CHAPTER 3

Identification of proteins showing differential expression under thiol stress in *Streptomyces coelicolor*

3.1. Introduction

Proteome analysis is analogous to genome analysis which defines the complete set of proteins expressed by an organism under defined set of conditions. Proteome analysis is much more complex and dynamic as compared to genome analysis since protein expression varies from one condition to another in order to adjust to the changing environment. In most of the cases, proteomics rely on 2D-PAGE (2 dimentional - polyacrylamide gel electrophoresis), which employs separation of all the proteins in two dimensions based on two properties, isoelectric point (pI) and molecular weight of individual proteins. Since the development of 2D-PAGE (Klose, 1975; O'Farrell, 1975), it has been used extensively to study various bacterial proteomes under different growth conditions (Linn and Losick, 1976; Reeh et al., 1977; Agabian and Unger, 1978) and under the influence of various stress conditions (Young and Neidhardt, 1978; Krueger and Walker, 1984; Gomes et al., 1986). Earlier the efforts were made to understand the global changes in the protein expression profile during growth using 2D PAGE. However, they did not have the advantage of complete genome sequence to identify the proteins using mass spectrometry (Puglia et al., 1995; Vohradsky et al., 1997; 2000). Major breakthrough came after complete genome sequencing of the first organism *Haemophilus influenzae* strain RD KW20 (Fleischmann et al., 1995). Based on the genome sequence and with the advancements in the Mass Spectrometry (MS) techniques it was possible to identify all the proteins and to study the global protein expression profile. These advancements led to a completely new era of protein expression profiling. The increase in high throughput and partial automation of 2D-PAGE analysis recently made it a very attractive tool to study cellular functions on a molecular level. The complete genomic sequences of more than 1000 bacteria are now publicly available (http://www.ncbi.nlm. nih.gov/genomes/lproks.cgi) that allow one to select among a variety of microorganisms for proteomic investigations.

Complete genome sequence of *Streptomyces* allows the study of global gene expression profile at the levels of mRNA abundance, typically using DNA microarrays (the transcriptome) (Lucchini et al., 2001) and protein profiling, by 2D-PAGE coupled with matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry (Blackstock and Weir, 1999; Mann et al., 2001).

3.2. Materials and methods

3.2.1. Protein expression profiling by 2D-PAGE

Protein expression profile was done on single dimension using SDS-PAGE as explained in Chapter 2 and for 2D gel electrophoresis, protein from supernatant of lysate was precipitated by adding three volumes of cold acetone and left overnight at -20 °C. The pellet collected by centrifugation was washed twice with cold acetone and allowed to dry at room temperature and analyzed by SDS-PAGE (Laemmli, 1970). Gels were stained with Coomassie brilliant blue. Same samples were also analyzed by 2D-PAGE using Bio-Rad Protein IEF (isoelectric focusing) Cell. Briefly ~200 µg of cell extract protein was dissolved in 185 µl of IEF buffer supplied by Bio-Rad. An 11 cm IPG strip pH 3-10 and pH 4-7 was rehydrated with the above sample in IEF buffer for 14 h at 20 °C. Following rehydration, IEF was performed on a Biorad 2-D gel electrophoresis system with the stepped program until the focusing reaches a voltage of 42 kVh and then the voltage was allowed to decrease from 8000 to 500 V in 1 h. The proteins in IPG strip were treated with DTT and alkylated with iodoacetamide. Second dimension electrophoresis was performed using Laemmli's method.

3.2.2. In gel trypsin digestion of protein bands/spots

Protein bands of interest were excised from the gel and trypsin digestion was carried out using standard protocol (Chait et al., 1993; Billeci and Stults, 1993; Dogruel et al., 1995; Ha, et al., 2002). Briefly, gel spots/bands from 2D-PAGE/SDS-PAGE gels were excised, minced and washed thrice with double

distilled water (ddH_2O). Silver stain was reduced by 250 µl of 50 mM sodium thiosulfate and 15 mM potassium ferricyanide for five minutes. The gel pieces were washed thrice with ddH_2O with mild agitation in order to remove silver stain. Gel pieces were equilibrated with 500 µL of 100 mM ammonium bicarbonate for 20 minutes at room temperature with gentle agitation followed by two washes of 50% and 100% acetonitrile respectively and the gels were allowed to air dry. The in gel proteins were reduced with 10 mM DTT in 100 mM ammonium bicarbonate for 30 minutes followed by alkylation with 50 mM iodoacetamide for 30 minutes in dark and washed twice with 50% and 100% acetonitrile respectively and acetonitrile was discarded and the gel was dried in speedvac. Gel pieces were rehydrated for 60 minutes at 4 °C in 30 µl of 0.02 µg/ml trypsin in 50 mM ammonium bicarbonate. Trypsin was removed and 30-50 µl of ammonium bicarbonate was added and incubated at 37 °C for 16-18 hours. Supernatant was transferred to a fresh tube and remaining peptides were extracted twice by adding 25-50 µl of extraction solution (60% acetonitrile, 1% trifluoroacetate) and the supernatant containing trypsin digested peptides was collected and used for MALDI-TOF detection.

3.2.3. MALDI-TOF analysis of the peptides

Tryptic digests of specific gel bands were analysed by mass spectrometry. MALDI-TOF mass spectra were recorded in reflectron mode on Tof-Spec 2E (Micromass, UK), fitted with a 337-nm laser. 1.0 µl of peptide digest in 70% acetonitrile, 0.1% trifluoroacetic acid (TFA) was mixed with equal volume of a solution of α-cyano (10 mg/ml solution in 70% acetonitrile/ H_2O containing 0.1% TFA) was used. Peptide mass fingerprint (PMF) of the digested protein was obtained in the reflectron mode at the operating voltage of 20 kV. 30-40 spectra were summed, smoothened and used for database search. Database search of the PMF was carried out on the website www.matrixscience.com using the Mascot program in the MSDB database.

3.2.4. Nano LC-ESI-MS/MS

Proteins in the gel slices were digested with trypsin and peptides were extracted and dried. Dried mixture of tryptic peptides was dissolved in 20 µl of 5% acetonitrile, 0.01% TFA. After injection (8µl) the peptides were concentrated on Trap C-18 enrichment column (0.3 x 5 mm, Agilent) and washed with 100 µl 5% acetonitrile, 0.01% TFA. The enrichment column was then switched into the nanoflow path (200 nl/min) and further separated on C-18 reversed phase ZORBAX 300SB-C-18 nanocolumn (0.075 x 150 mm; Agilent) coupled with the nanoelectrospray ionization (nESI) source of XCT Plus ion trap (Agilent 1100 Series LC/MSD Trap XCT Plus System). The nanoflow gradient started with 5% acetonitrile, 0.01% TFA and increased up to 55% acetonitrile, 0.01% TFA with a slope of 1%/min. Automated MS/MS spectra were acquired during the run in data dependent acquisition mode with selection of three of the most abundant precursor ions (0.5 min active exclusion; 2+ ions preferred). Spectra were extracted using Spectrum Mill software extractor (Agilent) and searched against NCBI database.

3.3 Results

3.3.1. MALDI-TOF analysis of the induced band (as observed by SDS-PAGE)

In order to identify the induced protein in SDS-PAGE (Figure 3.1), the band was excised out from control and DTT treated samples and was analyzed by MALDI-TOF and LC-ESI-MS/MS.

For characterizing the 55kDa protein, the protein band was in gel digested with trypsin and the peptides were extracted and subjected to MALDI analysis. Figure 3.2 shows the peptide mass spectra of the same.

For the identification of the protein, Peptide mass of 22 peptides as identified by MALDI-TOF were submitted to the MSDB database search. Out of these, 13 peptides mass values were found to match within ± 1.0 Da error with high

probability score of 63 and sequence coverage of 29% with CatalaseA (Q9RJK9) (Table 3.1, Figure 3.3) (NCBI: T42038; SwissProt: Q9RJK9) (EC 1.11.1.6).

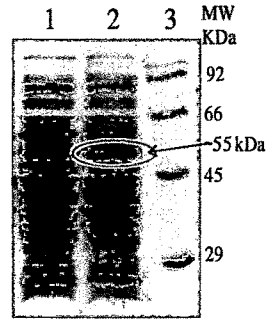

Figure 3.1: SDS-PAGE representing the induced band of 55 kDa in response to DTT. This band was excised out and analyzed using MALDI-TOF and nano-LC-ESI-MS/MS analysis. Lane 1: Control, Lane 2: 10 mM DTT, Lane 3: Molecular weight marker.

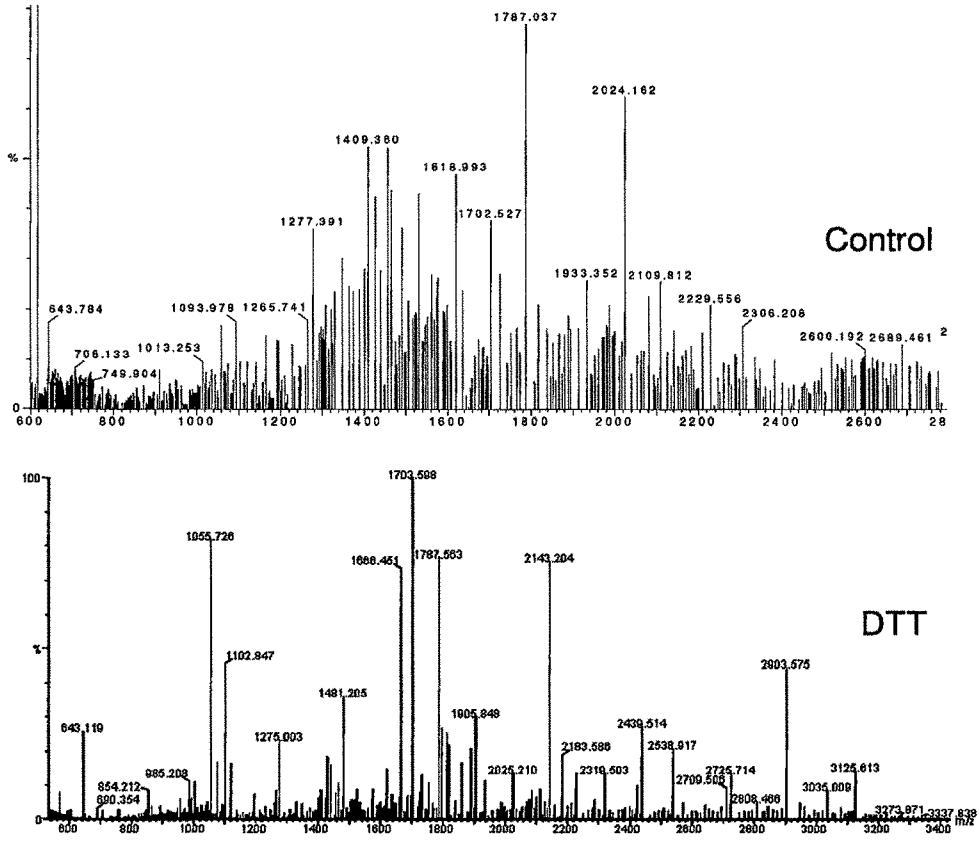

Figure 3.2: MALDI-TOF spectrum of control and DTT induced protein band of 55 kDa from log phase grown *S. coelicolor* cells.

Table 3.1: Results of peptide search for control and DTT treated samples in MASCOT database with purified Enolase from *P. falciparum* as a standard protein.

Sample	Standard protein Enolase (*P. falciparum*)	Control Without DTT (*S. coelicolor*)	Test DTT treated (*S. coelicolor*)
Number of peptides submitted	14	14	22
Number of peptides matched	9	6	13
Data base searched	Swiss Prot	MSDB	MSDB
Protein identified	Enolase (*P. falciparum*)	Lipoprotein (*S. coelicolor*)	CatalaseA (*S. coelicolor*)
Molecular Weight (Daltons)	49015	28926	55139
pI	6.21	6.09	5.84
Sequence courage	28%	27%	29%
Score	59	38	63
Expect value (E)	0.00025	1.2	0.0041

Control panel identified Lipoprotein with poor score where as DTT treated sample identified CatalaseA with significant score and "Expect" value. Under normal conditions catalaseA levels are much lower and do not fall in the range of detection under the experimental conditions chosen. Hence, search with control did not pick up any relevant peptides of catalaseA and the protein with nearest molecular weight match picked up was lipoprotein. The e-value is very high, indicating that the result is insignificant and the protein picked up is wrong as expected, emphasizing absence of peptides belonging to catalaseA.

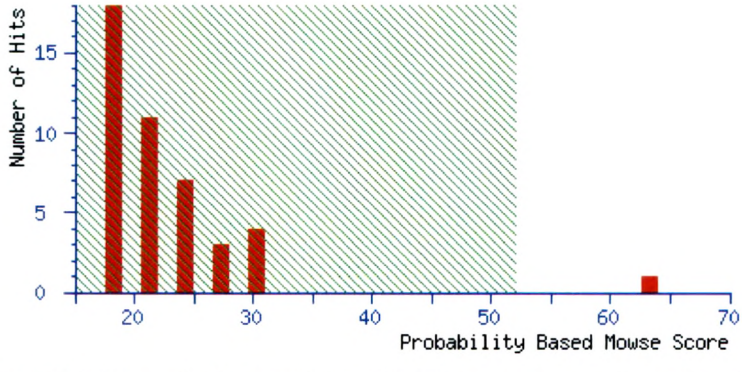

Figure 3.3: MASCOT search for peptides of DTT treated sample giving significant hit for CatalaseA (Q9RJK9) with the score of 63.

3.3.2. Nano-LC-ESI-MS/MS analysis of the induced band (as observed by SDS-PAGE)

For unequivocal identification of the induced protein(s), tryptic digests of the 55kDa band prepared from control and DTT treated samples, were subjected to nanoLC-ESI-MS/MS. Figure 3.4 presents the nano LC-ESI-MS/MS data set for a peptide sequence of the 55kDa band.

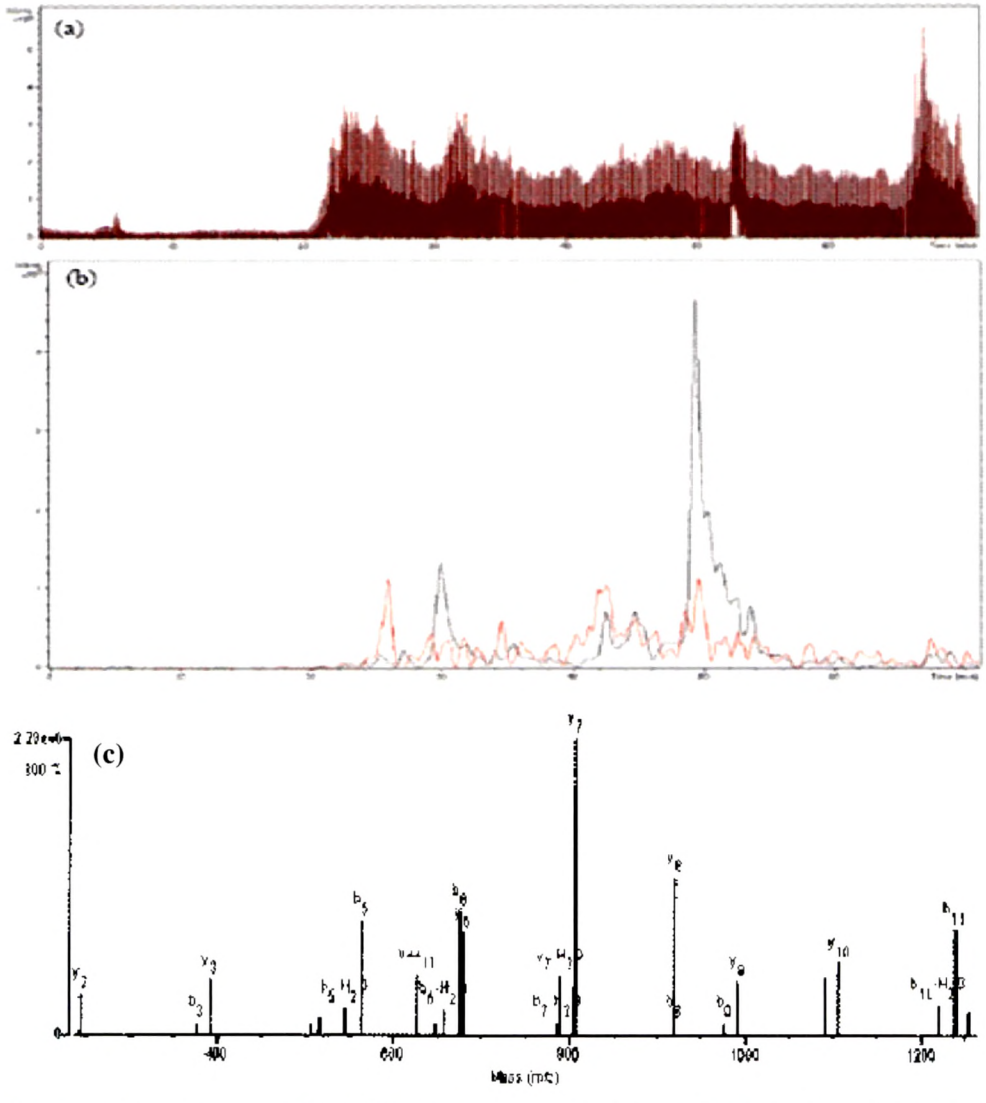

Figure 3.4: Nano-LC-ESI-MS/MS spectra of the 55 kDa induced band from DTT treated sample. (a) peptide spectra, (b) separation of m/z = 741.5 peak present in DTT treated sample, and absent in untreated sample, as identified by ion chromatogram and (c) partial sequencing of the peptides which gave the significant hit for sequence of *S. coelicolor* catalaseA (NCBI: T42038) (EC 1.11.1.6) (ENTREZ CAB58320).

After the search in the NCBI database using Spectrum Mill search engine, catalaseA with 42% sequence coverage was identified as unique and major component of treated sample. In contrast CatalaseA was present in substantially lower concentration in the control sample and led to the conclusion that CatalaseA is the major protein induced in *S. coelicolor* by DTT stress. Figure 3.5 presents the complete sequence of *S. coelicolor* catalaseA with the highlighted regions of protein sequence covered in MALDI-TOF spectra and/or sequenced by nano LC-ESI-MS/MS.

```
1   MPENNQKPLT TVAGAPVPDN QNSLTSGPRG PMLLQDVWFL EKLAHFDREV IPERRMHAKG
61  SGAFGTFTVT HDITRYTSAK IFSEIGKKTP LFTRFSTVAG ERGAADAERD IRGFAVKFYT
121 DEGNWDLVGN NTPVFFFRDP LKFPDLNHAV KRDPRTNLRN AENNWDFWTN LPEALHQVTI
181 VMSDRGIPAS YRHMHGFGSH TYSLINAEGE RFWVKFHHRT QQGIKNLTDA EAEALVGKDR
241 ESHQRDLFDA IEDGDFPKWK LFIQVMPEAD AENYRFHPFD LTKVWSKKDY PLIEVGEWEL
301 NRNPDNYFAD VEQAAFSPAN VVPGISFSPD RMLQGRLFSY GDAQRYRLGV NHHQIPVNAP
361 KNPVNSYHRD GAMRVDGNQG ATPGVEPNSY GRWQEQPAYR DPAQAVGAVA DRFNYREDDD
421 NYFEQPGNLF RQMSPEQQQV LFENTARAID GASAQTIERH IGNCTQADPA YGAGVRKAIE
481 ALAAGNL
```

Figure 3.5: Sequence of catalaseA (Q9RJK9) from *S. coelicolor* showing the peptides matched in MALDI-TOF and nano-LC-ESI-MS/MS. Peptides positively identified only by MALDI are shown in red, sequenced by nanoLC-ESI/MS/MS are shown in blue and the regions identified in both experiments are marked in green. There was 29% sequence coverage by peptides detected in MALDI and 42% of the protein was sequenced in MS/MS experiment.

Nano-LC-ESI-MS/MS peptide search (Table 3.2) showed that most of the proteins identified showed comparatively less number of peptide matches and sequence coverage as compared to catalaseA. Most of these identified proteins showed same level of expression in control and DTT treated samples and some of them showed differential expression which were of diverse activities. Two proteins Adenosylhomocysteinase and Glu-tRNA Gln amidotransferase showed increased expression in DTT treated samples where as Glutamyl tRNA synthetase, Inosine 5′ monophosphate dehydrogenase, chaperonin, and a

putative oxidoreductase were shown to have increased expression in control as compare to DTT treated samples.

3.3.3. Effect of DTT on protein expression profile of *S. coelicolor* (observed in 2D-PAGE)

In order to identify other proteins involved in DTT stress, the protein extracts from control and DTT treated samples were separated by 2D-PAGE. To identify the subtle changes in protein expression profile, initially proteins were separated in the broad pI range of 3 to 10 in which almost all of the visible protein spots appeared in the acidic range of pI 4 to pI 7 (Figure 3.6).

In order to get better separation subsequently the proteins were separated in the narrow range of pI 4 to pI 7 (Figure 3.7). Both the 2D-PAGE gels showed similar pattern in which there was an induction of nearly 55 kDa protein representing catalaseA earlier found in the SDS-PAGE and MALDI-TOF results. Differences in the levels of expression of several other proteins were also observed between sample and control.

Differential expression of CatalaseA was confirmed in 2D-PAGE. In addition to several proteins of low molecular weight (10-40 kDa range) (marked with arrows in Figure 3.6 and 3.7) were observed to show differential expression. These protein spots were excised and subjected to in gel digestion followed by MALDI-TOF analysis.

We tried to identify all the proteins but due to some technical problems the spectra of some of the proteins did not yield any result. 2D-PAGE and MALDI-TOF analysis showed that total 19 proteins were identified to be induced in the presence of DTT. Control specific proteins were also studied. Among the five differentially expressed proteins from control two of them could be identified and are listed at the end of table 3.3. The results also confirmed the presence of CatalaseA with significant search results. Other proteins identified were of

diverse categories (Table 3.3) which were classified based on their properties as follows:

Transcriptional regulators: Putative transcriptional regulator, Putative tetR-family transcriptional regulator.

Primary metabolism: N-acetyl-γ-glutamyl-phosphate reductase, probable acyl carrier protein synthase.

Membrane proteins: Probable integral membrane protein, ABC transporter protein.

Secondary metabolism: Actinorhodin polyketide ketoreductase.

DNA/RNA metabolism: Putative secreted nucleosidase, Putative DNA ligase, ATP/GTP binding protein.

Hypothetical proteins (without any known functions): F23a, SC3F7.05c

Besides these the two proteins Metalloproteinase and β chain of Putative DNA-polymerase III were shown to be down regulated under DTT stress.

Table 3.2: Proteins identified in *S. coelicolor* after DTT treatment by LC-ESI-MS/MS. Bands were in gel digested with trypsin and peptide extract was analyzed using LC-ESI-MS/MS. Sequenced peptides were searched in *S.coelicolor* data base for protein identification using Spectrum Mill software provided by Agilient.

Sr. No.	Accession Number	Protein MW (Da)	Protein Name	DTT Treated		Control	
				Distinct Peptides	% AA Coverage	Distinct Peptides	% AA Coverage
1	21221466	52940	Adenosylhomocysteinase	15	33	9	21
2	**21218922**	**55117**	**Catalse (EC 1.11.1.6)**	**14**	**42**	**None**	**None**
3	21220653	51998	Leucyl aminopeptidase	7	17	10	26
4	21224181	49046	Zinc protease	6	19	3	8
5	21220671	52568	Glutamine Synthetase I	6	19	6	21
6	21224008	51761	Aldehyde Dehydrogenase	5	15	2	5
7	21223902	54917	Glutamyl t-RNA Synthetase	5	14	11	30
8	21219210	49814	Ferredoxin-NADP Reductase	4	11	2	7
9	21223783	51167	Pyruvate Kinase	4	17	8	21
10	21224028	50661	Aldehyde Dehydrogenase	3	8	4	9
11	21223855	52271	Glu-tRNA Gln Amidotransferase	3	7	None	None
12	21223306	52777	Histidine Ammonia Lyase	3	8	4	11
13	21224182	48910	Putative protease	3	7	2	4
14	21223141	57119	60 kDa Chaperonin cpn 60	2	6	4	12
15	21223149	52476	Inosine 5' Monophosphate dehydrogenase	None	None	7	19
16	21222689	56830	Chaperonin	None	None	4	10
17	21223293	49816	Putative Oxido-reductase	None	None	3	12

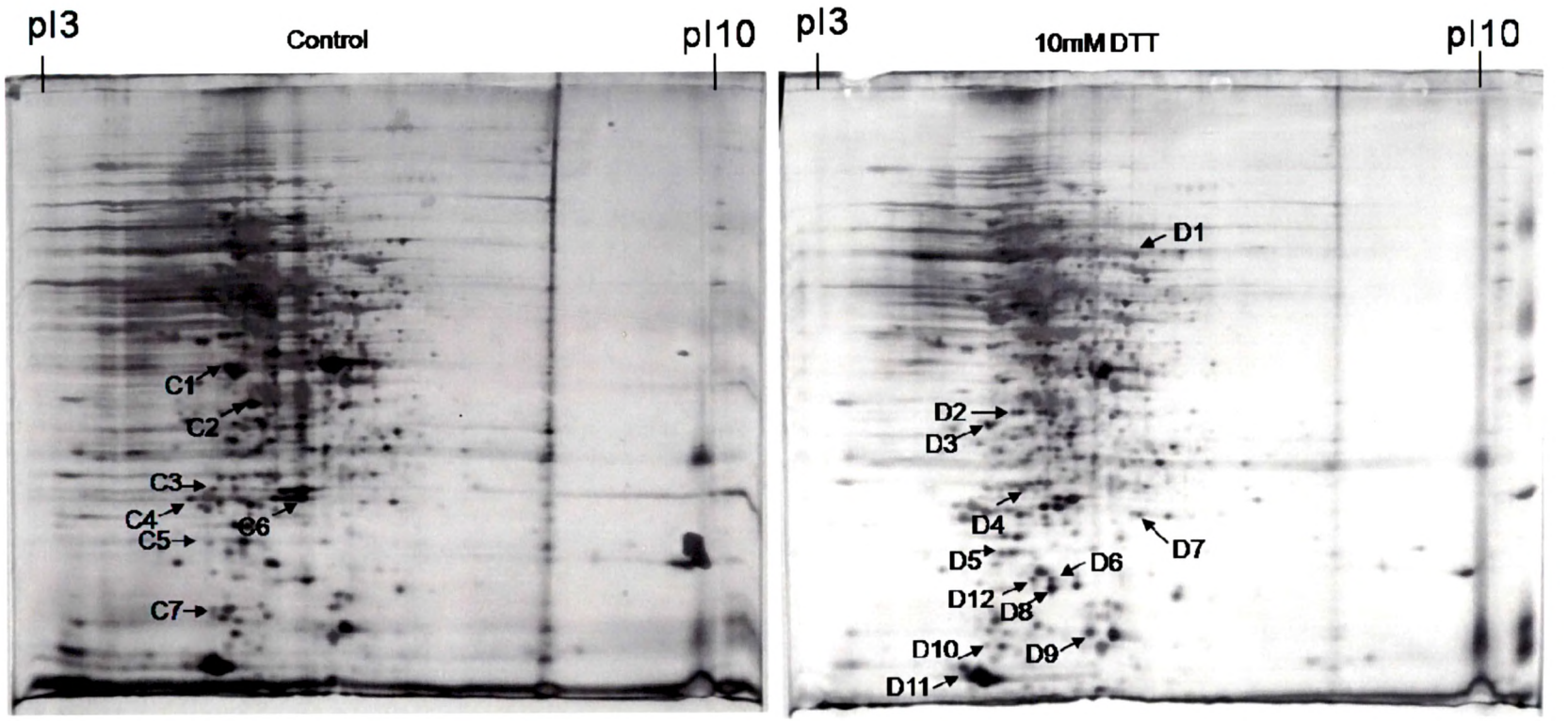

Figure 3.6: 2D PAGE profiles of proteins in the broad pI range 3-10. In the profiles several proteins show differential expression between Control and DTT induced cultures. Arrows indicating differentially expressed proteins in the two gels. Spots D1- D11 indicate proteins expressed in DTT treated sample only and Spots C1 to C7 indicate proteins expressed in control sample only.

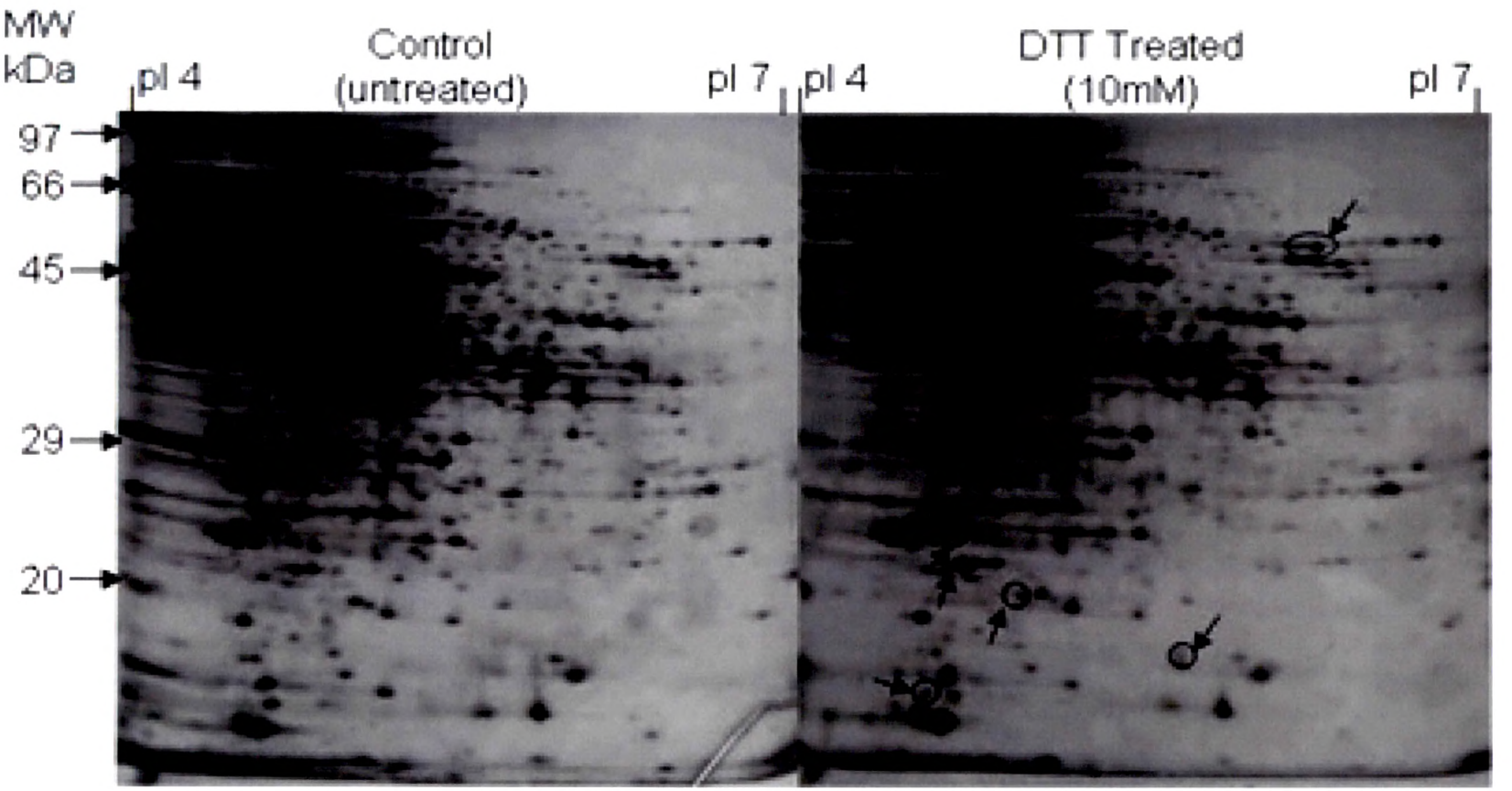

Figure 3.7 Separation of proteins by 2D PAGE in the narrow pI range 4-7. In the profiles several proteins show differential expression between Control and DTT induced cultures. Arrows indicate over expressed proteins in DTT treated samples only.

Table 3.3: Differentially expressed proteins under DTT stress as identified by 2D-PAGE, MALDI-TOF analysis and MASCOT search.

Spot No.	From 2D PAGE		Protein Identified from MASCOT search		Theoretical		Score
	MW kDa	pI			MW Da	pI	
D1a	55	6.36	Q9RJK9_STRCO	Catalase	55140	5.84	106
D2a	31.64	5.24	T36815	N-acetyl-γ -glutamyl-phosphate reductase	34710	5.88	43
D3a	30.17	5.13	T36500	hypothetical protein SCGD3.20	14705	6.18	29
D4a	29.28	4.84	Q9KXJ1_STRCO	probable membrane protein	34859	6.27	32
D5a	23.49	4.95	S22861	N-acetyl-γ-glutamyl-phosphate reductase	25360	5.88	36
D6a	22.7	5.53	Q9EWL0_STRCO	Putative transcriptional regulator	34109	11.47	39
D7a	22.5	5.53	T36196	Probable acyl carrier protein synthase	41295	6.35	32
D8a	22.7	5.3	Q9EWL0_STRCO	Putative transcriptional regulator	34109	11.47	40
D9a	21.32	5.9	T34884	Probable integral membrane protein	39529	6.16	33
D10a	21.04	4.84	Q9KXP4_STRCO	Conserved ATP/GTP binding protein	48049	6.04	30

D11a	20.54	4.59	Q9RIW3_STRCO	Putative integral membrane protein	11705	9.80	29
D12a	25.89	5.2	Q9RJF9_STRCO	ABC transporter protein, integral membrane subunit	28998	11.41	36
D1b	44.1	5.0	Q93IZ8_STRCO	Putative DNA ligase	76217	5.29	29
D2b	40.8	5.24	Q8CJN7_STRCO	Putative serine/threonine protein kinase	134332	4.8	31
D3b	34.5	4.62	Q9KIM3_STRCO	F23a	6272	9.69	35
D4b	27	4.88	T34911	Hypothetical protein SC3F7.05c	15659	6.06	37
D5b	20.6	4.73	A28788	Actinorhodin polyketide ketoreductase (EC 1.1.1.-) actIII	27476	5.27	35
D6b	14.9	4.55	Q9EX61_STRCO	Putative tetR-family transcriptional regulator	20769	8.15	60
D7b	13.7	4.38	Q9L1L3_STRCO	Putative secreted nucleosidase	30833	6.23	36
C1a	35.95	5.08	Q9AD31_STRCO	β chain of putative DNA-polymerase III	39156	4.75	37
C2a	31.75	5.3	S25187	Metalloproteinase	22309	5.63	45

3.4. Discussion

Our intention in subjecting *S. coelicolor* to thiol stress was to identify proteins which play a role in thiol, disulphide metabolism. The answer to the quest came from an interesting quarter. CatalaseA, an enzyme that converts hydrogen peroxide to water during oxidative stress, was found to be induced under thiol stress too. Hydrogen peroxide is generated as a byproduct of redox reactions inside a cell (Messner and Imlay, 1999; Gonzalez-Flecha and Demple, 1995). Hydrogen peroxide being an oxidizing agent is also a source of extremely potent hydroxyl radicals. Apart from damaging lipids and proteins, hydroxyl radicals can attack the bases and sugar phosphate backbone of DNA causing modifications and strand breaks (Hutchinson, 1985). Catalases, enzymes found in prokaryotes as well as eukaryotes are meant to convert endogenous hydrogen peroxide to omnipresent, harmless water. *S. coelicolor* has three different catalases. Two of these are monofunctional catalases, CatalaseA or CatA (Cho and Roe, 1997) and CatalaseB or CatB (Cho et al., 2000), while the third one is a catalase peroxidase, also referred to as CatalaseC or CatC (Hahn et al., 2000). Among these three catalases, CatA is the major catalase implicated in proper growth and resistance against H_2O_2.

As presumed earlier being a reducing agent, DTT would have opposite effects to that of oxidative stress. Surprisingly, we found induction of catalaseA.

From the results of MALDI-TOF and nano-LC-ESI-MS/MS analysis of the induced protein spots, the following conclusions can be drawn: Primarily, catalaseA is the major protein induced under DTT stress. Secondly, other proteins induced under DTT stress were of diverse categories. Many of them were having unknown or putative functions. In conclusion, DTT has profound and varied response which is likely to occur due to redox imbalance in the cell.