*
AtPDIL3-1	NM_104105	*A. thaliana*	OsPDIL7-1	AK069367	*O. sativa*
AtPDIL3-2	NM_112481	*A. thaliana*	OsPDIL8-1	AK099660	*O. sativa*
AtPDIL4-1	NM_130315	*A. thaliana*	PpPDIH	AF430644	*P. patens*
AtPDIL5-1	NM_100376	*A. thaliana*	PpPDIM	AF430645	*P. patens*
AtPDIL5-2	NM_128852	*A. thaliana*	RcPDIL1-1	U41385	*R. communis*
AtPDIL6-1	NM_202059	*A. thaliana*	ZmPDIL1-1	AY739284	*Z. mais*
AtPDIL7-1	NM_103262	*A. thaliana*	ZmPDIL1-2	AY739285	*Z. mais*
AtPDIL8-1	NM_112948	*A. thaliana*	ZmPDIL2-1	AY739286	*Z. mais*
AtPDIL8-2	NM_118842	*A. thaliana*	ZmPDIL2-2	AY739287	*Z. mais*
CrRB60	AF027727	*C. reinhardtii*	ZmPDIL3-1	AY739295	*Z. mais*
HvPDIL1-1	L33251	*H. vulgare*	ZmPDIL4-1	AY739288	*Z. mais*
MsPDIL1-1	Z11499	M. sativa	ZmPDIL4-2	AY739289	*Z. mais*
OsPDIL1-1	AK068268	*O. sativa*	ZmPDIL5-1	AY739290	*Z. mais*
OsPDIL1-2	AY739308	*O. sativa*	ZmPDIL6-1	AY739291	*Z. mais*
OsPDIL1-3	AK073161	*O. sativa*	ZmPDIL7-1	AY739292	*Z. mais*
OsPDIL2-1	AK071514	*O. sativa*	ZmPDIL7-2	AY739293	*Z. mais*
OsPDIL3-1	AK073970	*O. sativa*	ZmPDIL8-1	AY739294	*Z. mais*

4.10. Analisi Southern

4.10.1. Preparazione dei filtri di ibridazione (Southern blot)

Circa 15 µg di DNA della cv CS e delle sue linee nulli-tetrasomiche e ditelosomiche sono stati digeriti per 6 ore a 37°C con 8 U/µg di enzima di restrizione. Sono stati usati i seguenti enzimi di restrizione: *Bam*HI, *Eco*RI, *Eco*RV, *Hin*dIII, *Kpn*I, *Pst*I e *Sac*I. I frammenti di DNA digerito sono stati quindi separati in base alla loro lunghezza su gel di agarosio 1% in 1x TBE per 16-18 ore a 40V. Terminata l'elettroforesi, la digestione del DNA è stata verificata ai raggi UV dopo colorazione del gel in una soluzione di etidio bromuro (0,5 µg/mL) per circa 10 minuti e risciacquo con acqua distillata.

Per il trasferimento del DNA sul filtro di nylon (Hybond-N+ Amersham) è stato usato il metodo di Southern (1975). Il gel contenente il DNA è stato immerso per 45 minuti in una soluzione denaturante contenente NaOH 0,5 M e NaCl 1,5 M. E' seguito un ulteriore passaggio di 45 minuti in una soluzione neutralizzante (NaCl 1,5 M; Tris 0,5 M, portato a pH 8,0 con HCl). I frammenti di restrizione presenti nel gel sono stati sottoposti a trasferimento per capillarità utilizzando come tampone di trasferimento SSC 10X (NaCl 150 mM; citrato di sodio 15 mM). Il trasferimento dura circa 20 ore, quindi il DNA trasferito viene fissato alla membrana mediante irraggiamento UV (120mJ).

4.10.2. Marcatura delle sonde

Le sonde sono state marcate mediante incorporazione nel DNA di digossigenina-11-dUTP attraverso reazione di amplificazione PCR, utilizzando il kit PCR-DIG Probe Synthesis (Roche). Come templati sono stati usati circa 10 pg di DNA plasmidico, ottenuto, linearizzando mediante digestione enzimatica e purificando mediante precipitazione con 2,5 volumi di etanolo e 1/10 acetato di potassio 3M a -80°C, i rispettivi cloni cDNA ottenuti durante lo svolgimento della tesi.

Per l'amplificazione PCR sono state utilizzate le sotto elencate coppie di primer, attraverso le quali sono state marcate le differenti sonde usate nelle analisi Southern e northern. Per l'analisi Southern sono state usate la S-F4-R3, la S-F3-R5 e la S-F3-R4, mentre per l'analisi northern sono state usate tutte le sonde eccetto la S-F3-R5.

TaPDIL4-1

- Sonda **S-F4-R3**, lunga 763 pb, complementare alla seconda metà della sequenza codificante del gene e a parte della sequenza 3' non tradotta (117 pb):
  F4: 5'-TCTGACCGAGGAAACCTTTGA-3'
  R3: 5'-GCCGCTAAACTTTCACTGC-3'
- Sonda **S-F3-R5**, lunga 601 pb, complementare alla prima metà della sequenza codificante:
  F3: 5'-CACGGAGTCCACCTTCGAGA-3'
  R5: 5'-AGCTTTGTTACCCTTCGGGAAGAA-3'

TaPDIL5-1

- Sonda **S-F3-R4**, lunga 918 pb, complementare alla seconda metà della sequenza codificante e a parte della sequenza 3' non tradotta (150 pb):
  F3: 5'-TTGTGGAGTTCTTTGCACCG-3'
  R4: 5'-CCACCTTGCACATCAGGCTT-3'

TaPDIL2-1

- Sonda **S-F1-R1**, lunga 544 pb, complementare al quarto terminale della sequenza codificante e a parte della sequenza 3' non tradotta (109 pb):
  F1: 5'-CGTCAAAGTTGTTGTTGGCAA-3'
  R1: 5'-CAGCATACACCTCCAACGGC-3'

TaPDIL1-1

- Sonda **S-F2-F3**, lunga 516 pb, complementare ad una zona centrale rispetto ai due siti attivi:
  F2: 5'-CATGCCAACCATCTCCCACG-3'
  R3: 5'-AATAGGCTCGGACTTCCTGAATGG-3'

La reazione di amplificazione è stata condotta effettuando una denaturazione preliminare a 95°C per 2 minuti, seguita da 35 cicli di amplificazione: 95°C per 30 secondi, 62°C per 30 secondi, 72°C per 1 minuto. Al termine dell'ultimo ciclo è stato eseguito un ciclo di incubazione di 7 minuti a 72°C.

La reazione di amplificazione PCR è avvenuta in un volume di 50 µL contenente 2,6 unità di enzima (Expand High Fidelity; Roche), buffer di reazione 1x (Expand High Fidelity buffer 10X con $MgCl_2$ (15 mM); Roche), 1X PCR DIG probe synthesis mix (10X contiene: 2 mM di dATP, dGTP, dCTP; 1,3 mM dTTP, 0,7 mM DIG-11-dUTP (digossigenina-11-

uridina-5'-trifosfato); Roche), e 0,25 μM di ciascun primer. Terminata la reazione il DNA amplificato è stato precipitato con 2,5 volumi di etanolo e 1/10 di acetato di potassio 3M.

La presenza di digossigenina nel DNA rallenterà la corsa dello stesso su gel di agarosio. L'avvenuta incorporazione della digossigenina nella sonda è stato quindi visualizzata correndo la stessa accanto ad un controllo non marcato, su un gel di agarosio 1.5%.

4.10.3. Ibridazione molecolare

Questo procedimento è consistito in quattro fasi: a) preibridazione b) ibridazione c) lavaggio e d) rilevamento immunologico.

a) **Preibridazione.** La soluzione di preibridazione per i filtri contenenti DNA è composta da SSC 5X; 0,1 % di N-lauroylsarcosine; 0,02 % (w/v) di sodio dodecil solfato (SDS); 0,5% (w/v) di blocking reagent (Boehringer). La membrana è stata tenuta in preibridazione per 4 ore alla temperatura di 65°C.

b) **Ibridazione.** La sonda, diluita nella soluzione di ibridazione, è stata denaturata ponendola in acqua bollente per 10 minuti e quindi in ghiaccio per 5 minuti. I filtri, immersi in 50 mL di questa soluzione, sono stati ibridati per 16-20 ore a 65°C. Alla fine di questa fase la sonda è stata recuperata e conservata a 4°C per essere riutilizzata.

c) **Lavaggio.** Il filtro è stato lavato due volte a temperatura ambiente per 5 minuti in una soluzione a bassa stringenza contenente SSC 2x e SDS 0,1% (w/v) e quindi 2 volte per 15 minuti a 65°C con una soluzione ad alta stringenza contenente SSC 0,1x e SDS 0,1%.

d) **Rilevamento immunologico.** È stato eseguito seguendo la metodologia descritta da Lu et al. (1994):

- breve inumidimento dei filtri con un tampone costituito da TRIS pH 7,5 100 mM; NaCl 150 mM;
- incubazione a temperatura ambiente per almeno 30 minuti con un tampone analogo al precente ma contenente 0,5 % di blocking reagent;
- aggiunta di anticorpo coniugato in eccesso (Anti-digoxigenin-AP, Fab fragments; Roche) 75 mU/mL ed incubazione a temperatura ambiente per almeno 30 minuti;
- rimozione dell'anticorpo coniugato non legato mediante due risciacqui a temperatura ambiente con il tampone iniziale (TRIS pH 7,5 100 mM; NaCl 150 mM) per 15 minuti;
- risciacquo per 2 minuti con tampone alcalino (TRIS pH 9,5 100 mM; NaCl 100 mM; $MgCl_2$ 50 mM).

Infine i filtri sono stati tenuti al buio per 5 minuti in tampone alcalino contenente CSPD (Roche) 0,25 mM. Ciascun filtro, dopo essere stato asciugato all'aria per circa 45 minuti, è stato posto in camera oscura assieme ad una lastra fotografica per un tempo di esposizione sufficiente alla comparsa dei frammenti di restrizione in seguito allo sviluppo della lastra.

4.11. Analisi northern

L'RNA totale è stato quantificato spettrofotometricamente a 260 nm e successivamente controllato attraverso separazione in gel di agarosio 1% in MOPS 1X (20 mM MOPS, 5 mM acetato di sodio e 1 mM EDTA; pH 7,0) contenente formaldeide (2%) e colorato con etidio bromuro.

10 μg di RNA totale estratto dai diversi tessuti (radice, plantula, foglia, cariossidi e spighe) è stato separato su gel d'agarosio, alle medesime condizioni sopra descritte, e trasferito su membrana di nylon caricata positivamente (Roche, Basilea, Svizzera) usando SSC 20X in acqua DEPC. La membrana è stata quindi asciugata all'aria per 45 minuti e fissata agli UV (120mJ).

La reazione di preibridazione è stata condotta per 4 ore a 65°C in una soluzione contenente SSC 3X, 0,1% (w/v) di sodio dodecil solfato (SDS), 5X reagente di Denhart, 60 mM di Na_2HPO_4 pH 6,5 e 25 mM EDTA pH 8,0.

Le sonde sono state marcate mediante incorporazione nel DNA di digossigenina-11-dUTP attraverso reazione di amplificazione PCR, utilizzando il kit PCR-DIG Probe Synthesis (Roche), come precedentemente descritto nel paragrafo 4.10.2. Anche in questo caso, come per il Southern, la sonda prima dell'uso è stata diluita nella soluzione di preibridazione e denaturata ponendola in acqua bollente per 10 minuti e quindi in ghiaccio per 5 minuti. I filtri sono stati quindi immersi in 50 mL di questa soluzione ed ibridati per 16-20 ore a 65°C. Alla fine di questa fase la sonda è stata recuperata e conservata a 4°C per essere riutilizzata.

I lavaggi ed il rilevamento immunologico sono stati condotti come nel paragrafo 4.10.3. punti c e d, eccetto che invece del CSPD al tampone alcalino è stato aggiunto 0,55 mM nitroblu tetrazolio cloruro (NBT) e 0,4 mM 5-bromo-4-cloro-3-indolil-fosfato (X-fosfato).

4.12. Ottenimento dei costrutti molecolari

4.12.1. Preparazione ed analisi dei plasmidi donatori delle sequenze di interesse

La sequenza regolatrice in 5' della *TaPDIL**1**-1*, quella del promotore della 1Dx5 e quella del cDNA totale della *TaPDIL**1**-1*, dopo essere state amplificate, sono state inserite mediante ligasi nel sito di restrizione modificato *Eco*RV del vettore plasmidico pGEM-T (Promega), come descritto nel paragrafo 4.7., e quindi sequenziate. I primer usati per l'amplificazione dell'intera regione codificante per la *TaPDIL**1**-1* sono:

TaPDIL1-1-F6: 5'-GTCGACATCACTGCTCCCCAGT-3'

TaPDIL1-1-R5: 5'-TCTCGCAACGGGAAACAG-3'

I primer per l'amplificazione della regione regolatrice in 5' della 1Dx5 sono:

1Dx5-5'-F1: 5'-GGGAAGTTAGCCATGCCGA-3'

1Dx5-5'-R1: 5'-GGACTATCAGTGAATTGATCTCTAG-3'

Infine i primer usati per l'amplificazione della regione regolatrice in 5' della *TaPDIL**1**-1* sono:

TaPDIL1-1-5'-F4: 5'-GCTTTCCAGCTTGAGATGAGGC-3'

TaPDIL1-1-5'-R4: 5'-GGGAGCAGTGATGAGATTTC-3'

Per l'amplificazione della sequenza codificante e regolatrice della *TaPDIL**1**-1* è stato usato DNA estratto dalla cultivar Chinese Spring, mentre per l'amplificazione della regione promotrice della 1Dx5 è stato usato DNA estratto dalla cultivar Centauro. Tramite il software DNAMAN sono state quindi analizzati i diversi siti di restrizione dei cloni pGEM-T per confrontarli con quelli dei due plasmidi pAHC25 e pAHC20. Questi due plasmidi sono i donatori delle sequenze nucleotidiche relative al terminatore NOS, al gene *UidA* e al plasmide pUC8. Quest'ultima sequenza, in particolare, serve per formare lo scheletro dei nuovi plasmidi, ai quali conferisce contemporaneamente la resistenza all'ampicillina. Sono stati così individuati i siti di restrizione comuni usati nelle successive fasi di clonazione per l'ottenimento dei costrutti per la sovra-espressione e lo studio del promotore. Quando non erano presenti, siti di restrizione adatti sono stati inseriti alle estremità 5' e 3' dell'inserto tramite amplificazione con dei primer adattatori.

4.12.2. Amplificazione degli inserti con primer adattatori

Le sequenze promotrici della *TaPDIL**1**-1* e della 1Dx5 sono state amplificate mediante primer adattatori usando come templato 1 ng di DNA plasmidico, ottenuto dai rispettivi cloni

pGEM-T, linearizzato mediante digestione enzimatica con *Pst*I e purificato mediante precipitazione a -80°C con 2,5 volumi di etanolo e 1/10 di acetato di potassio 3M.

La reazione di amplificazione è stata condotta effettuando una denaturazione preliminare a 94°C per 2 minuti, seguita da 29 cicli di amplificazione: 94°C per 30 secondi, 64°C o 57°C per 1 minuto, 72°C per 2 minuti. Al termine dell'ultimo ciclo è stato eseguito un ciclo di incubazione di 7 minuti a 72°C. La reazione è stata eseguita in un volume finale di 25 µL contenente: 2mM di $MgCl_2$, 50 µM di ciascun dNTP, 0,2 µM di ciascun primer, 1X del buffer di reazione (CIMMYT *Taq* Buffer, Mg-free), 1 unità di *Taq* (CIMMYT *Taq* enzyme).

Sono state fatte diverse ripetizioni per ogni tipo di amplificazione. Queste, dopo essere state singolarmente controllate correndo 3 µL su gel di agarosio 1,2 %, sono state unite e purificate con il kit "Wizard PCR Preps DNA Purification system" della Promega. Il DNA è stato quindi risospeso in 50 µL di acqua deionizzata e filtrata e la concentrazione di DNA è stata stimata mediante separazione su gel di agarosio di una aliquota di 1-2 µL. Una quantità di circa 1-2 µg è stata quindi digerita con gli enzimi per i quali erano stati introdotti i siti di restrizione alle due estremità della sequenza. Il frammento di restrizione è stato quindi purificato da gel di agarosio ed eluito in 12-20 µL di acqua deionizzata e filtrata.

I primer adattatori sono stati disegnati in modo da avere 6 o 7 basi all'estremità 5' a monte del sito di restrizione, in modo da assicurare il perfetto funzionamento degli enzimi di restrizione.

Per la sovra-espressione è stata usata la seguente coppia di primer a 64°C:

AF1= 5'-ATCATAGACGTCTGCAGGGAAGTTAGCCATGCCGA-3'
AR1= 5'-CTGAGCGACGTCGGACTATCAGTGAATTGATCTCTA-3'

Per lo studio del promotore sono stati usati i seguenti primer a 57°C:

AF4=5'-ACGTCGACTGCAGGCTTTCCAGCTTGAGATGAGGC-3'
AF5=5'-ACGTCGACTGCAGAGGACTCATCAGAAACGCA-3'
AF6=5'-ACGTCGACTGCAGTCCGCCATATCTTCCTAGTC-3'
AF7=5'-ACGTCGACTGCAGAGGGTACGTGTTTGCAGC-3'
AR4=5'-ACAGTAGTCTAGAGGGAGCAGTGATGAGATTTC-3'

4.12.3. Ottenimento dei frammenti di restrizione dai plasmidi e dagli amplificati

I plasmidi (pAHC25, pAHC20 ed il clone pGEM-T contenente la sequenza cDNA totale della *TaPDIL1-1*) e le reazioni di amplificazione eseguite con i primer adattatori sono stati digeriti con gli enzimi di restrizione scelti per ottenere vettori ed inserti dotati di estremità adesive complementari.

Le reazioni di restrizione sono state eseguite in un volume di 40-50 µL, con una concentrazione finale 1X del buffer di restrizione e almeno 2,5 U di enzima per µg di DNA. Nel caso di DNA plasmidico sono stati digeriti 4-5 µg, mentre nel caso di DNA ottenuto mediante amplificazione sono stati digeriti 1-2 µg. La digestione è stata effettuata a 37°C per 2-3 ore.

I frammenti di restrizione sono stati separati mediante analisi elettroforetica per individuare i frammenti dalla grandezza attesa, che purificati da gel ed eluiti in 12-20 µL di acqua sterile sono pronti per la fase successiva di ligasi.

Nelle clonazioni non direzionali, il vettore, in quanto digerito con un unico enzima di restrizione e quindi dotato di estremità adesive compatibili, che durante la reazione di ligasi possono portare alla ricircolarizzazione, dopo digestione è stato defosforilato.

La reazione di defosforilazione è stata eseguita usando la SAP (Shrimp Alcaline Phosphatase; Promega) alle seguenti condizioni: 37°C per 30 minuti. La reazione è stata eseguita in un volume finale di 30-50 µL con 1X di SAP buffer e 2 unità di SAP per µg di DNA del vettore digerito. Dopo essere stato defosforilato il DNA è stato purificato mediante il kit "Wizard PCR Preps DNA Purification system" della Promega, per eliminare l'enzima SAP, e risospeso in 12-20 µL di acqua deionizzata e filtrata. Per controllare la completa defosforilazione del vettore, è stato inserita nell'esperimento una reazione di ligasi di controllo.

4.12.4. Ligasi e trasformazione

Per controllare la grandezza e la purezza dei frammenti di restrizione del vettore e dell'inserto e contemporaneamente stimarne la concentrazione di DNA, 2 µL di ciascun eluito sono stati separati su gel di agarosio. Per ogni esperimento sono stati provati diversi rapporti molari tra vettore ed inserto (1:3, 1:10, 1:25 e 1:50).

La reazione di ligasi è stata condotta in un volume finale di 20 µL con 4 µL di 5X ligase reaction buffer (Invitrogen), 1-2 µL di T4 DNA ligase (Invitrogen) ed una

concentrazione finale di DNA (vettore + inserto) non superiore ai 10 ng/µL. Le quantità relative di DNA di vettore e di inserto sono state calcolate secondo il rapporto molare scelto e la formula già vista precedentemente. Per il vettore sono stati usati 50-100 ng di DNA. La ligasi è stata tenuta per un'ora a temperatura ambiente (22°C) e quindi posta a 16°C per 16-18 ore. Nei casi particolari di clonazione direzionale, in cui vengono usati due differenti enzimi di restrizione per tagliare le due estremità 5' e 3' del vettore e dell'inserto, la reazione di ligasi è stata effettuata a temperatura ambiente (22°C) per 1 ora e quindi immediatamente usata per la trasformazione di cellule competenti. Le cellule trasformate sono quindi state piastrate su terreno di coltura LB contenente ampicillina (100 µg/ml).

L'esattezza dei diversi cloni è stata controllata innanzitutto mediante analisi di restrizione, usando per i casi di clonazione direzionale gli stessi enzimi usati per preparare il frammento da ligare, e per i casi di clonazione non direzionale enzimi capaci di discernere quei cloni in cui l'inserto si era inserito nel vettore nel giusto orientamento. L'esattezza dei plasmidi finali ottenuti è stata infine confermata mediante sequenziamento. In questo caso per il sequenziamento sono stati usati sia i primer universali per il plasmide pUC8 (F: 5'-TGTAAAACGACGGCCAGT-3' e R: 5'-AACAGCTATGACCATG-3') che i primer usati per l'amplificazione delle due sequenza regolatrice. Per la 1Dx5-5' questi sono **1Dx5-5'**-F1 e **1Dx5-5'**-R1 mentre per la *TaPDIL1-1* i primer sono **TaPDIL1-1**-F4 e **TaPDIL1-1**-R4.

Il DNA dei plasmidi finali per la sovra-espressione e lo studio del promotore è stato quindi prodotto in grandi quantità inoculando con le rispettive colonie batteriche 5 mL di terreno di coltura LB contenente 100 µg/mL di ampicillina. Dopo un giorno intero a 37°C e 225-250 rpm, con questi 5 mL sono stati inoculati altri 250 mL di terreno di coltura LB contenente 100 µg/mL di ampicillina, lasciando crescere il tutto a 37°C per 18-20 ore a 225-250 rpm. Quindi il liquido è stato trasferito e i batteri pellettati in una Corning da 50 mL. I plasmidi sono stati estratti e purificati con il kit di purificazione per maxi prep della Quiagen. Il DNA plasmidico è stato quindi controllato mediante analisi di restrizione e quantificato spettrofotometricamente. Sono state preparate diverse aliquote di 10 µL (0,5 µg/µL) in provette da 0,5 mL conservate a -20°C in attesa della trasformazione biolistica.

4.13. Trasformazione biolistica

Tutti i protocolli descritti di seguito provengono dal CIMMYT (Centro Internazionale per il Miglioramento del Mais e del Frumento, Messico) (Pellegrineschi et al., 1998; Pellegrineschi et al., 2002). I terreni di coltura sono basati sul terreno MS di Murashige e Skoog (1962).

4.13.1. Raccolta e sterilizzazione delle cariossidi immature

Per l'allevamento delle piante donatrici della cultivar di frumento tenero MBP-Bobwhite 26 vedere paragrafo 4.1. Le spighe vengono raccolte 12-16 giorni dopo l'antesi in seguito ad una valutazione visiva dello stadio di maturazione complessivo delle cariossidi. Le cariossidi che si trovano nello stadio di maturazione ideale vengono prese e sterilizzate immediatamente sotto cappa biologica. Per la sterilizzazione si pongono le cariossidi in un becker da 1 L contenente 500 mL di una soluzione 1,8% di ipoclorito di sodio (Cloralex, 6% di cloro libero) ed una goccia di Tween 80 e si lascia il tutto in agitazione con magnete per 30 minuti. In seguito si eseguono 3 lavaggi consecutivi di 5 minuti con acqua deionizzata ed autoclavata.

4.13.2. Isolamento e trattamento degli scutelli estratti da cariossidi immature

Dalle cariossidi così trattate vengono estratti gli embrioni immaturi usando microscopio, pinzette e bisturi. Con un taglio netto, effettuato all'estremità embrionale del lato dorsale della cariosside, avviene l'estrazione dell'embrione immaturo (0,7-1,2 mm) e la contemporanea rimozione dell'asse embrionale, necessaria a prevenire la germinazione precoce. Gli espianti destinati al bombardamento consistono, quindi, dell'unico cotiledone embrionale chiamato scudetto o scutello.

Gli scutelli vengono immediatamente adagiati su un terreno di coltura osmotico (E3M-Osmotic media) contenuto in piastre Petri, avendo cura di porli con la parte del taglio in diretto contatto con il terreno e di organizzarli in gruppi di 30-40 in una zona bersaglio centrale. Si prepara anche una piastra Petri di controllo, in cui si pongono 15-30 scutelli. Ogni esperimento è consistito in almeno 160 embrioni bombardati. Dopo qualche ora gli espianti vengono bombardati ed in seguito posti per tutta la notte a 26-27°C al buio.

E3M: 200 mL/L MS macroelementi (10X; Sigma), 100 mL/L MS microelementi (10X; Sigma), 1 mL/L MS vitamine (1000X; Sigma), 0,40 mg/L Tiamina HCl (Sigma), 0,150 g/L L-Asparagina (Sigma), 15% (w/v) Maltosio (Sigma), 2,5 mg/L 2,4-D (acido 2,4-diclorofenossiacetico; Sigma) e 9 g/L di Bacto-agar (Difco). Il pH va aggiustato a 5,7 prima di autoclavare il terreno.

4.13.3. Bombardamento

È stata lasciata scongelare a temperatura ambiente un'aliquota di 20 µL di particelle di oro con diametro di 0,6 µm (0,05 mg/mL). Per garantire la completa separazione le particelle di oro sono state sonicate per 15 secondi. Le aliquote di particelle di oro, conservate a -20°C, vengono preparate in precedenza lavando con etanolo e risospendendo per sonicazione le particelle d'oro più volte prima di aliquotarle in provette da 1,5 mL pronte per l'uso.

Quindi sono stati aggiunti i DNA di interesse nella quantità di 5 µg per ciascun plasmide usato. In ognuno degli esperimenti sono stati usati contemporaneamente due plasmidi, uno contenente il costrutto di interesse ed un altro conferente la resistenza all'erbicida BASTA (pAHC20 e pAHC25). Anche i DNA sono stati aliquotati in precedenza (vedi paragrafo 4.12.4.), e conservati a -20°C in attesa dell'uso. Ad ogni provetta contenente DNA e particelle d'oro sono stati aggiunti 50 µL di CaCl 2,5 M e 20 µL di spermidina 0,1 M ed il tutto è stato agitato per 15 minuti usando il vortex. Dopo una centrifugata di 30 secondi a 14.000 rpm il supernatante è stato eliminato ed il pellet risospeso con 200 µL di etanolo assoluto. Dopo questo lavaggio è stata eseguita un'ultima centrifugata di 30 secondi a 14.000 rpm ed il pellet è stato risospeso con 100 µL di etanolo assoluto e conservato in ghiaccio, mentre sono stati sterilizzati i "macrocarrier" (Bio-Rad) lasciandoli immersi per 20 minuti in etanolo al 75%. Contemporaneamente sono stati sterilizzati anche i "rupture discs" (900 psi; Bio-Rad) immergendoli in isopropanolo per 10 minuti. Una volta evaporato l'etanolo, i "macrocarrier" sono stati posti negli appositi sostegni. Con una pipetta e movimenti circolari 10 µL delle particelle d'oro ricoperte di DNA sono stati distribuiti al centro di ognuno dei "macrocarrier". È importante anche in questo caso asciugare i "macrocarrier" completamente, lasciando evaporare l'etanolo. Sono state quindi assemblate le varie parti della PDS-1000/He particle gun (Bio-Rad), in modo da lasciare 5 cm di distanza dallo "stopping plate", e quindi è stato creato il vuoto (91,4-94,8 kPa) all'interno della camera. Il vuoto è necessario per evitare che la resistenza dell'aria rallenti i microproiettili. A questo punto è stato aperto il flusso di elio all'interno del tubo di accelerazione. Questo crea una pressione crescente che, una volta

raggiunto il punto di rottura del "rupture disc" di 900 psi (1 psi = 6.895 kPa), viene rilasciata di colpo, creando una onda d'urto, che colpito il "macrocarrier" coperto di microproiettili, lo porta ad urtare contro lo "stopping screen", una rete metallica la cui funzione è bloccare il "macrocarrier", ma permettere ai microproiettili coperti di DNA di raggiungere il tessuto bersaglio sottostante. Questo processo è stato ripetuto per tutte le piastre Petri, eccetto per quella di controllo, utilizzando ovviamente per ogni "sparo" un nuovo "rupture disc" ed un nuovo "macrocarrier" ricoperto di microproiettili.

4.13.4. Induzione dei calli

Il giorno successivo al bombardamento gli scutelli di ogni esperimento sono stati trasferiti, lasciando la parte del taglio rivolta verso il basso, su un nuovo terreno di coltura per l'induzione di calli embriogenetici (E3-callus induction media) e sono stati quindi lasciati per 2-3 settimane al buio a 26-27°C.

E3: 200 mL/L MS macroelementi (10X; Sigma), 100 mL/L MS microelementi (10X; Sigma), 1 mL/L MS vitamine (1000X; Sigma), 0,40 mg/L Tiamina HCl (Sigma), 0,150 g/L L-Asparagina (Sigma), 20 g/L Saccarosio (Sigma), 2,5 mg/L 2,4-D (acido 2,4-diclorofenossiacetico; Sigma) e 9 g/L di Bacto-agar (Difco). Il pH va aggiustato a 5,7 prima di autoclavare il terreno.

4.13.5. Rigenerazione e selezione in vitro dei tessuti

Dopo 2-3 settimane dalla maggior parte degli scutelli si erano formati calli embriogenetici. I restanti scutelli o avevano dato origine a calli non embriogenetici o erano rimasti invariati. I calli embriogenetici sono stati quindi trasferiti su un terreno di rigenerazione (MSR-Regeneration media) a cui sono stati aggiunti 5 mg/L di PPT (DL-fosfinotricina; Sigma) e sono stati allevati con un fotoperiodo di 16 h luce: 8 h buio e 25±1°C di temperatura in una camera di crescita per la rigenerazione e la contemporanea selezione. Il periodo di selezione nei primi esperimenti è durato da 5 a 7 settimane ed ha avuto come riferimento principale dell'avvenuta selezione la piastra di controllo contenente i calli derivanti dagli scutelli non bombardati. Negli esperimenti successivi il periodo di selezione è stato allungato di altre 2-3 settimane.

Le piantine che sono risultate ancora verdi a fine selezione sono state trasferite ad un terreno MSR senza PPT, per permettere un ulteriore crescita delle stesse. La durata di questo periodo è stata molto variabile ed è dipesa dallo stadio di sviluppo e dal vigore delle singole piantine.

MSR: 4,4 g/L MS Sali (Sigma), 1 mL/L MS vitamine (1000X; Sigma), 0,40 mg/L Tiamina HCl (Sigma), 0,150 g/L L-Asparagina (Sigma), 20 g/L Saccarosio (Sigma) e 9 g/L di Bacto-agar (Difco). Il pH va aggiustato a 5,7 prima di autoclavare il terreno.

Le piantine, una volta raggiunto uno sviluppo sufficiente anche nell'apparato radicale, sono state trasferite in vasi contenenti una miscela di terriccio e quindi sono state poste direttamente in serra fino a completa maturazione delle cariossidi, le quali sono state quindi raccolte e conservate. Dalle foglie delle piante allevate in serra è stato estratto il DNA con il metodo CTAB.

4.14. Identificazione delle piante contenenti il costrutto per la sovra-espressione mediante PCR

La qualità e la concentrazione del DNA estratto con il metodo CTAB dalle potenziali piante transgeniche sono state valutate mediante separazione elettroforetica correndo 5 µL su un gel di agarosio 1% insieme a DNA lambda a concentrazioni note: 50 ng, 100 ng e 200 ng.

Per le reazioni PCR sono stati usati circa10 ng di DNA genomico di frumento e 20 ng di DNA plasmidico di controllo (p1Dx5-5'/TaPDIL1-1).

Inizialmente è stata usata la coppia di primer PDI-F e NOS-R. La reazione di amplificazione è stata condotta effettuando una denaturazione preliminare a 94°C per 2 minuti, seguita da 35 cicli di amplificazione così composti: 94°C per 30 secondi, 58°C per 1 minuto, 72°C per 2 minuti. Al termine dell'ultimo ciclo è stato eseguito un ciclo di incubazione di 7 minuti a 72°C. La reazione è stata eseguita in un volume finale di 25 µL contenente: 1mM $MgCl_2$, 50 µM di ciascun dNTP, 0,2 µM di ciascun primer, 1X buffer di reazione (CIMMYT *Taq* Buffer 10X, Mg-free), 1 unità di *Taq* Polymerase (CIMMYT *Taq* enzyme).

Le amplificazioni con la coppia di primer 1Dx5-F e PDI-R alle medesime condizioni descritte sopra, variando inoltre la temperatura tra 60 e 62°C e la concentrazione di magnesio

(1 mM, 1,5 mM, 2 mM e 2,5 mM), non hanno permesso di ottenere il frammento delle dimensioni attese. La reazione di amplificazione con questa coppia di primer è quindi stata eseguita usando la Platinum *Taq* DNA Polymerase della GIBCO, insieme ad un enhancer (10X PCRx Enhancer Solution). Sono state provate diverse concentrazioni dell'enhancer in combinazione con diverse temperature come consigliato dal produttore e alla fine la reazione è stata condotta in un volume finale di 12 µl, con una concentrazione 2X dell'enhancer ed una temperatura di annealing di 55°C. La reazione finale contiene inoltre 1X amplification buffer, 0,3 mM di ciascun dNTP, 1 mM $MgSO_4$, 0,3 µM di ciascun primer, 10 ng di DNA genomico o 20 ng di DNA plasmidico ed 1 unità di Platinum *Taq* DNA Polymerase.

PDI-F: 5'-TACTTCGTCACTCCTAGCG-3'
NOS-R: 5'-TTATCCTAGTTTGCGCGCTA-3'
1Dx5-F: 5'-TCCAATTGCTCCTTGCTTAT-3'
PDI-R: 5'-CCTGAAGATCTTGAGGGTAGG-3'

5. RISULTATI E DISCUSSIONE

5.1. Identificazione e clonazione di sequenze cDNA codificanti per proteine PDI-like in frumento

Allo scopo di identificare sequenze cDNA codificanti per proteine PDI-like in frumento, 11 sequenze di riso, assegnate alle 8 classi filogenetiche della famiglia delle PDI individuate nelle piante (Houston et al., 2005; Tab. 1.3 nella sezione 1.6. dell'introduzione), sono state utilizzate per effettuare una ricerca, mediante il programma BLAST, nella banca dati "TIGR wheat gene index" (TaGI, versione 10), contenente 580.155 ESTs di frumento. Questa analisi ha permesso di individuare 10 distinte sequenze consenso assemblate (TC, Tentative consensus sequence), contenenti almeno 5 EST ottenuti da librerie cDNA derivate preferenzialmente da cariossidi in maturazione, da spighe a diversi stadi di sviluppo e da diversi tessuti fiorali (Tab. 5.1). Una delle sequenze consenso identificate (TC250792), non riportata in tabella, ha evidenziato una significativa omologia con sequenze PDI-like di diverse specie vegetali appartenenti al primo gruppo filogenetico, i cui membri erano stati precedentemente clonati e caratterizzati in frumento (Ciaffi et al., 2006). Una ricerca condotta in "HarvEST Wheat" (versione 1.13) ha evidenziato, successivamente, la presenza di numerose sequenze EST omologhe a quelle degli 11 geni PDI-like di riso e 70 di esse sono state amplificate mediante RT-PCR di mRNA estratti da cariossidi immature della cv CS e successivamente clonate. Sulla base delle analisi di sequenza i 70 cloni sono stati classificati in 10 gruppi, che corrispondevano alle dieci sequenze consenso (TC), precedentemente identificate nel database "TIGR wheat gene index". Non considerando la sequenza consenso (TC250792) che ha evidenziato un'omologia significativa con i membri appartenenti al primo gruppo filogenetico della famiglia delle PDI, le rimanenti 9, relative a geni PDI-like non ancora clonati da frumento, sono state utilizzate come base per isolare le sequenze cDNA complete (full-length) mediante la tecnica RACE, sia in direzione 5' che 3', a partire da mRNA estratti da cariossidi in via di maturazione (8 GDA). Dopo separazione elettroforetica in gel d'agarosio, i diversi prodotti RACE sono stati purificati, clonati e caratterizzati determinandone la sequenza nucleotidica. Le sequenze delle estremità 5' e 3' ottenute mediante RACE sono state assemblate con le relative sequenze consenso (TC) ed i trascritti completi sono stati amplificati, utilizzando primer specifici che riconoscevano sequenze localizzate nelle regioni non tradotte alle estremità 5' e 3', mediante RT-PCR da mRNA

estratti da cariossidi immature (8 GDA). L'analisi elettroforetica dei prodotti RT-PCR ha mostrato la presenza di singoli prodotti di amplificazione delle dimensioni attese per ognuna delle nove sequenze cDNA complete, che sono stati purificati da gel e clonati. Per ciascuno dei nove cloni i cinque plasmidi ricombinanti analizzati avevano la stessa sequenza, le cui caratteristiche sono riportate in tabella 5.1.

Tab. 5.1 – Caratteristiche delle sequenze cDNA complete codificanti per proteine PDI-like di frumento. Nella tabella sono inoltre riportate le corrispondenti sequenze TC identificate in banca dati ed utilizzate per isolare mediante RACE le sequenze cDNA complete e l'omologia di sequenza con le proteine codificate dai relativi geni ortologhi di riso.

Clone	**Lunghezza della sequenza dopo RACE**				**Sequenza TC corrispondente (TaGI gene Index)**	**Gene ortologo di riso**	**Identità proteica**
	UTR		**ORF**	**PRO**			
	5' (nt)	**3' (nt)**	**(nt)**	**(aa)**			
TaPDIL2-1	37	118	1740	580	TC268016	OsPDIL1-4 (AK071514)	416/580 (71,7%)
TaPDIL3-1	138	236	1632	544	TC272054	OsPDIL1-5 (AK073970)	446/544 (82,0%)
TaPDIL4-1	56	120	1101	367	TC264921	OsPDIL2-1 (NM_185280)	316/367 (86,1%)
TaPDIL5-1	55	153	1320	440	TC251262	OsPDIL2-3 (AK062254)	393/440 (89,3%)
TaPDIL6-1	13	259	447	149	TC267704	OsPDIL5-1 (AK063663)	118/149 (79,2%)
TaPDIL6-2	13	36	486	162	TC272338	OsPDIL5-1 (AK063663)	121/149 (81,2%)
TaPDIL7-1	37	295	1242	414	TC236080	OsPDIL5-2 (AK069367)	363/414 (87,7%)
TaPDIL7-2	136	189	1254	418	TC252818	OsPDIL5-2 (AK069367)	387/412 (69,3%)
TaPDIL8-1	137	230	1455	485	TC235895	OsPDIL5-4 (AK099660)	446/485 (91,9%)

5.2. Analisi filogenetica

Allo scopo di determinare le relazioni evolutive tra le sequenze cDNA clonate da frumento e quelle di altri geni codificanti per proteine PDI-like isolate da diverse specie vegetali, è stata condotta un'analisi filogenetica basata sul confronto delle sequenze aminoacidiche dedotte di 52 geni codificanti per proteine PDI-like (Fig. 5.1). Per la

costruzione dell'albero filogenetico sono state utilizzate 13 sequenze isolate da *Arabidopsis*, 11 da riso e 12 da mais, rappresentative delle 8 classi filogenetiche della famiglia delle PDI identificate nelle piante (Houston et al., 2005; Tab. 4.3 nella sezione 4.9. dei Materiali e Metodi), 2 dal muschio *P. patens* e una sequenza ciascuno da erba medica, ricino, orzo ed alga unicellulare *C. reinhardtii*, insieme con le sequenze proteiche dedotte da 10 sequenze cDNA di frumento, nove delle quali isolate durante lo svolgimento della tesi e la decima relativa alla PDI classica (I gruppo filogenetico), localizzata nel cromosoma 4A e precedentemente clonata e caratterizzata (Ciaffi et al., 2006). L'analisi filogenetica ha dimostrato, come atteso, che le sequenze isolate in frumento possono essere chiaramente assegnate a sette delle otto sottofamiglie geniche identificate nelle piante: *TaPDIL**2**-1* (II sottofamiglia), *TaPDIL**3**-1* (III sottofamiglia), *TaPDIL**4**-1* (IV sottofamiglia), *TaPDIL**5**-1* (V sottofamiglia), *TaPDIL**6**-1* e *TaPDL**6**-2* (VI sottofamiglia), *TaPDIL**7**-1* e *TaPDIL**7**-2* (VII sottofamiglia) e *TaPDIL**8**-1* (VIII sottofamiglia). E' importante tener conto, nelle considerazioni che saranno svolte in seguito sull'evoluzione della famiglia genica delle PDI nelle piante, che, data la natura poliploide del frumento tenero, le sequenze isolate ed assegnate a ciascuno dei sette gruppi filogenetici corrispondono, probabilmente, ad uno dei tre geni omeologhi presenti nel suo genoma esaploide (AABBDD), come verrà dimostrato per le sequenze appartenenti al IV e V gruppo filogenetico e come è stato già documentato per il primo gruppo filogenetico, per il quale in frumento sono state identificate tre sequenze geniche localizzate sui cromosomi del gruppo di omeologia 4 (Ciaffi et al., 2006). E' importante sottolineare, inoltre, che il confronto delle tre sequenze omeologhe ha evidenziato un elevato livello di conservazione, sia nella regione codificante che nella struttura genomica complessiva, incluse le regioni promotrici. In particolare il confronto delle sequenze nucleotidiche delle regioni codificanti dei tre geni ha evidenziato un'identità compresa tra 96 e 97,5%. Solo 14 delle 66 singole sostituzioni nucleotidiche rilevate tra le regioni codificanti dei tre geni omeologhi determinano sostituzioni aminoacidiche e solo sei di esse potrebbero modificare le caratteristiche funzionali delle proteine mature. Quest'ultima considerazione è importante per evidenziare che l'entità delle possibili differenze che si potrebbero riscontrare fra i tre geni omeologhi, per ciascuna delle nove sequenze PDI-like isolate in frumento, potrebbe non essere tale da giustificare una loro diversificazione strutturale e funzionale.

L'analisi filogenetica condotta conferma che gli otto gruppi filogenetici si siano originati prima della divergenza delle dicotiledoni dalle monocotiledoni, infatti in ciascuno di essi è possibile riscontrare almeno una sequenza isolata da *Arabidopsis*, mais, riso e frumento. La presenza di più di una sequenza, nell'ambito di ciascuno degli otto gruppi filogenetici, per

una stessa specie vegetale (geni paraloghi) indica, inoltre, fenomeni di duplicazione genica che sono avvenuti, in maniera indipendente, non solo dopo la diversificazione delle dicotiledoni dalle monocotiledoni, ma anche, nell'ambito delle monocotiledoni, dopo la separazione delle tre linee evolutive che hanno determinato la speciazione di mais, riso e frumento. Questi fenomeni indipendenti di duplicazione genica, all'interno di ogni gruppo filogenetico, hanno agito diversamente nelle specie vegetali, essi sono avvenuti in tempi diversi e/o hanno subito, probabilmente, processi distinti di diversificazione nell'ambito di una particolare specie. Al primo gruppo filogenetico, ad esempio, sono state assegnate tre sequenze di riso, che derivano, probabilmente, da due eventi di duplicazione genica, due di *Arabidopsis* e due di mais, come risultato di un unico evento di duplicazione avvenuto in ciascuna di queste specie, ed una da ciascuna delle seguenti specie: ricino, medica, frumento e orzo. A differenza delle due sequenze identificate in *Arabidopsis* (*AtPDIL1-1* e *AtPDIL1-2*) e mais (*ZmPDIL1-1* e *ZmPDIL1-2*), che formano due coppie di geni paraloghi molto simili, le due sequenze di riso *OsPDL1-2* e *OsPDL1-3* sembrano abbastanza distinte dall'altra sequenza isolata in questa specie (*OsPDIL1-1*), indicando che il primo evento di duplicazione genica verificatosi in riso debba essere avvenuto molto prima di quello che ha interessato *Arabidopsis* e mais e/o che sia stato seguito da un processo di diversificazione più rapido. Il secondo gruppo filogenetico comprende due coppie di geni paraloghi molto simili, identificati in *Arabidopsis* e mais, e sequenze singole isolate in riso e frumento. Sequenze appartenenti a questa classe filogenetica sono state individuate anche nell'alga unicellulare *C. reinhardtii* (CrRB60) e nel muschio *P. patens* (PpPDIH e PpPDIM). Il quarto gruppo filogenetico contiene due coppie di geni paraloghi identificati in mais e riso, mentre nell'ambito del terzo, quinto ed ottavo gruppo l'evento di duplicazione genica è stato rilevato solo in *Arabidopsis*. E' infine importante sottolineare che per il frumento solo nel sesto e settimo gruppo filogenetico sono state identificate due distinte sequenze derivanti, in entrambi i casi, da un singolo evento di duplicazione genica indipendente dalla sua natura poliploide. Nell'ambito del VII gruppo, a differenza delle due sequenze isolate da mais (*ZmPDIL7-1* e *ZmPDIL7-2*) che formano una coppia di geni paraloghi molto simile, le due sequenze di frumento *TaPDIL7-1* e *TaPDIL7-2* sono abbastanza diversificate l'una dall'altra, suggerendo che le due coppie di geni paraloghi possano essere state sottoposte ad una dinamica evolutiva significativamente differente nelle due specie.

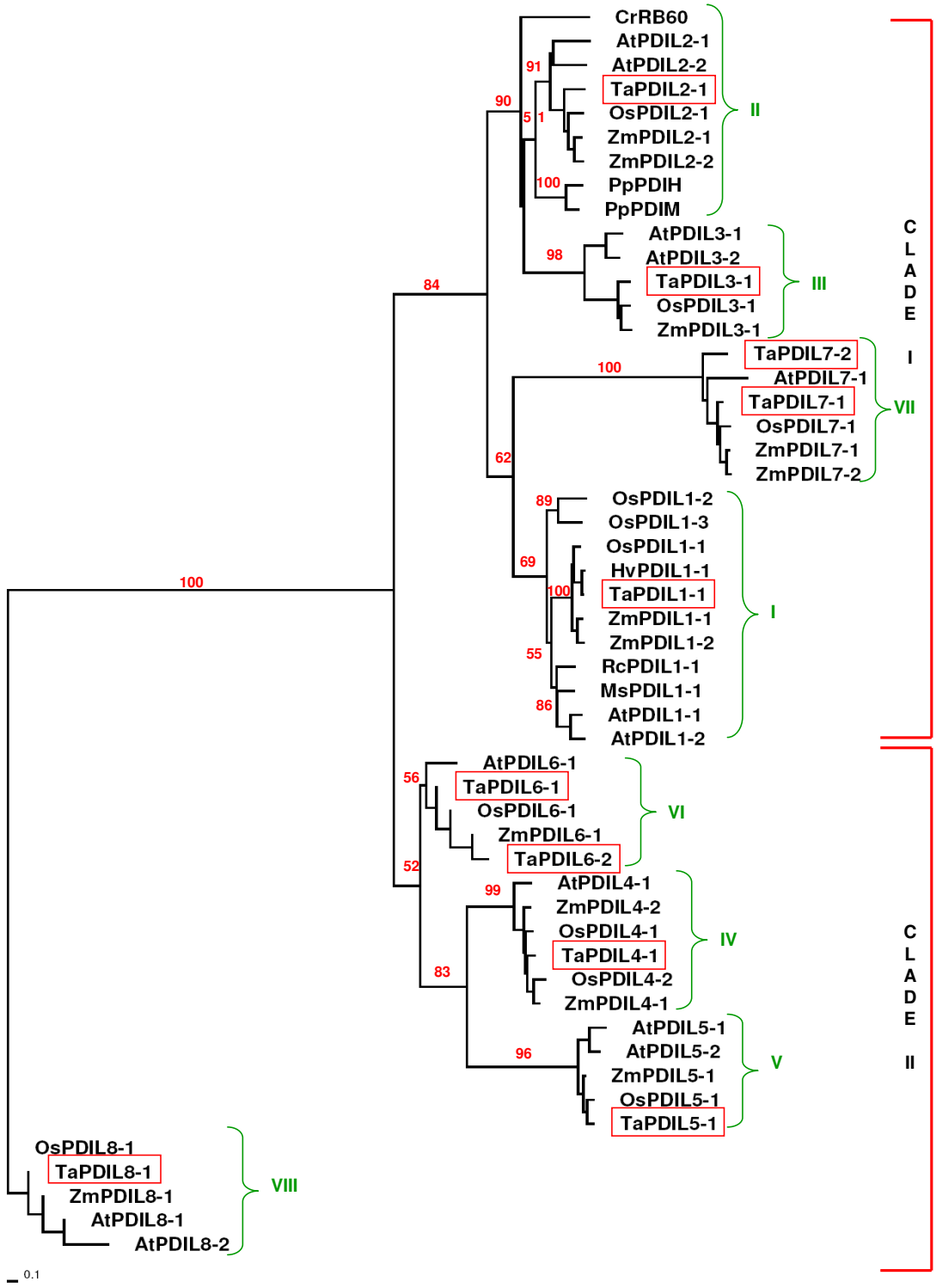

Fig. 5.1 - Albero filogenetico ottenuto utilizzando 52 sequenze aminoacidiche dedotte di geni PDI-like isolate da diverse specie vegetali: 13 da Arabidopsis, 11 da riso, 12 da mais, 2 dal muschio *P. patens*, 4 rispettivamente da erba medica, ricino, orzo ed alga unicellulare *C. reinhardtii* e 10 da frumento (evidenziate da un rettangolo rosso), nove isolate durante lo svolgimento della tesi. L'allineamento delle sequenze proteiche è stato ottenuto mediante il programma ClustalX versione 1.83 e l'albero filogenetico è stato costruito tramite il metodo neighbor-joining (NJ) e valutato mediante analisi bootstrap (PHYLIP versione 3.6). I numeri riportati sui nodi (ramificazioni) principali indicano i valori percentuali di boostrap per 1000 repliche. Le due cladi maggiori (I e II) e gli otto gruppi filogenetici (I-VIII) identificati nella famiglia delle PDI delle piante sono indicati rispettivamente da parentesi quadre in rosso e da parentesi graffe in verde.

Nel VI gruppo funzionale l'evento di duplicazione genica ha interessato solo il frumento, in quanto per tutte le altre specie considerate (*Arabidopsis*, riso e mais) è stata identificata un'unica sequenza.

A parte l'ottavo gruppo filogenetico, i cui membri rappresentano le proteine più diversificate nell'ambito della famiglia delle PDI delle piante, tanto che nella costruzione dell'albero sono stati considerati come sequenze outgroup, l'analisi filogenetica condotta indica chiaramente che la famiglia delle PDI può essere suddivisa in due cladi maggiori (Fig. 5.1). La prima clade è costituita dai gruppi I, II e III che posseggono proteine della classe strutturale 1 (proteine con due domini tioredossinici attivi, il primo localizzato nella regione N-terminale e il secondo in quella C-terminale), e dal gruppo VII, caratterizzato dalla presenza di proteine della classe strutturale 5 (proteine con un singolo dominio tioredossinico attivo). La seconda clade è composta dai gruppi IV e V, con proteine della classe strutturale 2 (proteine con due domini tioredossinici attivi adiacenti e disposti entrambi nella regione N-terminale), e dal gruppo VI, con proteine della classe strutturale 5. Questi risultati, ottenuti confrontando le sequenze aminoacidiche dedotte delle proteine PDI-like intere isolate nelle piante e non quelle derivanti dai singoli domini tioredossinici attivi come riportato in Houston et al., (2005), indicano che i primi tre gruppi filogenetici (I, II e III) appartenenti alla clade maggiore I potrebbero derivare da eventi di duplicazione genica che abbiano interessato un gene ancestrale, codificante per una proteina della classe strutturale 1, nel comune progenitore delle monocotiledoni e delle dicotiledoni e che la classe filogenetica VII si sia originata da un evento di duplicazione genica verificatosi prima della divergenza delle dicotiledoni dalle monocotiledoni in un gene ancestrale appartenente al I gruppo filogenetico, seguito dalla perdita di un dominio tioredossinico attivo. Più complesso risulta spiegare l'origine dei gruppi filogenetici appartenenti alla seconda clade maggiore. Indipendentemente dall'origine dei gruppi filogenetici IV e V, che è stato ipotizzato possano derivare da un evento di duplicazione genica che abbia interessato un gene ancestrale codificante per una proteina della classe strutturale 1, seguito dalla duplicazione del dominio tioredossinico N-terminale e successiva delezione di quello C-terminale (Kanai et al., 1998), i risultati ottenuti in questa ricerca stanno ad indicare che il gruppo filogenetico VI possa essersi originato in seguito ad un evento di duplicazione genica che ha interessato, prima della diversificazione delle monocotiledoni dalle dicotiledoni, un gene ancestrale codificante per una proteina della classe strutturale 2, seguita dalla delezione del dominio tioredossinico C-terminale. Poiché è stata riscontrata una maggiore omologia di sequenza tra i domini tioredossinici N-terminali del VI e IV gruppo filogenetico, che tra quelli del VI e V (Houston et al., 2005), la duplicazione

genica che ha dato origine alla VI classe filogenetica ha probabilmente interessato un gene ancestrale del IV gruppo presente nel comune progenitore delle monocotiledoni e dicotiledoni.

5.3. Struttura delle sequenze cDNA codificanti per proteine PDI-like in frumento

L'analisi delle sequenze aminoacidiche dedotte delle nove sequenze PDI-like isolate, condotta ricercando motivi conservati in diverse banche dati proteiche, ha evidenziato una notevole similitudine della loro struttura primaria con quella dei membri delle rispettive sottofamiglie.

La sequenza nucleotidica del clone *TaPDIL**2**-1* (II gruppo filogenetico) contiene un "Open Reading Frame" (ORF) di 1740 pb, con una capacità codificante di 580 aa (Tab. 5.1) ed un peso molecolare stimato per la proteina tradotta di 63,026 kDa. La proteina è ricca in aminoacidi acidi (9,60% D, acido aspartico, e 8,81% E, acido glutammico) con un valore stimato di pI pari a 4,42. La lunghezza di questa proteina (580 aa) è la maggiore tra quelle delle proteine PDI-like isolate in frumento e coincide, approssimativamente, con quella delle altre sequenze identificate in diverse specie vegetali (*Arabidopsis,* riso e mais) appartenenti al secondo gruppo filogenetico (561-597 aa) (Houston et al., 2005). La sequenza aminoacidica dedotta di *TaPDIL**2**-1* ha evidenziato una identità del 28,9% con quella della PDI classica di frumento (*TaPDIL**1**-1*, I gruppo filogenetico), pur presentando una struttura modulare ad essa molto simile, caratterizzata dalla presenza di due domini ***a*** e ***a'***, contenenti un sito tioredossinico attivo (CGHC), localizzati, rispettivamente, alle estremità N- e C-terminali e separati da due domini ***b*** e ***b'***, con struttura secondaria simile alla tioredossina ma privi del sito catalitico attivo (Fig. 5.2). *TaPDIL**2**-1* possiede, inoltre, un potenziale segnale peptidico ed il segnale KDEL, importanti, rispettivamente, per la traslocazione e la ritenzione delle proteine nel reticolo endoplasmatico (RE). La caratteristica distintiva di *TaPDIL**2**-1* rispetto a *TaPDIL**1**-1*, come di tutte le proteine appartenenti al secondo gruppo filogenetico rispetto a quelle del primo, è la presenza, nella regione N-terminale, di un dominio, lungo circa 60 aa, che contiene più del 40% di residui acidi (E+D). L'ERp72 e la PDI classica dei mammiferi contengono un dominio, localizzato rispettivamente alle estremità N- e C- terminale, ricco di residui prevalentemente acidi, caratteristica comune alle proteine che legano il calcio, come la maggior parte delle reticoloplasmine. Le proteine PDI-like appartenenti al secondo gruppo nelle piante potrebbero, pertanto, partecipare alla ritenzione del calcio all'interno della cellula

o interagire con particolari proteine o polipeptidi proprio mediante questo specifico dominio. L'analisi filogenetica (Fig. 5.1) ha evidenziato che le proteine delle piante appartenenti al II gruppo filogenetico risultano omologhe ad RB60, l'unica proteina PDI-like identificata nell'alga unicellulare *C. reinhardtii*, che è localizzata sia nel RE che nel cloroplasto (Levitan et al., 2005). Mentre la funzione della RB60 nel RE deve essere ancora chiarita, nel cloroplasto, come è stato precedentemente accennato, essa è coinvolta, in risposta alla luce ed a segnali fotosintetici, nella regolazione post-trascrizionale del gene *psbA*, il cui prodotto genico, la proteina D1, è uno dei componenti del fotosistema II. In particolare, RB60 è uno dei principali enzimi di un sistema di trasporto di elettroni che riduce una proteina, denominata cPABP, che nella sua forma ridotta fa parte di un complesso che ha la capacità di legarsi alla regione non trascritta a monte del codone di inizio AUG dell'mRNA di *psbA*, con la conseguente attivazione della traduzione del corrispondente prodotto D1 (Kim e Mayfield, 1997; Trebitsh et al., 2000). Nulla o poco si conosce sulla funzione svolta dalle proteine PDI-like delle piante superiori appartenenti al II gruppo filogenetico e, in particolare, se esse siano coinvolte nella regolazione dell'espressione genica nel cloroplasto, come la proteina omologa di *C. reinhardtii* RB60.

La sequenza nucleotidica del clone *TaPDIL**3**-1* (III gruppo filogenetico) possiede un ORF di 1632 pb codificante per una proteina di 544 aa (Tab. 5.1) con un peso molecolare stimato di 59,520 kDA e pI di 4,8. La sequenza proteica di *TaPDIL**3**-1* ha mostrato una identità di 21,1% e 29,0%, rispettivamente, con le sequenze aminoacidiche dedotte di *TaPDIL**1**-1* e *TaPDIL**2**-1*. La proteina codificata da *TaPDIL**3**-1*, che è caratterizzata dalla presenza di un potenziale segnale peptidico e del segnale KDEL, possiede una struttura multidominio di tipo ***c-a-b-b'-a'*** (Fig. 5.2), molto simile a quella codificata da *TaPDIL**2**-1*, ma con una ridotta dimensione (36 aa) della regione ricca in aminoacidi acidi localizzata nella regione N-terminale. La caratteristica distintiva più importante del prodotto di *TaPDIL**3**-1* e di tutte le proteine delle piante appartenenti al III gruppo filogenetico è la presenza di sequenze dei siti attivi tioredossinici nei domini ***a*** (CERS) e ***a'*** (CVDC) differenti da quella più comunemente riscontrata (CGHC) in proteine PDI-like non solo delle piante, ma anche di mammiferi e funghi, con probabili ripercussioni sull'attività ossido-riduttiva ed isomerasica di queste proteine. Sarebbe interessante, infatti, approfondire se queste sostituzioni aminoacidiche nel tetrapeptide del sito catalitico attivo tioredossinico mantengano l'ancestrale funzione o possano causare, come è probabile, una divergenza funzionale.

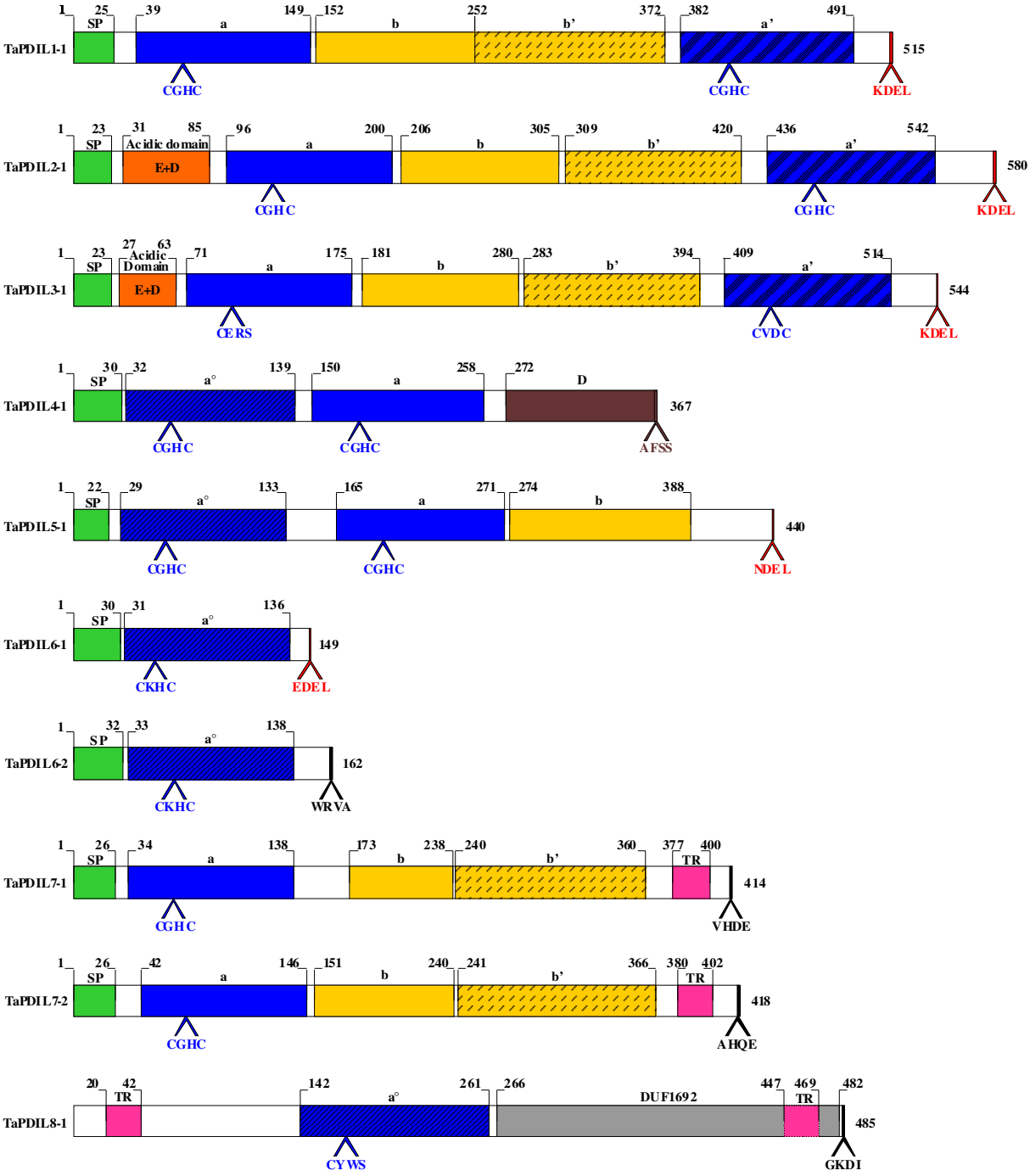

Fig. 5.2 – Struttura multidominio delle nuove proteine PDI-like identificate in frumento. In verde è evidenziato il segnale peptidico, in blu, differenziati da un diverso riempimento, i domini ***a***, ***a'*** e ***a***°. È riportato inoltre, sempre in blu, il relativo tetrapeptide presente al sito attivo. In giallo, differenziati dal riempimento, sono evidenziati i domini inattivi di tipo ***b*** e ***b'***. Con una colorazione arancione viene identificato il dominio acidico (E+D). In viola è indicato il dominio ***D***, mentre il dominio transmembrana (TR) è indicato in rosa. Nella *TaPDIL8-1* è inoltre presente un dominio DUF, indicato in grigio. Viene inoltre indicato, per ognuna delle sequenze, il tetrapeptide C-terminale. KDEL ed i suoi omologhi funzionali sono indicati in rosso.

La sequenza nucleotidica del clone *TaPDIL**4**-1* (IV gruppo filogenetico) ha un ORF di 1101 pb, con una capacità codificante di 367 aa ed un peso molecolare stimato per la proteina tradotta di 40,245 kDa e pI di 6,15. La proteina codificata da *TaPDIL**4**-1* ha evidenziato un'identità di 21,3%, 22,5% e 19,4%, rispettivamente, con le sequenze aminoacidiche dedotte di *TaPDIL**1**-1*, *TaPDIL**2**-1* e *TaPDIL**3**-1*, dalle quali differisce per la dimensione ridotta della sequenza proteica, per la disposizione e distanza dei due domini tioredossinici e per la presenza, all'estremità N-terminale, di un nuovo dominio, indicato come dominio ***D***, con una struttura secondaria completamente di tipo α-elica, la cui funzione non è ancora conosciuta (Fig. 5.2). E' importante sottolineare che, sebbene TaPDIL**4**-1, come tutte le proteine delle piante appartenenti al IV gruppo filogenetico, contenga un potenziale segnale di traslocazione nel reticolo endoplasmatico, in essa si riscontra l'assenza del segnale KDEL, importante per la ritenzione nel RE. Proteine con una struttura modulare ***a***°-***a***-***D***, definite proteine PDI-D, sono state identificate nel mixomicete *Dictyostelium discoideum* (Monnat et al., 1997) e nei funghi *Aspergillum niger* e *Neurospora crassa* (Jeenes et al., 1997), ma non nei mammiferi. Le proteine PDI-D delle piante differiscono da quelle isolate nei funghi per l'assenza di un segnale di ritenzione nel RE, tale assenza ha fatto ipotizzare che queste proteine potrebbero essere localizzate in un differente compartimento cellulare o potrebbero essere ritenute nel RE, come parte di un complesso eteromerico, attraverso l'interazione con subunità che contengano tale segnale. Recentemente è stato dimostrato che la parte C-terminale del dominio ***D*** è responsabile della ritenzione nel RE della PDI-D di *D. discoideum* (Monnat et al., 2000), rendendo plausibile che anche le proteine appartenenti al IV gruppo filogenetico delle piante, essendo tutte dotate di un dominio ***D***, siano localizzate nel RE (Lemaire e Miginiac-Maslow, 2004).

La sequenza proteica di *TaPDIL**5**-1* (V gruppo filogenetico), con valori stimati di peso molecolare e pI pari rispettivamente a 47,140 kDa e 5,06, ha evidenziato una identità di 17,4%, 18,8%, 17,2% e 29,8%, rispettivamente, con le sequenze aminoacidiche dedotte di *TaPDIL**1**-1*, *TaPDIL**2**-1*, *TaPDIL**3**-1* e *TaPDIL**4**-1*. La dimensione della proteina, che contiene un potenziale segnale peptidico e la sequenza C-terminale NDEL, è di 440 aa ed i due domini tioredossinici contenenti un sito catalitico attivo (CGHC) sono molto ravvicinati, essendo separati da 32 aa, mentre nella parte C-terminale è presente un unico dominio ***b*** di 114 aa (Fig. 5.2 e Tab. 5.1). I membri appartenenti al V gruppo filogenetico delle piante risultano strettamente correlati alle proteine P5 dei mammiferi in termini di identità di sequenza (43%), struttura modulare (***a***°-***a***-***b***), allineamento dei due potenziali domini tioredossinici e lunghezza delle proteine (439-443 aa).

Le sequenze nucleotidiche dei cloni *TaPDIL**6**-1* e *TaPDIL**6**-2* (VI gruppo filogenetico) contengono ORF, rispettivamente, di 447 e 486 pb, con una capacità codificante di 149 e 162 aa ed un peso molecolare stimato per le due proteine tradotte, rispettivamente, di 16,626 e 18,163 kDa e punti isoelettrici (pI) rispettivamente di 5,13 e 5,17 (Tab. 5.1). Le proteine codificate da *TaPDIL**6**-1* e *TaPDIL**6**-2*, che hanno un'identità di 81%, evidenziano in media un'identità di 27,8%, 19,5%, 16,9%, 33,4% e 26%, rispettivamente, con le proteine di frumento assegnate al I, II, III, IV e V gruppo filogenetico della famiglia delle PDI delle piante. La maggiore omologia riscontrata tra queste proteine e *TaPDIL**4**-1* ed in generale con le proteine delle piante appartenenti al IV gruppo filogenetico, rispetto a quelle assegnate agli altri nove gruppi, supporta l'ipotesi che i geni codificanti per proteine del VI gruppo derivino da un evento di duplicazione di un gene ancestrale del IV gruppo filogenetico, seguito dalla perdita della porzione del gene codificante la regione a valle del primo dominio tioredossinico (***a***^{0}). La lunghezza di queste due proteine è la minore tra quelle delle proteine PDI-like isolate in frumento e coincide, approssimativamente, con quella delle altre sequenze identificate in diverse specie vegetali (*Arabidopsis,* riso e mais) appartenenti al VI gruppo filogenetico (146-150 aa) (Houston et al., 2005). La struttura modulare di queste proteine è molto semplice, essendo caratterizzate dalla presenza di un unico dominio tioredossinico contenente una sequenza del sito attivo (CKHC) non comune rispetto a quella prevalentemente presente nelle proteine PDI-like delle piante e di altri organismi (CGHC). Sebbene queste proteine, come tutte quelle appartenenti al VI gruppo filogenetico, contengano un potenziale segnale di traslocazione (segnale peptidico) nel RE, esse mancano di un classico segnale di ritenzione nel RE, indicando che esse potrebbero essere localizzate in un altro compartimento cellulare o essere trattenute nel RE attraverso un meccanismo ancora sconosciuto. Recentemente è stato riportato che la proteina codificata da *ZmPDIL5-1*, l'unico membro di mais del VI gruppo filogenetico, viene particolarmente sovra-espressa, come la PDI classica (ZmPDIL1), la BiP e la calnexina, in mutanti di mais quali *floury-2* (*fl2*), *mucronate* (*Mu*) e *defective endosperm B30* (*de*-B30*), che manifestano condizioni di stress del RE a causa dell'espressione di proteine di riserva non correttamente processate (α- e γ-zeine), ma, a differenza degli altri chaperoni molecolari analizzati, non risulta localizzata nelle frazioni intracellulari di membrana corrispondenti alla porzione luminale del RE o a quella dei corpi proteici (Houston et al., 2005). Ulteriori studi saranno necessari per comprendere se le PDI-like appartenenti al VI gruppo filogenetico svolgano qualche funzione specifica all'interno del RE o se vi risiedono solo in maniera transiente, per essere poi traslocate in altri compartimenti cellulari.

Le sequenze nucleotidiche dei cloni *TaPDIL7-1* e *TaPDIL7-2* (VII gruppo filogenetico) contengono ORF, rispettivamente, di 1242 e 1254 pb, con una capacità codificante di 414 e 418 aa ed un peso molecolare stimato per le due proteine tradotte di 46,286 e 46,488 kDa e punti isoelettrici (pI) di 4,75 e 5,07 (Tab. 5.1). Le proteine codificate da *TaPDIL7-1* e *TaPDIL7-2*, che hanno un'identità di 72,6%, evidenziano in media un'identità di 24,05%, 22,15%, 19,95%, 22,95%, 22,3% e 28,62%, rispettivamente, con le proteine di frumento assegnate al I, II, III, IV, V e VI gruppo filogenetico della famiglia delle PDI delle piante. Le sequenze aminoacidiche dedotte di *TaPDIL7-1* e *TaPDIL7-2* presentano una struttura modulare, caratterizzata dalla presenza di un dominio ***a***, contenente un sito tioredossinico attivo (CGHC) localizzato alla estremità N-terminale, seguito da due domini adiacenti ***b*** e ***b'***, con struttura secondaria simile alla tioredossina, ma privi del sito catalitico attivo (Fig. 5.2). La struttura multidominio di tipo ***a-b-b'*** delle proteine di *TaPDIL7-1* e *TaPDIL7-2*, peraltro riscontrata nelle proteine codificate dai loro ortologhi isolati in *Arabidopsis,* riso e mais, supporta l'ipotesi che la classe filogenetica VII si sia originata in seguito ad un evento di duplicazione genica verificatosi prima della divergenza delle dicotiledoni dalle monocotiledoni in un gene ancestrale appartenente al I gruppo filogenetico, i cui membri mostrano una struttura modulare ***a-b-b'-a'***, seguito dalla perdita del dominio tioredossinico attivo C-terminale (***a'***). Le proteine codificate da *TaPDIL7-1* e *TaPDIL7-2* contengono, inoltre, un segnale peptidico ed un segmento transmembrana di circa 20 aa localizzato nella regione C-terminale, che in assenza del segnale KDEL potrebbe agire da punto di ancoraggio nella membrana del RE, in modo che queste proteine siano trattenute in questo compartimento cellulare. Eps1, una proteina PDI-like di lievito con una struttura modulare molto simile a quella dei membri del VII gruppo filogenetico delle piante, contenente un dominio transmembrana localizzato all'estremità C-terminale, è localizzata nel RE, dove sembra svolgere un ruolo importante nei processi di degradazione proteica che avvengono in questo compartimento cellulare (Wang e Chang, 1999). Recentemente proteine con una struttura analoga sono state identificate nell'uomo, dove è stata confermata la loro localizzazione nel RE (Ellgaard e Ruddock, 2005), in drosofila e nel nematode *C. elegans*. Le proteine PDI con un dominio transmembrana in queste specie sembrano svolgere un ruolo specifico durante diverse fasi dello sviluppo (Clissold e Bicknell, 2003).

La sequenza nucleotidica del clone *TaPDIL8-1* (VIII gruppo filogenetico) ha un ORF di 1455 pb, con una capacità codificante di 485 aa (Tab. 5.1) ed un peso molecolare stimato per la proteina tradotta di 54,358 kDa e pI di 6,96. TaPDIL**8**-1 è risultata, tra le proteine PDI-like isolate in frumento, la più diversificata, sia a livello di struttura che di omologia di

sequenza, avendo evidenziato un'identità solo di 17,2%, 14,4%, 15,2%, 16,5%, 15,8%, 22,2%, 21,0%, 17,4% e 16%, rispettivamente, con le sequenze aminoacidiche dedotte di *TaPDIL**1**-1*, *TaPDIL**2**-1*, *TaPDIL**3**-1*, *TaPDIL**4**-1*, *TaPDIL**5**-1*, *TaPDIL**6**-1*, *TaPDIL**6**-2*, *TaPDIL**7**-1* e *TaPDIL**7**-2*. La sequenza proteica di questa proteina, nella quale si riscontra l'assenza di un potenziale segnale peptidico, è caratterizzata dalla presenza di un dominio transmembrana, localizzato all'estremità N-terminale, di un dominio tioredossinico, contenente una sequenza del sito attivo (CYWS) non comune rispetto a quella prevalentemente presente nelle proteine PDI-like delle piante e di altri organismi (CGHC), e di un dominio, denominato DUF1692, contenente un segmento transmembrana alla sua estremità C-terminale, riscontrato in differenti proteine con funzione sconosciuta, due delle quali parzialmente caratterizzate. La prima di queste proteine, PTX1, è un fattore di trascrizione contenente un omeodominio che è coinvolto nella regolazione della trascrizione dei geni che codificano per gli ormoni prodotti dalla ghiandola pituitaria (Lanctôt et al., 1999). La seconda, denominata ERGIC-32, è una proteina di topo coinvolta nel trasporto delle proteine dal RE al Golgi (Breuza et al., 2004). In particolare questa proteina è localizzata nel cosiddetto "ER-Golgi intermediate compartement" (ERGIC), un insieme di vescicole tubulari che collegano il RE con il cis-Golgi e che rappresenta la prima fase dei movimenti anterogrado/retrogrado della via di secrezione. Diverse evidenze indicano che questo compartimento non esplica esclusivamente una funzione di veicolazione delle proteine dal RE al Golgi, ma le altre funzioni presunte non sono state ancora completamente chiarite. L'ERGIC sembra essere coinvolto nel controllo di qualità delle proteine neosintetizzate, come indicato dalla presenza del recettore del KDEL (Breuza et al., 2004), la cui funzione è quella di riportare nel RE chaperoni molecolari fuoriusciti per errore da questo compartimento cellulare, e nella traslocazione verso il citoplasma delle proteine non correttamente ripiegate per la loro degradazione (Raposo et al., 1995). E' stato inoltre ipotizzato che in questo compartimento cellulare possa avvenire, o essere completata, la maturazione di alcune proteine sintetizzate nel RE prima di essere veicolate nel Golgi (Yamamoto et al., 2001). E' interessante notare che la proteina ERGIC-32, al pari di TaPDIL**8**-1 e di tutte le proteine PDI-like dell'VIII gruppo filogenetico delle piante, possiede, oltre al dominio DUF1692, due segmenti transmembrana localizzati alle estremità N- e C-terminali, indicando che la maggior parte della proteina è localizzata nel lume dell'ERGIC. Ulteriori studi saranno necessari per comprendere se le proteine PDI-like appartenenti all'VIII gruppo filogenetico delle piante siano localizzate in un compartimento simile all'ERGIC dei mammiferi e quale funzione vi possano svolgere.

5.4. Isolamento delle sequenze genomiche mediante PCR

Al fine di isolare le sequenze genomiche relative a tre dei nove cloni cDNA isolati (*TaPDIL**2**-1*, *TaPDIL**4**-1* e *TaPDIL**5**-1)*, il DNA estratto dalla cultivar di frumento tenero CS è stato amplificato con le stesse coppie di primer utilizzate per l'isolamento dei relativi cDNA totali: ***2**-1*(F1-R1), ***4**-1*(F1-R1) e ***5**-1*(F1-R1) (Tab. 4.1 nella sezione 4.6. dei MM). Le tre combinazioni di primer sono state utilizzate in modo da amplificare l'intera regione codificante dei tre geni PDI-like e brevi segmenti delle regioni fiancheggianti alle estremità 5' e 3' (Fig. 5.3, 5.4 e 5.5). L'analisi elettroforetica dei prodotti PCR ha mostrato la presenza di singoli e specifici prodotti di amplificazione, per le tre combinazioni di coppie di primer, di circa 4600 pb (***2**-1,* F1-R1), 3800 pb *(**4**-1,* F1-R1) e 5300 pb (***5**-1,* F1-R1) (dati non mostrati). I tre prodotti di amplificazione sono stati clonati e utilizzati per le successive amplificazioni di regioni interne con diverse combinazioni di primer (Tab. 4.2 nella sezione 4.6. dei Materiali e Metodi). I primer usati per le amplificazioni interne sono stati sintetizzati prevalentemente sulla base delle tre sequenze cDNA precedentemente isolate e, in alcuni casi, sulla base delle sequenze genomiche ottenute dal sequenziamento progressivo dei cloni genomici parziali.

Come si può vedere nelle figure 5.3, 5.4 e 5.5, i primer sono stati disegnati al fine di amplificare regioni della sequenza genomica che si sovrappongono e di dimensioni progressivamente più piccole, in modo da ottenere prodotti che potessero essere completamente sequenziati.

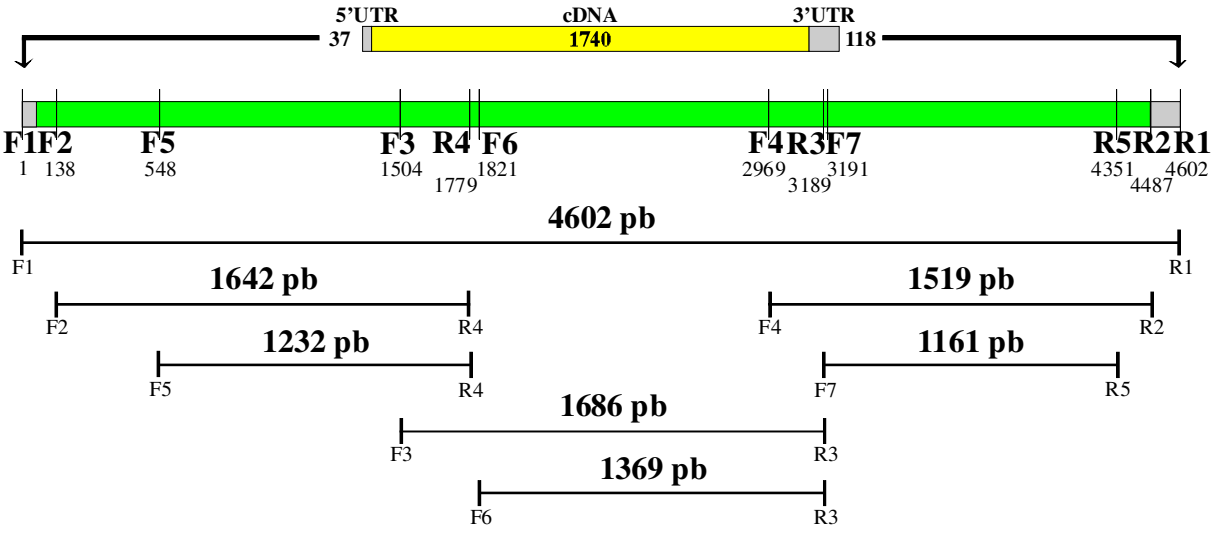

Fig. 5.3 – Rappresentazione schematica dei prodotti di amplificazione del DNA genomico ottenuti per il clone *TaPDIL**2**-1*. Sulla sequenza genomica complessiva, rappresentata in verde, sono indicate le posizioni relative dei diversi primer usati. In basso sono riportate le diverse combinazioni di primer usati e la dimensione e posizione dei prodotti di amplificazione. In alto, in giallo, è riportata la lunghezza della sequenza cDNA corrispondente, in grigio sono invece indicate le regioni non tradotte in 5' e 3'.

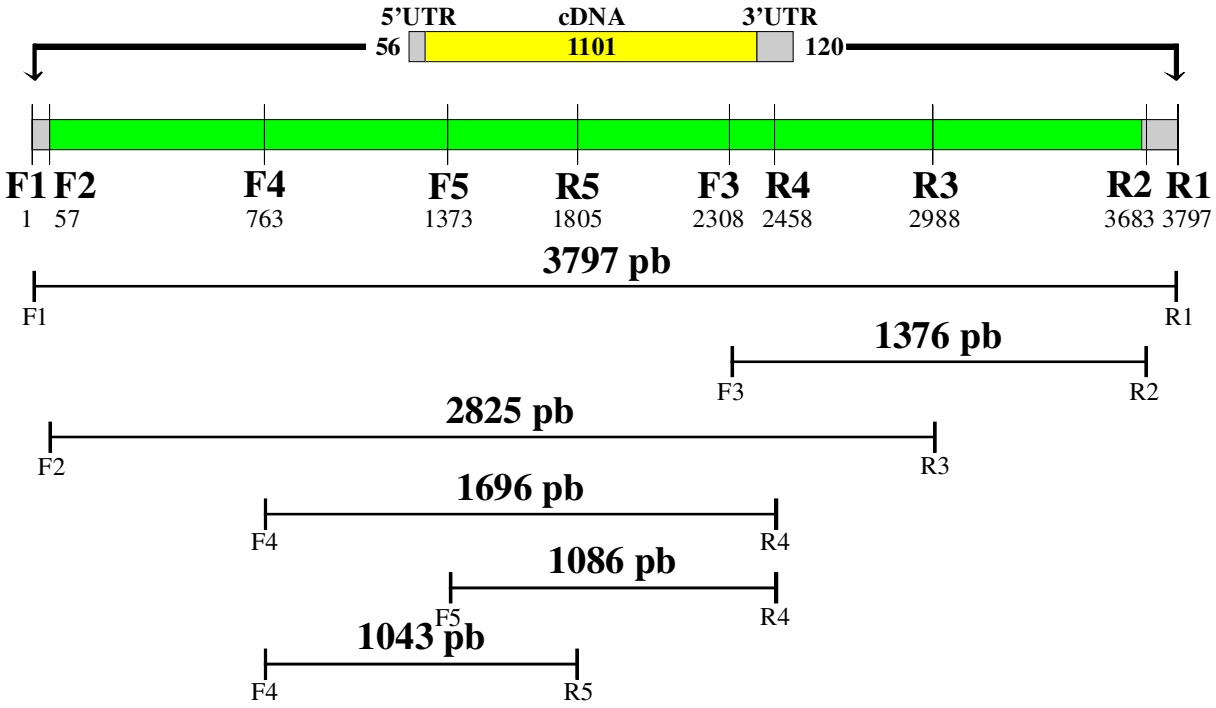

Fig. 5.4 – Rappresentazione schematica dei prodotti di amplificazione del DNA genomico ottenuti per il clone *TaPDIL4-1*. Sulla sequenza genomica complessiva, rappresentata in verde, sono indicate le posizioni relative dei diversi primer usati. In basso sono riportate le diverse combinazioni di primer usati e la dimensione e posizione dei prodotti di amplificazione. In alto, in giallo, è riportata la lunghezza della sequenza cDNA corrispondente, in grigio sono invece indicate le regioni non tradotte in 5' e 3'.

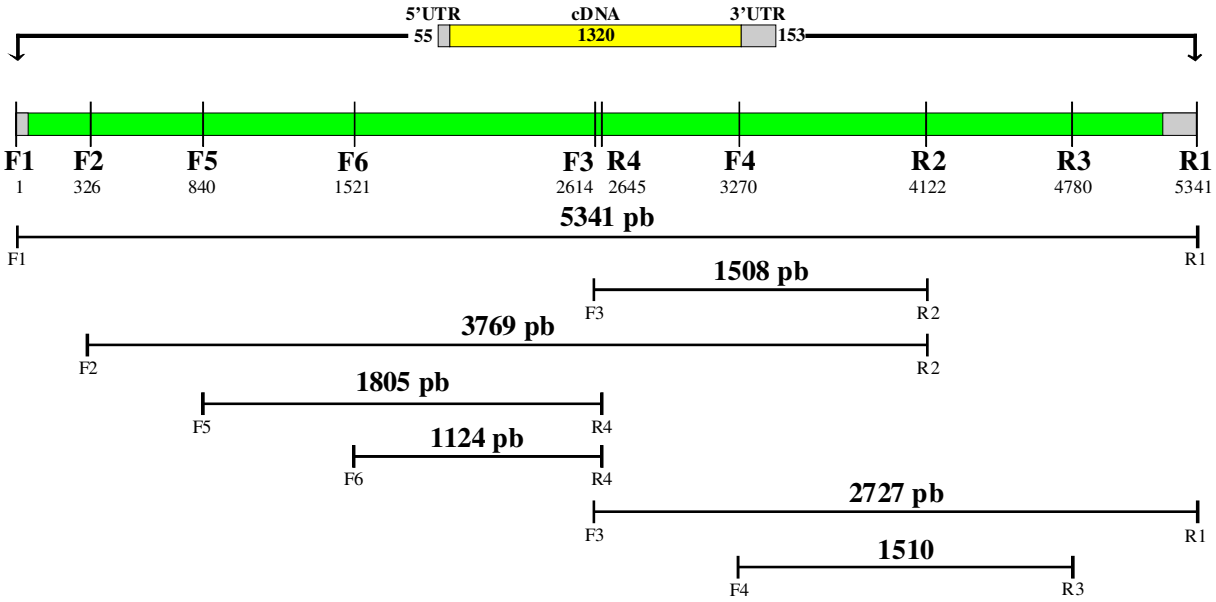

Fig. 5.5 – Rappresentazione schematica dei prodotti di amplificazione del DNA genomico ottenuti per il clone *TaPDIL5-1*. Sulla sequenza genomica complessiva, rappresentata in verde, sono indicate le posizioni relative dei diversi primer usati. In basso sono riportate le diverse combinazioni di primer usati e la dimensione e posizione dei prodotti di amplificazione. In alto, in giallo, è riportata la lunghezza della sequenza cDNA corrispondente, in grigio sono invece indicate le regioni non tradotte in 5' e 3'.

Nelle figure 5.3, 5.4 e 5.5 sono rappresentate in maniera schematica le posizioni dei diversi primer sulle relative sequenze genomiche e i prodotti di amplificazione ottenuti. Tutti i prodotti di amplificazione ottenuti dalle diverse coppie di primer sono risultati di dimensioni superiori a quelle attese in base alla lunghezza della relativa sequenza cDNA, confermando la presenza di introni all'interno dei geni codificanti per le tre proteine PDI-like. Ciascuno dei prodotti di amplificazione è stato purificato, clonato e sequenziato. Per ciascuno dei tre geni le sequenze parziali dei relativi cloni sono state allineate ed assemblate mediante il software DNAMAN, per ottenere le tre sequenze genomiche intere. Per la sequenza *TaPDIL**2**-1* la regione cDNA amplificata con la coppia di primer fiancheggianti l'intera regione codificante ***2**-1*(F1-R1) è lunga 1895 pb, mentre la sua sequenza genomica è lunga 4602 pb, ciò indica la presenza di 2707 pb di sequenze introniche. La sequenza cDNA *TaPDIL**4**-1*, amplificata con la coppia di primer ***4**-1*(F1-R1), è lunga 1277 pb, mentre la corrispondente sequenza genomica è lunga 3797 pb, le sequenze introniche coprono quindi 2520 pb. Infine le sequenze cDNA e genomica *TaPDIL**5**-1*, amplificate con la coppia di primer ***5**-1*(F1-R1), sono lunghe, rispettivamente, 1528 pb e 5341 pb, 3813 pb sono quindi relative agli introni. La presenza di una cosi alta percentuale di sequenze introniche, più del 50%, è stata riscontrata anche per *TaPDIL**1**-1* (Ciaffi et al., 2006). Per il gene omeologo localizzato sul cromosoma 4A, ad esempio, alla sequenza cDNA lunga 1545 pb corrisponde una sequenza genomica lunga 3361 pb.

5.5. Struttura dei geni PDI-like di frumento

Le sequenze genomiche relative ai tre geni *TaPDIL**2**-1*, *TaPDIL**4**-1* e *TaPDIL**5**-1*, lunghe, rispettivamente, 4602 pb, 3797 pb e 5341 pb, sono state allineate con le relative sequenze cDNA mediante il software DNAMAN per verificarne l'identità di sequenza e determinarne la struttura introni/esoni. L'analisi ha evidenziato una perfetta corrispondenza delle sequenze nucleotidiche nelle regioni relative agli esoni ed ha indicato che i tre geni PDI-like isolati in frumento presentano una struttura particolarmente complessa, caratterizzata dalla presenza di 12 (*TaPDIL**2**-1*), 11 (*TaPDIL**4**-1*) e 9 (*TaPDIL**5**-1*) esoni. Nelle figure 5.6, 5.7 e 5.8 sono rappresentate le strutture dei tre geni di frumento confrontate con quelle dei geni appartenenti al medesimo gruppo filogenetico isolati in riso e *Arabidopsis*. Sono state raffigurate le sequenze dal codone di inizio (ATG) a quello di terminazione (stop codon) della traduzione. La dimensione totale dell'intera sequenza genomica relativa al gene *TaPDIL**2**-1* (ATG-stop codon) è risultata di 4.450 pb. Gli 11 introni della sequenza *TaPDIL**2**-1* hanno evidenziato dimensioni molto variabili, con il più piccolo di 82 pb ed il più grande di 707 pb. Anche la dimensione degli esoni è risultata molto variabile, con il più piccolo lungo 57 pb (XI) ed il più grande lungo ben 569 pb (I). Questo primo esone inizia con una sequenza nucleotidica lunga 69 bp, che codifica per il segnale peptidico N-terminale di 23 aminoacidi, seguita, a poche paia di basi, da un sequenza lunga 165 pb, codificante per un segmento di 55 aa ricco in residui acidi, il dominio acidico. Nel primo esone sono comprese anche le 4 triplette codificanti per il primo dei due siti catalitici attivi (CGHC), quello relativo al dominio tioredossinico di tipo ***a***, presenti nella proteina. Per posizionare esattamente i codoni codificanti per il tetrapeptide -CGHC- è stata presa in considerazione l'ultima base della tripletta codificante per la glicina. Questa si trova esattamente a 375 pb dall'inizio e a 194 pb dalla fine del primo esone di *TaPDIL**2**-1*. L'ultima base della tripletta codificante per la glicina del secondo sito attivo, quello relativo al dominio tioredossina simile di tipo ***a'***, si trova invece a 21 pb dall'inizio del X esone. Come si può vedere in figura 5.6, escludendo gli esoni alle estremità 5' e 3'ed il terzo esone dell'ortologo di riso (*OsPDIL2-1*), lungo 169 pb invece che 166 pb, tutti gli altri esoni sono perfettamente conservati tra i geni ortologhi. La perfetta conservazione degli esoni è rispecchiata anche dalla conservazione della posizione del secondo sito attivo, che in tutti i geni si trova a 21 pb dall'inizio del X esone. I primi esoni dei diversi geni ortologhi, localizzati all'estremità 5', benché non perfettamente conservati, presentano tutti una dimensione relativamente simile, compresa tra 503 e 587 pb. La diversa

lunghezza è dovuta, probabilmente, a segnali peptidici e/o domini acidici di lunghezza variabile.

La posizione dei codoni codificanti per il tetrapeptide del primo sito attivo risulta parzialmente conservata tra i diversi ortologhi. Anche in riso, infatti, nonostante in questa specie le triplette codificanti per la sequenza consenso del tetrapeptide (CAHC) siano leggermente diverse, esse si trovano a 194 pb dalla fine del primo esone come in frumento. In entrambi i geni di *Arabidopsis AtPDIL2-1*, che ha una sequenza consenso del tetrapeptide di tipo –CGAC-, e *AtPDIL2-2*, che ha invece la sequenza conservata di tipo –CGHC-, i codoni codificanti per il sito attivo si trovano a 188 pb (Fig. 5.6) dalla fine del primo esone. Sostituzioni aminoacidiche a carico dei due aminoacidi centrali del tetrapeptide –CGHC- potrebbero influenzare la forza ossido-riduttiva del sito attivo. Tutte le quattro sequenze ortologhe contengono, inoltre, una sequenza codificante per il tetrapeptide C-terminale KDEL nell'ultimo esone (XII). Il segnale KDEL è importante per la ritenzione delle proteine residenti nel lume del reticolo endoplasmatico ed è presente anche nella PDI classica (Denecke et al., 1992), in posizione immediatamente adiacente alla tripletta di terminazione della traduzione. Il dominio tioredossinico di tipo ***a*** è codificato da una sequenza che, in frumento (Fig. 5.6), include parte del primo esone (284 pb terminali) e parte del secondo esone (31 pb iniziali). Il secondo dominio tioredossinico di tipo ***a'*** è codificato da una sequenza che comprende parte del IX esone (72 pb finali), X e XI esone per intero e parte del XII esone (57 pb iniziali). Il dominio inattivo di tipo ***b*** è codificato, invece, da una sequenza che comprende parte del II esone (84 pb finali) tutto il III esone e parte del IV esone (50 pb iniziali). Infine il dominio di tipo ***b'*** è codificato da una sequenza che comprende parte del IV esone (60 pb finali), tutto il V, VI e VII esone e parte dell'VIII esone (33 pb iniziali).

La sequenza genomica di *TaPDIL**4**-1*, la cui regione delimitata dall'ATG e dallo stop codon è lunga 3.624 pb, comprende 11 esoni, come i corrispondenti geni ortologhi di riso, mentre il gene ortologo di *Arabidopsis* ne possiede solo 10 (Fig. 5.7). Anche in questo caso gli introni hanno mostrato una dimensione molto variabile, con il più piccolo di 85 pb (III introne) ed il più grande, localizzato all'estremità 5' della sequenza (I introne), di 1142 pb. Come si può notare in figura 4.7, 8 esoni risultano perfettamente conservati tra le diverse specie. La lunghezza del primo esone di ciascuno dei geni ortologhi risulta leggermente variabile (152 pb-188 pb), mentre l'ultimo esone del gene di *Arabidopsis* è più lungo (175 pb) rispetto a quello degli ortologhi di riso e frumento e corrisponde, approssimativamente, alla somma dei due esoni terminali conservati negli altri geni, il X (100 pb) e l'XI (96 pb).

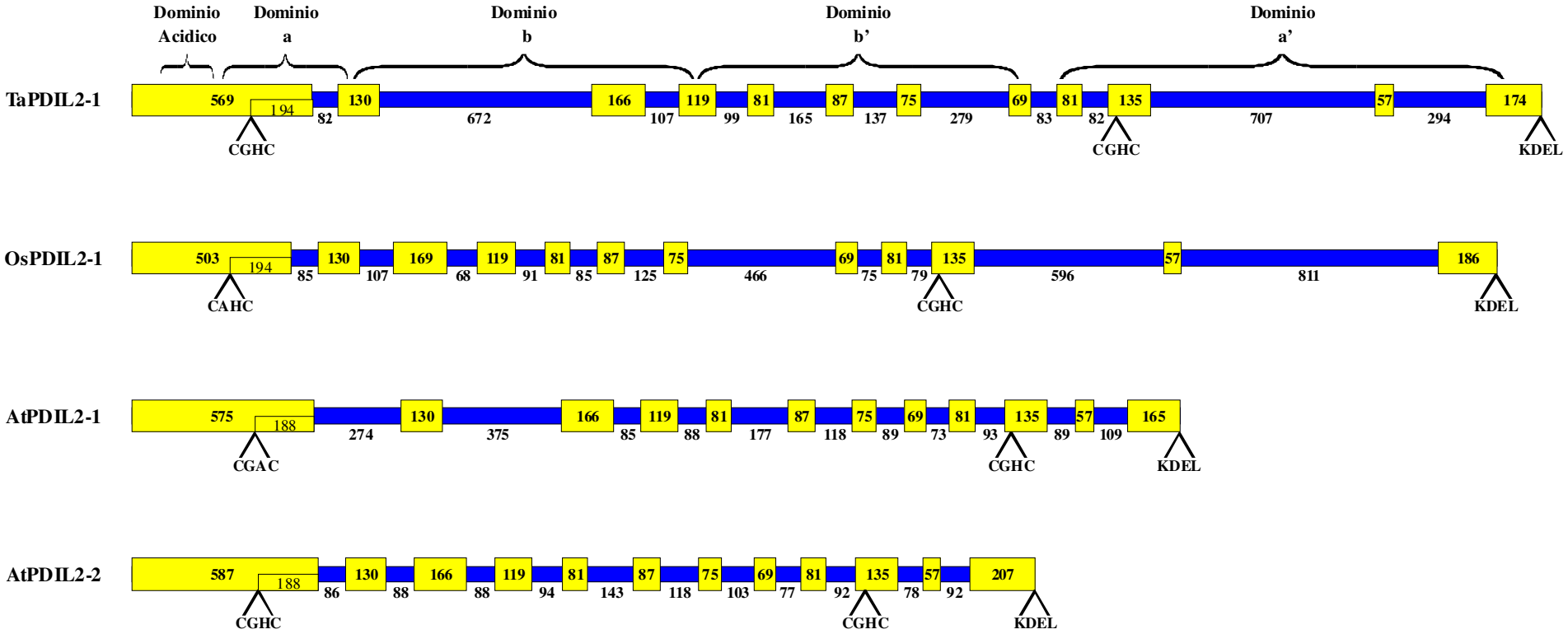

Fig. 5.6 - Rappresentazione schematica della struttura esoni/introni dei geni appartenenti al II gruppo filogenetico. In giallo sono rappresentati gli esoni, mentre gli introni sono rappresentati in blu. Sono inoltre indicate le sequenze dei tetrapeptidi presenti ai siti attivi e quelle del tetrapeptide C-terminale KDEL. Le parentesi graffe indicano i limiti dei diversi domini.

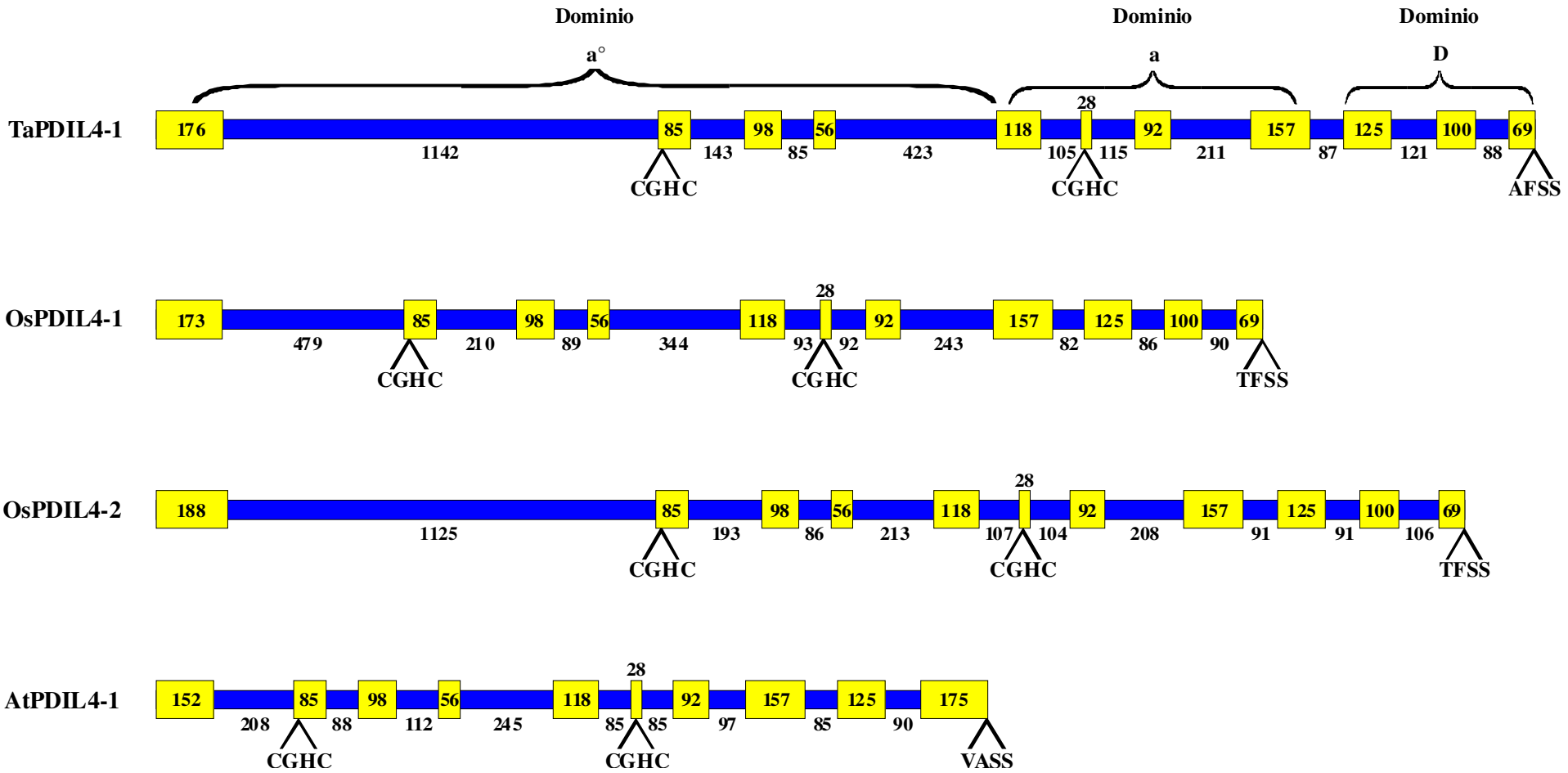

Fig. 5.7 - Rappresentazione schematica della struttura esoni/introni dei geni appartenenti al IV gruppo filogenetico. In giallo sono rappresentati gli esoni, mentre gli introni sono rappresentati in blu. Sono inoltre indicate le sequenze del tetrapeptide –CGHC-, presente ai siti attivi, e quella dei vari tetrapeptidi presenti all'estremità C-terminale delle diverse sequenze. Le parentesi graffe indicano i limiti dei diversi domini.

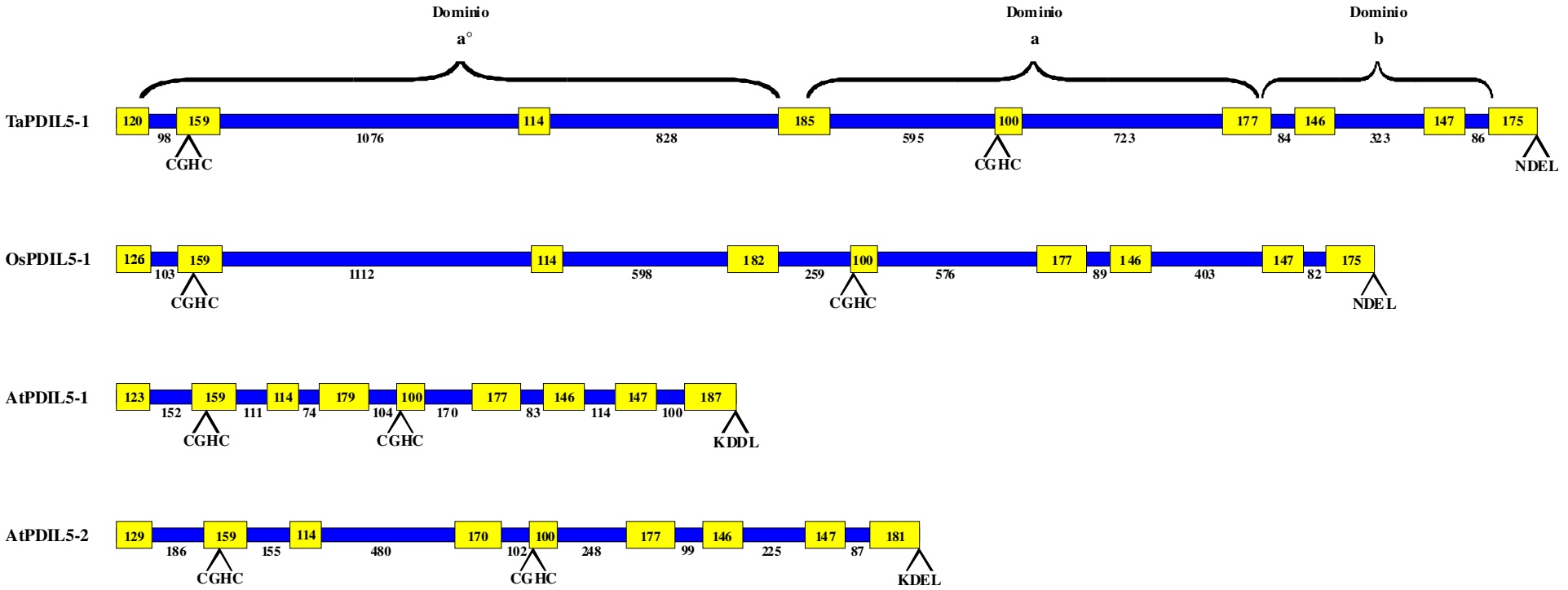

Fig. 5.8 - Rappresentazione schematica della struttura esoni/introni dei geni appartenenti al V gruppo filogenetico. In giallo sono rappresentati gli esoni, mentre gli introni sono rappresentati in blu. Sono inoltre indicati i tetrapeptidi presenti ai siti attivi e il tetrapeptide C-terminale KDEL o uno dei suoi derivati funzionali. Le parentesi graffe indicano i limiti dei diversi domini.

Ciò fa supporre che esso si sia formato per fusione dei due esoni, in seguito alla perdita del X introne. Perfettamente conservate tra i diversi geni ortologhi risultano anche le posizioni dei codoni relativi ai siti attivi tioredossinici, che si trovano dopo due triplette dall'inizio del II e VI esone. Le diverse sequenze dei geni ortologhi terminano con i tetrapeptidi AFSS in frumento, TFSS in riso e VASS in *Arabidopsis*, che non sono segnali di ritenzione nel RE. Anche altre proteine appartenenti al gruppo filogenetico 4, in particolare le proteine del tipo PDI-Dα di erba medica e di *Dictyostelium*, per le quali è nota la presenza nel reticolo endoplasmatico (Shorrosh e Dixon, 1992; Monnat et al., 2000), non presentano un segnale per la ritenzione in questo compartimento. Tutte queste proteine presentano, comunque, un segnale di secrezione N-terminale e, recentemente, è stato dimostrato che la parte C-terminale del dominio ***D*** è responsabile della ritenzione nel RE della proteina PDI-Dα in *Dictyostelium* (Monnat et al., 2000). Questo dominio ***D*** terminale è codificato, in frumento, da una sequenza di 288 pb che include la parte finale del IX esone (122 pb finali), tutto il X esone e parte dell'XI esone (66 pb iniziali). Il dominio tioredossinico di tipo ***a***° è codificato da una sequenza di 324 pb, che comprende parte del primo esone (83 pb finali), tutto il II, il III ed il IV esone ed una piccolissima parte del V esone (2 pb iniziali). Infine il dominio tioredossinico di tipo ***a*** è codificato da una sequenza di 327 pb che include parte del V esone (86 pb finali), tutto il VI ed il VII esone e parte dell'VIII esone (121 pb iniziali).

Come per i geni sopra descritti, anche per il gene *TaPDIL**5**-1* gli 8 introni risultano molto variabili, da 84 pb del VI introne a 1076 pb del II. Tutte i geni ortologhi presentano 9 esoni 6 dei quali risultano perfettamente conservati. Gli esoni variabili sono il primo, l'ultimo e il IV. Anche in questo caso le posizioni dei codoni che codificano per i siti attivi tioredossinici sono perfettamente conservate tra i diversi geni ortologhi. Nel primo di essi l'ultima base della tripletta codificante per la glicina si trova infatti a 54 pb dall'inizio del II esone, mentre nel secondo si trova a 7 pb dall'inizio del V esone. Tutti i geni presentano una sequenza codificante per il tetrapeptide C-terminale di ritenzione nel RE, NDEL in frumento e riso, KDEL e KDDL per i due geni di *Arabidopsis*. La sequenza codificante per il dominio ***a***° include parte del I esone (36 pb finali), il II ed il III ed una piccola parte del IV esone (6 pb iniziali), mentre quella per il dominio ***a*** comprende parte del IV esone (86 pb finali), tutto il V e parte del VI (135 pb iniziali). Infine la sequenza codificante per il dominio ***b*** include parte del VI (36 pb finali), gli esoni VII ed VIII e parte del IX (16 pb iniziali). La dimensione totale dell'intera sequenza genomica relativa al gene *TaPDIL**5**-1* (ATG-stop codon) è risultata di 5.136 pb.

Il confronto delle sequenze genomiche dei geni ortologhi isolati in *Arabidopsis*, riso e frumento nell'ambito dei gruppi filogenetici II, IV e V, come precedentemente rilevato anche per il primo gruppo filogenetico (Ciaffi et al., 2006), ha evidenziato in generale una discreta conservazione nell'organizzazione introni/esoni, anche se è stato rilevato che i geni PDI-like di riso e frumento tendenzialmente hanno una dimensione degli introni relativamente più grande rispetto a quelli di *Arabidopsis*, come risulta particolarmente evidente per il V gruppo filogenetico, a cui appartiene il gene *TaPDIL**5**-1* (Fig. 5.8). Le sequenze genomiche confrontate evidenziano, invece, una quasi perfetta conservazione delle dimensioni degli esoni. In particolare gli esoni in cui sono localizzati i codoni codificanti per i siti attivi tioredossinici sono risultati perfettamente conservati in tutti i casi, eccetto il primo esone del gruppo filogenetico II, in cui è probabilmente inclusa una sequenza di lunghezza variabile codificante per un dominio ricco in aminoacidi acidi. E' interessante notare che in tutti i geni appartenenti al IV e V gruppo filogenetico la sequenza codificante per il dominio di tipo ***a***° sembra avere i propri confini coincidenti con le giunzioni esoni/introni. Sia nella proteina *TaPDIL**4**-1* che nella *TaPDIL**5**-1*, infatti, questo dominio presenta, rispettivamente, solo 2 e 6 pb che cadono nel seguente esone. I domini di tipo ***a***, ***b*** e ***b'*** sembrano essere, invece, unità funzionali i cui limiti estremi coincidono, approssimativamente, con giunzioni esoni/introni, mentre se questi domini vengono presi a coppie tendono a terminare ed ad iniziare a cavallo di un esone intermedio. Questa potrebbe essere una reminescenza dell'organizzazione strutturale della proteina originale, in cui questi domini avevano questa organizzazione modulare. Questa condizione è infatti presente anche in *TaPDIL**1**-1*, la cui sequenza codificante per il dominio di tipo ***a*** include le 86 pb finali del primo esone, l'intero secondo esone e termina nel terzo esone, nel quale, a poche paia di basi di distanza, inizia la sequenza codificante per il dominio ***b***. Il dominio ***b***, a sua volta, termina nel IV esone esattamente dove inizia il dominio di tipo ***b'***, che a sua volta termina nel VI esone, poche basi prima dell'inizio del dominio di tipo ***a'***.

La netta separazione del dominio di tipo ***a***° dagli altri domini tioredossinici nelle proteine *TaPDIL**4**-1* e *TaPDIL**5**-1* potrebbe quindi essere il risultato del riarrangiamento di esoni che ha portato alla duplicazione del dominio ***a*** e, quindi, alla presenza dei due domini ***a*** e ***a***° adiacenti. Questo supporta l'ipotesi, emersa dai risultati ottenuti da Kanai et al. (1998), che il dominio tioredossinico N-terminale tenda a duplicarsi, mentre il dominio tioredossinico C-terminale tenderebbe a perdersi. Una considerazione simile può essere fatta anche per il dominio di tipo ***D*** presente nella *TaPDIL**4**-1*, anche in questo caso, infatti, il limite del dominio riflette perfettamente le giunzioni esoni/introni. Per il dominio acidico presente nella

*TaPDIL**2**-1* non si possono, invece trarre conclusioni di questo tipo in quanto, pur essendo probabile che il primo esone si sia originato dalla fusione di almeno due esoni distinti, il punto o i punti di fusione e, quindi, anche i limiti tra domini adiacenti, ossia tra il dominio acidico e quello di tipo ***a***, non sono ipotizzabili con i dati a disposizione. A differenza di quello che succede nella *TaPDIL**1**-1*, nella *TaPDIL**2**-1* il dominio di tipo ***a'*** risulta strutturalmente separato dai domini ***a***, ***b*** e ***b'***. Una situazione simile potrebbe essere stata quella presente nella proteina originaria dalla quale, in seguito a riarrangiamento degli esoni per crossing-over ineguale, si è verificata la perdita del dominio tioredossinico terminale ***a'*** (*TaPDIL**7**-1* e *-2*) o dell'introne intermedio VIII e la fusione degli esoni VIII e IX, con la conseguente perdita dell'indipendenza strutturale del dominio ***a'*** (*TaPDIL**1**-1*). Secondo queste deduzione, la forma più arcaica sarebbe quindi più simile alla *TaPDIL**2**-1*, nella quale viene mantenuta la netta separazione del dominio ***a'*** da quello ***b'***, mentre gli altri tre domini, ***a***, ***b*** e ***b'*** mantengono la stessa organizzazione modulare presente nella forma classica della PDI.

5.6. Localizzazione cromosomica delle sequenze dei geni *TaPDIL4-1* e *TaPDIL5-1*

Inizialmente il DNA genomico estratto dalla cultivar Chinese Spring (CS) è stato digerito con diversi enzimi di restrizione (*Eco*RI, *Bam*HI, *Sac*I, *Pst*I, *Hin*dIII, *Eco*RV, *Kpn*I), separato tramite elettroforesi, blottato su membrana ed ibridato con sonde specifiche per le sequenze *TaPDIL**4**-1* e *TaPDIL**5**-1*. In questo modo sono state individuate le combinazioni sonda ed enzima di restrizione che permettono una localizzazione cromosomica di questi due geni immediata ed univoca.

Per la *TaPDIL**4**-1*, in particolare, sono state provate due differenti tipi di sonda. La prima, S-F3-R5, è lunga 601 pb ed è complementare alla prima metà circa della regione codificante, mentre la seconda, S-F4-R3, è lunga 763 pb ed è complementare alla seconda metà della regione codificante ed a parte della sequenza non tradotta in 3'. Per entrambe le sonde, in combinazione con i differenti enzimi di restrizione provati, sono stati ottenuti gli stessi profili di ibridazione, con l'unica differenza che la sonda S-F4-R3 ha prodotto un segnale più intenso. Il livello di "background" per entrambe le sonde è risultato molto basso, quasi insignificante. Per la localizzazione cromosomica di *TaPDIL**4**-1* è stata quindi scelta la sonda S-F4-R3 in combinazione con l'enzima di restrizione *Sac*I. Il profilo di ibridazione

ottenuto tramite la digestione del DNA di CS con questo enzima è risultato, infatti, molto chiaro, in quanto costituito da tre sole bande, tutte della stessa intensità e ben separate tra loro.

Per la sequenza *TaPDIL**5**-1* è stata provata una sola sonda, la S-F3-R4, lunga 918 pb e complementare a circa la metà in 3' della regione codificante ed a parte della sequenza non tradotta in 3'. Questa sonda è risultata molto specifica, con un livello di background quasi assente, ed in combinazione con l'enzima di restrizione *Eco*RV ha prodotto un profilo di ibridazione consistente di 3 bande ben definite.

Una volta individuata la combinazione ottimale tra sonda ed enzima di restrizione per entrambi i geni è stata eseguita la loro localizzazione cromosomica usando il DNA estratto dalla cultivar Chinese Spring e dalle relative linee nulli-tetrasomiche. La sequenza *TaPDIL**4**-1* è stata localizzata sui cromosomi del gruppo di omeologia 1, mentre la sequenza *TaPDIL**5**-1* è stata localizzata su quelli del gruppo 5.

Poiché solo le linee nulli-tetrasomiche del gruppo di omeologia 1, per la sequenza *TaPDIL**4**-1*, e quelle del gruppo di omeologia 5, per la sequenza *TaPDIL**5**-1*, differivano da CS nel profilo di ibridazione, ulteriori analisi sono state effettuate utilizzando solo le linee ditelosomiche di questi due gruppi di omeologia. In questo modo è stato possibile localizzare le sequenze geniche codificanti per TaPDIL**4**-1 sui bracci corti dei tre cromosomi del gruppo di omeologia 1 (1AS, 1BS e 1DS), mentre quelle codificanti per TaPDIL**5**-1 sono state localizzate sui bracci lunghi dei tre cromosomi del gruppo di omeologia 5 (5AL, 5BL e 5DL) (Fig. 5.9 e 5.10). In particolare la sonda S-F4-R3 ibrida con tre distinti frammenti prodotti dalla restrizione del DNA totale di CS con l'enzima *Sac*I, la cui dimensione stimata è 12,7, 10,8 e 8,2 kb, e che corrispondono a sequenze geniche localizzate, rispettivamente, sul braccio corto del cromosoma 1D (profilo n°11 in Fig. 5.9), sul braccio corto del cromosoma 1B (profilo n°9) e sul braccio corto del cromosoma 1A (profilo n°7). Per le sequenze geniche codificanti per la TaPDIL**5**-1 (Fig. 5.10), invece, l'analisi Southern con la sonda S-F3-R4, in combinazione con l'enzima di restrizione *Eco*RV, ha indicato la loro localizzazione sui bracci lunghi dei cromosomi appartenenti al gruppo di omeologia 5 e ha prodotto tre frammenti di restrizione con una dimensione stimata superiore a 21 kb. Il frammento di restrizione di maggiori dimensioni è localizzato sul braccio cromosomico 5BL (profilo n°8 in Fig. 5.10), quello intermedio sul 5DL (profilo n°10) e quello minore sul 5AL (profilo n°6). In Fig. 5.9 e 5.10 appare chiaro che i profili di ibridazione ottenuti per entrambi i geni sono ben definiti e specifici, con segnali ben identificabili per tre frammenti di restrizione per gene e assenza di background.

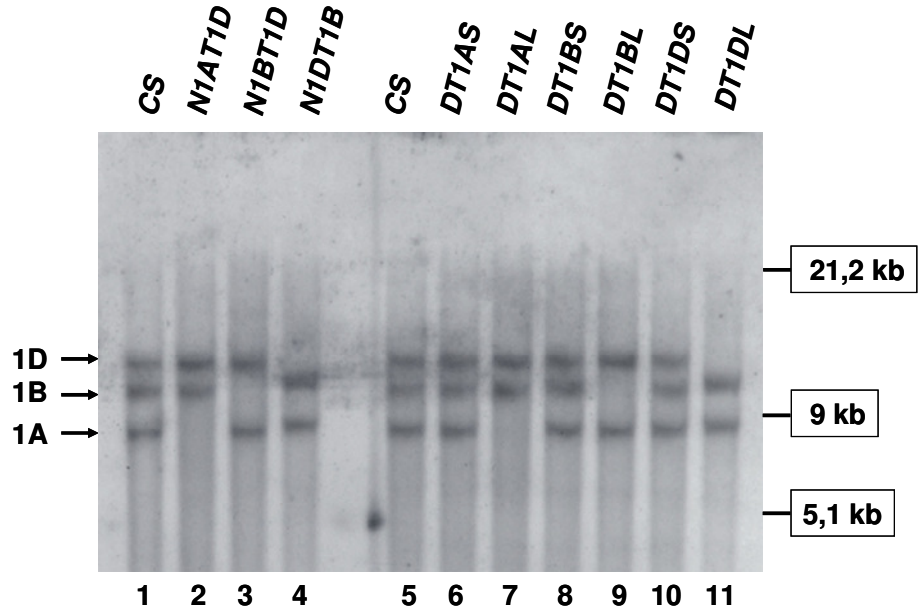

Fig. 5.9 - Analisi Southern del DNA genomico, digerito con *Sac*I, di Chinese Spring (CS) e delle sue linee nulli-tetrasomiche e ditelosomiche del gruppo di omeologia 1 ibridato con la sonda S-F4-R3, specifica per *TaPDIL4-1*. La localizzazione cromosomica di ogni frammento è indicata nella parte sinistra di ogni foto, a lato del profilo di ibridazione della cultivar CS, mentre i pesi molecolari del marker sono riportati nella parte destra della foto.

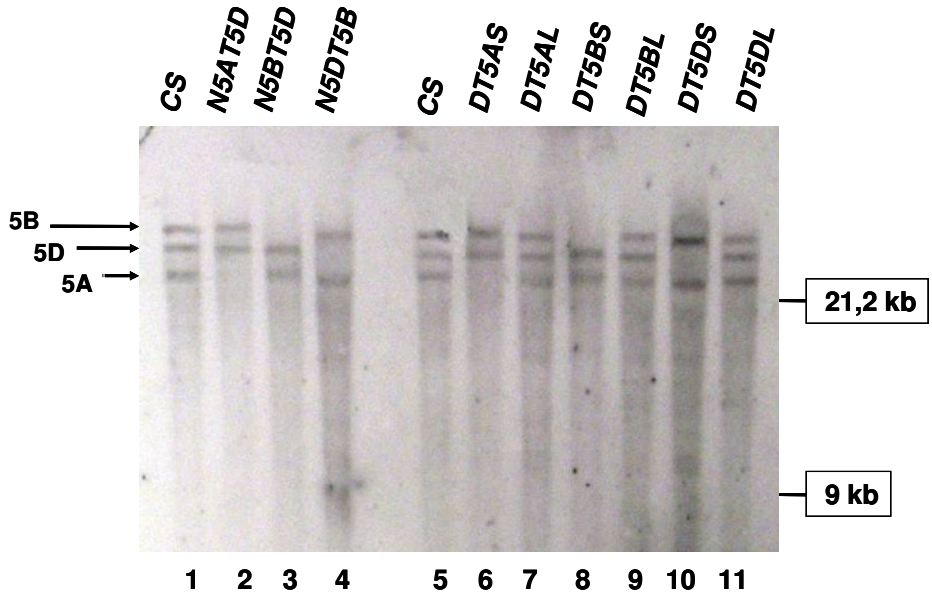

Fig. 5.10 - Analisi Southern del DNA genomico, digerito con E*co*RV, di Chinese Spring (CS) e delle sue linee nulli-tetrasomiche e ditelosomiche del gruppo di omeologia 5 ibridato con la sonda S-F3-R4, specifica per *TaPDIL5-1*. La localizzazione cromosomica di ogni frammento è indicata nella parte sinistra di ogni foto, a lato del profilo di ibridazione della cultivar CS, mentre i pesi molecolari del marker sono riportati nella parte destra della foto.

Il numero di tre sequenze geniche codificanti per ciascuna delle due proteine PDI-like identificate nella cultivar esaploide di frumento tenero Chinese Spring, una per genoma, è in linea con i risultati ottenuti dalle ricerche effettuate nelle banche dati EST di frumento, che hanno evidenziato la presenza di una sequenza appartenente al IV e una al V gruppo filogenetico. Per il V gruppo filogenetico i risultati ottenuti con l'analisi Southern per la sequenza *TaPDIL**5**-1* sono in linea anche con i dati disponibili per riso e mais, in entrambe queste specie, infatti, è stata identificata una sola sequenza appartenente al quinto gruppo filogenetico, mentre in *Arabidopsis* ne sono state identificate due. Per il quarto gruppo filogenetico, invece, è stata identificata una sola sequenza in *Arabidopsis* e due sia in riso che in mais. Dal profilo di ibridazione ottenuto per *TaPDIL**4**-1* e *TaPDIL**5**-1* in CS e nelle sue linee aneuploidi appare evidente che le sequenze geniche identificate per ciascuna di esse sono relative a sequenze omeologhe localizzate, rispettivamente, sui cromosomi dei gruppi 1 e 5. Va sottolineato, inoltre, che per *TaPDIL**4**-1* sono state usate due differenti sonde, che insieme coprono l'intera sequenza codificante di questo gene, e che il profilo di ibridazione ottenuto con entrambe le sonde era identico e costituito di 3 bande localizzate nel gruppo di omeologia 1. Questi risultati suggeriscono che la presenza di due sequenze del IV gruppo filogenetico in riso e in mais possa derivare da eventi di duplicazione genica avvenuti indipendentemente in queste specie.

5.7. Analisi northern dei trascritti relativi ai geni *TaPDIL4-1*, *TaPDIL5-1* e *TaPDIL2-1*

Per studiare il profilo di espressione di tre delle sequenze geniche isolate codificanti per proteine PDI-like in frumento, il loro livello di trascrizione è stato confrontato con quello della PDI classica (*TaPDIL**1**-1*), identificata e caratterizzata precedentemente da Ciaffi e collaboratori (2006). I profili di espressione relativi a *TaPDIL**1**-1, TaPDIL**4**-1*, *TaPDIL**5**-1* e *TaPDIL**2**-1* sono stati analizzati mediante northern di RNA totali estratti da diversi tessuti della varietà di frumento tenero CS: radici e piantine (fusto e prime foglie vere) raccolte da semenzali 20 giorni dopo la germinazione; foglie a bandiera raccolte dalle piante mature allo stadio di spigatura completa; spighette in formazione a diversi stadi di sviluppo raccolte, ad intervalli regolari di 5-7 giorni, a partire da quando l'infiorescenza raggiunge la dimensione di 1-1,5 cm fino a spigatura completa; cariossidi in via di sviluppo raccolte tra 4 e 36 giorni dopo l'antesi (GDA).

Come atteso, la *TaPDIL**1**-1* (Fig. 5.11) risulta presente costitutivamente in tutti i tessuti analizzati, ma l'espressione risulta particolarmente elevata nelle cariossidi, con il livello di trascrizione più alto tra l'8° ed il 12° giorno dopo l'antesi. Nelle spighe il livello di espressione della PDI classica è risultato relativamente più elevato negli stadi intermedi del loro sviluppo. Il livello di espressione più basso è stato riscontrato per i tessuti vegetativi, quali foglie mature e piantine.

La *TaPDIL**4**-1* (Fig. 5.11) è espressa costitutivamente in tutti i tessuti analizzati, come la forma classica della PDI, ma mentre *TaPDIL**1**-1* è espressa preferenzialmente nelle cariossidi in via di sviluppo, il livello di espressione più elevato per *TaPDIL**4**-1* è stato rilevato nelle spighe in formazione. Livelli paragonabili di trascrizione sono presenti, inoltre, nella cariossidi raccolte nei primi giorni successivi all'antesi (4 e 8 GDA), con tempi leggermente spostati in avanti rispetto alla forma classica, il cui livello massimo di espressione si riscontra tra l'ottavo ed il dodicesimo giorno dopo l'antesi. Il livello di trascrizione della *TaPDIL**4**-1* nei tessuti vegetativi è risultato, come per la *TaPDIL**1**-1*, più elevato nelle radici che nelle foglie a bandiera e nelle piantine. Il livello relativo dei trascritti della *TaPDIL**4**-1*, rispetto alla forma classica della PDI, risulta maggiore nelle spighe in formazione, mentre il livello dei trascritti nelle prime fasi di sviluppo delle cariossidi risulta più basso. Questo è in accordo con quanto riportato da Shorrosh e Dixon (1992) per la PDI di erba medica appartenente a questo gruppo filogenetico. Questa PDI risulta espressa con livelli di trascrizione approssimativamente simili in tutti i tessuti analizzati.

Nelle spighe in via di sviluppo e nelle cariossidi in formazione la *TaPDIL**5**-1* (Fig. 5.11) ha mostrato un profilo simile a quello della *TaPDIL**4**-1*, anche se il livello di espressione complessivo è risultato più basso. Non sono stati rilevati trascritti nelle piantine e nelle foglie mature, mentre l'espressione nelle radici era di intensità simile a quella rilevata nelle spighe e nelle cariossidi.

La *TaPDIL**2**-1* (Fig. 5.11), infine, ha mostrato un profilo di espressione simile a quello della *TaPDIL**1**-1*. Nelle spighe l'intensità è risultata paragonabile a quella della PDI classica, se non addirittura leggermente superiore, ed il massimo livello di trascrizione si verifica nei primi stadi di sviluppo. Il profilo di espressione nelle cariossidi è simile a quello della PDI classica, con un massimo tra 8 e 12 giorni dopo l'antesi, ma il livello di trascrizione nei singoli stadi di sviluppo risulta molto più basso rispetto alla PDI classica. Nei tessuti vegetativi, infine, come per la *TaPDIL**5**-1* si riscontrano trascritti solo nella radice.

L'analisi di espressione di un gene codificante per una proteina PDI-like appartenente al II gruppo filogenetico è stata condotta in diversi tessuti di carota da Xu et al. (2002). Tra i

tessuti vegetativi analizzati, i trascritti di questo gene risultavano particolarmente abbondanti nelle radici, mentre il segnale di ibridazione era molto debole nelle foglie mature. Un segnale più intenso si riscontrava nelle foglie giovani, materiale che non è stato analizzato per il frumento, rendendo impossibile un confronto specifico. Le piantine raccolte 20 giorni dopo la germinazione sono formate, infatti, dalle prime 2-3 foglie vere, che ancora non svolgono un'attività fotosintetica completa in quanto sfruttano per la loro crescita le riserve del seme. È quindi probabile che questi tessuti siano troppo giovani per manifestare l'espressione di *TaPDIL2-1*. Sarebbe interessante analizzare l'espressione di questo gene di frumento in foglie raccolte a differenti stadi di sviluppo e sottoposte a diverse intensità di luce. La luce gioca un ruolo essenziale, infatti, nell'incrementare i livelli di trascrizione dell'ortologo di *Arabidopsis AtPDIL2-1* (At3g54960) (Lu e Christopher, 2006). L'espressione di questo gene aumenta in risposta alla luce, durante lo sviluppo dei cloroplasti, mentre non è indotto nelle foglie mature. Per questo gene è stato proposto un ruolo nella biogenesi transitoria dell'amido, mentre un'altra PDI-like del genoma di *Arabidopsis* potrebbe essere coinvolta, come la RB60 di *C. reinhardtii*, nella regolazione del gene *psbA* (Shen et al., 2001). Ciò potrebbe spiegare la mancanza di espressione di *TaPDIL2-1* in piantine giovani e fotosinteticamente inattive di frumento e la presenza relativamente elevata di un suo ortologo in giovani foglie fotosintetizzanti di carota (Lu e Christopher, 2006). Un livello di espressione relativamente elevato è stato riscontrato anche in gemme apicali ed embrioni somatici di carota, tessuti per i quali non è possibile fare confronti diretti con quelli analizzati in questa ricerca. È ipotizzabile che il livello di trascrizione più elevato nelle cariossidi in via di sviluppo del frumento possa essere dovuto alla presenza degli embrioni. Non si può escludere, tuttavia, che l'elevato livello di trascrizione ottenuto negli embrioni somatici di carota sia dovuto alle condizioni di induzione scelte o rappresenti una caratteristica peculiare della specie analizzata.

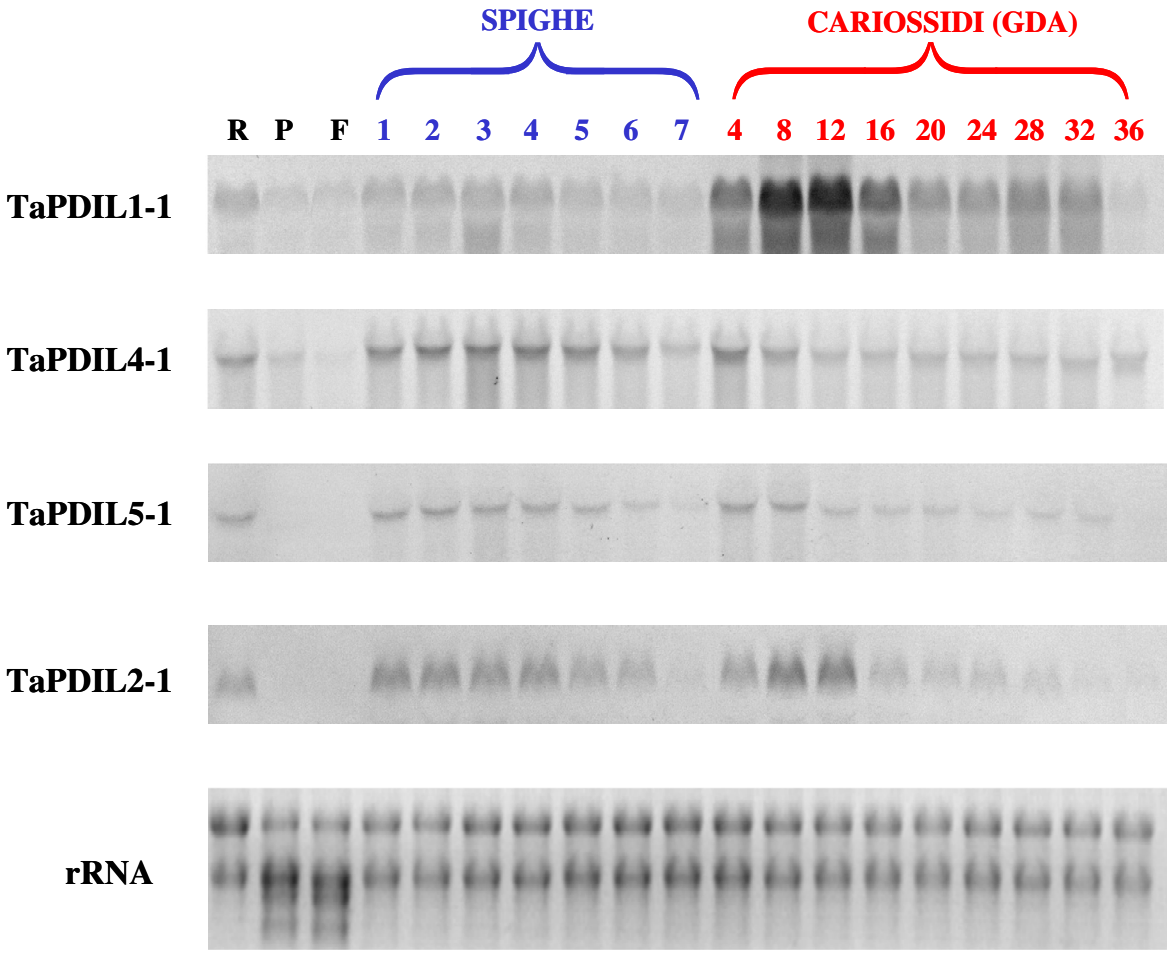

Fig. 5.11 – Analisi northern dei trascritti dei geni codificanti per TaPDIL**1**-1, TaPDIL**4**-1, TaPDIL**5**-1 e TaPDIL**2**-1 in differenti tessuti. R: radice di piantine 20 giorni dopo la germinazione; P:piantine (fusto e prime foglie vere) 20 giorni dopo la germinazione; F: foglie a bandiera raccolte a spigatura completa; Spighe 1-7 (blu): spighette a 7 differenti stadi di sviluppo, raccolte ad intervalli regolari di 5-7 giorni, a partire da quando l'infiorescenza è lunga 1-1,5 cm fino a spigatura completa; Cariossidi 4-36 giorni dopo l'antesi (GDA): cariossidi a differenti stadi di sviluppo raccolti ad intervalli regolari di 4 giorni a partire dal 4° giorno dopo l'antesi fino al 36° giorno dopo l'antesi. Quantità uguali di RNA totale (10µg) sono state caricate nel gel colorato con etidio bromuro.

5.8. Ottenimento del costrutto per la sovra-espressione della PDI classica

Per la sovra-espressione del gene della PDI classica di frumento è stata utilizzata la sequenza del promotore del gene *Glu-1D1*, codificante per la subunità gluteninica ad elevato peso molecolare 1Dx5 (HMW-GS) di frumento. Esperimenti in frumento transgenico, in cui il gene reporter *UidA* era stato posto sotto il controllo di parte di questa sequenza regolatrice, in particolare quella compresa tra le posizioni -1191 e +58, hanno dimostrato che questa regione conferisce un'elevata espressione endosperma specifica (Lamacchia et al., 2001). Analisi di espressione hanno infatti confermato che il gene *UidA* posto sotto il controllo di questa regione regolatrice è espresso limitatamente all'endosperma di cariossidi di frumento in via di sviluppo a partire da 10-12 giorni dopo l'antesi. Halford e collaboratori (1989) hanno inoltre dimostrato che anche la regione compresa tra -277 e +39 di questa sequenza regolatrice è sufficiente, almeno in tabacco, ad assicurare un'espressione endosperma specifica del gene *UidA*. In questo caso il massimo livello di espressione si ha intorno a 20-22 giorni dopo l'antesi. Sulla base della sequenza nucleotidica della regione in 5' del gene *Glu*-1D1, isolata dalla cv. Cheyenne e depositata in banca dati EMBL con numero di accessione AJ301618 da Lamacchia et al. (2001), è stata disegnata una coppia di primer (**1Dx5-5'**-F1: 5'-GGGAAGTTAGCCATGCCGA-3' e **1Dx5-5'**-R1: 5'-GGACTATCAGTGAATTGATCTCTAG-3') per l'amplificazione di questa sequenza regolatrice da DNA della cv Centauro che, come la cv Cheyenne, possiede la subunità gluteninica 1Dx5 ad elevato peso molecolare. Il frammento di amplificazione ottenuto è lungo 1094 pb ed è compreso tra le posizioni -1039 e +55, come mostrato in figura 5.12. Dopo essere stata amplificata, questa sequenza di 1094 pb, che per comodità indicheremo con 1Dx5-5', è stata clonata nel vettore pGEM-T Easy Vector (Promega) e sequenziata. Dall'allineamento della sequenza ottenuta con la corrispondente sequenza della cv. Cheyenne sono state riscontrate solo 5 singole sostituzioni nucleotidiche, nessuna delle quali ha interessato sequenze conservate di elementi regolatori identificati fino ad oggi in questa sequenza regolatrice (Thomas e Flavell, 1990; Lamacchia et al., 2001; Norre et al., 2002) (Fig. 5.12). Risultano perfettamente conservati, quindi, oltre al TATA box e al sito di inizio della trascrizione, anche l'elemento regolatore bi-fattoriale identificato da Norre et al., (2002), che consiste nel motivo G-box like e nel prolamin box in posizione prossimale e rappresenta il principale motivo regolatore responsabile dell'espressione endosperma specifica della sequenza regolatrice a monte del gene codificante per la subunità 1Dx5.

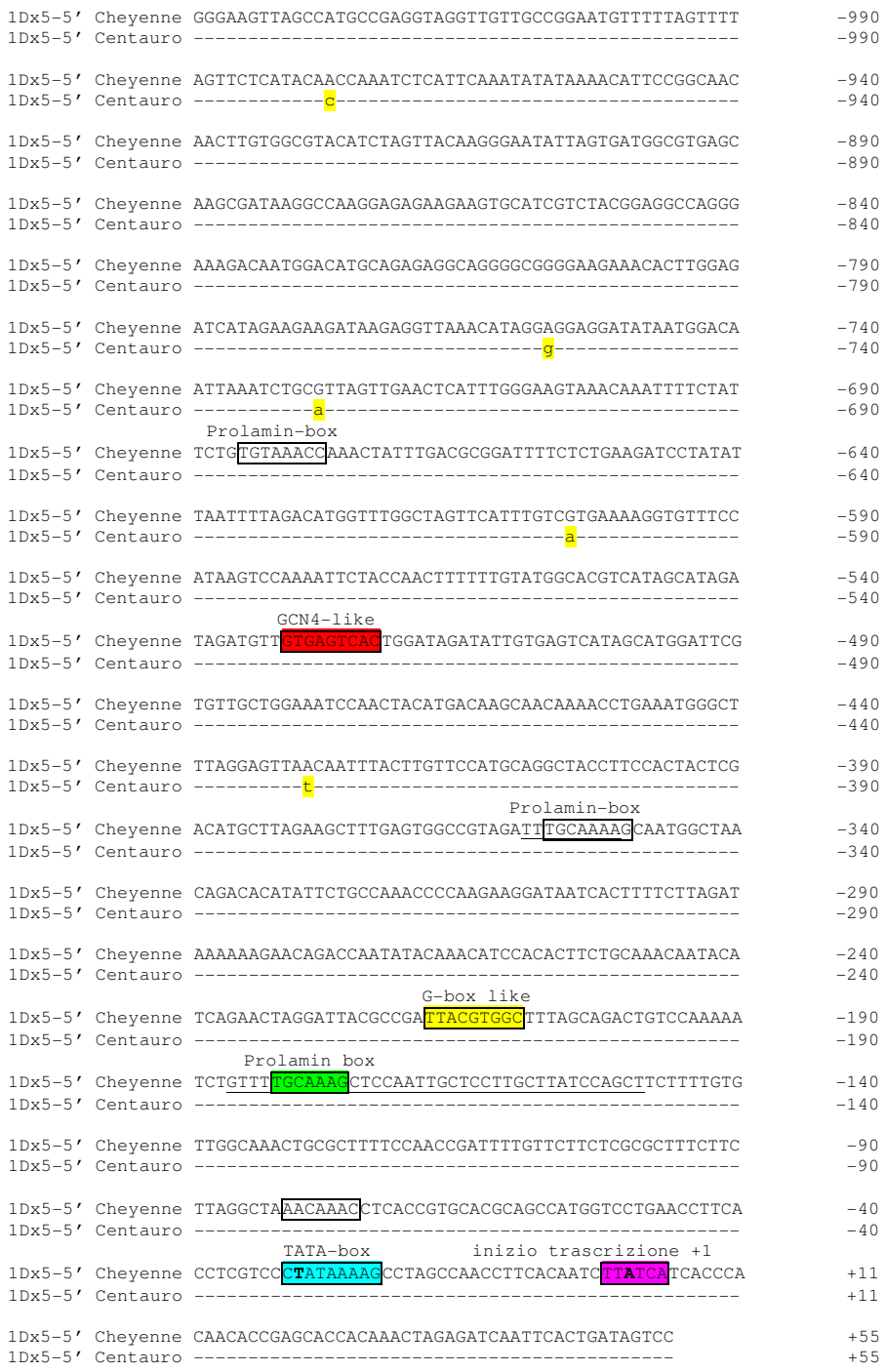

Fig. 5.12 - Allineamento delle sequenze nucleotidiche, fiancheggianti in 5' il codone ATG della regione codificante per la subunità della prolammina HMW 1Dx5, isolate dalla cv. Cheyenne (Lamacchia et al., 2001) e dalla cv. Centauro. Sono indicati gli elementi regolatori. L'intero elemento regolatore identificato da Thomas e Flavel (1990), la cui posizione è conservata in tutti i promotori delle prolammine HMW (Shewry et al., 1999), è sottolineato. Il TATA box, il sito di inizio della trascrizione, il prolamin box, il motivo GCN4 like, il motivo G box like e due altri potenziali prolamin boxes sono racchiusi da rettangoli. Le 5 differenze nucleotidiche della sequenza regolatrice isolata nella cv. Centauro rispetto alla sequenza regolatrice originale isolata in Cheyenne sono evidenziate con lettere minuscole, i restanti nucleotidi conservati sono rappresentati da trattini.

La sequenza regolatrice in 5' di 1094 pb del gene *Glu-1D1*, isolata dalla cv. Centauro, è stata quindi riamplificata, usando come DNA stampo il DNA del plasmide pGEM-T, nel quale era stata inizialmente clonata. È stata usata una coppia di primer-adattatori (AF1-AR1) allo scopo di aggiungere ad entrambe le estremità 3' e 5' della sequenza un sito di restrizione per l'enzima *Aat*II. Nella regione 5' del promotore, internamente al sito *Aat*II, è stato inserito anche un sito *Pst*I, necessario nella seconda fase di clonazione (Fig. 5.13).

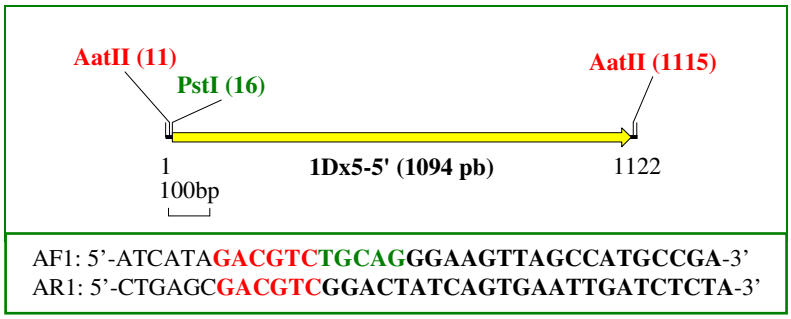

Fig. 5.13 - Mappa di restrizione del frammento di amplificazione della regione regolatrice in 5' della subunità 1Dx5 dalla posizione -1039 a +55 con i primer adattatori AF1 e AR1. AF1 inserisce, all'estremità 5', un sito di restrizione esterno per l'enzima *Aat*II ed uno più interno per l'enzima *Pst*I, mentre AR1 inserisce un sito di restrizione *Aat*II all'estremità 3'. In figura sono indicate le esatte posizioni dei siti di restrizione e la lunghezza della sequenza regolatrice, indicata come 1DX5-5'. In figura sono riportate, inoltre, le sequenze nucleotidiche dei primer-adattatori. Le sequenze dei siti di restrizione sono evidenziate nel seguente modo: *Aat*II*:* (GACGTC) e *Pst*I*:* (CTGCAG).

Il prodotto di questa amplificazione, dopo essere stato tagliato con l'enzima di restrizione *Aat*II e purificato da gel, è stato sub-clonato, all'interno del sito *Aat*II (20 pb) del plasmide pGEM-T Easy Vector (Promega), a monte del frammento cDNA comprendente l'intera regione codificante del gene della *TaPDIL1-1*, precedentemente clonato nel sito *Eco*RV (Fig. 5.14 e Fig. 5.15).

Tramite analisi di restrizione, usando l'enzima *Pst*I, sono stati identificati i cloni in cui la sequenza regolatrice 1Dx5-5' era inserita nell'orientamento desiderato rispetto alla sequenza codificante della *TaPDIL1-1* (Fig. 5.15) ed il corrispondente frammento di restrizione, contenente sia la sequenza regolatrice 1Dx5-5' che la sequenza codificante per la *TaPDIL1-1* nel giusto orientamento, è stato purificato da gel.

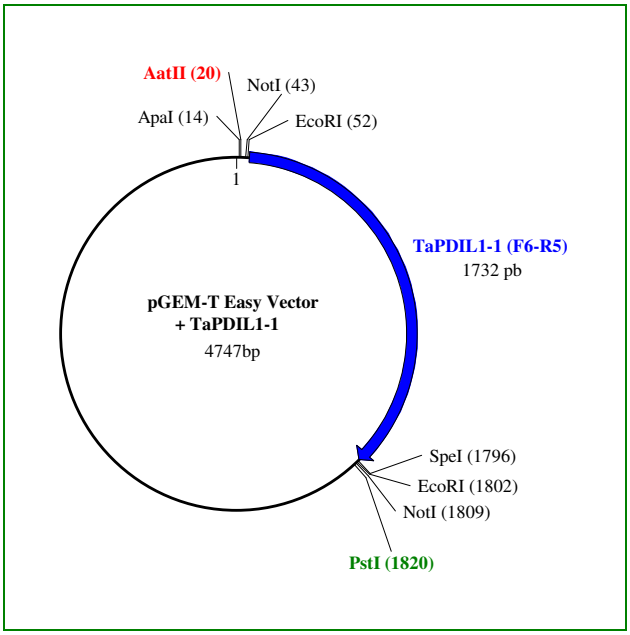

Fig. 5.14 - Mappa di restrizione del plasmide ricombinate pGEM-T Easy Vector nel cui sito *Eco*RV è stata clonata la sequenza codificante della TaPDIL**1**-1 amplificata con i primer F6-R5 e lunga 1732 pb. La sequenza codificante e la sua orientazione sono rappresentate dalla freccia di colore blu. I siti di restrizione usati nella diverse fasi di clonazione, *Aat*II e *Pst*I, sono indicati a colori ed in grassetto. Il sito *Aat*II, al cui interno verrà clonata la sequenza promotrice, è evidenziato in rosso Sono inoltre indicati alcuni ulteriori siti di restrizione del polylinker del plasmide pGEM-T Easy Vector.

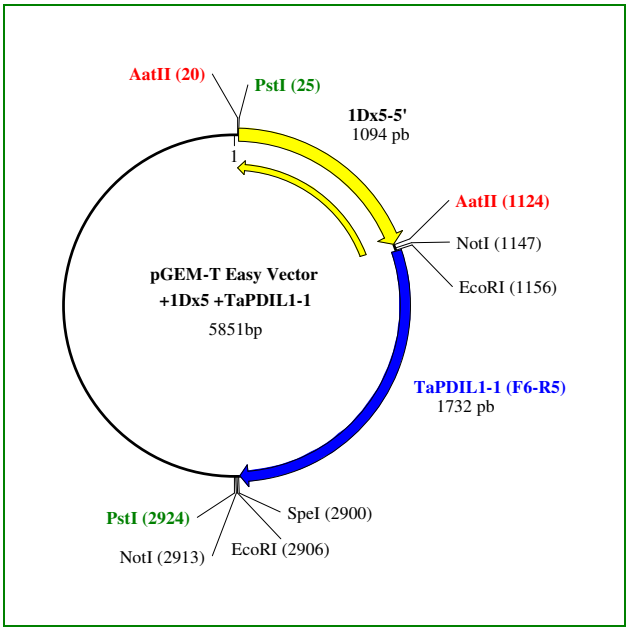

Fig. 5.15 - Mappa di restrizione del plasmide ricombinate pGEM-T Easy Vector (Promega) contenente la sequenza codificante della TaPDIL**1**-1 a monte della quale, all'interno del sito di restrizione *Aat*II (20) del polylinker è stata inserita la sequenza regolatrice della subunità 1Dx5, lunga 1094 pb, nel giusto orientamento. La freccia gialla più interna indica la possibilità di orientamento inverso della sequenza regolatrice dopo il suo inserimento. Sono indicati, inoltre, alcuni ulteriori siti di restrizione del polylinker del plasmide pGEM-T Easy Vector. In grassetto ed a colori sono indicati i siti di restrizione usati nelle fasi di sub-clonaggio. In verde è indicato il sito *Pst*I, che verrà usato per l'inserimento del frammento contenente la sequenza regolatrice 1Dx5-5' assieme alla sequenza codificante della TaPDIL**1**-1 all'interno del plasmide pAHC20, digerito a sua volta con *Pst*I.

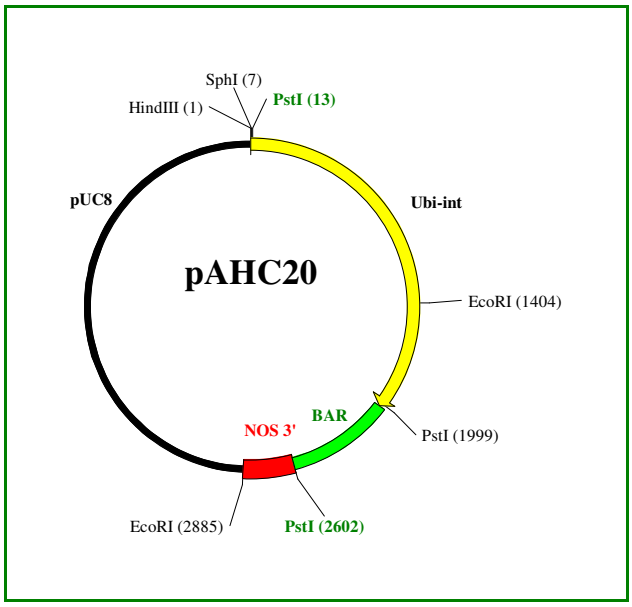

Fig. 5.16 - Mappa di restrizione del plasmide pAHC20. Ubi-int: promotore del gene codificante per l'ubiquitina in mais (Christensen et al., 1992) associato al primo introne di *Ubi*, che ne aumenta l'efficienza di trascrizione. Bar: gene conferente la resistenza alla fosfinotricina. NOS 3': sequenza di terminazione della trascrizione derivata dal gene per la nopalina sintetasi di agrobatterio.

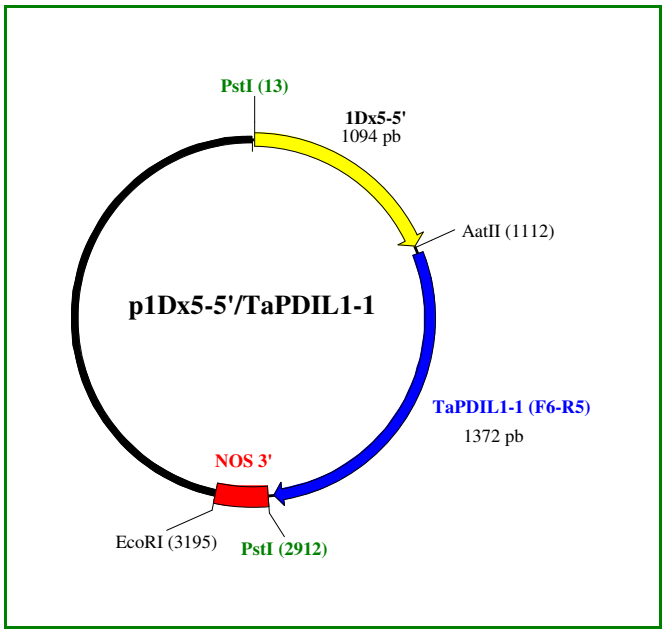

Fig. 5.17 - Mappa di restrizione del plasmide di sovra-espressione finale, p1Dx5-5'/*TaPDIL**1**-1*. All'interno di uno scheletro di base fornito dal plasmide pUC8 si trova inserito il costrutto chimerico costituito dalla regione regolatrice 1Dx5-5' di 1094 pb, dalla posizione -1039 a +55, l'intera regione codificante per la TaPDIL**1**-1 di 1732 pb (F6-R5) ed il terminatore NOS (Nopaline Synthase).

Contemporaneamente è stato digerito, con l'enzima *Pst*I, anche il plasmide pAHC20 (Fig. 5.16) per identificare e purificare da gel il frammento di restrizione contenente sia la sequenza del plasmide pUC8 che la sequenza del terminatore NOS 3'(derivante dal gene nopalina sintetasi di agrobatterio). Quest'ultimo frammento di restrizione è stato quindi ligato

al frammento di restrizione *Pst*I-*Pst*I, contenente la sequenza regolatrice 1Dx5-5' e la sequenza codificante per la TaPDIL**1**-1 precedentemente purificato. Il plasmide, in cui i tre elementi promotore, gene e terminatore sono orientati correttamente gli uni rispetto agli altri, all'interno del plasmide pUC8, è stato identificato tramite analisi di restrizione usando l'enzima *Eco*RI. Il plasmide ricombinante così identificato è stato denominato p1Dx5-5'/*TaPDIL**1**-1* (Fig. 5.17), e la sua autenticità è stata verificata mediante sequenziamento.

5.9. Ottenimento dei costrutti per lo studio funzionale del promotore del gene della PDI classica

Allo scopo di studiare i meccanismi regolatori responsabili dell'espressione della *TaPDIL**1**-1* e, in particolare, di individuare il frammento minimo del promotore necessario a conferire l'elevata espressione nelle cariossidi in via di sviluppo, il gene reporter *UidA*, codificante per la β-glucuronidasi (GUS), è stato posto sotto il controllo di 4 frammenti di delezione progressivi di una delle tre sequenze regolatrici in 5' isolate e caratterizzate in precedenza (Ciaffi et al., 2006). La scelta di usare una delle tre sequenze regolatrici è dettata da due diverse considerazioni, in primo luogo le sequenze regolatrici in 5' dei tre geni omeologhi della PDI classica (TaPDIL**1**-1) isolate da Ciaffi e collaboratori (2006) presentano un elevato grado di omologia e di conservazione dei diversi motivi regolatori potenzialmente coinvolti nell'espressione endosperma specifica (Tab. 1.4 e Fig. 5.18). In secondo luogo l'analisi dell'espressione dei tre geni omeologhi della PDI aveva indicato che i relativi trascritti non presentano differenze significative nei profili e livelli di espressione rilevati nelle cariossidi in via di sviluppo (Ciaffi et al., 2006). Considerando quindi l'elevata omeologia di sequenza e la conservazione, in tutte le tre sequenze isolate, dei diversi motivi regolatori individuati, nonché il profilo di espressione quasi identico durante la fase di riempimento delle cariossidi è stata scelta, per le successive analisi, la sequenza promotrice di uno dei geni *TaPDIL**1**-1*, quello localizzato sul cromosoma 4D. In Fig. 5.18 è riportata la sequenza del promotore del gene codificante per la TaPDIL**1**-1 e sono indicati i diversi motivi, individuati da Ciaffi e collaboratori (2006), coinvolti nella regolazione di geni espressi specificatamente nell'endosperma. È indicato anche un ulteriore elemento regolatore coinvolto nell'espressione endosperma specifica, definito elemento ACGT, individuato in una successiva analisi del promotore della *TaPDIL**1**-1* (Fig. 5.18).

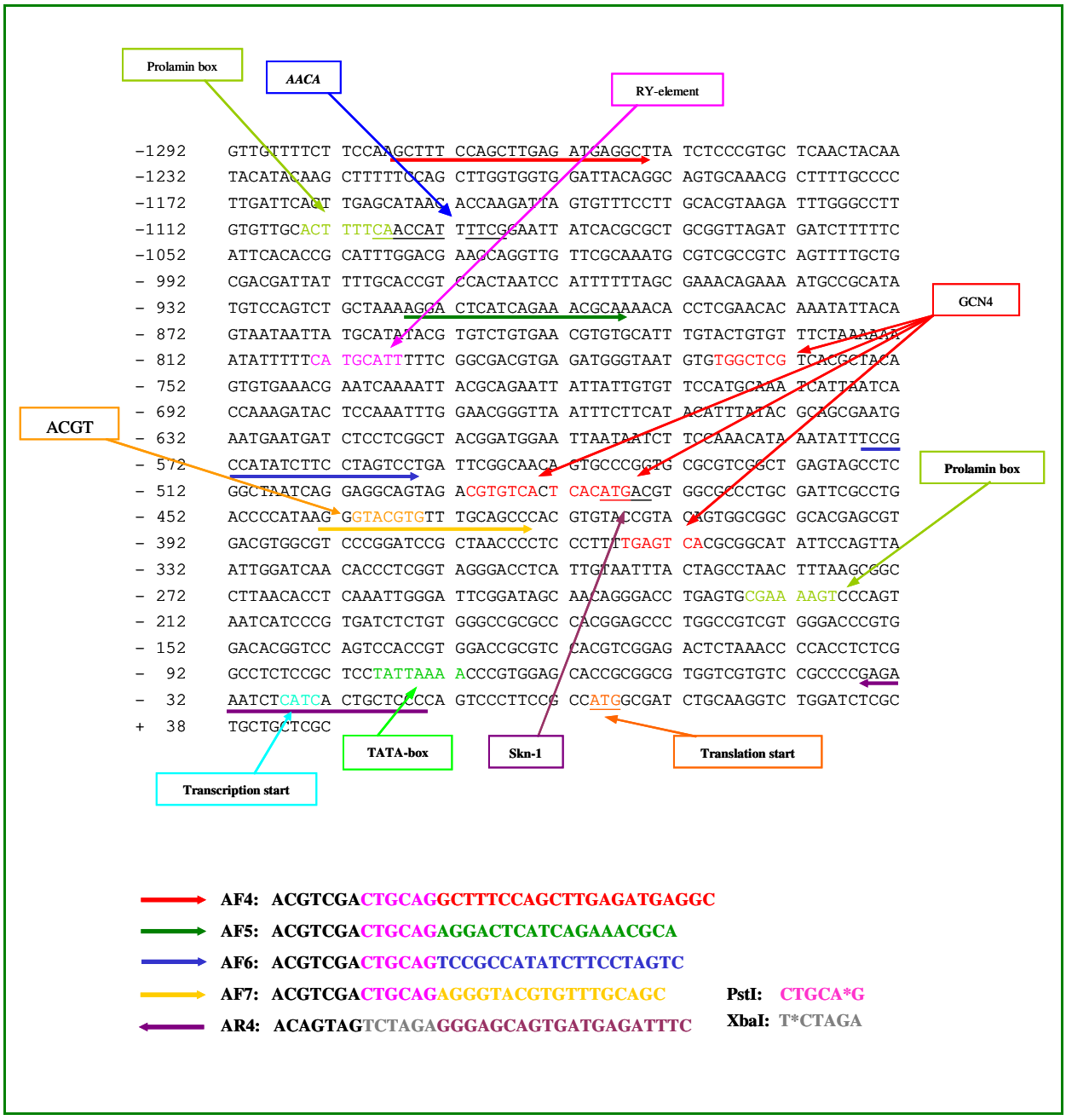

Fig. 5.18 – Promotore del gene codificante per la TaPDIL**1**-1 localizzato sul cromosoma 4D. Sono evidenziati i motivi regolatori identificati. Sono inoltre indicati i primer usati per l'amplificazione dei diversi frammenti di delezione. Le sequenze e le frecce identificate dai colori rosso, verde, blu, giallo e viola corrispondono alla parte del primer complementare alla sequenza regolatrice, in rosa ed in grigio sono indicati i siti di restrizione inseriti rispettivamente alle estremità 5' e 3' di ognuno dei 4 frammenti di amplificazione. Sono inoltre indicate in nero le 7 basi aggiunte all'estremità 5' dei primer per assicurare il corretto funzionamento degli enzimi di restrizione.

5.9.1. Ruolo funzionale dei motivi regolatori individuati nel promotore di TaPDIL1-1

Il ruolo di alcuni degli elementi regolatori di tipo cis individuati nel promotore di *TaPDIL**1**-1* nell'espressione endosperma specifica è stato dimostrato attraverso l'analisi funzionale, basata su esperimenti di perdita di funzione (loss of function) e di ripristino di funzione (gain of function). Inizialmente le analisi dei promotori di geni preferenzialmente

espressi nei semi e nell'endosperma di specie cerealicole sono state eseguite in sistemi eterologhi, usando piante dicotiledoni come il tabacco, più facili da trasformare e rigenerare. Va considerata, tuttavia, la possibilità che la regolazione spaziale e temporale dell'espressione genica possa non essere mantenuta in un background genetico eterologo (Hirano et al., 1995b; Grosset et al., 1997; Leisy et al., 1989). Bisogna considerare, infatti, che la regolazione dell'espressione genica dipende non solo da vari elementi "cis-acting" prossimali alla sequenza codificante, ma anche da elementi "trans-acting" localizzati in posizioni anche molto lontane dalla regione codificante e dalle sue sequenze regolatrici prossimali. È quindi più affidabile ed auspicabile l'uso di sistemi omologhi. La messa a punto di protocolli efficienti di trasformazione anche per alcune specie monocotiledoni ha permesso di eseguire l'analisi funzionale anche in queste piante. Il riso, in particolare, è diventato la specie modello per le monocotiledoni anche per questo tipo di studi (Digeon et al., 1999). Avendo a disposizione un protocollo di trasformazione efficiente per il frumento, è stato deciso di effettuare l'analisi funzionale del promotore della PDI classica di frumento nella stessa specie per ridurre al minimo la possibilità di ottenere profili di espressione aberranti e per verificare direttamente il funzionamento dei diversi frammenti regolatori, ottenuti per delezioni successive della regione distale, in frumento.

L'abbondanza di informazioni relative al funzionamento in piante transgeniche dei motivi regolatori individuati permette di fare alcune ipotesi sul possibile ruolo dei diversi frammenti di delezione della sequenza promotrice della *TaPDIL1-1* nel regolare l'espressione del gene reporter *UidA* in frumento transgenico e, sulla base di queste, di impostare un adeguata strategia sperimentale. I diversi motivi regolatori sembrano contribuire in maniere diversa al profilo di espressione finale del promotore di cui fanno parte. GCN4 e prolamin box sono motivi conservati nei promotori di molti geni codificanti per proteine di riserva di diversi cereali, quali mais, riso, frumento, orzo, avena e sorgo (Müller e Knudsen, 1993). Nei promotori di molti geni codificanti per prolammine è presente un numero limitato di motivi conservati, tra i quali il più conservato è un elemento di 25 pb situato a circa 300 pb a monte del codone di inizio della traduzione (Kreis et al., 1985). Questo elemento conservato, definito "endosperm box", consiste di un prolamin-box e di un motivo GCN4-like situati a pochi nucleotidi di distanza l'uno dall'altro. L'analisi funzionale dell'endosperm box di una γ-zeina ha dimostrato l'importanza di entrambi i motivi per la massima espressione endosperma specifica di questo gene in mais (Marzàbal et al., 1998). È interessante notare che nella regione prossimale del promotore della HMWG della subunità Dx5 (PrHMWG-Dx5) è presente un "endosperm box" atipico, in cui manca il motivo GCN4-like ed è invece presente

un motivo G-box like. In mais transgenico è stato dimostrato che un frammento di 89 pb, comprendente questo "endosperm box" atipico, è responsabile dell'espressione endosperma specifica (Norre et al., 2002). AACA è un terzo motivo coinvolto nell'espressione endosperma specifica e risulta conservato in tutti i promotori dei geni codificanti per le gluteline isolati fino ad ora in riso (Takaiwa et al., 1996; Washida et al., 1999). Anche questo motivo risulta strettamente associato al motivo GCN4-like. Questi due elementi insieme sono capaci di conferire un'espressione endosperma specifica al promotore 35S di CaMV troncato a -90 (Yoshihara et al., 1996). È importante sottolineare che questo frammento del promotore virale contiene, in aggiunta al TATA box, anche un elemento G-box like (as-1) (Washida et al., 1999). Il motivo ACGT presente nel promotore della PDI classica ha la sequenza consenso GTACGTG, identica a quella presente nel promotore del gene *GluB-1* di riso (Washida et al., 1999). Da analisi funzionali di questo promotore è emerso che una regione promotrice minima di 197 pb a monte della sequenza trascritta e contenente 4 motivi regolatori (GCN4, AACA, ACGT e prolamin-box), è sufficiente a garantire un'espressione endosperma specifica in piante di riso transgeniche (Wu et al., 1998). I diversi elementi sembrano comunque contribuire in maniera diversa all'espressione endosperma specifica, il motivo GCN4, in particolare, risulta di centrale importanza. La mutazione di una base all'interno della sequenza consenso di questo motivo nella regione promotrice minima di 197 pb ha determinato una riduzione significativa dell'attività del promotore e l'alterazione del profilo di espressione nell'endosperma (Wu et al., 1998). È stato dimostrato, inoltre, che copie multiple ripetute a tandem di questo elemento hanno prodotto, in piante di riso transgenico, un profilo di espressione simile al promotore nativo (Wu et al., 1998). Anche i tre elementi restanti, AACA, prolamin-box e ACGT sono necessari a garantire la massima attività della regione del promotore di 197 pb, esercitando un effetto quantitativo sulla regolazione dell'espressione. Sostituzioni di basi in questi motivi hanno avuto come conseguenza una diminuzione o addirittura la perdita (AACA) dell'attività della regione promotrice di 197pb, evidenziando l'importanza di questi motivi per la sua massima attività (Wu et al., 2000). Nessuno di questi motivi, in forma multimerica ed associato al promotore -46 di CaMV 35 S, contenente il solo TATA box, è però in grado di conferire un'attività trascrizionale individuabile, verosimilmente perché essi hanno un effetto quantitativo sull'espressione nell'endosperma, ma non sono in grado di determinare da soli, a differenza del motivo GCN4, un'espressione endosperma specifica. GCN4 è l'unico motivo, quindi, necessario e sufficiente per determinare l'espressione endosperma specifica quando è presente in copie multiple. Anche la presenza di un motivo GCN4 associato ad AACA o ad un prolamin-box risulta

insufficiente per conferire livelli individuabili di espressione nell'endosperma, mentre una loro dimerizzazione, anche quando c'è l'inversione di una coppia, lo è. Visto che un dimero di GCN4 non è sufficiente a conferire un livello individuabile di espressione in riso transgenico, l'espressione dovuta alla presenza di una doppia coppia GCN4–AACA o GCN4-prolamin box è legata ad un effetto combinato tra i motivi. La combinazione AACA, ACGT e prolamin box, invece, è inattiva in riso transgenico (Wu et al., 1998). Dai risultati ottenuti con questi esperimenti in riso transgenico è quindi emerso che una combinazione lineare di almeno tre motivi, uno dei quali deve essere GCN4, è la condizione minima per ottenere un'espressione endosperma specifica (Wu et al., 2000). Ad esempio una combinazione di AACA, ACGT e GCN4 è risultata sufficiente a garantire un'espressione del gene GUS in riso transgenico, anche se il livello è risultato molto basso (Wu et al., 2000). Skn-1 è un ulteriore motivo regolatore identificato nel promotore di *TaPDIL**1**-1*. Questo motivo, che si trova sovrapposto ad uno degli elementi GCN4 nel promotore della *TaPDIL**1**-1*, è anche presente nel promotore del gene *GluB-1* e, anche in questo caso, la sua posizione è risultata molto vicina al motivo GCN4. Analisi in tabacco transgenico hanno mostrato che una combinazione di GCN4, AACA e Skn-1 assieme al frammento del promotore 35 S contenete oltre al TATA box l'elemento as-1, è sufficiente a garantire un elevato livello di espressione nei semi in maturazione. Anche in orientamento inverso questa combinazione mantiene la sua funzionalità, nonostante l'attivazione dell'espressione seme specifica risulti ridotta del 60% (Washida et al., 1999). Skn-1 ha un effetto quantitativo sul livello di espressione, in quanto mutazioni introdotte nella sua sequenza provocano una diminuzione del livello di espressione sostanziale, come dimostrato in tabacco (Washida et al., 1999). Infine c'è un ultimo elemento, RY-element, la cui sequenza consenso è CATGCA(C/T) (Dickinson et al., 1988). Questi elementi sono frequenti in promotori seme-specifici di monocotiledoni e dicotiledoni e sono noti anche come "legumin box" (Gatehouse et al., 1986) o "Sph element" (Kao et al., 1996).

5.9.2. Analisi funzionale dei frammenti di delezione del promotore di TaPDIL1-1

Sulla base della posizione dei diversi motivi individuati sono stati disegnati quattro diversi primer senso (AF4, AF5, AF6 e AF7) ed un unico primer antisenso (AR4), localizzato a 14 pb dall'ATG (Fig. 5.18). I 4 primer senso sono stati disegnati in modo da inserire, all'estremità 5' della sequenza del promotore della *TaPDIL**1**-1* a cui sono complementari, un sito di restrizione per *Pst*I (CTGCA*G), mentre il primer antisenso è stato disegnato in modo da inserire un sito di restrizione per *Xba*I (T*CTAGA) all'estremità 3' dei frammenti del

promotore (Fig. 5.18). L'aggiunta di questi due siti di restrizione era necessaria per le successive fasi di clonaggio perchè, nella sequenza del promotore non erano presenti siti di restrizione adatti. Le combinazioni di primer usate sono state: AF4-AR4 (1278 pb), AF5-AR4 (901 pb), AF6-AR4 (561 pb) e AF7-AR4 (430 pb); il numero tra parentesi si riferisce alla lunghezza in pb dei 4 differenti frammenti di delezione progressivi del promotore amplificati dalle relative coppie di primer (Fig. 5.18).

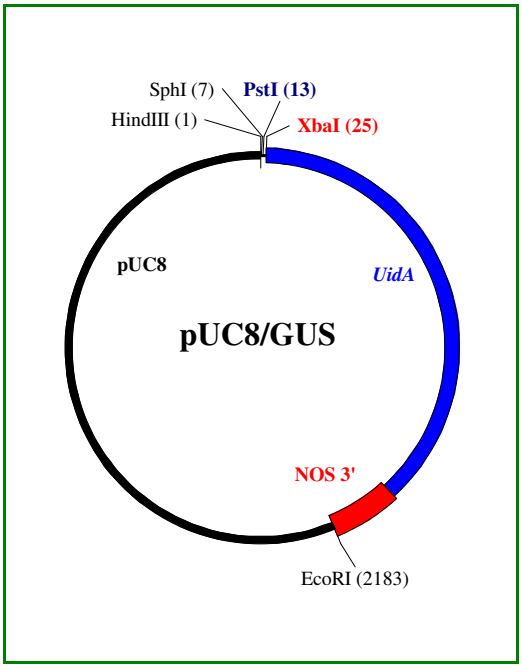

Fig. 5.19 - Mappa del plasmide ricombinante pUC8/GUS ottenuto ligando il frammento contenente la sequenza del plasmide pUC8 e il frammento contenente la sequenza del gene *UidA* assieme a quella del terminatore NOS 3'. Sono stati riportati i siti di restrizione usati durante le fasi di subclonaggio e due ulteriori siti, *Hin*dIII (1) e *Sph*I (7). Il sito *Pst*I (13) e il sito *Eco*RI (2170) sono stati usati nella costruzione del plasmide pUC8/GUS. I siti di restrizione *Pst*I (13)e *Xba*I (25) sono stati usati in seguito per subclonare i diversi frammenti di delezione del promotore e sono evidenziati, rispettivamente, in blu e rosso.

I 4 diversi frammenti di delezione del promotore, dopo essere stati amplificati con le quattro coppie di primer e tagliati con gli enzimi di restrizione *Pst*I e *Xba*I, sono stati purificati da gel e inseriti uno per volta, tramite clonazione direzionale (5' *Pst*I-*Xba*I 3'), nel plasmide pUC8/GUS a monte del gene *UidA* (Fig. 5.19).

Il plasmide pUC8/GUS è stato ottenuto digerendo il plasmide pAHC25 (Christensen e Quail, 1996) con gli enzimi *Eco*RI e *Pst*I e ligando due dei frammenti di restrizione ottenuti e purificati da gel, quello di circa 2800 pb, comprendente la sequenza del plasmide pUC8, e quello di circa 2200 pb, contenente la sequenza del gene *UidA* seguita dal terminatore NOS (Fig. 5.20).

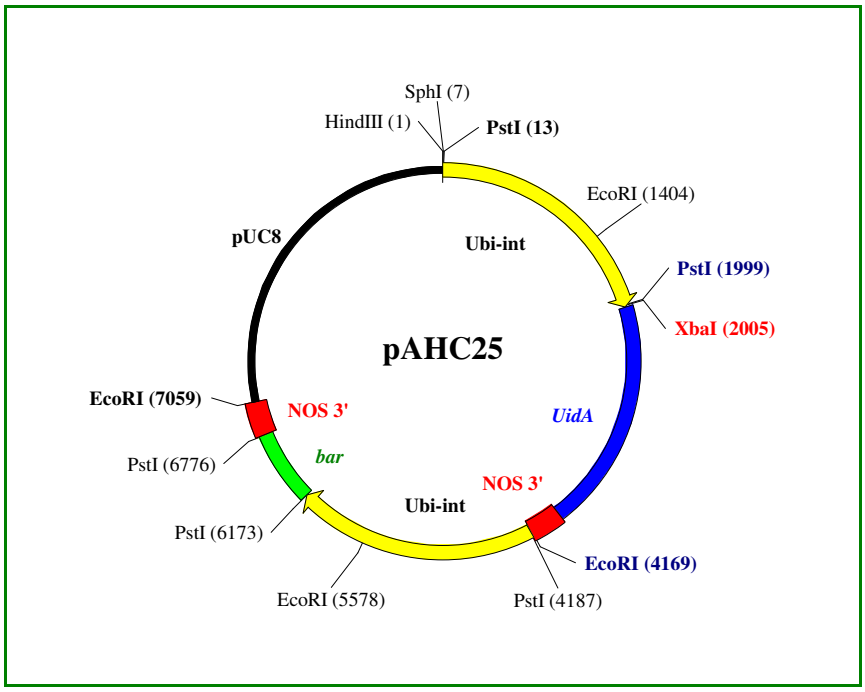

Fig. 5.20 - Mappa del plasmide pAHC25 (Christensen e Quail, 1996). Sono riportati tutti i siti di restrizione *Eco*RI e *Pst*I presenti nel plasmide oltre al sito *Xba*I (2005), che nel nuovo plasmide pUC8/GUS si trova 25 pb dopo il sito *Hin*dIII (1). I siti di restrizione *Pst*I (13) e *Eco*RI (7059) delimitano il frammento di 2800 pb circa contenente la sequenza del plasmide pUC8, utilizzata come scheletro di base nel plasmide pUC8/GUS, mentre il sito *Pst*I (1999) e il sito *Eco*RI (4169) delimitano il frammento di 2170 pb, contenente la sequenza del gene *UidA* e quella del terminatore NOS.

Sono stati così ottenuti i 4 plasmidi finali: pPDI**1278**-GUS (Fig. 5.21), pPDI**901**-GUS, pPDI**561**-GUS e pPDI**430**-GUS, la cui autenticità è stata confermata mediante analisi di sequenza. I numeri in grassetto si riferiscono alla lunghezza della sequenza regolatrice usata. Questi quattro plasmidi possiedono, quindi, tutti la medesima struttura di base, costituita dalla sequenza del plasmide pUC8, dalla sequenza codificante del gene *UidA* e da quella del terminatore NOS, mentre differiscono solo per la diversa lunghezza della sequenza regolatrice derivata dal promotore di *TaPDIL1-1*.

Le quattro sequenze regolatrici di diversa lunghezza, prodotte mediante delezione progressiva della sequenza del promotore all'estremità distale 5', pur essendo caratterizzate dalla presenza di un numero variabile di elementi regolatori, comprendono tutte le medesime 430 pb dell'estremità prossimale 3'. In particolare la sequenza regolatrice di dimensioni minori, ottenuta dall'amplificazione del promotore con la coppia di primer AF7-AR4, è costituita solo da queste 430 pb prossimali. Questo permetterà di identificare la regione minima della sequenza regolatrice necessaria per l'espressione nelle cariossidi in via di sviluppo ed eventuali effetti additivi esplicati dagli elementi progressivamente eliminati.

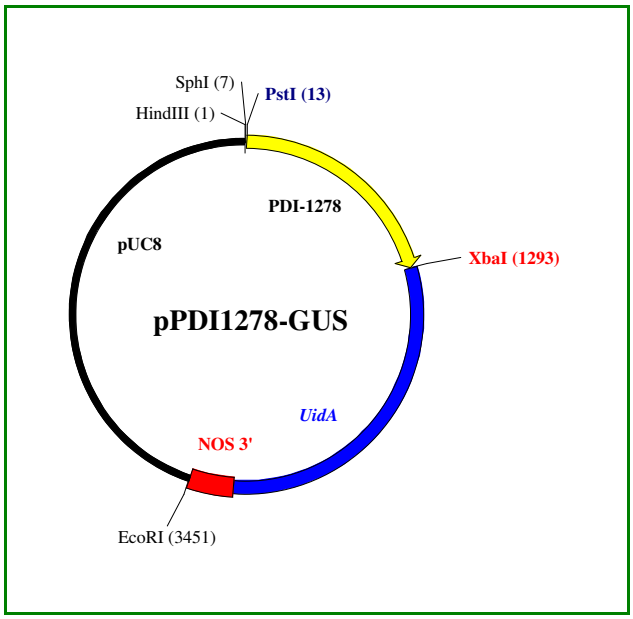

Fig. 5.21 - Mappa del più grande dei quattro plasmidi per lo studio del promotore, pPDI1278-GUS. In giallo è rappresentata la sequenza regolatrice di 1278 pb, contenente tutti gli elementi regolatori ad oggi identificati. I restanti tre plasmidi hanno una mappa di restrizione identica, eccetto per la lunghezza della sequenza regolatrice inserita tra i siti di restrizione *Pst*I e *Xba*I.

Il primo frammento di delezione, amplificato con la prima coppia di primer (AF4-AR4) è quello di dimensioni maggiori (1278 pb) e contiene tutti gli elementi regolatori identificati nel promotore (4 elementi GCN4-like, Skn-1, ACGT, Ry-element, AACA e due prolamin box; Fig. 5.18). Al secondo frammento di delezione, amplificato con la coppia di primer AF5-AR4 e lungo 901 pb, manca, rispetto al primo, il prolamin box distale e il motivo AACA. Il terzo frammento di delezione è lungo 561 pb ed è stato amplificato con la coppia di primer AF6-AR4; oltre al prolamin box distale e al motivo AACA mancano il RY-element e il primo dei quattro motivi GCN4-like. Infine l'ultimo frammento, amplificato con la coppia di primer AF7-AR4, è lungo 430 pb e mancano il prolamin box distale, il motivo AACA, il RY-element, il motivo Skn-1 e tre motivi GCN4-like, mentre contiene i seguenti elementi regolatori: un elemento GCN4-like, un elemento ACGT ed il prolamin box prossimale.

È possibile ipotizzare che il frammento di 430 pb (AF7-AR4), con i suoi motivi (GCN4, prolamin box, ACGT), possa conferire una significativa espressione nell'endosperma. Ciò potrebbe anche non succedere se la mancanza del motivo AACA non fosse compensata dalla presenza, accanto al motivo GCN4, del prolamin box e del motivo ACGT. In questo caso, infatti, il minimo frammento funzionale potrebbe essere quello di 560 pb (AF6-AR4), in quanto contiene tre copie del motivo GCN4, ammesso, ovviamente, che tutti i tre motivi siano funzionali. Questo secondo frammento potrebbe anche avere una forza maggiore, rispetto a quello di 430 pb, per la presenza dell'elemento Skn-1. Un aumento di

forza del promotore potrebbe inoltre derivare dalla presenza, nei restanti due frammenti, di un ulteriore motivo GCN4-like e del prolamin box distale. È tuttavia possibile che solo il motivo di dimensioni maggiori (F4-R4) sia funzionale, in quanto questo contiene l'unico motivo AACA identificato mediante il programma PlantCARE. Questo motivo risulta più conservato nella sequenza regolatrice del gene della PDI localizzato sul cromosoma 4B, mentre nei cromosomi 4A e 4D presenta alcune sostituzioni di basi. Anche altri motivi presentano differenze nucleotidiche tra le regioni regolatrici dei tre geni omeologhi, alcuni sono invece perfettamente conservati. Considerando però che i tre promotori non differiscono significativamente nella regolazione della trascrizione a livello delle cariossidi in via di sviluppo, è possibile ipotizzare che le mutazioni del motivo AACA non ne riducano la funzionalità o che questo motivo non sia essenziale in questo promotore. Lo stesso vale per i motivi GCN4 ripetuti. Ad oggi queste sono solo congetture e, vista la complessità di un'analisi di questo genere, l'individuazione di una regione minima funzionalmente attiva per l'elevata espressione nelle cariossidi in via di sviluppo permetterà una prima valutazione delle ipotesi fatte. Successive analisi funzionali potranno essere effettuate sulla regione minima, provocando mutazioni di basi mirate a compromettere la funzionalità dei singoli motivi individuati. Sarà anche interessante, una volta identificati i motivi essenziali, verificare eventuali differenze presenti a livello di sequenza nucleotidica tra le tre sequenze regolatrici della PDI. Questo permetterebbe di individuare, infatti, eventuali variazioni dei motivi responsabili del diverso profilo di espressione della PDI in tessuti quali spighe, foglie e radici. In tutto questo è importante considerare la possibilità che ci siano ulteriori motivi regolatori, non ancora identificati, responsabili dell'espressione endosperma specifica.

5.10. Co-trasformazione della cultivar MPB-Bobwhite 26 per la sovra-espressione di *TaPDIL1-1* e per l'analisi funzionale del suo promotore

Per la sovra-espressione della PDI il plasmide p1Dx5-5'/*TaPDIL**1**-1* è stato usato in esperimenti di co-trasformazione biolistica con il plasmide PAHC25 (Christensen e Quail, 1996), contente il gene *bar* per la resistenza all'erbicida BASTA (DL-phosphinothricin PPT) ed il gene reporter *UidA*, entrambi sotto il controllo del promotore del gene di mais per l'ubiquitina. I 4 differenti plasmidi ottenuti per l'analisi del promotore**,** pPDI**1278**-GUS, pPDI**901**-GUS, pPDI**561**-GUS e pPDI**430**-GUS, sono stati usati in esperimenti di co-trasformazione biolistica con il plasmide pAHC20 (Christensen e Quail, 1996), contenente il

gene *bar* per la resistenza all'erbicida BASTA (DL-phosphinothricin PPT) sotto il controllo del promotore del gene di mais per l'ubiquitina.

Per la sovra-espressione sono stati bombardati complessivamente 1639 embrioni immaturi, mentre per lo studio del promotore ne sono stati trasformati complessivamente 4257, in particolare 1252 per il plasmide pPDI**1278**-GUS, 1141 per pPDI**901**-GUS, 1071 per pPDI**561**-GUS e 793 per pPDI**430**-GUS (Tab. 5.2).

Tab. 5.2 – Per ogni plasmide sono riportati il numero totale di embrioni bombardati, esclusi quelli usati per il controllo, il numero totale di calli embriogenetici ottenuti, il rapporto tra i calli differenziatisi ed il numero di embrioni bombardati ed il numero di piante ottenute dalla selezione in vitro e trasferite in serra.

Plasmide	N° embrioni	Calli embriogenetici	_Calli/embrioni	Piante
p1Dx5-5'/*TaPDIL1-1*	1639	1208	0,74	951
pPDI**1278**-GUS	1252	986	0,79	336
pPDI**901**-GUS	1141	969	0,85	351
pPDI**561**-GUS	1071	860	0,80	344
pPDI**430**-GUS	793	646	0,81	126

La linea di frumento tenero MPB-Bobwhite 26, definita "super-transformable", è caratterizzata da un'elevata capacità di differenziare calli embriogenetici a partire da scutelli di embrioni immaturi (72%±13), da un'elevata attitudine rigenerativa e da un'accentuata sensibilità al PPT (Pellegrineschi et al., 2002). I risultati relativi all'induzione dell'embriogenesi somatica ottenuti per questi esperimenti (Tab. 5.3 e 5.4) sono in accordo con quelli ottenuti da Pellegrineschi e collaboratori (2002). Nella maggior parte degli esperimenti di trasformazione, infatti, più del 70% di embrioni ha prodotto calli embriogenetici. Per la sovra-espressione, come si può vedere in tabella 5.3, solo in 3 dei 7 esperimenti meno del 70% di embrioni ha prodotto calli embriogenetici, mentre nell'analisi del promotore (Tab. 5.4) su 10 esperimenti solo in due la resa di calli embriogenetici è stata inferiore al 70%. Un caso particolare è quello dell'esperimento numero 7, relativo alla sovra-espressione della *TaPDIL**1**-1* (Tab. 5.3) in cui solo il 59% degli embrioni ha prodotto calli embriogenetici. Infatti non solo l'efficienza di induzione di calli è stata bassa, ma solo uno di questi 248 calli ha dato origine, durante la rigenerazione e selezione, ad alcuni espianti verdi, i quali, peraltro, non hanno superato il periodo di selezione. La ridotta formazione di calli embriogenetici e la loro mancata rigenerazione nell'esperimento numero 7 potrebbe essere dovuta al trattamento antiparassitario effettuato, contro gli afidi, 2 giorni prima della raccolta delle spighe e dell'estrazione degli embrioni immaturi. Nei restanti esperimenti, durante il periodo di selezione, un numero soddisfacente di calli ha differenziato piantine (Fig. 5.22), di cui solo una parte ha superato però l'intera fase di selezione.

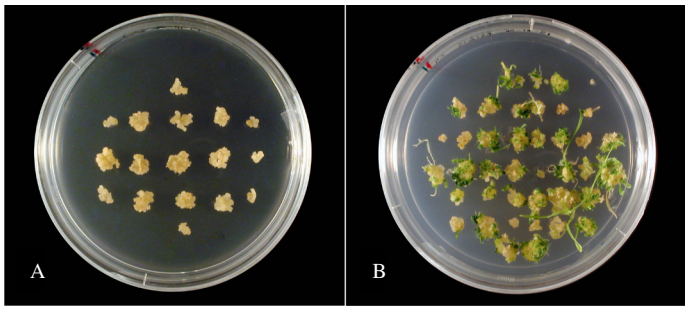

Fig. 5.22 – Selezione in vitro. (A) Controllo: calli derivanti da embrioni non bombardati; (B) Esperimento: calli derivanti da embrioni trasformati, durante la fase di selezione su terreno di coltura MSR contenente 5 mg/L di PPT. Come appare chiaramente dalla figura, i calli derivanti dagli embrioni trasformati con pAHC25 presentano una buona percentuale di rigenerazione, mentre i calli di controllo sono morti.

Alla fine della selezione sono state trasferite nel terreno 951 piante per la sovra-espressione ed un totale di 1157 piante per lo studio del promotore. La mortalità delle piantine in seguito al trapianto si è aggirata intorno a 8-10,8%. Particolarmente alta risulta la percentuale del 19% di piante morte relative al costrutto pPDI**430**-GUS, probabilmente si tratta di un effetto casuale, dovuto al numero minore di trasformazioni complessive fatte per questo costrutto. Ci sono infatti casi particolari in cui i calli di una particolare piastra Petri producono piante non ben sviluppate, la cui probabilità di superare lo stress da trapianto è bassa. L'incidenza complessiva di questi casi particolari diminuisce all'aumentare del numero di esperimenti indipendenti effettuati. Nel caso particolare del costrutto pPDI**430**-GUS, infatti, 10 delle 24 piante morte derivavano da una stessa piastra Petri, che da sola aveva prodotto 23 piante. Escludendo questa piastra Petri dal conteggio complessivo, la mortalità delle piantine dopo il trapianto, anche per questo costrutto, scende al 11,4%. Questo permette di escludere la possibilità remota che questo particolare costrutto abbia un effetto letale nelle piante che lo esprimono.

Tab. 5.3 – Per ciascuno dei 7 esperimenti indipendenti effettuati per la sovra-espressione sono riportati il numero di embrioni bombardati, esclusi quelli destinati al controllo, i rispettivi calli embriogenetici sviluppatisi e passati a rigenerazione ed il rapporto trai due.

N° esperimento	N° embrioni bombardati (E)	N° calli embriogenetici (C)	C/E
1	161	99	0,61
2	206	175	0,85
3	193	130	0,67
4	160	127	0,79
5	205	191	0,93
6	292	238	0,82
7	422	248	0,59
totale	**1639**	**1208**	**0,74**

Tab. 5.4 – Per ciascuno dei 10 esperimenti indipendenti effettuati per lo studio del promotore sono riportati il numero di embrioni bombardati, escluso quelli destinati al controllo, i rispettivi calli embriogenetici sviluppatisi e passati a rigenerazione ed il rapporto tra i due. In ogni esperimento è stata bombardata una percentuale variabile di scutelli per ciascuno dei 4 differenti costrutti.

N° esperimento	N° embrioni bombardati (E)	N° calli embriogenetici (C)	C/E
1	320	147	0,46
2	287	235	0,82
3	455	266	0,58
4	523	410	0,78
5	202	188	0,93
6	502	434	0,86
7	471	437	0,93
8	546	459	0,84
9	435	392	0,90
10	516	493	0,96
totale	**4257**	**3461**	**0,81**

Come appare evidente dalla tabella 5.2, dagli esperimenti di sovra-espressione è stato ottenuto un numero maggiore di piantine rispetto a quelli relativi allo studio del promotore. La particolare abbondanza delle piante prodotte per la sovra-espressione è dovuta al fatto che i primi esperimenti sono stati trasferiti ad un nuovo terreno di crescita privo di PPT quando i calli di controllo erano completamente ed irreversibilmente imbruniti, sottoponendoli, quindi, ad un massimo di 7 settimane di selezione. In seguito, per evitare falsi positivi, ossia piante sfuggite alla selezione pur non contenendo il transgene per la resistenza, i restanti esperimenti sono stati sottoposti ad un ulteriore periodo di selezione, della durata di 2-3 settimane. Questo ha ridotto moltissimo il numero complessivo di piante ottenute ma, contemporaneamente, ha aumentato la probabilità di ottenere un'elevata percentuale di piante trasformate. Prima del trapianto in terra, le piantine sono stata poste in un terreno di coltura senza PPT per permettere la loro ulteriore crescita. La scelta di lasciare le piantine nella fase finale in un terreno privo di agente selettivo è stata fatta allo scopo di ottenere piante sufficientemente forti da garantire un'elevata percentuale di sopravvivenza durante il trapianto in serra, visto che, per motivi tecnici, non è stato possibile effettuare un periodo di acclimatamento in condizioni controllate.

Uno dei vantaggi legati all'uso della linea di frumento tenero MPB-Bobwhite 26 è la sua elevata sensibilità al PPT (Pellegrineschi et al., 2002). Questa caratteristica, assieme all'elevata dose di erbicida usata (5 mg/L), letale sia per i calli in rigenerazione che per i tessuti organizzati, sono un efficiente strumento di selezione. Nonostante ciò, ci sono stati alcuni rari esperimenti in cui uno o due calli del controllo hanno differenziato alcuni stentati abbozzi di germogli, che sono rimasti verdi a lungo. Alla fine del periodo di selezione,

comunque, tutti i calli di controllo e gli eventuali germogli formatisi da essi erano morti. Non si può escludere, tuttavia, la presenza di alcuni falsi positivi tra le piante portate a maturazione in serra, in particolare per i primi esperimenti, nei quali la selezione è stata più breve.

La maggior parte delle piante allevate in serra sono risultate almeno parzialmente fertili, mentre solo alcune di esse erano completamente sterili.

5.11. Identificazione mediante PCR delle piante contenenti i costrutti per la sovra-espressione della PDI

Sono state analizzate tramite PCR 384 delle 951 piante rigenerate dopo bombardamento con il costrutto recante la sequenza codificante per la PDI per ottenere la sua sovra-espressione. Tali piante sono una parte di quelle ottenute nei primi tre esperimenti.

Considerando che sia il promotore sia il gene del costrutto per la sovra-espressione sono endogeni, cioè sequenze già presenti nella pianta di frumento, sono state disegnati un primer senso nella regione 3' del gene della PDI classica (PDI-F: 5'-TACTTCGTCACTCCTAGCG-3') ed un primer antisenso nella regione 3' del terminatore NOS (NOS-R: 5'-TTATCCTAGTTTGCGCGCTA-3').

Tra le 384 piante analizzate 5 sono risultate positive, infatti l'amplificazione con i primer precedentemente indicati ha prodotto un frammento atteso di 570 pb (Fig. 5.23). In particolare una (27741) delle 5 piante è stata rigenerata dai calli ottenuti dai 161 embrioni bombardati nel primo esperimento, 2 (27922, 28022) dai 206 embrioni del secondo e 2 (28305, 28365) dai 193 embrioni del terzo.

Le piante derivanti dai primi due esperimenti di trasformazione sono state analizzate anche con una seconda coppia di primer, disegnata per amplificare la sequenza tra il promotore e la regione 5' del gene. Il primer senso è stato disegnato nella regione 3' del promotore della 1Dx5 (1Dx5-F: 5'-TCCAATTGCTCCTTGCTTAT-3') e il primer antisenso nella regione 5' del gene della PDI classica (PDI-R: 5'-CCTGAAGATCTTGAGGGTAGG-3'). Con questa seconda coppia di primer sono risultate positive 2 piante rigenerate nel primo esperimento e 7 nel secondo (Fig. 5.24 e 5.25). Le 3 piante dei primi due esperimenti risultate positive nell'amplificazione con la coppia di primer PDI-F e NOS-R sono risultate positive anche nell'amplificazione con la coppia di primer 1Dx5-F e PDI-R, probabilmente, quindi, contengono l'intero costrutto.

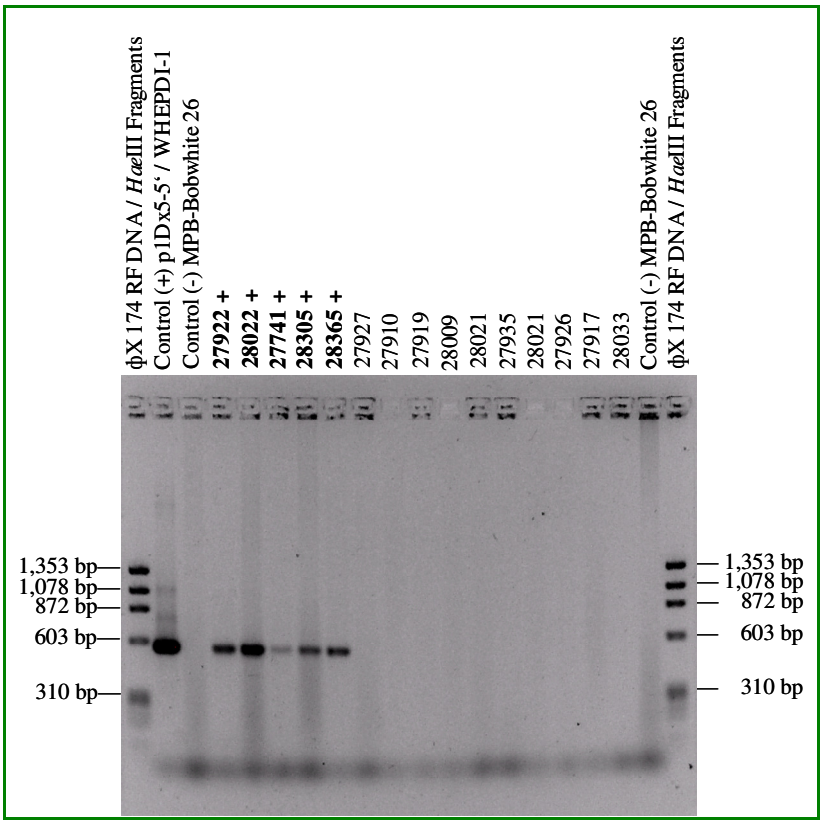

Fig. 5.23 - Prodotti di amplificazione di 15 piante T_0 per la sovra-espressione e 2 piante di MPB-Bobwhite 26 (10 ng) non trasformate come controllo negativo. Come controllo positivo è stato usato il plasmide p1Dx5-5'/*TaPDIL1-1* (20 ng). La coppia di primer usata (PDI-F e NOS-R) amplifica un frammento di 570 pb.

L'analisi PCR condotta con la coppia di primer 1Dx5-F e PDI-R amplifica una regione del gene della *TaPDIL1-1* con un elevato contenuto di G+C che potrebbe determinare la formazione di probabili loop che non permettono di ottenere il frammento di dimensioni attese (Fig. 5.24).

Risulta difficile, quindi, interpretare i risultati dell'amplificazione e, in particolare, verificare se i prodotti ottenuti siano artefatti della PCR o bande reali, derivanti dall'inserimento del costrutto chimerico incompleto in seguito alla trasformazione. Per ovviare a questo problema sono state adottate condizioni di amplificazione PCR particolari, utilizzando la "Platinum *Taq* DNA Polymerase" della GIBCO insieme ad un cosolvente che facilita l'efficiente amplificazione delle sequenze ricche in G e C, chiamato "Enhancer Solution", della stessa ditta. Queste analisi hanno confermato che per 5 delle 9 piante il frammento amplificato è delle dimensioni attese (673 pb), mentre per le restanti 4 piante sono necessarie ulteriori prove, in quanto non è stato possibile rilevare né la banda attesa di 673 pb, né quelle di dimensioni minori ottenute senza l'utilizzo delle particolari condizioni di PCR (Fig. 5.25).

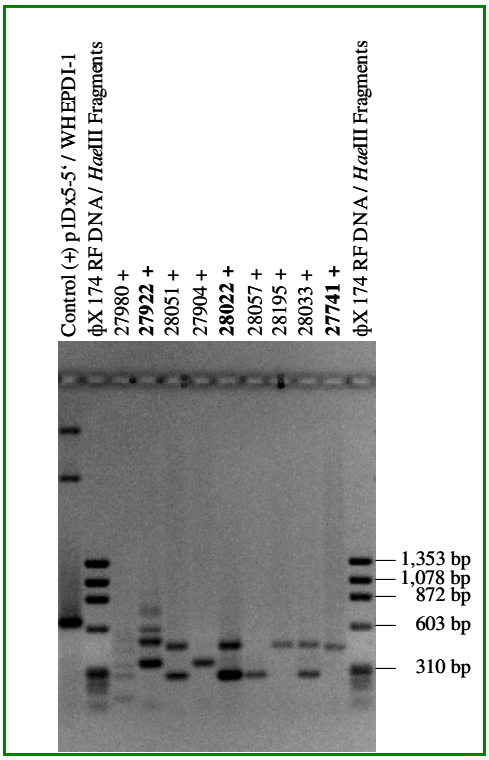

Fig. 5.24 - Prodotti di amplificazione di 9 piante T_0 per la sovra-espressione risultate positive all'amplificazione con la coppia di primer 1Dx5-F e PDI-R. Il frammento ottenuto nel controllo (20 ng del plasmide p1Dx5-5'/*TaPDIL1-1*) è di 673 pb, mentre nelle 9 piante transgeniche il frammento o i frammenti ottenuti sono di dimensione inferiore.

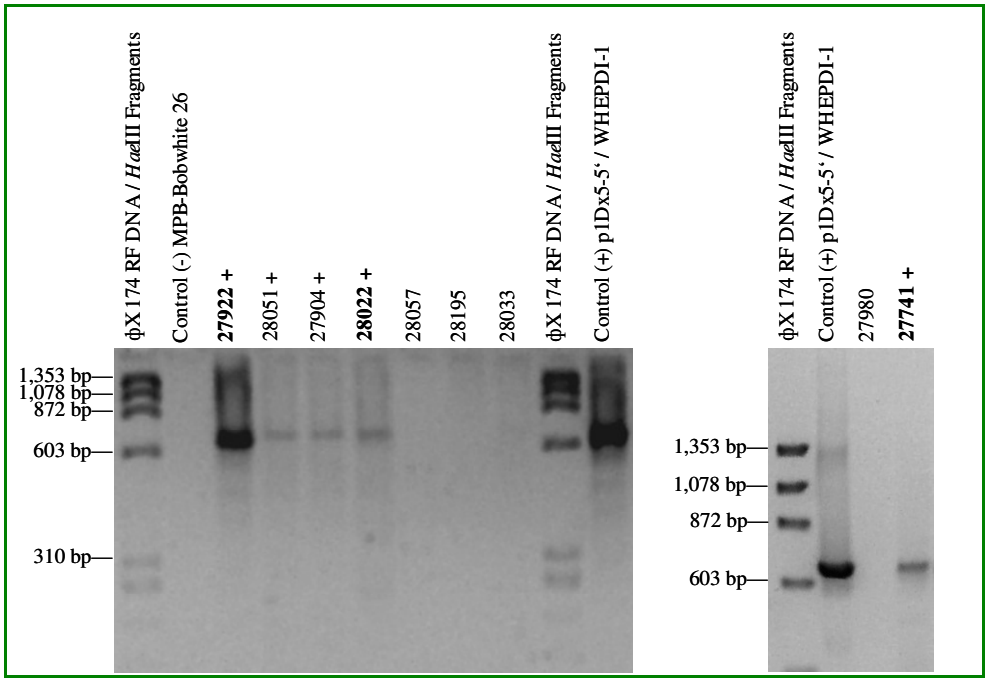

Fig. 5.25 - Prodotti di amplificazione delle 7 piante transgeniche ottenute nel secondo esperimento e delle 2 piante derivate dal primo esperimento. Per 5 delle 9 piante T_0 per la sovra-espressione il frammento ottenuto dall'amplificazione con l'aiuto dell'enhancer è delle dimensioni attese di 673 pb, confermando che i risultati ottenuti senza queste particolari condizioni PCR sono dovuti ad artefatti legati alla formazione di loop nella zona ricca in G e C della sequenza della *TaPDIL1-1*. Per le restanti 4 piante non è stato ottenuto alcun frammento. Come controllo negativo è stato usato il DNA genomico di una pianta di MPB-Bobwhite 26 (10 ng) non trasformato, come controllo positivo è stato usato il DNA del plasmide p1Dx5-5'/*TaPDIL1-1* (20 ng). La coppia di primer usata è 1Dx5-F e PDI-R.

6. CONCLUSIONI

Le proteine appartenenti alla famiglia PDI, comprendente la PDI classica e le PDI-like, rivestono un ruolo fondamentale nel metabolismo cellulare degli organismi animali e vegetali. Esse svolgono numerose funzioni, tra le quali è molto importante, tanto da conferire il nome all'intera famiglia genica, l'intervento nella formazione ed isomerizzazione dei legami disolfuro e nel folding delle proteine. Il recente completamento delle sequenze nucleotidiche dei genomi di *Arabidopsis* e di riso ha permesso di ottenere la prima descrizione completa della complessità e diversità di questa famiglia genica nelle piante, ma, nonostante questi studi approfonditi, poco o nulla si conosce sul ruolo fisiologico dei singoli membri ad essa appartenenti. Nelle piante la forma classica della PDI, sulla quale sono stati condotti studi di caratterizzazione molecolare, regolazione trascrizionale e localizzazione intracellulare, è la meglio caratterizzata. I risultati di alcuni di questi studi, condotti in mais e riso, suggeriscono che anche nelle piante, come nei vertebrati, la PDI classica possa giocare un ruolo importante nel folding delle proteine di secrezione ed in particolare nella formazione dei legami disolfuro e nell'assemblaggio delle proteine di riserva. Questo aspetto riveste un significato pratico di notevole importanza in frumento, se si considera che la qualità delle farine e delle semole dei frumenti coltivati dipende prevalentemente dalla composizione proteica del glutine e dalle interazioni chimiche che si instaurano tra le diverse componenti. La possibilità di formare legami disolfuro inter- ed intra-molecolari è considerato uno dei principali fattori responsabili delle peculiari caratteristiche fisico-chimiche e tecnologiche del glutine. Queste interazioni dipendono non solo dalla composizione aminoacidica delle proteine di riserva, ma anche da processi biochimici che avvengono nelle cellule dell'endosperma delle cariossidi durante la granigione. Una migliore comprensione dei meccanismi attraverso i quali si esplicano la sintesi, la maturazione e soprattutto l'assemblaggio di alcune proteine di riserva in grandi aggregati durante la granigione aprirebbe nuove prospettive di intervento per la manipolazione delle proprietà funzionali del glutine, mediante tecniche di ingegneria genetica. Lo studio della famiglia genica codificante per proteine PDI-like in frumento è motivato, quindi, non solo da un interesse teorico per i molteplici ruoli che le proteine PDI-like potrebbero svolgere nel metabolismo cellulare e per comprendere l'evoluzione molecolare di questa famiglia multigenica in un contesto poliploide, ma anche da un aspetto meramente applicativo, visto il loro potenziale coinvolgimento nella sintesi e deposizione delle proteine di riserva e quindi nella determinazione delle proprietà reologiche del glutine.

Il primo obiettivo del lavoro effettuato durante lo svolgimento della tesi di dottorato è stato lo studio della complessità e diversità della famiglia multigenica PDI, comprendente proteine PDI e PDI-like, in frumento tenero. L'iniziale ricerca effettuata nella banca dati "TIGR wheat gene index", utilizzando 11 sequenze di riso, assegnate alle 8 classi filogenetiche della famiglia PDI nelle piante individuate da Houston e collaboratori (2005), ha permesso di individuare in frumento 10 distinte sequenze consenso (TC, Tentative Consensus) codificanti per altrettanti geni PDI-like. Questi risultati sono stati confermati anche dalla ricerca di sequenze EST omologhe a quelle relative agli 11 geni di riso in un ulteriore banca dati, "HarvEST Wheat", che è in diretta connessione con la divisione EST delle banche dati di GeneBank, EMBL, DDBJ e NCBI, e dalla successiva analisi di 70 delle sequenze EST identificate, che sono state amplificate mediante RT-PCR e clonate. Una di queste dieci sequenze consenso, la TC250792, corrisponde alla sequenza codificante per la PDI classica, precedentemente clonata e caratterizzata (Ciaffi et al., 2006). Sono state quindi clonate le sequenze cDNA complete relative alle restanti 9 sequenze consenso relative a geni codificanti per proteine PDI-like. L'analisi filogenetica si è basata sul confronto delle sequenze aminoacidiche dedotte da 52 sequenze geniche della famiglia PDI, comprendenti quelle di alcune specie vegetali, le 9 sequenze PDI-like clonate durante questa ricerca e la sequenza codificante per la PDI classica di frumento (Ciaffi et al., 2006), appartenente al primo gruppo filogenetico. L'analisi ha permesso di assegnare le nove sequenze PDI-like clonate a sette delle otto classi filogenetiche della famiglia genica PDI identificate nelle piante. Grazie alla ricerca incrociata in diverse banche dati EST, è verosimile che la maggior parte dei membri della famiglia PDI di frumento siano stati identificati. È possibile che non siano state individuate le sequenze cDNA di altri geni paraloghi codificanti per proteine PDI-like in frumento, sono state clonate, tuttavia, sequenze relative ad almeno un membro per ciascuna delle otto classi filogenetiche individuate da Houston e collaboratori (2005) in base alle sequenze presenti in *Arabidopsis*, riso e mais. I risultati ottenuti in frumento avvalorano l'ipotesi che gli otto gruppi filogenetici abbiano avuto origine da altrettanti geni presenti nella specie ancestrale da cui si sono evolute le dicotiledoni e le monocotiledoni. L'analisi filogenetica ha inoltre evidenziato la presenza di geni paraloghi derivanti da duplicazioni geniche che hanno avuto luogo, in maniera indipendente, nelle singole specie analizzate, *Arabidopsis*, mais, riso e frumento. L'analisi filogenetica condotta in questa ricerca, a differenza del lavoro svolto da Houston e collaboratori (2005) in cui venivano considerati i singoli domini tioredossinici, si è basata sulle sequenze aminoacidiche dedotte relative alle intere sequenze codificanti. La principale conseguenza derivante da questa scelta è che

nell'albero filogenetico i membri dell'ottavo gruppo sono risultati nettamente diversificati da quelli delle altre sottofamiglie di proteine PDI-like, perciò le sequenze di questo gruppo sono state inserite come outgroup. Escludendo l'ottava classe filogenetica, che diverge nettamente dalle altre, è possibile dividere la famiglia delle PDI in due cladi principali. La prima e più ampia di esse comprende tutte le proteine appartenenti alla prima classe strutturale (due domini tioredossinici attivi situati uno nella regione N-terminale e l'altro in quella C-terminale) individuate nelle piante (gruppi filogenetici I, II e III) e alcune proteine della quinta classe strutturale (gruppo filogenetico VII), contenenti un solo dominio attivo. Per i gruppi filogenetici I, II e III è possibile ipotizzare una loro origine da eventi di duplicazione genica, che abbiano interessato un gene ancestrale codificante per una proteina di classe strutturale 1, avvenuti nella specie progenitrice delle monocotiledoni e delle dicotiledoni. Il gruppo filogenetico VII potrebbe essersi originato, invece, da un analogo evento di duplicazione, seguito dalla perdita di uno dei due domini tioredossinici attivi. I risultati relativi alla prima clade sono in accordo con quelli ottenuti da Houston e collaboratori (2005). L'albero filogenetico ottenuto non permette, tuttavia, di trarre conclusioni altrettanto lineari sull'origine e sull'evoluzione dei membri della seconda clade principale, le cui sequenze risultano molto più eterogenee.

Le sequenze aminoacidiche dedotte di geni appartenenti alla stessa sottofamiglia presentano un elevato livello di conservazione, le sequenze di frumento e di riso, in particolare, hanno un'identità compresa tra 69,3% a 91,9%. L'analisi della struttura esoni/introni, effettuata per le sequenze genomiche di tre geni PDI-like di frumento (*TaPDIL**2**-1*, *TaPDIL**4**-1* e *TaPDIL**5**-1*) e per i loro ortologhi di riso e *Arabidopsis*, hanno dimostrato che anche l'organizzazione genomica, a livello di giunzioni esone/introne, lunghezza degli esoni e posizione dei siti attivi, risulta altamente conservata fra i membri di una stessa sottofamiglia. L'elevata conservazione dell'organizzazione genomica riflette, verosimilmente, l'importanza funzionale dei prodotti proteici codificati. L'analisi Southern di CS e delle sue linee aneuploidi utilizzando come sonde sequenze specifiche per i cloni *TaPDIL**4**-1*, appartenente al IV gruppo filogenetico, e *TaPDIL**5**-1*, del V gruppo, ha permesso di effettuare la localizzazione cromosomica delle sequenze genomiche corrispondenti. Il clone *TaPDIL**4**-1* ha rilevato la presenza di tre geni omeologhi localizzati sul braccio corto dei tre cromosomi del gruppo 1; i tre geni omeologhi individuati dal clone *TaPDIL**5**-1* sono localizzati, invece, sul braccio lungo dei tre cromosomi del gruppo 5. Questi risultati sono coerenti con la natura alloesaploide del frumento tenero ed è presumibile che anche per ciascuno degli altri cloni PDI-like ottenuti siano presenti tre copie di geni omeologhi, conferiti

dalle specie donatrici dei tre genomi. L'analisi di espressione genica effettuata per *TaPDIL2-1*, *TaPDIL4-1* e *TaPDIL5-1* ed il confronto con quella relativa alla PDI classica ha permesso di verificare le differenze nei livelli di trascrizione dei singoli geni nei diversi tessuti analizzati. Questi dati non permettono di desumere le funzioni svolte dai relativi prodotti genici. I profili di espressione suggeriscono che tra i geni analizzati potrebbe essere presente una certa ridondanza funzionale, associata, tuttavia, ad una diversificazione nei livelli e nei tempi di espressione.

Un primo passo verso la caratterizzazione funzionale dei membri di questa grande famiglia genica ha riguardato la produzione di linee transgeniche di frumento nelle quali la PDI classica viene sovra-espressa sotto il controllo di un forte promotore endosperma specifico. Durante lo svolgimento della tesi di dottorato sono state ottenute numerose piante potenzialmente transgeniche, in alcune delle quali la presenza del transgene è stata confermata tramite analisi PCR. E' stato intrapreso, inoltre, lo studio funzionale del promotore di uno dei tre geni omeologhi della PDI classica. Per caratterizzare meglio il ruolo funzionale dei motivi regolatori presenti in questo promotore, il gene reporter *UidA* (GUS) è stato posto sotto il controllo di 4 frammenti di delezione progressivi in 5'. I quattro costrutti così ottenuti sono stati introdotti in altrettante linee di frumento. Dai diversi esperimenti di trasformazione biolistica sono state prodotte numerose piante di frumento potenzialmente transgeniche, che dovranno essere analizzate per rilevare l'espressione del gene *UidA*. L'analisi funzionale del promotore potrebbe essere interessante per un'analisi approfondita del profilo di espressione spaziale (tessuto specifica) e temporale di uno dei geni omeologhi codificanti per la PDI classica. L'interesse di questa ricerca deriva anche dalla possibilità di disporre di una sequenza regolatrice, caratterizzata funzionalmente, che potrebbe essere utilizzata per l'espressione tessuto specifica di altri geni di interesse, arricchendo la disponibilità di promotori di questo tipo, attualmente piuttosto limitata.

7. PROSPETTIVE FUTURE

Sarà interessante, in futuro, caratterizzare meglio i profili di espressione relativi ai rimanenti 6 membri di questa famiglia genica per evidenziare eventuali differenze nella regolazione della trascrizione. Sarà importante definire, inoltre, la loro localizzazione cromosomica, anche per evidenziare la presenza di eventuali geni paraloghi non ancora identificati. L'individuazione di regioni altamente conservate tra le diverse sequenze geniche o tra gruppi di esse permetterebbe di ottenere il loro silenziamento mediante la tecnica dell'interferenza dell'RNA (RNAi), superando il problema derivante dalla ridondanza funzionale. Questo approccio consentirebbe l'analisi funzionale dei geni codificanti le proteine PDI e PDI-like e di chiarire il loro ruolo nella deposizione delle proteine di riserva e nella formazione degli aggregati proteici.

Per quanto riguarda gli esperimenti relativi alla sovra-espressione della PDI classica e all'analisi funzionale del suo promotore, la maggior parte delle piantine rigenerate dagli embrioni bombardati devono essere ancora analizzate mediante la tecnica PCR per verificare l'avvenuta integrazione del costrutto di interesse e del gene *bar* che conferisce la resistenza alla fosfinotricina. Questa ultima informazione rivelerà il numero di falsi positivi e l'effettiva efficienza dell'intero processo di trasformazione, in particolare della selezione con l'erbicida, permettendo di valutare l'efficienza di co-trasformazione. Una variazione significativa rispetto all'efficienza di cotrasformazione attesa del 25-30% (Pellegrineschi et al., 2001), per un particolare costrutto, potrebbe indicare un suo potenziale effetto deleterio sul metabolismo e quindi sulla sopravvivenza della pianta. Sarà importante, infine, quantificare il numero reale di falsi positivi, ossia delle piante sopravvissute alla selezione, ma non contenenti il gene di resistenza *bar*, per un dosaggio più accurato dell'erbicida, ai fini della selezione, nei futuri esperimenti di trasformazione biolistica. Anche le piante relative allo studio del promotore, rigenerate dagli embrioni bombardati, saranno analizzate mediante PCR con primer specifici per il gene reporter *UidA* e per il gene di resistenza *bar*.

La progenie delle piante T_0 che saranno risultate positive dalle analisi PCR verranno moltiplicate ed analizzate mediante PCR per verificare l'integrazione stabile del relativo costrutto transgenico di interesse. Analisi Southern permetteranno di verificare il numero di copie inserite e la presenza dell'intero costrutto.

Le linee di frumento contenenti una o più copie del transgene saranno autofecondate e nelle progenie saranno identificate le piante omozigoti. Mediante RT-PCR, si verificherà l'espressione dei trascritti del transgene nelle cariossidi a diversi stadi di sviluppo ed il

confronto con piante della linea MPB-Bobwhite 26 non trasformate permetterà di verificare sia la sovra-espressione della PDI classica che l'espressione del gene reporter. Analisi western, infine, rileveranno la presenza del prodotto genico.

Per la sovra-espressione della PDI sarà interessante verificare l'effetto sulla deposizione delle proteine di riserva e sugli aggregati proteici e, quindi, sulla qualità delle farine. Per lo studio del promotore sulle piante con una o poche copie del transgene saranno eseguite anche analisi istochimiche per evidenziare la presenza di GUS nei differenti tessuti. Queste analisi permetteranno di definire la regolazione spaziale e temporale determinata dall'attività del promotore e di valutare, mediante il confronto del livello di espressione determinato dai diversi frammenti di delezione, il ruolo dei motivi regolatori identificati.

8. BIBILIOGRAFIA

Alanen HI, Williamson RA, Howard MJ, Lappi AK, Jäntti HP, Rautio SM, Kellokumpu S, Ruddock LW (2003) J Biol Chem 278: 28912-28920.

Albani D, Hammond-Kosack MCU, Smith C, Conlan S, Colot V, Holdsworth M, Bevan MW (1997) Plant Cell 9: 171-184.

Altieri F, Maras B, Eufemi M, Ferraro A, Turano C (1993) Biochem Biophys Res Commun 194: 992-1000.

Altschuler Y, Rosenberg N, Harel R, Galili G (1993) Plant Cell 5: 443-450.

Altschuler Y, Galili G (1994) J Biol Chem 269: 6677-6682.

Anelli T, Alessio M, Mezghrani A, Simmen T, Talamo F, Bachi A Sitia R (2002) EMBO J 21: 835-844.

Anelli T, Alessio M, Bachi A, Bergamelli L, Bertoli G, Camerini S, Mezghrani A, Ruffato E, Simmen T, Sitia R (2003) EMBO J 22: 5015-5022.

Apweiler R, Attwood TK, Bairoch A, Bateman A, Birney E, Biswas M, Bucher P, Cerutti L, Corpet F, Croning MD, et al, (2001) Nucleic Acids Res 29: 37-40.

Argos P, Pedersen K, Marks MD, Larkins BA (1982) J Biol Chem 257: 9984-9990.

Bader M, Muse W, Ballou DP, Gassner C, Bardwell JC (1999) Cell 98: 217-227.

Bader MW, Xie T, Yu CA, Bardwell JC (2000) J Biol Chem 275: 26082-26088.

Bader MW, Hiniker A, Regeimbal J, Goldstone D, Haebel PW, Riemer J, Metcalf P, Bardwell JC (2001) EMBO J 20: 1555-1562.

Bagga S, Adams HP, Rodriguez FD, Kemp JD, Sengupta-Gopalan C (1997) Plant Cell 9: 1683-1696.

Baksh S, Burns K, Andrin C, Michalak M (1995) J Biol Chem 270: 31338-31344.

Bass R, Ruddock LW, Klappa P, Freedman RB (2004) J Biol Chem 279: 5257-5262.

Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, Sonnhammer ELL (2000) Nucleic Acids Res 28: 263-266.

Bechtel DB, Juliano BO (1980) Ann Bot 45:503-509.

Bechtel DB, Wilson JD, Shewry PR (1991) Cereal Chem 68: 573-577.

Belton PS (1999) J Cereal Sci 29: 103-107.

Benham AM, Cabibbo A, Fassio A, Bulleid N, Sitia R, Braakman I (2000) EMBO J 19: 4493-4502.

Bennett CF, Balcarek JM, Varricchio A, Crooke ST (1988) Nature 334: 268-270.

Bennett TA, Edwards BS, Sklar LA, Rogelj S (2000) J Immunol 164: 4120-4129.

Berezney R (1991) J Cell Biochem 47: 109-123.

Bjelland S (1987) Comp Biochem Physiol B 87: 907-914.

Bohring C, Krause E, Habermann B, Krause W (2001) Mol Hum Reprod 7: 113-118.

Boston RS, Fontes EBP, Shank BB, Wrobel RL (1991) Plant Cell 3: 497-505.

Breuza L, Halbeisen R, Jenö P, Otte S, Barlowe C, Hong W, Hauri HP (2004) J Biol Chem 279(45): 47242-47253.

Brodsky JL, Werner ED, Dubas ED, Goeckeler JL, Kruse KB, McCracken AA (1999) J Biol Chem 274: 3453-3460.

Bulleid NJ, Freedman RB (1988a) Nature 335: 649-651.

Bulleid NJ, Freedman RB (1988b) Biochem J 254:805-810.

Bulleid NJ, Shewry PR, Freedman RB (1992) In Shewry PR, Gutteridge S, eds Plant Protein Engineering. Cambridge University Press, Cambridge, UK: 201-208.

Burgess JK, Hotchkiss KA, Suter C, Dudman NP, Szollosi J, Chesterman CN, Chong BH, Hogg PJ (2000) J Biol Chem 275: 9758-9766.

Cabibbo A, Pagani M, Fabbri M, Rocchi M, Farmery MR, Bulleid NJ, Sitia R (2000) J Biol Chem 275: 4827-4833.

Cabral CM, Liu Y, Sifers RN (2001)Trends Biochem Sci 26: 619-624.

Cai H, Wang CC, Tsou CL (1994) J Biol Chem 269: 24550-24552.

Chakravarthi S, Bulleid NJ (2004) J Biol Chem 279: 39872-39879.

Chang JY, Li L, Bulychev A (2000) J Biol Chem 275: 8287-8289.

Chaudhuri MM, Tonin PN, Lewis WH, Srinivasan PR (1992) Biochem J 281: 645-650.

Chen F, Hayes PM (1994) Plant Physiol 106: 1705- 1706.

Chen K, Lin Y, Detwiler TC (1992) Blood 79: 2226-2228.

Cheng SY, Gong QH, Parkinson C, Robinson EA, Appella E, Merlino GT Pastan I (1987) J Biol Chem 262: 11221-11227.

Cherry JM, Ball C, Weng S, Juvik G, Schmidt R, Adler C, Dunn B, Dwight S, Riles L, Mortimer RK, Botstein D (1997) Nature 387: 67-73.

Chivers PT, Laboissière MCA, Raines RT (1996) EMBO J 15: 2659-2667.

Chivers PT, Prehoda KE, Raines RT (1997) Biochemistry 36: 4061-6406.

Christensen AH, Sharrock RA, Quail PH (1992) Plant Mol Biol 18: 675-689.

Christensen AH, Quail PH (1996) Transgenic Res 5: 213-218.

Ciaffi M, Tozzi L, Lafiandra D (1996) Cereal Chem 73: 346-351.

Ciaffi M, Dominici L, Tanzarella OA, Porceddu E (1999) Theor Appl Genet 98: 405-410.

Ciaffi M, Paolacci AR, Umana E, Tanzarella OA, Corbellini M, Borghi B, Porceddu E (1998) In: Troccoli A, De Vita P, Di Fonzo N, eds Prospettive di una moderna cerealicoltura alle porte del terzo millennio: 141-165.

Ciaffi M, Dominici L, Umana E, Tanzarella OA, Porceddu E (2000) Theor Appl Genet 101: 220-226.

Ciaffi M, Paolacci AR, Dominici L, Tanzarella OA, Porceddu E (2001) Gene 265: 147-156.

Ciaffi M, Tanzarella OA, Porceddu E (2005) Frontiers of Wheat Bioscience: 61-75.

Ciaffi M, Paolacci AR, d'Aloisio E, Tanzarella OA, Porceddu E (2006) Gene 366: 209-218.

Clissold PM, Bicknell R (2003) BioEssays 25: 603-611.

Coleman CE, Lopes MA, Gillikin JW, Boston RS, Larkins BA (1995) Proc Nat Acad Sci USA 92: 6828-6831.

Coleman CE, Hewrman EH, Takasaki K, Larkins BA (1996) Plant Cell 9: 2335-2345.

Coleman CE, Larkins BA (1999) In: Shawry PR, Casey R, eds Seed proteins. Dordrecht: Kluwer Academic Publisher: 109-139.

Colot V, Bartels D, Thompson R, Flavell R (1989) Mol Gen Genet 216: 81-90.

Coppari S, Altieri F, Ferraro A, Chichiarelli S, Eufemi M, Turano C (2002) J Cell Biochem 85: 325-333.

Coughlan SJ, Hasting C, Winfrey RJ Jr (1996) Eur J Biochem 235: 215-224.

Creighton TE (1979) J Mol Biol 129: 411-434.

Creighton TE, Hillson DA, Freedman RB (1980) J Mol Biol 142: 43-62.

Creighton TE (1991) Science 256: 111-113.

Cunnea PM, Miranda-Vizuete A, Bertoli G, Simmen T, Damdimopoulos AE, Hermann S, Leinonen S, Huikko MP, Gustafsson JA, Sitia R, Spyrou G (2003) J Biol Chem 278: 1059-1066.

Cuozzo JW, Kaiser CA (1999) Nat Cell Biol 1: 130-135.

D'Amico L, Valsasina B, Daminati MG, Fabbrini MS, Nitti G, Bollini R, Ceriotti A, Vitale A (1992) The Plant Journal 2: 443-455.

Dai Y, Wang CC (1997) J Biol Chem 272: 27572-27576.

Darby NJ, Creighton TE (1995a) Biochemistry 34: 3576-3587.

Darby NJ, Creighton TE (1995b) Biochemistry 43: 16770-16780.

Darby NJ, Creighton TE (1995c) Biochemistry 34: 11725-11735.

Darby NJ, Kemming J, Creighton TE (1996) Biochemistry 35: 10517- 10528.

Darby NJ, Penka E, Vincentelli R (1998) J Mol Biol 276: 239-247.

Darby NJ, van Straaten M, Penka E, Vicentelli R, Kemming J (1999) FEBS Letters 448: 167-172.

DeBurman SK, Reymond GJ, Caughey B, Lindquist S (1997) Proc Natl Acad Sci USA 94: 13938-13943.

Deleage G, Blanchet C, Geourjon C (1997) Biochemie 79(11): 681-686.

Delom F, Mallet B, Carayon P, Lejeune PJ (2001) J Biol Chem 276: 21337-21342.

Demmer J, Zhou MC, Hubbard MJ (1997) FEBS lett 402: 145-150.

Denecke J, De Rycke R, Botterman J (1992) EMBO J 11: 2345-2355.

Denecke J (1996) Plant Physiol Biochem 34: 197- 205.

Desilva MG, Notkins AL, Lan MS (1997) DNA Cell Biol 16: 269-274.

Devos KM, Dubcovsky J, Dvorak J, Chinoy CN, Gale MD (1995) Theor Appl Genet 91: 282-288.

Dickinson CD, Evans RP, Nielsen NC (1988) Nucl Acids Res 16: 371.

Digeon JF, Guideroni E, Alary R, Michaux-Ferrière N, Joudrier P, Gautier MF (1999) Plant Mol Biol 39: 1101-1112.

Dobson CM (2001) Biochem Soc Symp 68: 1-26.

Doolittle RF (1995) Ann Rev Biochem 64: 287-314.

Dorner AJ, Wasley LC, Raney P, Haugejorden S, Green M, Kaufman RJ (1990) J Biol Chem 265: 22029-22034.

Dyson HJ, Jeng MF, Tennant LL, Slaby I, Lindell M, Cui DS, Kuprin S, Holmgren A (1997) Biochemistry 36: 2622-2636.

Edman JC, Ellis L, Blacher RW, Roth RA, Rutter WJ (1985) Nature 317: 267-270.

Ellgaard L, Frickel EM (2003) Cell Biochem Biophys 39: 223-247.

Ellgaard L, Ruddock LW (2005) EMBO *Rep* 6: 28-32.

Elliott JG, Oliver JD, High S (1997) J Biol Chem 272:13849-13855.

Fassio A, Sitia R (2002) Histochem Cell Biol 117: 151-157.

Felstein J (1989) Cladist 5:164-166.

Fenouillet E, Barbouche R, Courageot J, Miquelis R (2001) J Infect Dis 183: 744-752.

Ferrari DM, Van Nguyen P, Kratzin HD, Söling HD (1998) Eur J Biochem 255: 570-579.

Ferrari DM, Söling HD (1999) Biochem J 339: 1-10.

Ferraro A, Altieri F, Coppari S, Eufemi M, Chichiarelli S, Turano C (1999) J Cell Biochem 72:528-539.

Field JM, Shewry PR, Burgess SR, Forde J, Parmar S, Miflin BJ (1983a) J cereal Sci 1:33-41.

Field JM, Shewry PR, Miflin BJ (1983b) J Sci Food Agric 34: 370-377.

Frand AR, Kaiser CA (1998) Mol Cell 1: 161-170.

Frand AR, Kaiser CA (1999) Mol Cell 4: 469-477.

Frand AR, Cuozzo JW, Kaiser CA (2000) Trends Cell Biol 10(5): 203-210.

Freedman RB (1984) Trends Biochem Sci 9: 438-441.

Freedman RB (1987) Nature 329: 196-198.

Freedman RB (1989) Cell 57: 1069-1072.

Freedman RB, Hirst TR, Tuite MF (1994) Trends Biochem Sci 19: 331-336.

Freedman RB, Hawkins HC, McLaghlin SH (1995) Methods EnzyMol 251: 397-406.

Freedman RB, Klappa P, Ruddock LW (2002) EMBO *Rep* 3: 136-140.

Frickel EM, Frei P, Bouvier M, Stafford WF, Helenius A, Glockshuber R, Ellgaard L (2004) J Biol Chem 279: 18277-18287.

Füllekrug J, Sönnischen B, Wunsch U, Arseven K, Van Nguyen P, Söling HD, Mieskes G (1994) J Cell Sci 107: 2719-2727.

Fullington JG, Miskelly DM, Wringley CW, Kasarda DD (1987) J Cereal Sci 5: 233.

Galili G, Sengupta- Gopalan C, Ceriotti A (1998) Plant Mol Biol 38: 1-29.

Gatehouse JA, Evans IM, Croy RRD, Boulter D (1986) Philos Trans R Soc B314: 367-384.

Gerber J, Muhlenhoff U, Hofhaus G, Lill R, Lisowsky T (2001) J Biol Chem 276: 23486-23491.

Gerner C, Holzmann K, Meissner M, Gotzmann J, Grimm R, Sauermann G (1999) J Cell Biochem 74: 145-151.

Gilbert HF (1998) Methods Enzymol 290: 26-50.

Gillikin JW, Zhang F, Coleman CE, Bass HW, Larkins BA, Boston RS (1997) Plant Physiol 114: 345-352.

Goldberger RF, Epstein CJ, Anfisen CB (1963) J Biol Chem 238: 628-635.

Goldberg ME, Rudolph R, Jaenicke R (1991) Biochemistry 30: 2790-2797.

Grauschopf U, Winther JR, Korber P, Zander T, Dallinger P, Bardwell JC (1995) Cell 83: 947-955.

Greene FC, Anderson OD, Litts JC, Gautier MF (1985) Cereal Chem 62: 398-405.

Grimwade B, Tatham AS, Freedman RB, Shewry PR, Napier JA (1996) Plant Mol Biol 30: 1067-1073.

Grosset J, Alary R, Gautier MF, Menossi M, Martinez-Izquierdo JA, Joudrier P (1997) Plant Mol Biol 34: 331-338.

Guilfoyle TJ (1997) In: Shetlow JK, eds Genetic engineering vol 19, Plenum Press, New York: 15-47.

Günther R, Srinivasan M, Haugejorden S, Green M, Ehrbrecht IM, Küntzel H (1993) J Biol Chem 268: 7728-7732.

Gupta RB, Khan K, Mac Ritchie F (1993) J Cereal Sci 18: 23.

Gupta RB, Masci S, Lafiandra D, Bariana HS, MacRitchie F (1996) J Exp Bot 47: 1377.

Halford NG, Forde J, Shewry PR, Kreis M (1989) Plant Sci 62: 207-216.

Halford NG, Field JM, Blair H, Urwin P, Moore K, Robert L, Thompson R, Flavel RB, Tatham AS, Shewry PR (1992) Theor Appl Genet 83: 373-378.

Hammond C, Helenius A (1994) J Cell Biol 126: 41-52.

Hammond C, Helenius A (1995) Curr Opin Cell Biol 7: 523- 529.

Haugejorden SM, Srivinasan M, Green M (1991) The Journal of Biol Chem 266: 6015-6018.

Harding HP, Zhang Y, Zeng H, Novoa I, Lu PD, Calfon M, Sadri N, Yun C, Popko B, Paules R, Stojdl DF, Bell JC, Hettmann T, Leiden JM, Ron D (2003) Mol Cell 11: 619-633.

Harper JD, Lansbury Jr. PT (1997) Annu Rev Biochem 66: 385-407.

Hausman RE, Moscona AA (1975) Proc Natl Acad Sci USA 72: 916-920.

Hawkins HC, Blackburn EC, Freedman RB (1991) Biochem J 275: 349-353.

Hayano T, Kikuchi M (1995) FEBS Lett 372: 210-214.

Hayano T, Hirose M, Kikuchi M (1996) FEBS Lett 377: 505-511.

Hebsgaard SM, Korning PG, Tolstrup L, Engelbrecht J, Rouzè P, Brunak S (1996) Nucl Acids Res 24: 3439-3452.

Hirano N, Shibasaki F, Sakai R, Tanaka T, Nishida J, Yazaki Y, Takenawa T, Hirai H (1995a) Eur J Biochem 234: 336-342.

Hirano HY, Tabayashi N, Matsumura T, Tanida M, Komeda Y, Sano Y (1995b) Plant Cell Physiol 36:37-44.

Hoagland DR, Arnon DI (1950) Circ. 347. Berkeley, CA: Agric Exp Stn, Univ of Calif

Holaska JM, Black BE, Love DC, Hanover JA, Leszyk J, Paschal BM (2001) J Cell Biol 152: 127-140.

Holst B, Tachibana C, Winther JR (1997) J Cell Biol 138: 1229-1238.

Hoober KL, Glynn NM, Burnside J, Coppock DL, Thorpe C (1999) J Biol Chem 274: 31759-31762.

Horibe T, Gomi M, Iguchi D, Ito H, Kitamura Y, Masuoka T, Tsujimoto I, Kimura T, Kikuchi M (2004) J Biol Chem 279: 4604-4611.

Hosoda A, Kimata Y, Tsuru A, Kohno K (2003) J Biol Chem 278: 2669-2676.

Hotchkiss KA, Matthias LJ, Hogg PJ (1998) Biochim Biophys Acta 1388: 478-488.

Houston NL, Fan C, Xiang QY, Schulze JM, Jung R, Boston RS (2005) Plant Physiol 137: 762-778.

Huebner FR, Bietz JA (1972) J of Cereal Sci 4: 379.

Huth JR, Norton SE, Lockridge O, Shikone T, Hsueh AJW, Ruddon RW (1994) J Endocrinology 135: 911-918.

Hwang C, Sinskey AJ, Lodish HF (1992) Science 257: 1496- 1502.

Jackson EA, Holt LM, Payne PI (1985) GeneT Res 46: 11-17.

Jacobi A, Huber-Wunderlich M, Hennecke J, Glockshuber R (1997) J Biol Chem 272: 21692-21699.

Jeenes DJ, Pfaller R, Archer DB (1997) Gene 193: 151-156.

Jessop CE, Bulleid NJ (2004) J Biol Chem 279: 55341-55347.

Jiang XM, Fitzgerald M, Grant CM, Hogg PJ (1999) J Biol Chem 274: 2416-2423.

John DC, Grant ME, Bulleid NJ (1993) EMBO J 12: 1587-1595.

Johnson JC, Bhave M (2004) Plant Sci 167: 397-410.

Johnson S, Michalak M, Opas M, Eggleton P (2001) Trend Cell Biol 11: 122-129.

Kadokura H, Katzen F, Beckwith J (2003) Annu Rev Biochem 72: 111-135.

Kanai S, Toh H, Hayano T, Kikuchi M (1998) J Mol Evol 47: 200-210.

Kao CY, Cocciolone SM, Vasil IK, McCarty DR (1996) Plant Cell 8: 1171-1179.

Katti SK, LeMaster DM, Eklund H (1990) J Mol Biol 212: 167-184.

Kawagoe Y, Suzuki K, Tasaki M, Yasuda H, Akagi K, Katoh E, Nishizawa NK, Ogawa M, Takaiwa F (2005) Plant Cell 17: 1141-1153.

Kelley LA, MacCallum RM, Sternberg MJE (2000) J Mol Biol 299(2): 499-520.

Kemmink J, Darby NJ, Dijkstra K, Scheek RM, Creighton TE (1995) Protein Sci 4: 2587-2593.

Kemmink J, Darby NJ, Dijkstra K, Nilges M, Creighton TE (1996) Biochemistry 35: 7648-7691.

Kemmink J, Darby NJ, Dijkstra K, Nilges M, Creighton TE (1997) Curr Biol 7: 239-245.

Kim CS, Hunter BG, Kraft J, Boston RS, Yang S, Jung R, Larkins BA (2004) Plant Physiol 134: 380-387.

Kim JM, Mayfield SP (1997) Science 278: 1954-1957.

Klappa P, Hawkins HC, Freedman RB (1997) Eur J Biochem 248: 37-42.

Klappa P,Ruddock LW, Darby NJ, Freedman RB (1998) EMBO J 17: 927-935.

Knoblach B, Keller BO, Groenendyk J, Aldred S, Zheng J, Lemire BD, Li L, Michalak M (2003) Mol Cell Proteomics 2: 1104-1119.

Ko FC, Chow KL (2002) Development 129: 1185-1194.

Koivunen PN, Helaakoski T, Karvonen P, Jaakkola M, Palotie A, Kivirikko KI (1997) Genomics 42: 397-404.

Koo EH, Lansbury PT, Kelly JW (1999) Proc Natl Acad Sci USA 96: 9989-9990.

Kortemme T, Darby NJ Creighton TE (1996) Biochemistry 35: 14503-14511.

Kramer B, Ferrari DM, Klappa P, Pohlmann N, Söling HD (2001) Biochem J 357: 83-95.

Kreis M, Forde BG, Rahman S, Miflin BJ, Shewry PR (1985) J Mol Biol 183: 499-502.

Krishna Rao ASM, Hausman RE (1993) Proc Natl Acad Sci USA 90: 2950-2954.

Krishnan HB, Franceschi VR, Okita TW (1986) Planta 169: 471-480.

Krishnan HB, White JA, Pueppke SG (1992) Plant Sci 81: 1-11.

Krishnan HB, White JA (1995) Plant Physiol 109: 1491-1495.

Kulp MS, Frickel EM, Ellgaard L, Weissman JS (2005) J Biol Chem 281: 876-884.

Kuznetsov G, Chen LB, Nigam SK (1994) J Biol Chem 269: 22990-22995.

Laboissière MC, Sturley SL, Raines RT (1995) J Biol Chem 270: 28006-28009.

Lamacchia C, Shewry PR, Di Fonzo N, Forsyth JL, Harris N, Lazzeri PA, Napier JA, Halford NG, Barcelo P (2001) J Exp Bot 355(52): 243-250.

LaMantia ML, Lennarz WJ (1993) Cell 74: 899-908.

Lanctôt C, Gauthier Y, Drouin J (1999) Endocrinology 140: 1416-1422.

Lange H, Lisowsky T, Gerber J, Muhlenhoff U, Kispal G, Lill R (2001) EMBO Rep 2: 715-720.

Lappi AK, Lensink MF, Alanen HI, Salo KEH, Lobell M, Juffer AH, Ruddok LW (2004) J Mol Biol 335: 283-295.

Larkins BA, Hurkman WJ (1978) Plant Physiol 62: 256-263.

Lawrence DA, Song R, Weber P (1996) J Leukoc Biol 60: 611-618.

Lee J, Hofhaus G, Lisowsky T (2000) FEBS lett 477: 62-66.

Leisy DJ, Hnilo J, Zhao Y, Okita TW (1989) Plant Mol Biol 14: 41-50.

Lemaire SD, Miginiac-Maslow M (2004) Photosynthesis Res 82: 203-220.

Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouze P, Rombauts S (2002) Nucleic Acids Res 30: 325-327.

Letunic I, Gootstadt L, Dickens NJ, Doerks T, Schultz J, Mott R, Ciccarelli F, Copley RR, Ponting CP, Bork P (2002) Nucleic Acids Res 30: 242-244.

Levanony H, Rubin R, Altschuler Y, Galili G (1992) J Cell Biol 119: 1117-1128.

Levitan A, Trebitsh T, Kiss V, Pereg Y, Dangoor I, Danon A (2005) Proc Natl Acad Sci USA 102: 6225-6230.

Lew E J-L, Kuzmicky DD, Kasard DD (1992) Cereal Chem 69: 508.

Lewis MJ, Mazzarella RA, Green M (1986) Arch Biochem Biophys 245: 389-403.

Li CP, Larkins BA (1996) Plant Mol Biol 30: 873-882.

Li X, Wu Y, Zhang DZ, Gillikin JW, Boston RS, Franceschi VR, Okita TW (1993) Science 262: 1054-1056.

Liepinsh E, Baryshev M, Sharipo A, Ingelman-Sundberg M, Otting G, Mkrtchian S (2001) Structure (Camb) 9: 457-471.

Lisowsky T, Weinstat-Saslow DL, Barton N, Reeders St, Schneider MC (1995)Genomics 29: 690-697.

Livesley MA, Bulleid NJ, Bray CM (1992) Seed Sci Res 2: 97-103.

Lodish HF, Kong N (1993) J Biol Chem 268: 20598-20605.

Lu YH, Merlino M, Isaac PG, Stacci J, Bernard R, Leroy P (1994) Agronomie 14:33-39.

Lu DP, Christopher DA (2006) Int J Plant Sci 167(1): 1-9.

Lucero HA, Kaminer B (1999) J Biol Chem 274: 3243-3251.

Lucero HA, Lebecke E, Kaminer B (1994) J Biol Chem 269: 23112-23119.

Lundström J, Krause G, Holmgren A (1992) J Biol Chem 267: 9047-9052.

Lundström J, Holmgren A (1993) Biochemistry 32: 6649-6655.

Macer DRY, Koch GLE (1988) Journal of Cell Science 92: 61-70.

MacRitchie F (1987) J of Cereal Science 6: 259-268.

MacRitchie F, Gupta R (1993) Aust J Agric Res 41: 603.

Mandel R, Ryser HJ, Ghani F, Wu M, Peak D (1993) Proc Natl Acad Sci USA 90: 4112-4116.

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning. A Laboratory Manual. Cold Spring Harbor, NY.

Marzàbal P, Busk PK, Ludevid MD, Torrent M (1998) Plant J 16: 41-52.

Masci S, Lafiandra D, Porceddu E, Lew EJL, Ta HP, Kasarda D (1993) Cereal Chem 70:581.

Matsuo Y, Akiyama N, Nakamura H, Yodoi J, Noda M, Kizaka-Kondoh S (2001) J Biol Chem 276: 10032-10038.

Mazzarella R, Srinivasan M, Haugejorden M, Green M (1990) J Biol Chem 265:1094-1101.

McArthur AG, Knodler LA, Silberman JD, Davids BJ, Gillin FD, Sogin ML (2001) Mol Biol Evol 18: 1455-1463.

Meng X, Zhang C, Chen J, Peng S, Cao Y, Ying K, Xie Y, Mao Y (2003) Biochem Genet 41: 99-106.

Mezghrani A, Fassio A, Benham A, Simmen T, Braakman I, Sitia R (2001) EMBO J 20: 6288-6296.

Miernyk JA, Duck NB, Shatters RGJ, Folk WR (1992) Plant Cell 4: 821-829.

Mkrtchian S, Fang C, Hellman U, Ingelman-Sundberg M (1998) Eur J Biochem 251: 304-313.

Molteni SN, Fassio A, Ciriolo MR, Filomeni G, Pasqualetto E, Fagioli C, Sitia R (2004) J Biol Chem 279: 32667-32673.

Monnat J, Hacker U, Geissler H, Rauchenberger R, Neuhaus EM, Maniak M, Soldati T (1997) FEBS Lett 418: 357-362.

Monnat J, Neuhaus E, Pop MS, Ferrari DM, Kramer B, Soldati T (2000) Mol Biol Cell 11: 3469-3484.

Morjana NA, Gilbert HF (1991) Biochemistry 30: 4985-4990.

Müller M, Knudsen S (1993) Plant J 4: 343-355.

Murashige T, Skoog F (1962) Physiol Plant 15: 473-497.

Nakamoto H, Bardwell JC (2004) Biochim Biophys Acta 1694: 111-119.

Ndubuisi MI, Guo GG, Fried VA, Etlinger JD, Sehgal PB (1999) J Biol Chem 274: 25499-25509.

Nigam SK, Goldberg AL, Ho H, Rohde ME, Bush KT, Sherman MY (1994) J Biol Chem 269: 1744-1749.

Noiva R, Lennarz WJ (1992) J Biol Chem 6: 3553-3556.

Norgaard P, Westphal V, Tachibana C, Alsoe L, Holst B, Winther JR (2001) J Cell Biol 152: 553-562.

Norre F, Peyrot C, Garcia C, Rancé I, Drevet J, Theisen M, Gruber V (2002) Plant Mol Biol 50: 699-712.

Oliver JD, Van der Wal FJ, Bulleid NJ, High S (1997) Science 275: 86-87.

Orth RA, Buishuk W (1972) Cereal Chem 49: 268.

Osborne TB (1924) The Vegetable Proteins 2nd edn, , Longmans, Green and co, London.

Pagani M, Fabbri M, Benedetti C, Fassio A, Pilati S, Bulleid NJ, Cabibbo A, Sitia R (2000) J Biol Chem 275: 23685-23692.

Pandya MJ, Sessions RB, Williams PB, Dempsey CE, Tatham AS, Shewry PR, Clarke AR (2000) Proteins 38: 341-349.

Pantoja-Uceda D, Bruix M, Gimenez-Gallego G, Rico, M Santoro J (2003) Biochemistry 42: 13839-13847.

Pantoja-Uceda D, Shewry PR, Bruix M, Tatham AS, Santoro J, Rico M (2004) Biochemistry 43: 6976-6986.

Panzner S, Dreier L, Hartmann E, Kostka S, Rapoport T (1995) Cell 81: 561- 570.

Pariser HP, Zhang J, Hausman RE (2000) Exp Cell Res 258: 42-52.

Parkkonnen T, Kivirikko KI, Pilhajaniemi T (1988) Biochem J 256: 1005-1011.

Payne PI, Corfield KG (1979) Planta 145:83.

Payne PI, Corfield KG, Blackman JA (1979) Theor Appl Genet 55: 153-159.

Payne PI, Corfield KG, Holt LM, Blackman JA (1981) J Sci Fd Agric 32: 51-60.

Payne PI, Lawrence GD (1983) Cereal Research Communication 11:29-35.

Pedrazzini E, Giovinazzo G, Bollini R, Ceriotti A, Vitale A (1994) Plant J 5: 103-110.

Pelham HRB (1988) EMBO J 7: 913-918.

Pellegrineschi A, Fennel S, McLean S, Brito RM, Velazquez L, Salgado M, Olivares JJ, Hernandez R, Hoisington D (1998) Wheat Biotechnology International Workshop-Colonia, Uruguay 18-21 November 1998.

Pellegrineschi A, McLean S, Salgado M, Velazquez L, Hernandez R, Brito RM, Noguera M, Medhurst A, Hoisington D (2001) Euphytica 119: 133-136.

Pellegrineschi A, Noguera LM, Skovmand B, Brito RM, Velazquez L, Salgado MM, Hernandez R, Warburton M, Hoisington D(2002) Genome 45: 421-430.

Pihlajaniemi T, Helaakoski T, Tasanen K, Mylylla R, Huhtala ML, Koivu JG, Kivirikko KI (1987) EMBO J 6: 643-649.

Pirneskoski A, Klappa P, Lobell M, Williamson RA, Byrne L, Alanen HI, Salo KEH, Kivirikko KI, Freedman RB, Ruddock LW (2004) J Biol Chem 279: 10374-10381.

Pollard MG, Travers KJ, Weissman JS (1998) Mol Cell 1: 171-182.

Pollock S, Kozlov G, Pelletier MF, Trempe JF, Jansen G, Sitnikov D, Bergeron JJ, Gehring K, Ekiel I, Thomas DY (2004) EMBO J 23: 1020-1029.

Primm TP, Walker KW, Gilbert HF (1996) J Biol Chem 271: 33664-33669.

Puig A, Gilbert HF (1994a) J Biol Chem 269: 7764-7771.

Puig A, Gilbert HF (1994b) J Biol Chem 269: 25889-25896.

Quan H, Fan GB, Whang CC (1995) J Biol Chem 270: 17078-17080.

Raposo G, van Santen HM, Leijendekker R, Geuze HJ, Ploegh HL (1995) J Cell Biol 131: 1403-1419.

Rechinger KB, Simpson DJ, Svendson I, Cameron-Mills V (1993) Plant J 4: 841-853.

Reeves CD, Krishnan HB, Okita TW (1986) Plant Physiol 82: 34-40.

Roden LT, Miflin BJ, Freedman RB (1982) FEBS Lett 138: 121-124.

Roderick HL, Campbell AK, Llewllyn DH (1997) FEBS Lett 405: 181-185.

Rothwarf DM, Li YJ, Scheraga HA (1998) Biochemistry 37: 3760-3766.

Rubin R, Levanony H, Galili G (1992) Plant Phisiol 99: 718-724.

Ruddock LW, Freedman RB, Klappa P (2000) Protein Sci 9: 758-764.

Rupp K, Birnbach U, Lundström J, Van Nguyen P, Söling HD (1994) J Biol Chem 269: 2501-2507.

Russell SL, Ruddock LW, Salo KEH, Oliver JD, Roebuck QP, Llewellyn DH, Roderick HL, Koivunen P, Myllyharju J, High S (2004) J Biol Chem 279: 18861-18869.

Ryser HJ, Levy EM, Mandel R, DiSciullo GJ (1994) Proc Natl Acad Sci USA 91: 4559-4563.

Sahrawy M, Hecht V, Lopez-Jaramillo J, Chueca A Chartier Y, Meyer Y (1996) J Mol Evol 42: 422-431.

Saitou N, Nei M (1987) Mol Biol Evol 4:406-425.

Scherens B, Dubois E, Messenguy F (1991) Yeast 7: 185-193.

Schultz J, Copley RR, Doerks T, Ponting CP, Bork P (2000) Nucleic Acids Res 28: 231-234.

Schwaller M, Wilkinson B, Gilbert HF (2002) J Biol Chem 278: 7154-7159.

Sears ER (1966) In: Riley R, Lewis KR, eds Chromosome manipulation and plant genetics. Oliver and Boyd, London: 29-45.

Sevier CS, Cuozzo JW, Vala A, Aslund F, Kaiser CA (2001) Nat Cell Biol 3: 874-882.

Sevier CS, Kaiser CA (2006) Antioxid Redox Signal 8: 797-811.

Shani N, Rosenberg N, Kasarda DD, Galili G (1994) J Biol Chem 269: 8924-8930.

Shen Y, Danon A, Christopher DA (2001) Plant Cell Physiol 42: 1071-1078.

Shewry PR, Tatham AS, Forde J, Kreis M, Miflin BJ (1986) J Cereal Sci 4: 97-100.

Shewry PR, Halford NG, Tatham AS (1992) Journal of Cereal Science 15: 105-120.

Shewry PR, Miles MJ, Tatham AS (1993) Prog. Biophys Molec Biol 61: 37-59.

Shewry PR, Tatham AS (1997) J Cereal Sci 25: 207-227.

Shewry PR (1999) Cereal Foods World 44: 587-589.

Shewry PR, Tatham AS, Halford NG (1999) in: Casey R, Shewry PR, eds Seed Proteins. Dordrecht: Kluwer Academic Press: 35-78.

Shewry PR, Halford NG (2002) J Exp Bot 53(370): 947-958.

Shimoni Y, Segal G, Zhu X, Galili G (1995a) Plant Physiol 107: 281.

Shimoni Y, Zhu X, Levanoy H, Segal G, Galili G (1995b) Plant Physiol 108: 327-335.

Shimoni Y, Galili G (1996) J Biol Chem 271: 18869-18874.

Shorrosh BS, Dixon RA (1991) Proc Natl Acad Sci USA 88: 10941-10945.

Shorrosh BS, Dixon RA (1992) Plant J 2: 51-58.

Shorrosh BS, Wen L, Zen KC, Huang JK, Pan JS, Hermodson MA, Tanaka K, Muthukrishnan S, Reeck GR (1992) Plant Mol Biol 18: 151-154.

Shotwell M, Larkins BA (1989) In: Marcus E, eds. The Biochemistry of Plant: A comprehensive Treatise. Academic Press, Orlando; FL: 296-345.

Sideraki V, Gilbert HF (2000) Biochemistry 39: 1180-1188.

Singh NK, Sheperd KW (1988) Theor Appl Genet 75: 628-641.

Smith T, Ferreira LR, Hebert C, Norris K, Sauk JJ (1995) J Biol Chem 270: 18323-18328.

Solovyov A, Xiao R, Gilbert HF (2004) J Biol Chem 279: 34095-34100.

Song JL, Wang CC (1995) Eur J Biochem 231: 312-316.

Song JL, Quan H, Wang CC (1997) Biochem J 328: 841-846.

Southern E, (1975) J Mol Biol 98: 503.

Stein G, Lisowsky T (1998) Yeast 14: 171-180.

Stein GS, van Wijnen AJ, Stein J, Lian JB, Montecino M (1995) Int Rev Cytol 162A: 251-278.

Stevenson SG, Preston KR (1996) J Cereal Sci 23: 121.

Suh WC, Lu CZ, Gross CA (1999) J Biol Chem 274:30534-30539.

Sullivan DC, Huminiecki L, Moore JW, Boyle JJ, Poulsom R, Creamer D, Barker J, Bicknell R (2003) J Biol Chem 278: 47079-47088.

Schwaller M, Wilkinson B, Gilbert HF (2002) J Biol Chem 278:7154-7159.

Tachibana C, Stevens TH (1992) Mol Cell Biol 12: 4601-4611.

Tachikawa H, Takeuchi Y, Funahashi W, Miura T, Gao XD, Fujimoto D Mizunaga T, Onodera K (1995) FEBS Lett 369: 212-216.

Tachikawa H, Funahashi W, Takeuchi Y, Nakanishi H, Nishihara R, Katoh S, Gao XD, Mizunaga T, Fujimoto D (1997) Biochem Biophys Res Commun 239: 710-714.

Takaiwa F, Kikuchi S, Oono K (1987) Mol Gen Genet 208: 15-22.

Takaiwa F, Yamanouchi U , Yoshihara T, Washida H, Tanabe F, Kato A, Yamada D (1996) Plant Mol Biol 30: 1207-1221.

Takemoto Y, Coughlan SJ, Okita TW, Satoh H, Ogawa M, Kumamaru T (2002) Plant Physiol 128: 1212-1222.

Tan JT, Bardwell JC (2004) Chembiochem 5: 1479-1487.

Tanaka K, Sugimoto T, Ogawa M, Kasai Z (1980) Agric Biol Chem 44: 1633-1639.

Tanaka S, Uehara T, Nomura Y (2000) J Biol Chem 275: 10388-10393.

Taniyama Y, Seko C, Kikuchi M (1990) J Biol Chem 265: 16767-16771.

Terada K, Manchikalapudi P, Noiva R, Jauregui HO, Stockert RJ, Schilsky ML (1995) J Biol Chem 270: 20410-20416.

Thomas MS, Flavell RB (1990) Plant Cell 2: 1171-1180.

Thomas PJ, Qu BH, Pedersen PL (1995) Trends Biochem Sci 20: 456-459.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) Nucleic Acids Res 25: 4876-4882.

Trebitsh T, Levitan A, Sofer A, Danon A (2000) Mol Cell Biol 20: 1116-1123.

Tsai B, Rodighiero C, Lencer WI, Rapoport TA (2001) Cell 104: 937-948.

Tsibris JMC, Hunt LT, Ballejo G, Barker WC, Toney LJ, Spellacy WN (1989) J Biol Chem 264: 13967-13970.

Tu BP, Ho-Schleyer SC, Travers KJ, Weissman JS (2000) Science 290: 1571-1574.

Tu BP, Weissman JS (2002) Mol Cell 10: 983-994.

Tu BP, Weissmann JS (2004) J Cell Biol 164(3): 341-346.

Turano C, Coppari S, Altieri F, Ferraro A (2002) J Cell Physiol 193: 154-163.

Upadahya MD, Swaminthan MS (1963) Cromosoma 14: 589-600.

Urade R, Kito M (1992) FEBS Lett 312: 83-86.

Urade R, Okudo H, Kato H, Moriyama T, Arakaki Y (2004) Biochemistry 43: 8858-8868.

van Lith M, Hartigan N, Hatch J, Benham AM (2005) J Biol Chem 280: 1376-1383.

Van Nguyen P, Rupp K, Lampen A, Söling HD (1993) Eur J Biochem 213: 789-795.

VanderWaal RP, Spitz DR, Griffith CL, Higashikubo R, Roti JL (2002) J Cell Biochem 85: 689-702.

Verde C, Pascale MC, Martire G, Lotti LV, Torrisi MR, Helenius A, Bonatti S (1995) Eur J Cell Biol 67: 267-274.

Walhund KG, Gustaffson M, MacRitchie F, Nylander T, Wannerberger L (1996) Cereal Sci 23: 113.

Walker KW, Gilbert HF (1995) Biochemistry 34: 13642-13650.

Walker KW, Lyles MM, Gilbert HF (1996) Biochemistry 35: 1972-1980.

Walker KW, Gilbert HF (1997) J Biol Chem 272: 8845-8848.

Wang Q, Chang A (1999) EMBO J 18: 5972-5982.

Washida H, Wu CY, Suzuki A, Yamanouchi U, Akihama T, Harada K, Takaiwa F (1999) Plant Mol Biol 40: 1-12.

Weegels PL, Hamer RJ, Schofield JD (1996) J Cereal Sci 23:1.

Weisbart RH (1992) J Immunol 148: 3958-3963.

Weissman JS, Kim PS (1991) Science 253: 1386-1390.

Weissman JS, Kim PS (1993) Nature 365: 185-188.

Wetterau JR , Combs KA, Spinner SN, Joiner BC (1990) J Biol Chem 265: 9801-9807.

Wetterau JR, Combs KA, McLean LR, Spinner SN, Aggerbeck LP (1991) Biochemistry 30: 9728-9735.

Whiteley EM, Hsu TA, Betenbaugh MJ (1997) J Biol Chem 272: 22556-22563.

Wilkinson B, Gilbert HF (2004) Biochim Biophys Acta 1699: 35-44.

Wittke I, Wiedemeyer R, Pillmann A, Savelyeva L, Westermann F, Schwab M (2003) Cancer Res 63: 7742-7752.

Woychik JH, Boundy JA, Dimler RJ (1961) Arch Biochem Biophys 94: 4770-4820.

Wrigley CW (1996) Nature 381: 738.

Wrobel R (1996) PhD Thesis. North Carolina State University, Raleigh, NC.

Wroblewski VJ, Masnyk M, Khambatta SS, Becker GW (1992) Diabetes 41: 539-547.

Wu CY, Suzuki A, Washida H, Takaiwa F (1998) Plant J 14: 673-683.

Wu CY, Washida H, Onodera Y, Harada K, Takaiwa F (2000) Plant J 23: 415-421.

Wunderlich M, Glockshuber R (1993) Protein Sci 2: 717-726.

Xiao R, Wilkinson B, Solovyov A, Winther JR, Holmgren A, Lundström-Ljung J, Gilbert HF (2004) J Biol Chem 279: 49780-49786.

Xu ZJ, Ueda K, Masuda K, Ono M, Inoue M (2002) Gene 284: 225-231.

Yamagata H, Sugimoto T, Tanaka K, Kasai Z (1982) Plant Physiol 70: 1094-1100.

Yamagata H, Tanaka K (1986) Plant Cell Physiol 27: 135-145.

Yamamoto K, Fujii R, Toyofuku Y, Saito T, Koseki H, Hsu VW, Aoe T (2001) EMBO J 20: 3082-3091.

Yao Y, Zhan JC, Whang CC (1997) EMBO J 16:651-658.

Yoshihara T, Washida H, Takaiwa F (1996) FEBS Lett. 383: 213-218.

Yoshimori T, Semba T, Takemoto H, Akagi S, Yamamoto A, Tashiro Y (1990) J Biol Chem 265: 15984-15990.

Zapun A, Creighton TE, Rowling PJ, Freedman RB (1992) Proteins 14: 10-15.

Zapun A, Bardwell JC, Creighton TE (1993) Biochemistry 32: 5083-5092.

Zapun A, Missiakas D, Raina S, Creighton TE (1995) Biochemistry 34: 5075-5089.

Zapun A, Petrescu SM, Rudd PM, Dwek RA, Thomas DY, Bergeron JJ (1997) Cell 88: 29-38.

Zapun A, Darby NJ, Tessier DC, Michalak M, Bergeron JJM, Thomas DY (1998) J Biol Chem 273: 6009-6012.

Zhang F, Boston RS (1992) Protoplasma 171: 142-152.

Zhao WM, Gatehouse JA, Boulter D (1983) FEBS Lett 162: 96-102.

Zou Y, Crowley DJ, Van Houten B (1998) J Biol Chem 273: 12887-12892.

RINGRAZIAMENTI

Ringrazio il Prof. Enrico Porceddu, Coordinatore di questo dottorato, ed il Prof. Oronzo Antonio Tanzarella per l'interesse con il quale hanno seguito la fase sperimentale del lavoro e per la lettura critica della presente tesi di dottorato. In particolare voglio ringraziare il Prof. Mario Ciaffi che in questi tre anni è stato sempre disponibile come amico e come docente, la cui preziosa guida mi ha permesso di realizzare il presente lavoro. Ringrazio tutti i colleghi che con me hanno condiviso quest'esperienza: Andrea, Cristiano, Giulio, Rosita ed con particolare affetto e gratitudine Anna Rita. Ringrazio i miei numerosi amici, agronomi e forestali dell'Università della Tuscia per il loro affetto e per la loro comprensione. Ringrazio i miei amici e colleghi del CIMMYT per la loro amicizia ed i loro numerosi consigli: Christelle, Debra, Eblin, Eric, Gary, Janet, Joost, Lorena, Luis Miguel, Margherita, Maria, Mario, Marta, Maru, Nick, Pablo, Peter, Rosa Maria, Rosaura, Silverio, Stefanie, Susanne e moltissimi altri. Ringrazio per l'opportunità offertami ed il suo aiuto, il Dott. Jonathan Crouch. Ringrazio il Dott. Alessandro Pellegrineschi, il Dott. Manilal William e la Dott.sa Monica Mezzalama per il loro aiuto. In particolar modo voglio ringraziare anche il Dott. Jean Philippe Vielle Calzada ed il Dott. Peter R. Matthews per aver messo a mia disposizione in diverse discussioni scientifiche la loro preziosa esperienza e per l'interesse mostrato per il mio lavoro. Ringrazio infine la mia famiglia per il loro amore e il loro appoggio. Ad alcuni amici straordinari va tutto il mio affetto per la loro infinita pazienza e per il loro costante aiuto: Arianna, Christelle, Debra e Fabio.