

CLOWING, CHARACTERIZATION and EXPRESSION of a NOVEL  
METALLOTHIONEIN GENE from *Triticum durum*  
and  
THEORETICAL and EXPERIMENTAL  
STRUCTURE-FUNCTION RELATIONSHIP PREDICTION

by  
KIVANÇ BİLECEN

Submitted to the Graduate School of Engineering and Natural Sciences  
in partial fulfillment of  
the requirements for the degree of  
Master of Science

Sabancı University

June 2003

CLONING, CHARACTERIZATION and EXPRESSION of a NOVEL  
METALLOTHIONEIN GENE from *Triticum durum*  
and  
THEORETICAL and EXPERIMENTAL  
STRUCTURE-FUNCTION RELATIONSHIP PREDICTION

APPROVED BY:

Assoc. Prof. Zehra Sayers .....  
(Dissertation Supervisor)

Dr. Uğur Sezerman .....  
(Dissertation Co-Supervisor)

Prof. Michel Koch .....

Prof. Yuda Yürüm .....

Asist. Prof. Alpay Taralp .....

DATE OF APPROVAL: .....

© Kıvanç BİLECEN 2003

All Rights Reserved

## ABSTRACT

Metallothioneins (MTs) are small, cystein rich, low molecular mass polypeptides found in almost all organisms. They are thought to be involved in heavy-metal detoxification and metabolism of essential trace elements like copper and zinc. Unlike their mammalian counterparts, plant MTs have not been thoroughly characterized in terms of cellular regulation and function.

A novel gene, from *Triticum durum* (pasta wheat), coding for plant MT type 1 protein was isolated and characterized. The *durum mt* gene was cloned in *E. coli* for solution X-ray scattering studies to obtain the first experimental structural data on a plant MT in the literature. *Triticum durum mt* gene was shown to contain 2 exons and a non-coding intron region.

The coded MT protein, showing high similarity to mammalian MTs in its cystein residue distribution pattern, forms two metal binding domains bridged with an exceptionally long connecting region. This hinge region was shown to be highly conserved among plant MTs using sequence alignment algorithms on data available in the literature.

Homology modeling and heuristic fragment assembly approaches were used to predict a 3D structure for the *durum* MT (dMT). Guided by the predicted structures, functional motif and structure searches were performed yielding a possible DNA binding and/or protein interaction function for dMT. High probability of wMT to form dimers or trimers inside the solution was also speculated.

dMT was expressed in *E. coli* cells as a fusion protein with GST and preliminary X-ray solution scattering measurements were carried out on the purified recombinant

protein. These measurements indicated the high tendency of the protein to form aggregates in solution. Theoretical predictions and solution scattering measurements were also supported by the results of polyacrylamide gel electrophoresis and size exclusion chromatography analysis of expressed and purified recombinant dMT and GSTdMT proteins.

Further, sequence and structure analyses showed a high structure and sequence similarity between dMT hinge region and the DNA binding domain of a cyanobacterial metallothionein suppressor protein (SmtB). Indeed, the results indicate that dMT metal binding domains would also bind to DNA with very high probability. These results, altogether, point to a new role for plant MTs other than metal scavenging such as being a transcription factor or a gene suppressor.

## ÖZET

Metallotioninler (MT'ler) hemen hemen tüm organizmalarda bulunan sistin bakımından zengin, düşük molekül ağırlıklı polipeptitlerdir. Hücre içerisinde bulunan ağır metallerin detoksifikasyonu ve bakır, çinko gibi yaşam için gerekli eser elementlerin metabolize edilmesi görevlerini yürüttükleri düşünülmektedir. Memelilerdeki eşleniklerinin tersine, bitki metallotioninleri hücre içerisindeki fonksiyonları ve regülasyonları açısından tam olarak karakterize edilmemişlerdir.

*Triticum durum*'dan (makarnalık buğday) tip 1 bitki metallotionini kodlayan yeni bir gen tanımlanmıştır. Gen karakterizasyonu çalışmalarının sonuçlarına göre *Triticum durum mt* geni 2 ekzon ve 1 kodlamayan intron kısımlarını içerir. *mt* geninin kodladığı rekombinant protein literatürde bitki metallotionein proteinleri ile ilk deneysel yapı verilerini elde etmek amacıyla *E. coli*'de sentezletirilmiş ve rekombinant MT ile solüsyon X-ışını saçılımı deneyleri başlatılmıştır.

Kodlanan MT proteininin amino asit dizisindeki sistin gruplarının dağılım düzeni açısından memeli MT'lerine çok büyük bir benzerlik gösterdiği ancak durum MT'sinin (dMT) metal bağlayan iki bölümünün alışılmadık uzunlukta bir köprü bölgesiyle birbirlerine bağlandıkları gösterilmiştir. Mevcut protein verileri üzerine uygulanan dizi eşleştirme algoritmaları yardımıyla bu köprü bölgesinin bitki MT'leri arasında korunduğu ispatlanmıştır.

dMT'nin üç-boyutlu yapısını tahmin etmek için homolog modelleme ve iz sürücü (heuristic) parça bütünleştirici yaklaşımları kullanılmıştır. Tahmin edilen yapılar kullanılarak, işlevsel motif ve yapı aramaları yapılmış ve bu aramalar sonucunda dMT için DNA veya proteine bağlanma fonksiyonları öngörülmüştür. Yapılan tahminler

çerçevesinde dMT'nin çözelti içerisinde dimer ve trimer oluşturma ihtimalleri de tartışılmış ve öngörülmüştür.

dMT ekspresyonu *E. coli* içerisinde GST'ye bütünleşik füzyon proteini olarak gerçekleştirilmiştir. Yukarıda sözü edilen teorik sonuçlar rekombinant dMT ve GSTdMT ile solüsyon X-ışını saçılımı ölçümleri, poliakrilamid jel elektroforezi ve jel filtrasyon kromatografisi yöntemleri kullanılarak elde edilen deneysel bulgularla desteklenmiştir.

Detaylı ve geniş kapsamlı dizi ve yapı analizleri, bitki MT'sinin köprü bölgesiyle sayanobakteriel MT baskılayıcı protein (SmtB) DNA bağlanma bölgesi arasında oldukça yüksek oranda bir benzerlik ortaya çıkarmışlardır. Aynı zamanda dMT'nin metal bağlayan bölümlerinin de yüksek oranda DNA'ya bağlanma olasılıkları gösterilmiştir. Tüm sonuçlar birlikte değerlendirildiklerinde, bitki MT proteinleri için metal bağlama yanısıra transkripsiyon faktörü ya da gen baskılayıcı olmak gibi yeni bir işlevi de işaret etmektedirler.

*To my family with all my heart...*

*"...eventually reach the point at which  
we become not the shadows but the light itself..."*



## ACKNOWLEDGEMENTS

I would like to thank, my supervisor and “master” Dr. Zehra Sayers who is one of the most valuable person in my life. I was very lucky to find the chance of working under her supervision in this project. She showed a great patience being the light on my way throughout the study even in very desperate moments. She brought the order to the disorder when every result pointed different directions. Thank you for being with me.

I would also give my tanks to my co-supervisor, Dr. O. Ugur Sezerman, whose help gave a clear view and strong awareness in a totally challenging area, structural bioinformatics.

Dr. Alpay Taralp, on the other hand, was the “Master Yoda” for me and with his advice I learnt to stay calm and prepared even under most challenging conditions. He was the one who thought us scientific realities within the theory. I wish him a long and wonderful life with his beautiful wife Emel.

I’m grateful to Prof. Michel Koch who helped us during the solution X-ray scattering measurements at European Molecular Biology Laboratory, Hamburg Outstation which was a completely new area for me. His opinions and counterarguments forced me to questioning my-self in a way so that I did not see what I wanted to see, but what it was. He also kindly accepted to be a jury member in my thesis defense at Sabanci University, Turkey. I should also express my thanks to Dr. Dimitri Svergun and Margret Fischer for their help and support during our visit in Hamburg.

I'm also thankful to my friend, colleague, and lab partner H. Umit Ozturk, who was the part of the project. Suphan Bakkal, being the “unofficial post doctorate” student of our group was the other essential trace element in the laboratory with my other labmates Mert Şahin, Erinç Şahin and Çağdaş Seçkin. I'm sure they will be successful faculty members in the near future, as they are already successful scientists. Other lab members Adil Doganay Duru and Tolga Sutlu were fortunately present during our hard experiments, being the rescue button.

I will never forget Yavuz Darendelioğlu due to his help and support while I was a computer consultant under his supervision during my undergraduate years at Middle East Technical University. He showed me the world of bioinformatics and opened the door to pass through. Prof. Hüseyin Avni Öktem gave me “tips and tricks” about my future academic life while I was preparing to “take wing” and he was always there like a brother when I needed.

While I was writing these endless pages Itır Ürünay Ürünsak was always with me. Her “beautiful voice” kept me awake during nights and days. I will never forget your help and support. With all my heart, I wish you a wonderful life because you deserve all beauties...

I should also express my special thanks to Prof. İsmail Çakmak for his advice throughout the project and to our valuable faculty members Prof. Huveyda Basaga, Assist. Prof. Damla Bilgin and Assist. Prof. Metin Bilgin. I should not forget the great effort of Tugba Baytekin while we were trying to deal with Murphy's Rules in the lab.

Administrative issues were not a problem for me due to three beautiful ladies; Saila Kurtbay, Işıl Önal, and Zehra Tuğlu. Thank you for your patience and help while we were always forgetting something to submit on time.

During the endless night and days of our beautiful city Istanbul while we were working in the lab or battling in the “fy\_ice”; Dr. Özgür Kütük, Mert Şahin and Özgür Gül, you were there and you will always be...

## TABLE OF CONTENTS

1	INTRODUCTION .....	1
2	OVERVIEW .....	4
2.1	Metallothionein (MT) Proteins .....	4
2.1.1	General information .....	4
2.1.1.1	Nomenclature of MT .....	5
2.1.1.2	Classes and types of MT .....	5
2.1.1.2.1	Cystein residue distribution or source organism.....	7
2.1.1.2.2	Human MTs .....	8
2.1.1.2.3	Non-mammalian MTs.....	9
2.1.1.3	Cellular localization and function of MTs.....	9
2.1.1.3.1	Mammalian systems .....	9
2.1.1.3.2	Function of MTs in systems other than mammalian.....	11
2.2	Plant responses to metal toxicity and metal chelators produced by plants .....	12
2.2.1	Metal uptake by roots .....	12

2.2.2	Metal storage in plants.....	12
2.2.2.1	Metal storage proteins.....	13
2.2.2.2	MT-IIIs (phytochelatins).....	13
2.2.2.3	Metal chelation by small molecules.....	14
2.3	Plant MTs.....	14
2.3.1	Types of plant MTs.....	14
2.3.2	Localization of plant MTs.....	16
2.3.3	Functions of plant MTs.....	19
2.4	Isolation and purification of plant MTs .....	20
2.5	Structural characteristics of MT protein .....	20
2.5.1	Thiol bonds stabilize the protein structure.....	21
2.5.2	Known MT structures .....	23
3	MATERIALS and METHODS .....	26
3.1.1	Chemicals.....	26
3.1.2	Primers .....	26
3.1.3	Enzymes.....	26
3.1.3.1	Restriction enzymes.....	26
3.1.3.2	Ligase.....	26
3.1.3.3	Taq Polymerase.....	27
3.1.3.4	Reverse Transcriptase .....	27
3.1.3.5	Commercial Kits.....	27

3.1.4	Vectors .....	27
3.1.5	Cells .....	28
3.1.6	Buffers and solutions .....	28
3.1.6.1	Culture medium .....	28
3.1.6.1.1	Liquid medium.....	28
3.1.6.1.2	Solid medium.....	29
3.1.6.2	Buffers for gel electrophoresis.....	29
3.1.6.2.1	Denaturing PAGE.....	29
3.1.6.2.2	Non-Denaturing PAGE.....	29
3.1.6.2.3	Agarose gel electrophoresis.....	29
3.1.7	Sequencing.....	29
3.1.8	Equipment.....	30
3.2	Methods .....	31
3.2.1	Plant growth.....	31
3.2.2	DNA and mRNA isolation from plant.....	31
3.2.2.1	Genomic DNA isolation .....	31
3.2.2.2	mRNA isolation .....	32
3.2.3	Bacterial cell growth.....	32
3.2.4	PCR and RT-PCR .....	33
3.2.4.1	PCR.....	33
3.2.4.2	RT-PCR .....	34

3.2.4.3	Purification of PCR products .....	34
3.2.5	Cloning.....	34
3.2.5.1	Subcloning .....	35
3.2.5.2	Ligation .....	35
3.2.5.3	Transformation.....	35
3.2.5.4	Colony Selection.....	35
3.2.5.5	Plasmid isolation.....	36
3.2.5.6	Restriction enzyme digestion.....	36
3.2.5.7	DNA and cDNA analysis.....	36
3.2.5.8	Frozen stocks of cells.....	36
3.2.5.9	Sequence verification.....	37
3.2.5.10	Cloning into expression vector .....	37
3.2.6	Expression and induction.....	37
3.2.7	Purification of the recombinant protein .....	38
3.2.7.1	Batch purification .....	38
3.2.7.2	Column purification.....	38
3.2.8	Cleavage of GSTdMT by thrombin protease.....	39
3.2.9	Size exclusion .....	39
3.2.9.1	Column calibration .....	39
3.2.9.2	Analysis dMT and GSTdMT preparations .....	39
3.2.10	Solution X-ray scattering on dMT and GSTdMT.....	40

3.2.11	Sequence alignments.....	40
3.2.12	Modeling.....	41
3.2.13	Motif search.....	41
4	RESULTS.....	42
4.1	Genomic DNA isolation from <i>Triticum aestivum</i> and <i>Triticum durum</i> .....	42
4.2	Amplification of the target metallothionein ( <i>mt</i> ) gene.....	43
4.3	Cloning of <i>mt-a</i> and <i>mt-d</i> in <i>E. coli</i> .....	45
4.4	Characterization of <i>mt-a</i> and <i>mt-d</i> genes.....	45
4.5	Cloning of <i>mt-a</i> cDNA in <i>E. coli</i> .....	50
4.6	Cloning of <i>mt-d</i> cDNA in <i>E. coli</i> .....	52
4.7	Expression of dMT in <i>E. coli</i> .....	55
4.7.1	Insertion into the expression vector.....	55
4.7.2	Induction of dMT expression in <i>E. coli</i> .....	58
4.8	Purification of dMT protein.....	64
4.8.1	Batch purification.....	64
4.8.2	Purification using GST (Glutathione-S-transferase) affinity chromatography.....	66
4.8.3	Size exclusion.....	69
4.8.3.1	Column calibration.....	69
4.8.3.2	Analysis of dMT preparations by size exclusion chromatography.....	71
4.8.3.3	Analysis of dMT preparations by size exclusion chromatography.....	72
4.9	Solution X-ray scattering on dMT and GSTdMT.....	74

4.9.1	Solution X-ray scattering on GSTdMT .....	74
4.9.2	Solution X-ray scattering on dMT .....	76
4.10	Prediction of wheat MT structure and function .....	78
4.10.1	wMT structure prediction .....	78
4.10.1.1	Secondary structure prediction .....	79
4.10.1.2	Modeling of $\alpha$ - and $\beta$ -domains .....	81
4.10.1.3	Modeling the hinge region.....	84
4.10.1.4	Completing the puzzle .....	87
4.10.2	Function prediction .....	89
4.10.2.1	Conservation of hinge region among plant species .....	89
4.10.2.2	Functional motif search .....	91
4.10.2.3	Search for similar folds.....	92
4.10.2.4	Is wMT a “Natively Unfolded” protein? .....	97
4.10.2.5	Is wMT a DNA binding protein? .....	100
5	DISCUSSION.....	104
6	CONCLUSION.....	113
7	REFERENCES .....	115
	APPENDIX A.....	122
	APPENDIX B.....	126



## LIST OF FIGURES

Figure 2.1: Proposed classification of the Family 1: vertebrate metallothionein proteins (Binz and Kagi, 1997).....	8
Figure 2.2: Plant metallothionein proteins' multiple alignment, showing the 4 types of MTs. Conserved cystein residues are marked with a star. Protein sequences are derived from known gene sequences of Arabidopsis (At), <i>Brassica napus</i> (Bn), rice (Os), pea (Ps), alfalfa (Ms), <i>Brassica oleracea</i> (Bo), petunia (Ph), <i>Silene vulgaris</i> (Sv), banana (Ma), kiwifruit (Ad), cotton (Gh), <i>Picea glauca</i> (Pg), maize (Zm), and wheat (Ta) (Cobbett and Goldsbrough, 2002). .....	15
Figure 2.3: Crsytal structures of rat liver Cu-metallothionein, 4MT2 (above) and NMR structures of Cu-metallothionein of <i>Saccharomyces cerevisiae</i> , 1AQR (below). .....	21
Figure 2.4: Tetrahedral coordination geometry and S-Cys distances for Cd1 in sea urchin metallothionein beta domain, 1QJL_A.....	22
Figure 4.1: Agarose gel electrophoresis analysis of isolated genomic DNA of <i>Triticum aestivum</i> , Bezostaja (left) and <i>Triticum durum</i> , Balcali (right). .....	42
Figure 4.2: Results of optimization studies on PCR conditons. Magnesium ion concentration and annealing temperature were varied as described.....	44
Figure 4.3: ~ 450bp long PCR products of metallothionein gene from <i>T. aestivum</i> ( <i>mt-a</i> ) (left) and <i>T. durum</i> ( <i>mt-d</i> ) (right). .....	44
Figure 4.4: Pairwise aligment of <i>T. aestivum</i> and <i>T. durum</i> genomic MT gene sequences. ....	46
Figure 4.5: Multiple alignment of maize, durum and aestivum metallothionein gene DNA sequences.....	47

Figure 4.6: Multiple alignments of durum, aestivum and maize metallothionein proteins. .....	48
Figure 4.7: <i>T. durum</i> metallothionein gene has 2 exons (blue shaded) and 1 intron (red shaded). .....	48
Figure 4.8: <i>T. aestivum</i> metallothionein gene has 2 exons (blue shaded) and 1 intron (red shaded). .....	48
Figure 4.9: Multiple alignments of AAA50846 (L11879_whe), Balcali (durum_MT) and Bezostaja (aestivum_M) MT protein sequence. ....	49
Figure 4.10: Multiple sequence alignments of Bezostaja (aestivum_M), Balcali (durum_MT_g) and AAA50846 (wheat_MT_L). ....	50
Figure 4.11: Electrophoretic analysis of RT-PCR results showing amplification of <i>T. aestivum</i> cDNA for <i>mt</i> gene.....	51
Figure 4.12: Electrophoretic analysis of digestion results for pGEMaMT constructs. Undigested constructs (lanes 5 and 6). <i>mt-a</i> cDNA bands migrate between 200 and 300 bp bands of the low range DNA ladder. ....	52
Figure 4.13: Electrophoretic analysis of RT-PCR products showing amplification Balcali and Cesit-1252 <i>mt</i> cDNA. ....	53
Figure 4.14: Electrophoretic analysis of digestion results for pGEMdMT constructs. Undigested (undig) and digested (dig) constructs. <i>mt-d</i> cDNA bands migrate between 200 and 300 bp bands of the low range DNA ladder.....	54
Figure 4.15: Electrophoretic analysis of digestion results for pGEMdMT constructs. Undigested (undig) and digested (dig) constructs. <i>mt-d</i> cDNA bands migrate between 200 and 300 bp bands of the low range DNA ladder.....	54
Figure 4.16: Preparative agarose gel analysis for isolation of the <i>d-MT</i> cDNA. Undigested construct (lane 3) and linearized construct with <i>SpeI</i> digestion (lane 4). <i>mt-d</i> cDNA bands migrate between 200 and 300 bp bands of the marker DNA.....	55
Figure 4.17: Electrophoretic analysis of amplified durum cDNA using primers designed for pGEX-4T2 (lane 1) and designed for pGFPuv (lane 2) vectors.....	56

Figure 4.18: Electrophoretic analysis of digestion results for pGEXdMT constructs. <i>mt-d</i> cDNA bands migrate between 200 and 300 bp bands of the low range DNA ladder.	57
Figure 4.19: Growth curve of 0.7 mM IPTG induced (+) and non-induced (-) <i>E. coli</i> BL21 cells (BL21+ and BL21-) containing pGEX-4T-2 (+,-) and pGEXdMT (+,-) vectors.	58
Figure 4.20: SDS-PAGE analysis to check GST and GSTdMT productions in IPTG induced and non-induced cells. (for legend, Table 4.4).	59
Figure 4.21(a): Growth curves of GSTdMT (#4) and GST (#9) expressing BL21(DE3) cells at 0.05 mM CdSO <sub>4</sub> concentration.	61
Figure 4.22: SDS-PAGE analysis to check effects of different IPTG concentrations during induction of recombinant dMT protein expression (for legend, Table 4.6).	63
Figure 4.23: SDS-PAGE analysis of eluted GSTdMT fusion proteins (lanes 2-7). Protein marker 3 (lane 1).	64
Figure 4.24: SDS-PAGE analysis of different elution fractions of batch purified recombinant GST (29 kDa) and GST-dMT (36 kDa). Molecular masses of the marker are indicated on the left.	65
Figure 4.25: Absorbance spectra of batch purified GST and GST-dMT.	66
Figure 4.26: Elution of dMT (1 <sup>st</sup> peak) and GST (2 <sup>nd</sup> peak) proteins from GSTrap FF affinity column.	67
Figure 4.27: SDS-PAGE analysis of cleaved dMT and GST recombinant proteins. Sample numbers indicate different purification batches.	68
Figure 4.28: Native PAGE analysis of cleaved dMT and GST recombinant proteins ...	68
Figure 4.29: Figure 4.30: Elution of proteins (Table 4.6) used in the calibration of size exclusion column. Albumin (1 <sup>st</sup> peak), ovalbumin (2 <sup>nd</sup> peak), chymotrypsinogen (3 <sup>rd</sup> peak), Rnase A (4 <sup>th</sup> peak), aprotinin (5 <sup>th</sup> peak), and vitamin B12 (6 <sup>th</sup> peak).	70
Figure 4.30: Calibration curve for size exclusion column, the equation used for the molecular mass determination of dMT and GSTdMT is given on the chart.	70
Figure 4.31: Elution profile of the dMT size exclusion chromatography.	71

Figure 4.32: GSTdMT purification with HiTrap column. ....	72
Figure 4.33: Elution profile of the GSTdMT size exclusion chromatography. ....	73
Figure 4.34 (a) and (b): Scattering patterns for GST-dMT solutions at 1.44 and 5.9 mg/ml concentrations respectively. The buffer was 50mM HEPES, pH 8.0 and 150 mM NaCl. Note that the scattering curves indicate aggregated protein in the range of the scattering vector $0.027 < s < 0.14 \text{ nm}^{-1}$ . ....	74
Figure 4.35 (a) and (b): Scattering patterns for GST-dMT solutions at 1.0 and 3.0 mg/ml concentrations. The buffer was 50mM HEPES, pH 8.0 and 150 mM NaCl. Note that the scattering curves indicate aggregated protein in the range $0.027 \leq s \leq 0.14 \text{ nm}^{-1}$ . ....	75
Figure 4.36: 4-20% native tris-glycine gel analysis of different fractions of column purified GST-dMT. Low and high molecular mass markers are shown in 1 <sup>st</sup> and 9 <sup>th</sup> lanes. ....	76
Figure 4.37 (a) and (b): Scattering pattern (a) and the Guinier plot (b) for the 16 kDa dMT fraction. ....	77
Figure 4.38: Rat MT (4MT2) and wheat MT protein sequences. Two metal binding domains and hinge regions are indicated. ....	79
Figure 4.39: Predicted secondary structure features for wMT. Used algorithms are indicated (right) and described in Materials and Methods. ....	80
Figure 4.40: Pairwise alignments of wheat MT alpha with sea urchin MT beta domains; and wheat MT beta with rat liver MT beta domains. ....	81
Figure 4.41: Wheat MT alpha domain (above-left) with sea urchin MT beta domain (above-right). ....	82
Figure 4.42: Wheat MT beta domain (above-left) with rat liver MT beta domain (above-right). ....	83
Figure 4.43: Predicted coordinates were processed and visualized with DeepView. ....	86
Figure 4.44: Two cluster of predicted hinge regions, one representative from each cluster shown and indicated. ....	87

Figure 4.45(a): Ribbon representation of the modeled wMT protein with the relaxed hinge region. ....	88
Figure 4.46: Multiple alignment of plant MT hinge regions. ....	89
Figure 4.47: A rooted tree for plant MT hinge regions based on sequence similarities. ....	90
Figure 4.48: Multiple alignment of selected plant MT hinge regions. ....	91
Figure 4.49: Hydrophobicity diagram for wMT according to Kyte and Doolittle approximation with corrected amino acid scale to 0 to 1. ....	97
Figure 4.50: Comparison of the mean net charge and the mean hydrophobicity of folded (open circles) and natively unfolded proteins (gray circles) (Uversky, 2002). ....	98
Figure 4.51: wMT disordered regions, diagram generated by “Dunker’s Lab Predictor of Natural Disordered Regions Server”. ....	99
Figure 4.52: wMT hinge region disordered regions, diagram generated by “Dunker’s Lab Predictor of Natural Disordered Regions Server”. ....	99
Figure 4.53: <i>Synechococcus</i> metallothionein repressor (SmtB) protein, DNA binding dimer form. DNA binding alpha helices are indicated in ribbon structure.....	100
Figure 4.54: The two alpha helices of wMT aligned with those of SmtB. Similar regions are indicated with arrows. ....	101
Figure 4.55(a): Surface accessible possible DNA binding residues of wMT $\alpha$ -domain .....	102
Figure 5.1: Structure-function studies require input from several disciplines.....	105

## LIST OF TABLES

Table 2.1: Defined MT families and subfamilies (Binz and Kagi, 1997). .....	6
Table 2.2: Types, cystein motifs, and localizations of plant metallothioneins (Rausser, 1999). .....	17
Table 2.3: 27 known MT protein entries from EMB, Swiss-Prot Database (12.03.2003). .....	23
Table 4.1: Primers designed for mt gene identification on <i>T. aestivum</i> and <i>T. durum</i> genomic DNA. ....	43
Table 4.2: Temperature and magnesium concentrations that were tried during PCR to find optimum conditions. ....	43
Table 4.3: Designed primers with RE sites for pGEX-4T-2 vector. ....	56
Table 4.4: Legend for Figure 4.21 .....	59
Table 4.5: Cell types and given Cd concentration during the induction. ....	60
Table 4.6: Legend for figure 4.23. ....	63
Table 4.7: Protein samples used for the column calibration and as the low molecular mass marker in the native PAGE analysis. ....	69
Table 4.8: Elution volume and corresponding molecular mass calculated according to the calibration curve. The 3 <sup>rd</sup> column indicates; (calculated mass / dMT mass). ....	71

Table 4.9: Elution volume and corresponding molecular mass calculated according to the calibration curve. The 3<sup>rd</sup> column indicates; (calculated mass / GSTdMT mass). ... 73

Table 4.10: Detected sequence fragments, PDB file name, referring local structure, clustering group and confidence value given respectively. .... 84

Table 4.11: Proteins sharing structural similarity with the wMT hinge region. .... 93

## ABREVIATIONS

**Asn:** Asparagine

**BSA:** Bovine serum albumin

**cDNA:** Complementary DNA

**Cys:** Cystein

**FPLC:** Fast perfusion liquid chromatography

**Gln:** Glutamine

**Gly:** Glycine

**kDa:** Kilodalton

**Lys:** Lysine

***mt-a:*** Metallothionein gene from *Triticum aestivum* (bread wheat)

***mt-d:*** Metallothionein gene from *Triticum durum* (pasta wheat)

**OD<sub>x</sub>:** Optical density (absorbance) measured at “x” wavelength



**PAGE:** Polyacrylamide gel electrophoresis

**pGEMaMT:** *mt-a* gene inserted into pGEM-Teasy vector

**pGEXdMT:** *mt-d* gene inserted into pGEX-4-T2 vector

**SDS:** Sodium dodecyl sulfate

## 1 INTRODUCTION

Metallothioneins (MTs) are a group of low molecular weight (8000-10,000 Daltons) polypeptides rich in cysteine residues (25-33 %). Cysteine residues in MTs form thiol bonds with metal ions to scavenge toxic heavy metals (cadmium, mercury, etc.), to store biologically essential metals (copper and zinc) and to regulate metal dependent processes fundamental to cellular pathways (Vasak and Hasler, 2000).

MTs, present in a wide range of organisms from fungi to mammals, have been conventionally classified into three groups: class I MTs are those with sequences similar to mammalian renal MTs, class II are all other MTs and class III consists of phytochelatins; enzymatically synthesized polypeptides that bind metals in plants (Rausser, 1999). More recent sequence analyses using computational analyses have shown the diversity of MT proteins and a more detailed classification scheme based on the number and location of cysteine residues has been proposed (Coyle *et al.*, 2002; Vasak and Hasler, 2000).

Plant MTs, comprise a very large family of proteins and are difficult to classify (Yu, *et al.*, 1998; Yeh, *et al.*, 1995). These are mainly grouped into class II MTs, and ~65 genes have so far been identified as corresponding to MT-like proteins (Clemens, 2001; Rausser, 1999). However, presence of Class I-like plant MTs, e.g. in wheat early cysteine labelled (EC) protein and in rice and barley have also been reported (Rausser, 1999). Early cysteine labeled protein was one of first MTs to be isolated from wheat and maize (Kawashima *et al.*, 1992; Ma *et al.*, 2003), and the metal binding domains of the wheat MT gene and rat liver MT gene show 75% sequence similarity (Kawashima *et al.*, 1992; Braun *et al.*, 1992).

Although the exact role of MTs in metal detoxification and scavenging in plant systems is not clear, both monocots and dicots have MT-like protein genes (Yu *et al.*, 1998), and evidence, at the level of gene expression, for MT-protein synthesis and metal detoxification has been reported (Briat and Lebrun, 1999). Plant MTs, probably with some other mechanisms, maintain internal concentrations of essential metals between limits of deficiency and toxicity, and of nonessential metals below their toxicity thresholds. There are also reports in the literature of involvement of NO in metal binding to MT (Katakai *et al.*, 2001; Zangger *et al.*, 2001) and an antioxidant role for MT has been proposed (Ebadi *et al.*, 1996).

No structural studies are reported in the literature on plant MTs. There exists, however, 30 PDB entries for MTs from various organisms. Studies on MTs in general indicate that the protein structure is stabilized by metals and that there are no detectable secondary structural features, which stabilize the whole protein in the absence of metals. This appears to be a characteristic feature of also some DNA binding proteins (or regions of proteins) such as transcription factors (Capoli *et al.*, 2001) and zinc fingers (Blindauer *et al.*, 2001).

In this thesis isolation and characterization of a novel metallothionein gene from *Triticum durum* (*mt-d*), which shows over 90% DNA similarity with *Triticum aestivum* metallothionein gene (*mt-a*) (Kawashima *et al.*, 1992) is reported. *mt-d* has been expressed using pGEX-4T-2 vector system (Amersham Biosciences) as a recombinant GST-fusion protein in a prokaryote, *Escherichia coli*(*BL21*). The recombinant protein has been purified and characterized using biochemical methods. Three dimensional structure of wheat MTs have been predicted using detailed sequence similarity analyses with structurally known MT proteins and heuristic fragment assembly approach. Preliminary X-ray scattering measurements have been carried out to verify the predicted structural model.

These detailed and multi-approach studies on plant MTs showed that heavy-metal detoxification and/or essential-metals regulation need not be the only metabolic activities of plant metallothionein proteins. Possible DNA binding and protein-protein

interaction functions according to computational analysis may help to classify these proteins as regulatory proteins like transcription factors, transcription suppressors, etc.

## 2 OVERVIEW

### 2.1 Metallothionein (MT) Proteins

#### 2.1.1 General information

Metallothionein (MT) superfamily includes low molecular weight, intracellular metal-binding proteins that are found in almost all organisms. The first MT was isolated from horse kidney in 1957 by Margoshes and Vallee (Margoshes & Vallee, 1957). Since then MTs have been isolated from various organisms including plants, vertebrates, invertebrates, fungi, unicellular eukaryotes and some prokaryotes (Coyle *et al.*, 2002).

These proteins have a large number of cysteine residues which form thiolate bonds with transition metals ( $d^{10}$  metal ions) stabilize the protein 3D structure and result in high metal content. MTs have been isolated from different organisms in bound forms to Cd, Cu and Zn; on the other hand, they can also bind Hg, Pt, Bi, Ag, and Au *in vitro* experiments (Vasak & Kagi, 1994). Their capacity to bind both the essential and non-essential metals point to another function of these proteins in heavy-metal detoxification in addition to regulation of the biological activities of essential trace elements.

### **2.1.1.1 Nomenclature of MT**

The plenum of the First International Meeting on Metallothionein and Other Low Molecular Weight Metal-Binding Proteins in 1978 generated the first nomenclature of MTs. The adapted version was presented in 1985 by the Committee on the Nomenclature of Metallothionein during the Second International Meeting on Metallothionein and Other Low Molecular Weight Metal-Binding Proteins (Kojima *et al.*, 1997).

Characteristics of the first protein that was isolated from horse kidney were as follows: low molecular weight, high metal content, high cysteine content, no aromatic residues, no histidine residues, unique cysteine residue distribution, spectroscopic features of mercaptides, and metal thiolate cluster. According to these features committee made a definition for MT proteins in 1985:

“Polypeptides resembling equine renal metallothionein in several of their features can be designated as “metallothionein”” (Kojima *et al.*, 1997).

### **2.1.1.2 Classes and types of MT**

The committee established in 1985 decided to divide the metallothionein superfamily of proteins into 3 classes. Class-I MTs include all proteins that share a similar cysteine distribution with the horse kidney or mammalian MTs. Class II MTs, are those that have characteristics of MT without a similar cysteine residue distribution throughout the protein. Finally, Class-III MTs include all other similar polypeptides that are enzymatically synthesized. Today the search term “metallothionein” will return with 4721 nucleotide and 1113 protein sequences in the NCBI Database during a normal keyword search. Since 1957, when the first horse kidney metallothionein was isolated, or even since 1985 protein and nucleotide sequences in databases have increased logarithmically, and so did MT sequences.

The use of heuristic algorithms and powerful computers enable us to deal with the thousands of protein/nucleotide sequences. Using multiple sequence analysis software all MT sequences could be aligned and placed into a phylogenetic tree. At this point should the important consideration be the taxonomic relations (how source organisms are related to each other evolutionarily) or just the properties of their MT proteins (cystein residue distribution, amino acid composition, length of the protein, etc.)?

Table 2.1: Defined MT families and subfamilies (Binz and Kagi, 1997).

<b>Family 1:</b>	<b>vertebrate MTs</b>	<b>Family 7:</b>	<b>ciliata MTs</b>
m1:	mammalian MT-1	c1:	ciliate MT
m2:	mammalian MT-2	<b>Family 8:</b>	<b>fungi-I MTs</b>
m3:	mammalian MT-3	f1:	fungi-I MT
m4:	mammalian MT-4	<b>Family 9:</b>	<b>fungi-II MTs</b>
m:	n.d. mammalian MT	f2:	fungi-II MT
a1:	avian MT-1	<b>Family 10:</b>	<b>fungi-III MTs</b>
a2:	avian MT-2	f3:	fungi-III MT
a:	n.d. avian MT	<b>Family 11:</b>	<b>fungi-IV MTs</b>
b:	batracian MT	f4:	fungi-IV MT
t:	teleost MT	<b>Family 12:</b>	<b>fungi-V MTs</b>
<b>Family 2:</b>	<b>mollusk MTs</b>	f5:	fungi-V MTs
mo1:	mussel MT-1	<b>Family 13:</b>	<b>fungi-VI MTs</b>
mo2:	mussel MT-2	f6:	fungi-VI MTs
mog:	gastropod MT	<b>Family 14:</b>	<b>prokaryota MTs</b>
mo:	n.d. mollusk MT	pr:	prokaryota MT
<b>Family 3:</b>	<b>crustacean MTs</b>	<b>Family 15:</b>	<b>planta MTs</b>
c1:	crustacean MT-1	p1:	plant MT type 1
c2:	crustacean MT-2	p2:	plant MT type 2
c:	n.d. crustacean MT	p2v:	plant MT type 2 variant, described as a clan of p2
<b>Family 4:</b>	<b>echinodermata MTs</b>	p3:	plant MT type 3
e1:	echinodermata MT type 1	p21:	plant MT type 2x1
e2:	echinodermata MT type 2	pec:	plant EC MT-like protein
<b>Family 5:</b>	<b>diptera MTs</b>	<b>Family 99:</b>	<b>phytochelatins and other non-proteinaceous MTs</b>
d1:	diptera MT type 1		
d2:	diptera MT type 2		
<b>Family 6:</b>	<b>nematode MTs</b>		
n1:	nematode MT type 1		
n2:	nematode MT type 2		

#### 2.1.1.2.1 Cystein residue distribution or source organism

Two main features of MT proteins are their high cystein residue content and the unique distribution of these residues in the protein structure. Two major types of cystein distributions can be seen in the MT super-family. In one type cystein residues form two distinct and easily observable domains like in the case of horse kidney MT, which are named as metal binding domains. The number of residues in the region that connects the two metal binding domains greatly varies among organisms. The second type of MT proteins have the cystein residues distributed relatively equally throughout the sequence, without forming distinct domains when compared to the first type. Differences between MT proteins coming from different source organisms lie in their cystein distribution pattern; x-x-Cys-x-Cys-x-x, or x-Cys-Cys-x-x-Cys-x, etc., in both of the two types.

Although cystein distribution is the key feature, the evolutionary connections between organisms should also be considered as Theodosius Dobzhansky (1900-1975) says “Nothing in biology makes sense except in the light of evolution”. A detailed and general classification of the metallothionein superfamily was carried out by Binz and Kagi (Binz & Kagi, 1997). This classification is available at; <http://www.unizh.ch/~mtpage/classif.html>.

As stated in 1985, **metallothionein superfamily** contains any polypeptide that resembles horse kidney MT in several of their features as stated above. The **family**, **subfamily**, **subgroup** and **isoform** form other steps in the hierarchical system of this classification (Table 2.1, Figure 2.1). In the family proteins are thought to be evolutionarily related and share a particular set of sequence specific properties. The subfamily defines more sequence specific properties like the conservation of repetition sequences in the gene, the resemblance of non-coding regions in the genome. Some clearly distinguishable branches are then formed in the re-constructed phylogenetic tree and named as subgroup (Binz and Kagi, 1997). Isoforms or allelic forms, on the other hand, are used to define all MTs occurring naturally in a single species (Kojima *et al.*, 1997). They are located on different chromosomes and their differences are due to



variations in their metal compositions and/or posttranslational acetylation (Coyle *et al.*, 2002).

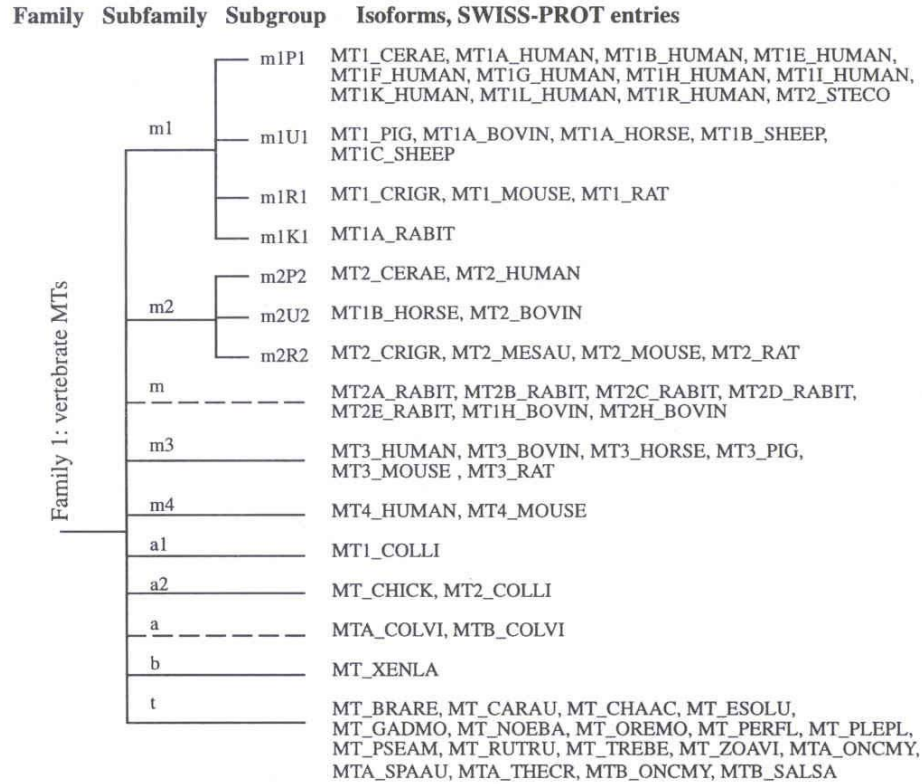


Figure 2.1: Proposed classification of the Family 1: vertebrate metallothionein proteins (Binz and Kagi, 1997).

#### 2.1.1.2.2 Human MTs

Mammalian MTs show highly conserved features both in the length of the protein and in the unique cystein residue distribution. They are generally 61-62 amino acids in length and contain 20 conserved cystein residues with a sequence identity more than 85%. Four types of MTs, MT-1 through MT-4, have been isolated and characterized both in mammalian systems and most of other vertebrates.

MT-1 and MT-2 are widely expressed in all organs and tissues of the human body. They are shown to be inducible by stress factors like glucocorticoids, cytokines, reactive oxygen species, metal ions, and increased or decreased temperature levels (Beattie *et al.*, 1996; Sato *et al.*, 1996). MT-3 is expressed mainly in the brain and is thought to be a neuronal factor (Coyle *et al.*, 2002). Although, little is known about its function, MT-4 is only abundant in certain stratified tissues (Quaife *et al.*, 1994).

#### 2.1.1.2.3 Non-mammalian MTs

MTs are also found in non-mammalian systems including invertebrates, fungi, plants, unicellular eukaryotes, and some prokaryotes (Valentine and Gralla, 1997; Tanguy *et al.*, 2001;). In recent years, MTs in other systems have been isolated and tracked. Two isoforms of MT were found in the snail *Helix pomatia*, which are thought to be specific for the cadmium detoxification and copper regulation. Indeed, these organisms attracted attention due to their high capacity to tolerate very high amount of cadmium in the soil (Dallinger *et al.*, 1997). It was known that prokaryotes have MT like protein and *mt* gene regulation similar to that of eukaryotes (Silver and Phung, 1996; Turner *et al.*, 1996). A very recent study showed the relation between the MT protein and CPx-ATPase to prevent filamentous cyanobacterium *Oscillatoria brevis* from heavy metal detoxification (Liu *et al.*, 2003).

#### 2.1.1.3 Cellular localization and function of MTs

##### 2.1.1.3.1 Mammalian systems

In humans the MT concentration has been found to be high in the liver, kidney, intestine and pancreas with a concentration range of 400 to 700 µg/g of tissue. Although the main localization of MTs in human cell is the cytoplasm, rapidly proliferating cells have a high concentration of MTs in their nucleus (Coyle *et al.*, 2002).

Obvious functions for any MT seem to be essential trace element regulation and heavy-metal detoxification. *In vitro* metal binding assays are not sufficient to show involvement of MTs specifically in protection against high concentration of metal ions; high cysteine residue content makes them accessible to any bonding that can be established. However, studies that have showed positive correlations between the increased MT expression in response to increased metal concentration, proved MTs as metal scavengers (Tanguy *et al.*, 2001; Ma *et al.*, 2003; Liu *et al.*, 2003).

MTs role in the regulation of cytoplasmic trace elements have been clearly demonstrated by two studies on nitric oxide mediated release of bound metals from the proteins. At inflammatory sites stimuli such as interleukin-1, tumor necrosis factor alpha, and lipopolysaccharide affect inducible nitric oxide synthase (iNOS) where these factors also increase the level of MT expression (Beattie *et al.*, 1996; Sato *et al.*, 1996). NO then affects the MT  $\beta$  domain causing the release of bound metals like zinc and copper where these essential metals are used by other antioxidant defense enzymes and act as coenzymes for many other vital enzymes (Katakai *et al.*, 2001; Zangger *et al.*, 2001).

Besides heavy metal detoxification, mammalian MTs have been shown to be involved in resistance against oxidative stress by scavenging free radicals (Sato and Bremer, 1993). Indeed, an enhanced sensitivity to oxidative stress has been shown in transgenic mice deficient in MT-I and MT-II genes (Lazo *et al.*, 1995). Agents such as iron, hydrogen peroxide and alcohols generate free radicals and cause oxidative stress. MT transcription level was positively affected when cells were administered such agents, although the induction mechanism is unknown (Ebadi *et al.*, 1996). A possible mechanism can be through cytokines that are produced as a result of the inflammatory response due to extensive protein damage. Cytokines such as interleukin-1, interleukin-6, tumor necrosis factor alpha, and gamma interferon may act as inducers for MT transcription and/or expression (Andrews, 2000). On the other hand, comparisons of wild type (MT(+/+)) and MT-null (MT(-/-)) mice, have clearly shown that the normal tissue levels of metallothionein do not protect mice *in vivo* against oxidative stress. Lack of metallothionein in MT-null mice did not cause any alteration in the antioxidant

defense system (superoxide dismutase, catalase, of glutathione peroxidase and glutathione levels) (Conrad *et al.*, 1997 and 2000).

MT-III has been characterized as a brain specific growth inhibitory factor and was discovered in 1988 during research aimed at understanding the pathogenesis of Alzheimer's disease. The  $\alpha$ -domain of MT-III is the functional domain for the growth inhibitory effect where the  $\beta$ -domain binds zinc or copper ions. MT-III is not inducible with agents that are known to increase MT-I and MT-II cellular levels. Indeed, MT-III competes for available zinc and copper in case of their depletion; whereas, MT-I and -II release these metals under the same circumstances indicating another function for MT-III. Although, this function is not very clear, significant downregulations of MT-III in Alzheimer's disease have been detected (Oz and Armitage, 2001; Yu *et al.*, 2001).

#### 2.1.1.3.2 **Function of MTs in systems other than mammalian**

The snail *Helix pomatia* has two isometallothionein one is mainly expressed in the midgut-gland and the other in the mantle. The mantle MT selectively binds Cu(I) needed for the biosynthesis of the oxygen-carrying protein haemocyanin. Cd, on the other hand, is accumulated in the snail soft tissues by midgut gland specific MTs. These two MT isoforms showed to be 60% similar in their amino acid compositions (Dallinger *et al.*, 1997).

## **2.2 Plant responses to metal toxicity and metal chelators produced by plants**

Organisms can not synthesize all materials that they need, and they have to fulfill this requirement from their environment. Aquatic life forms continuously filter water and terrestrial ones acquire those either by means of feeding on others or direct intake from soil. Such necessary elements for organisms are called micronutrients and they include copper (Cu), nickel (Ni), zinc (Zn), etc., and in trace amounts, these are necessary for plant survival. These elements may be present in soil in varying concentrations and other metal ions, which do not have any essential role for plants, such as cadmium (Cd), lead (Pb), and mercury (Hg) may also be present. Plants, being sessile organisms, are faced with such fluctuating conditions and they have several mechanisms that maintain internal levels of non-essential metals below toxicity and of essential metals between deficiency and toxicity.

### **2.2.1 Metal uptake by roots**

Essential and non-essential metals enter into plants via root systems and root tissue maintains metal ions up to a certain concentration depending on the plant species. Apoplast, especially in root tissue, is important in the transport and distribution of metal ions between tissues and cells. Metal uptake from the soil is selectively controlled by specific carriers located on the plasma membrane of root cells (Briat and Lebrun, 1999).

### **2.2.2 Metal storage in plants**

Once inside the plant cell, metal ions are scavenged, chelated, or stored by means of various mechanisms. Metal binding to cell wall, reduced transport across cell membranes, active efflux, intracellular compartmentalization and intracellular chelation are among the possible defence mechanisms that are dependent on plant species. MTs and ferritins are intracellular proteins that scavenge heavy-metal ions and prevent

toxicity. MTs are also thought to be involved in essential metal regulation, which will be discussed in later sections. Phytochelatins or MT-IIIs, are also intracellular polypeptides that carry out similar scavenging functions, however they are not gene products but are synthesized by enzymes. Organic acids like citrate, malate, and oxaloacetate; aminoacids like free histidine, nicotinamine (NA); and phosphate derivatives like phytate (myo-inositol hexakiphosphate) are other chelators used by plants in the defence mechanisms against metal toxicity.

#### **2.2.2.1 Metal storage proteins**

Ferritins and MTs store cellular metal ions and detoxify cytoplasm by decreasing their availability to other proteins.

#### **2.2.2.2 MT-IIIs (phytochelatins)**

Phytochelatins are generally named as “MT-III” mistakenly; they actually belong to a subgroup of class III MTs. MT-IIIs are polypeptides with repeating units of  $\gamma$ -Glu-Cys and there is no gene encoding for them. They are produced by non-ribosomal enzymes in the cytoplasm and then carried into the vacuole.

Class III MTs have 5 families of  $\gamma$ -Glu-Cys peptides. The C terminal amino acid determines the type of the peptide and can be; glycine,  $\beta$ -alanine, cysteine, serine, or glutamine. The main peptide chain is formed by 2-7 times repeating units of  $\gamma$ -Glu-Cys (Rauser, 1999).

### **2.2.2.3 Metal chelation by small molecules**

MTs, ferritins, and class III MTs scavenge free metal-ions and detoxify the cytoplasm. On the other hand, small molecules like organic acids, amino acids, and phosphate derivatives are also commonly used by plants to chelate cations (Briat and Lebrun, 1999). Citric acid, malic acid, and oxalic acid are present both in the cytoplasm and the vacuole. There are several proposed mechanisms for the organic acid scavenging system in the cytoplasm. According to mostly accepted mechanism, malate chelates zinc in the cytosol and moves it into the vacuole, where oxalate (more abundant) chelates zinc to free malate for returning to the cytosol (Rausser, 1999).

## **2.3 Plant MTs**

### **2.3.1 Types of plant MTs**

MTs in plant kingdoms are widely expressed in all tissues and the first MT-like gene in plant was described in 1990 for a copper-tolerant ecotype of *M. guttatus*. Although some have the same architecture as mammalian MTs in terms of cystein residue distribution, other types also exist. Several different classification schemes for plant MTs have been proposed (Rausser, 1999; Cobbett and Goldsbrough, 2002; Robinson, 1993).

MTs in plant kingdom are categorized into 4 types according to the cystein residue distribution pattern (Figure 2.2). Type 1 and type 2 MTs are similar to mammalian MTs in terms of the organization of cystein residues except the hinge region that connects the two metal binding domains. Different MTs that are expressed during fruit ripening belong to the type 3 plant MTs. Finally, the type 4 MTs are wheat early-cystein labeled (Ec) proteins which were the first plant MT to be characterized (Cobbett and Goldsbrough, 2002).

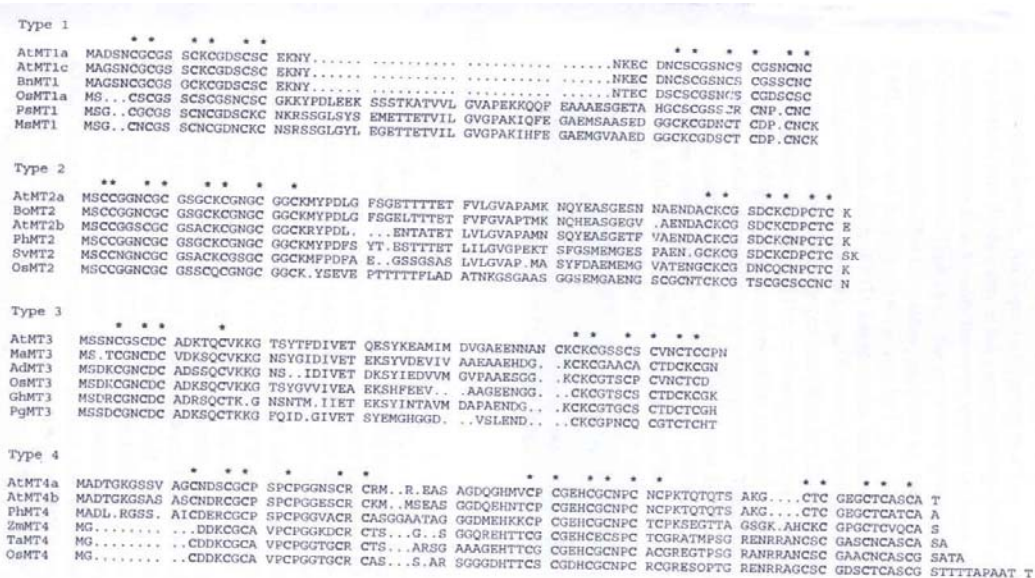


Figure 2.2: Plant metallothionein proteins' multiple alignment, showing the 4 types of MTs. Conserved cysteine residues are marked with a star. Protein sequences are derived from known gene sequences of *Arabidopsis* (At), *Brassica napus* (Bn), rice (Os), pea (Ps), alfalfa (Ms), *Brassica oleracea* (Bo), petunia (Ph), *Silene vulgaris* (Sv), banana (Ma), kiwifruit (Ad), cotton (Gh), *Picea glauca* (Pg), maize (Zm), and wheat (Ta) (Cobbett and Goldsbrough, 2002).

Type 1 MTs contain two metal binding domains and a ~40 amino acid long of hinge region that connects them. Cys-X-Cys motif is repeated three times in each of these domains summing up a total number of 6 cysteine residues per domain and 12 per protein. The type 2 MTs are similar to type 1 in terms of domain structure and they also have a hinge region of the same length. The  $\alpha$ -domain amino acid motif is highly conserved among type 2 MTs. The main difference comes from the cysteine residue distribution pattern. Here the first cysteine residues are paired "Cys-Cys" at positions 3 and 4. Another distinctive point is the Cys-Gly-Gly-Cys motif that is found at the end of the  $\alpha$ -domain (Rauser, 1999; Cobbett and Goldsbrough, 2002).

Type 3 MTs have a total of 10 cysteine residues and the difference comes from the  $\alpha$ -domain that contains only 4 cysteines. The  $\beta$ -domain contains Cys-X-Cys motifs and there are two conserved motifs as seen in figure 2.2. The hinge region of ~40 amino acid residues is again present in the type 3 MTs.



Although the hinge region that exists in all three types of plant MTs show variations among species, there are also common features. All include several aromatic amino acids, and show no stable secondary structural features. Any computer modeling attempt for the hinge region using either homology modeling and/or other heuristic methods ends with helix-turn-helix type of DNA binding structure, which is found in suppressor or repressor proteins and some transcription factors (Cook *et al.*, 1998; Giedroc *et al.*, 2001; Morita *et al.*, 2002; Blindauer *et al.*, 2001).

The wheat Ec protein is a very classical example of type 4 MTs and it is the first EC protein that was characterized in plant kingdom. Three genes for the EC protein from wheat and one from maize are well characterized (Kawashima *et al.*, 1992; White and Rivin, 1995) although several more are known from cDNA libraries (see figure 2.2).

### **2.3.2 Localization of plant MTs**

In plant systems excess cations are generally scavenged and then stored in the vacuole. Transportation of metal complexes are done by means of several specific transporter proteins that are located on the plasma and vacuole membrane (Rauser, 1999; Coyle *et al.*, 2002). MT-III's-Cd complexes are known to be found in vacuole of Cd<sup>+2</sup> treated seedlings (Robinson *et al.*, 1993). On the other hand, in the literature, there is no work stating cellular localization of plant MTs type 1, 2, and 3.

Although cellular localization of plant MTs is not clearly known, information on tissue specific expression of them is abundant. Type 4 MTs are expressed only in developing seeds and their expression is regulated by abscisic acid. Type 1 and 2 MTs are transcribed in roots, stems, leaves, flowers, fruits, and seeds (Table 2.2). mRNAs for type 1 MTs are found to be higher in roots but for type 2 MTs higher in shoots. Plants like banana, kiwi, and apple highly express type 3 MTs during fruit ripening (Cobbett and Goldsbrough, 2002).

Table 2.2: Types, cystein motifs, and localizations of plant metallothioneins (Rausser, 1999).

MT-II Genes in Plants: Predicted Number of Cys Motifs, Distribution, and Regulation											
Plant	Gene	Cys motif			Residues between domains	Transcript abundance				Regulation	Ref.
		CxC	CxxC	Other		Root	Leaf	Seed	Other		
<b>Type 1</b>											
<i>Mimulus</i>		6	0	–	39	High	Low	–	–	Down by Cu, Zn, Cd	(53)
Pea	<i>PsMT<sub>A</sub></i>	6	0	–	42	High	Low	–	–	–	(54)
Barley	<i>ids-1</i>	6	0	–	43	High	–	–	–	Up with Fe deficiency	(55)
Maize		Sequence corroborated			45	High	Low	Low	Pith	–	(56)
		6	0	–		Yes	–	–	–	–	Up with root tip excision, Up with glucose starvation
Wheat	<i>wali1</i>	6	0	–	44	High	High	–	–	Up by Al in roots, not by Cd, constitutively high in leaves	(58)
White clover		6	0	–	42	–	–	–	–	–	(59)
<i>Arabidopsis</i>	<i>MT1/MT1a</i>	6	0	1 lone C	7	High	Low	–	–	Up by Cu > Zn, Cd	(60)
		Sequence corroborated as <i>AtMT-q</i>									(61,62)
	<i>MT1c</i>	6	0	1 lone C	7	Yes	Yes	–	Siliques	–	(62)
<i>Brassica napus</i>		6	0	1 lone C	7	None	High	–	Flower	Up with leaf senescence	(64)
Rice	<i>OsMT-1</i>	6	0	–	43	High	Some	–	Cultured cells	Up by Cu, heat stress, sucrose starvation, leaf senescence, down with ABA	(65)
Cotton	<i>MT1-A</i>	Sequence corroborated			42	High	Low/nil	–	–	Evidence for two further <i>MT1</i> genes	(66)
		6	0	–							
Fava bean	<i>MT1a</i>	6	0	–	45	High	High	–	Stem	Not by Cu, Cd, Zn, Fe	(67)
	<i>MT1b</i>	6	0	–	43	–	–	–	–	–	(68)
Red fescue		6	0	–	39	–	–	–	–	–	(69)
<b>Type 2</b>											
Soybean		5	1	1 CC	41	Low	High	–	–	Down by Cu	(70)
<i>Arabidopsis</i>		5	1	1 CC	43	–	–	–	–	–	(71)
	<i>MT2/MT2a</i>	Sequence corroborated			43	Low	High	–	–	Up by Cu > Zn, Cd	(61,62)
		Sequence corroborated as <i>AtMT-k</i>									
	<i>MT2b</i>	5	1	1 CC	39	Yes	Yes	–	Siliques	Slightly up by Cu	(62)
Castor bean		5	1	1 CC	42	–	–	–	–	–	(72)
Fava bean		5	1	1 CC	39	Least	High	–	Flower Trichomes	Down by cold, salt, salicylic acid not by Cu, Zn, Cd	(73)
Kiwi fruit	<i>pKIWI504</i>	5	1	1 CC	40	None	–	–	Fruit	Up early in fruit development	(74)

Table 2.2 (continues): Types, cystein motifs, and localizations of plant metallothioneins (Rausser, 1999).

Coffee		5	1	1 CC	42	-		From young leaves	-	(75)	
Chinese cabbage		5	1	1 CC	42	-		From the inflorescence	-	(76)	
<i>Sambucus</i>		5	1	1 CC	39	-	Yes	Abscission zone	Up with ethylene, leaflet senescence	(72)	
Tobacco		5	1	1 CC	37	-	-	-	Modulated by cytokinin	(78)	
		5	1	1 CC	42	-	Yes	-	Up with Cu, virus, wounding	(79)	
White clover		5	1	1 CC	39	-	-	-	-	(80)	
<i>Brassica campestris</i>		5	1	1 CC	42	-	-	Shoot apex	Up with vernalization	(81)	
Rice	<i>OsMT-2</i>	5	1	1 CC	43	Low	High	Cultured cells	Up by heat shock, sucrose starvation	(82)	
Tomato									Down by Cu, Cd	(82)	
(three sequences)		5	1	1 CC	35,36,43	Yes	-	-	-	(83)	
	<i>LeMT<sub>A</sub></i>	5	1	1 CC	35	Low	High	-	-	(84)	
	<i>LeMT<sub>B</sub></i>	5	1	1 CC	43	Low	High	-	-	(84)	
<i>Brassica napus</i>		5	1	1 CC	19	-	Yes	-	-	Up with senescence	(85)
<i>Brassica juncea</i>											
(5 sequences)		5	1	1 CC	42,43	Yes	-	-	-	(86)	
Apricot		5	1	1 CC	40	-	-	-	Fruit	(87)	
Common ice plant		5	1	1 CC	42	-	-	-	-	Confers Cu tolerance	(88)
Type 3											
Rice		4	2	2 CC 1 lone C	40	-	-	-	-	(89)	
		5	1	2 CC 1 lone C	37	-	-	-	Stem	(90)	
		5	1	2 CC 1 lone C	37	-	-	-	-	(91)	
Type 4											
Kiwi fruit	clone 503	4	0	2 lone C	32	None	-	-	Fruit	Up late in fruit development	(74)
Apple		4	0	2 lone C	34	-	-	-	-	Up with cold storage	(92)
Papaya		4	0	2 lone C	33	-	-	-	-	(93)	
Banana	clone 3-6	4	0	2 lone C	34	No	Yes	-	Fruit	Up with fruit ripening	(94)
Rice (two sequences)		4	0	2 lone C	30,33	-	-	-	-	(95)	
Sweet cherry		4	0	2 lone C	34	-	-	-	Fruit	(95a)	
Others											
<i>Arabidopsis</i>	<i>MT1b</i>	4	0	2 lone C	7	-	-	-	-	Inactive gene	(62)
<i>Arabidopsis</i>	<i>MT3</i>	4	0	4 lone C	34	-	-	-	Seedlings	Up by Cu	(96)
Tomato		5	-	3 lone C	40	Yes	-	-	-	(97)	
<i>Brassica campestris</i>		4	1	1 CC 1 lone C	42	-	-	-	Shoot apex	Up with vernalization	(81)
Douglas fir	<i>PM 2.1</i>	1	0	3 lone C	38	-	-	Low	-	Maximal at midembryogenesis	
										Up by osmoticum, ABA, Zn	(98)
Strawberry	<i>FMET1</i>	5	1	1 CC 1 lone C	40	-	-	-	Fruit	(99)	
Banana	clone 3-23	5	1	2 lone C	41	No	Low	-	Fruit, corm	Down with fruit ripening	(94)

Number and type of cysteine motifs are shown with x denoting any amino acid other than Cys. The Cys-rich domains 1 and 2 are separated by a Cys-deficient region varying in the number of amino acid residues. The listing for a type of Cys motif is chronological.

### 2.3.3 Functions of plant MTs

Plant MTs, due to their cysteine residues, are thought to be primarily involved in detoxification and homeostasis of metal ions in the cytoplasm like the mammalian counterparts. This theory has also been proven by showing the expression of Arabidopsis MT gene in the phloem tissue, although they have been characterized in phloem tissue of other plant species (Cobbett and Goldsbrough, 2002).

Expression of a plant MT in another system like *E. coli*, a prokaryotic organism does not give a clear answer (Kille *et al.*, 1991). Although such an approach is suitable for protein isolation, *in situ* role of MT remains unknown.

On the other hand, in plant systems there are other proteins and molecules to scavenge and chelate metal ions. Cadmium is easily chelated by phytochelatins and excess iron is scavenged by ferritins. Zinc ions are transported to the vacuole by forming complexes with organic acids and phosphate derivatives (Briat and Lebrun, 1999; Rauser, 1999). MTs are not the only way for a plant to protect itself from metal toxicity, in fact it may be costly for a plant to produce MT type 1 and 2 proteins instead of organic acids, phosphate derivatives, or phytochelatins.

## 2.4 Isolation and purification of plant MTs

Mammalian and fungi MTs are widely characterized both biochemically and structurally and most of the information on plant MTs is obtained from the work on mammalian system (Tanguy and Moraga, 2001; Blindauer *et al.*, 2001; Sayers *et al.*, 1999; Huang *et al.*, 2002).

Isolation of MTs directly from plant tissue is almost impossible as the proteins are readily oxidized and have very small molecular weight, 6-10 kDa. The most convenient way is to clone a plant MT gene into a host and then to purify the over-expressed protein (Kille *et al.*, 1991). Using a fusion protein provides a convenient way for protection from protease attacks, for stabilization of the structure *in vitro* and for simplifying purification procedures (Evans *et al.*, 1992). Indeed, during these procedures, detection and handling of MT protein will be much easier (Huang *et al.*, 2002; Yu *et al.*, 2002).

## 2.5 Structural characteristics of MT protein

Cystein residues in MTs form thiol bonds with metal ions stabilizing the structure of the whole protein. It appears that there is no average solution structure for MTs in their metal free states; apo-MT (Zangger *et al.*, 2001). There are two main structures (Figure 2.3) that have been proposed for metallothioneins based on their cystein residue distribution. In the rat MT cystein residues are distributed forming 2 domains and there are 2 lysine residues in the hinge region. Such kind of a cystein arrangement generates a 2 domain structure with an elongated 3D shape for the protein. On the other hand, yeast MT has cystein residues scattered into the whole protein generating a relatively more globular single domain structure (Sayers *et al.*, 1999; Furey *et al.*, 1986).



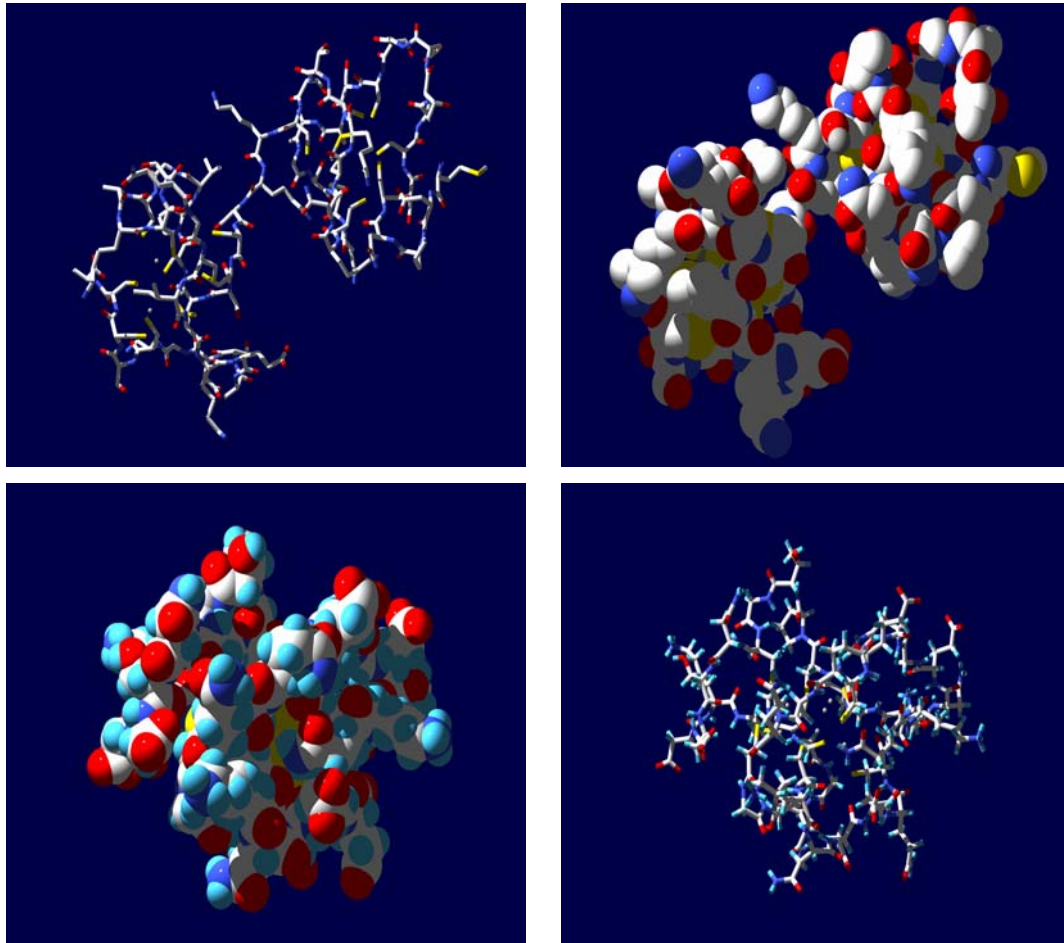


Figure 2.3: Crsytal structures of rat liver Cu-metallothionein, 4MT2 (above) and NMR structures of Cu-metallothionein of *Saccharomyces cerevisiae*, 1AQR (below).

### 2.5.1 Thiol bonds stabilize the protein structure

Cystein residues in MTs form thiol bonds with metal ions and their metal binding properties are dependent on the ionic characteristics of the metal.  $\text{Cd}^{+2}$  shows tetrahedral coordination geometry, whereas  $\text{Zn}^{+2}$  exhibits trigonal geometry in binding to MTs (Munoz *et al.*, 2000; Oz *et al.*, 2001).  $\text{Cu}^{+1}$  and  $\text{Ag}^{+1}$  have also different coordination properties. While  $\text{Cu}^{+1}$  favors trigonal coordination,  $\text{Ag}^{+1}$  prefers bicoordination (Bertini *et al.*, 2000).

Bond geometry and lengths for  $\text{Cd}^{+2}$  in the tetrahedral coordination have been well characterized. In the chemically synthesized native  $^{113}\text{Cd}_3\text{S}_9$  beta domain of *Lobster* MT,  $\text{Cd}^{+2}$  was shown to be tetraordinated to four cysteines with some cysteines bridging between two  $\text{Cd}^{+2}$  ions (Munoz *et al.*, 2000; Riek *et al.*, 1999). While three cysteines of the tetrahedral tend to be close ( $\sim 2\text{-}3$  Å) the fourth one stabilizes the geometry from a distance of  $\sim 4$  Å. This type of information on bonding properties of cysteines and metal ions, is used during homology modeling of unknown MT structures.

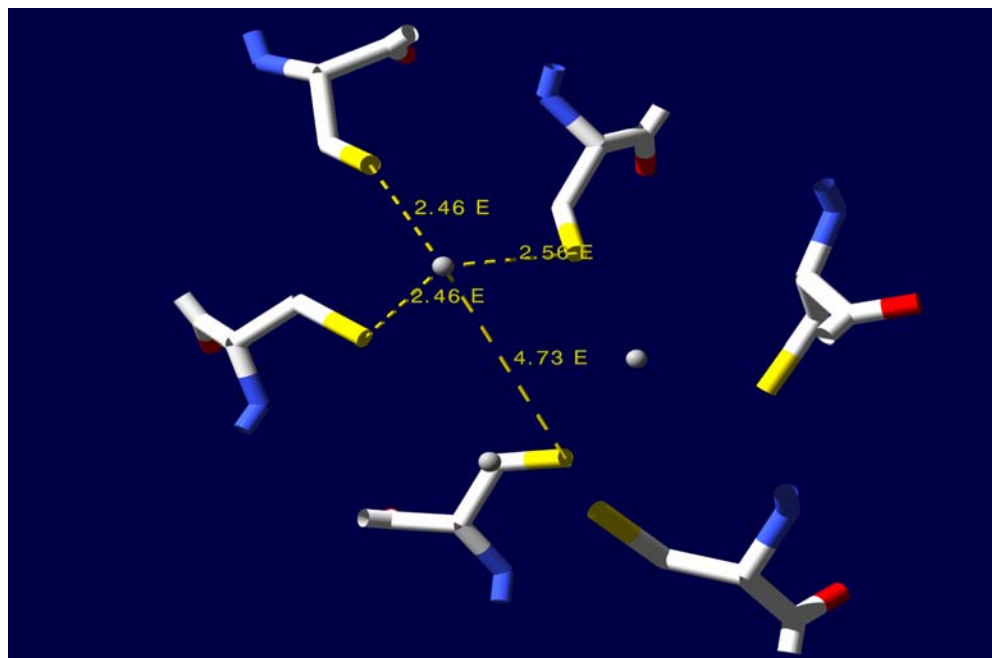


Figure 2.4: Tetrahedral coordination geometry and S-Cys distances for Cd1 in sea urchin metallothionein beta domain, 1QJL\_A.

## 2.5.2 Known MT structures

Swiss-Prot Database contains 27 entries for MT proteins structures (Table 2.3), which have been obtained using small angle X-ray solution scattering, X-ray diffraction and nuclear magnetic resonance (NMR) techniques. Crystallization of MTs is a challenging task due to the high number of cysteine residues in a relatively small molecular weight and oxidation sensitivity of the protein (An *et al.*, 1999; Melis *et al.*, 1983; Robbins *et al.*, 1991). Besides, due to high salt and varying pH conditions used for crystallization, proteins in crystals may not reflect their native *in vivo* structure. Solution X-ray scattering or NMR provide good alternatives, as these two techniques do not force protein structures. Small angle X-ray solution scattering, with good resolution, can reveal secondary structure arrangements, but atomic interactions and localizations will still be questionable. NMR, on the other hand, is only applicable in the presence of NMR-active metals such as  $\text{Cd}^{+2}$  and  $\text{Ag}^{+2}$ , and for small molecular weight proteins. Silver substitution is generally performed for proteins with metal ions other than these two (Peterson *et al.*, 1996).

Table 2.3: 27 known MT protein entries from EMB, Swiss-Prot Database (12.03.2003).

<b>1J5M</b> <b>Solution Structure Of The Synthetic 113cd_3 Beta_n Domain Of Lobster Metallothionein-1</b> [19479]
<b>1J5L</b> <b>Nmr Structure Of The Isolated Beta_c Domain Of Lobster Metallothionein-1</b> [19478]
<b>1JI9</b> <b>Solution Structure Of The Alpha-Domain Of Mouse Metallothionein-3</b> [17228]
<b>1JJD</b> <b>Nmr Structure Of The Cyanobacterial Metallothionein Smta</b> [17143]
<b>1FMY</b> <b>High Resolution Solution Structure Of The Protein Part Of Cu7 Metallothionein</b> [14427]



**1DFT**  
**Solution Structure Of The Beta-Domain Of Mouse Metallothionein-1**  
**[12320]**

**1DFS**  
**Solution Structure Of The Alpha-Domain Of Mouse Metallothionein-1**  
**[12319]**

**1QJL**  
**Metallothionein Mta From Sea Urchin (Beta Domain)**  
**[11258]**

**1QJK**  
**Metallothionein Mta From Sea Urchin (Alpha Domain)**  
**[11257]**

**1A00**  
**Ag-Substituted Metallothionein From Saccharomyces Cerevisiae, Nmr, Minimized Average Structure**  
**[6836]**

**1AQS**  
**Cu-Metallothionein From Saccharomyces Cerevisiae, Nmr, 10 Structures**  
**[6788]**

**1AQR**  
**Cu-Metallothionein From Saccharomyces Cerevisiae, Nmr, Minimized Average Structure**  
**[6787]**

**1AQQ**  
**Ag-Substituted Metallothionein From Saccharomyces Cerevisiae, Nmr, 10 Structures**  
**[6786]**

**1SMT**  
**Smtb Repressor From Synechococcus Pcc7942**  
**[6866]**

**1DMF**  
**Cd-6 Metallothionein-1 (Cd-6 Mt) (Beta Domain) (Nmr, 18 Structures)**  
**[769]**

**1DME**  
**Cd-6 Metallothionein-1 (Cd-6 Mt) (Beta Domain) (Nmr, Minimized Average Structure)**  
**[768]**

**1DMD**  
**Cd-6 Metallothionein-1 (Cd-6 Mt) (Alpha Domain) (Nmr, 18 Structures)**  
**[767]**

**1DMC**  
**Cd-6 Metallothionein-1 (Cd-6 Mt) (Alpha Domain) (Nmr, Minimized Average Structure)**  
**[766]**

**4MT2**

**Metallothionein Isoform Ii**

**[3191]**

**2MRT**

**Cd-7 Metallothionein-2 (Beta Domain) (NMR)**

**[2817]**

**2MRB**

**Cd-7 Metallothionein-2a (Beta Domain) (NMR)**

**[2816]**

**2MHU**

**Cd-7 Metallothionein-2 (Beta Domain) (NMR)**

**[2810]**

**1MRT**

**Cd-7 Metallothionein-2 (Alpha Domain) (NMR)**

**[1720]**

**1MRB**

**Cd-7 Metallothionein-2a (Alpha Domain) (NMR)**

**[1709]**

**1MHU**

**Cd-7 Metallothionein-2 (Alpha Domain) (NMR)**

**[1673]**

**1HZQ**

**Isolated Beta\_C Domain of Lobster Metallothionein-1 (NMR)**

**1HZR**

**Solution Structure of the Synthetic  $^{113}\text{Cd}$ \_3 Beta\_N Domain of Lobster Metallothionein**

## **3 MATERIALS and METHODS**

### **3.1.1 Chemicals**

All chemicals were supplied by Fluka (Switzerland), Merck (Germany), Riedel de Hën (Germany), and SIGMA (USA).

### **3.1.2 Primers**

Primers were designed according to the *Triticum aestivum* cDNA for *mt* gene (NCBI; L11879) (Snowden and Gardner, 1993) and synthesized by Integrated DNA Technologies, USA. Primers with restriction sites (stated within results part) were purchased from SeqLab (Germany).

### **3.1.3 Enzymes**

#### **3.1.3.1 Restriction enzymes**

*EcoRI*, *XhoI*, *SpeI*, *BamHI*, *Sall* (Promega and Fermentas).

#### **3.1.3.2 Ligase**

T4 DNA Ligase (Promega and Fermentas)

### **3.1.3.3 Taq Polymerase**

Taq DNA Polymerase in Storage Buffer A (Promega)

### **3.1.3.4 Reverse Transcriptase**

OneStep RT-PCR Enzyme Mix (QIAGEN)

### **3.1.3.5 Commercial Kits**

PCR Core System II (Promega)

pGEM-Teasy Vector Systems (Promega)

Qiaquick<sup>®</sup> PCR Purification Kit (250) (QIAGEN)

Qiaquick<sup>®</sup> Gel extraction Kit (250) (QIAGEN)

Qiaprep<sup>®</sup> Spin Miniprep Kit (250) (QIAGEN)

QIAGEN<sup>®</sup> Plasmid Midi Kit (100) (QIAGEN)

TOPO<sup>®</sup> TA Cloning Kit (Invitrogen)

### **3.1.4 Vectors**

Maps of all vectors can be found in Appendix A.

pGEM<sup>®</sup>-T Easy (Promega)

pGEX-4T2 (Amersham Pharmacia)

pCR<sup>®</sup>-II- TOPO<sup>®</sup> (Invitrogen)

### **3.1.5 Cells**

Different *E. coli* strains containing TOP10, XL1 Blue, BL21 (DE3), BL21(DE3)pLysE, Rosetta(DE3), Rosetta(DE3)pLysS were kindly provided by EMBL, Hamburg.

### **3.1.6 Buffers and solutions**

All buffers and solutions, except those providing with commercial kits, were prepared according to Sambrook and Russell, 2001.

#### **3.1.6.1 Culture medium**

##### **3.1.6.1.1 Liquid medium**

LB (Luria-Bertani) Broth from SIGMA was used to prepare liquid culture media for bacterial growth.

### **3.1.6.1.2 Solid medium**

LB (Luria-Bertani) Broth Agar from SIGMA was used for the preparation of solid culture media for bacterial growth.

### **3.1.6.2 Buffers for gel electrophoresis**

#### **3.1.6.2.1 Denaturing PAGE**

1 X Tris-Glycine SDS

#### **3.1.6.2.2 Non-Denaturing PAGE**

1 X Tris-Glycine

#### **3.1.6.2.3 Agarose gel electrophoresis**

1 X Tris-Acetate EDTA (TAE)

1 X Formaldehyde (FA)

### **3.1.7 Sequencing**

Sequencing service was commercially provided by SEQLAB (Germany) or MWG-The Genomic Company (Germany).

### **3.1.8 Equipment**

Please see Appendix C for a complete list of all equipments that were used during this study.

## 3.2 Methods

### 3.2.1 Plant growth

*Triticum aestivum* (Bezostaja) and *Triticum durum* (Balcali, C-1252) seeds were surface sterilized with 10% H<sub>2</sub>O<sub>2</sub> for 20 minutes and then rinsed thoroughly with distilled water. Germination was done in perlite moistened with saturated CaSO<sub>4</sub> and after 4 days seedlings were transferred into pots containing nutrient solution. Day/night cycles were adjusted as 16/8 hours under continuous aeration of pots at ~25°C and each 3 days the solution was changed.

The nutrient solution was containing; 0.88 mM K<sub>2</sub>SO<sub>4</sub>, 2.0 mM Ca(NO<sub>3</sub>)<sub>2</sub>, 0.25 mM KH<sub>2</sub>PO<sub>4</sub>, 1.0 mM MgSO<sub>4</sub>, 0.1 mM KCl, 100 µM FeEDTA, 1 µM H<sub>3</sub>BO<sub>3</sub>, 0.5 µM MnSO<sub>4</sub>, 0.2 µM CuSO<sub>4</sub>.

### 3.2.2 DNA and mRNA isolation from plant

#### 3.2.2.1 Genomic DNA isolation

100 mg fresh shoot tissue of 12 days old seedlings was disrupted using a mortar containing liquid nitrogen. QIAGEN DNeasy Plant Mini Kit was then used for the genomic DNA isolation and manufacturer's protocol was followed without further modifications.



### **3.2.2.2 mRNA isolation**

A week old seedlings were given 5 and 10  $\mu\text{M}$  cadmium ( $\text{CdSO}_4$ ) and harvested 3 days after. 100 mg fresh shoot tissue was disrupted using a RNase free mortar containing liquid nitrogen. QIAGEN RNeasy Plant Mini Kit was used for the total RNA isolation from plant tissue, which after QIAGEN Oligotex® suspension was used to isolate mRNAs.

All equipments and buffer solutions using during RNA isolation procedure were RNase free. Plastic wares were rinsed with chloroform thoroughly and then autoclaved for 15 minutes at 125°C. Glass wares were first cleaned with detergent and then put into an oven for more than 5 hours at 240°C. 0.1% diethyl pyrocarbonate (DEPC) was used to prepare RNase free water. The appropriate amount of DEPC was pour into water and waited for at least 12 hours. The solution was then autoclaved for 15 minutes at 125°C. Benches and other equipments were first cleaned with chloroform and then with RNase free water. RNase free tips and tubes were used during the whole RNA isolation and RT-PCR procedure.

### **3.2.3 Bacterial cell growth**

Cells were grown overnight (12-16h) in LB Broth (Luria Bertani) medium prior to any application. LB Agar (Miller's LB agar) solid medium was used as selective and unselective solid medium for the growth of bacteria.

Protocols for liquid and solid culture growth and the other applications including competent cell preparation, glycerol stocks were done according to Maniatis *et al.*, 1989 and Sambrook J and Russell DW, 2001.

### 3.2.4 PCR and RT-PCR

#### 3.2.4.1 PCR

Recommended reaction volumes and final concentrations of the PCR Core System II components were used for PCR reaction mixture. Annealing temperatures for primers were determined experimentally (see result part).

95°C, 1 minute	
95°C, 1 minute	a total of 40 cycles
53.5°C, 1 minute	
72°C, 1 minute	
72°C, 1 minute	
22°C, hold for ~	

### 3.2.4.2 RT-PCR

QIAGEN OneStep RT-PCR enzymes and reagents were used, and concentrations for primers and template mRNA were determined according to the available manual.

50°C, 30 minutes	
95°C, 15 minutes	
94°C, 1 minute	a total of 40 cycles
53.5°C, 1 minute	
72°C, 1 minute	
72°C, 10 minutes	
22°C, on hold	

### 3.2.4.3 Purification of PCR products

PCR product was purified either from the 1.5% Agarose gel with Qiaquick<sup>®</sup> Gel extraction Kit (250) (QIAGEN) or directly with Qiaquick<sup>®</sup> PCR Purification Kit (250) (QIAGEN).

### 3.2.5 Cloning

Basic procedures were carried out according to Manniatis *et al.*, 1989.

### **3.2.5.1 Subcloning**

*mt-d* gene was amplified and subcloned into pGEM-Teasy (Promega) and pCR®-II-TOPO (Invitrogen) vectors according to the given protocols.

### **3.2.5.2 Ligation**

PCR products were ligated into pGEM®T-Easy vector (Promega) in such a way that 3:1 and 10:1 insert:vector ratios were used. The reaction mixture was incubated for 8 hours at room temperature.

PCR products were ligated into pCR® II- TOPO® (Invitrogen) vectors. Reaction mixture was incubated at least 30 min. at room temperature (~25<sup>0</sup> C) and 1µl 6X TOPO® Cloning Stop Solution was added to stop the ligation reaction.

### **3.2.5.3 Transformation**

Ligation mixtures were transformed into different endonuclease deficient strains of *E. coli*- XL1 Blue, TOP10. Transformed cells and controls were plated on appropriate antibiotic selective LB plates prepared according to the ligation vectors.

### **3.2.5.4 Colony Selection**

Positive colonies were selected and grown on liquid LB culture containing appropriate antibiotic for both preparing glycerol stocks and plasmid isolation.

### **3.2.5.5 Plasmid isolation**

Plasmid isolation was done either with Qiaprep<sup>®</sup> Spin Miniprep Kit (250) (QIAGEN) or following to the alkaline lysis protocol from Maniatis *et al.*, 1989.

### **3.2.5.6 Restriction enzyme digestion**

Purified plasmids containing *mt-d*, *mt-a*, and *mt-a*, and *mt-d* cDNAs were digested with appropriate restriction enzymes according suppliers instructions to verify the presence of corresponding genes (Enzyme/reaction mix) v/v ratio was kept at 1/10 or smaller in all digestions. cDNAs were further isolated from 1.5% agarose gel to be cloned into appropriate expression vector.

### **3.2.5.7 DNA and cDNA analysis**

Purified plasmids and digested plasmids were analyzed by agarose gel electrophoresis. Appropriate DNA markers were used for size and concentration determination. In addition, concentration and OD<sub>260/280</sub> ratio were monitored by absorption measurements.

### **3.2.5.8 Frozen stocks of cells**

Frozen stocks of *E. coli* containing different plasmids with cDNAs and genomic DNA sequences were prepared in 15% glycerol in LB with antibiotics and kept at -80°C according to the protocol from Maniatis *et al.*, 1989.

### **3.2.5.9 Sequence verification**

QIAGEN® Plasmid Mini Kit (250) and QIAGEN® Plasmid Midi Kit (100) (QIAGEN) purified corresponding plasmids containing DNA sequences were sent for sequence analysis. Plasmids were checked by restriction and electrophoretic analysis before sequencing.

### **3.2.5.10 Cloning into expression vector**

Purified subcloning vectors were digested with corresponding restriction enzymes. Resulted fragments containing necessary sites for directional cloning were then ligated into pGEX-4T-2 expression vector.

The expression vector containing *mt-d* and *mt-a* cDNAs was cloned into *E. coli* XL1-Blue cells for vector amplification and storage purposes. Then *E. coli* BL21(DE3) expression cells were transformed with the purified vector. The insert was verified by restriction enzyme digestion after each transformation step, and the insert containing multi-cloning site of the expression vector was then sequenced for nucleotide deletions and additions.

### **3.2.6 Expression and induction**

Expression of the recombinant GSTdMT proteins was performed according to the protocol from Maniatis et al.,(1989) and Sambrook and Russell (2001). *E. coli* BL21(DE3) cells containing pGEXdMT were assayed under differentiating IPTG, 0.5 to 1.9 mM, and Cd, 0.05 to 4.0 mM, concentration for the induction and expression optimization. Aliquots corresponding a total OD<sub>600</sub> of 1.4 were taken from induced cells.

Cells were pelleted and stored at  $-20^{\circ}$  C. Pellets were prepared for SDS-PAGE gel according to the protocol from Maniatis *et al.*, 1989 and Sambrook and Russell, 2001. 10  $\mu$ l of the samples were loaded on 5%-12% SDS-PAGE gels. The gels were run at 30mA constant current for 1 h and stained with Coomassie blue for protein bands were observed after destaining solution overnight.

Protein molecular weight markers (Fermentas) were used to identify the molecular weights of expressed proteins observed on the gel.

### **3.2.7 Purification of the recombinant protein**

1 and 2 liters of bacterial cells containing recombinant plasmid were induced with 0.7 mM of IPTG and harvested 4-5 hours after the induction. The starting point for the induction was determined as  $OD_{600} = 1.2$  to 1.4.

#### **3.2.7.1 Batch purification**

Batch purification using Glutathione Sepharose 4 Fast Flow® (Amersham) medium was performed according to the supplied protocol without further modifications with an exception at the last elution step. 6 successive elutions were performed instead of 3.

#### **3.2.7.2 Column purification**

Bacterial cell lysate was loaded onto the 5 ml GSTrap® FF prepacked affinity column (Amersham) using a benchtop peristaltic pump having a constant flow rate of 1 ml per minute. Two volumes of bacterial lysate were run through the column to ensure complete binding of the recombinant GSTdMT proteins.

The GSTrap column was then connected to an AKTA®-FPLC system during the washing and elution steps which were performed according to the supplied manual. The only difference was at the elution step, in which, 20 mM reduced glutathione was used instead of 10 mM.

Elutions of both batch and column purifications were then analyzed using denaturing and denaturing PAGE, and spectrum analysis at 280 nm for the purification efficiency.

### **3.2.8 Cleavage of GSTdMT by thrombin protease**

Purified recombinant GSTdMT proteins were cleaved both on the column and in the solution following the supplied protocol by manufacturer.

### **3.2.9 Size exclusion**

#### **3.2.9.1 Column calibration**

Amersham Biosciences HiLoad® 26/60, 75 prep grade column was calibrated using a set of protein (Table 4.7).

#### **3.2.9.2 Analysis dMT and GSTdMT preparations**

dMT and GSTdMT recombinant proteins was run on Amersham Biosciences HiLoad® 26/60, 75 prep grade column with 2 different buffer; 1X PBS Sigma®, pH 7.5, 150 mM NaCl and 50 mM Tris-HCl, 150 mM NaCl. The system was connected to a fraction collector and 1, 1.5, 2 ml of fractions were collected.



### 3.2.10 Solution X-ray scattering on dMT and GSTdMT

Small angle x-ray solution scattering patterns of dMT and GSTdMT were recorded on the double-focusing monochromator-mirror camera X33 (Koch and Bordas, 1983) of the EMBL in HASYLAB on the storage ring DORIS of the Deutsches Elektronen Synchrotron (DESY) using delay line readout detectors (Gabriel and Dauvergne, 1982) and standard data acquisition and evaluation systems (Boulin *et al.*, 1986). Data were collected at 4°C in 20 frames of 1 min. to monitor radiation damage. The observation range was  $0.03 < S < 0.2 \text{ nm}^{-1}$ , where  $S$  is the modulus of the momentum transfer vector ( $S = 4\pi \sin\theta/\lambda$ ,  $2\theta$  is the scattering angle and  $\lambda$  the wavelength). Data reduction, background subtraction and correction for detector response followed standard procedures (Koch, 1991) using the program SAPOKO (Svergun, D. and Koch, M.H.J., unpublished). Radii of gyration and distance distributions were calculated using the indirect transform program GNOM (Svergun, 1991).

Measurements were performed at 1 to 5 mg/ml protein concentration in either 50 mM HEPES, pH 8.0, 150 mM NaCl buffer or in PBS. No radiation damage could be detected either in the X-ray data or in the SDS gel electrophoresis analysis of the protein obtained before and after irradiation. Bovine serum albumine (BSA) was also measured as a standard at about 7 mg/ml in 50 mM HEPES, pH 8.0, 150 mM NaCl buffer.

### 3.2.11 Sequence alignments

Both pairwise and multiple alignments were done using algorithms available at the web server “Biology Workbench” (<http://biowb.sdscs.edu>). ALIGN (Pearson, ©1997) was used for pairwise alignments, where CLUSTALW (Thompson *et al.*, 1994) was preferred for multiple alignments. Both algorithms were run with their default parameters.

### **3.2.12 Modeling**

Sequence similarity based homology modeling using aligned protein sequences was performed with “DeepView – The Swiss-Pdb Viewer, V.3.7” which is available through the web site; <http://www.expasy.org/spdbv>

wMT hinge region, on the other hand, was modeled using ROSETTA (<http://honduras.bio.rpi.edu/~sites/hmmstr/>).

### **3.2.13 Motif search**

Motif search for wMT hinge region was done in PROSITE database (<http://us.expasy.org/prosite/>) both selecting the “ScanProsite” and “MotifScan” options.

## 4 RESULTS

### 4.1 Genomic DNA isolation from *Triticum aestivum* and *Triticum durum*

Bezostaja and Balcali are two subcultivars of *Triticum aestivum* and *Triticum durum*, relatively. Genomic DNA from both subcultivars was isolated using Qiagen® DNeasy Plant Mini Kit following manufacturers' instructions without any modification. The yield of DNA was 9 µg from 100 mg of fresh plant shoot tissue.



Figure 4.1: Agarose gel electrophoresis analysis of isolated genomic DNA of *Triticum aestivum*, Bezostaja (left) and *Triticum durum*, Balcali (right).

## 4.2 Amplification of the target metallothionein (*mt*) gene

The same primers (Table 4.1), designed according to the known sequence of *T. aestivum*, were also used for amplification of *mt* gene both from Bezostaja and Balcali.

Table 4.1: Primers designed for *mt* gene identification on *T. aestivum* and *T. durum* genomic DNA.

Oligo 1	5'-ATGTCTTGCAACTGTGGA-3'	forward primer
Oligo 2	5'-TTAACAGTTGCAGGGGTT-3'	reverse primer with stop codon
Oligo 3	5'-ACAGTTGCAGGGGTTGCA-3'	reverse primer without stop codon

Three different  $Mg^{+2}$  and three different temperature, as shown in table 4.2, were tried to find the optimum PCR conditions for *mt-a* and *mt-d*.

Table 4.2: Temperature and magnesium concentrations that were tried during PCR to find optimum conditions.

		temperatures		
		51.7 °C	53.7 °C	56.4 °C
Mg <sup>+2</sup> concentrations	1.0 mM	S1	S4	S7
	1.5 mM	S2	S5	S8
	2.0 mM	S3	S6	S9

PCR results showed that the optimum annealing temperature is 53.7 °C and magnesium concentration is 1.5 mM for our primer and equipment (Figure 4.2).

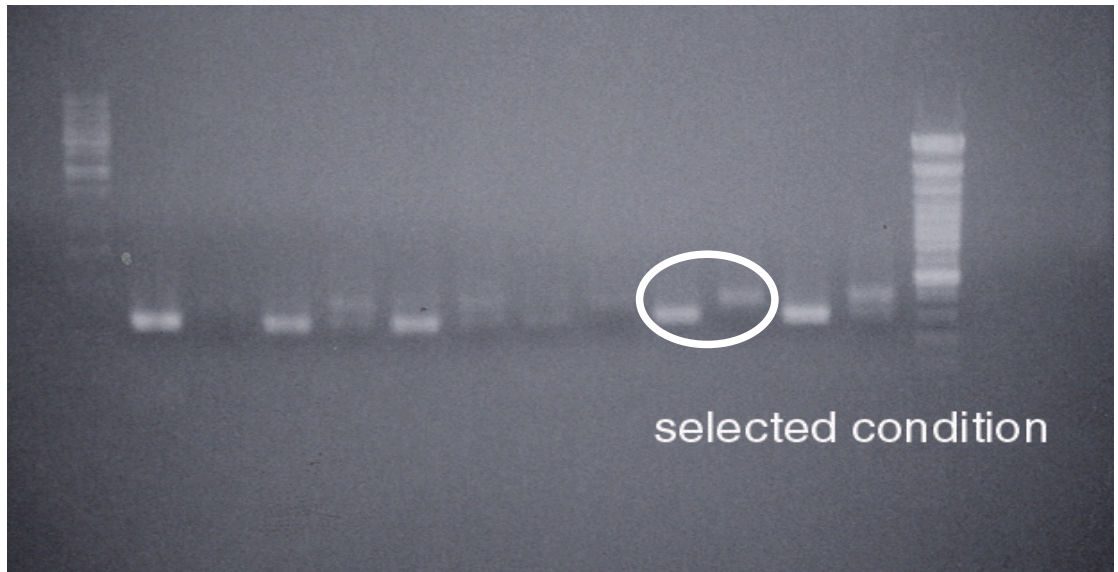


Figure 4.2: Results of optimization studies on PCR conditons. Magnesium ion concentration and annealing temperature were varied as described in table 4.2. from left to right. PCR control (left) and the amplified *mt* gene (right) indicated.

PCR was carried out at 55 °C and products corresponding to ~450bp length were observed for both (Figure 4.3).

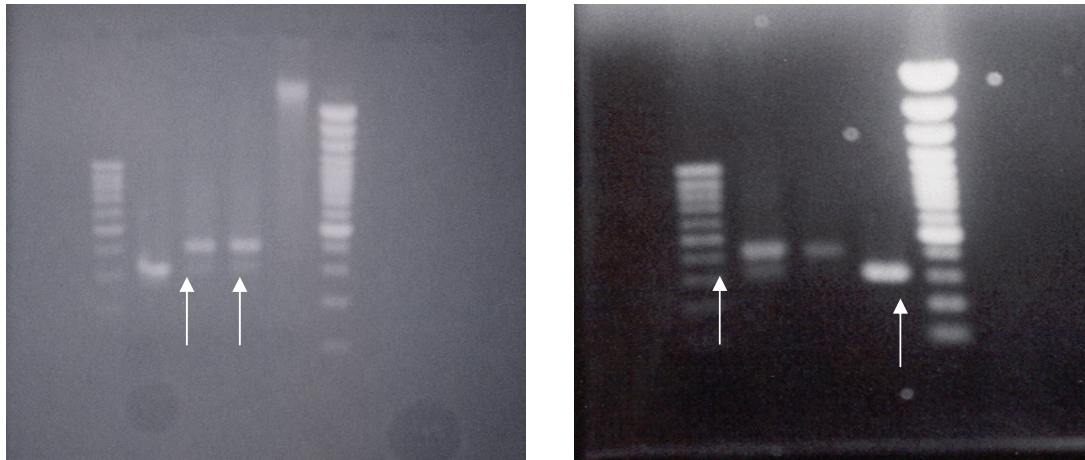


Figure 4.3: ~ 450bp long PCR products of metallothionein gene from *T. aestivum* (*mt-a*) (left) and *T. durum* (*mt-d*) (right).

### 4.3 Cloning of *mt-a* and *mt-d* in *E. coli*

PCR products of *mt-a* were ligated into pGEM-Teasy vector which makes use of the 3'-A overhangs produced by *Taq* DNA Polymerase. *mt-d*, on the other hand, was ligated into TOPO-II-PCR vector. *E. coli* Top10 cells were then transformed with the constructs into where plasmid multiplication was achieved. The pGEM-Teasy+*mt-a* construct was digested with *Bam*HI and *Sal*I to check the insert, *mt-a*. For the control digestion of *mt-d*, *Eco*RI was used. The constructs were sequenced for verification of *mt-a* sequence and determination of the *mt-d* DNA sequence.

### 4.4 Characterization of *mt-a* and *mt-d* genes

Repeated adenine and thymine nucleotides especially in the intron sites of both *T. durum* and *T. aestivum* is a characteristic for metallothionein genes. Pairwise alignment of *mt-a* and *mt-d* shows a similarity of 95% (Figure 4.4). *mt-a* is 400 bp long, whereas *mt-d* is 416 bp; and the difference comes from the repetitive “TTTTA” part in the intron part of the *mt-d* gene.

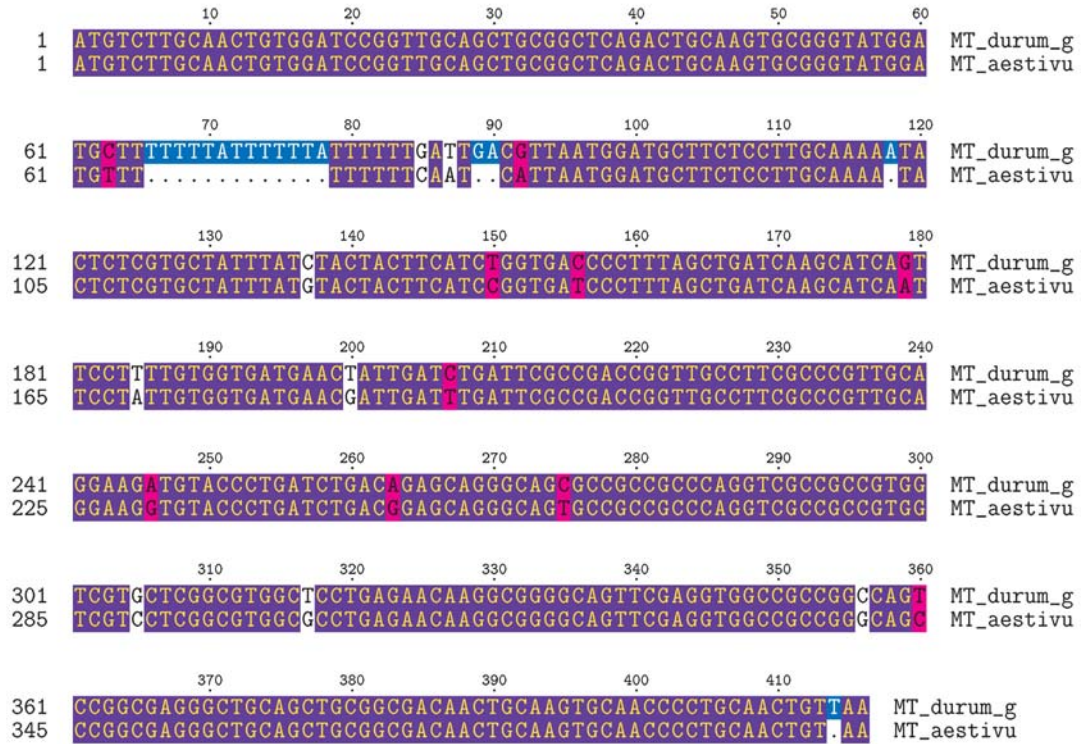


Figure 4.4: Pairwise alignment of *T. aestivum* and *T. durum* genomic MT gene sequences.

There is no available metallothionein genomic sequence for *T. aestivum* or *T. durum* so alignment with maize (*Zea mays*) genomic sequence (NCBI entry #: S57628) was performed for comparison. Differences are observed in the intron regions between wheat and maize (Figure 4.5). However, the three show high level of similarity at the level of protein (Figure 4.6).



```

      10      20      30      40      50      60
1  ATGCTTTGCAACTGTGGATCCGGTTGCAGCTGCGGGCTCAGACTGCAAGTGCGGGTATGGA MT_aestivu
1  ATGCTTTGCAACTGTGGATCCGGTTGCAGCTGCGGGCTCAGACTGCAAGTGCGGGTATGGA MT_durum_g
1  ATGCTTTGCACTGCGGATCAAGCTGCGGCTGCGGGCTCAAGCTGCAAGTGCGGGTAATA maize_MT_g

      70      80      90      100     110     120
61 TGTTTT.....TTTTCAAT...CATTAATGGATGCTTCTCCTTGCAAAAAT MT_aestivu
61 TGCCTTTTTTATTTTATTTTTCATTGA.CGTTAATGGATGCTTCTCCTTGCAAAAAT MT_durum_g
60 TATAATA.....ATAATAAGTGCAACCGTGCATGATTAATTTCTCCAGCCTTCTT maize_MT_g

      130     140     150     160     170     180
104 ACTCTCGTG.CTATTTA..TGTACTACTTCATCCGGTGATCCCTTTAGCTGATCAAGCA. MT_aestivu
120 ACTCTCGTG.CTATTTA..TCTACTACTTCATCTGGTGAACCCCTTTAGCTGATCAAGCA. MT_durum_g
110 CTTGTCTTCTCTAGTTAATTCCTTCTTTATTATTTTTTCCATT.GCAAAACAACAATA maize_MT_g

      190     200     210     220     230     240
160 .TCAATTCCTATTGTGGTGATGAACGATTGATTGATTCGCCGACCGGTTGCCTTCGCCC MT_aestivu
176 .TCAGTTCCTTTTGTGGTGATGAACATTGATCGATTCCGCCGACCGGTTGCCTTCGCCC MT_durum_g
169 ACAAATAAAAGTAAATCTGGATCGAGTAGTTCAATCCATTTCGCCGCTGCTCC..... maize_MT_g

      250     260     270     280     290     300
219 GTTGCAAGGAGCTGTACCCTGATCTGACGAGCAGGGCAGTGGCCCGCCAGGTGCGCG MT_aestivu
235 GTTGCAAGGAGATGTACCCTGATCTGACAGAGCAGGGCAGTGGCCCGCCAGGTGCGCG MT_durum_g
224 .TTTCAGCAAGAAAGTACCCTGACTGAGAGACAGCA...CCGCCGCGCAGCCACC. maize_MT_g

      310     320     330     340     350     360
279 CCGTGGTCTGCTCCTCGGCGTGGCGCCTGAGAAACAAGCGGGGC...AGTTCGAGGTGGCCG MT_aestivu
295 CCGTGGTCTGCTCCTCGGCGTGGCGCCTGAGAAACAAGCGGGGC...AGTTCGAGGTGGCCG MT_durum_g
279 .....GTCGTCCTCGGCGTGGCCCGGAGAAAGAAGCCCGCCCGAGTTCGTCGAGGCCG maize_MT_g

      370     380     390     400     410     420
336 CCGGCAGCCCGCGAGG.....GCTGCAGCTGCGGCGAACAACTGCAAGTGCAACC MT_aestivu
352 CCGGCAGTCCCGCGAGG.....GCTGCAGCTGCGGCGAACAACTGCAAGTGCAACC MT_durum_g
334 CCGGCAGTCCCGCGAGGCCGCCCACGGCTGCAGCTGCGGTAGCGGCTGCAAGTGCCACC maize_MT_g

      430
387 CCTGCAACTGTAA MT_aestivu
403 CCTGCAACTGTAA MT_durum_g
394 CCTGCAACTGCTGA maize_MT_g

```

Figure 4.5: Multiple alignment of maize, durum and aestivum metallothionein gene DNA sequences.



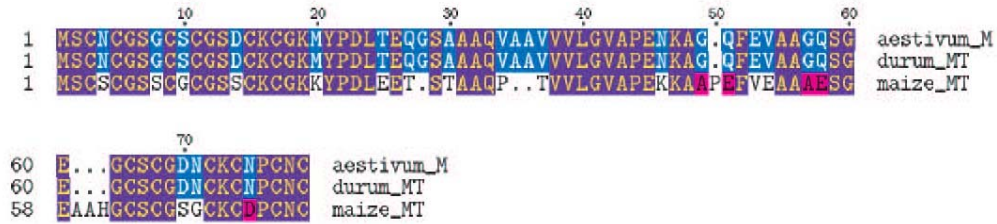


Figure 4.6: Multiple alignments of durum, aestivum and maize metallothionein proteins.

Aestivum and durum genomic DNA sequences were also aligned with the metallothionein cDNA of *T. durum* (alignment not shown). According to these alignments both *T. durum* and *T. aestivum mt* genes have 2 exons and 1 intron (Figure 4.7 and 4.8).



Figure 4.7: *T. durum* metallothionein gene has 2 exons (blue shaded) and 1 intron (red shaded).



Figure 4.8: *T. aestivum* metallothionein gene has 2 exons (blue shaded) and 1 intron (red shaded).

The only available information in the literature is on type I wheat MT which is given as “metallothionein-like protein from *Triticum aestivum*” with an NCBI accession number, AAA50846 (Snowden and Gardner, 1993). The sequenced mRNAs for durum and aestivum MTs were translated into protein sequence and aligned with each other as well as AAA50846. The three protein sequences from Balcali, Bezostaja and AAA50846 were shown to be identical (Figure 4.9).

```

      10      20      30      40      50      60
1  MSCNCGSGCSCGSDCKCGKMPDLTEQGSAAAQVAAVVVLGVAPENKAGQFEVAAGQSGE L11879_whe
1  MSCNCGSGCSCGSDCKCGKMPDLTEQGSAAAQVAAVVVLGVAPENKAGQFEVAAGQSGE durum_MT
1  MSCNCGSGCSCGSDCKCGKMPDLTEQGSAAAQVAAVVVLGVAPENKAGQFEVAAGQSGE aestivum_M

      70
61  GCSCGDNCKCNPNC L11879_whe
61  GCSCGDNCKCNPNC durum_MT
61  GCSCGDNCKCNPNC aestivum_M

```

Figure 4.9: Multiple alignments of AAA50846 (L11879\_whe), Balcali (durum\_MT) and Bezostaja (aestivum\_M) MT protein sequence.

Although all three wheat MT proteins seem to be identical, their mRNA sequences show differences (Figure 4.10). These differences, however, are not reflected in the protein sequences due to the degeneracy of the genetic code.

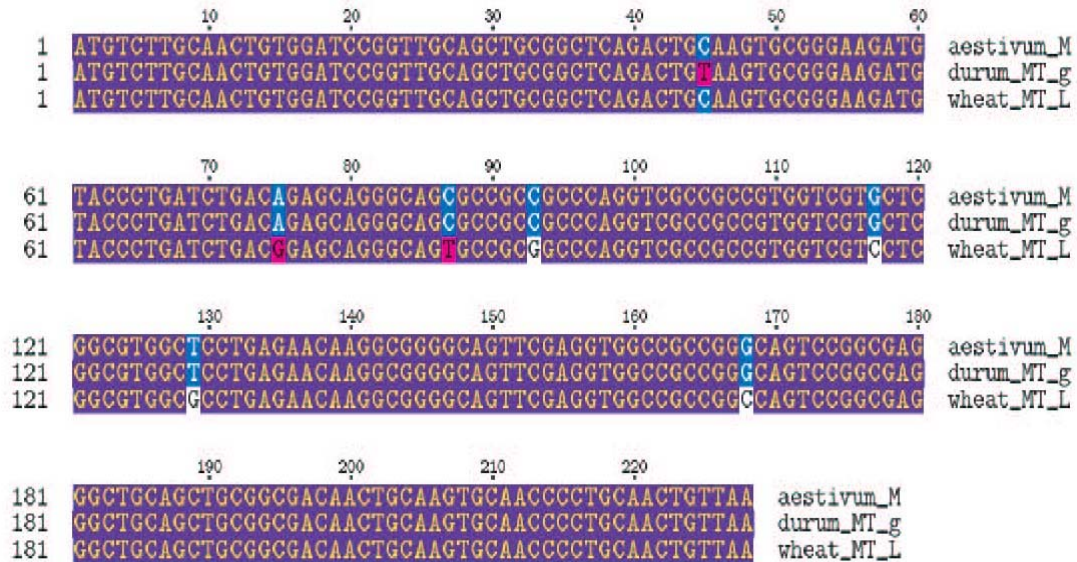


Figure 4.10: Multiple sequence alignments of Bezostaja (*aestivum\_M*), Balcali (*durum\_MT\_g*) and AAA50846 (*wheat\_MT\_L*).

#### 4.5 Cloning of *mt-a* cDNA in *E. coli*

After characterization of *mt* genomic sequences it was necessary to obtain *mt* cDNA which could be used in recombinant protein expression studies. For this purpose total RNA and mRNA were isolated from plants treated with 5 and 10  $\mu$ M Cd to induce heavy metal stress and *mt* gene expression.

RT-PCR was performed on the mRNA using QIAGEN One Step RT-PCR enzyme mix with primers; Oligo 1 and Oligo 2 (Table 4.1) with an annealing temperature of 53.5 °C. Metallothionein cDNAs could be detected at both cadmium concentrations (Figure 4.11).

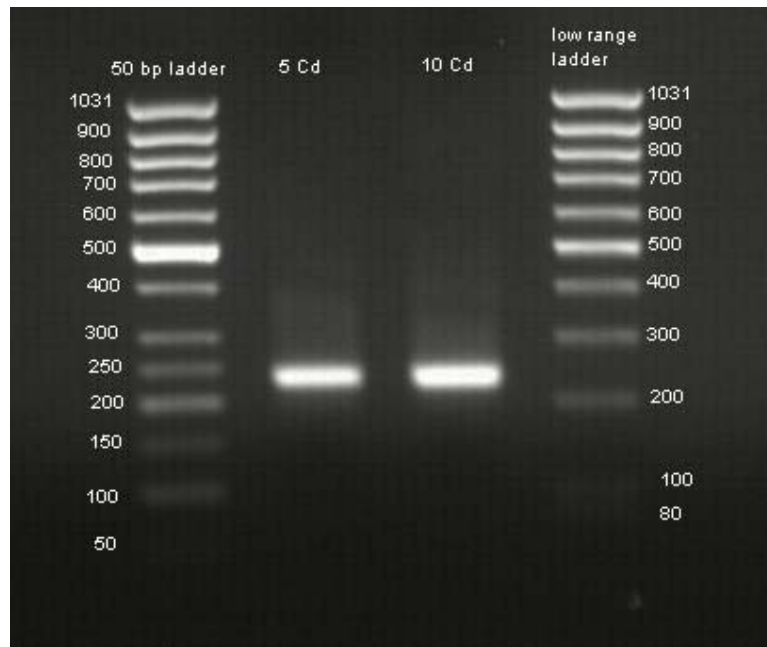


Figure 4.11: Electrophoretic analysis of RT-PCR results showing amplification of *T. aestivum* cDNA for *mt* gene.

Purified cDNAs were then ligated into pGEM-Teasy vector (pGEMaMT) and *E. coli* Top10 cells were transformed with the constructs.

Colonies that grew on selective ampicillin plates were screened for *mt-a* cDNA insert by plasmid isolation from cultures of transformed *E. coli*. Screening was carried out by restriction enzyme digestion using *EcoRI* whose recognition sites are at the two flanking regions of the insert on the pGEM-Teasy vector. *mt-a* cDNA can be seen migrating between 200 and 300 bp bands on agarose gels (Figure 4.12).

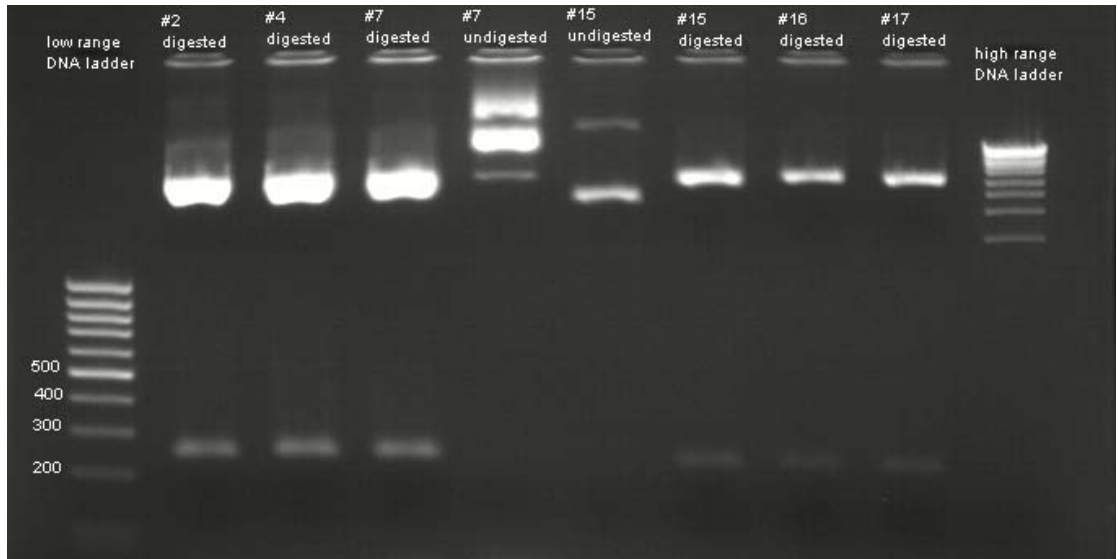


Figure 4.12: Electrophoretic analysis of digestion results for pGEMaMT constructs. Undigested constructs (lanes 5 and 6). *mt-a* cDNA bands migrate between 200 and 300 bp bands of the low range DNA ladder.

#### 4.6 Cloning of *mt-d* cDNA in *E. coli*

*mt-d* mRNA was isolated from *T. durum* cultivars Balcali and Cesit-1252 which were grown under 5 $\mu$ M and 10 $\mu$ M CdSO<sub>4</sub> stress and RT-PCR was carried out as described in Materials and Methods. Electrophoretic analysis of RT-PCR products again showed DNA bands migrating at a position between 200 and 300 bp (Figure 4.13). These bands were expected to be *mt-d* cDNA.

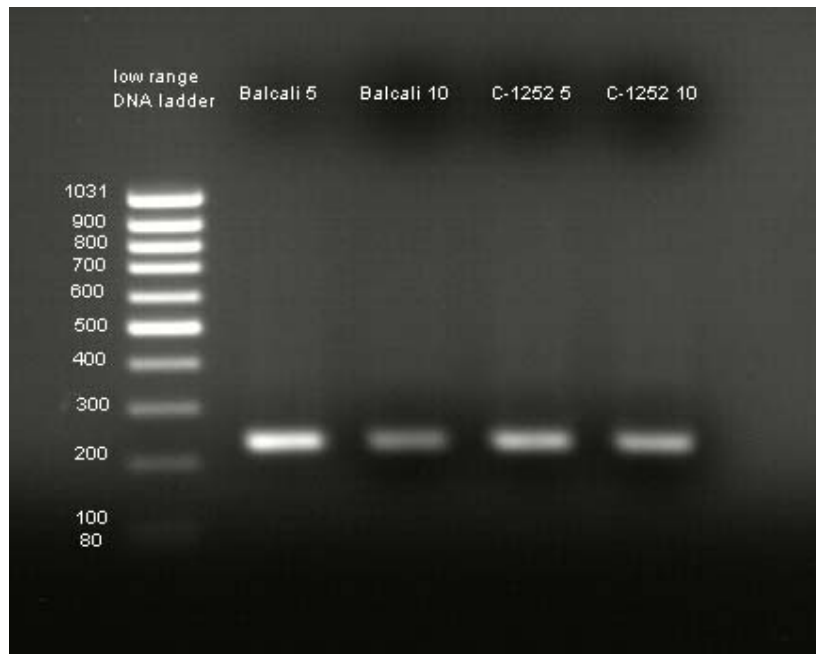


Figure 4.13: Electrophoretic analysis of RT-PCR products showing amplification Balcali and Cesit-1252 *mt* cDNA.

Similar to *T. aestivum* case described in section 4.5 *mt-d* cDNAs were then ligated with pGEM-Teasy vectors (pGEMdMT) and cloned into *E. coli* Top10 bacterial cells.

Control digestions to check the insert *d-mt* cDNAs were carried out on plasmids isolated from cells grown from colonies that grew on ampicillin plates (Figure 4.14 and 4.15).

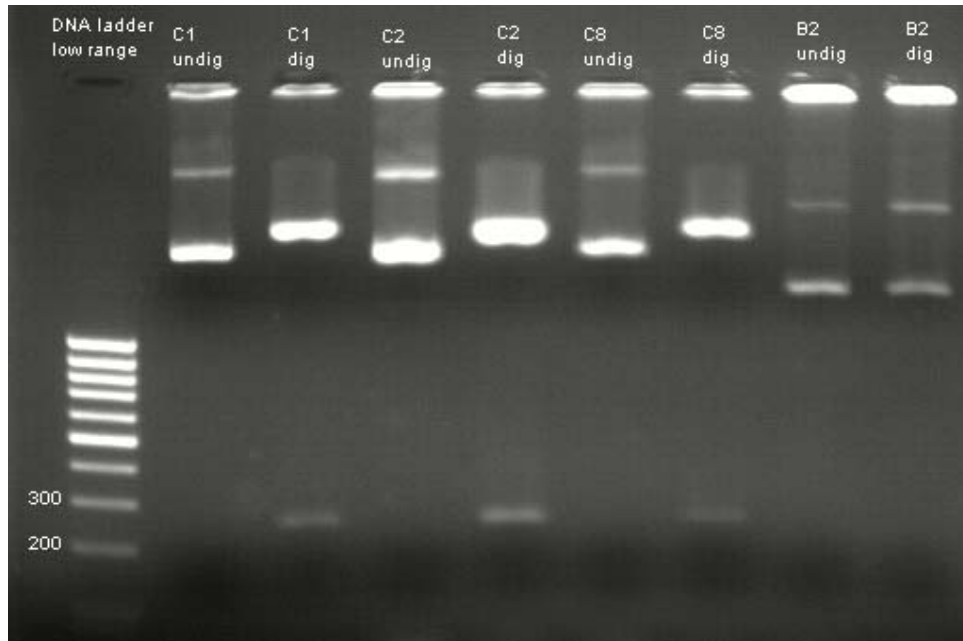


Figure 4.14: Electrophoretic analysis of digestion results for pGEMdMT constructs. Undigested (undig) and digested (dig) constructs. *mt-d* cDNA bands migrate between 200 and 300 bp bands of the low range DNA ladder.

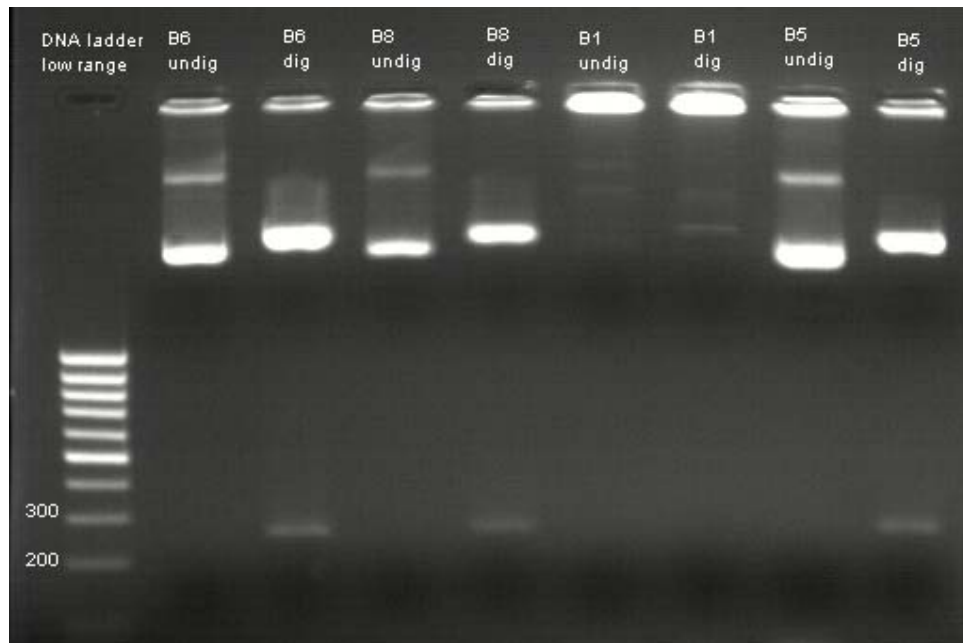


Figure 4.15: Electrophoretic analysis of digestion results for pGEMdMT constructs. Undigested (undig) and digested (dig) constructs. *mt-d* cDNA bands migrate between 200 and 300 bp bands of the low range DNA ladder.

## 4.7 Expression of dMT in *E. coli*

### 4.7.1 Insertion into the expression vector

Positions of the *SpeI* and *EcoRI* restriction sites on pGEM-Teasy vectors allowed the *mt-d* cDNA insert to be conveniently digested out of the pGEM-dMT construct for further manipulations (Figure 4.16).

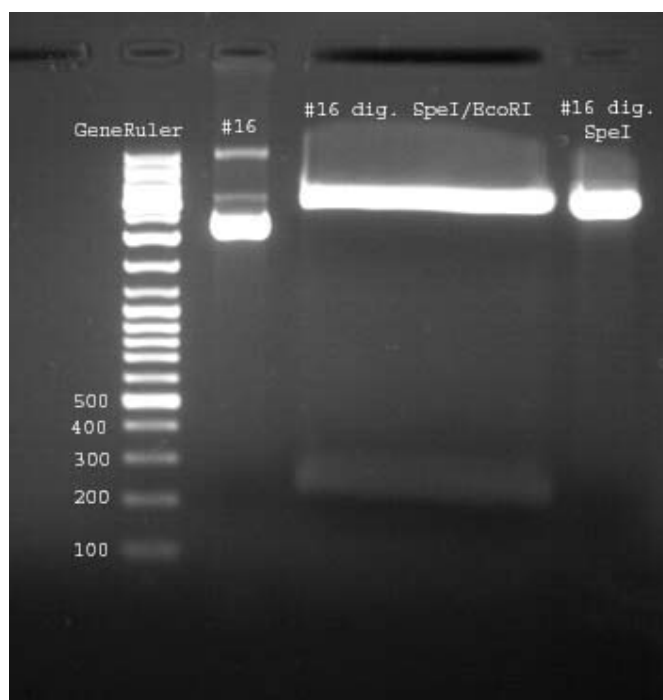


Figure 4.16: Preparative agarose gel analysis for isolation of the *d-MT* cDNA. Undigested construct (lane 3) and linearized construct with *SpeI* digestion (lane 4). *mt-d* cDNA bands migrate between 200 and 300 bp bands of the marker DNA.

*mt-d* genes were amplified from cDNA inserts isolated from pGEM-dMT constructs using primers (Table 4.1) that do not contain restriction enzyme sites and then with those that contain restriction sites for *EcoRI* at 5' end, and *XhoI* at 3' end (Table 4.3) which would facilitate ligation with the pGEX-4T-2 expression vector. Electrophoretic analysis of the PCR products indicated the high efficiency of the amplification process (Figure 4.17).



Table 4.3: Designed primers with RE sites for pGEX-4T-2 vector.

P#6_F	5'-CTATGGAATTCCCATGTCTTGCAAC -3'	forward primer
P#3_Rstop	5'-CTATGCTCGAGTTAACAGTTGCAGG -3'	reverse primer

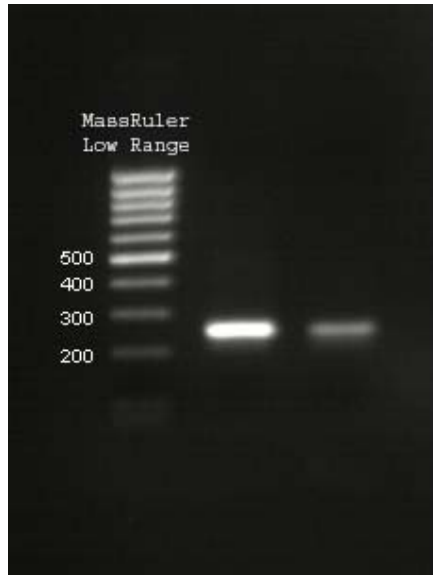


Figure 4.17: Electrophoretic analysis of amplified durum cDNA using primers designed for pGEX-4T2 (lane 1) and designed for pGFPuv (lane 2) vectors.

The amplified and purified durum cDNAs were then ligated into pCR-II-TOPO vector (pCRdMT) and *E. coli* XL1-Blue bacterial host cells were transformed with the construct. Presence of the insert was screened by a double digestion with *EcoRI* and *XhoI* of plasmids isolated from *E. coli* cultures prepared from cells that grew on ampicillin plates. This subcloning step was performed in order to prepare stocks of *mt-d* cDNA constructs and to achieve high plasmid yield for sequence verification and further manipulations.

Preparative purification of pCRdMT constructs were carried out for restriction digestion of the *mt-d* cDNA with restriction enzymes *EcoRI* and *XhoI*. Similarly pGEX-4T-2 vector was prepared by digestion with the same restriction enzymes and durum cDNA was ligated with pGEX-4T-2 expression vector (pGEXdMT) as described in Materials and Methods.

*E. coli* XL1-Blue host cells were transformed with pGEXdMT constructs and screened for the plasmid and then for the insert by a double digestion with *EcoRI* and *XhoI* (Figure 4.18).

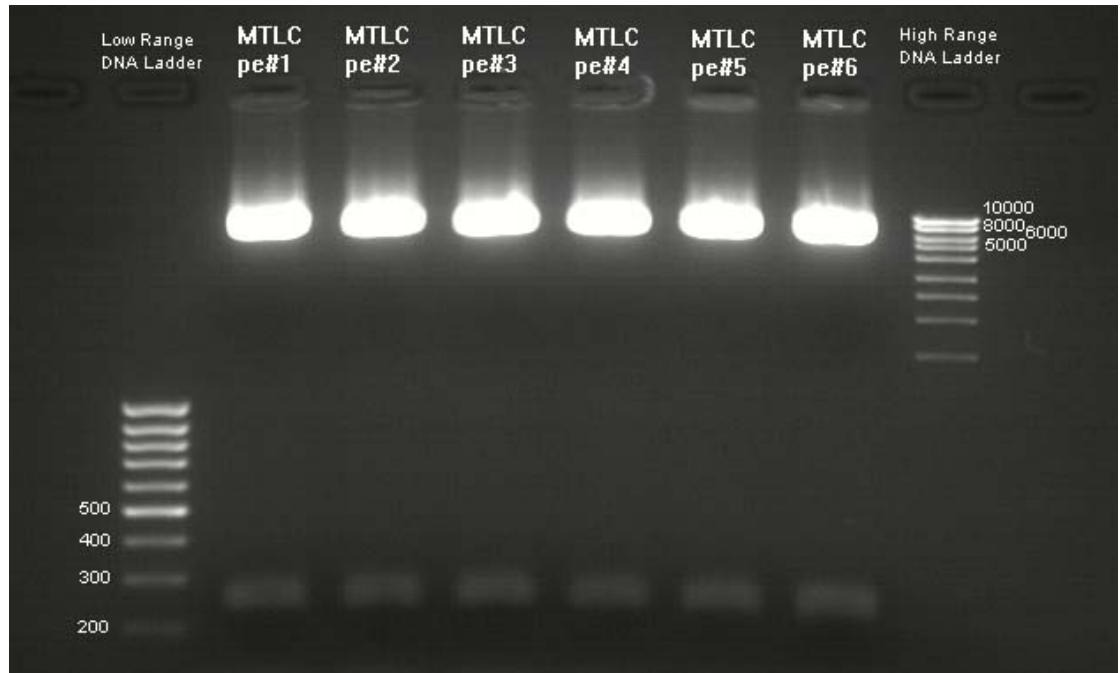


Figure 4.18: Electrophoretic analysis of digestion results for pGEXdMT constructs. *mt-d* cDNA bands migrate between 200 and 300 bp bands of the low range DNA ladder.

After confirmation of the insert expression constructs were purified and *E. coli* BL21 expression host cells were transformed.

#### 4.7.2 Induction of dMT expression in *E. coli*

Expression of *T. durum mt* gene in *E. coli* was followed in induction experiments. Liquid cultures were inoculated with BL21(DE3), pGEX-4-T2 in BL21(DE3), and pGEXdMT in BL21(DE3) cells and expression was induced with 0.7 mM of IPTG. Cells with unmodified pGEX-4T-2 would express the fusion protein GST alone and serve as control for expression and untransformed BL21(DE3) cells would provide controls for expression and untransformed BL21(DE3) cells would provide controls for growth of unmodified cells as well as providing the protein profile of unmodified cells. Results showed that the presence of neither the expression vector nor the expressed dMT protein affected the growth of the bacterial host cell (Figure 4.19).

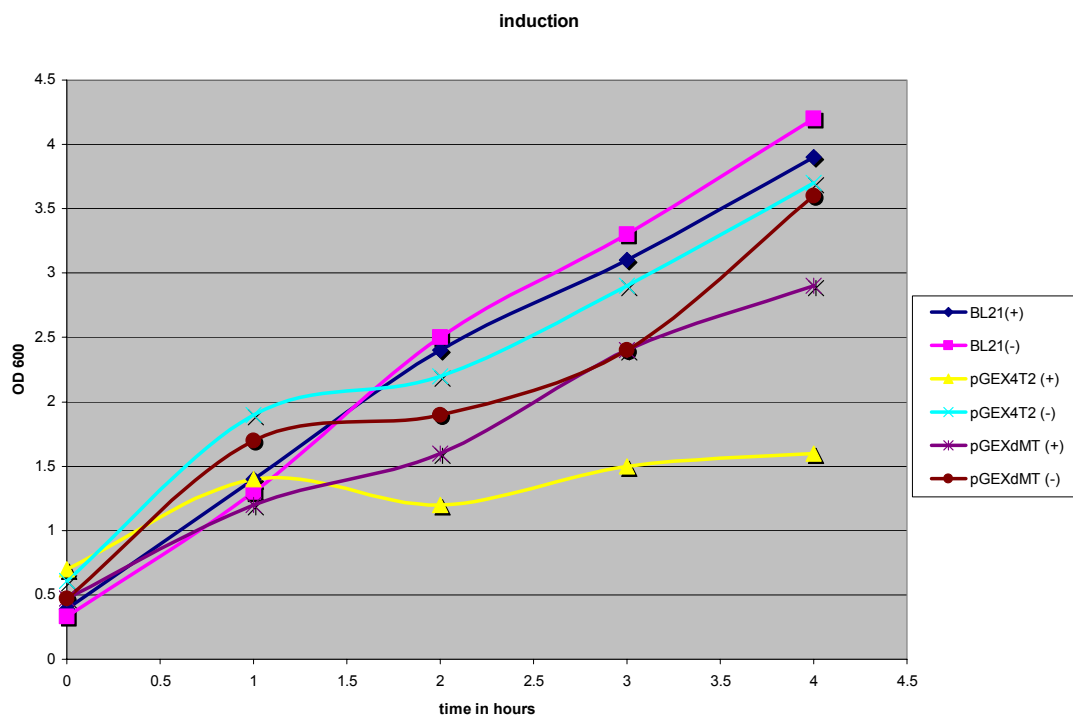


Figure 4.19: Growth curve of 0.7 mM IPTG induced (+) and non-induced (-) *E. coli* BL21 cells (BL21+ and BL21-) containing pGEX-4T-2 (+,-) and pGEXdMT (+,-) vectors.

SDS-PAGE analysis of the cellular lysates showed that although the protein profiles of induced and non-induced cells with the construct were not very different, the molecular weight distinction of the GST protein and the GSTdMT fusion protein could be easily seen (Figure 4.20).

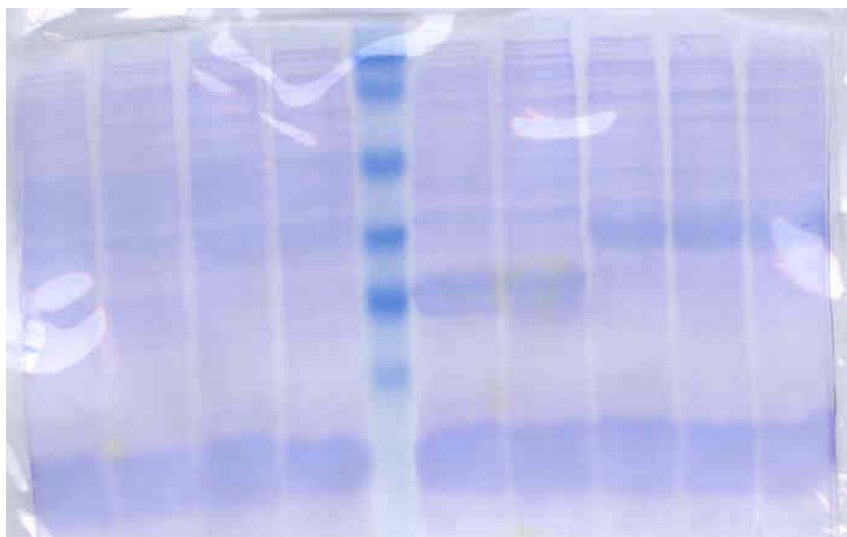


Figure 4.20: SDS-PAGE analysis to check GST and GSTdMT productions in IPTG induced and non-induced cells. (for legend, Table 4.4).

Table 4.4: Legend for Figure 4.21

Lane #	sample	Lane #	sample
1	pGEX4T2(-) at T <sub>3</sub>	6	pGEX4T2(+) at T <sub>3</sub>
2	pGEX4T2(-) at T <sub>4</sub>	7	pGEX4T2(+) at T <sub>4</sub>
3	pGEXdMT(-) at T <sub>3</sub>	8	pGEXdMT(+) at T <sub>2</sub>
4	pGEXdMT(-) at T <sub>4</sub>	9	pGEXdMT(+) at T <sub>3</sub>
5	protein marker 1	10	pGEXdMT(+) at T <sub>4</sub>

In order to stabilize the expressed MT structure it was decided to include metal ions in the growth medium of bacteria and the effect of cadmium (CdSO<sub>4</sub>) on host cell growth was assayed at 5 different Cd concentrations (Table 4.5) on control bacterial cells transformed with the unmodified expression vector and those expressing the recombinant dMT protein.

Table 4.5: Cell types and given Cd concentration during the induction.

sample #	cell type	Cd concentration
#1	BL21	
#2	BL21/pGEX4T2	
#3	BL21/pGEXdMT	
#4	BL21/ pGEXdMT	0.05 mM
#5	BL21/ pGEXdMT	0.30 mM
#6	BL21/ pGEXdMT	0.60 mM
#7	BL21/ pGEXdMT	2.00 mM
#8	BL21/ pGEXdMT	4.00 mM
#9	BL21/ pGEX4T2	0.05 mM
#10	BL21/ pGEX4T2	0.30 mM
#11	BL21/ pGEX4T2	0.60 mM
#12	BL21/ pGEX4T2	2.00 mM
#13	BL21/ pGEX4T2	4.00 mM

Cells were grown as described in materials and methods and were induced at  $t=0$ . 2 and 4 mM of  $\text{CdSO}_4$  was shown to be lethal both for dMT expressing and control cells. Growth curves (Figure 4.21) showed that in the range 0.05 – 0.3 mM  $\text{CdSO}_4$  in the presence of recombinant GSTdMT, cells tolerated Cd in the growth medium better than the controls and optimum harvesting time for protein extraction was found to be 6-7 hours after induction with 0.7 mM IPTG.

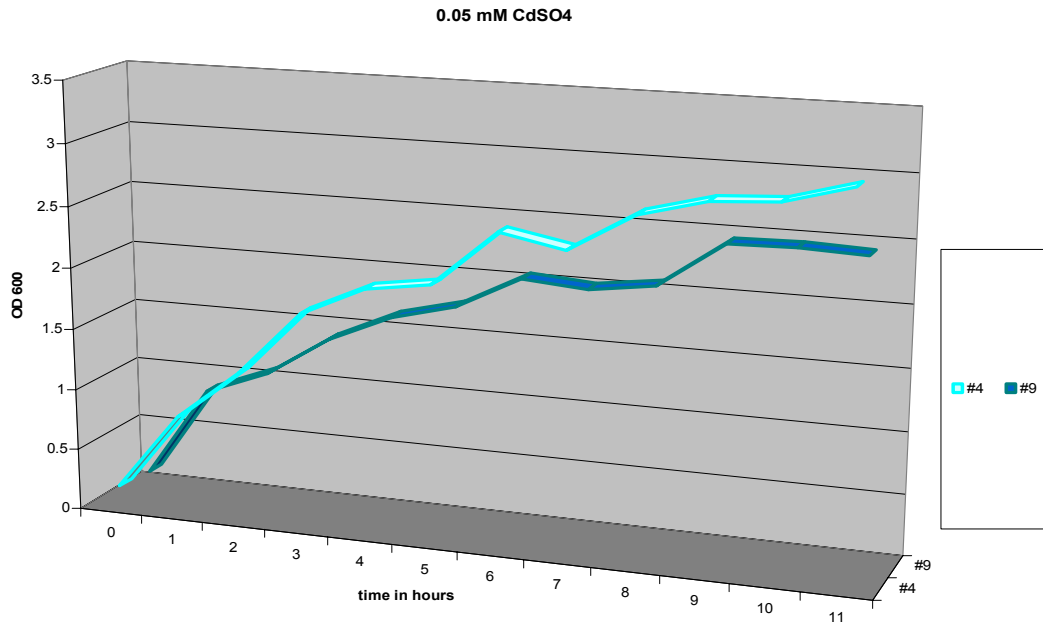


Figure 4.21(a): Growth curves of GSTdMT (#4) and GST (#9) expressing BL21(DE3) cells at 0.05 mM CdSO<sub>4</sub> concentration.

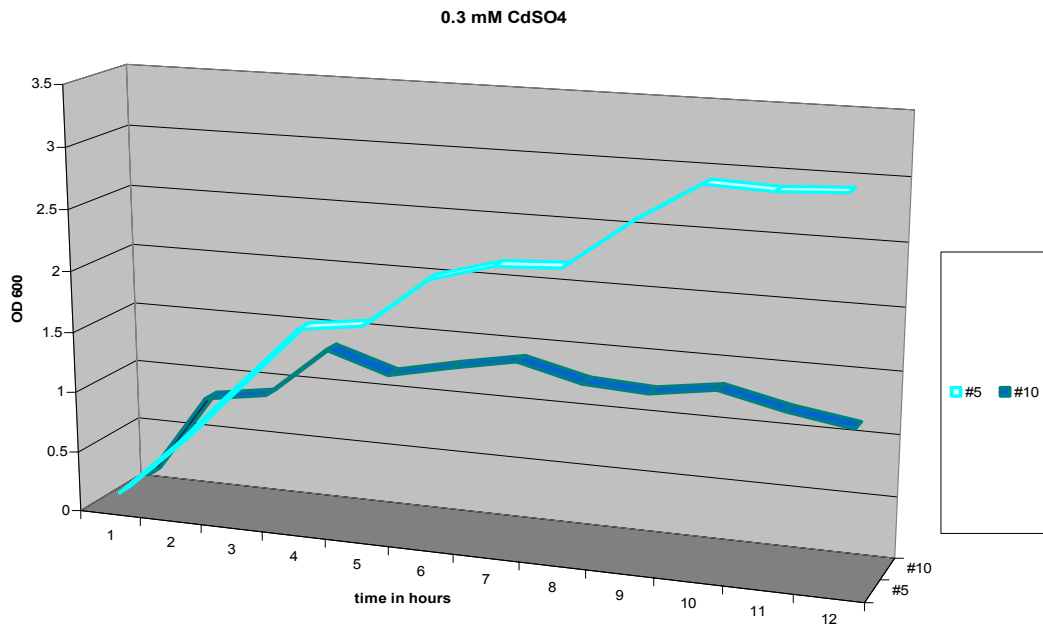


Figure 4.21(b): Growth curves of GSTdMT (#5) and GST (#10) expressing BL21(DE3) cells at 0.3 mM CdSO<sub>4</sub> concentration.

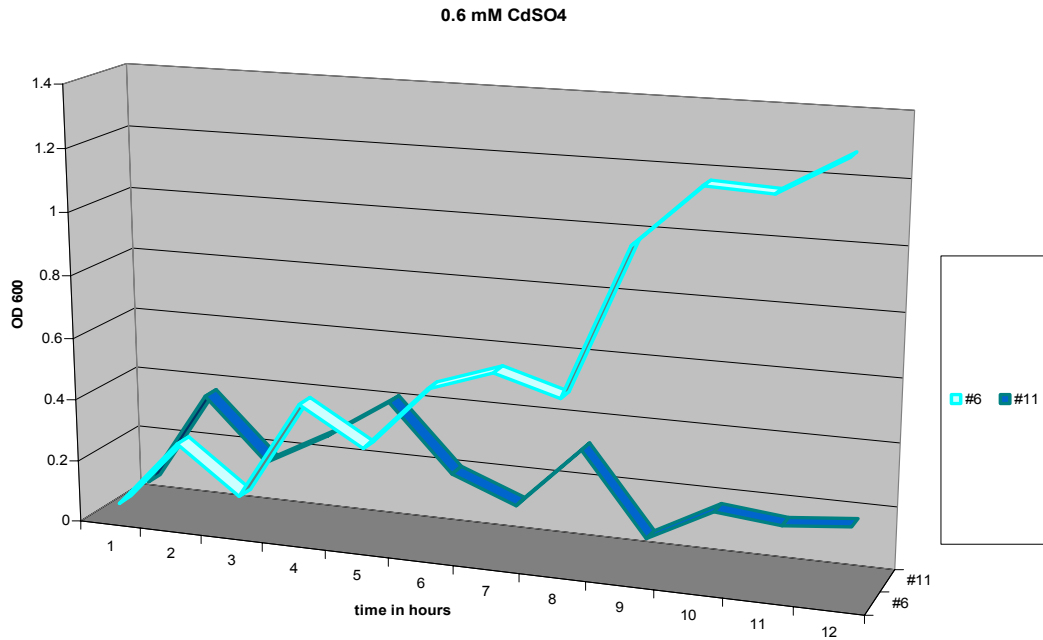


Figure 4.21(c): Growth curves of GSTdMT (#6) and GST (#11) expressing BL21(DE3) cells at 0.6 mM CdSO<sub>4</sub> concentration.

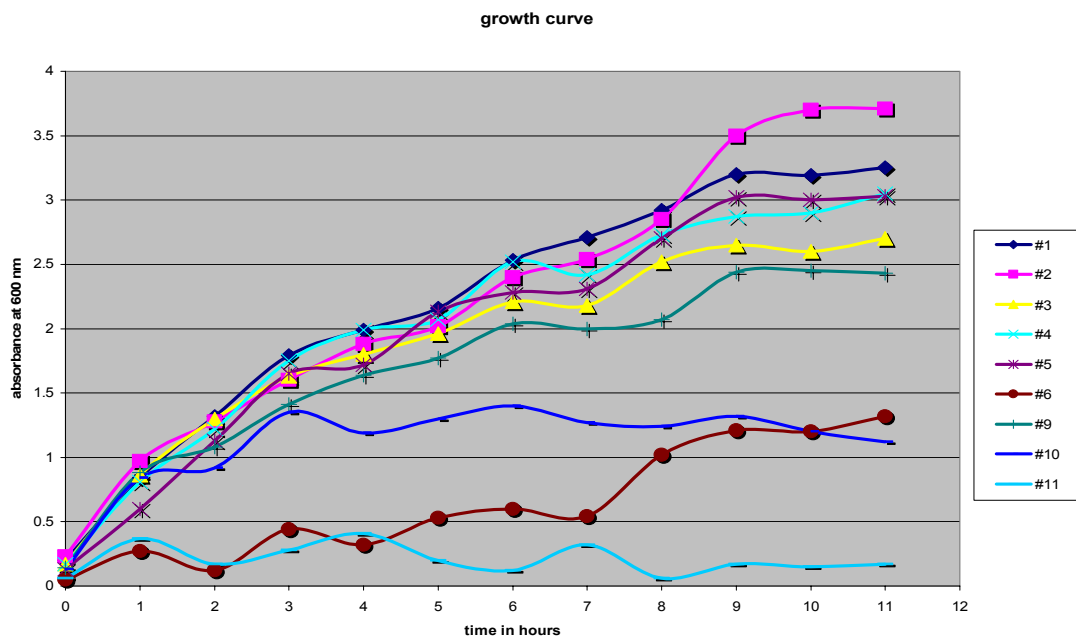


Figure 4.21(d): Summary growth curves at all tried Cd concentrations. See Table 4.5 for legend.

As well as varying CdSO<sub>4</sub> concentration, IPTG concentrations at 0.5, 0.7, 1.0, 1.3, 1.6, and 1.9 mM were also tried in order to find the optimum induction level. Cellular proteins were extracted and checked for recombinant protein expression differences. No significant difference due to varied IPTG concentrations was observed (Figure 4.22). 0.7 mM IPTG and 0.05 mM CdSO<sub>4</sub> concentrations were chosen for all further induction experiments.

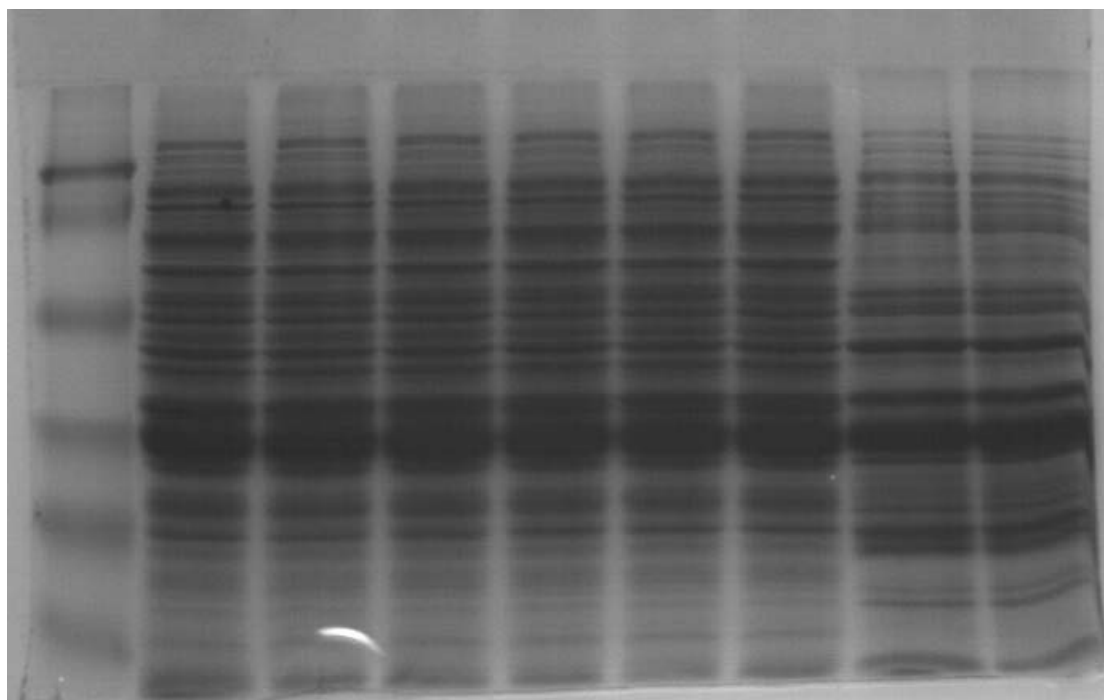


Figure 4.22: SDS-PAGE analysis to check effects of different IPTG concentrations during induction of recombinant dMT protein expression (for legend, Table 4.6).

Table 4.6: Legend for figure 4.23.

Lane #	sample	Lane #	sample
1	protein marker 1	6	1.6 mM IPTG at T <sub>6</sub>
2	0.5 mM IPTG at T <sub>6</sub>	7	1.9 mM IPTG at T <sub>6</sub>
3	0.7 mM IPTG at T <sub>6</sub>	8	pellet, 0.7 mM IPTG at T <sub>6</sub>
4	1.0 mM IPTG at T <sub>6</sub>	9	pellet, 1.0 mM IPTG at T <sub>6</sub>
5	1.3 mM IPTG at T <sub>6</sub>		



## 4.8 Purification of dMT protein

### 4.8.1 Batch purification

Recombinant protein was extracted from pellets of induced *E. coli* BL21(DE3) cells containing pGEXdMT construct as well as with unmodified pGEX-4T-2 which expressed GST only. GSTdMT and GST protein purification was performed using Glutathione Sepharose 4 Fast Flow resin and manufacturer's "batch purification protocol" was followed without any modification.

Six successive elution steps yielded a total of 0.405 mg protein from 3.6 g cell pellet obtained from 1 liter of culture. Isolated proteins could be characterized by absorption measurements at 280 nm which measured the amount of recombinant GST in the fusion protein directly, by comparison of absorption spectra of GST and GSTdMT as well as SDS-PAGE analysis. SDS-PAGE analysis (Figure 4.23) showed the presence of impurities in the final protein solutions.

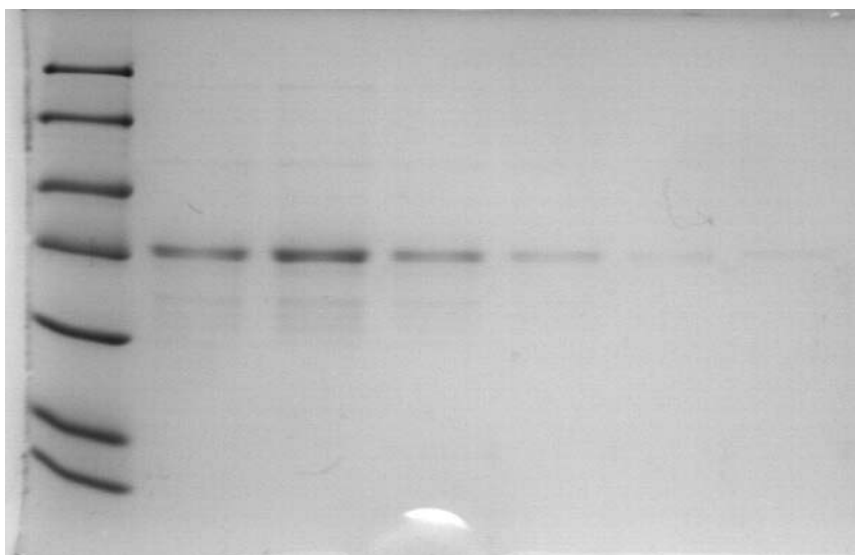


Figure 4.23: SDS-PAGE analysis of eluted GSTdMT fusion proteins (lanes 2-7). Protein marker 3 (lane 1).

On the other hand, GST and GSTdMT preparations could be clearly distinguished from the molecular mass difference between the recombinant proteins visualized by SDS-PAGE analysis (Figure 4.24).

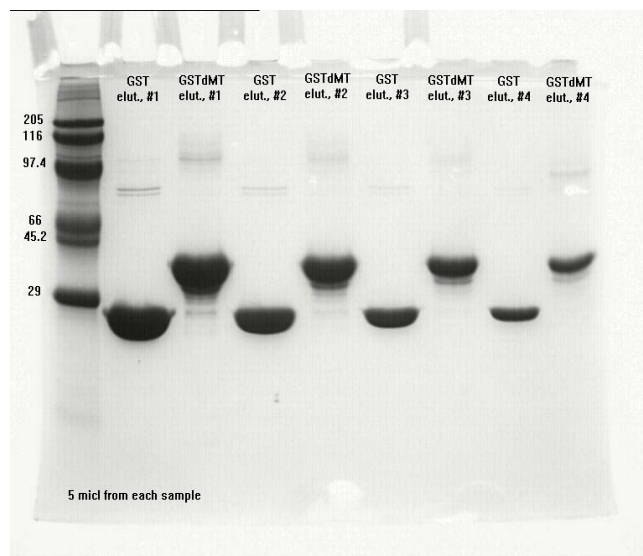


Figure 4.24: SDS-PAGE analysis of different elution fractions of batch purified recombinant GST (29 kDa) and GST-dMT (36 kDa). Molecular masses of the marker are indicated on the left.

Further characterization of GST and GSTdMT preparations was carried out by recording the absorption spectra. Spectra were recorded in the wavelength range between 300 to 230 nm. Comparison of the spectra shows increased absorbance for the GSTdMT sample in the range of 245-265 nm, which corresponds to charge transfer due to presence of cadmium (Figure 4.25).

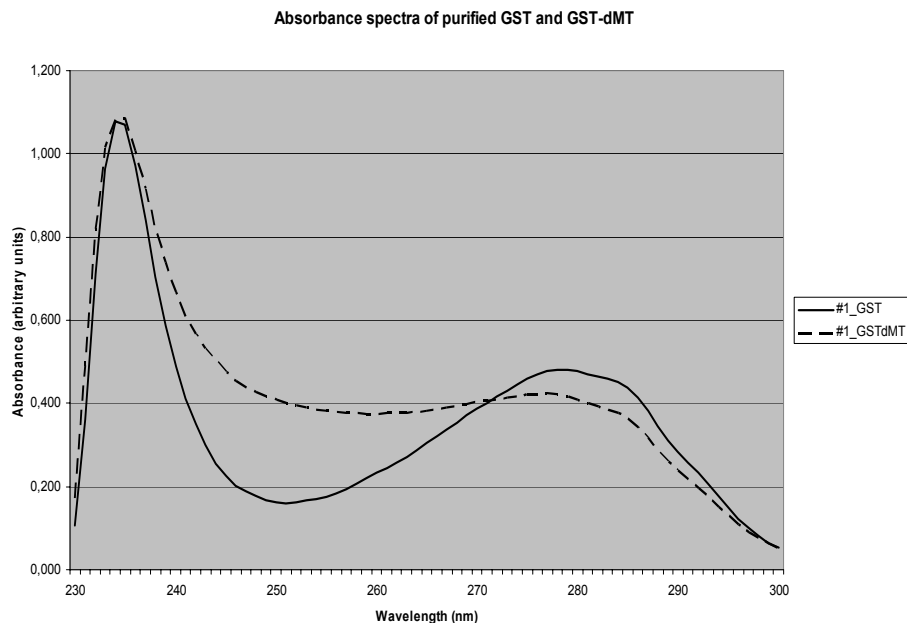


Figure 4.25: Absorbance spectra of batch purified GST and GST-dMT.

#### 4.8.2 Purification using GST (Glutathione-S-transferase) affinity chromatography

In order to improve homogeneity of the purified GST-dMT a prepacked Amersham GStrap FF affinity column was used. After loading on the affinity column d-MT was cleaved from GST fusion protein on the column using thrombin protease. Cleaved d-MT and GST proteins were then eluted from the column which was connected to AKTA-FPLC<sup>(R)</sup> system (Figure 4.26).

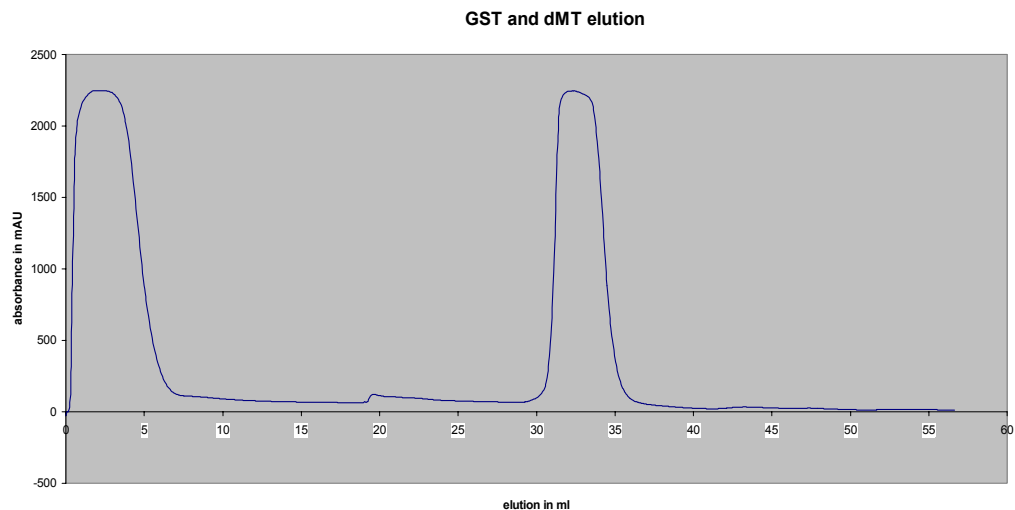


Figure 4.26: Elution of dMT (1<sup>st</sup> peak) and GST (2<sup>nd</sup> peak) proteins from GSTrap FF affinity column.

SDS-PAGE analysis was carried out for the separated dMT and GST recombinant proteins. As expected, although GST was seen on the gel, dMT could not be visualized. Comparison of different samples loaded on the gel at different total protein showed that the major impurities in dMT fraction were contamination from GST (Figure 4.27). Samples were also analyzed on native polyacrylamide gels, run under nondenaturing conditions, and on this overloaded gel bands corresponding to the contaminants as well as 2 bands which can be distinguished on the background of a smear were seen. Positions of these bands appear to indicate complex formation by dMT leading to dimers, trimers, etc. (Figure 4.28).

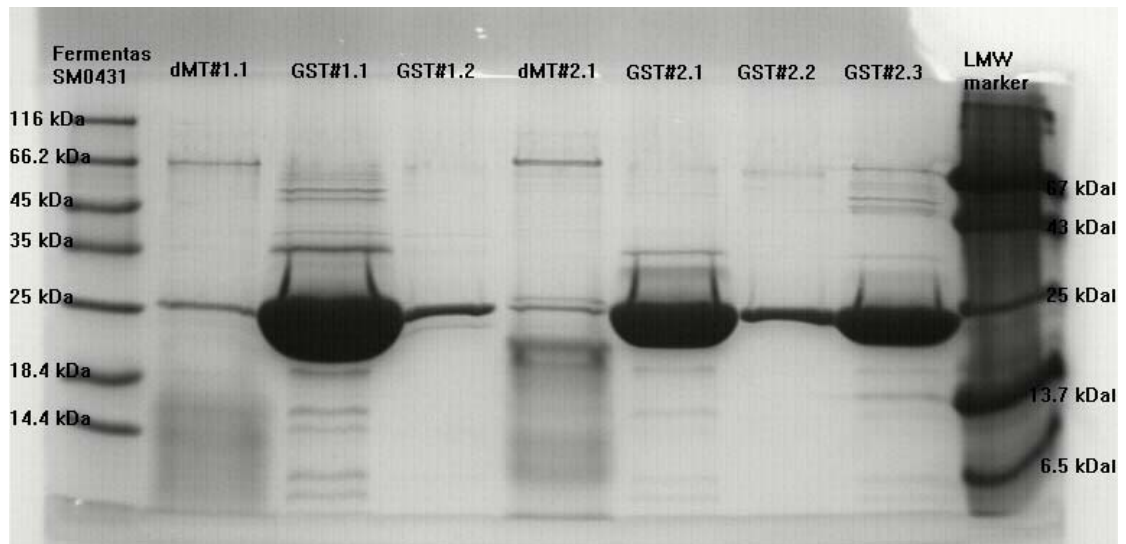


Figure 4.27: SDS-PAGE analysis of cleaved dMT and GST recombinant proteins. Sample numbers indicate different purification batches.

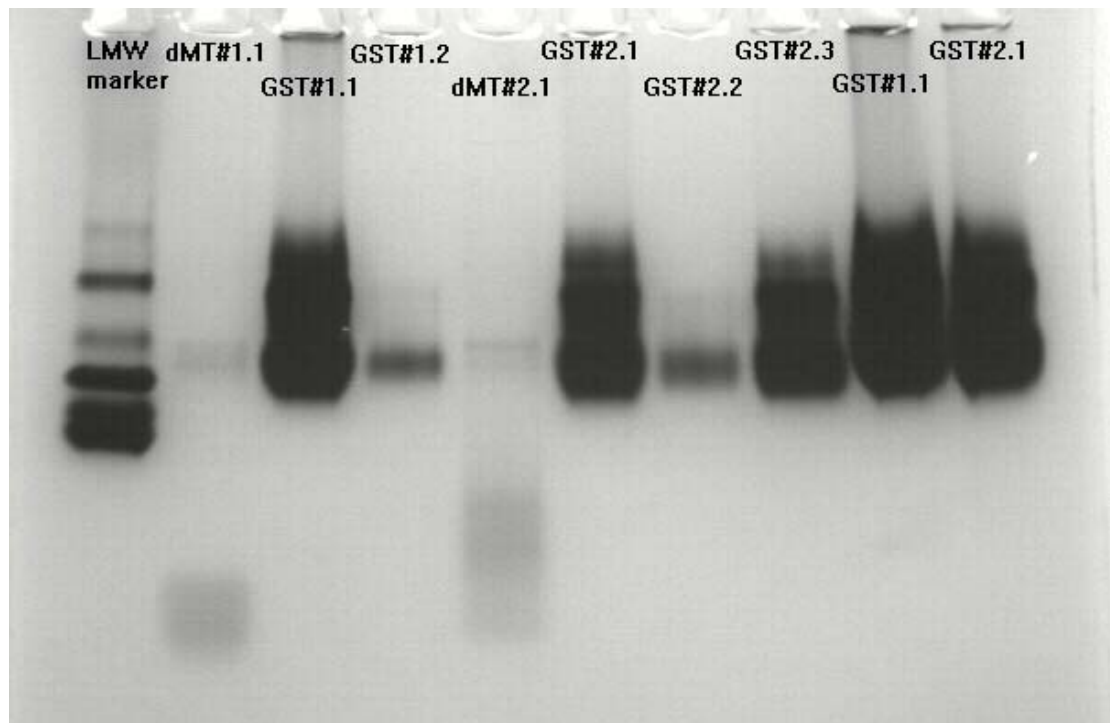


Figure 4.28: Native PAGE analysis of cleaved dMT and GST recombinant proteins (lanes 2-10). Low molecular mass marker (lane 1, see appendix). Sample numbers indicate different purification batches.

### 4.8.3 Size exclusion

Size exclusion chromatography was carried out to estimate the molecular mass of the purified d-MT and GST-dMT fractions and to monitor non-specific aggregation or complex formation in the protein preparations.

#### 4.8.3.1 Column calibration

Amersham Biosciences HiLoad<sup>(R)</sup> 26/60 column was calibrated with a set of protein (Table 4.7 and Figure 4.29) and a calibration curve was calculated (Figure 4.30).

Table 4.7: Protein samples used for the column calibration and as the low molecular mass marker in the native PAGE analysis.

<b>Protein</b>	<b>MW (kDa)</b>	<b>V<sub>e</sub> (ml)</b>	<b>loaded (mg)</b>
Albumin	67.0	133.57	20
Ovalbumin	43.0	150.21	20
Chymotrypsinogen A	25.0	177.69	20
RNase A	13.7	200.31	20
Aprotinin	6.5	238.77	5
Vitamin B12	1.35	277.76	30

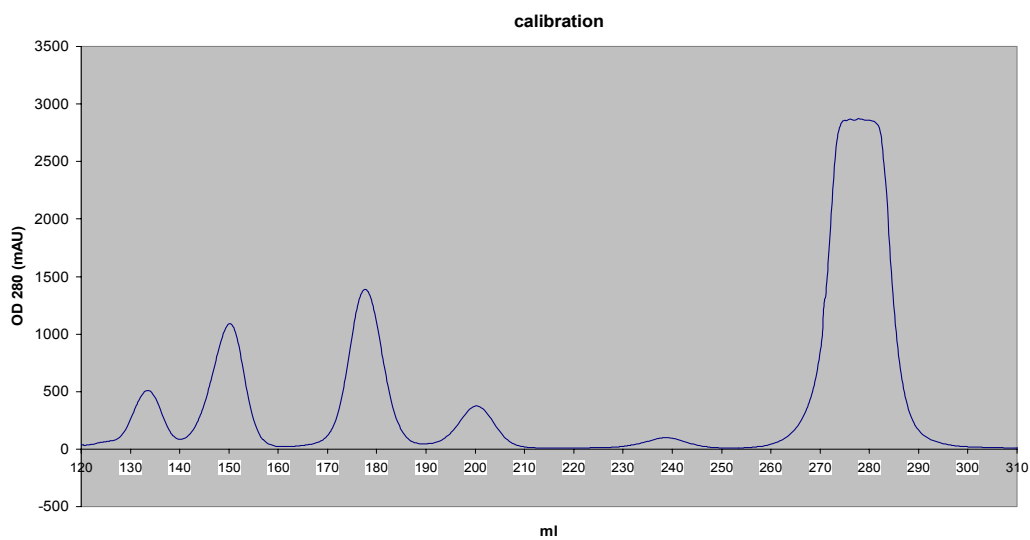


Figure 4.29: Figure 4.30: Elution of proteins (Table 4.6) used in the calibration of size exclusion column. Albumin (1<sup>st</sup> peak), ovalbumin (2<sup>nd</sup> peak), chymotrypsinogen (3<sup>rd</sup> peak), Rnase A (4<sup>th</sup> peak), aprotinin (5<sup>th</sup> peak), and vitamin B12 (6<sup>th</sup> peak).

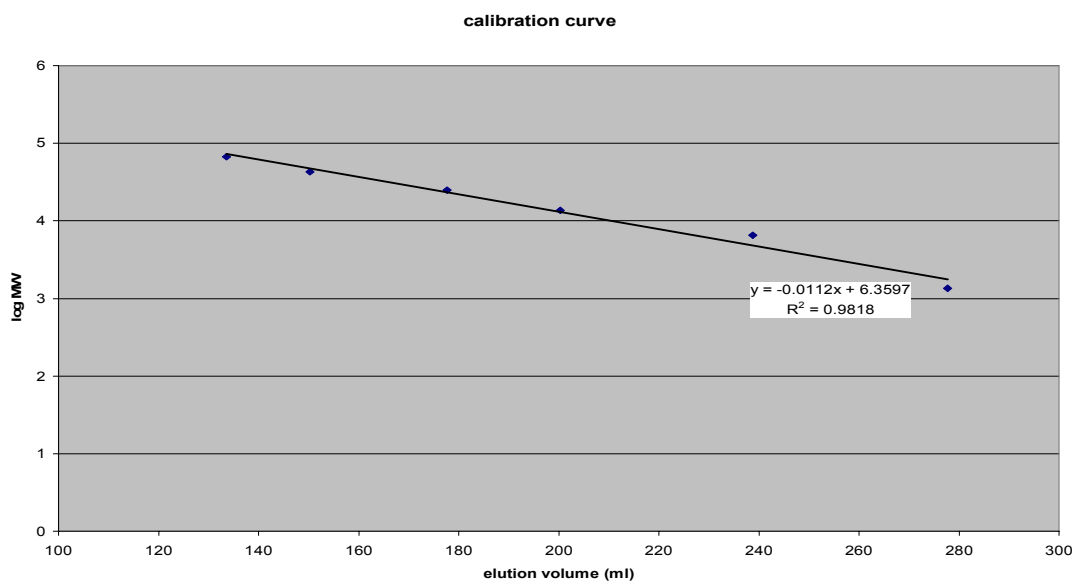


Figure 4.30: Calibration curve for size exclusion column, the equation used for the molecular mass determination of dMT and GSTdMT is given on the chart.

### 4.8.3.2 Analysis of dMT preparations by size exclusion chromatography

The 4 main peaks were detected at 118.61, 178.26, 292.57, and 334.47 ml in the elution pattern (Figure 4.31).

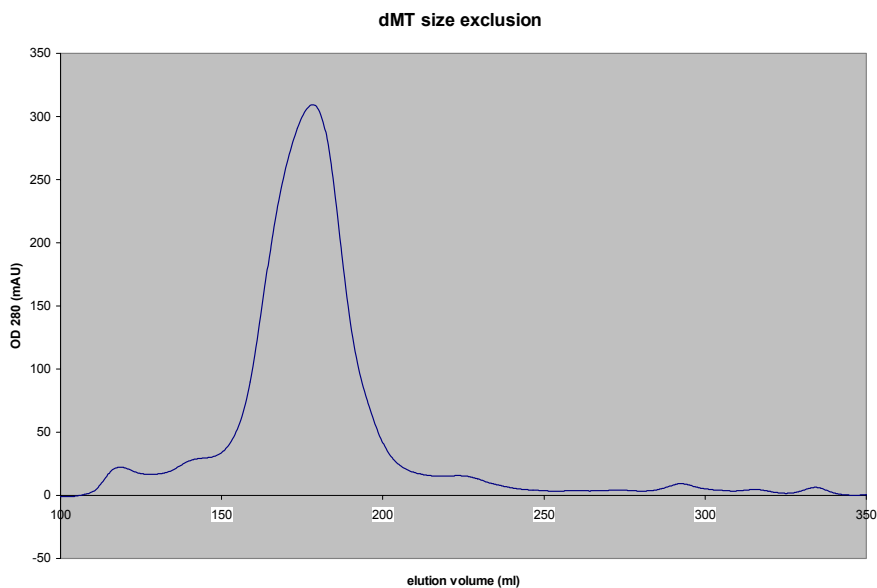


Figure 4.31: Elution profile of the dMT size exclusion chromatography.

Table 4.8: Elution volume and corresponding molecular mass calculated according to the calibration curve. The 3<sup>rd</sup> column indicates; (calculated mass / dMT mass).

<b>Elution volume</b>	<b>Calculated Molecular Mass (Da)</b>	<b>x dMT Molecular Mass</b>
118.61	107465	14.32867
178.26	23077	3.076933
292.57	1210	0.161333
334.47	410	0.054667



### 4.8.3.3 Analysis of dMT preparations by size exclusion chromatography

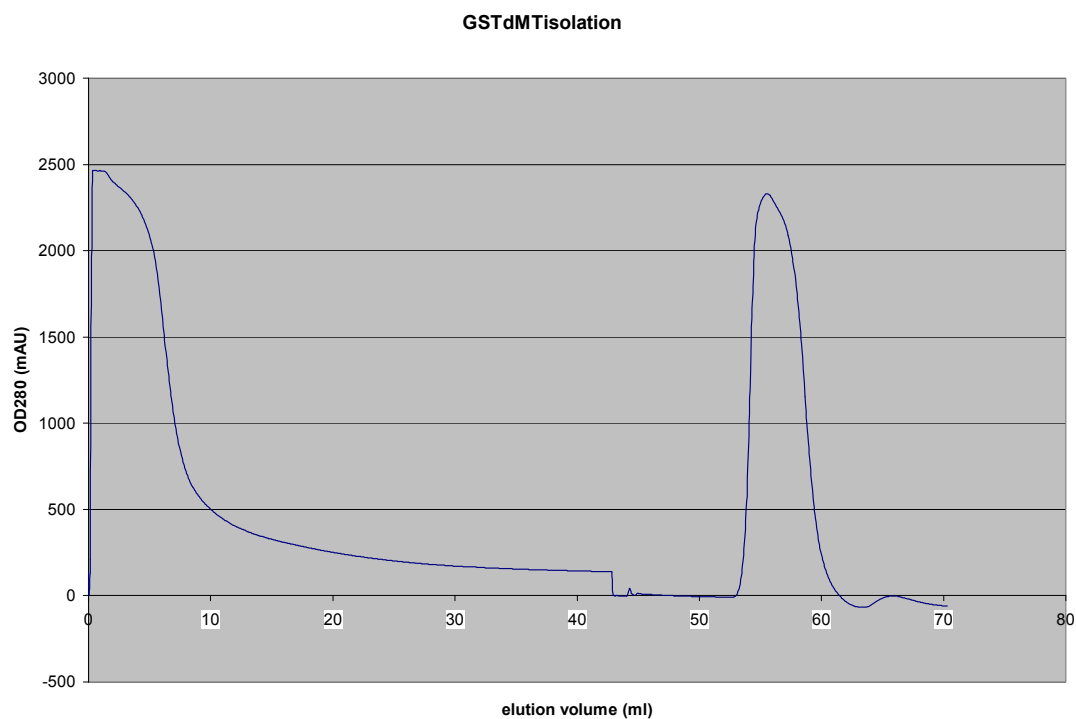


Figure 4.32: GSTdMT purification with HiTrap column.

There were 2 main peaks in the elution diagram of the GSTdMT size exclusion at 116.16 and 133.71 ml (Figure 4.33 and Table 4.9).

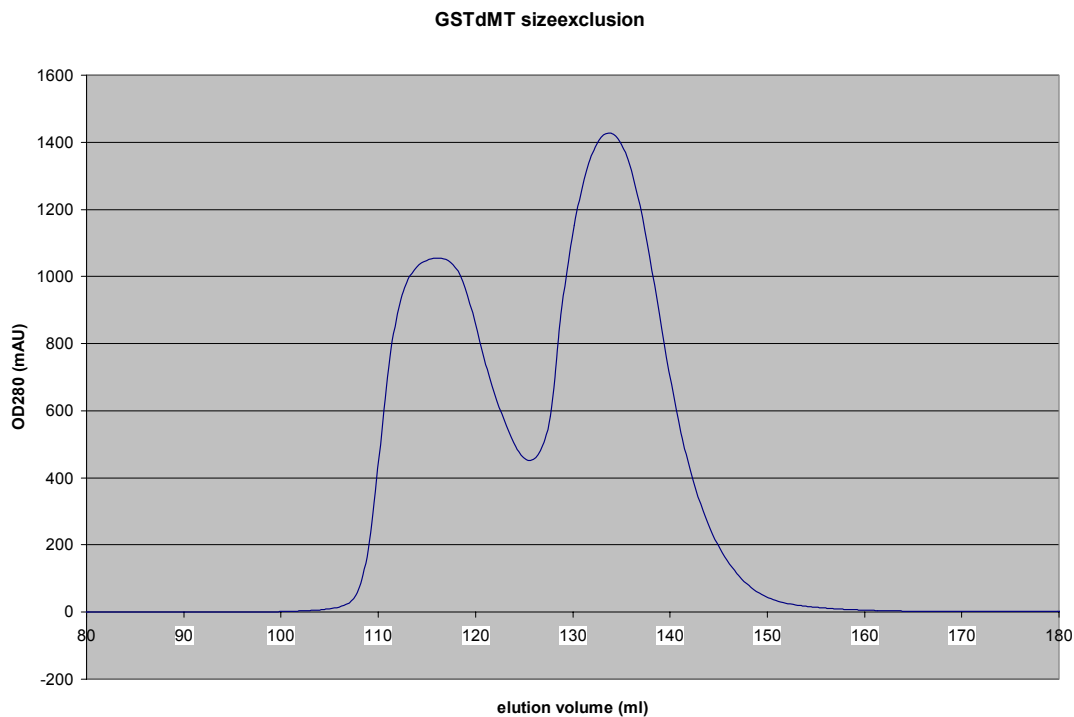


Figure 4.33: Elution profile of the GSTdMT size exclusion chromatography.

Table 4.9: Elution volume and corresponding molecular mass calculated according to the calibration curve. The 3<sup>rd</sup> column indicates; (calculated mass / GSTdMT mass).

<b>Elution volume</b>	<b>Calculated MW (Da)</b>	<b>x dMT MW</b>
116.16	114474	3.16
133.71	72802	1.99

## 4.9 Solution X-ray scattering on dMT and GSTdMT

### 4.9.1 Solution X-ray scattering on GSTdMT

The scattering curves of 5.85 and 1.44 mg/ml GSTdMT solutions (Figure 4.34) indicate that the purified protein is highly aggregated and the Guinier region ( $0.027 < s < 0.14 \text{ nm}^{-1}$ ) did not yield a meaningful value for the radius of gyration for the monomer.

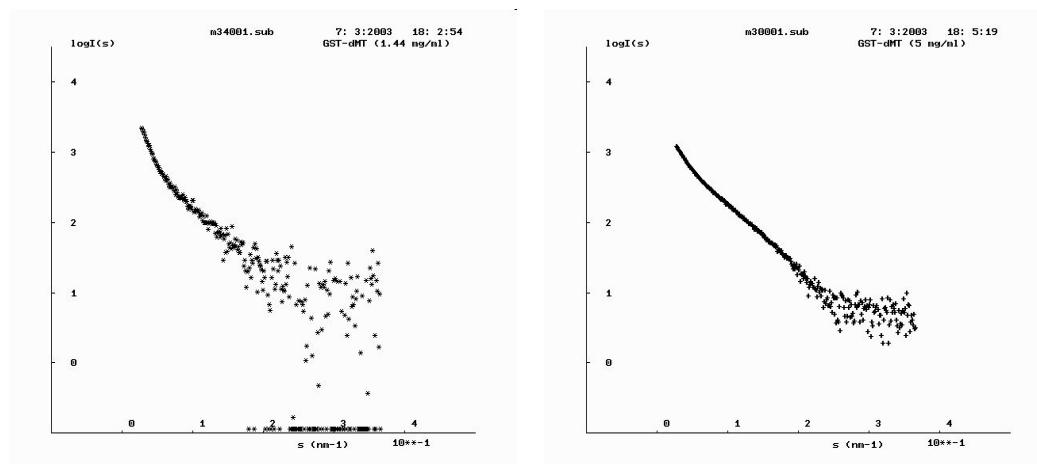


Figure 4.34 (a) and (b): Scattering patterns for GST-dMT solutions at 1.44 and 5.9 mg/ml concentrations respectively. The buffer was 50mM HEPES, pH 8.0 and 150 mM NaCl. Note that the scattering curves indicate aggregated protein in the range of the scattering vector  $0.027 < s < 0.14 \text{ nm}^{-1}$ .

To investigate if aggregation was due to preparation conditions and high protein concentration, GST-dMT fusion protein was isolated a second time using a HiTrap GST Fast Flow (5 ml) column (Amersham-Pharmacia) and the purified protein was first analyzed using SDS PAGE. A single band was observed on 12% polyacrylamide gels indicating that the sample was homogeneous (data not shown). Solutions for X-ray scattering measurements were prepared at lower protein concentrations. However, the

scattering patterns were dominated by aggregates (Figure 4.35) and an estimate for the radius of gyration of GST-dMT could not be obtained.

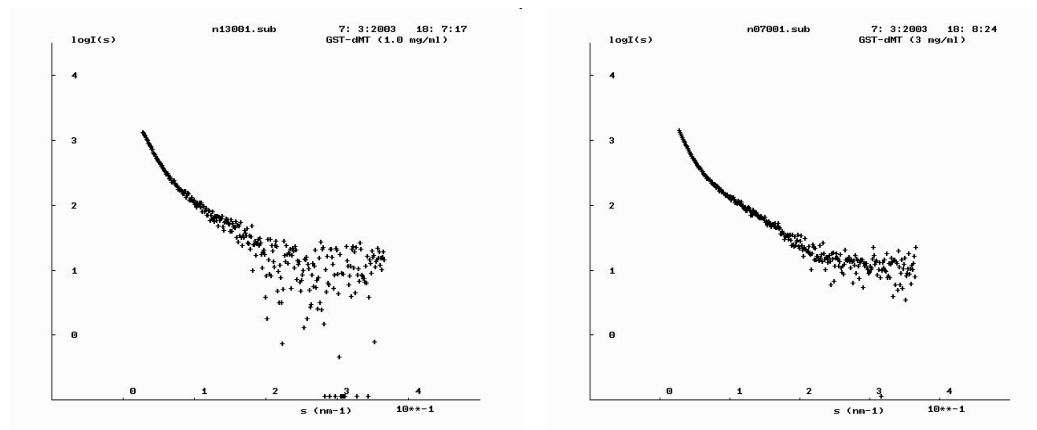


Figure 4.35 (a) and (b): Scattering patterns for GST-dMT solutions at 1.0 and 3.0 mg/ml concentrations. The buffer was 50mM HEPES, pH 8.0 and 150 mM NaCl. Note that the scattering curves indicate aggregated protein in the range  $0.027 \leq s \leq 0.14 \text{ nm}^{-1}$ .

This result was confirmed by native gel electrophoresis (Figure 4.36) a minimum of 4 bands were identified indicating complexes of different size and aggregation states of GST-dMT in solution.

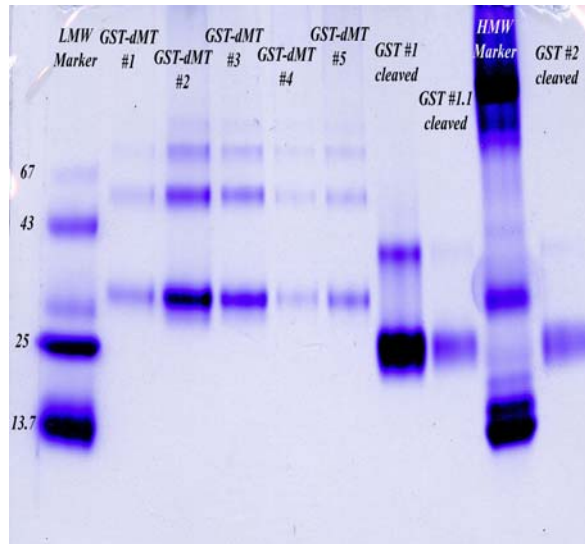


Figure 4.36: 4-20% native tris-glycine gel analysis of different fractions of column purified GST-dMT. Low and high molecular mass markers are shown in 1<sup>st</sup> and 9<sup>th</sup> lanes.

#### 4.9.2 Solution X-ray scattering on dMT

To obtain a monodisperse solution of dMT for X-ray scattering measurements, the overexpressed GST-dMT recombinant protein was cleaved and the two components were separated. Cleavage was carried out using thrombin protease on the affinity column and the dMT component eluting in the flow-through was collected in different fractions. These fractions were combined according to native gel analysis. X-ray scattering measurements carried out on this sample at a concentration of 2.5 mg/ml again indicated aggregation (data not shown).

Complexes of dMT were fractionated by gel filtration chromatography with a G75 Superdex 16/60 column (Pharmacia). Three peaks eluting at 52 ml, 68 ml, and 78 ml correspond to 80 kDa, 29 kDa, and 16 kDa respectively. Native gel analysis indicated a single major component in the 16 kDa fraction. However, the 16 kDa fraction needs to be further characterized to determine if the major component is a monomer or a dimer of dMT.

The signal to noise ratio of the scattering pattern (Figure 4.37) was very low due to the low protein concentration in this sample. So far, concentration determination of recombinant protein was based on the extinction coefficient of GST. After the separation of dMT and GST as there are no extinction coefficient known for dMT estimation of concentration can only be made on the basis of relative absorbance measurements or relative intensity of the bands observed on native gels.

Radius of gyration for the major component in the 16 kDa fraction was found to be 2.2 nm.

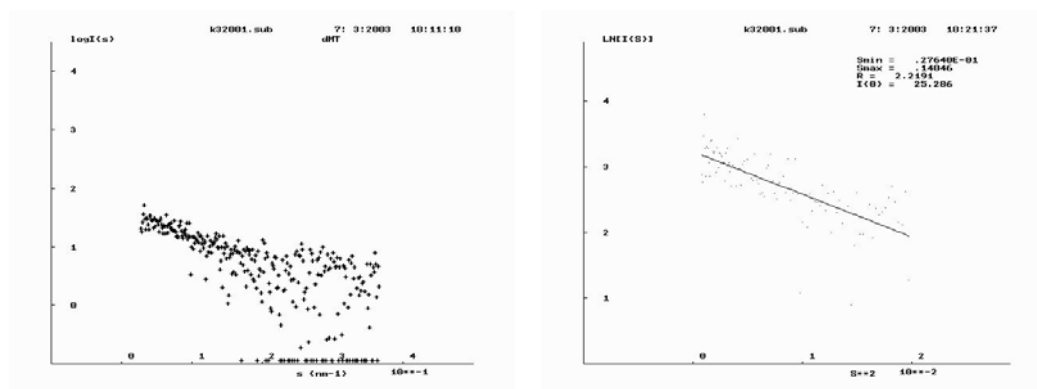


Figure 4.37 (a) and (b): Scattering pattern (a) and the Guinier plot (b) for the 16 kDa dMT fraction.

#### 4.10 Prediction of wheat MT structure and function

In the literature there are 27 entries for MT structure (Table 2.3) of various organisms; however there is no entry for any plant MT. Except type 4, all other type of plant MTs have cystein residue distribution similar to that of mammalian class I MTs.

wMT has a similar cystein distribution pattern with rat liver MT (4MT2), which is a structurally characterized mammalian MT protein. Indeed, wMT's  $\alpha$ - and  $\beta$ -domains have high sequence similarity with metal binding domains of 4MT2.

4MT2 has 2 metal binding domains and a small hinge region containing 2 lysine residues (Figure 4.38). This two domains structure should also be seen in wMT, where structure for metal binding domains could be easily predicted using homology between sequences of metal binding domains.

Homology modeling and heuristic approaches were used to predict the structure of metal binding domains and hinge region, respectively.

The predicted structure with sequence data were used to predict wMT functions other than metal scavenging. Prediction results indicate a possible DNA binding and/or protein-protein interaction role for wMT.

##### 4.10.1 wMT structure prediction

A high sequence similarity between wheat and rat liver MT (4MT2) has been observed except in the hinge region connecting the two metal binding domains (Figure 4.38). Wheat MT (wMT) hinge region contains 42 residues, whereas 2-3 residues exist in mammalian MTs. wMT, therefore, was divided into 3 functional parts; metal binding  $\alpha$ - and  $\beta$ -domains, and hinge region. Structure predictions were done for each of these functional parts separately.

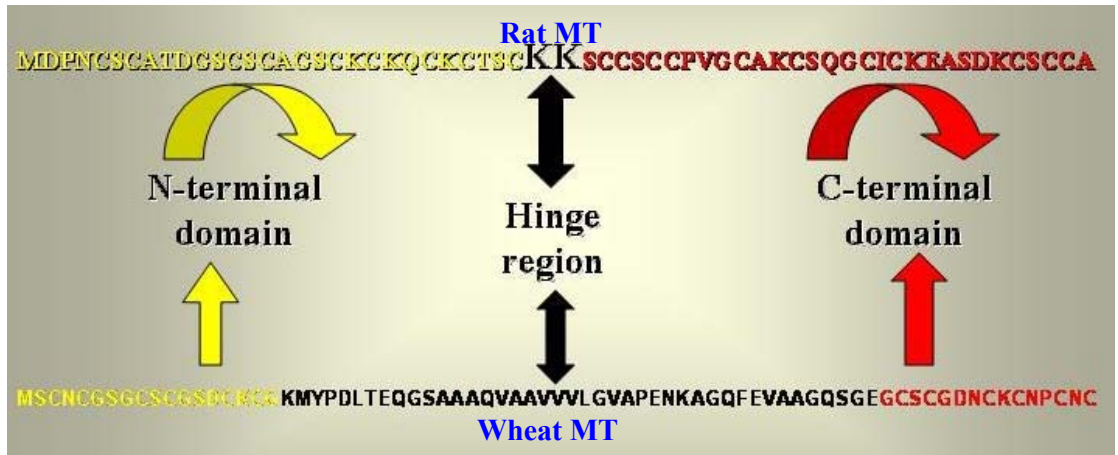


Figure 4.38: Rat MT (4MT2) and wheat MT protein sequences. Two metal binding domains and hinge regions are indicated.

#### 4.10.1.1 Secondary structure prediction

Strong secondary structure features were searched within the sequence of wMT. Such features would be reference points during the validation of the predicted wMT structure.

Secondary structure features for wMT were predicted using several algorithms available at Biology Workbench 3.2 (<http://biowb.sdsc.edu/>) (Figure 4.39).



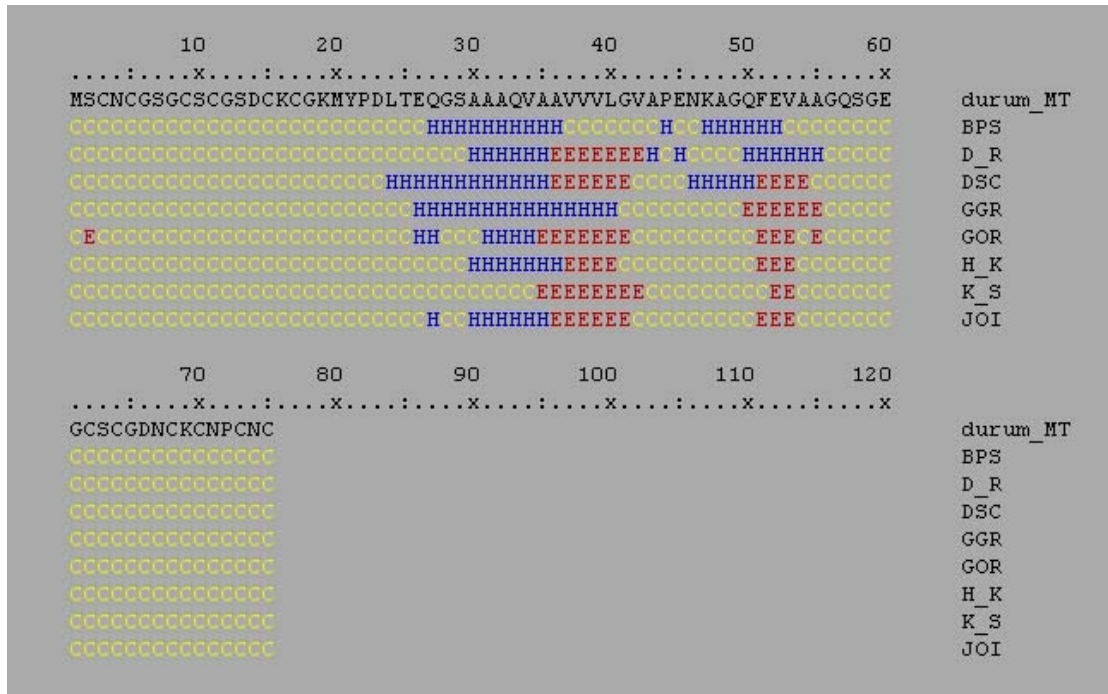


Figure 4.39: Predicted secondary structure features for wMT. Used algorithms are indicated (right) and described in Materials and Methods. (H for alpha helices; E for beta strands; C for coils)

Although a helix-turn-helix motif is quite obvious for the wMT hinge region, used algorithms could not give a proper prediction for the metal binding domains. These results, non-existence of strong secondary features, indicate the importance of metal binding for the stabilization of  $\alpha$ - and  $\beta$ -domains. This is also a key properties for DNA binding and interacting proteins like transcription factors, suppressors, repressors.

#### 4.10.1.2 Modeling of $\alpha$ - and $\beta$ -domains

As stated above there is no similar sequence within the known MT structures that could be used as template structure for the homology modeling studies. wMT metal binding domains, therefore, were separately searched in protein databases.

In accordance with the alignment results (Figure 4.40) metal-cystein distances were taken from sea urchin MT  $\beta$ -domain (1QJL) and rat liver MT  $\beta$ -domain (2MRT).

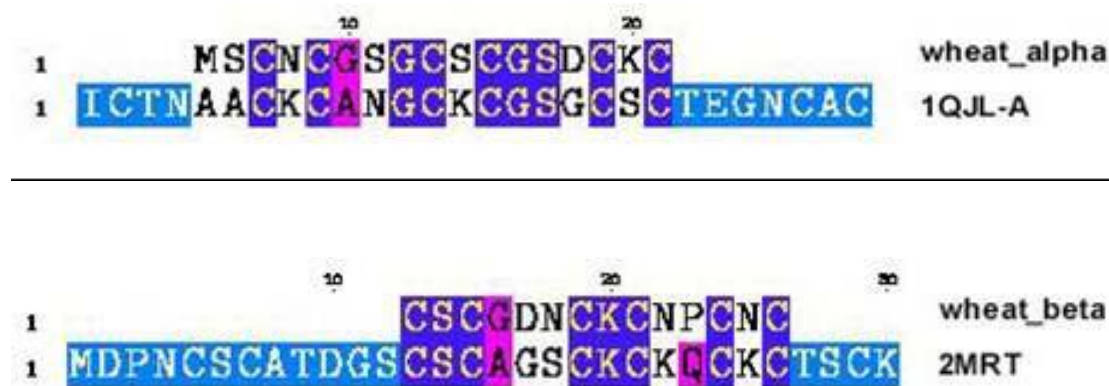


Figure 4.40: Pairwise alignments of wheat MT alpha with sea urchin MT beta domains; and wheat MT beta with rat liver MT beta domains.

These sequence similarities were used for the modeling of wheat  $\alpha$ - and  $\beta$ -domains. In order to model the protein in its native form, functional residues must have been preserved. For this purpose, the cystein-metal distances were fixed and this was the only constraint that was applied in the model.

The residues in the template sequence were mutated to the ones in target sequence. The side chains of the new sequence were placed in a way to keep the global energy of the overall structure at minimum (Figure 4.41 and Figure 4.42).

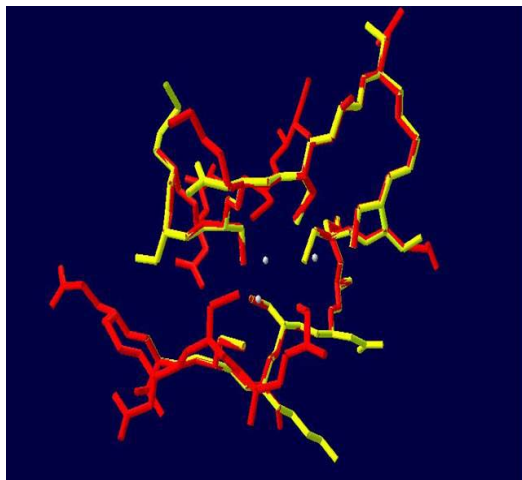
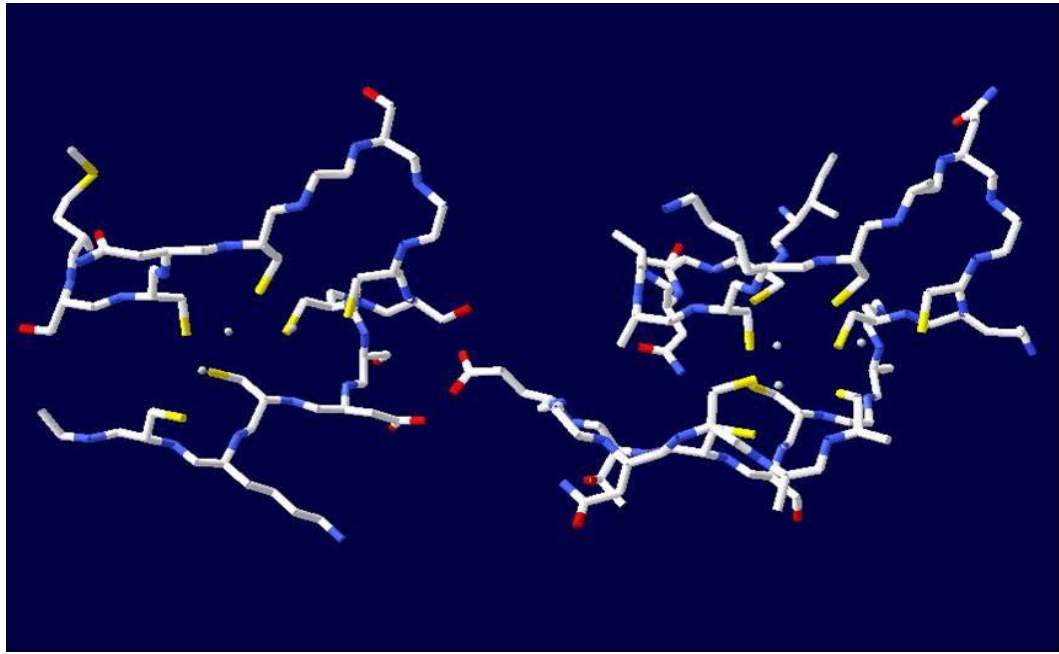


Figure 4.41: Wheat MT alpha domain (top-left) with sea urchin MT beta domain (top-right). Superimposed image of modeled wheat MT alpha domain (yellow) onto sea urchin MT beta domain (red).

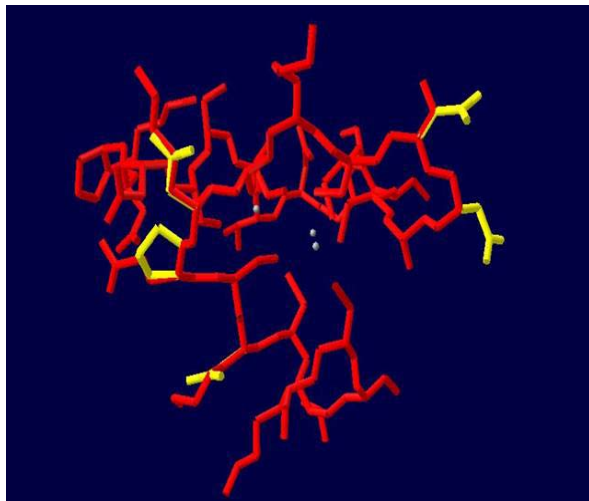
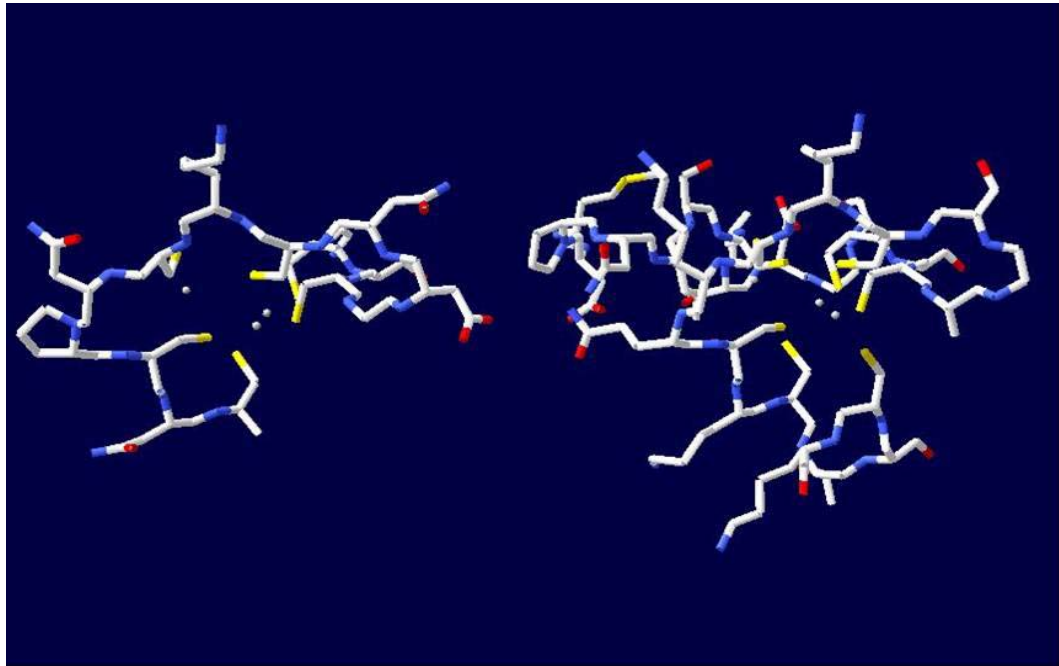


Figure 4.42: Wheat MT beta domain (top-left) with rat liver MT beta domain (top-right). Superimposed image of modeled wheat MT beta domain (yellow) onto rat liver MT beta domain (red).

#### 4.10.1.3 Modeling the hinge region

The structure of the wMT hinge region could not be predicted using homology modeling approach because no similar sequence was detected within known MT structures and sequence similarity searches in Protein Structure Databank did not give an answer with high confidence.

The hinge region sequence was also searched against folding libraries to find small but similar sequences from known structures. However, a confident and applicable solution was not extracted from results (data not shown).

Heuristic structure prediction methods, on the other hand, do not need any template, but predict the structure from the available sequence information. HMMSTR is a hidden Markov model for local and secondary structure prediction, based on the I-sites library of sequence-structure motifs. The given protein sequence is divided into small fragments and each fragment is searched against a set of fragments with known folding. Generated and sorted possible fragments are then processed with ROSETTA, which is a Monte Carlo Fragment Insertion protein folding program (Simons *et al.*, 1997).

wMT hinge region sequence was upload to the “I-sites/HMMstr/Rosetta Prediction Server”. Local structures were determined by HMMSTR (Table 4.10) and structure predictions for wMT hinge region were obtained by ROSETTA (Figure 4.43).

Table 4.10: Detected sequence fragments, PDB file name, referring local structure, clustering group and confidence value given respectively.

SeqNo	Sequence	Paradigm	LocalStruct	Cluster	Confidence
17	AAVVV	3tgl _ 25	HHHHH	5252	0.60
16	VAAVV	3tgl _ 25	HHHHH	5252	0.59
18	AVVVL	3tgl _ 25	HHHHH	5252	0.58
17	AAVVVLG	2cmd _ 70	BHEEEeG	7136	0.57
15	QVAAV	3tgl _ 25	HHHHH	5252	0.57
14	AQVAA	3tgl _ 25	HHHHH	5252	0.56

19	VVVLGVA	lice A 278	EBdEEEE	7428	0.55
18	AVVVLGV	2dld A 228	BBLBEEE	7312	0.55
13	AAQVA	3tgl _ 25	HHHHH	5252	0.55
18	AVVVLGV	lice A 278	EBdEEEE	7428	0.53
12	AAAQV	3tgl _ 25	HHHHH	5252	0.53
27	ENKAGQFE	lhgx A 18	BHHHHHHH	8055	0.52
19	VVVLG	3tgl _ 25	HHHHH	5252	0.52
4	PDLTEQGS	lhgx A 18	BHHHHHHH	8055	0.52
31	GQFEVAAG	lhgx A 18	BHHHHHHH	8055	0.51
18	AVVVL	lytb A 154	bELBE	5116	0.51
14	AQVAAVVV	lhgx A 18	BHHHHHHH	8055	0.51
3	YPDLT	3pmg A 260	lEBbE	5108	0.51
33	FEVAA	3tgl _ 25	HHHHH	5252	0.50
29	KAGQFEVAAGQS	1frp A 36	HHHHHHHHHHH	12005	0.50
17	AAVVVLG	lice A 278	EBdEEEE	7428	0.50
19	VVVLGV	lice A 278	EBdEEEE	6288	0.49
7	TEQGSAAA	lhgx A 18	BHHHHHHH	8055	0.49
22	LGVAP	lisd A 499	HGdBB	5107	0.48
20	VVLGVAP	2dld A 228	BBLBEEE	7312	0.48
20	VVLGV	3tgl _ 25	HHHHH	5252	0.48
19	VVVLGVA	2dld A 228	BBLBEEE	7312	0.48
19	VVVLGV	lice A 280	dEEEEb	6139	0.48
19	VVVLGV	lhqa A 46	EEEEEE	6037	0.48
19	VVVLG	lisd A 499	HGdBB	5107	0.48
17	AAVVV	lytb A 154	bELBE	5116	0.48
11	SAAAQ	3tgl _ 25	HHHHH	5252	0.48
6	LTEQGSAAAQVAA	lubs A 78	BEHHHHHHHHHHH	13010	0.48
32	QFEVA	lytb A 154	bELBE	5116	0.47
21	VLGVAP	lhqa A 46	EEEEEE	6037	0.47
16	VAAVVVLG	lhgx A 18	BHHHHHHH	8055	0.47
16	VAAVVVL	lice A 278	EBdEEEE	7428	0.47
15	QVAAVVVL	lhgx A 18	BHHHHHHH	8055	0.47
15	QVAAV	litg _ 164	HHHGH	5032	0.47
37	AGQSG	3tgl _ 25	HHHHH	5252	0.46
33	FEVAA	larv _ 213	HHHHH	5040	0.46
28	NKAGQFEV	lhgx A 18	BHHHHHHH	8055	0.46
26	PENKAGQFEVAAG	lubs A 78	BEHHHHHHHHHHH	13010	0.46
17	AAVVVL	lice A 278	EBdEEEE	6288	0.46
9	QGSAA	3tgl _ 25	HHHHH	5252	0.46
4	PDLTE	lvhh _ 154	HGLbH	5236	0.46
32	QFEVA	lsbp _ 35	lBEEE	5097	0.45
20	VVLGVAP	lice A 278	EBdEEEE	7428	0.45
19	VVVLG	lsbp _ 35	lBEEE	5097	0.45
10	GSAAA	3tgl _ 25	HHHHH	5252	0.45
8	EQGSAAAQVAAV	1frp A 36	HHHHHHHHHHH	12005	0.45

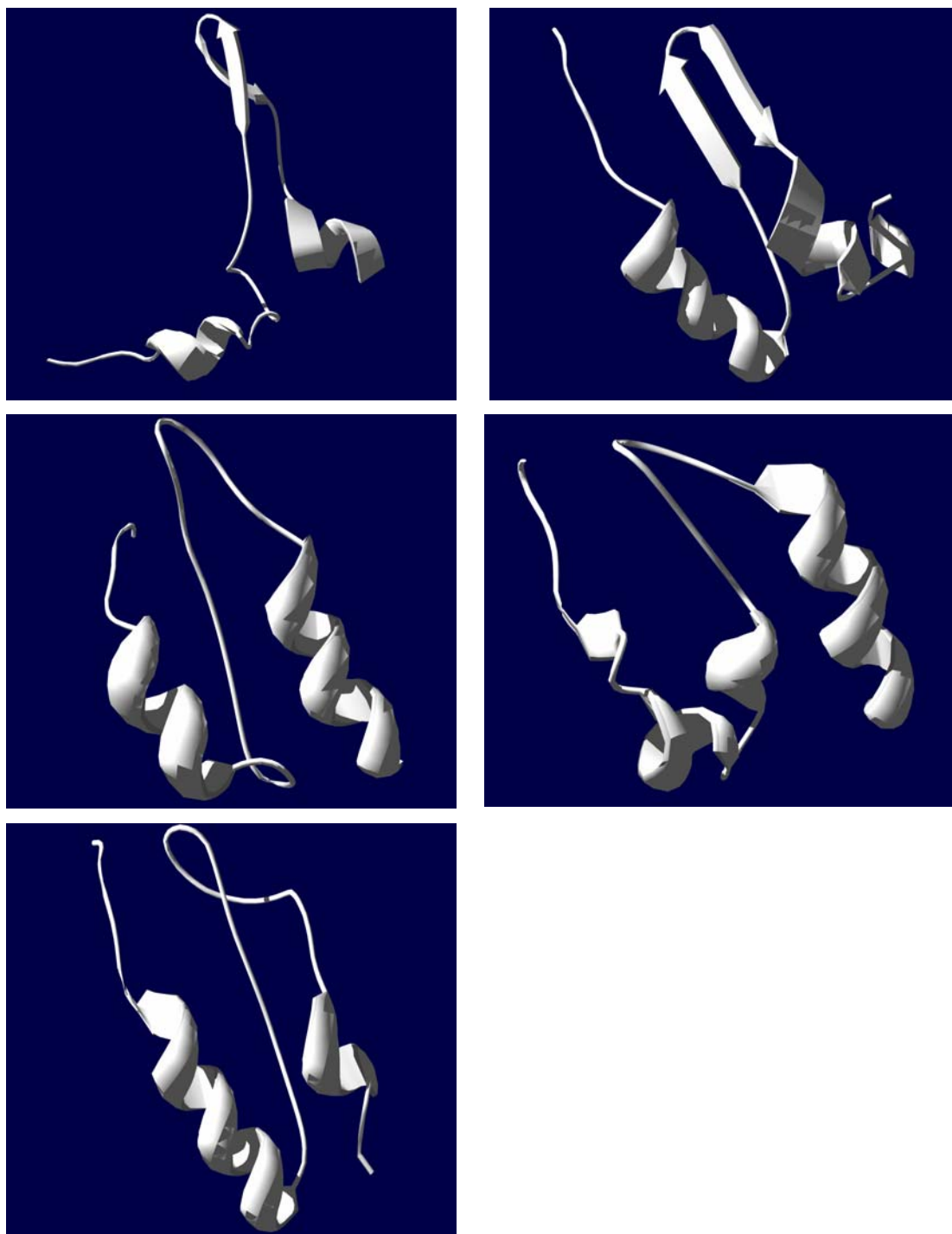


Figure 4.43: Predicted coordinates were processed and visualized with DeepView.

Predicted structures converged into two groups; one contained larger loop regions that allow the metal centers to fold easily, and the other with beta sheets that resulted in rigid structures (Figure 4.44).

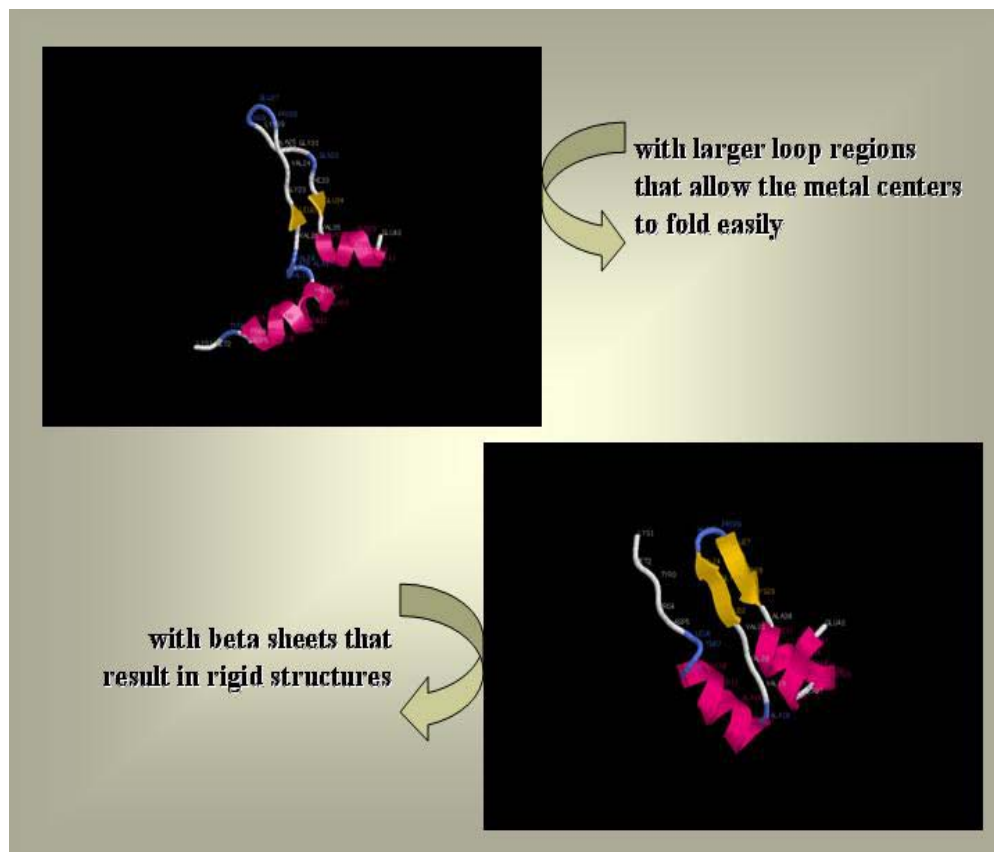


Figure 4.44: Two cluster of predicted hinge regions, one representative from each cluster shown and indicated.

#### 4.10.1.4 Completing the puzzle

Separately modeled wMT  $\alpha$ - and  $\beta$ -domains were connected with the selected hinge region structures using Deep View – The Swiss-Pdb Viewer 3.7 (Figure 4.45).



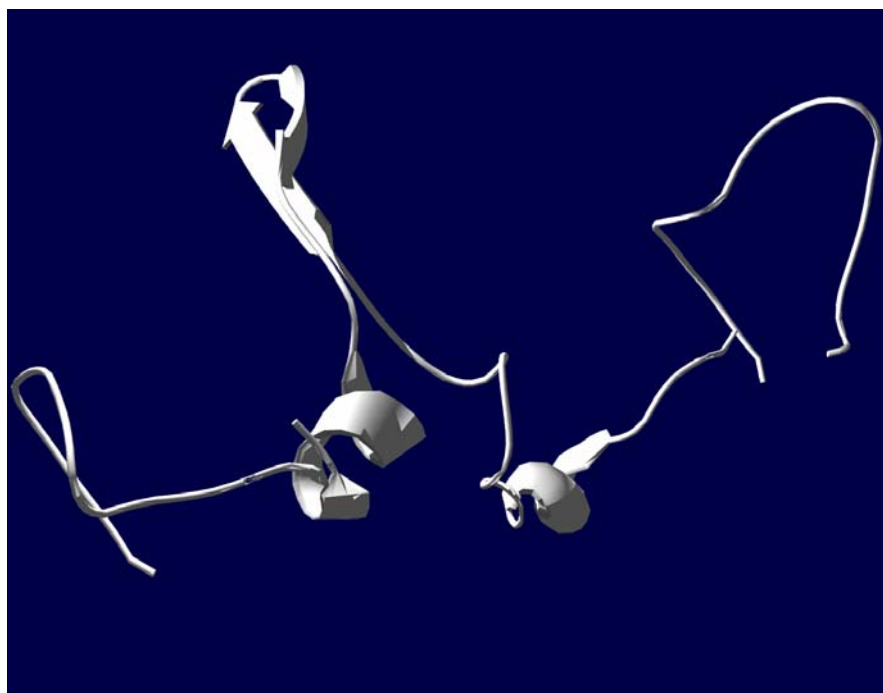


Figure 4.45(a): Ribbon representation of the modeled wMT protein with the relaxed hinge region.

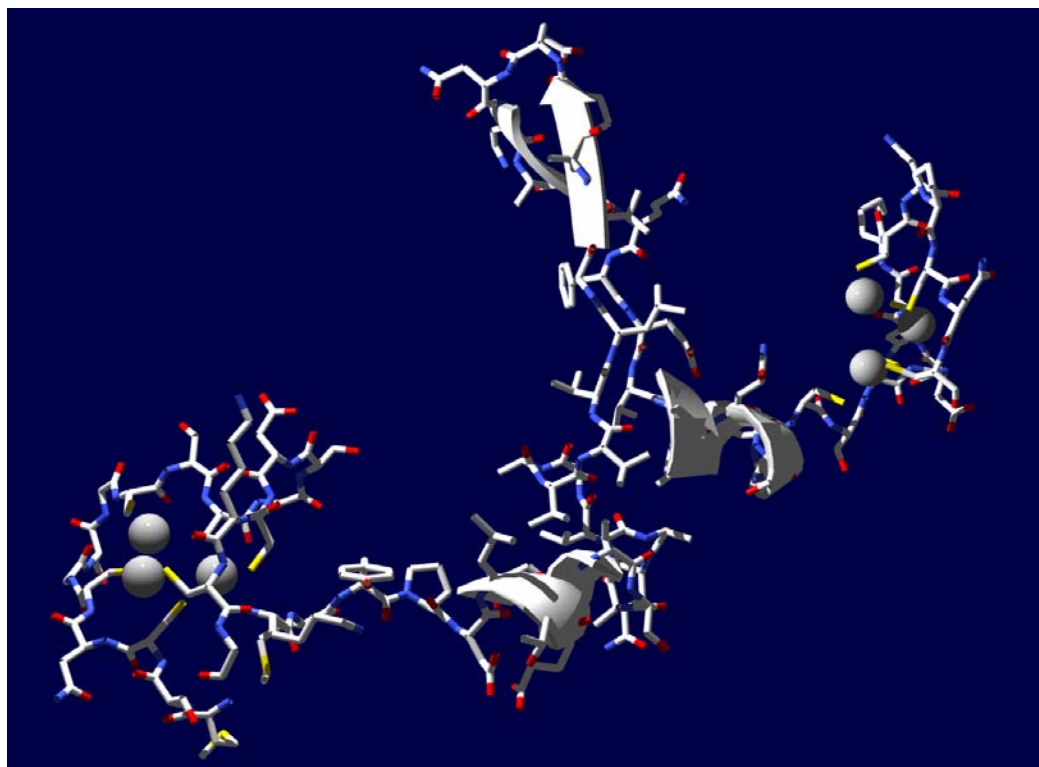


Figure 4.45(b): Molecular representation of the modeled wMT protein with the relaxed hinge region.

## 4.10.2 Function prediction

Functional roles for wMT were investigated using the predicted model structures.

### 4.10.2.1 Conservation of hinge region among plant species

As stated in previous parts, the main difference between plant and mammalian MTs is the long hinge region that connects the two metal centers in plants. A multiple alignment was run for known plant MT hinge regions in order to see conserved motifs (Figure 4.46) and a rooted tree was constructed from this alignment (Figure 4.47).

```

      10      20      30      40      50
1  KMYPDLGFSGETTTET..FVLGVAPAMKNQYEASGES.NNAES..DA  arabidopsi
1  KMYPDLGFSGETTTET..FVLGVAPAMKNQYEASGES.NNAEN..DA  arabidop_1
1  KRYPDL....ENTATET..LVLGVAPAMNSQYEASGET.FVAEN..DA  arabidop_2
1  KRYPDL....ENTATET..LVLGVAPAMNSQYEASGET.FVAEN..DA  arabidop_3
1  KMYPNLSYS.E.AAEP..LVLGVAPQKNTY.GGSVEE.VTAEN..G  avicennia_
1  MYPELSYT.ENTAAET..LVLGVAPPKTTY.LEGAGEE.AAAEN..GG  coffee_P43
1  GKRYPDLEET.STAAQP..TVVLGVAPEK..KAAPEFVE.AAAES..GEAAHG  maize_S175
1  KMYPDLTEQGSAAAQVAAVVVLGVAPEN..KAG.QFEV.AAGQS..GE  MT_wheat_h
1  KMFPDVEATAGAAA....MVMPTASHK..GSSGGFEM.AGCE..GG  barley_S16
1  NMYPDLGYSEATAPAE..ALVLGVAPLKF.YEGAESVE.GAAEN..G  avicenni_9
1  GMYPDLEEKSGATMQVT.VIVLGVGSAKV.QFEEAAEF.GEAH..G  barley_2_S
1  KMYPDLSYN.ESTTTET..LVLGVGHEKTSFGTMEMGESPAEN..G  tobacco_2_
1  KMYPDLSYT.ESTTTET..LVLGVGPEKTSFDAMEFGESLIAEN..G  belladona_
1  NKRSSGLNYV.EAETTET..VILGVGPAKIQFEDAEMG..VAAED..SG  clover_CAA
1  NKRSSGLSYS.EMETTET..VILGVGPAKIQFEGAEMS..AASED..GG  pea_S09098
1  KMNADLSYT.ESTTTET..LVMGVGSAKAQFEGAEMG...AES..GG  clover_2_C
1  GMYPDL....EKSTTET..IIQGVAP.MNNFEELGEKA.A.E.GGNG  tobacco_AA
1  GMYPDV....EKSTTFS..VVQGVAP.MNNYGRVEESA.A.AKGGNG  bellado_17
1  SMYPDM....ETNTVT..MIEGVAP.LKMYSEGSEKS.FGAE.GGNG  monkey_flo
1  NRYPDLSSSGESTTSET..LIMGVAPQKRHFD.GAEME.AGAEN..G  avocado_CA
1  GMYPDIS.MEQSTTET..LIMGVAPQTSHE.ASEM.VVAEN..G  peach_CAB5
1  KMYPDMANGASST.AT..LVTGVAPKISYFDNGSEMG.VGAEN.DG  iceplant_T
1  KMYPDIA.EKSTTTSET..MLLGVAPQKGFEG.FEMV.AEKED..GG  cattail_AA
1  FMLTDLG.EERSSTSQT..MIMGVAPQKGFEE.LETA.AGSDN..G  banana_022
1  VKKGNS..YGDIVVETEKSYVATVVMVPAAGH.EGS  grape_CAB8
1  VKKGNS..YGDIVVETEKSYVDEVIVAAEAH.DGK  banana_2_Q
1  VKKGTS..YTFDIVVETQESYKEAMIMDVGAEN.NAN  arabodopsi

```

Figure 4.46: Multiple alignment of plant MT hinge regions.

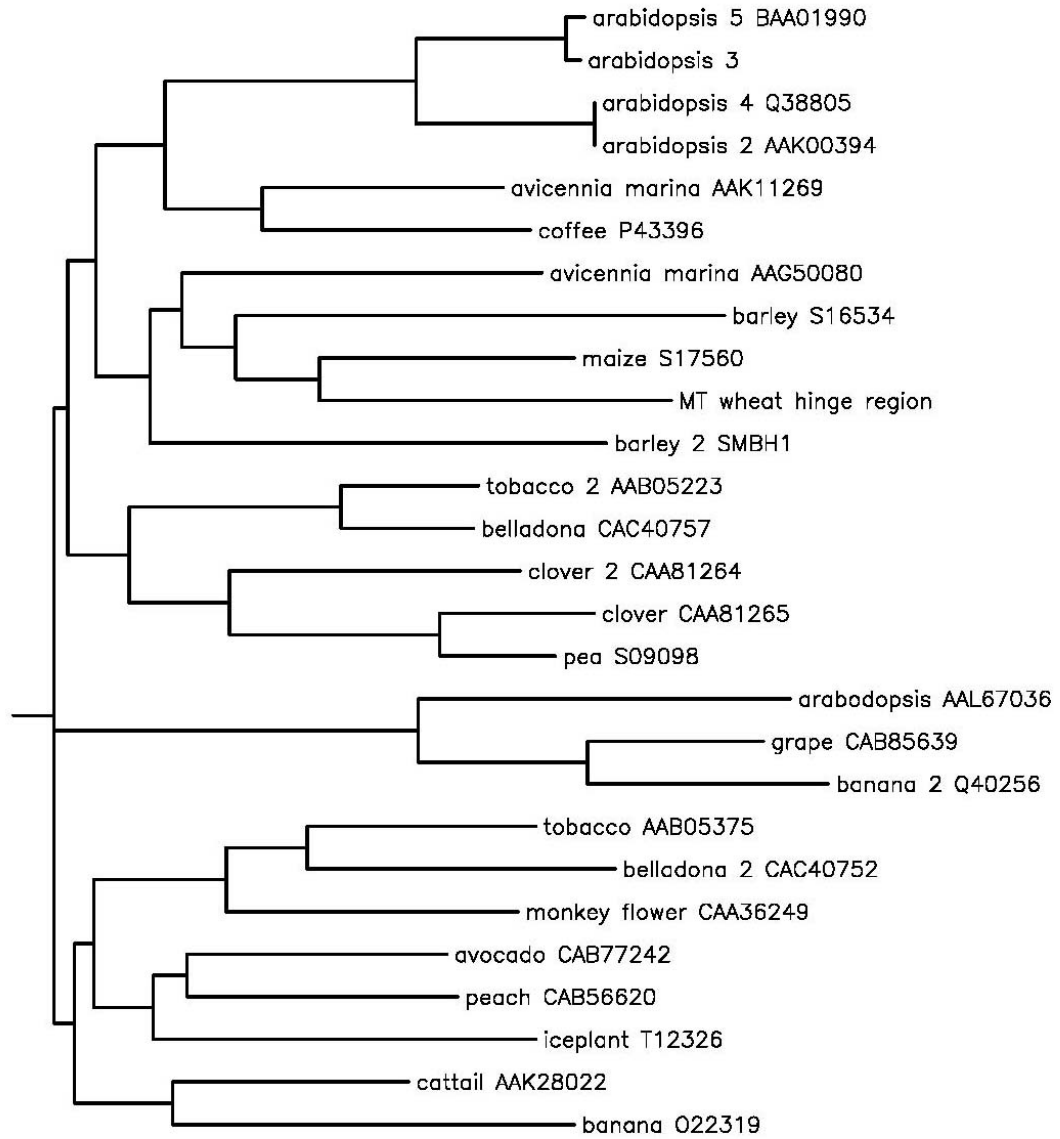


Figure 4.47: A rooted tree for plant MT hinge regions based on sequence similarities.

#### 4.10.2.2 Functional motif search

Sequences similar to wMT hinge region (Figure 4.47) were selected and a multiple alignment was performed for those selected sequences (Figure 4.48) in order to find out most conserved residues.



Figure 4.48: Multiple alignment of selected plant MT hinge regions.

This multiple alignment was refined and an amino acid pattern was constructed;

K-x-[FY]-P-D-x(1,2)-E-x(2,3)-A-[AGT]-x(2)-[ATV]-x-[AV]-[ILMV]-[GPV]-x-[AGL]

with 100% identity and another pattern with 80% identity;

K-x-[FY]-P-D-[LV]-[ET]-[AE]-x-[AGS]-[GST]-A-[AT]-x(2)-[TV]-x-[APV]-x-[AGV]

Pattern search in PROSITE database gave only one hit with acceptable confidence level; “MYRISTYL” pattern. The N-myristoylation site has a conserved pattern of

G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}.

This pattern, however, is too unspecific and can not be considered as a real match.

It was thus not possible to deduce a functional role for the hinge region based sequence analysis.

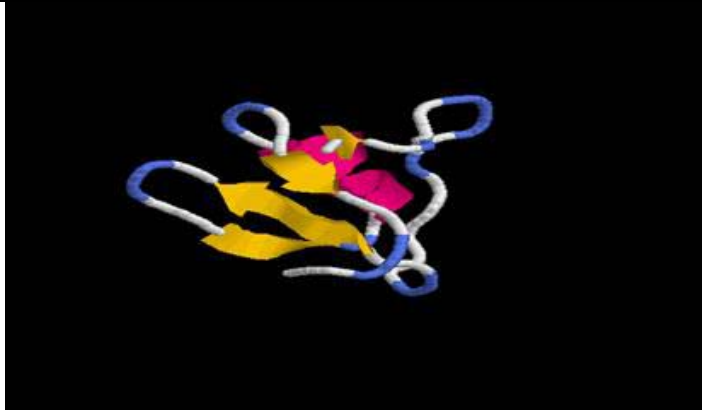
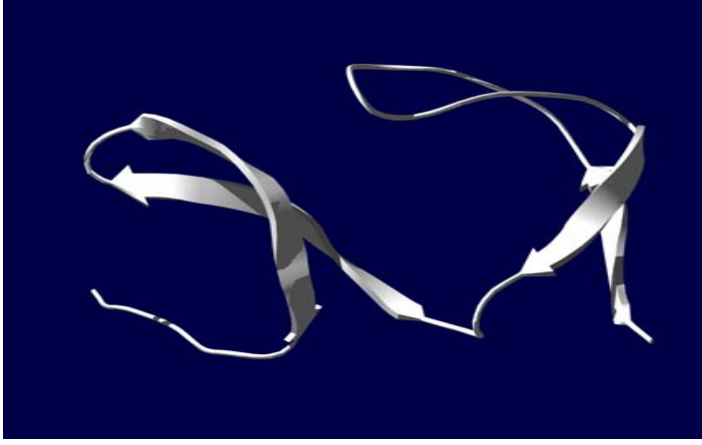
#### **4.10.2.3 Search for similar folds**



Folds that share similar structural characteristics with the wMT hinge region were found by the two following methods.

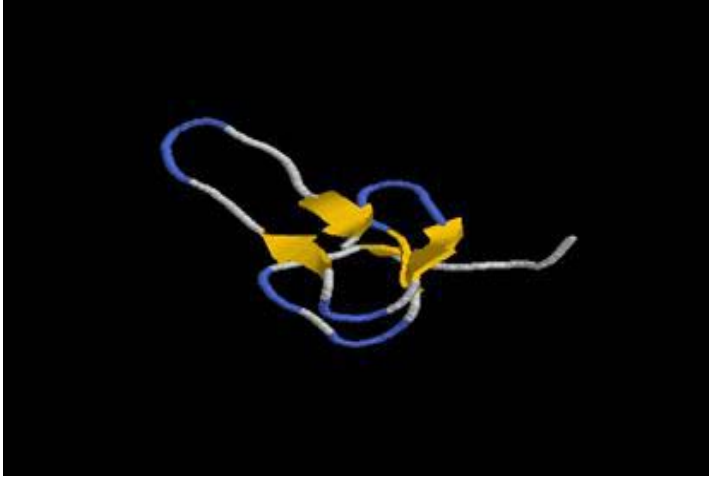
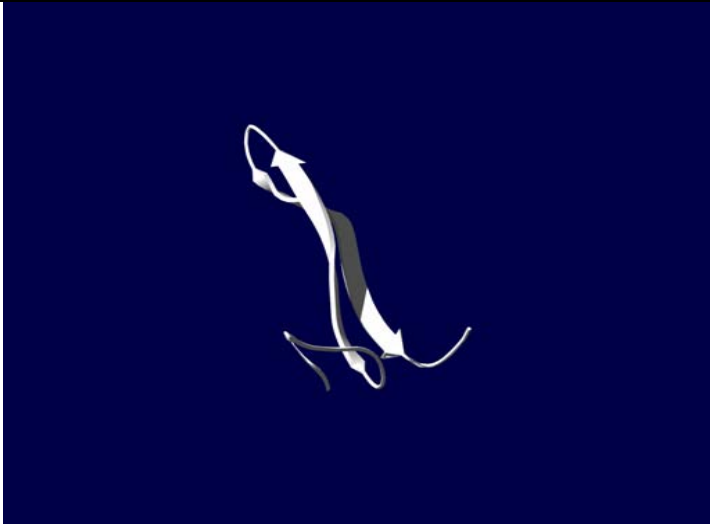
wMT hinge region sequence was submitted to “3D-PSSM Imperial College Fold Recognition Server” (Kelley *et al.*, 2000) and “3D Protein Structure Comparison and Alignment Server” (Shindyalov and Bourne, 1998).

Results were refined considering their calculated e-values for both entries (Table 4.11).


Table 4.11: Proteins sharing structural similarity with the wMT hinge region.

Identity, %	Certainty, %	Fold	Super Family	PDB Id	Structure
24	61.3	Protein binding	Trypsin inhibitor	1IW4	
21	60	SH3-like barrel	SH3-domain	1AOJ	

14	55.1	Transcription	Transcription factor wstf	1F62	 A 3D ribbon diagram of the transcription factor wstf, showing a complex, multi-domain protein structure with several loops and helices, rendered in white and light blue against a dark blue background.
17	48.9	Ribosome	Ribosomal protein 136	1KPJ_4	 A 3D ribbon diagram of ribosomal protein 136, showing a single, elongated protein chain with several loops and helices, rendered in white against a dark blue background.

18	39	Small inhibitors, toxins	Leech antihemostatic proteins	1DEC	
12	33	Protein binding	Proteinase C inhibitor	2PAI	



16	31	TNF receptor like	TNF receptor like	1D4V	
----	----	----------------------	----------------------	------	---

#### 4.10.2.4 Is wMT a “Natively Unfolded” protein?

Hydrophobicity calculations for wMT (Figure 4.49) and its hinge region was performed using Kyte and Doolittle (Kyte and Doolittle, 1982) approximation. Mean hydrophobicity for wMT was -0.037, where 0.05 for the hinge region.

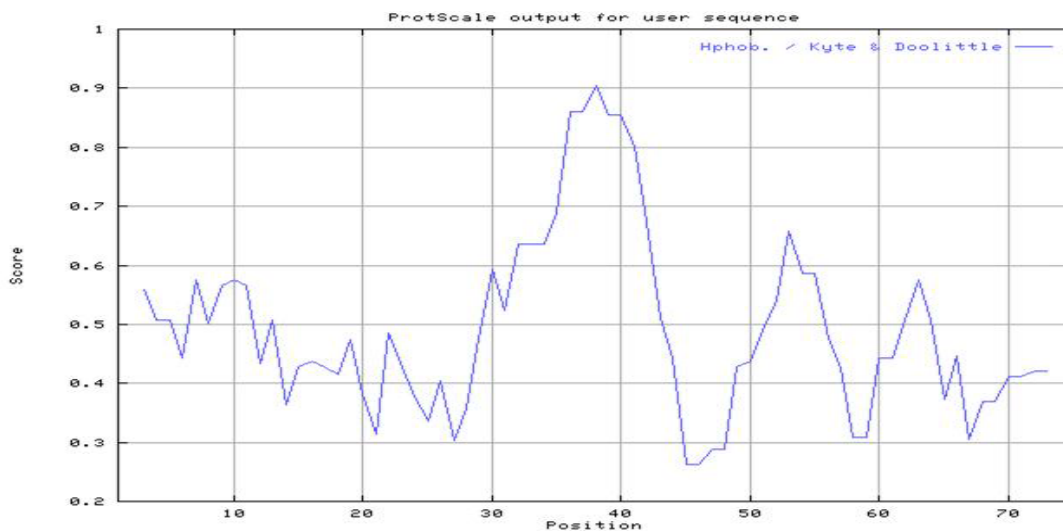


Figure 4.49: Hydrophobicity diagram for wMT according to Kyte and Doolittle approximation with corrected amino acid scale to 0 to 1.

In the wMT there were a total of 11 charged residues where 4 of them were positively charged (Arg + Lys) and 7 were negatively (Asp + Glu); where these numbers were 2 and 5 for the hinge region, respectively. Mean net charge of wMT, on the other hand, was 0.04 and of hinge region was 0.07.

These values were used to calculate the “boundary” mean hydrophobicity value,  $\langle H \rangle_b$ , which determines the probability of unfolding according to the equation generated by Uversky (2000).

$\langle H \rangle_b = (\langle R \rangle + 1.151)/2.785$ ; where  $\langle R \rangle$  is the mean net charge for the given polypeptide.

Calculations gave  $\langle H \rangle_b$  values of 0.427 for wMT and 0.438 for hinge region.

According to the comparison graph of natively unfolded and folded proteins (Uversky, 2002), wMT fall just onto the boundary region (Figure 4.50).

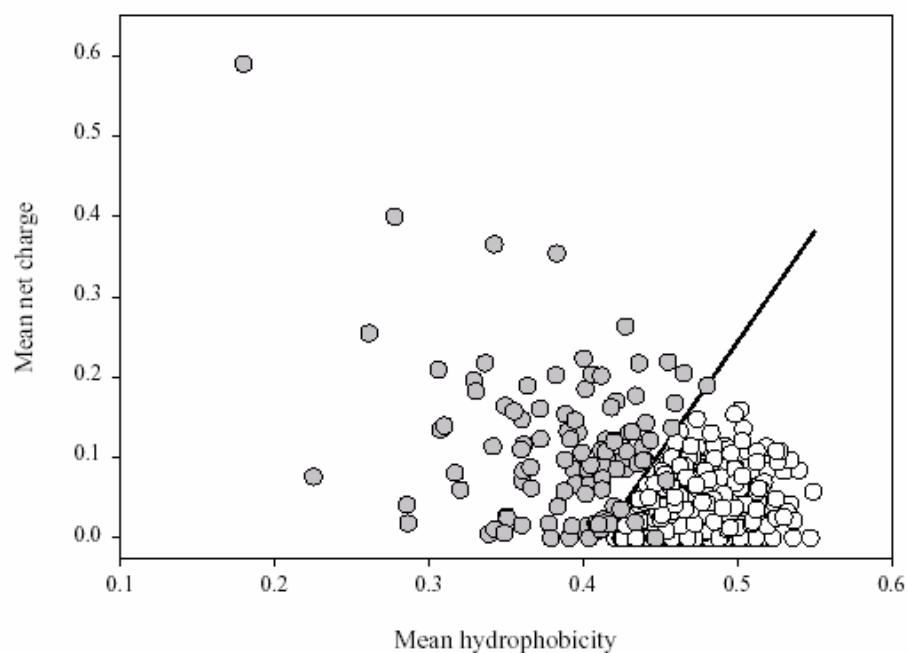


Figure 4.50: Comparison of the mean net charge and the mean hydrophobicity of folded (open circles) and natively unfolded proteins (gray circles) (Uversky, 2002).

Further calculation using “Dunker’s Lab Predictor of Natural Disordered Regions Server” (<http://www.pondr.com>) showed that the possible protein binding flanking region were disordered where this result was correlated with function prediction based on similar structure search (Table 4.11).

**Predictor of Natural Disordered Regions**  
**Keith Dunker's lab**  
<http://www.pondr.com>

MSCNCGSGCSCGSDCKCGKMYPDLTEQG**SAAAQVA**AVVVLGVAPENKAG**QFEVAAG**QSGEGCSCGDNCKCNPCNC

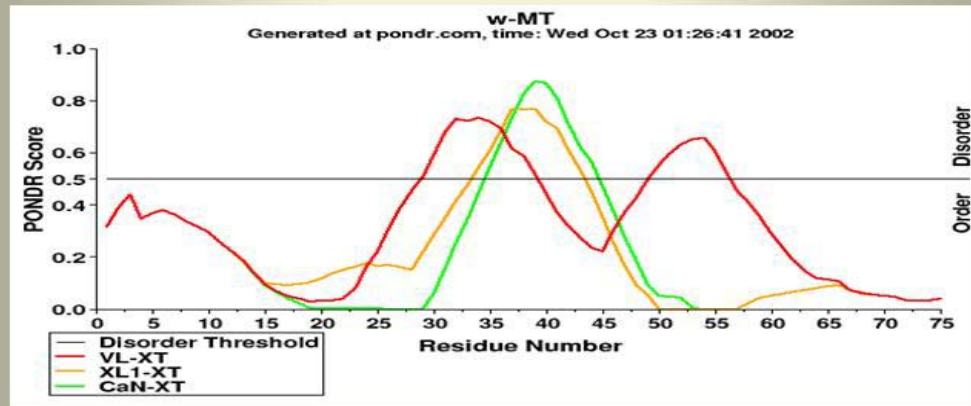


Figure 4.51: wMT disordered regions, diagram generated by “Dunker’s Lab Predictor of Natural Disordered Regions Server”.

KMYPDLTEQG**SAAAQVA**AVVVLGVAPENKAG**QFEVAAG**QSGE

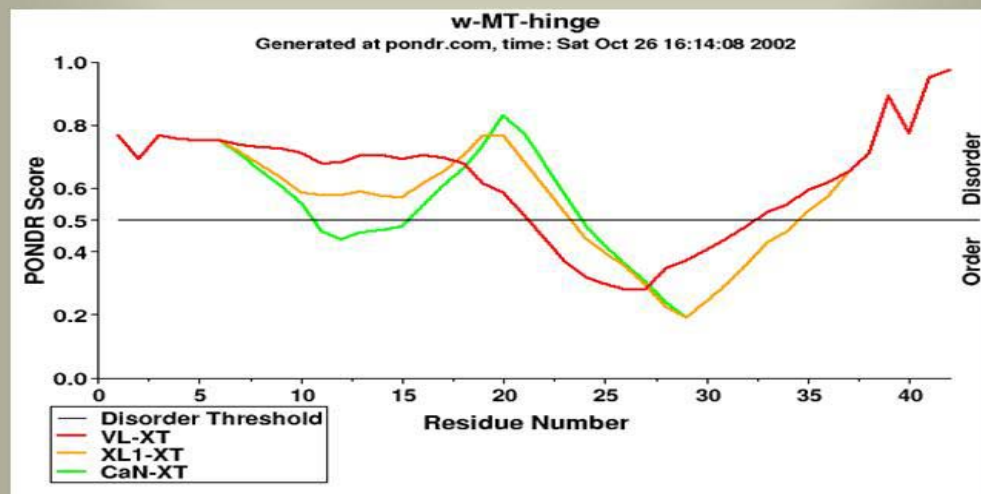


Figure 4.52: wMT hinge region disordered regions, diagram generated by “Dunker’s Lab Predictor of Natural Disordered Regions Server”.

#### 4.10.2.5 Is wMT a DNA binding protein?

Cyanobacterial trans-acting regulator SmtB protein (Figure 4.53) binds promoter region of SmtA gene region by its 2 alpha helices and alpha helices that are found in the wMT hinge region showed high sequence similarity with those (Figure 4.54).

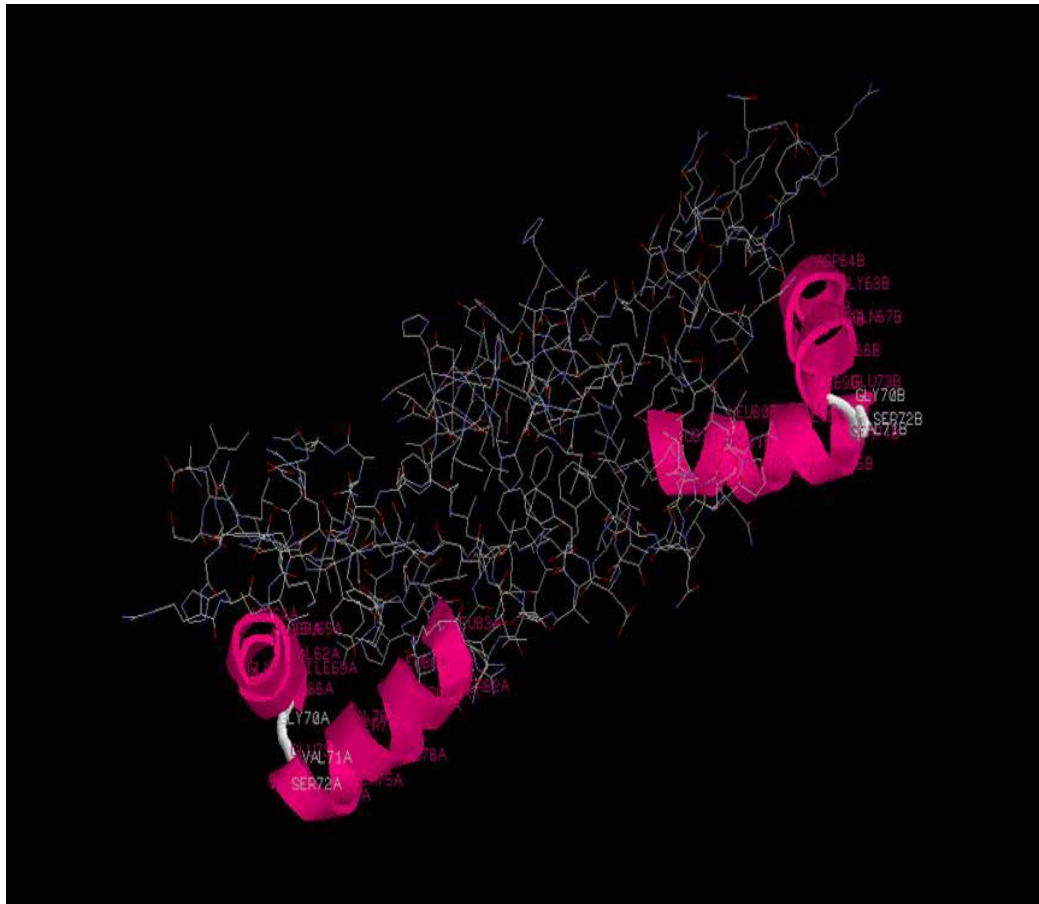


Figure 4.53: *Synechococcus* metallothionein repressor (SmtB) protein, DNA binding dimer form. DNA binding alpha helices are indicated in ribbon structure.

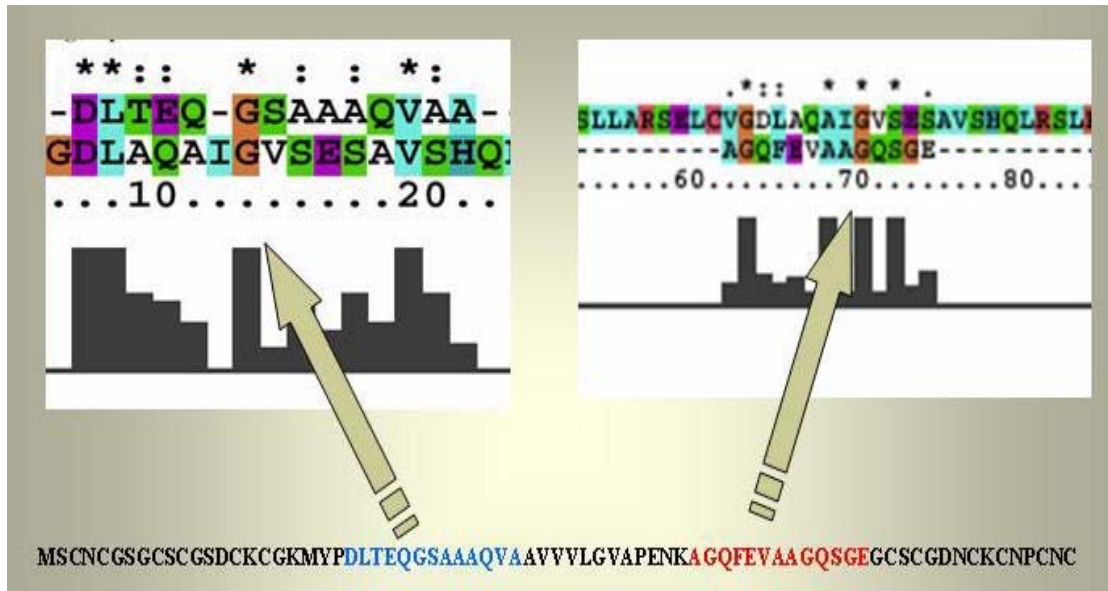


Figure 4.54: The two alpha helices of wMT aligned with those of SmtB. Similar regions are indicated with arrows.

The two metal binding domains of wMT have also DNA binding capacity. Upon metal binding several amino acids become available for such a purpose (Figure 4.55). These surface accessible amino acids were shown to have DNA binding capacities both water mediated bonds and amino acid-nucleotide base contacts (Luscombe *et al.*, 2001).

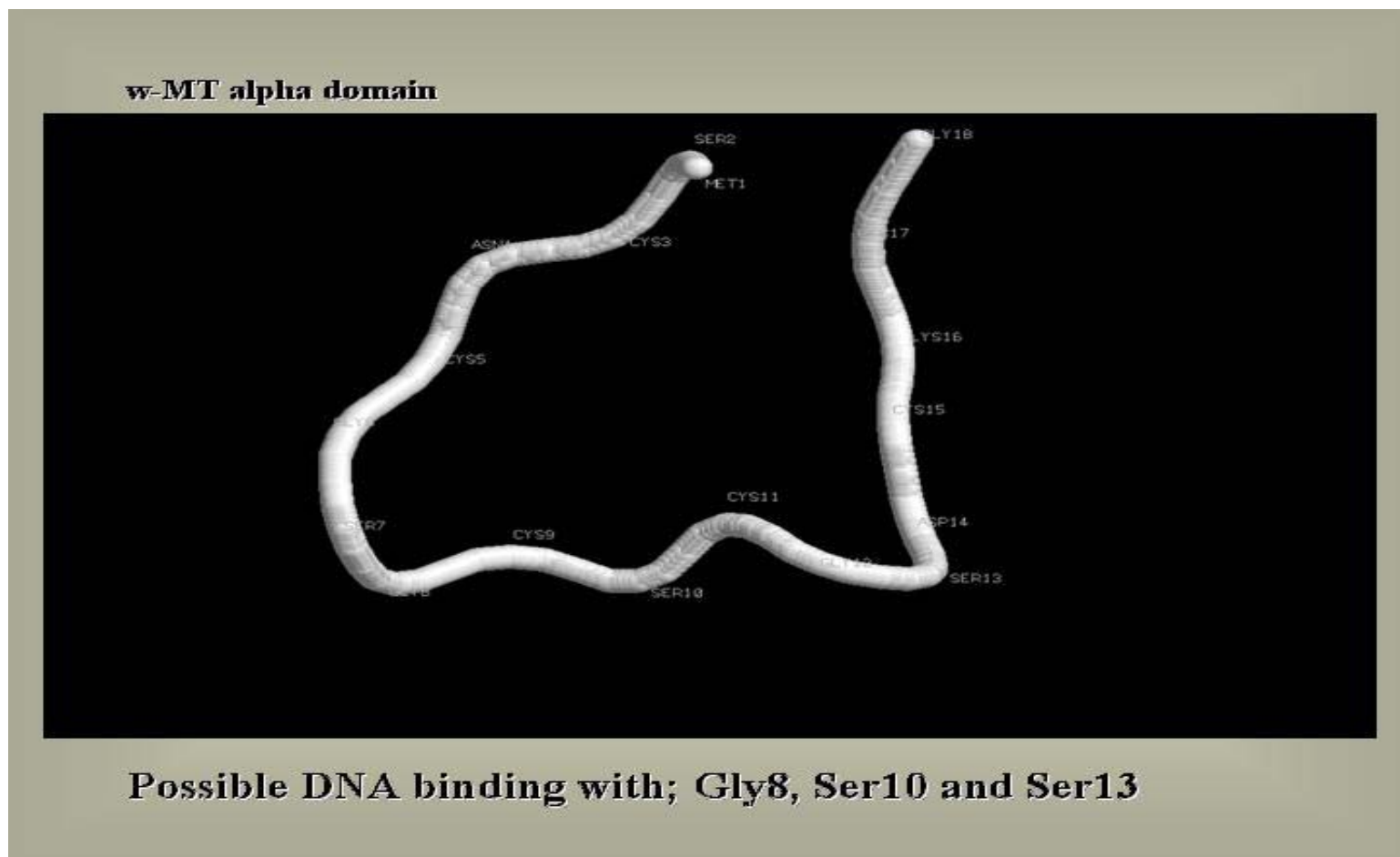
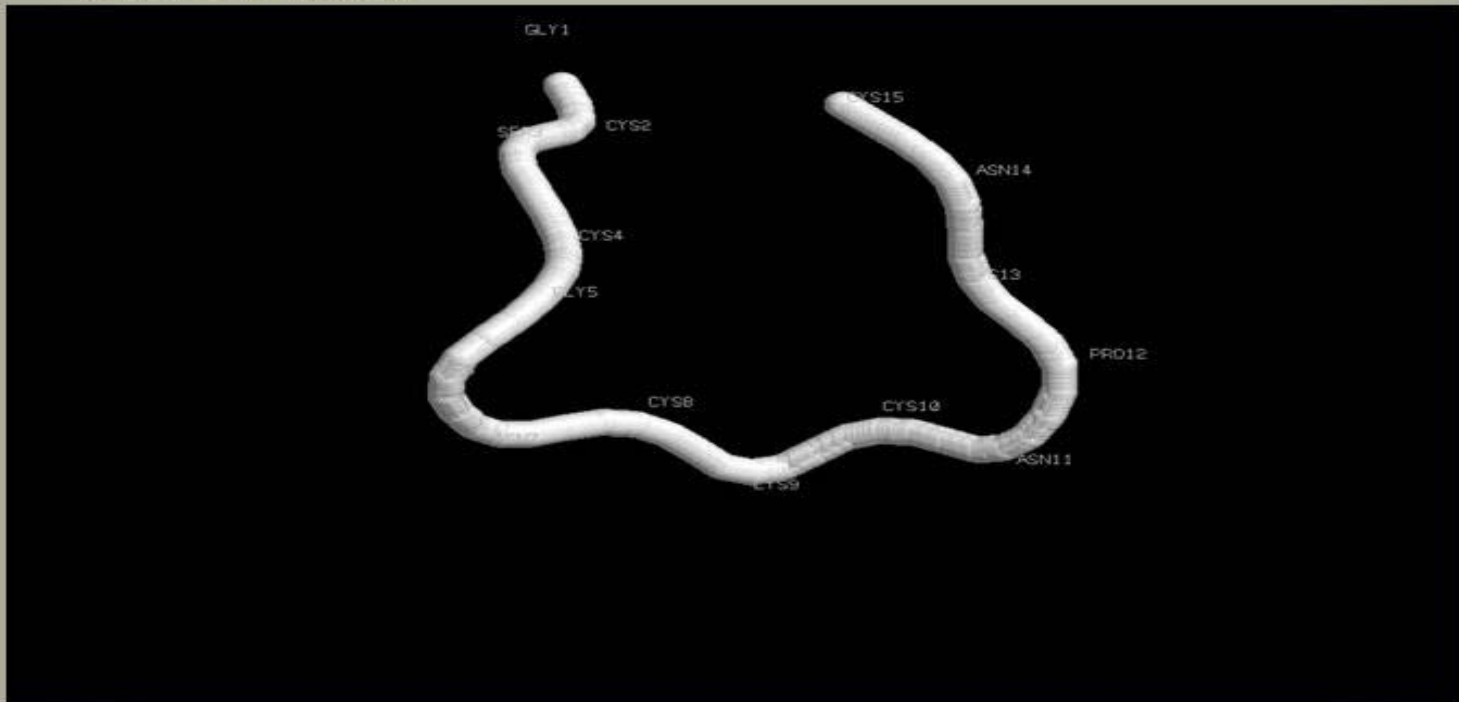


Figure 4.55(a): Surface accessible possible DNA binding residues of wMT  $\alpha$ -domain.

**w-MT beta domain**



**Possible DNA binding with; Asn7, Lys9 and Asn11**

Figure 4.55(b): Surface accessible possible DNA binding residues of wMT  $\beta$ -domain.



## 5 DISCUSSION

Structure-function relationship is commonly referred to as the “holy dream of structural bioinformatics” due to the complex nature of the phenomenon. Although the main issue seems to be the prediction of the protein function from its structure, if the structure itself is not meaningful, all possible predictions will be condemned to remain as yet another computation. Finding the right, “native” structure, on the other hand, is another but very sophisticated issue which requires many parameters to be specified; e.g., pH, temperature, ligand interactions, etc. All these factors make this endeavour highly challenging, as Pam and Dill stated:

“Protein folding is one of the most intriguing intellectual challenges in molecular biology.”

Pam R.H. (1994), *Nature*

“Protein folding problem is a problem for us, not for proteins. They just fold...”

Dill K. (1998), *Prot. Str. Analysis*

Protein folding predictions start from the DNA sequence and end with the predicted atomic coordinates of proteins and include several approaches which fall into two general classes; homology modeling and *ab initio* predictions. Homology modeling approaches are used when a high sequence similarity exists between the template and the target proteins. *ab initio* methods, on the other hand, directly use the sequence information and try to predict coordinates of atoms in the protein.

At a glance, one can easily conclude that these are just algorithms and once they get the input, structure will appear with two possibilities; right or wrong. However, structure and/or function prediction requires also input from molecular biology, biochemistry, physics and evolution (Figure 5.1). Such an approach was very helpful for us while predicting the structure and function of the wMT.

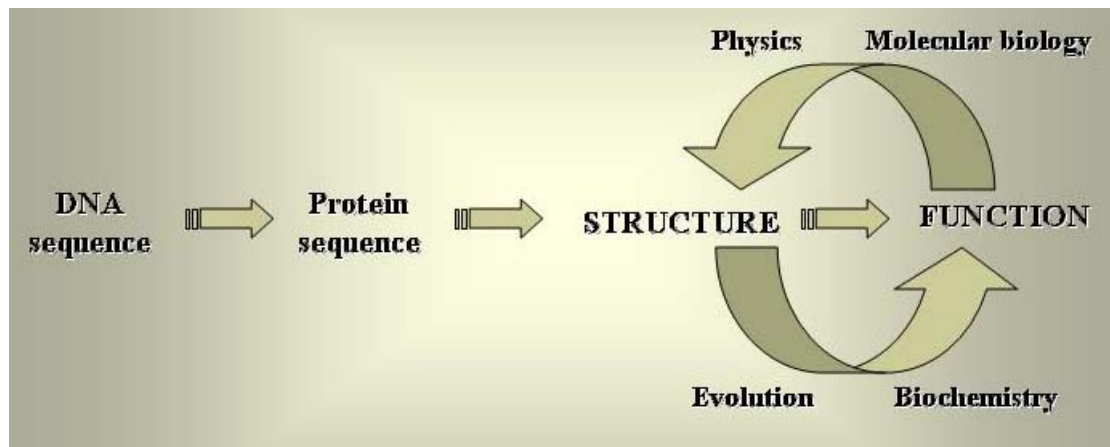


Figure 5.1: Structure-function studies require input from several disciplines.

The work described in this thesis consists mainly of two parts;

- Molecular biology work to produce material for experimental proof of the results of predictions and to obtain additional information for prediction studies.
- Theoretical part to predict structure and/or function of wMT and provide information for experimental studies.

## 5.1 Characterization of *mt* gene in *Triticum durum*

Primers, used during the *mt* gene amplification from *T. durum*, were designed according to the *T. aestivum* MT mRNA sequence which is available through NCBI nucleotide database. MT superfamily proteins are found in nearly all plant species and they are highly conserved among themselves. Indeed, alpha and beta metal binding domains are very similar in terms of amino acid composition. Multiple alignments of MTs from closely related species such as rice, maize and barley, proved that this approach was correct.

Sequencing of *mt* cDNAs and genomic DNAs from both species, *T. durum* and *T. aestivum* showed that these two proteins have the same amino acid composition although their mRNAs had dissimilarities at several points. Third codon degeneracy, sequencing errors, or both of them could be possible reasons for such “point mutations”. Intron sequences, on the other hand, were different mostly in their repeated “TTTTA” regions, where *d-mt* had an extra repetition in addition to several nucleotide base dissimilarities. Both species have 2 exons that are connected with an intron in their genomic DNA that code for type 2 MT. This seems to be a common feature among MT superfamily proteins of plants and mammalians.

## 5.2 Induction of MT expression in wheat

MTs are cytoplasmic metal scavengers and their mRNA levels increase under metal stress. For this reason several Cd doses were tried on Bezostaja and Balcali, and optimum concentration was determined to be 5 to 10  $\mu\text{M}$ .. Higher metal concentrations can stimulate some other plant defense mechanisms like antioxidant defense systems and necrosis.

At this point, our aim was neither to screen plant responses to applied Cd doses nor to analyze other defense mechanisms. Therefore, we could not speculate on MT

involvement in heavy metal detoxification mechanisms. mRNA isolation was performed at the two selected Cd concentrations.

### **5.3 Cloning and expression of *mt-d* in *E. coli***

Prokaryotic expression systems are generally chosen when a high protein yield is aimed. We wanted to obtain maximum possible protein yield with most possible purity and therefore a GST fusion system was preferred. A direct expression of MT was not preferred due to difficulties likely to be faced during the isolation and purification steps. Some of these difficulties are listed below.

Firstly, dMT is a very small protein having a molecular weight of 7350 Da. Therefore, its detection with conventional gel analysis is almost impossible. Secondly, due to the high cysteine content, MTs are readily oxidized even in the presence of trace amount of oxidizing agents such as O<sub>2</sub>. Finally, the correct folding of the MT is another challenge. Proteins that bind to ligands for proper function do not have stable structures in their free forms (Uversky, 2002). MTs can only form a stable structure when they form thiol bonds between their cysteines and metal ions.

Expression of dMT as a fusion protein in the form of GSTdMT, first of all, overcame isolation and detection problems. During the expression optimization steps, cellular proteins were screened using SDS-PAGE analysis and due to the high expression level GSTdMT was easily detected on the gel when compared with non-transformed cell lysate.

pGEX-4T-2 vector has a strong promoter and a total of 38mg GSTdMT protein could be obtained from 1.5 L of bacterial culture under optimized conditions.

GST can possibly protect dMT from proteolytic attacks and help for a proper folding (Huang *et al.*, 2002). In addition, cadmium was present in LB media and solutions that were used in isolation procedures to support proper folding. The selection of appropriate Cd concentration was done in a way similar to that of wheat plants. As we expected, although it was not our primary aim, bacterial cells expressing GSTdMT

showed higher growth rate in the presence of the metal. This difference became obvious at higher Cd concentrations. This result, on the other hand, does not reflect the primary function of MTs in plant, but may be suggestive.

After optimizing expression, purification steps were performed using sepharose 4 fast flow GST affinity resin. In addition to batch and prepacked column purification procedures which were supplied by the manufacture, a home made column was also used. Batch and home made column purification, although they were successful in terms of purification, did not provide us with the desired homogeneity for solution X-ray scattering experiments. The prepacked HiTrap® column, on the other hand, gave the expected purity, but still some impurities. The high molecular weight impurities probably interact with GST and/or dMT, as will be discussed later.

Quantification of GST, GSTdMT, and dMT was done by spectroscopic analysis at a wavelength of 280 nm. At this wavelength mostly aromatic groups absorb the incoming light and GST has an experimentally determined extinction coefficient value which gives 1 absorption unit for a concentration of 0.5 mg protein per ml. Knowing the extinction coefficient, purified GST and GSTdMT concentrations were determined with direct measurements. dMT, on the other hand, could not be quantified directly with spectroscopic measurements. Firstly because it contains very few aromatic carbon chains and does not absorb at 280 nm. Secondly, there is no determined extinction coefficient value for dMT. The only signal came from Cd-thiol bonds within the range of 235-250 nm which was clearly seen by a comparison between the GST and GSTdMT spectra.

Non-denaturing polyacrylamide gel (native-PAGE) analysis of GST and GSTdMT showed a strong interaction between monomers in solution; although, single bands were seen on denaturing SDS gels. These results were reproducible and correlated with a unit size, indicating possible protein-protein interactions but not unspecific aggregation and/or precipitation. These results are further supported by preliminary solution X-ray scattering measurements and size exclusion chromatography results of GSTdMT and dMT.

Size exclusion chromatography of dMT gave a very strong peak at an elution volume corresponding to ~23 kDa which may possibly be a trimeric form of dMT, considering that the dMT molecular weight is 7.35 kDa. According to the GSTdMT chromatogram, two different molecular weight protein species existed in the solution; a possible tetramer with 114 kDa and a possible trimer with 72.8 kDa, where the molecular weight of the GSTdMT monomer is expected to be ~ 36.35 kDa.

Solution X-ray scattering measurement results indicated that the protein solutions were heterogeneous, i.e. containing more than one type of structural entity, pointing to high molecular weight assemblies.

Purification was performed using several different buffers, but using neither solution scattering nor with size exclusion chromatography the monomer form of GSTdMT and dMT could be obtained and screened. This was an unsuccessful result in terms of structural biology, because we could not get the structure for the “wheat type 2 metallothionein protein”. On the other hand, these findings had been expected according to modeled wMT, and these were strongly suggested by the predicted functional (structural) motifs on the protein.

#### **5.4 Structure and function prediction of wMT**

Two methods, homology modeling and ROSETTA which uses simulated annealing, were used for the structure prediction of wMT. Cystein residue distribution pattern of wMT was found to be very similar to rat liver MT (4MT2). The wMT was expected to have two metal binding domains similar to those of 4MT2 and finding structures with high sequence similarities was the input for the homology modeling studies. On the other hand, unlike 4MT2 or any other mammalian MT a relatively very long hinge region was present in wMT.

Sequence similarity searches within known MT structures and even other PDB entries did not give a result that could be used for homology modeling. For this reason,

the hinge region sequence was submitted to ROSETTA server. This approach includes a sequence similarity search within known protein structures by dividing the protein into small fragments. Then a simulated annealing algorithm connects obtained structures and generates the whole protein by keeping the global energy at minimum. This approach generated several structures due to its intrinsic properties. The selection of the right structure, on the other hand, is another matter which should include the consideration of protein function and most importantly experimental results.

The  $\alpha$ - and  $\beta$ -metal-binding domains' structure prediction was relatively less challenging due to the presence of similar sequences within known MT structures. Two selected structures were used as template and other residues were mutated according to the target protein sequence. During homology modeling studies the critical point is the cysteine-metal thiol bond length conservation between the template and the target structure. This constraint was applied due to the lack of available software which would take into account the presence of metals during the homology modeling. An algorithm that could calculate the global energy minimization by considering the presence of metal ions or any cation, would probably have given better structures for wMT alpha and beta domains.

Results coming from homology modeling and ROSETTA server were put together and the predicted structure for the wMT was reconstructed. The presence of more than one predicted structure gave flexibility for the functional predictions.

Structural characteristics are based on the primary structure of any protein and with a generalization; “similar sequences form similar structures in proteins” could be a correct expression. Under the light of this theory, presence of conserved motifs like DNA binding or protein-protein interactions in the hinge region, could give clues about wMT function. At this point, only the hinge region was considered as a possible source of a different function for wMT, because alpha and beta domains would bind metal in any case.

Conserved pattern searches in Pfam and PROSITE were not successful, and the only hit was the presence of a N-myristoylation site, but the pattern was too unspecific to speculate on.

Interestingly, results of structure similarity searches for the hinge region indicated proteins that have DNA binding and protein interaction roles although their sequence based similarities with wMT hinge region were below standard thresholds levels. These proteins included transcription factors, specific SH3-like barrel fold (DNA binding structure), trypsin and proteinase C inhibitors related domains. These structures were strong evidences for the probably correct relative positioning of helices and beta sheets in the predicted structure.

The main information came from the *Syanococcus metallothionein suppressor protein (SmtB)* which is a cyanobacterial DNA binding protein (REFERENCE). SmtB normally binds to the promoter/operator region of the cyanobacterial metallothionein protein (SmtA) preventing its transcription. In the presence of metals, however, a structural change results with the dissociation of SmtB from its DNA binding site making further transcription possible. SmtB binds to DNA in its dimer form, and DNA binding domain is formed from two alpha helices which share a high sequence similarity with the wMT alpha helices in the hinge region. In the SmtB these two alpha helices are next to each other, in the wMT, on the other hand, these two helices are apart from each other in the primary structure. According to the predicted structure, interestingly, these two helices come next to each other the metal bound state.

In addition cyanobacterial SmtA gene and its promoter region contain several inverted repeats which are target sequences for the proper binding of SmtB. wMT gene sequence, similarly, contains such repeated regions. However, we could not speculate on this as we do not have the operator/promoter region of the wheat *mt* gene.

Such DNA or metal binding proteins generally fall generally into “natively unfolded” protein group (Uversky, 2002). According to a calculation that is based on mean hydrophobicity and mean net charge proportions which was generated by Uversky, wMT could be classify as a natively unfolded protein. These proteins do not



generally have a stable structure unless they bind to their ligand; DNA, protein, or metal ions. Further calculations specifically indicate disordered regions at the flanking part of the hinge region supporting the predicted structure.

Metal binding domains get stable structures in their metal bound states, and upon binding metal ions several residues became surface available at the lower region of the protein. These residues include Glycine (8), Serine (10), and Serine (13) in the alpha domain; and Asparagine (7), Lysine (9), Asparagine (11) in the beta domain of the wMT. These amino acids are shown to be involved in nucleic acid interactions with very high probability (Luscombe *et al.*, 2001).

When put together, these results indicate a most possible DNA binding function for wMT. When the protein binding part is included to this picture, a probable activator, repressor or transcription factor role appear on the scene.

Further work is needed, however, to:

- Characterize the recombinant dMT fractions in terms of size and metal content.
- Determine structural characteristics and shape of the protein in solution.
- Compared the predicted shape with that experimentally determined.
- Study interactions of dMT with itself, other proteins and DNA *in vitro* and *in vivo* (using mutations, reporter proteins, markers etc).

Determine structural characteristics of the complexes.

## 6 CONCLUSION

Metallothionein superfamily proteins, although they are well characterized in mammalian and yeast systems, generate many questions about their functions in plant systems, indeed , there is no structural work done on plant MTs. Our motivation was to find the solution structure of the wMT by using X-ray scattering and compare the obtained structure with the theoretical one. Function prediction, on the other hand, was speculated and predicted under the light of the theoretical model.

Predicted structure and experimental data were well correlated giving clues about the function of wMTs as a possible DNA and/or protein interacting protein. These experimental and theoretical results could add several possible function to wMT such as transcription factor, and gene suppressor or activator protein. These theoretical speculations should be proved through experimental works, which will probably cover *in vivo* and *in vitro* DNA/protein interaction assays, generations of transgenic wheat cultivars having silenced or overexpressed MTs, and their physiologic and proteomic characterizations, and for sure diffraction or NMR studies to find the whole structure of the protein.

On the other hand, in terms of experimental results, a novel MT gene was found, characterized, and isolated from *Triticum durum*, which is an economically important wheat species especially used in pasta making. Sequencing results will be soon submitted to NCBI database being the first *mt* gene isolated and characterized from durum wheat.

A total characterization of MT superfamily proteins will provide us with information on how plants tolerate to heavy metal toxification, regulate their essential trace element metabolism, and behave under oxidative stress conditions.

## 7 REFERENCES

- An,Y., Li,G. & Ru,B. (1999). Crystallization and preliminary X-ray studies of metallothionein II from rabbit liver. *Acta Crystallographica*, **D55**:1242-1243.
- Andrews,G.K. (2000). Regulation of metallothionein gene expression by oxidative stress and metal ions. *Biochem.Pharmacol.*, **59**: 95-104.
- Beattie,J.H., Black,D.J., Wood,A.M. & Trayhurn,P. (1996). Cold-induced expression of the metallothionein-1 gene in brown adipose tissue of rats. *American Journal of Physiology*, **270**:971-977.
- Bertini,I., Hartmann,H.J., Klein,T., Liu,G., Luchinat,C. & Weser,U. (2000). High resolution solution structure of the protein part of Cu<sub>7</sub> metallothionein. *Eur.J.Biochem.*, **267**: 1008-1018.
- Binz,P.A. & Kagi,H.R. (1997) Metallothionein: Molecular evolution and classification. *In*: Klaassen (ed): Metallothionein IV. Birkhauser Verlag, Basel Boston Berlin, 7-21.
- Blindauer,C.A., Harrison,M.D., Parkinson,J.A., Robinson,A.K., Cavet,J.S., Robinson,N.J. & Sadler,P.J. (2001). A metallothionein containing a zinc finger within a four-metal cluster protects a bacterium from zinc toxicity. *Proc.Natl.Acad.Sci.U.S.A.*, **98**: 9593-9598.

Boulin,C., Kempf,R., Koch,M.H.J. & McLaughlin, S. (1986). Data acquisition systems for linear and area X-ray detectors using delay line readout. *Nucl. Instrum. Methods* **249**:399-407.

Braun,W., Vasak,M., Robbins,A.H., Stout,C.D., Wagner,G., Kagi,J.H. & Wuthrich,K. (1992). Comparison of the NMR solution structure and the X-ray crystal structure of rat metallothionein-2. *Proc. Natl. Acad. Sci.*, **89(21)**:10124-10128.

Briat,J.F. & Lebrun,M. (1999). Plant responses to metal toxicity. *C.R.Acad.Sci.III*, **322**: 43-54.

Clemens,S. (2001). Molecular mechanisms of plant metal tolerance and homeostasis. *Planta*, **212**: 475-486.

Cobbett,C. & Goldsborough,P. (2002). Phytochelatins and Metallothioneins: Roles in Heavy Metal Detoxification and Homeostasis. *Annual Reviews of Plant Biology*, **53**:159-182.

Cook,W.J., Kar,S.R., Taylor,K.B. & Hall,L.M. (1998). Crystal structure of the cyanobacterial metallothionein repressor SmtB: a model for metalloregulatory proteins. *J.Mol.Biol.*, **275**: 337-346.

Coyle,P., Philcox,J.C., Carey,L.C. & Rofe,A.M. (2002). Metallothionein: the multipurpose protein. *Cell Mol.Life Sci.*, **59**: 627-647.

Dallinger,R., Berger,B., Hunziker,P. & Kagi,J.H. (1997). Metallothionein in snail Cd and Cu metabolism. *Nature*, **388**: 237-238.

Ebadi,M., Leuschen,M.P., el Refaey,H., Hamada,F.M. & Rojas,P. (1996). The antioxidant properties of zinc and metallothionein. *Neurochem.Int.*, **29**: 159-166.

Evans,K.M., Gatehouse,J.A., Lindsay,W.P., Shi,J., Tommey,A.M. & Robinson,N.J. (1992). Expression of the pea metallothionein-like gene PsMTA in *Escherichia coli* and *Arabidopsis thaliana* and analysis of trace metal ion accumulation: implications for PsMTA function. *Plant Mol.Biol.*, **20**: 1019-1028.

Furey,W.F., Robbins,A.H., Clancy,L.L., Winge,D.R., Wang,B.C. & Stout,C.D. (1986). Crystal structure of Cd,Zn metallothionein. *Science*, **231**: 704-710.

Gabriel,A. and Dauvergne,F. (1982). The localization method used at EMBL. *Nucl. Instrum. Methods* **201**:223-224.

Giedroc,D.P., Chen,X. & Apuy,J.L. (2001). Metal response element (MRE)-binding transcription factor-1 (MTF-1): structure, function, and regulation. *Antioxid.Redox.Signal.*, **3**: 577-596.

Huang,Z.X., Gao,Y., Yu,W.H., Zhang,S.Y. & Yang,P.Y. (2002). Construction of alpha-alpha domains structure in recombinant monkey metallothionein-1. *J.Inorg.Biochem.*, **92**: 183-192.

Katakai,K., Liu,J., Nakajima,K., Keefer,L.K. & Waalkes,M.P. (2001). Nitric oxide induces metallothionein (MT) gene expression apparently by displacing zinc bound to MT. *Toxicol.Lett.*, **119**: 103-108.

Kawashima,I., Kennedy,T.D., Chino,M. & Lane,B.G. (1992). Wheat Ec metallothionein genes. Like mammalian Zn<sup>2+</sup> metallothionein genes, wheat Zn<sup>2+</sup> metallothionein genes are conspicuously expressed during embryogenesis. *Eur.J.Biochem.*, **209**: 971-976.

Kelley,L.A., MacCallum,R.M. & Sternberg M.J.E. (2000). Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J. Mol. Biol.* **299(2)**:499-520.

Koch,M.H.J. (1991) in Handbook on synchrotron radiation Vol. 4, Ebashi, S., Koch, M. and Rubenstein, E. eds,Elsevier Science Publishers BV, Amsterdam, pp 241-268.

Koch,M.H.J. and Bordas,J. (1983). X-ray diffraction and scattering on disordered systems using synchrotron radiation. *Nucl. Instrum. Methods* **208**:461-469.

Kojima,Y., Binz,P.A. & Kagi,H.R. (1997) Nomenclature of metallothionein: Proposal for a revision. *In: Klaassen (ed): Metallothionein IV*. Birkhauser Verlag, Basel Boston Berlin, 3-6.

Kyte,J. and Doolittle,R.F. (1982). A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* **157(1)**:105-132.

Lazo,J.S., Kondo,Y., Dellapiazza,D., Michalska,A.E., Choo,K.H.A. and Pitt,B.R. (1995). Enhanced sensitivity to oxidative stress in cultured embryonic cells from transgenic mice deficient in metallothionein I and II genes. *Journal of Biological Chemistry*, **270**:5506-5510.

Liu,T., Nakashima,S., Hirose,K., Uemura,Y., Shibasaki,M., Katsuhara,M. & Kasamo,K. (2003). A metallothionein and CPx-ATPase handle heavy-metal tolerance in the filamentous cyanobacterium *Oscillatoria brevis*. *FEBS Lett.* **8**:159-63.

Ma,M., Lau,P-S., Jia,Y-T., Tsang,W-K., Lam,S.K.S., Tam,N.F.Y. & Wong,Y-S. (2003). The isolation and characterization of Type I metallothionein (MT) cDNA from a heavy-metal-tolerant plant, *Festuca rubra* cv. Merlin. *Plant Science*, **164**:51-60.

Margoshes,M. & Vallee,B.L. (1957) A cadmium protein from equine kidney cortex. *Journal of American Chemical Society*, **79**:4813-4814.

Melis,K.A., Carter,D.C., Stout,C.D. & Winge,D.R. (1983). Single crystals of cadmium, zinc metallothionein. *Journal of Biological Chemistry*, **258(10)**:6255-6257.

Morita,E.H., Wakamatsu,M., Uegaki,K., Yumoto,N., Kyogoku,Y. & Hayashi,H. (2002). Zinc ions inhibit the protein-DNA complex formation between cyanobacterial transcription factor SmtB and its recognition DNA sequences. *Plant Cell Physiol*, **43**: 1254-1258.

Munoz,A., Petering,D.H. & Shaw,C.F., III (2000). The requirements for stable metallothionein clusters examined using synthetic lobster domains. *Mar.Environ.Res.*, **50**: 93-97.

Oz,G., Zangger,K. & Armitage,I.M. (2001). Three-dimensional structure and dynamics of a brain specific growth inhibitory factor: metallothionein-3. *Biochemistry*, **40**: 11433-11441.

Peterson,C.W., Narula,S.S. & Armitage,I.M. (1996). 3D solution structure of copper and silver-substituted yeast metallothioneins. *FEBS Letters*, **379**:85.

Quaife,C.J., Findley,S.D., Erickson,J.C., Froleick,G.J., *et al.* (1994). Induction of a new metallothionein isoform (MT-IV) occurs during differentiation of stratified squamous epithelia. *Biochemistry*, **33**:7250-7259.

Rausser,W.E. (1999). Structure and function of metal chelators produced by plants: the case for organic acids, amino acids, phytin, and metallothioneins. *Cell Biochem.Biophys.*, **31**: 19-48.

Robbins,A.H., McRee,D.E., Williamson,M., Collett,S.A., Xuong,N.H., Furey,W.F., Wang,B.C. & Stout,C.D. (1991). Refined crystal structure of Cd, Zn metallothionein at 2.0 Å resolution. *Journal of Molecular Biology*, **221(4)**:1269-1293.

Robinson,N.J., Tommey,A.M., Kuske,C. & Jackson,P.J. (1993). Review article: Plant metallothionein. *Biochemical Journal*, **295**:1-10.



Sato,M. & Bremer,I. (1993). Oxygen free radicals and metallothionein. *Free Radicals in Biology and Medicine*, **14**:325-337.

Sato,M., Apostolova,M.D., Hamaya,M., Yamaki,J., Choo,K.H.A., Michalska,A.E., Kodama,N. & Tohyama,C. (1996). Susceptibility of metallothionein-null mice to paraquat. *Environmental Toxicology and Pharmacology*, **1**:221-225.

Sayers,Z., Brouillon,P., Svergun,D.I., Zielenkiewicz,P. & Koch,M.H. (1999). Biochemical and structural characterization of recombinant copper- metallothionein from *Saccharomyces cerevisiae*. *Eur.J.Biochem.* , **262**: 858-865.

Silver,S. & Phung,L.T. (1996). Bacterial heavy metal resistance: new surprises. *Annu.Rev.Microbiol.*, **50**: 753-789.

Svergun,D.I. (1991). Mathematical methods in small-angle scattering data analysis. *J.Appl.Cryst.* **24**:485-492.

Tanguy,A. & Moraga,D. (2001). Cloning and characterization of a gene coding for a novel metallothionein in the Pacific oyster *Crassostrea gigas* (CgMT2): a case of adaptive response to metal-induced stress? *Gene*, **273**: 123-130.

Tanguy,A., Mura,C. & Moraga,D. (2001). Cloning of a metallothionein gene and characterization of two other cDNA sequences in the Pacific oyster *Crassostrea gigas* (CgMT1). *Aquat.Toxicol.*, **55**: 35-47.

Thompson,J.D., Higgins,D.G. & Gibson,T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673-4680.

Turner,J.S., Glands,P.D., Samson,A.C. & Robinson,N.J. (1996). Zn<sup>2+</sup>-sensing by the cyanobacterial metallothionein repressor SmtB: different motifs mediate metal-induced protein-DNA dissociation. *Nucleic Acids Res.*, **24**: 3714-3721.

Valentine,J.S. & Gralla,E.B. (1997). Delivering copper inside yeast and human cells. *Science*, **278**: 817-818.

Vasak,M. & Hasler,D.W. (2000). Metallothioneins: new functional and structural insights. *Curr.Opin.Chem.Biol.*, **4**: 177-183.

Vasak,M. and Kägi,J.H.R. (1994) in: Encyclopedia of Inorganic Chemistry (King, R.B., Ed.), pp. 2229–2241, J. Wiley and Sons, New York.

White,C.N. & Rivin,C.J. (1995). Characterization and expression of a cDNA encoding a seed-specific metallothionein in maize. *Plant Physiology*, **108**:831-832.

Yeh,S-C., Hsieh,H-M. & Huang,P.,C. (1995). Transcripts of metallothionein genes in *Arabidopsis thaliana*. *The Journal of Sequencing and Mapping*, **5**:141-144.

Yu,L.H., Umeda,M., Liu,J.Y., Zhao,N.M. & Uchimiya,H. (1998). A novel MT gene of rice plants is strongly expressed in the node portion of the stem. *Gene*, **206**: 29-35.

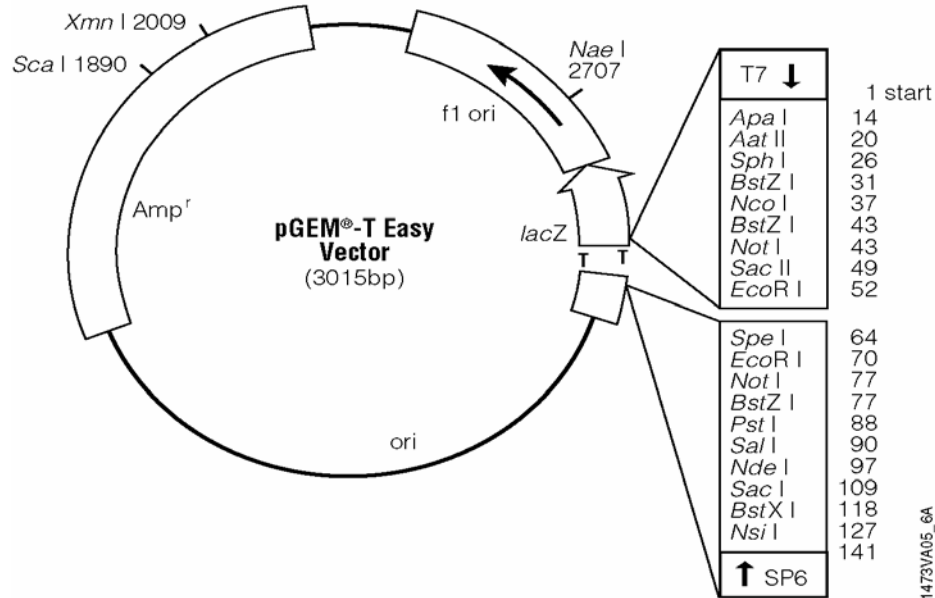
Yu,W.H., Cai,B., Gao,Y., Xie,Y. & Huang,Z.X. (2002). Expression, characterization, and reaction of recombinant monkey metallothionein-1 and its C33M mutant. *J.Protein Chem.*, **21**: 177-185.

Yu,W.H., Lukiw,W.J., Bergeron,C., Niznik,H.B. & Fraser,P.E. (2001). Metallothionein III is reduced in Alzheimer's disease. *Brain Res.*, **894**: 37-45.

Zangger,K., Oz,G., Haslinger,E., Kunert,O. & Armitage,I.M. (2001). Nitric oxide selectively releases metals from the amino-terminal domain of metallothioneins: potential role at inflammatory sites. *FASEB J.*, **15**: 1303-1305.

## APPENDIX A

### A1. pGEM-Teasy vector circle map showing multiple cloning sites with restriction sites on it (top) and vector regions (bottom).



**pGEM-T Easy Vector Sequence reference points:**

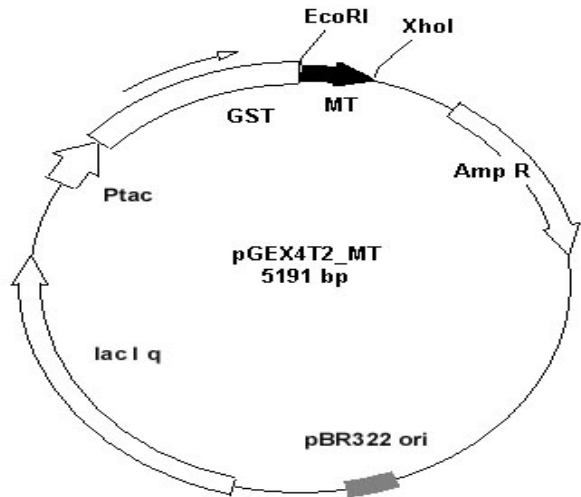
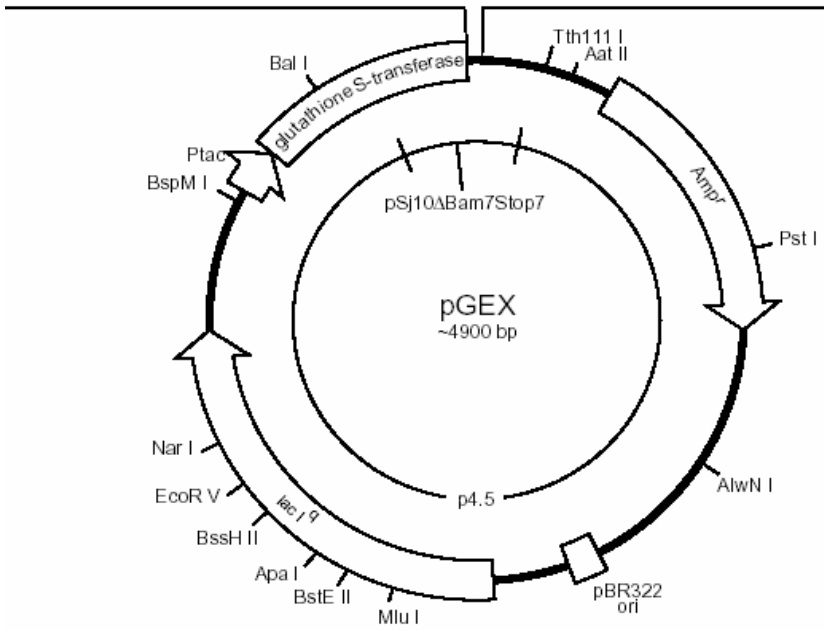
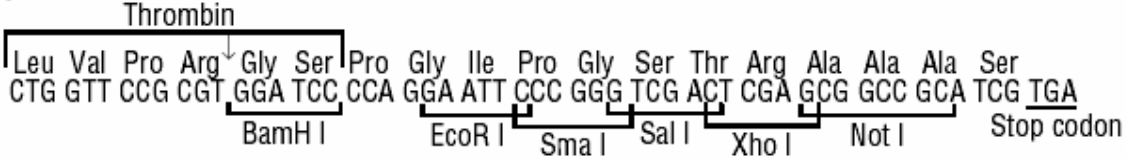
T7 RNA Polymerase transcription initiation site	1
SP6 RNA Polymerase transcription initiation site	141
T7 RNA Polymerase promoter (-17 to +3)	2999-3
SP6 RNA Polymerase promoter (-17 to +3)	139-158
multiple cloning region	10-128
lacZ start codon	180
lac operon sequences	2836-2996, 166-395
lac operator	200-216
β-lactamase coding region	1337-2197
phage f1 region	2380-2835
binding site of pUC/M13 Forward Sequencing Primer	2956-2972
binding site of pUC/M13 Reverse Sequencing Primer	176-192

Enzyme	# of Sites	Location	Enzyme	# of Sites	Location
<i>Aat</i> II	1	20	<i>Bst</i> Z I	2	31, 62
<i>Acc</i> I	1	76	<i>Cfr</i> 10 I	2	1475, 2690
<i>Acy</i> I	2	17, 1932	<i>Dde</i> I	4	777, 1186, 1352, 1892
<i>Afr</i> III	2	99, 502	<i>Dra</i> I	3	1261, 1280, 1972
<i>Alw</i> 26 I	2	1456, 2232	<i>Dra</i> III	1	2589
<i>Alw</i> 44 I	2	816, 2062	<i>Drd</i> I	2	610, 2544
<i>Alw</i> N I	1	918	<i>Dsa</i> I	2	37, 43
<i>Apa</i> I	1	14	<i>Eag</i> I	2	31, 62
<i>Asp</i> HI	4	94, 820, 1981, 2066	<i>Ear</i> I	3	386, 2190, 2878
<i>Ava</i> II	2	1533, 1755	<i>Eci</i> /HK I	1	1395
<i>Ban</i> I	3	246, 1343, 2626	<i>Eco</i> 52 I	2	31, 62
<i>Ban</i> II	3	14, 94, 2664	<i>Eco</i> CR I	1	92
<i>Bbu</i> I	1	26	<i>Eco</i> RV	1	51 (see above)
<i>Bgl</i> I	3	39, 1515, 2833	<i>Fok</i> I	5	119, 1361, 1542, 1829, 2919
<i>Bsa</i> I	1	1456	<i>Fsp</i> I	2	1617, 2840
<i>Bsa</i> A I	1	2589	<i>Hae</i> II	4	380, 750, 2740, 2748
<i>Bsa</i> HI	2	17, 1932	<i>Hga</i> I	4	613, 1191, 1921, 2806
<i>Bsa</i> J I	5	37, 43, 241, 662, 2936	<i>Hinc</i> II	1	77
<i>Bsp</i> 120 I	1	10	<i>Hind</i> II	1	77
<i>Bsp</i> HI	2	1222, 2230	<i>Hsp</i> 92 I	2	17, 1932
<i>Bsp</i> MI	1	62	<i>Mae</i> I	5	56, 997, 1250, 1585, 2740
<i>Bss</i> S I	2	675, 2059	<i>Mlu</i> I	1	99
<i>Bst</i> O I	5	242, 530, 651, 664, 2937			
<i>Bst</i> X I	1	103			

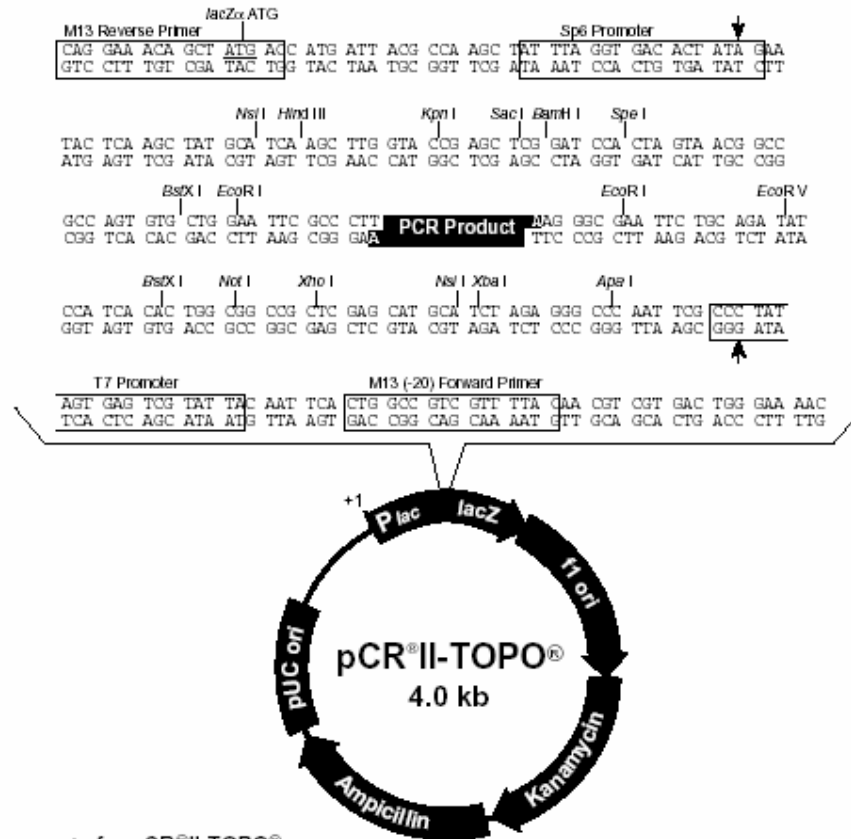
Total number of restriction enzyme hits on pGEM-Teasy vector sequence.

**A2. Vector map of pGEX-4T-2 and multiple cloning site showing available restriction enzyme hits.**

**pGEX-4T-2 (27-4581-01)**



### A3. pCR-II-TOPO vector map and sequence reference points



Comments for pCR<sup>®</sup>II-TOPO<sup>®</sup>  
3973 nucleotides

- LacZ*α gene: bases 1-589
- M13 Reverse priming site: bases 205-221
- Sp6 promoter: bases 239-256
- Multiple Cloning Site: bases 269-383
- T7 promoter: bases 406-425
- M13 (-20) Forward priming site: bases 433-448
- f1 origin: bases 590-1027
- Kanamycin resistance ORF: bases 1361-2155
- Ampicillin resistance ORF: bases 2173-3033
- pUC origin: bases 3178-3851

## APPENDIX B

Autoclave:	Hirayama, Hiclave HV-110, JAPAN
	Certoclav, Table Top Autoclave CV-EL-12L, AUSTRIA
Balance:	Sartorius, BP211D, GERMANY
	Sartorius, BP221S, GERMANY
	Sartorius, BP610, GERMANY
	Schimadzu, Libror EB-3200 HU, JAPAN
Centrifuge:	Eppendorf, 5415C, GERMANY
	Eppendorf, 5415D, GERMANY
	Eppendorf, 5415R, GERMANY
	Kendro Lab. Prod., Heraeus Multifuge 3L, GERMANY
	Hitachi, Sorvall RC5C Plus, USA
	Hitachi, Sorvall Discovery 100 SE, USA
Deepfreeze:	-70° C, Kendro Lab. Prod., Heraeus Hfu486 Basic, GERMANY
	-20° C, Bosch, TÜRKİYE

Distilled Water: Millipore, Elix-S, FRANCE  
Millipore, MilliQ Academic, FRANCE

Electrophoresis: Biogen Inc., USA  
Biorad Inc., USA

Gel Documentation: UVITEC, UVIdoc Gel Documentation System, UK  
Biorad, UV-Transilluminator 2000, USA

Ice Machine: Scotsman Inc., AF20, USA

Incubator: Memmert, Modell 300, GERMANY  
Memmert, Modell 600, GERMANY

Laminar Flow: Kendro Lab. Prod., Heraeus, HeraSafe HS12, GERMANY

Magnetic Stirrer: VELP Scientifica, ARE Heating Magnetic Stirrer, ITALY  
VELP Scientifica, Microstirrer, ITALY

Microliter Pipette: Gilson, Pipetman, FRANCE  
Mettler Toledo, Volumate, USA

Microwave Oven: Bosch, TÜRKİYE

pH meter: WTW, pH540 GLP MultiCal<sup>®</sup>, GERMANY



Power Supply: Biorad, PowerPac 300, USA  
Wealtec, Elite 300, USA

Refrigerator: +4° C, Bosch, TÜRKİYE

Shaker: Forma Scientific, Orbital Shaker 4520, USA  
GFL, Shaker 3011, USA  
New Brunswick Sci., Innova™ 4330, USA

Spectrophotometer: Shimadzu, UV-1208, JAPAN  
Shimadzu, UV-3150, JAPAN  
Secoman, Anthelie Advanced, ITALY

Speed Vacuum: Savant, Speed Vac® Plus Sc100A, USA  
Savant, Refrigerated Vapor Trap RVT 400, USA

Thermocycler: Eppendorf, Mastercycler Gradient, GERMANY

Vacuum: Heto, MasterJet Sue 300Q, DENMARK

Water bath: Huber, Polystat cc1, GERMANY