The discovery of haemoglobins (Hbs) in invertebrates, eukaryotes, plants and bacteria was a startling revelation in the history of molecular biology because it challenged the age-old notion of haemoglobins being the O_2 transport molecules. Studies on microbial and invertebrate Hbs revealed large variations in their structures and functions, apart from the conservation of few key residues. It is suggested that the structural variations are reflective of the diverse functions of the haemoglobin molecules that they play in their respective hosts. With newer findings adding to our knowledge about the roles of Hbs, the exact functionalities of these molecules are difficult to generalise and need to be explored extensively. Single domain, two-domain and truncated haemoglobins (trHbs) have been identified in several important bacterial species and many of them are highly pathogenic. *Mycobacterium tuberculosis* is one of the most important human pathogen that carries three distinct Hbs.

Amongst all infectious organisms, because of its extraordinary adaptive response, *Mycobacterium tuberculosis* still remains a challenge for modern day medical investigators. Rich literature is available suggesting the strategies adopted by the bacterium to survive in the inhospitable conditions. The bacterium broadly adopts two strategies to limit the hostile environment inside the macrophages; the first is restricting the phagosomal acidification to pH 6.4 and limiting fusion with lysosomes (Russel, 2007). The second strategy it adopts is the setting up of the bacterial components as effectors to combat the immune response directed against the bacterial antigens u(Gengenbacher and Kaufmann, 2012). *M. tuberculosis* is known to possess multiple Hb genes including both group I and II trHbs, the HbN and the HbO. Recent findings about these trHbs suggest that they might be an important weapon in the arsenal of *Mycobacteria* for survival during oxidative and nitrosative stress.

The trHbs, HbN and HbO, belonging to the group I and II of the trHb family possess the non-classical two-over-two globin fold and show conservation of key amino acids required for the structural stabilization of the trHb fold. The sequence identity shared by the HbO and the the HbN is only about $\leq 30\%$. The HbN possesses an N-terminal extension of 12 amino acids which is lacking in the HbO. These two trHbs exhibit distinctly different tunnel structures connecting the protein surface to their respective active sites. This tunnel is long and continuous in the case of the HbN, and for the

HbO it is short and restricted (Bolognesi et al., 1982; Milani et al., 2003; Perutz, 1989; Scott et al., 2002). Although, their elaborate physiological functions are yet to be established, trHbN and trHbO have been identified to perform some important functions in *M. tuberculosis*. Studies on the HbN of *M. tuberculosis* in heterologous hosts *Escherichia coli* and *Mycobacterium smegmatis* demonstrate that it is capable of metabolising NO and convert it into a non-toxic form (Pathania et al., 2002a). Subsequently, it has been demonstrated that the pre-helical region of the HbN plays a vital role in NO-dioxygenation (NOD) and the management of nitrosative stress (Lama et al., 2009). The NOD activity of the HbN is much higher than the HbO (Ascenzi et al., 2006; Ouellet et al., 2003; Ouellet et al., 2002; Pawaria et al., 2007). This activity is comparable to that of flavoHbs which have an additional reductase domain to facilitate the process of electron cycling (Singh et al., 2014). Another important finding about the HbN is that it gets post-translationally modified via glycosylation in its native host. Mannose-bound HbN is suggested to affect the virulence and pathogenicity in *M. tuberculosis* (Arya et al., 2013). The HbO has been identified as a membrane associated protein. Experimental evidences suggest its role in O_2 uptake, aerobic respiration and cellular respiration during hypoxia and latency (Liu et al., 2004; Pathania et al., 2002b). The HbO of *M. leprae* exhibits peroxynitrite detoxification (Ascenzi et al., 2009). Apart from these studies, little is known about the physiological functions of trHbs.

On one hand some literature is available about the structure and functions of trHbs, HbO and HbN, of *M. tuberculosis,* essentially nothing consequential is known about the trHbs, SgO and SgN, of the non-pathogenic *M. smegmatis*. The SgO and SgN of *M. smegmatis* show around 70-80% sequence identity with the HbO and HbN of the pathogenic *M. tuberculosis*. This can be suggestive of similarities in their functionalities; however, no significant research is available on these trHbs so far. Study on the SgN shows that it does not exhibit an efficient NOD activity possibly due to lack of the pre-A region, which modulates the NOD function of the HbN of *M. tuberculosis* (Lama et al., 2009). This indicates that in spite of close similarities, the functions of these trHbs in different mycobacterial species might be variable. Study of trHbs of *M. smegmatis* holds relevance because it adds to our understanding of the variegated functions performed by the members of the Hb family of proteins in the

non-pathogenic *Mycobacteria.* Such studies might also shed light on the molecular mechanisms of *M. smegmatis* can therefore help us expand the scope of research on *M. tuberculosis* as well.

The trHbs, SgO and SgN, of *M. smegmatis* are coded by the *glb*O (Msmeg_4695) and the *glb*N (Msmeg_5765) genes, respectively. Western blot analyses showed that the SgO gets expressed in its native host, *M. Smegmatis,* during the stationary phase of growth, unlike, trHbO of *M. tuberculosis* that is expressed during all stages of growth in its host. On the other hand the SgN was found to be expressed during the stationary phase; similar to the HbN of *M. tuberculosis* (Couture et al., 1999; Pathania et al., 2002a). It was also observed that the expression of the SgO and the SgN in their native host, *M. Smegmatis,* is relatively low. Western blot analysis of the cellular fractions show that trHbs, SgO and SgN, were expressed in the cytoplasmic as well as the membrane fractions of *M. smegmatis.* The trHbs, HbN and HbO, of *M. tuberculosis* also exhibit a similar pattern of expression. This could indicate that the SgO and the SgN might be membrane associated proteins like the HbO and the HbN. One of the most important physiological properties of Hbs is their ability to bind to various ligands. The O_2-binding ability of the HbN is considerably high which suggests that this molecule is not apt for O_2 transport or storage roles, while that of the HbO is relatively lower than that of the HbN which makes HbO a better candidate for such functions in *M. tuberculosis* (Lama et al., 2006; Pathania et al., 2002a). The rate constants of O_2-association were determined for the SgO and the SgN of *M. smegmatis* by laser-flash photolysis method. The results show that both these trHbs have similar rates of association with O_2, which is however significantly lower than that of the HbN of *M. tuberculosis*. The CO association rate constants were also determined for both these trHbs and it was observed that this rate was similar for SgO and HbN, but for SgN the rate constant value was double that of SgO and HbN. Although the consequences of the different association rates are not fully understood, this represents a remarkable variation in the behaviour and activity of *M. smegmatis* trHbs. The Q-bands observed in UV-Visible spectra of the SgO and the SgN do not show splitting which is indicative of a pentacoordinate state of heme-Fe in these Hbs. From the gel-filtration chromatography profile and MALDI-TOF analyses it was observed that the SgO and the SgN were purified in their dimeric form. To understand

the physiological implications of trHbs of *M. smegmatis* and *M. tuberculosis*, studies were performed on *M. smegmatis* cells overexpressing Hbs SgO, SgN, HbO and HbN. The trHb-overexpressing cells showed better growth under hypoxic conditions. The increase was maximum for cells overexpressing the HbO and HbN of *M. tuberculosis*, followed by cells overexpressing SgO and SgN of *M. smegmatis*. In Wayne model of hypoxia, both SgO and SgN demonstrate better oxygen uptake indicating that the expression of these trHbs allows their host to metabolize O_2 more efficiently. Also trHbs, overexpressing *M. smegmatis* cells exhibit better viability in hypoxia. This effect was more pronounced for cells overexpressing HbO and HbN of *M. tuberculosis*, followed by the cells overexpressing SgO and SgN of *M. smegmatis*. These results indicate a protective effect of trHbs in *M. smegmatis* under hypoxic stress.

TrHbs exhibit large variations in their structural organisation as compared to the conventional Hbs. Few members of the Hb superfamily proteins in eukaryotes *Paramecium* and *Tetrahymnea* exhibit post-translational modifications, which were identified to be acetylation (Iwaasa et al., 1989; Iwaasa et al., 1990). Recently, a study on the HbN in our lab showed that HbN gets post translationally modified at the C-terminal region via glycosylation and the attached glycan was identified as mannose (Arya et al., 2013). *In silico* analyses of the sequences of the HbN and the SgN showed significant difference in C-terminal regions of these proteins. SgN exhibits the absence of the region that is involved in the glycan linkage in the HbN. However, further *in silico* analyses suggested the possibility of post-translational phosphorylation in the SgO, the HbO and the SgN. Auto-phosphorylation activity in GlbO proteins was confirmed by radio-activity based auto-phosphorylation assay using γP^{32}-ATP. While auto-kinase assay confirmed the *in vitro* auto-phosphorylation of SgO; positive phosphostaining of the spot representing the SgO in the 2D-gel of *M. smegmatis* cell lysate confirmed the *in vivo* phosphorylation of the SgO in *M. smegmatis*. Bioinformatic analyses predicted Thr-106 to be the site of phosphorylation and loss of auto-kinase activity in the point mutant T106A-SgO confirmed this prediction. On examining the sequence of trHb, SgO for identification of the ATP-binding site, it was observed that SgO lacked the consensus ATP-binding motif represented by G-x(4)-GK-[TS]-LhhhhD-, where x denotes any amino acid. However, it possessed a sequence similar to the Walker B motif, -hhhhDE-, where h

denotes any hydrophobic amino acid and Asp and Glu have important roles to play (Hanson and Whiteheart, 2005; Walker et al., 1982). Considering this information, the point mutant D10A-SgO was created and it was found to lack auto-kinase activity, proving our assumption that Asp at the 10^{th} position forms an important part of the ATP-binding site in the SgO. SgN, on the other hand did not exhibit any auto-kinase activity. This observation brings to light a fundamental difference between the SgO and the SgN and indicates their different functionalities within the same host. After establishing auto-kinase activity in the SgO, another study elucidated the effect of change in heme-Fe oxidation state on the auto-kinase activity of the SgO. It was noticed the like few other heme-proteins, the auto-kinase activity displayed by Fe^{+3}-SgO was maximum. Aerobically purified SgO and O_2-bound Fe^{+2}-SgO showed nearly similar auto-kinase activities, and these were significantly less than that of Fe^{+3}-SgO. The unbound Fe^{+2}-SgO exhibited least auto-kinase activity (Singh and Kumar, 2015). This observation suggests the role of phospho-bound SgO in oxygen/redox sensing.

Literature shows that the Walker motif is present in several ATP and GTP-binding proteins (Walker et al., 1989; Hanson and Whiteheart, 2005). Radioactivity based auto-kinase assay confirmed the auto-kinase activity of both the SgO and the HbO using γP^{32}-GTP. This proved that both the HbO and the SgO could be phosphorylated by ATP and GTP. It was further investigated as to which out of ATP and GTP could be preferred as a phosphate group donor in the case of SgO of *M. smegmatis* and HbO *M. tuberculosis*. A series of ATP-GTP competition based auto-phosphorylation assays clearly confirmed that the SgO shows a preferential binding to GTP whereas the HbO shows preferential binding to ATP. Molecular basis and the functional relevance of SgO selectivity of GTP over ATP is not clear at present. No reports of such kind exist in literature so far. It was interesting to note that in spite of a significant sequence identity, the SgO and the HbO show variation in regard to their preferred phosphate donor. Phosphorylation is considered to be one of the most significant post translational modifications, regulating various physiological processes and metabolic pathways. Phosphorylation of trHbs, SgO and HbO, could be of great significance as it might be capacitating the trHbs to function in numerous other ways which are not yet acknowledged. The results also indicate the intrinsic differences between these two orthologous proteins.

TrHbs in *Mycobacteria* appear to orchestrate some unanticipated and explained physiological changes. It requires a deeper probe to understand their wholesome implications in the mycobacterial system. In pursuit of further understanding of the roles of trHbs and their involvement in oxygen metabolism, 2D proteomic profiling was performed for *M. smegmatis* during overexpression of trHbs SgO, SgN, HbO and HbN in *M. smegmatis,* under aerobic conditions and in hypoxia. Previous studies undertaken on mycobacterial proteomics have been largely targeted towards understanding the proteomics of *M. tuberculosis* system, the biomarkers of the tuberculosis disease progression and identifying new therapeutic targets (Betts et al., 2000; López et al., 2016; Murugasu-Oei et al., 1999; Prasad et al., 2013; Rao and Li, 2009; Rosenkrands et al., 2002; Starck et al., 2004). Not many studies are available in case of *M. smegmatis* and none exists so far focussed towards elaborating our understanding of the trHbs in *Mycobacteria.* In our proteomic profiling studies conducted for each case of trHb overexpression around 90-110 proteins were detected by the PDQuest 8.0.1 software and we could successfully identify nearly 25-40 proteins with an altered expression using Mass Spectroscopy. It was observed that a majority of the identified proteins were upregulated as compared to that in the wild type (and control) *M. smegmatis.* Because of the limitation of resolution and identification of lowly expressed proteins in 2D gels was not possible to identify the proteins with low expression levels. These studies conducted on *M. smegmatis* showed that overexpression of trHbs causes distinct changes in the expression pattern of several proteins. These proteins could be broadly classified as the metabolic and bioenergetics pathways enzymes, oxidoreductase and redox family proteins, fatty acid biosynthetic pathway enzymes and stress proteins. Certain proteins involved in protein folding and transport antigenic proteins, DNA binding proteins, and those involved in transcription and translation were also found to exhibit an altered expression. The alterations observed in the protein expression levels give us an insight of the umpteen metabolic changes taking place in the *M. smegmatis* system. Previous literature indicates that the upregulation of metabolic enzymes indicates enhanced cellular metabolism and the cells' preparatory phase for adapting to the stationary phase (Hampshire et al., 2004). The increase in the fatty acid synthesis enzymes correlate to the biological event of thickening of the mycobacterial cell membrane,

which is another event marking the cells' entry into the dormant phase and an adaptation to stress (Cunningham and Spreadbury, 1998; Starck et al., 2004). The upregulation of stress proteins is usually a characteristic feature during oxidative stress, heat shock and hypoxia. The overexpression of trHbs also affects the expression of several essential proteins like Electron Transfer Flavoproteins α and β, Elongation Factor Tu, DNA directed RNA Polymerase, S-Adenosine Methionine Synthatase, ssDNA-binding protein, β-keto adipyl CoA Thiolase and Antigen 85A. Several key proteins like Enolase, Fructose-1, 6-bispospho aldolase, Phosphoglygerate kinase, HpCE protein, Alcohal dehydrogenases Msmeg_6242 and Msmeg_1977 were also found to be down regulated in case of trHb overexpression. There were few anomalies reported with regard to the expression of certain proteins. In case of overexpression of the HbO and the HbN of *M. tuberculosis* in *M. smegmatis,* the upregulation of antigenic proteins and proteins involved in antibiotic metabolism pathway was an intriguing observation. This was not observed in case of overexpression of the SgO and the SgN of *M. smegmatis* in the native host. In case of SgO overexpression, the high levels of the protein Trigger factor were observed. Literature studies indicate a functional cooperation between Trigger Factor, DnaK and GroEL in *E. coli* (Hoffmann et al., 2010). This similar observation might indicate certain changes in this particular metabolic pathway in case of *M. smegmatis* as well. Nucleotide diphosphate kinase was another protein found to be upregulated only in case of overexpression of the SgO. There were certain anomalies observed in case of overexpression of the HbN *M. smegmatis.* The levels of DNA directed RNA polymerase were found to have nearly a three-fold increase that all the other cases of trHb overexpression. Similarly the Elogation Factor Tu and Thiosulfate sulphur transferase showed a two-fold increase as compared to the other cases. Fructose-1,6-bisphosphatase was found to be upregulated in case of overexpression of the HbN in *M. smegmatis* grown in both aerobic condition and hypoxia. This protein could not be identified in any other case.

In case of hypoxia grown trHb overexpressing *M. smegmatis* cells, Thiol Peroxidase, ABC transporter substrate binding protein and ABC type amino acid transport system and Chain A, Main Porin protein were upregulated in case of both the SgO and the HbO overexpression. Universal Stress Protein Family Protein UspA was found to be

downregulated during hypoxia in both cases of SgO and HbO overexpression. This relative decrease could be attributed to an increase in expression of stress proteins in wild type *M. smegmatis* during hypoxia. Similarly, the alcohol dehydrogenase Msmeg_6242, that showed a significant decrease during aerobic growth conditions, showed nearly a similar expression during hypoxia in wild type *M. smegmatis* and SgO and HbO overexpressing *M. smegmatis*. In case of hypoxia grown cells, the overexpression of *glb*O and *glb*N genes showed almost a similar pattern of protein expression, however; in aerobic and hypoxic growth of the HbN overexpressing *M. smegmatis*, Fructose-1, 6-bisphosphatase showed an increased expression. Glycerol kinase was downregulated during aerobic growth whereas it was upregulated during hypoxia; on the other hand, Molecular Chaperone GroEL and Phosphoglycerate Kinase were upregulated during aerobic growth and downregulated during hypoxic growth. HpCE protein was upregulated during hypoxia in case of the SgN overexpression.

When a similar study was conducted to check the effects of overexpression of trHb, HbO and HbN in *M. tuberculosis* H37Ra grown in aerobic conditions, the PDQuest software could identify around 100 spots in each case out of which around 20-25 proteins with altered expression were identified by Mass spectroscopy. These included metabolically important enzymes and proteins like Electron Transfer Flavoproteins α and β, Elongation Factor Tu, Peptidyl-prolyl-cis-trans Isomerase, Nucleoside Diphosphate Kinase, β-keto acyl CoA Thiolase (FadA3), 3-keto-acyl-ACP Reductase (FABG4). It was also observed that the expression of antigenic and immunogenic proteins like Antigen 85B, Immunogenic protein Mpt 63, and Chain A, Mtb Antigen 85b were upregulated. Stress related proteins like 60 kDa Chaperonin 2 GroEL and Iron regulated Universal Stress Protein Family Protein were also upregulated. Overall, these proteomic studies suggest that there are multifarious implications of trHbs overexpression in *Mycobacteria*. We observe that more than one metabolic pathway are being affected due to the overexpression of trHbs in the mycobacterial systems. This could be due to the ability of trHbs to bind O_2 and facilitate several redox changes in the cells or trHbs could be executing such complex cellular responses with the help of other unidentified interacting partners. But a befitting explanation to these phenomena would require further extensive experimentation for which proteomic research could be instrumental.

To conclude, the present study on trHbs of *M. smegmatis* has provided some new insights into their possible physiological functions. The auto-phosphorylation ability, their protective effects on the cells during hypoxia and the differential protein expression cause by the overexpression of trHbs give clear evidence on functional versatility and complex roles of these proteins in the cellular metabolism, suggesting that trHbs may not be assessed as simple O_2-binding molecules anymore. Co-existence of more than one trHbs in the same host highlights distinctive and cooperative functions of trHbs. After establishing the auto-kinase ability of the *glb*O genes, it would be interesting to pursue further studies on the implications of phosphorylation of trHbs in *Mycobacteria.* Identification of the interacting partners of trHbs and the metabolic pathways they affect would also be challenging and yet instrumental in our fight against the notorious mycobacterial infections.