Available online at www.sciencedirect.com
SCIENCE DIRECT
Carbohydrate Research 341 (2006) 1164–1173
Carbohydrate RESEARCH

Conformational analyses of mycothiol, a critical intracellular glycothiol in *Mycobacteria*

Christine E. Hand,[a] France-Isabelle Auzanneau[b] and John F. Honek[a,*]

[a]*Department of Chemistry, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1*
[b]*Department of Chemistry, University of Guelph, 50 Stone Road East, Guelph, Ontario, Canada N1G 2W1*

Received 18 November 2005; received in revised form 6 March 2006; accepted 11 March 2006
Available online 21 April 2006

Abstract—Intracellular thiols are essential biomolecules, which play several critical roles in living organisms including controlling intracellular redox potential and acting as cofactors for several vital detoxification enzymes including *S*-transferases and formaldehyde dehydrogenases. The tripeptide γ-L-glutamyl-L-cysteinylglycine, more commonly known as glutathione, is well known as the major intracellular thiol in eukaryotes and in some bacteria. However, glutathione is absent in the *Actinomycetales* bacteria such as *Mycobacteria* and *Streptomyces* and is believed to be replaced by 1-D-*myo*-inosityl-2-(*N*-acetyl-L-cysteinyl)amido-2-deoxy-α-D-glucopyranoside, mycothiol, in these organisms. Although much is known about the chemistry and biochemistry of glutathione, currently much less is known concerning mycothiol and its properties. The structure of mycothiol is composed of a glycoside linkage between *myo*-inositol and D-glucosamine with an *N*-acetyl-L-cysteine linked to the 2′-amino group of the D-glucosamine moiety. Mycothiol is currently of intense interest due to its essential role in the cellular physiology of *Mycobacteria*, such as *Mycobacterium tuberculosis*, and its possible role in antimycobacterial drug resistance. A detailed investigation of its chemistry is therefore essential in ameliorating our knowledge of this key glycothiol, and in shedding additional light on its biochemical role in these pathogenic organisms. This report presents a detailed conformational analysis of mycothiol utilizing a variety of force fields and stochastic search protocols. Cluster analyses of energetically low lying conformations have indicated the presence of several key conformations that are populated in the gas phase and with implicit water solvation. These conformations are compared to recent NMR studies on a derivative of mycothiol. This information should be an important contribution to our basic understanding of the chemistry of this glycothiol and critical in the design of novel inhibitors of pathogen enzymes that require it.

Keywords: Mycothiol; Conformational analyses; *Mycobacteria*; Glycothiol

1. Introduction

Intracellular thiols are essential for the maintenance of cellular redox homeostasis, as well as for the protection of the cell from reactive oxygen and nitrogen species.[1,2] The most well studied of these thiols is the tripeptide glutathione (γ-L-glutamyl-L-cysteinylglycine, GSH, **1**). GSH is the major thiol in eukaryotes and most Gram-negative bacteria including the cyanobacteria, and the purple bacteria.[3] Not all organisms use GSH as their major thiol: the halobacteria use γ-L-glutamyl-L-cysteine,[4] the pathogenic protozoa of the genera *Trypanosoma* and *Leishmania* use trypanothione (N^1,N^8-bis(glutathionyl)spermidine, **2**),[5] and the *Actinomycetales* bacteria use mycothiol (1-D-*myo*-inosityl-2-(*N*-acetyl-L-cysteinyl)amido-2-deoxy-α-D-glucopyranoside, MSH, **3**).[6]

The *Actinomycetales* bacteria include *Streptomyces* and *Mycobacteria*; the former being an industrially important producer of a large number of antibiotics, and the latter being of great clinical and agricultural importance. The most well known disease associated with *Mycobacteria* is tuberculosis (TB), which is caused by *Mycobacterium tuberculosis* in humans[7] and *Mycobacterium bovis* in mammals.[8] The emergence of antibiotic resistant strains of *M. tuberculosis* has brought this disease back to the forefront of clinical research: strains

* Corresponding author. Tel.: +1 519 888 4567x5817; fax: +1 519 746 0435; e-mail: jhonek@uwaterloo.ca

0008-6215/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2006.03.020

resistant to single drugs have been reported in every country surveyed and strains that are resistant to all of the major TB drugs have been documented.[7] According to the World Health Organization, a third of the world's population is infected with TB.[7] *Mycobacteria* are linked to a number of other clinically and agriculturally important diseases including Johne's disease (*Mycobacterium avium* subspecies *paratuberculosis*, MAP), which affects cattle and causes losses to the dairy industry in excess of $1.5 billion (US) per year.[9] The ability of *Mycobacteria* to resist the oxidative stresses placed upon it by macrophages during host infection may be due to MSH and its biochemistry.[10] Mutants of *Mycobacterium smegmatis* deficient in MSH production have been found to be hypersensitive to a host of toxic agents; therefore, MSH and its biosynthetic pathways act as an attractive target for the development of novel antimicrobial agents specific for *Mycobacteria*.[11]

Despite the medical importance of the organisms that utilize MSH, very little is known concerning MSH biochemistry.[†] Much of the work published thus far has focused on the enzymes believed to be involved in the MSH biosynthetic pathway (MshA through MshD).[1] In addition, three enzymes have been characterized to date as requiring MSH as a cofactor; all of these enzymes have parallel functions to GSH-utilizing detoxification enzymes: MSH-dependent formaldehyde dehydrogenase,[1] MSH-disulfide reductase[1] and MSH-*S*-conjugate amidase (MCA). Several partial or total chemical syntheses of MSH have been reported.[12,13]

There is very little information available with regards to the chemical reactivity and conformational preferences of MSH, information that could be of benefit to the design of novel inhibitors of MSH-utilizing enzymes. Based upon these considerations, a detailed conformational study of MSH was undertaken. The findings reported herein result in the elucidation of important characteristics of the major conformations of MSH predicted by these methods. The results of these investigations should not only aid in our fundamental understanding of this glycothiol but also should serve to aid in the design of MSH analogues, including conformationally restricted derivatives as important leads in the development of novel antimycobacterial agents.

2. Computational methods

2.1. General methods

A conformational study of MSH was undertaken using MacroModel (versions 8.1 and 9.0, Schrödinger Inc., Portland, OR, USA) and the Molecular Operating Environment program suite MOE2004 (Chemical Computing Group Inc., Montreal, QC, Canada). MacroModel and MOE use the Generalized Born/Surface Area (GB/SA) continuum model[14] to create an implicit solvation environment and this model was used as implemented by each software package. The MacroModel calculations were performed on an SGI O_2 workstation or the University of Waterloo multi-CPU Silicon Graphics Origin 3800 system (Flexor) in the gas phase and with implicit water solvation. MOE2004 calculations were performed using implicit water solvation on a PC Pentium 4, 2.40 GHz CPU.

2.2. MacroModel calculations

All MacroModel minimizations were performed using the same protocol. The electrostatic treatment was dependent upon the force field used, with a dielectric constant of 1.0; the charges were defined by the force field library. Hydrogen-bonding interactions were cut off at 4 Å. Van der Waals and charge–charge interactions were cut off at 7 and 12 Å, respectively, in the gas phase, at 8 and 20 Å under implicit solvation. A Polak–Ribiere Conjugate Gradient (PRCG)[15] was used and each structure was subjected to 5000 iterations. The minima converged on a gradient with a threshold of 0.05.

An initial survey of the force fields available in MacroModel was undertaken to determine, which were appropriately parameterized for application to the MSH conformational study. Initial minimizations of MSH with each of the following force fields,[‡] AMBER*,[16] AMBER94,[17] MM2*,[18] MM3*,[19] MMFF,[20] OPLS[21] and OPLS-AA,[22] were performed, which returned information on the number and quality of the bend, stretch and torsional parameters for the MSH structure (Table 1, see also Results and discussion). Of the force fields available in MacroModel v. 8.1, the OPLS-AA[§] force field appeared best suited to model MSH; therefore, OPLS-AA was chosen as the main force field for the conformational study performed using MacroModel.

An X-ray structure of MSH is not currently available; therefore, the corrected structure of MSH,[¶,23] was drawn in Maestro (Schrödinger Inc.) and was minimized in the gas phase and with implicit water solvation (GB/SA) using the OPLS-AA force field. The minimized structures were then used as the starting molecular structures for all MacroModel conformational searches. The conformational space of MSH was searched using

[†] A recent survey of the PubMed database retrieved 65 publications on MSH, as compared to 69,370 publications on GSH (February 21, 2006).

[‡] As implemented by MacroModel v. 8.1.

[§] OPLS-AA was used as implemented by MacroModel v. 8.1.

[¶] There was initial confusion regarding the stereochemistry of the inositol moiety of MSH as it was often written as 1-D but drawn as 1-L; MSH has since been confirmed to be 1-D.[23]

Table 1. Number of each quality of the various force field parameters available in MacroModel for bond stretching and bending and torsion angle rotation in MSH

Force field[a]	Stretch			Bend			Torsion		
	High	Med	Low	High	Med	Low	High	Med	Low
AMBER*	37	0	0	111	1	0	152	11	0
AMBER94[b]	—	—	—	—	—	—	—	—	—
MM2	75	8	0	140	23	0	143	108	0
MM3	35	14	14	58	73	137	82	20	1
MMFF	63	0	0	113	0	0	173	0	6
OPLS	63	0	0	112	0	1	170	9	0
OPLS-AA	**63**	**0**	**0**	**113**	**0**	**0**	**179**	**0**	**0**

[a] As implemented by MacroModel v. 8.1.
[b] Insufficient force field constants were available—no constants were available for the O_{ring}–C1–$O_{linkage}$ bend.

the Monte Carlo multiple minimization (MCMM) method.[24] The general procedure for a MCMM conformational search involves the choice of a stable starting structure, which then undergoes random variations of dihedral angles, followed by minimization and comparison of the new conformer to previously determined minima. If the structure contains a ring, this ring is opened prior to dihedral variation and then closed before minimization. The structure is either stored as new and unique or rejected as identical and the cycle is repeated. One cycle is termed a Monte Carlo (MC) step.[24] In this study the structures were allowed 500,000 MC steps and structures within 50 kJ/mol of the global minimum were retained.

The starting structure for each step of this study was selected from the conformations thus far generated (in the case of the initial step, this was the inputted structure). This starting geometry was the 'least used' structure based on the number of times each conformer had been used as the starting structure; this structure would only be used if it is within 50 kJ/mol of the lowest energy structure so far identified. For each starting structure, x number of degrees of freedom, or in this case torsional angles, were altered, where x is a randomly generated number between 2 and 5. The bonds would be rotated plus or minus a randomly generated number between 0° and 180°. To adjust the torsional angles within the D-glucosamine and *myo*-inositol rings, the bonds between C3 and C4 and C2′ and C3′ were temporarily broken to open each ring. The dihedral angles were then modified and the rings closed, with the newly formed bonds constrained to between 0.5 and 2.5 Å in length. The absolute values of the torsion angles of the amide bonds were monitored and the structure was discarded if its angles were outside of 0–90° for C2–N1″–C2″–O (D-glucosamine) and 90–180° for H–N5″–C6″–O (*N*-acetyl-cysteine).

The resultant conformation was then minimized using the OPLS-AA force field. After 1666 iterations, a preliminary energy test was performed and the minimization was aborted if the structure's energy was more than 100 kJ/mol higher than the current minima structure; speeding the overall conformational search. The chiralities of the stereogenic atoms, the carbon atoms of the D-glucosamine and *myo*-inositol rings and the α-carbon of the *N*-acetyl-cysteine moiety, were saved by the computation of an improper torsion with the first three substituents and then rejection of any structure whose chirality was altered by minimization; the structure was then compared to previously saved conformers for redundancy. Structures whose least squares superimposition was less than 0.25 Å were considered identical and the higher energy conformer was then discarded. Initially, all heavy atoms and polar hydrogen atoms were used for comparison. After these initial conformational searches 13,074 and 137,764 unique structures were found in the gas phase and with implicit water solvation, respectively. These conformers were then subjected to the redundant conformation elimination (RCE) module of MacroModel to remove conformers whose heavy atom placement differed by less than 0.25 Å.

A second search was performed using AMBER*[‖] to account for any inadequacies of the OPLS-AA force field (see Results and discussion). MSH was minimized using AMBER* with implicit solvation (GB/SA) and this was used as the starting point for the conformational search. The MCMM conformational search was performed similar to the OPLS-AA search except that the torsional angles within the *myo*-inositol and D-glucosamine rings were not modified nor were these rings opened; the torsional angles between these rings were subject to rotation (see Results and discussion). After the initial conformational search, 33,124 conformers were found, which were then subjected to RCE.

Implicit water conformers within 3 kcal/mol of the global minimum (OPLS-AA: 198 conformers, AMBER*: 116 conformers) were then analyzed using the XCluster module of MacroModel v. 9.0. The conformers were grouped into families based upon the ϕ (H–C1–O–C1′) and ψ (C1–O–C1′–H) angles (for numbering, see Fig. 1, **3**) of the disaccharide portion of MSH. Conformers containing five axial substituents on the *myo*-inositol

[‖] AMBER* was used as implemented by MacroModel 9.0.

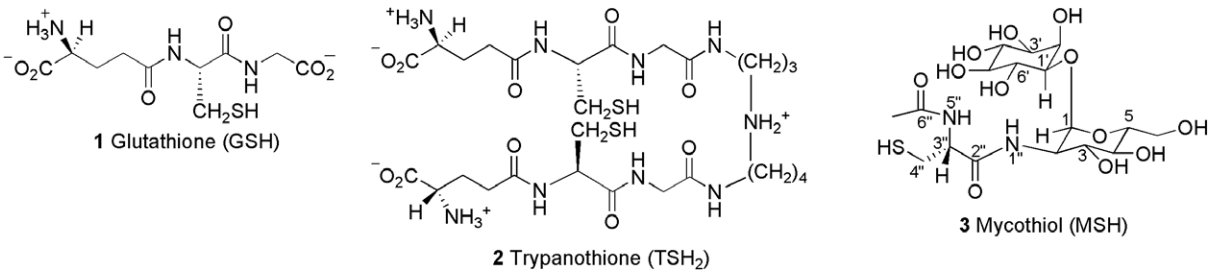

Figure 1. Important intracellular thiols.

ring were excluded from further analysis (OPLS-AA: 50 conformers, see Results and discussion).

2.3. MOE2004 calculations

Stochastic conformational searches were performed using the corrected structure of MSH[¶,23] as a starting structure. These searches are similar to random incremental pulse searches (RIPS),[25] where new molecular conformations are generated by random rotation of all bonds (including ring bonds), instead of only Cartesian coordinate perturbation. This method led to alterations of the absolute configuration of the *myo*-inositol carbon atoms during the perturbations; therefore, the heavy atoms of this ring were fixed after initial minimization with the PEF95SAC[††] force field.[26] For all other linkages that were allowed to rotate, a bias around 30° was selected; therefore, the dihedral angles were rotated randomly with a sum-of-gaussians distribution with peaks at multiples of 30°. To make the search even more efficient, a simultaneous Cartesian perturbation of 0.4 Å was also applied. These random perturbations were repeated 250,000 times per conformational search and the structures were subsequently optimized using the PEF95SAC force field using implicit solvation (GB/SA). Two separate searches were performed and the databases were combined for analysis. Initially each new conformation was checked against the previously found conformations using a root-mean-square (RMS) tolerance of 0.1 Å on all atoms, including the hydrogen atoms to account for possible hydrogen bonding. The second search was carried out using a more stringent RMS value (0.01 Å) applied to the heavy atoms only. In both searches the calculations terminated at the end of the 250,000 iterations as the number of failures criteria to find new conformations (10,000 iterations in a row) was never met for this large conformational search. Structures that did not have the $^4C_1$ conformation for D-glucosamine were eliminated (see Results and discussion). Conformations with energies over 10 kcal/mol above the global minima were rejected resulting in 25,852 structures.

The combined database was also re-minimized using the AMBER94 force field[17] with implicit solvation (GB/SA), which selected 10,656 structures within 10 kcal/mol of the AMBER94 determined global minimum. The structures within 3 kcal/mol of the global minima for each force field were then further analyzed for relationships between the ϕ and ψ angles: 1860 and 1067 conformations for the PEF95SAC and AMBER94 force fields, respectively.

3. Results and discussion

MSH is a clinically important glycothiol believed to be essential for *Mycobacteria's* survival however there is very little known regarding its biochemistry and conformational preferences.[1,27] Information regarding the conformational preferences of MSH may be useful in the design of novel inhibitors aimed at MSH-utilizing enzymes. MCA, for example, is an attractive target for the development of inhibitors as it is believed to be involved in the removal of electrophilic species from *Mycobacteria*, and may be involved in the detoxification of thiol reactive antibiotics, such as rifamycin S and cerulenin.[28] The design and synthesis of MSH-utilizing enzyme inhibitors is ongoing, with some of these being synthetic MSH analogues.[29–32] The results of our modelling study may aid in the design of inhibitors by identifying key dihedral angles in MSH, which are highly populated and found in the solution.

MSH is an unusual thiol in that it contains a pseudo-disaccharide moiety coupled to cysteine (Fig. 1, **3**), this results in a wider variety of atom and bond types within MSH, complicating force field selection, as many force fields are not parameterized for all of the MSH bond and atom types. This parameterization is critical for accurate energetic information. MacroModel returns information regarding the quality (high, medium, or low) of the parameters in use for each structure (Table 1). The lower the quality of the parameter, the less experimental or ab initio data that have been utilized in fitting the force field equation. A low quality rating indicates that generalized parameters are in use, which can lead to inaccurate conformational energy differences

[††] AMBER94 and PEF95SAC were used as implemented by MOE2004.

and geometries. The use of low quality stretch parameters, in particular, can give crude partial charges, which in turn can cause computed charges and solvation energies to be inaccurate.[33] AMBER*, a well known force field used for modelling carbohydrates, returned a medium bend parameter for the bend of the O_{ring}–C1–$O_{linkage}$ angle and medium torsion parameters for the cysteine portion of MSH (Table 1). Based upon the quality of the parameters returned by MacroModel (Table 1), the OPLS-AA force field appeared to be best suited for modelling MSH (Table 1) and therefore was used as the main force field for the MacroModel search. The OPLS-AA force field has traditionally been used to model amino acids, but it has also been parameterized for,[34] and successfully used with carbohydrates including disaccharides[35–37] and cyclodextrin.[38] In addition, as implemented by MacroModel, OPLS-AA is parameterized for sulfur in thiols.[22]

For a conformational study on a complex molecule such as MSH to be undertaken, various conformations must be generated that explore all or at least a substantial portion of the conformational space of the molecule. One technique to examine torsional angles is the dihedral angle drive or adiabatic mapping method;[39] however, due to the complexity of MSH, the number of angles to be probed appeared prohibitive. If one ignored the numerous dihedral angles in the disaccharide rings themselves, there would remain nine dihedral angles to evaluate, assuming only heavy atoms were considered. If these dihedral angles were modified in 20° increments this would result in 18^9 or over 198 billion possible starting conformations to examine, which would become a major computational undertaking. More tractable approaches are those of stochastic or MCMM methods, which utilize random structural perturbations to generate starting conformers for minimization. If a sufficient number of conformers are generated, this should sample a large portion of the conformational space available to the molecule and hence was chosen for our MacroModel conformational search. This method has been useful in many studies including the study of the coil-to-helix transition in polyalanine,[40] and the determination of the low energy conformations of synthesized carbohydrates.[41]

During the MCMM conformational search using the OPLS-AA force field, the rings were opened, followed by torsional angle variation and ring closure. This yielded an *unbiased* search of the ring conformations. The results indicated that all of the structures within 3 kcal/mol of the implicitly solvated global minimum had a $^4C_1$ orientation for the D-glucosamine moiety; this is consistent with determined crystal[42] and aqueous phase[43] structures of *N*-acetyl-D-glucosamine. Structures having *myo*-inositol in the one equatorial/five axial (1eq/5ax) conformation were eliminated (50 structures within 3 kcal/mol of the global minimum), as it is known from solution phase NMR[44] and X-ray crystallography[45] that *myo*-inositol prefers the 5eq/1ax conformation.

MSH was modelled in the gas phase as well as with implicit water solvation using the GB/SA model, which has been shown to be efficient for the modelling of carbohydrates,[46] and has been successfully applied to conformational analyses of ganglioside head groups[47] and mannobiosides and triosides.[48] The overlay of the global minima of MSH in the gas phase and in implicit water demonstrates the effect that solvent shielding can cause (Fig. 2) as the importance of hydrogen bonding and electrostatics can be exaggerated in the gas phase and folded or collapsed structures become more probable.[49]

Figure 2. Stereoview overlay of the global minimum conformations of MSH found for the gas (coral) and implicit water (cyan) phases using the OPLS-AA force field.

Table 2. Hydrogen-bonding interactions for the global minimum conformations of MSH found for the gas phase and with implicit water solvation (OPLS-AA)

Phase	No. of intramolecular H-bonds	Atoms involved[a]
Gas	3	C6″=O···H–O–C2′ C6″=O···H–N1″ C2″=O···H–O–C6
Water	2	C2″=O···H–S C6″=O···H–N1″

[a] See Figure 1 for numbering.

As can be seen in Table 2, the gas phase and implicitly solvated global minimum structures differ in hydrogen bond placement with the gas phase having a more closed structure with the amino acid portion hydrogen bonding to the disaccharide (Fig. 2).

When analyzing the results of a conformational study, it is important to note that while the global minimum structure may be the most energetically favourable structure, it may not necessarily be the structure required for biological activity and protein binding.[50] Studies have indicated that the conformations of bound ligands may be those within 3 kcal/mol of the global minimum structure.[49] In the case of highly flexible ligands, the upper energy limit for binding conformers has been suggested to be as high as 5 kcal/mol above the global minimum.[51] While MSH has a high degree of flexibility in the amino acid portion, there is much less flexibility around the disaccharide linkage of D-glucosamine and *myo*-inositol; hence a 3 kcal/mol cut off may be sufficient for MSH and only those conformers within 3 kcal/mol of the implicitly solvated global minimum were studied further.

Clustering analysis was performed using XCluster, which has been utilized to study a wide range of systems including antibiotic mutacin 1140,[52] thrombin inhibitors,[53] agouti-related peptides[54] and peptide-tannin systems.[55] Clustering of the implicit water OPLS-AA data, using ϕ (H–C1–O–C1′) and ψ (C1–O–C1′–H) as the angles of interest, indicated the presence of three distinct clusters (Fig. 3 and Table 3). An overlay of local minimum structures from each cluster highlights the differences between these angles (Fig. 4). Cluster 1 is the most highly populated, with 131 members and contains the OPLS-AA implicitly solvated global minimum structure, while Clusters 2 and 3 contain 13 and 4 members, respectively (Fig. 3 and Table 3).

While OPLS-AA was the initial choice for modelling MSH, several additional force fields known for modelling carbohydrates[46] AMBER*, AMBER94 and PEF95SAC, were used to detect conformational modelling inadequacies in the OPLS-AA results. Of the 148 structures kept from the OPLS-AA conformational search very little variance in their ring torsional angles was observed (data not shown). The subsequent AMBER* search therefore froze the internal torsion angles of the D-glucosamine and *myo*-inositol rings after an initial minimization. The AMBER94 and PEF95SAC conformational searches did not hold the D-glucosamine ring of MSH rigid, which allowed for ring flipping. Any conformations not having the $^4C_1$ D-glucosamine orientation were eliminated from the databases as it has been shown that *N*-acetyl-D-glucosamine prefers the $^4C_1$ orientation.[42,43]

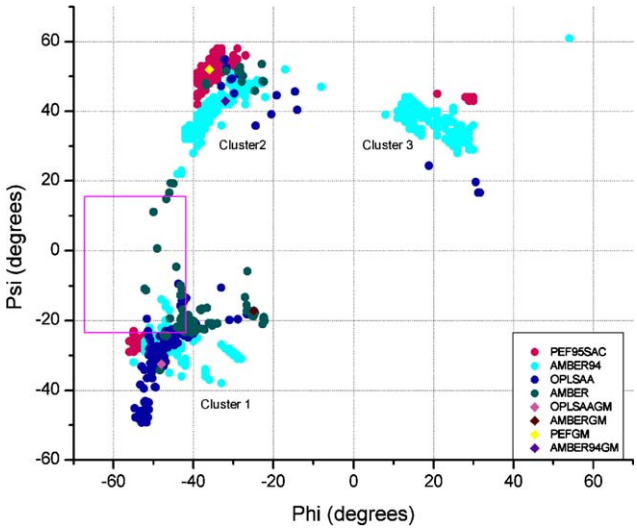

Figure 3. Clustering of the ϕ and ψ angles of MSH conformers found to be within 3 kcal/mol of their respective global minima. These conformers were generated using the OPLS-AA, AMBER*, AMBER94 and PEF95SAC force fields. The purple box represents the ϕ and ψ angle ranges of MSH-*S*-bimane as determined by NMR data.[27] There are two conformers found using AMBER*, which are not shown in this figure with $(\phi,\psi) = (-27°, -155°)$ and $(-28°, -155°)$.

The data obtained using the random torsion method combined with Cartesian perturbation, implemented in the MOE stochastic search, was in good agreement with the MCMM results, indicating that both methods adequately covered the conformational space of MSH. There was some variation among the force fields with respect to their cluster populations (Table 3). As mentioned earlier, the majority of the OPLS-AA derived structures are found in Cluster 1, including the global minimum. In the case of AMBER*, the majority of the structures, including the global minimum, were found in Cluster 1 (101 conformers), with only 10 conformers being found in Cluster 2 and no conformers were found with the ϕ and ψ angle combination of Cluster 3. There were some outlying AMBER* structures that belonged to Cluster 1 with a ϕ angle of ~50° and ψ angles ranging from ~−20° to 20°, as well as two conformers with (ϕ, ψ) = (−27°, −155°) and (−28°, −155°). These structures are not seen with any other force field and may be due to the medium quality parameter seen for the O_{ring}–C1–$O_{linkage}$ angle, which may allow for additional flexibility of that angle, affecting the overall ϕ and ψ torsional angles. The AMBER94 data had a more even distribution

Table 3. Comparison of the three clusters and global minima of the implicitly solvated MSH conformers generated in this study

	Cluster 1	Cluster 2	Cluster 3	GM[a]
OPLS-AA[b] *(MacroModel)*				
ϕ Dihedral	−55° to −26°	−33° to −14°	19° to 32°	−48°
ψ Dihedral	−49° to −10°	36° to 55°	17° to 24°	−33°
No. of members	131	13	4	—
AMBER*[c,d] *(MacroModel)*				
ϕ Dihedral	−52° to −22°	−37° to −23°	—	−25°
ψ Dihedral	−34° to 19°	46° to 54	—	−17°
No. of members	101	13		
AMBER94[e] *(MOE2004)*				
ϕ Dihedral	−55° to −28°	−44° to −8°[f]	8° to 30°[g]	−31°
ψ Dihedral	−38° to −13°	22° to 51°	28° to 43°[g]	43°
No. of members	391	390	285	—
PEF95SAC[e] *(MOE2004)*				
ϕ Dihedral	−56° to −47°	−39° to −27°	20° to 30°	−36°
ψ Dihedral	−29° to −23°	42° to 58°	43° to 45°	52°
No. of members	66	1782	12	—
Mycothiol-S-bimane (NMR data)[h]				
ϕ Dihedral	−66° to −42°	—	—	—
ψ Dihedral	−22° to 18°	—	—	—

[a] Implicitly solvated global minimum.
[b] As implemented by MacroModel v. 8.1.
[c] As implemented by MacroModel v. 9.0.
[d] Two outlying angle combinations were found with $(\phi,\psi) = (-27°, -155°)$ and $(-28°, 155°)$, no angle combinations corresponding to Cluster 3 were found.
[e] As implemented by MOE2004.
[f] Only two conformers were found to have ϕ angles of −8°, all others in this cluster were below −17°.
[g] One outlying conformer was determined with $(\phi,\psi) = (54°, 61°)$.
[h] Angle ranges suggested by NOE data reported for mycothiol-*S*-bimane.[27]

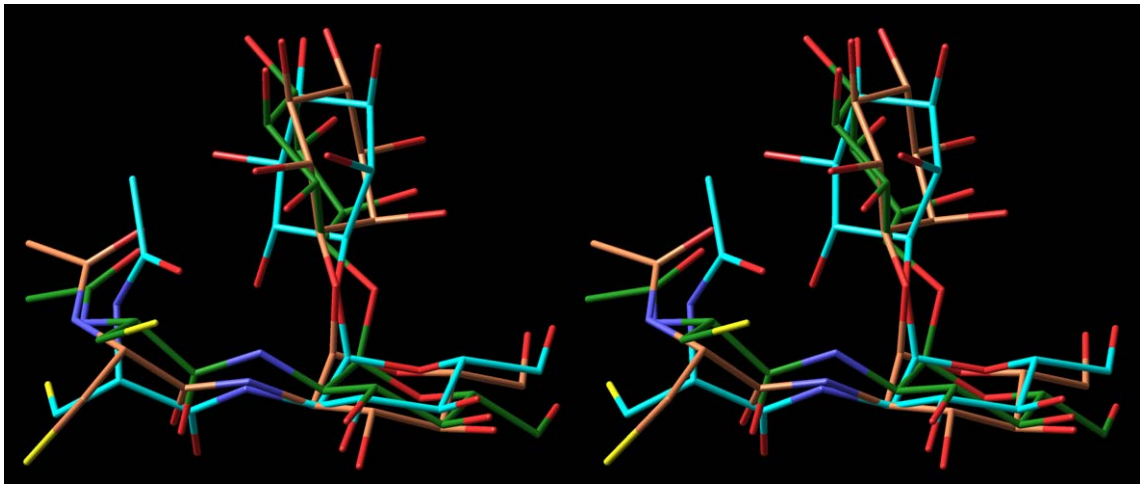

Figure 4. Stereoview of the superimposition of the lowest energy conformers from each of the MSH clusters in implicit water determined using the OPLS-AA force field: cyan—Cluster 1 (also the global minimum); coral—Cluster 2; green—Cluster 3.

between the clusters with the distribution slightly favouring those of Clusters 1 and 2. In contrast, Cluster 2 from the PEF95SAC data had 1782 members with Clusters 1 and 3 having populations of 66 and 12, respectively. Unlike the OPLS-AA and AMBER* data, the global minima of AMBER94 and PEF95SAC were found in Cluster 2. While the energetics of the structures differed among the various force fields, the overall ϕ and ψ angle ranges were in agreement among the four force fields (Fig. 3 and Table 3) indicating the conformations obtained for MSH are independent of the force field chosen. Based upon the population densities, there is some flexibility around the ϕ and ψ angles, however MSH will most frequently be found with ϕ and ψ angle

combinations consistent with conformations in Clusters 1 and 2.

As previously mentioned, a crystal structure of MSH is not available for comparison to our calculated values. However, a solution phase structure of a bimane derivative of MSH (MSmB, Fig. 5, **4**), has recently been determined using NMR techniques.[27] MSmB is a substrate for MCA, the enzyme responsible for the cleavage of MSH-*S*-conjugates formed from MSH and various electrophiles.[28] The structure of MSmB was therefore used as a comparison against the implicitly solvated MSH conformers found in this study. The local minimum structures (C1–C3) of the three clusters, as found by OPLS-AA (seen in Fig. 4), were used as representative structures from Clusters 1–3; in the case of Cluster 1, the local minimum was also the OPLS-AA global minimum structure.

A comparison of the cysteinyl torsional angles (Table 4) reported for the MSmB are in good agreement with those measured from C1. The MSmB structure was determined using key inter-hydrogen distances obtained from NOE data.[27] A comparison of these distance ranges to the distances observed for the implicitly solvated representative structures C1–C3 are shown in Table 5. An initial comparison of the distances found for the three representative structures indicates that C2 and C3 are more similar to each other than to C1. The similarity in structure can be seen in Figure 3 where the *myo*-inositol ring of C1 (cyan) is oriented away from the *myo*-inositol rings of C2 (coral) and C3 (green).

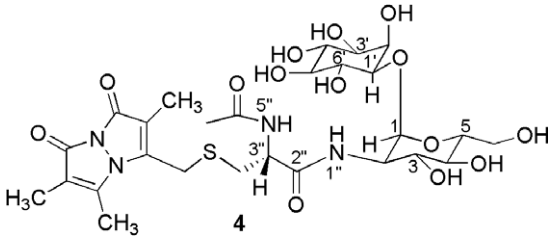

Figure 5. MSmB, a substrate for MSH-*S*-conjugate amidase.

Table 4. Comparison of selected cysteinyl torsional angles from implicitly solvated MSH (OPLS-AA) and MSmB (NMR)

Torsion angles	NMR data[a]	GM (aq)[b]
H–C3″–C2″–O	Trans	125.3°
O–C2″–N1″–H	Trans	174.0°
H–N5″–C3″–H	−150°	−140.1

[a] Mahadevan et al. (2003).[27]
[b] Global minimum.

The representative structures from all three clusters satisfy the inter-hydrogen distance ranges determined by NMR[27] for H1–1′, H1″–3″, H5″–3″, H5–1′ and CH_3–H5″. However, only C1 satisfies the distance ranges reported for H1–6′, H5–2′ and CH_3–H6′ as the distances for C2 and C3 are outside of the ranges reported for the NOE determined conformations. In particular, the CH_3 group of the acetyl group is distal from the protons of the *myo*-inositiol ring; this can be clearly seen in Figure 3 (C1: cyan, C2: coral, C3: green) and Table 5. However, C2 and C3 satisfy the distance range reported for H1–2′. The orientation of the *myo*-inositol ring brings H6′ close to H1 in C1 and places H2′ away from H1, such that it cannot satisfy the H1–2′ NOE predicted distance range; the reverse is true for C2 and C3. Thus, these results indicate that in solution MSH does not exist as a static structure, but is likely conformationally flexible, leading to variable distances between the various protons during the time frame of an NMR experiment. This is supported by the inability of one representative structure to satisfy all of the distances determined using NOE experiments. This is also supported by a comparison of the ϕ and ψ angles calculated in this study and those determined by Mahadevan et al.[27] There is some overlap between the NOE predicted angle ranges and those found in Cluster 1, as seen in Figure 3 and Table 5.

The structure predicted by the NOE data is best fit to the calculated distance and angle data as an average of structures; this is consistent with the time scale of the NMR experiment. Structures determined using NMR are usually a composite of conformations that can exist in solution and do not necessarily represent a single conformation.[56]

Table 5. Comparison of key inter-hydrogen distances for implicitly solvated MSH conformers (OPLS-AA) to those of MSmB (NMR)

Structure	Atoms[a]								
	H1–1′	H1–2′	H1–6′	H5–1′	H5–2′	H1″–3″	H5″–3″	CH_3–H6′[b]	CH_3–H5″[b]
NOE[c]	S[d]	M[d]	M	W[d]	S	S′[d]	W′	W″[d]	S″
C1[e]	2.6	4.4	3.4	3.2	2.7	2.5	2.9	3.1	3.0
C2	2.3	2.7	4.5	4.3	3.6	2.2	2.9	7.1[e]	2.9
C3	2.3	2.7	4.5	4.6	4.7	2.2	2.9	7.8	2.9

[a] See Figure 1 for numbering.
[b] The CH_3 distance was obtained by averaging the distance from each H of the CH_3 group to either H6′ or H5″.
[c] Mahadevan et al.[27]
[d] S: strong (1.8–3.0 Å) M: medium (1.8–4.0 Å), W: weak (1.8–5.0 Å), S′: (1.8–3.2 Å), W′: (1.8–5.2 Å), S″: (1.8–3.5 Å), W″: (1.8–5.5 Å).
[e] Representative structures from Clusters 1 (C1), 2 (C2) and 3 (C3). These are the implicitly solvated local minima from each cluster determined using the OPLS-AA force field.

4. Conclusions

A detailed conformational investigation was performed on the novel intracellular glycothiol MSH (**3**), which is found solely in the *Actinomycetales* bacteria, including medically important *M. tuberculosis*. The OPLS-AA force field was used for MCMM calculations performed on MSH in vacuo and with implicit water solvation. The global minimum structures of MSH in the two phases were found to be similar; however, intramolecular hydrogen-bonding played a greater role in the conformation of MSH in vacuo than in implicit water; this is believed to be due to the lack of shielding and competing interactions present when MSH is solvated in water.

To detect any inadequacies in the OPLS-AA results, MSH was also modelled using AMBER* and the MCMM method, and using AMBER94 and PEF95SAC and a stochastic conformational search method. Conformers within 3 kcal/mol of the respective global minima were compared and it was found that these conformations were independent of the force fields used. All conformers were clustered into three distinct families based upon the ϕ and ψ angles of the disaccharide. Based upon the population of these clusters it is expected that MSH will be found predominantly in Clusters 1 and 2. Cluster 1 has ϕ and ψ angle ranges of roughly $-55°$ to $-26°$ and $-49°$ to $-10°$, while Cluster 2 has ϕ and ψ angles ranges of approximately $-44°$ to $-14°$ and 22–58°. This data compares favourably to a previous NMR conformational study on MSmB, an MSH-*S*-conjugate containing a bimane moiety. The knowledge of the structures of these possible conformers of MSH should be important in the future design and optimization of inhibitors of MSH-utilizing enzymes.

Acknowledgements

The authors wish to thank the Natural Sciences and Engineering Research Council of Canada (J.F.H. and F.-I.A.), the University of Waterloo (C.E.H. and J.F.H.) and the University of Guelph (F.-I.A.) for funding. The authors gratefully acknowledge the Canada Foundation for Innovation and the Ontario Innovation Trust for funding of the software and computer systems utilized in this study. The technical assistance of Mr. Robyn Landers (University of Waterloo, Flexor) was especially appreciated.

References

1. Hand, C. E.; Honek, J. F. *J. Nat. Prod.* **2005**, *62*, 293–308, and references cited therein.
2. Fahey, R. C. *Annu. Rev. Microbiol.* **2001**, *55*, 333–356, and references cited therein.
3. Newton, G. L.; Fahey, R. C. *Meth. Enzymol.* **1995**, *251*, 148–166.
4. Newton, G. L.; Javor, B. *J. Bacteriol.* **1985**, *161*, 438–441.
5. Fairlamb, A. H.; Blackburn, P.; Ulrich, P.; Chait, B. T.; Cerami, A. *Science* **1985**, *227*, 1485–1487.
6. Newton, G. L.; Arnold, K.; Price, M. S.; Sherrill, C.; Delcardayre, S. B.; Aharonowitz, Y.; Cohen, G.; Davies, J.; Fahey, R. C.; Davis, C. *J. Bacteriol.* **1996**, *178*, 1990–1995.
7. World Health Organization. Tuberculosis: www.who.int/mediacentre/factsheets/fs104/en/print.html.
8. Food and Agriculture Organization of the United Nations. Tuberculosis: www.fao.org/ag/againfo/subjects/en/health/diseases-cards/cards/tuberculosis.html.
9. Rice, D. N.; Rogers, D. G. Johne's Disease (Paratuberculosis): http://ianrpubs.unl.edu/animaldisease/g977.htm, University of Nebraska at Lincoln Institute of Agriculture and Natural Resources.
10. Newton, G. L.; Bewley, C. A.; Dwyer, T. J.; Horn, R.; Aharonowitz, Y.; Cohen, G.; Davies, J.; Faulkner, D. J.; Fahey, R. C. *Eur. J. Biochem.* **1995**, *230*, 821–825.
11. Rawat, M.; Newton, G. L.; Ko, M.; Martinez, G. J.; Fahey, R. C.; Av-Gay, Y. *Antimicrob. Agents Chemother.* **2002**, *46*, 3348–3355.
12. Jardine, M. A.; Spies, H. S.; Nkambule, C. M.; Gammon, D. W.; Steenkamp, D. J. *Bioorg. Med. Chem.* **2002**, *10*, 875–881.
13. Lee, S.; Rosazza, J. P. *Org. Lett.* **2004**, *6*, 365–368.
14. Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005–3014.
15. Polak, E.; Ribiere, G. *Rev. Fr. Inform. Rech. Operation., Ser. Rouge* **1969**, *16*, 35–37.
16. McDonald, D. Q.; Still, W. C. *Tetrahedron Lett.* **1992**, *33*, 7743–7746.
17. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
18. Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127–8134.
19. Allinger, N. L.; Yuh, Y. H.; Lii, J. H. *J. Am. Chem. Soc.* **1989**, *111*, 8551–8566.
20. Halgren, T. A. *J. Comput. Chem.* **1999**, *20*, 730–748.
21. Jorgensen, W. L.; Tiradorives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
22. Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474–6487.
23. Nicholas, G. M.; Kovác, P.; Bewley, C. A. *J. Am. Chem. Soc.* **2002**, *124*, 3492–3493.
24. Chang, G.; Guida, W. C.; Still, W. C. *J. Am. Chem. Soc.* **1989**, *111*, 4379–4386.
25. Ferguson, D. M.; Raber, D. J. *J. Am. Chem. Soc.* **1989**, *111*, 4371–4378.
26. Fabricius, J.; Engelsen, S. B.; Rasmussen, K. *J. Carbohydr. Chem.* **1997**, *16*, 751–772.
27. Mahadevan, J.; Nicholas, G. M.; Bewley, C. A. *J. Org. Chem.* **2003**, *68*, 3380–3386.
28. Steffek, M.; Newton, G. L.; Av-Gay, Y.; Fahey, R. C. *Biochemistry* **2003**, *42*, 12067–12076.
29. Fetterolf, B.; Bewley, C. A. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3785–3788.
30. Nicholas, G. M.; Eckman, L. L.; Newton, G. L.; Fahey, R. C.; Ray, S.; Bewley, C. A. *Bioorg. Med. Chem.* **2003**, *11*, 601–608.
31. Knapp, S.; Gonzalez, S.; Myers, D. S.; Eckman, L. L.; Bewley, C. A. *Org. Lett.* **2002**, *4*, 4337–4339.

32. Knapp, S.; Amorelli, B.; Darout, E.; Ventocilla, C. C.; Goldman, L. M.; Huhn, R. A.; Minnihan, E. C. *J. Carbohydr. Chem.* **2005**, *24*, 103–130.
33. *MacroModel 9.0 User Manual*; Schrodinger, 2005.
34. Damm, W.; Frontera, S.; Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1997**, *18*, 1955–1970.
35. Vishnyakov, A.; Widmalm, G.; Kowalewski, J.; Laaksonen, A. *J. Am. Chem. Soc.* **1999**, *121*, 5403–5412.
36. Conrad, P. B.; de Pablo, J. J. *J. Phys. Chem. A* **1999**, *103*, 4049–4055.
37. Umemura, M.; Hayashi, S.; Nakagawa, T.; Urakawa, H.; Kajiwara, K. *J. Mol. Struct. (THEOCHEM)* **2003**, *636*, 215–228.
38. Starikov, E. B.; Bräsicke, K.; Knapp, E. W.; Saenger, W. *Chem. Phys. Lett.* **2001**, *336*, 504–510.
39. Leach, A. R. *Molecular Modelling, Principles and Applications*; Addison Wesley Longman Ltd: Harlow, Essex, England, 1996.
40. van Giessen, A. E.; Straub, J. E. *J. Chem. Phys.* **2005**, *122*, 024904.
41. DeMatteo, M. P.; Snyder, N. L.; Morton, M.; Baldisseri, D. M.; Hadad, C. M.; Peczuh, M. W. *J. Org. Chem.* **2005**, *70*, 24–38.
42. Johnson, L. N. *Acta Crystallogr.* **1966**, *21*, 885–891.
43. Boyd, J.; Porteous, R.; Soffe, N.; Delepierre, M. *Carbohydr. Res.* **1985**, *139*, 35–46.
44. Abraham, R. J.; Byrne, J. J.; Griffiths, L.; Koniotou, R. *Magn. Reson. Chem.* **2005**, *43*, 611–624.
45. Rabinowitz, H. N.; Kraut, J. *Acta Crystallogr.* **1964**, *17*, 159–168.
46. Perez, S.; Imberty, A.; Engelsen, S. B.; Gruza, J.; Mazeau, K.; Jimenez-Barbero, J.; Poveda, A.; Espinosa, J. F.; van Eyck, B. P.; Johnson, G.; French, A. D.; Louise, M.; Kouwijzer, C. E.; Grootenuis, P. D. J.; Bernardi, A.; Raimondi, L.; Senderowitz, H.; Durier, V.; Vergoten, G.; Rasmussen, K. *Carbohydr. Res.* **1998**, *314*, 141–155.
47. Brocca, P.; Bernardi, A.; Raimondi, L.; Sonnino, S. *Glycoconjugate J.* **2000**, *17*, 283–299.
48. Bernardi, A.; Colombo, A.; Sánchez-Medina, I. *Carbohydr. Res.* **2004**, *339*, 967–973.
49. Boström, J.; Norrby, P.-O.; Liljefors, T. *J. Comput. Aided Mol. Des.* **1998**, *12*, 383–396.
50. Nicklaus, M. C.; Wang, S.; Driscoll, J. S.; Milne, G. W. A. *Bioorg. Med. Chem.* **1995**, *3*, 411–428.
51. Perola, E.; Charifson, P. S. *J. Med. Chem.* **2004**, *47*, 2499–2510.
52. Smith, L.; Azachariah, C.; Thirumoorthy, R.; Rocca, Jim; Novák, J.; Hillman, J. D.; Edison, A. S. *Biochemistry* **2003**, *42*, 10372–10384.
53. Greenidge, P. A.; Mérette, S. A. M.; Beck, R.; Dodson, G.; Goodwin, C.; Scully, M. F.; Spencer, J.; Wiser, J.; Deadman, J. J. *J. Med. Chem.* **2003**, *46*, 1293–1305.
54. Thirumoorthy, R.; Holder, J. R.; Bauzo, R. M.; Richards, N. G. J.; Edison, A. S.; Haskell-Luevano, C. *J. Med. Chem.* **2001**, *44*, 4114–4124.
55. Simon, C.; Barathieu, K.; Laguerre, M.; Schmitter, J.-M.; Fouquet, E.; Pianet, I.; Dufourc, E. J. *Biochemistry* **2003**, *42*, 10385–10395.
56. Neuhaus, D.; Williamson, M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, 2nd ed.; John Wiley and Sons: Toronto, 2000.