The Dissertation Committee for Jamie Nicole Jones certifies that this is the approved version of the following dissertation:

Group 14 Metallocenes and Structure—Function Relationships in

Metalloenzymes

Committee:

_____
Alan H. Cowley, Supervisor

_____
Alan Campion

_____
John G. Ekerdt

_____
Richard A. Jones

_____
Richard J. Lagow

Group 14 Metallocenes and Structure—Function Relationships in Metalloenzymes

by

Jamie Nicole Jones, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August, 2004

To My Mother

For teaching me to multiply using dried beans

For your patience when my science experiment ruined the dining room table

For showing me how to respect myself and others

For instilling in me the value of education

But most importantly, for your endless love and support

Acknowledgements

Without friends no one would choose to live, thought he had all other goods --Aristotle

First and formost, I would like to thank Professor Alan Cowley for his insight, patience, and instruction. His enthusiasm for chemistry is inspirational and I can only hope that I too can demonstrate the same level of zeal throughout my career. The Cowley group always seems to have a plethora of talented, energetic, and beer-drinking scientists in and out of the lab. Whether you played a direct role in the work presented here or were simply there for conversation and/or support, I thank you, Dr. John Gorden, Dr. Robert Wiacek, Dr. Jeffery Pietryga, Silvia Filliponi, Lucy Mullins, Zheng Lu, Dragoslav Vidovic, Christopher Entwistle, and Michael Findlater. When I entered graduate school I had the distinct privilege of being mentored by Dr. Charles Macdonald, whose work ethic and love of chemistry was contaigous. Dr. Piyush Shukla continues to be one of my biggest supporters and always there whenever I need a kind word. One person has played a key role in helping me maintain my sanity, but just as often contributed to my insanity, Jennifer Moore; thanks for pushing me to swim, entertaining me while we run, and yelling at me to bike.

I would like to thank the Dr. Katherine A. Brown and the Brown group for providing me the opportunity to transition my career path from organometallic

synthesis to molecular biology. Special thanks goes to Dr. Nigel Eady who patiently beat biochemistry into this strictly organometallic-thinking brain of mine and Dr. Karen Keith who made certain I always had a place to vent and a full pint of beer.

To all the people outside of the Cowley lab who have motivated, inspired, and provided hours of entertaining conversation, thanks Jessica White, Monica Collins, Sabin Davis, Justine Miller, Jill Lang, and Josh and Valerie Watson. Dr. William Murray must have the patience of a saint, as he tolerates my attention deficit disorder with grace and continues to love me just the same.

Finally, I want to thank the people who made certain I had a good southern upbringing, my family. For instilling in me at a very early age the value of an education, I would like to thank my Aunt Ruth. For your endless love and support, thanks Mom, Dad, and Sara, I would not be receiving this degree without all the encouragement you have provided me.

Group 14 Metallocenes and Structure—Function Relationships in

Metalloenzymes

Publication No. _____

Jamie Nicole Jones, Ph.D.

The University of Texas at Austin, 2004

Supervisor: Alan. H. Cowley

A general synthetic method has been developed for the synthesis of post-transition metal, triple-decker cations. The addition of positively charged half-sandwich cations to neutral metallocenes affords triple-decker cations of the formula, $[\eta^5\text{-}(C_5Me_5)_3E_2]^+$. Examples of both group 13 and group 14 complexes have been prepared and structurally characterized. These triple-decker cations provide a platform, which shows promise for the development of nanomaterials, such as molecular wires. Group 14 metallocene complexes containing the 1,3-bis(trimethylsilyl)indenide ligand have also been synthesized and structurally

characterized.   Rational syntheses for complexes of the formulas [1,3-(Me_3Si)$_2$(η^5-C_9H_5)]$_2$E   and   [1,3-(Me_3Si)$_2$(η^5-C_9H_5)](η^5-C_5Me_5)E   have   been developed. The lability of the indenide ligand renders these complexes promising potential precursors for group 14 nanoparticle formation.

The merger of organometallic chemistry with molecular biology to provide creative solutions to difficult problems opened an opportunity to apply the background I have acquired in structure chemistry to the study of three metalloenzymes, 3-dehydroquinate synthase, 5-*enol*pyruvylshikimate 3-phosphate synthase, and *Mycobacterium tuberculosis* catalase-peroxidase (*M. tuberculosis* CP). Computational analsysis of the crystal structure of each system provided valuable insight into residues that play a key role in the activity of the enzymes. This knowledge can in turn be applied to the development of the small-molecule inhibitors of the enzymes.   In addition to computational analysis of *M. tuberculosis* CP, an X-ray crystal structure of this enzyme was determined to a resolution of 2.4 Å.

Table of Contents

CHAPTER 1: GROUP 13 AND 14 MULTI-DECKER CATIONS

INTRODUCTION

The pentamethylcyclopentadienide ligand, $[C_5Me_5]^-$, is one of the most widely used bulky substituents for the stabilization of low valent and electron-deficient species. By accommodating a wide range of bonding modes, the $[C_5Me_5]^-$ ligand can effectively adjust to its electronic environment.[1] This flexibility in bonding modes congruently with the bulky nature of the σ- or π-bound C_5Me_5 moiety permits the kinetic and thermodynamic stabilization of otherwise highly reactive species.[1] Transition and lanthanide metal complexes featuring the $[C_5Me_5]^-$ ligand have been investigated extensively; however, it was not until recently that analogous complexes of the main group elements have been studied.[2, 3]

Since coinage of the term in 1952, metallocenes have typically been associated with (bis-cyclopentadienyl)metal derivatives; however the scope of the definition has expanded subsequently to include half-sandwich, multi-decker, and polymeric cyclopentadiene complexes. The $[C_5Me_5]^-$ ligand exhibits fluxionality and a tendency for ring slippage. There are six plausible bonding modes for the C_5Me_5 ligand, which are shown in Figure 1.1. Due to the hapticity changes, the

C_5Me_5 ligand is able to adjust its effective electron contribution to the element center.
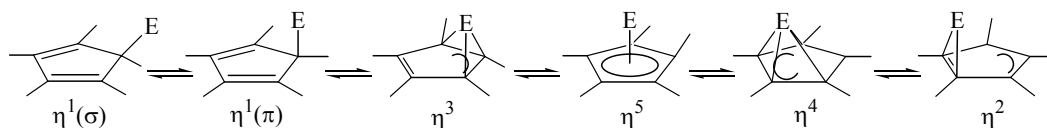
Figure 1.1 The six possible bonding modes for C₅Me₅.

In addition to their tendency toward ring slippage, the cyclopentadienyl family of ligands displays a facile migratory behavior. NMR spectroscopic studies have evidenced the equivalence of all proton and all carbon and methyl carbon resonances for many C_5Me_5 species. The migratory process is accomplished by a series of 1,2-sigmatropic shifts as illustrated in Figure 1.2.[1]
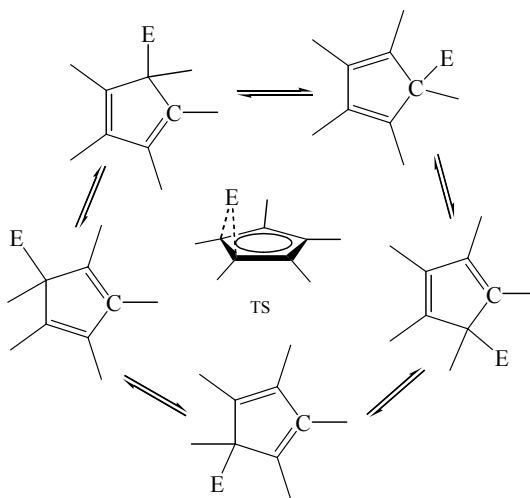
Figure 1.2 Sigmatropic rearrangements in C₅Me₅—E compounds where C designates a C-atom in the C₅Me₅ ring.

The π-bonding mode of the $[C_5Me_5]^-$ ligand can be utilized to create metallocenes that are similar in structure to ferrocene, $Fe(\eta^5-C_5H_5)_2$.[4-7] Several pentamethylcyclopentadienide complexes of the p-block elements have been structurally characterized including those of $Sn(\eta^5-C_5Me_5)_2$[8] and $Pb(\eta^5-C_5Me_5)_2$.[9] In each case, the $[C_5Me_5]^-$ rings of these +II oxidation state group 14 metallocenes are not coplanar due to the presence of a lone pair on the heteroatom. The presence of the lone pair causes a rather large bend in the ring centroid—metal—ring centroid bond angles: ($154.9°$ in $Sn(\eta^5-C_5Me_5)_2$[8] and $151°$ in $Pb(\eta^5-C_5Me_5)_2$[9]).

Since the discovery of the first simple sandwich complex, ferrocene[4-7] in 1952, an intense interest has been generated in the syntheses, structures, and properties of multi-decker sandwich compounds. The first genuine triple-decker complex, $[(\eta^5-C_5H_5)_3Ni_2]^+$, was synthesized by Werner and Salzer in 1972.[10] Several transition metal multi-decker complexes have since been synthesized since that time including $[(\eta^5-C_5Me_5)_3Co_2]^+$,[11] $[(\eta^5-C_5Me_5)_3Ru_2]^+$,[12] and $[(\eta^5-C_5H_5)Mo(CO)_3(\mu-\eta^5-C_5H_5)Fe(CO)_2(\eta^5-C_5H_5)Mn(CO)_3]^+$.[13] Until recently only two main group multi-decker species had been characterized, namely $[(\eta^5-C_5H_5)_3Tl_2]^-$ [14] and $[(\eta^5-C_5H_5)_3Cs_2]^-$.[15] Both anions have a trans-geometry with respect to the terminal cyclopentadienide groups (Figure 1.3).
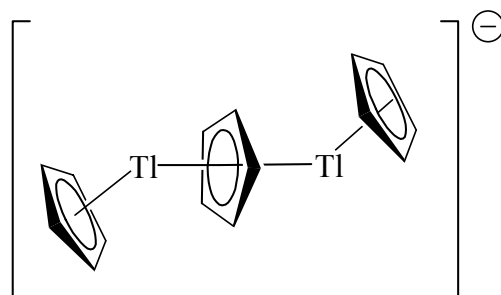
Figure 1.3   The [(η^5-C$_5$H$_5$)$_3$Tl$_2$]$^-$ [14] **and [(η^5-C$_5$H$_5$)$_3$Cs$_2$]$^-$** [15] **anions have a trans-geometry with respect to the terminal cyclopentadienide groups.**

An upsurge of interest in the chemistry of main group metallocenes has occurred over the past decade.[3]   Apart from the fact that much less is known about s- and p-block metallocenes than their d- and f-block counterparts, interest in the main group metallocenes has been stimulated by structure and bonding considerations,[3] their utility as reagents and chemical vapor desposition sources,[16] and the possibility of serving as useful catalysts for alkene polyerizations.[17, 18]   Additionally, the use of post transition metal compounds as precursors for novel materials has driven the growth of synthesis and characterization of new classes of main group compounds.

Cation-π interactions provide a powerful tool for constructing novel molecular architectures and hold promise for the development of potentially new electronic materials.   The main group multi-decker, sandwich-type complexes have not received a similar level of attention in comparison with transition metal analogues.   The main group multidecker anions of thallium[14] and cesium[15] were

4

identified in 1995 and 1996, respectively. An important step in the assembly of multidecker anions of the heavier main group elements consists of the addition of cyclopentadienide anions to neutral metallocenes in the presence of weakly coordinating cations.[14, 15, 19, 20] However, the inverse of this approach, *viz.* the synthesis of homonuclear multidecker *p*-block cations by the addition of positively charged fragments to neutral metallocenes in the presence of appropriate anions, had not been reported until Cowley *et al.*[21] reported the synthesis of $[(\eta^5\text{-}C_5Me_5)_3Sn_2)][Ga(C_6F_5)_4]$ in 2001. In an attempt to create a Ga→Sn donor-acceptor bond, it was found that the reaction of $Sn(\eta^5\text{-}C_5Me_5)_2$ with the Lewis acid $Ga(C_6F_5)_3$ seredipidously afforded the triple-decker cation $[(\eta^5\text{-}C_5Me_5)_3Sn_2)]^+$ (Figure 1.4). It was hypothesized that the formation of the triple-decker cation took place by the *in situ* formation of $[Sn(\eta^5\text{-}C_5Me_5)]^+$ *via* abstraction of a $[C_5Me_5]^-$ moiety by $Ga(C_6F_5)_3$, followed by addition of the mononuclear tin cation to the neutral stannocene, $[Sn(\eta^5\text{-}C_5Me_5)_2]$.[21]
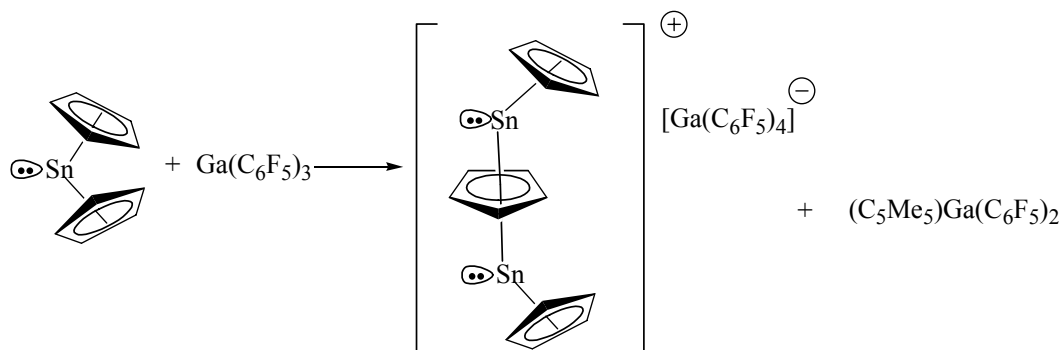
Figure 1.4 Reaction of [Sn(η^5-C$_5$Me$_5$)$_2$] with Ga(C$_6$F$_5$)$_4$. Methyl groups have been omitted for clarity.

In order to extend the triple-decker cation chemistry to the group 13 elements, it was necessary to recognize the isolobal relationship between [Sn(η^5-C$_5$Me$_5$)]$^+$ and [In(η^6-C$_6$H$_6$)]$^+$ (Figure 1.5). As summarized in Figure 1.5, the protolytic cleavage of In(η^5-C$_5$Me$_5$) by the Brønsted acid, H$_2$O·B(C$_6$F$_5$)$_3$ in the presence of an arene solvent provides a potential source of [In(arene)]$^+$ cations for addition to neutral In(η^5-C$_5$Me$_5$). Addition of an equimolar quantity of the Lewis acid, B(C$_6$F$_5$)$_3$,[22, 23] was used for complexation to the conjugate base of H$_2$O·B(C$_6$F$_5$)$_3$. Thus, treatment of In(η^5-C$_5$Me$_5$) with equimolar quantities of B(C$_6$F$_5$)$_3$,[22, 23] and H$_2$O·B(C$_6$F$_5$)$_3$[24] in toluene solution at 0ºC, afforded colorless crystals that were identified by X-ray diffraction as [(η^6-toluene)In(η^5-C$_5$Me$_5$)In(η^6-toluene)][HO—(B(C$_6$F$_5$)$_3$)$_2$] (Figure 1.6).[21]
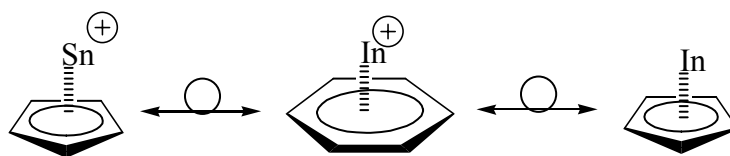
Figure 1.5 The isolobal relationships between $[Sn(\eta^5\text{-}C_5H_5)]^+$, $[In(\eta^6\text{-}C_6H_6)]^+$, and $[In(\eta^5\text{-}C_5H_5)]$.
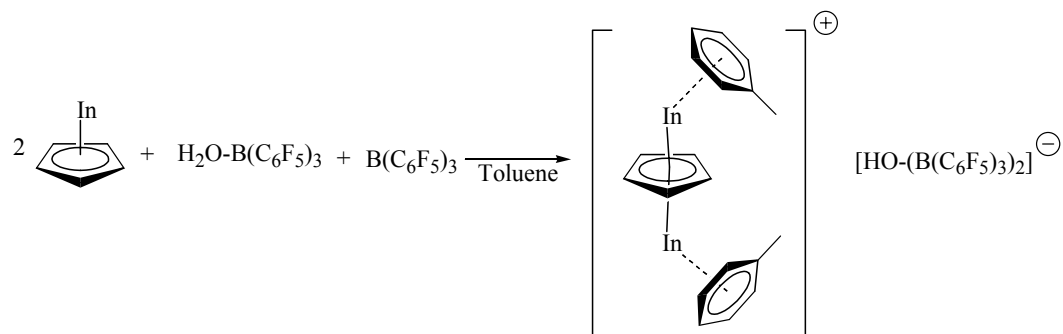
Figure 1.6   Protolytic cleavage of $In(\eta^5\text{-}C_5Me_5)$ with $H_2O\cdot B(C_6F_5)_3$ to give $[(\eta^6\text{-}toluene)In(\eta^5\text{-}C_5Me_5)In(\eta^6\text{-}toluene)]^+$.   Methyl groups of the $\eta^5\text{-}C_5Me_5$ have been omitted for clarity.

The bonding in $[(\eta^6\text{-}toluene)In(\eta^5\text{-}C_5Me_5)In(\eta^6\text{-}toluene)]^+$ can be interpreted in two different ways, namely (*i*) as a triple-decker sandwich cation or (*ii*) a base-stabilized inverse sandwich cation. Density functional theory (DFT) calculations indicate that η^6-coordination of the two toluene molecules to the $[In(\mu\text{-}C_5H_5)In]^+$ moiety causes only a slight perturbation of the core and the arene. Moreover, the arene—In bond dissociation energy (6.6 kcal/mol) suggests a very weak interaction, thus lending credence to identification of the complex as the first example of a main group inverse sandwich structure.

The present work focuses on validating and generalizing the concept of the synthesis of triple-decker, main group metallocene cations by the addition of positively charged fragments to neutral main group metallocenes. The scope of the general method is tested further by attempting to assemble heterogeneous metal multi-decker sandwich compounds. Additionally, investigations into the significance of the role of the gegenion will also be discussed.

RESULTS AND DISCUSSION

Synthesis and Characterization of [(η^5-Me$_5$C$_5$)Ge][B(C$_6$F$_5$)$_4$], [1][B(C$_6$F$_5$)$_4$]

The pentamethylcyclopentadienylgermanium(II) cation, **1**$^+$, was synthesized in low yield *via* the metathetical reaction of pentamethylcyclopentadienylgermanium(II) chloride with one equivalent of LiB(C$_6$F$_5$)$_4$. The observed equivalence of the ^{1}H and ^{13}C NMR spectral peaks of [**1**][B(C$_6$F$_5$)$_4$] supported the proposed C$_{5v}$ symmetry structure for the cation. Insight into the proposed structure of the anion was provided by the detection of the *ortho*, *meta,* and *para* phenyl ring carbons in the ^{13}C NMR spectrum, and the presence of a singlet at δ -20.23 in the ^{11}B NMR spectrum. The three peaks representing the *ortho*, *meta,* and *para* fluorine atoms on the phenyl rings were evidenced in the ^{19}F NMR spectrum.

Recrystallization of [**1**][B(C$_6$F$_5$)$_4$] by the slow diffusion of hexane and dichloromethane yielded crystals suitable for X-ray diffraction experiments. An X-ray diffraction study identified the crystals as the title compound shown in Figure 1.7. Details of the data collection, structure solution, and refinement are compiled in Table 1.3 and selected metrical parameters are listed in Table 1.4 and Table 1.5.

Figure 1.7 Molecular structure of [(η^5-C$_5$Me$_5$)Ge][B(C$_6$F$_5$)$_4$], [1][B(C$_6$F$_5$)$_4$]), showing the close contacts of the germanium cation with some of the fluorine atoms of the anion. Carbon, boron, fluorine, and germanium atoms are shown in black, orange, green, and red, respectively. Hydrogen atoms have been omitted for clarity.

As illustrated in Figure 1.7, the germanium atom is centered above the C$_5$Me$_5$ ring and shows close contact interactions with three of the fluorine atoms of the tetrakis(pentafluorophenyl)borate anion. The Ge—C$_{ring}$ bond distances range from 2.234(3)—2.250(3) Å with an average bond length of 2.241(3) Å, which is slightly shorter that that reported by Winters *et al.*,[25] for [(η^5-C$_5$Me$_5$)Ge][BF$_4$] (av. 2.255(2) Å). The ring-centroid to metal bond distance is 1.883(4) Å. The [η^5-C$_5$Me$_5$] ring exhibits delocalized π-bonding as evidenced by the C$_{ring}$—C$_{ring}$ bond distances of 1.368(4), 1.371(4), 1.382(4), 1.378(3), and 1.396(3) Å. Based on the pattern of bond distances, germanium is considered to

be pentahapto bonded to the pentametylcyclopentadienide ring. The ring itself is planar and the methyl groups extend out of the plane and away from the germanium center. Evidence for this is provided by the angles between the ring plane and the least-squares planes defined by the two adjacent ring-carbon atoms and the attached methyl carbon atoms, which have an average value of 2.9°. Three of the fluorine atoms (F44, F12, and F11) of the tetrakis(pentafluorophenyl)borate anion are located at distances which are less than the sum of the van der Waals radii of germanium and fluorine (van der Waals radii of Ge = 1.97 Å and F = 1.47 Å).[26] The close contacts between the fluorine atoms and the germanium atom range in distance from 3.089—3.312 Å. Although the cation is very similar to [Cp*Ge][BF$_4$] reported by Winters *et al.*, [25] there is no discussion of coordination of the germanium to the fluorine atoms of the anion in that publication.

Synthesis and Characterization of [(η^5-Me$_5$C$_5$)Sn][B(C$_6$F$_5$)$_4$], [2][B(C$_6$F$_5$)$_4$]

The pentamethylcyclopentadienyltin(II) cation, **2**$^+$, was synthesized *via* the metathetical reaction of pentamethylcyclopentadienyltin(II) chloride with one equivalent of LiB(C$_6$F$_5$)$_4$ as described by Rhodes, *et al.*[27] As reported by these authors, the ^{1}H and ^{13}C NMR spectra of [**2**][B(C$_6$F$_5$)$_4$] evidence the C$_{5v}$ symmetry of the cation. Insight into the structure of the anion was provided by the detection of the *ortho*, *meta,* and *para* phenyl ring carbon resonances in the ^{13}C NMR

11

spectrum. For further characterization of the proposed structure of the cation and anion, ^{119}Sn, ^{11}B, and ^{19}F NMR spectra were recorded. As expected, examination of the ^{119}Sn NMR spectrum revealed the presence of only one peak located at δ -2219. A singlet at δ -20.23 in the ^{11}B NMR spectrum and three peaks representing the *ortho*, *meta,* and *para* fluorine atoms on the phenyl rings in the ^{19}F NMR spectrum were consistent with the tetrahedral geometry of the $[B(C_6F_5)_4]^-$ anion.

Although the synthesis of [**2**][$B(C_6F_5)_4$] had been reported previously,[27] no structural data was available for this compound. Recrystallization of [**2**][$B(C_6F_5)_4$] by the slow diffusion of hexane and dichloromethane yielded crystals suitable for X-ray diffraction experiments. An X-ray diffraction study identified the crystals as the title compound shown in Figure 1.8. Details of the data collection, structure solution, and refinement are compiled in Table 1.6 and selected metrical parameters are listed in Table 1.6 and Table 1.7.

Figure 1.8   Molecular structure of $[(\eta^5\text{-}C_5Me_5)Sn][B(C_6F_5)_4]$, $[2][B(C_6F_5)_4]$, showing the close contacts of the tin cation with some of the fluorine atoms on the anion.   Carbon, boron, fluorine, and tin atoms are shown in black, orange, green, and red, respectively.   Hydrogen atoms have been omitted for clarity.

As illustrated in Figure 1.8, the C_{5v} symmetry of the $[(\eta^5\text{-}C_5Me_5)Sn]^+$ cation, which was initially indicated by NMR spectroscopy, was confirmed by the X-ray diffraction experiment.   The tin atom is centered above the C_5Me_5 ring and exhibits close contact interactions with four fluorine atoms of the neighboring tetrakis(pentafluorophenyl)borate anion (F51, F52, F54, and F55).   The Sn—C_{ring} bond distances range from 2.428(2)—2.446(3) Å with an average bond length of 2.438(2) Å.   The ring-centroid to metal bond distance is 2.114(3) Å.   The $[\eta^5\text{-}C_5Me_5)]$ ring exhibits delocalized π-bonding as evidenced by the C_{ring}—C_{ring} bond

distances of 1.426(4), 1.427(4), 1.428(4), 1.432(4), and 1.432(4) Å. Based on the pattern of bond distances, the tin atom is pentahapto bonded to the pentametylcyclopentadienide ring. The ring itself is planar and the methyl groups extend out of the plane and away from the tin center. Evidence for this is provided by the angles between the ring plane and the least-squares planes defined by the two adjacent ring-carbon atoms and the attached methyl carbon atoms, which have an average value of 5.7°.

Four of the fluorine atoms (F51, F52, F54, and F55) of the tetrakis(pentafluorophenyl)borate anion are located at distances which are less than the sum of the van der Waals radii of germanium and fluorine (van der Waals radii of Sn = 2.26 Å and F = 1.47 Å).[26] The close contacts between the fluorine atoms and the tin atom range in distance from 3.172(2)—3.294(1) Å. Overall, the structure of 2^+ closely compares with that of $[(\eta^5\text{-C}_5\text{Me}_5)\text{Sn}][\text{BF}_4]$ reported by Jutzi et $al.$, [8] in which the average Sn—C_{ring} bond distance is 2.462 Å. However, there was no discussion of tin interactions between the cation and the anion in this publication.

Synthesis and Characterization of $[(\eta^5\text{-Me}_5\text{C}_5)_3\text{Sn}_2][B(C_6F_5)_4]$, $[3][B(C_6F_5)_4]$

Tris(pentamethylcyclopentadienyl)bis(tin) cation, 3^+, was synthesized *via* the addition reaction of $[3][B(C_6F_5)_4]$ to one equivalent of bis(pentamethylcyclopentadienyl)tin. The ^{1}H NMR spectrum of $[3][B(C_6F_5)_4]$

exhibits only a single proton signal thus indicating the equivalence of all forty-five methyl protons on the three C_5Me_5 rings. Likewise, the ^{13}C NMR spectrum of 3^+ showed only two signals for the carbon atoms of the C_5Me_5 rings, namely one for the methyl carbons and one for the ring carbons. Three ^{13}C NMR signals were detected for the $[B(C_6F_5)_4]^-$ anion and correspond to the *ortho, meta,* and *para* ring carbon atoms. Examination of the ^{119}Sn NMR indicated the presence of only one peak located at δ -2223. Analysis of the ^{11}B and ^{19}F NMR spectra confirmed the presence of the $[B(C_6F_5)_4]^-$ anion. Overall, however, it was not possible to assign the structure of the cation unambiguously hence an X-ray diffraction study was undertaken.

Recrystallization of $[3][B(C_6F_5)_4]$ by the slow diffusion of hexane and dichloromethane yielded crystals suitable for X-ray diffraction experiments. An X-ray diffraction study identified the crystals as the title compound shown in Figure 1.9 and clarified the structure of the cation in the solid state. Details of the data collection, structure solution, and refinement are compiled in Table 1.9 and selected metrical parameters are listed in Table 1.10 and Table 1.11.
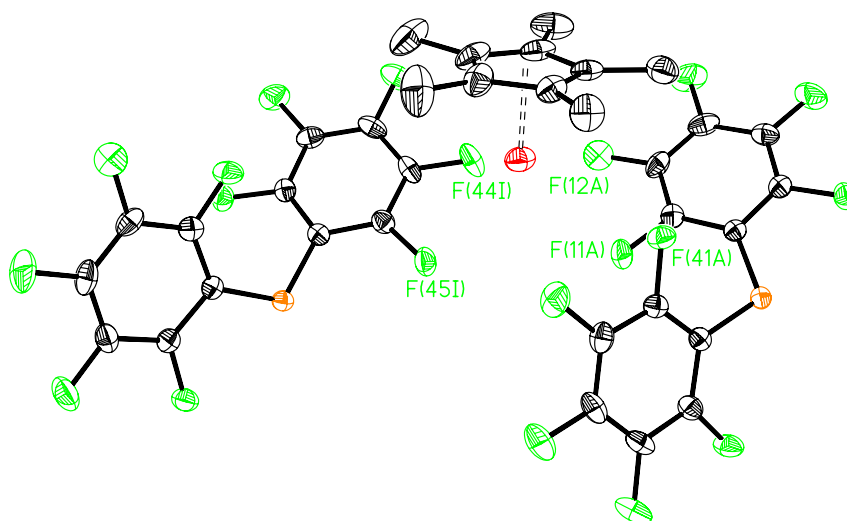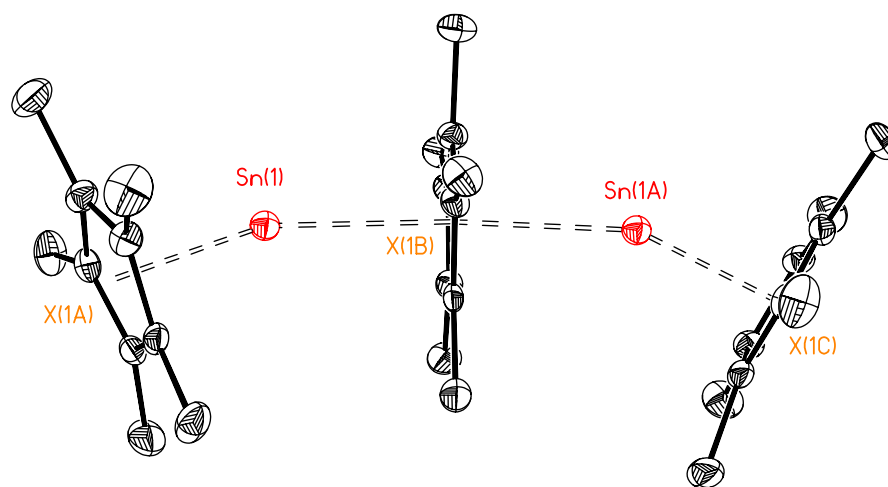
Figure 1.9   Molecular structure of [(η⁵-C₅Me₅)₃Sn₂]⁺, [3]⁺, showing *cis*-geometry of the cation. Carbon and tin atoms are shown in black and red, respectively. Hydrogen atoms have been omitted for clarity.

In the solid state, **3**⁺, exists as a cationic, triple-decker, sandwich-type structure (Figure 1.9). A pentahapto C_5Me_5 ring serves as a bridging group for two $Sn(\eta^5\text{-}C_5Me_5)$ units. Within experimental error, the two tin atoms are located equidistantly from the bridging C_5Me_5 ring-centroid (X1B), and are located in a nearly linear fashion on either side of the bridging C_5Me_5 group (Sn(1)—X(1B)—Sn(1A) = 175.5º). The terminal C_5Me_5 rings are positioned in a mutually *cis*-arrangement with a Sn—ring-centroid (X1A/X1C) distance of 2.233 Å, which is shorter than those to the bridging $\eta^5\text{-}C_5Me_5$ moiety (2.614 Å). The values lie between the value reported for $Sn(\eta^5\text{-}C_5Me_5)_2$ (2.396 Å)[8] and that determined for $[Sn(\eta^5\text{-}C_5Me_5)]^+$ (2.114(3) Å). The X(1A)—Sn—X(1C) angle of 154.7º is similar

to the values reported for $Sn(\eta^5\text{-}C_5Me_5)_2$ (av. 154.9°).[8] The *cis*-type geometry of the overall structure is intriguing because, as noted earlier, the triple-decker, main group anions, $[(\eta^5\text{-}C_5H_5)_3Tl_2]^{-[14]}$ and $[(\eta^5\text{-}C_5H_5)_3Cs_2]^{-[15]}$, possess transoid conformations.

Overall, the solution state NMR data were not in accord with the solid state structure as determined by X-ray diffraction. Thus the ^1H, ^{13}C, and ^{119}Sn NMR spectra evidenced only one type of C_5Me_5 group and a unique tin center suggesting rapid, reversible dissociation of **3$^+$** into $Sn(\eta^5\text{-}C_5Me_5)_2$ and $[Sn(\eta^5\text{-}C_5Me_5)]^+$ (Figure 1.10). In an attempt to identify the triple-decker cation in solution, low temperature ^1H and ^{119}Sn were attempted; however, the low-temperature spectra were identical to those obtained at room temperature.
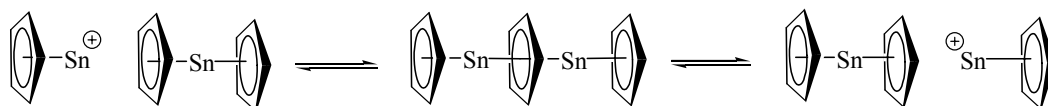
Figure 1.10 The proposed model for the facile exchange processes of [Sn(η^5-C$_5$Me$_5$)]$^+$ based on the analysis of NMR spectra. Methyl groups have been omitted for clarity.

Synthesis and Characterization of [(η^5-Me$_5$C$_5$)Pb][B(C$_6$F$_5$)$_4$], [4][B(C$_6$F$_5$)$_4$]

The pentamethylcyclopentadienyllead(II) cation, **4$^+$**, was synthesized in good yield *via* the metathetical reaction of pentamethylcyclopentadienyllead(II) chloride with one equivalent of $LiB(C_6F_5)_4$. As in the cases of the analogous

germanium and tin cations, the ^{1}H and ^{13}C NMR spectra of [**4**][B(C$_6$F$_5$)$_4$] were indicative of the C$_{5v}$ symmetry of the cation. Insight into the structure of the anion was provided by the detection of the *ortho*, *meta*, and *para* phenyl ring carbons in the ^{13}C NMR spectrum, a singlet at δ -20.23 in the ^{11}B NMR spectrum, and three peaks representing the *ortho*, *meta*, and *para* fluorine atoms on the phenyl rings in the ^{19}F NMR spectrum.

Recrystallization of [**4**][B(C$_6$F$_5$)$_4$] by the slow diffusion of hexane and dichloromethane yielded crystals suitable for X-ray diffraction experiments. An X-ray diffraction study identified the crystals as the title compound shown in Figure 1.11**.** Details of the data collection, structure solution, and refinement are compiled in Table 1.12 and selected metrical parameters are listed in Table 1.13 and Table 1.14.
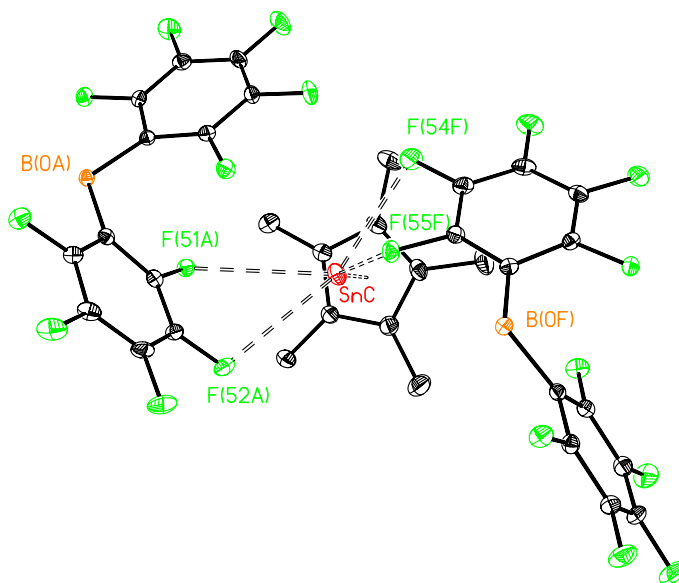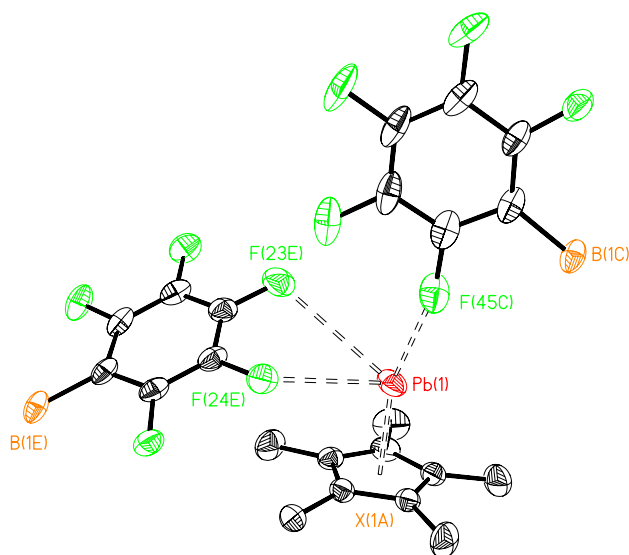
Figure 1.11 Molecular structure of [(η^5-C$_5$Me$_5$)Pb][B(C$_6$F$_5$)$_4$], [4][B(C$_6$F$_5$)$_4$], showing the close contacts of the tin cation with some of the fluorine atoms of the anion. Carbon, boron, fluorine, and lead atoms are shown in black, orange, green, and red, respectively. Hydrogen atoms have been omitted for clarity.

As illustrated in Figure 1.11, the lead atom in the [(η^5-C$_5$Me$_5$)Pb]$^+$ cation is centered above the C$_5$Me$_5$ ring. The C$_{5v}$ symmetry of **4**$^+$, previously indicated on the basis of NMR data, was verified by X-ray diffraction analysis. The Pb—C$_{ring}$ bond distances range from 2.538(4)—2.559(4) Å with an average bond length of 2.548 Å. The ring-centroid to metal bond distance is 2.240 Å. The [η^5-C$_5$Me$_5$)] ring exhibits a delocalized π-bonding as evidenced by the C$_{ring}$—C$_{ring}$ bond distances of 1.424(6), 1.428(6), 1.429(6), 1.435(6) and 1.503(6) Å. Based on the pattern of bond distances, the lead atom is pentahapto bonded to the

pentametylcyclopentadienide ring. The ring itself is planar and the methyl groups extend out of the plane and away from the tin center. Evidence for this arrangement is provided by the values of the angles between the ring plane and the least-squares planes defined by the two adjacent ring-carbon atoms and the attached methyl carbon atoms, which have an average value of 6.5°.

Three of the fluorine atoms (F23, F24, and F45) of the neighboring tetrakis(pentafluorophenyl)borate anions are located at distances which are less than the sum of the van der Waals radii of germanium and fluorine (van der Waals radii of Pb = 2.34 Å and F = 1.47 Å).[26] The close contacts between the fluorine atoms and the lead atom range in distance from 2.944—3.327 Å.

Overall, the structure of 4^+ is similar to that of $[(\eta^5\text{-}C_5Me_5)Pb][BF_4]$ reported by Jutzi et al., [28] which has an average Pb—C_{ring} bond distance of 2.565(10) Å. In this case, there is a distinct interaction of the lead atom with one fluorine atom on each of the two adjacent tetrafluoroborate anions.

Synthesis and Characterization of $[(\eta^5\text{-}Me_5C_5)_3Pb_2][B(C_6F_5)_4]$, $[5][B(C_6F_5)_4]$

The tris(pentamethylcyclopentadienyl)bis(lead) cation, 5^+, was synthesized *via* the reaction of $[5][B(C_6F_5)_4]$ with one equivalent of bis(pentamethylcyclopentadienyl)lead. The ^1H NMR spectrum of $[5][B(C_6F_5)_4]$ exhibited only a single proton signal, indicating the equivalence of all forty-five methyl protons on the three C_5Me_5 rings. The ^{13}C NMR spectrum of 5^+ showed

two signals for the carbon atoms on the C_5Me_5 rings, namely one for the methyl carbons and one for the ring carbons. ^{13}C NMR signals were detected for the *ortho*, *meta*, and *para* phenyl ring carbons of the anion. Analysis of the ^{11}B and ^{19}F NMR confirmed the presence of the expected $[B(C_6F_5)_4]^-$ anion.

Recrystallization of $[\textbf{5}][B(C_6F_5)_4]$ by the slow diffusion of hexane and dichloromethane yielded crystals suitable for X-ray diffraction experiments. An X-ray diffraction study identified the crystals as the title compound shown in Figure 1.12. Details of the data collection, structure solution, and refinement are compiled in Table 1.15 and selected metrical parameters are listed in Table 1.16 and Table 1.17.

Figure 1.12 **Molecular structure of $[(\eta^5\text{-}C_5Me_5)_3Pb_2][B(C_6F_5)_4]$, $[\textbf{5}][B(C_6F_5)_4]$, depicting the *cis*-geometry of the cation. Carbon and lead atoms are shown in black and red, respectively. Hydrogen atoms have been omitted for clarity.**

In the solid state, $\mathbf{5}^+$ possesses a triple-decker, sandwich-type structure (Figure 1.12). A pentahapto bonded C_5Me_5 ring serves as the bridging group between two $Pb(\eta^5\text{-}C_5Me_5)$ units. Within experimental error, the two lead atoms are located equidistantly from the bridging C_5Me_5 ring-centroid (X1B), and are located in a nearly linear fashion on wither side of the bridging C_5Me_5 groups, Pb(1)—X(1B)—Pb(1A) = 177.2°. The terminal C_5Me_5 rings are positioned in a mutually *cis*-arrangement with a Pb—ring-centroid (X1A/X1C) distance of 2.339 Å, which is shorter than that to the bridging $\eta^5\text{-}C_5Me_5$ moeity (2.672 Å) and lies between the value reported for $Pb(\eta^5\text{-}C_5Me_5)_2$ (av. 2.501(3) Å)[8] and that determined for $[Pb(\eta^5\text{-}C_5Me_5)]^+$ (2.240(3) Å). The X(1A)—Pb—X(1C) angle of 152.6° is similar to the value reported for $Pb(\eta^5\text{-}C_5Me_5)_2$ (151.3°).[9] The *cis*-type geometry of the overall structure is similar to that obtained for $\mathbf{5}^+$, but contrasts the transoid conformations of the triple-decker, main group anions, $[(\eta^5\text{-}C_5H_5)_3Tl_2]^-$ [14] and $[(\eta^5\text{-}C_5H_5)_3Cs_2]^-$.[15]

Overall, the solution state NMR data were not in accord with the solid state structure as determined by X-ray diffraction. Thus the ^1H, ^{13}C, and ^{207}Pb NMR spectra evidenced only one type of C_5Me_5 group and a unique lead center thus suggesting rapid, reversible dissociation of $\mathbf{5}^+$ into $Pb(\eta^5\text{-}C_5Me_5)_2$ and $[Pb(\eta^5\text{-}C_5Me_5)]^+$ (refer to Figure 1.10). In an attempt to identify the triple-decker

cation in solution, low temperature ^{1}H and ^{207}Pb were attempted; however, the low temperature spectra were identical to those obtained at room temperature.

Reactions of half-sandwich cations and neutral metallocenes

In an attempt to synthesize mixed metal, triple-decker metallocenes, a series of addition reactions between the half-sandwich cations and the neutral metallocenes were conducted. Initially, the group 14 half-sandwich cations were treated with transition metal metallocenes as summarized in Table 1.1 and Figure 1.13. However, each attempt afforded an oxidation/reduction product, which was identified by X-ray crystallography, rather than the desired mixed metal, triple-decker metallocene.

Half-sandwich cation	Neutral metallocene	Reaction product
$[Sn(\eta^5\text{-}C_5Me_5)]^+$	$Fe_2(\eta^5\text{-}C_5Me_5)$	$[(\eta^5\text{-}C_5Me_5)_2Fe]^+$
$[Sn(\eta^5\text{-}C_5Me_5)]^+$	$Co_2(\eta^5\text{-}C_5Me_5)$	$[(\eta^5\text{-}C_5Me_5)_2Co]^+$
$[Sn(\eta^5\text{-}C_5H_5)]^+$	$Ni_2(\eta^5\text{-}C_5H_5)$	$[(\eta^5\text{-}C_5H_5)_2Ni]^+$

Table 1.1 Reactions between group 14 half-sandwich cations and transition metal metallocenes.

23
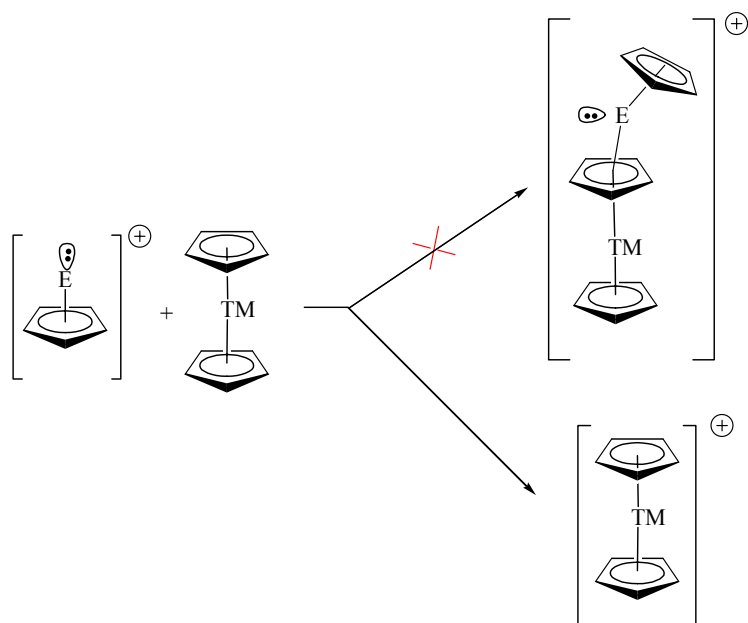
Figure 1.13   Potential reaction scheme for the synthesis of mixed metal, triple-decker metallocenes, where E = group 14 metal and TM = transition metal. All methyl groups have been omitted for clarity.

To avoid the oxidation/reduction problems that were encountered using group 14 elements and transition metals, reactions involving only group 14 metals were attempted (Figure 1.14) and the products identified *via* X-ray crystallography.   The results (Table 1.2) provided important information regarding the reactivity and stability of the metal cations. The stability of the cations of general formula, $[E(\eta^5\text{-}C_5Me_5)]^+$, appears to increase moving up the group 14 elements from Pb<Sn<Ge. The abstraction of a $[C_5Me_5]^-$ anion from germanocene by $[Sn(\eta^5\text{-}C_5Me_5)]^+$ to yield decamethylstannocene and the $[Ge(\eta^5\text{-}C_5Me_5)]^+$ cation evidences the electrophilic nature of the $[Sn(\eta^5\text{-}C_5Me_5)]^+$ cation.

24

The mechanism for the formation of $[(\eta^5-C_5Me_5)_3Pb_2]^+$ in the reaction of $[Sn(\eta^5-C_5Me_5)]^+$ with decamethylplumbocene, is hypothesized to proceed by the abstraction of a $[C_5Me_5]^-$ anion from $Pb_2(\eta^5-C_5Me_5)$ to form the less stable cation, $[Pb(\eta^5-C_5Me_5)]^+$. In turn, this electrophilic cation adds to any decamethylplumbocene that remains in the solution to form the non-stoichiometric product $[(\eta^5-C_5Me_5)_3Pb_2]^+$.

Half-sandwich cation	Neutral metallocene	Reaction product
$[Sn(\eta^5-C_5Me_5)]^+$	$Ge(\eta^5-C_5Me_5)_2$	$[Ge(\eta^5-C_5Me_5)]^+$
$[Sn(\eta^5-C_5Me_5)]^+$	$Pb(\eta^5-C_5Me_5)_2$	$[(\eta^5-C_5Me_5)_3Pb_2]^+$

Table 1.2   Reactions between group 14 half-sandwich cations and group 14 metallocenes.
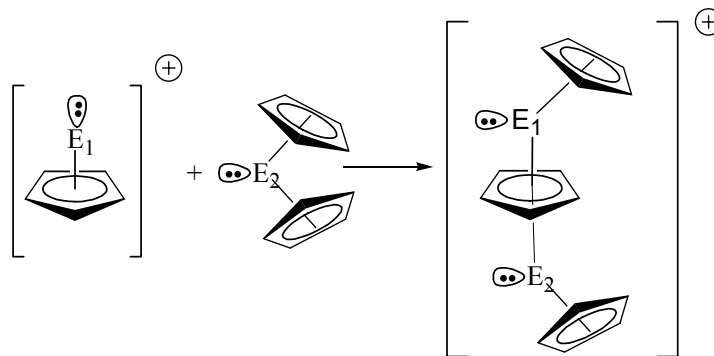
Figure 1.14   Potential reaction scheme for the synthesis of mixed metal, group 14, triple-decker metallocenes, where E_1 and E_2 are group 14 elements. Methyl groups are omitted for clarity.

In the reaction of $[Sn(\eta^5-C_5Me_5)]^+$ with $Ge(\eta^5-C_5Me_5)_2$, the $[Sn(\eta^5-C_5Me_5)]^+$ cation abstracted a $[C_5Me_5]^-$ anion from decamethylgermanocene to

yield the $[Ge(\eta^5\text{-}C_5Me_5)]^+$ cation and stannocene. A similar reaction between $[Sn(\eta^5\text{-}C_5Me_5)]^+$ and plumbocene yielded slightly different results, namely the formation of the triple-decker cation, $[(\eta^5\text{-}C_5Me_5)_3Pb_2]^+$.

Synthesis and Characterization of $[(\eta^5\text{-}\mu\text{-}Me_5C_5)In_2][B(C_6F_5)_4]$, $[6][B(C_6F_5)_4]$

In order to extend the triple-decker cation chemistry to group 13 elements, the isolobal relationship between $[Sn(\eta^5\text{-}C_5Me_5)]^+$ and $[In(\eta^5\text{-}C_6H_6)]^+$ was utilized. Previous work had structurally characterized $(\eta^6\text{-toluene})In(\mu\text{-}\eta^5\text{-}C_5Me5)In(\eta^6\text{-toluene})$, which as pointed out earlier (p. 7) can be construed as a triple-decker system or it can equally well be thought of as an "inverse sandwich" cation. The distinction between these possibilities depends on whether the capping η^6-interacting toluene molecules are bonded or not. In order to gain further insight into this question, it was decided to replace the voluminous $[(C_6F_5)_3BO(H)B(C_6F_5)_3]^-$ anion by a counter anion of more modest size. Utilizing the acidity of $[(toluene)H]^+$ to protonate a C_5Me_5 group, a second approach to the synthesis of multi-decker group 13 sandwich compounds was designed. The reaction of pentamethylcyclopentadienylindium with one equivalent of $[(toluene)H][B(C_6F_5)_4]$ at -78 °C resulted in the inverse sandwich cation, $[(\eta^5\text{-}\mu\text{-}Me_5C_5)In_2]$, **6**$^+$. The ^{1}H NMR spectrum of $[\mathbf{6}][B(C_6F_5)_4]$ evidenced the C_{5v} symmetry of the C_5Me_5 ring. Unfortunately, the impurity of the sample prevented the acquisition of a useful ^{13}C NMR spectrum.

26

Crystallization of [**6**][B(C$_6$F$_5$)$_4$] occurred during the initial work up of the sample and afforded crystals suitable for X-ray diffraction experiments. An X-ray diffraction study identified the colorless crystals as the title compound shown in Figure 1.15. Details of the data collection, structure solution, and refinement are compiled in Table 1.18 and selected metrical parameters are listed in Table 1.19 and Table 1.20.
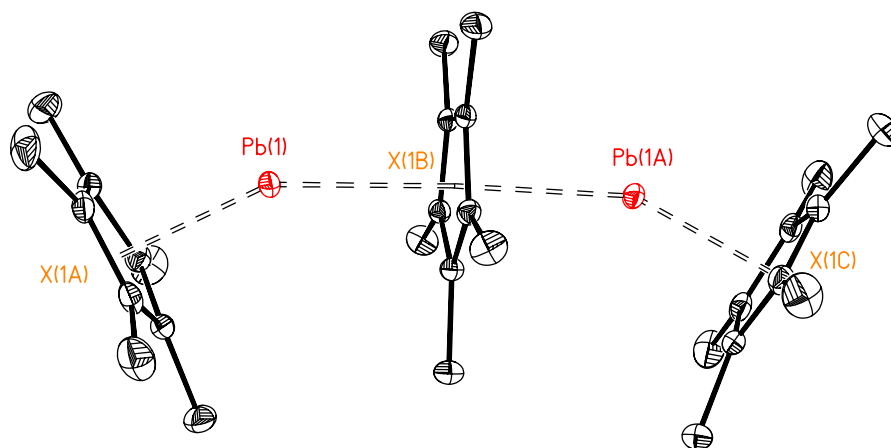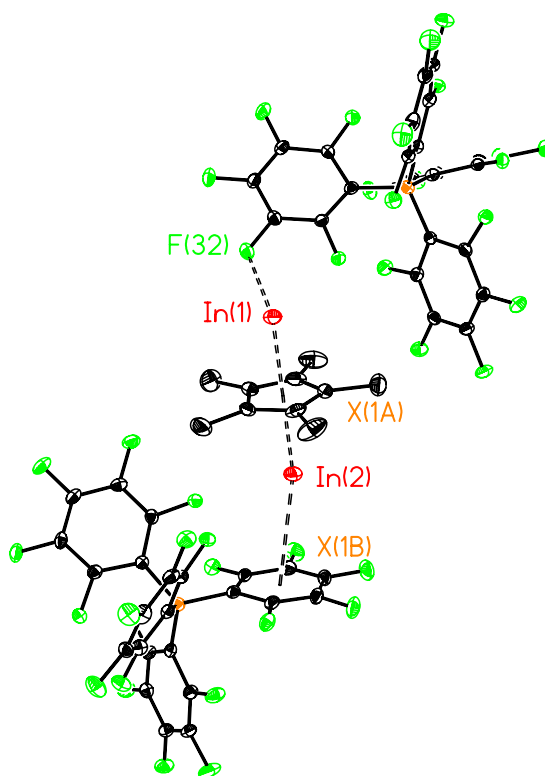

Figure 1.15 **Molecular structure of [(η^5-μ-C$_5$Me$_5$)In$_2$][B(C$_6$F$_5$)$_4$], [6][B(C$_6$F$_5$)$_4$], showing the close anion contacts of an indium atom with a neighboring fluorine atom and possible In—C$_6$F$_5$ π-interaction with the second indium atom. Carbon, boron, fluorine, and indium atoms are shown in black, orange, green, and red, respectively. Hydrogen atoms have been omitted for clarity.**

As illustrated in Figure 1.15, the indium atoms are centered above and below the bridging C_5Me_5 ring. The In—C_5Me_5 bond distances range from 2.698(8)—2.708(8) Å with an average bond length of 2.701 Å. The ring-centroid (X1A) to metal bond distances of 2.4350(7) and 2.4603(7) Å are longer than those reported[29] for monomeric (2.288(4) Å) and hexameric (2.302(4) Å) In(C_5Me_5). The In(1)—X(1A)—In(2) for the bridging C_5Me_5 ring is close to linear (177.2º). The η^5- C_5Me_5 ring exhibits delocalized π-bonding as evidenced by the C_{ring}—C_{ring} bond distances of 1.379(13), 1.404(11), 1.416(10), 1.417(13) and 1.490(13) Å. The foregoing bond distances indicate both indium atoms are pentahapto bonded to the pentamethylcyclopentadienide ring. The ring itself is planar with a mean deviation from the plane of 0.004 Å while the methyl groups have a mean deviation from the plane of 0.0189 Å. Essentially, the methyl carbon atoms are located in the same plane as the ring carbon atoms. Evidence for this conclusion stems from the angles between the ring plane and the least-squares planes defined by two adjacent ring-carbon atoms and the attached methyl carbon atoms, which have an average value of 1.3°.

Close cation-anion interactions are a consequence of the electrophilic condition of the indium atoms in the inverse sandwich cation. A *meta*-fluorine atom (F32) of a neighboring tetrakis(pentafluorophenyl)borate anion is located 3.150(3) Å from In(1). Although this distance is less than the sum of van der

28

Waals radii for In and F (3.90 Å) (van der Waals radius of In = 2.43 Å and F = 1.47 Å) [26], shorter In· · ·F contacts have, in fact, been reported for In(2,4,6-$(CF_3)_3C_6H_2)_3$ (av. 2.762(7) Å) and $In_2(2,4,6-(CF_3)_3C_6H_2)_4$ (range 2.801(10)-2.957(13) Å). [30] Interestingly, the other indium atom of 6^+ undergoes a weak η^6-interaction with one of the C_6F_5 moieties of a second $[B(C_6F_5)_4]^-$ anion. The In(2) to ring-centroid distance is 3.588(4) Å. This represents an unprecedented mode of interaction for the $[B(C_6F_5)_4]^-$. To determine whether the close contacts between the indium atoms and the anion were due to coordination or crystal packing forces, density functional theory calculations were undertaken. All calculations were performed with the Gaussian 94 package of programs. Each molecule/ion was fully optimized in the indicated symmetry. Density functional theory (DFT) optimization of the model system $[(C_6F_6)—In—(\eta^5-\mu-C_5H_5)—In]^+$ predicts a D_{5h} symmetric structure with a computed In—C_6F_6 bond distance of 3.218 Å, which is much shorter than the value observed experimentally for 6^+. The In—C_6F_6 bond dissociation energy (~ 7 kcal/mol) suggest an extremely weak interaction, which could easily be overcome by crystal packing forces.

Synthesis and Characterization of [(η^5-μ-H_5C_5)Tl_2][B(C_6F_5)_4], [7][B(C_6F_5)_4]

The inverse sandwich cation $[(\mu-H_5C_5)Tl_2]$, 7^+, was synthesized *via* the reaction of cyclopentadienylthallium(I) with one equivalent of $[(toluene)H][B(C_6F_5)_4]$ at -78 °C. A pure crystalline sample was never obtained;

29

however, high resolution, chemical ionization mass spectrometry indicated the presence of **7**[+].

Synthesis and Characterization of [(Me$_5$C$_5$)H$_2$][B(C$_6$F$_5$)$_4$], [8][B(C$_6$F$_5$)$_4$]

An attempt to prepare a gallium inverse sandwich cation by treatment of [(η[5]-C$_5$Me$_5$)Ga]$_6$ with [(η[6]-toluene)H][B(C$_6$F$_5$)$_4$] resulted unexpectedly in formation of the pentamethylcyclopentenyl cation, **8**[+], as its [B(C$_6$F$_5$)$_4$]$^-$ salt, [**8**][B(C$_6$F$_5$)$_4$]. The identity of **8**[+] was confirmed by X-ray diffraction studies. Following the identification of [**8**][B(C$_6$F$_5$)$_4$], a rational synthesis of this compound was developed by protonation of pentamethylcyclopentadiene with an equimolar quantity of [(toluene)H][B(C$_6$F$_5$)$_4$]. The ^{1}H NMR and ^{13}C NMR spectrum of [**8**][B(C$_6$F$_5$)$_4$] matched those reported previously by Campbell *et al.* [31]

Recrystallization of [**8**][B(C$_6$F$_5$)$_4$] by the slow diffusion of toluene and dichloromethane yielded crystals suitable for X-ray diffraction experiments. X-ray diffraction studies identified the crystals as the title compound shown in Figure 1.16. Details of the data collection, structure solution, and refinement are compiled in Table 1.21 and selected metrical parameters are listed in Table 1.22 and Table 1.23.
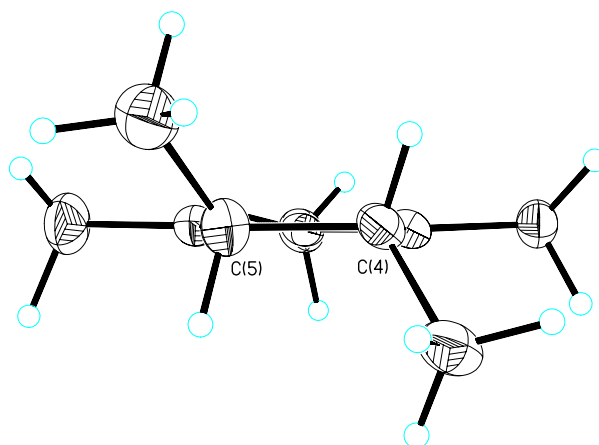
Figure 1.16 Side-on view of [(Me$_5$C$_5$)H$_2$]$^+$, 8$^+$, showing the hydrogen atoms attached to C(4) and C(5) and the *trans* arrangement of the methyl groups attached to these carbon atoms. Carbon and hydrogen atoms are shown in black and cyan, respectively.

The structure of the pentamethylcyclopentenyl cation, **8$^+$,** is illustrated in Figure 1.16.   The C$_{ring}$—C$_{ring}$ bond distances in **8$^+$** are 1.396(3), 1.385(3), 1.473(4), 1.478(3), and 1.517(4) Å indicating a localized cyclopentadiene-type structure for the ring. The C(4)—C(5) bond length is close to the value of 1.54 Å anticipated for a single bond between sp^3 hybridized carbon atoms. Moreover, tetrahedral geometries are evident at carbon atoms C(4) and C(5) and the methyl groups attached to these carbon atoms are arranged above and below the five-carbon ring in a *trans* fashion. The CH$_3$—C(4)—C(5)—H dihedral angle is 135.0(2)°.   The cyclopentenyl ring is somewhat non-planar, the internal ring dihedral angles being -2.1(3), 0.9(3), 2.7(3), -2.1(3), and 0.7(3)°.

The serendipitous synthesis and characterization of 8^+ occurred at the ideal moment; Lambert et al.[32] reported the crystal structure of what was thought be to be the pentamethylcyclopentadienyl cation, $[Me_5C_5]^+$. However, after meticulous inspection of the data presented in the Lambert et al. communication, the strong similarities to 8^+ were noted. The intra-ring C—C bond lengths of 8^+ and the structure of Lambert et al.[32] were determined to be nearly identical (Figure 1.17). A retraction letter[33] and a communication from our laboratory[34] were soon published identifying the 'pentamethylcyclopentadienyl cation' as the pentamethylcyclopentenyl cation, 8^+.
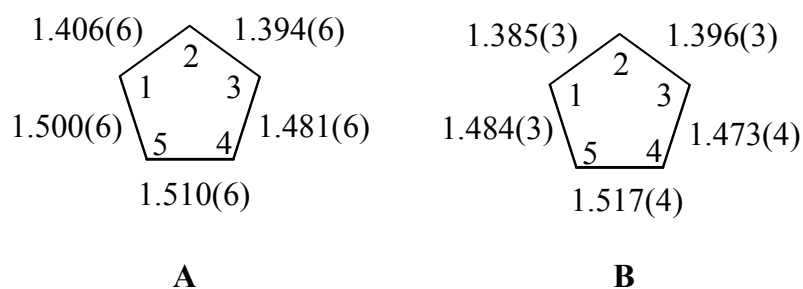
Figure 1.17   Comparison of intra-ring C—C bond lengths (Å) for A) Lambert et al.[32] and B) 8⁺.

CONCLUSIONS AND FUTURE WORK

A rational synthesis of triple-decker main group metallocene cations has been developed. The addition of positively charged half-sandwich cations to neutral metallocenes affords triple-decker cations of the formula, $[\eta^5\text{-}(C_5Me_5)_3E_2]^+$. The size of the gegenion plays a critical role in terms of

maximizing the lattice energies of the resulting salts, thus facilitating crystallization. Interestingly, the solid state and solution phase structures of the triple-decker cations are not identical. In the solid state, the cations have the structure shown in Figure 1.9; however, in solution phase, the triple-decker cations undergo facile exchange *via* reversible dissociation of $[\eta^5\text{-}(C_5Me_5)_3E_2]^+$ moieties. The exchange process is rapid with respect to the NMR spectroscopy time scale; the triple-decker species can not be identified by NMR spectroscopy at low temperature.

Unfortunately, attempts to produce mixed metal triple-decker metallocene cations were unsuccessful; however, valuable information regarding the reactivity of the $[\eta^5\text{-}(C_5Me_5)_3E_2]^+$ cations emerged from the studies. Based on the reactions of $[\eta^5\text{-}(C_5Me_5)Sn]^+$ with group 14 metallocenes, the stability of the half-sandwich cations is $[\eta^5\text{-}(C_5Me_5)Ge]^+ > [\eta^5\text{-}(C_5Me_5)Sn]^+ > [\eta^5\text{-}(C_5Me_5)Pb]^+$. Although the $[\eta^5\text{-}(C_5Me_5)_3Ge_2]^+$ cation has yet to be synthesized, the predicted stability of the triple-decker group 14 cations is $[\eta^5\text{-}(C_5Me_5)_3Ge_2]^+ > [\eta^5\text{-}(C_5Me_5)_3Sn_2]^+ > [\eta^5\text{-}(C_5Me_5)_3Pb_2]^+$.

One potential method to overcome the difficulties encountered in synthesizing mixed metal, triple-decker cations is by the use of the ansa-type ligands (Figure 1.18). The use of a ligand system where the aromatic moieties are conjoined will prevent the abstraction of a ligand by the cationic species.

However, to date, no such main group metallocene complexes have been structurally characterized.
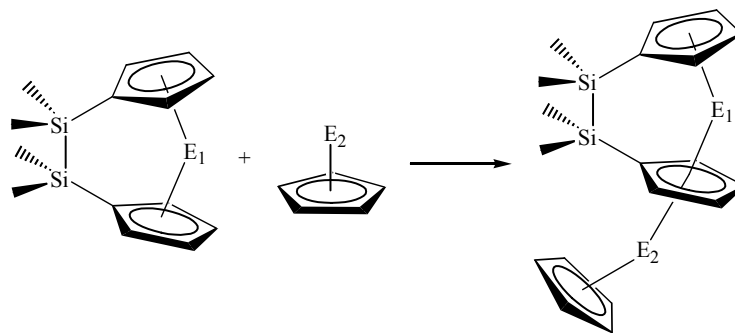
Figure 1.18   Potential reaction scheme to for the synthesis of mixed metal group 14 triple-decker cations using ansa-type metallocenes.

The synthesis of the first inverse sandwich main group cation has been accomplished by the reaction of $[\eta^5\text{-}(C_5Me_5)In]$ and $[(toluene)H][B(C_6F5)_4]$. The strong Brønsted acid, $[(toluene)H][B(C_6F5)_4]$, was also used for the synthesis of $[(C_5Me_5)H_2][B(C_6F5)_4]$. The ability of the acid to remove a $[C_5Me_5]^-$ anion by protolysis in the presence of a large, weakly coordinating anion, such as $[B(C_6F5)_4]^-$, may prove to be useful in future reactions.

EXPERIMENTAL SECTION

General Procedures

All solvents were distilled over sodium benzophenone ketyl, except dichloromethane, which was distilled over CaH_2, and degassed prior to use. An M-Braun or VAC Vacuum Atmosphere argon-filled drybox was used for the manipulation of all solid reagents. All reactions were performed under dry, oxygen-free conditions using standard Schlenk or drybox techniques. To ensure the absence of water, all glassware was dried overnight in a 140° C oven before use.

The group 14 dihalides were purchased from a commercial source and used without further purification. The compounds $(Me_5C_5)Li$[35], $LiB(C_6F_5)_4$[22], $(\eta-Me_5C_5)_2Ge$[36], $(\eta-Me_5C_5)_2Sn$,[36] and $(\eta-Me_5C_5)_2Pb$[9] were prepared according to literature procedures.

Physical Measurements

Low-resolution CI mass spectra were collected on a Finnigan MAT TSQ-700 machine; high resolution mass spectra were measured on a VG Analytical ZAB-VE sector instrument. Mass spectral samples were flame sealed in glass capillaries to prevent exposure to oxygen. Unless otherwise noted, solution phase ^{13}C and ^1H NMR spectra were recorded at 295 K on a GE EQ-300 instrument (^1H,

300 MHz; ^{13}C, 75 MHz). Some ^{13}C and ^{1}H NMR spectra and all ^{11}B, ^{19}F, ^{119}Sn, and ^{207}Pb solution-phase NMR spectra were recorded at 295 K, unless otherwise noted, on a Varian Inova-500 spectrometer (^{1}H, 500 MHz; ^{11}B, 160 MHz; ^{13}C, 125 MHz; ^{19}F, 470 MHz; ^{27}Al, 130 MHz) instrument. NMR samples were flame-sealed or recorded immediately following their removal from the drybox. Deuterodichloromethane was obtained in sealed vials from a commercial source and used without further purification. ^{1}H, ^{11}B, ^{13}C, and ^{19}F spectra are reported relative to tetramethylsilane (0.00 ppm) and referenced to the solvent. ^{119}Sn chemical shifts are reported relative to Me$_4$Sn (0.00 ppm), while ^{207}Pb spectra are reported relative to Me$_4$Pb (0.00 ppm). Unless otherwise noted, ^{13}C spectra were obtained under conditions of broadband proton decoupling. Melting points were obtained on a Fisher-Johns apparatus after flame-sealing the samples in glass capillaries under argon; the reported values are uncorrected.

X-Ray Crystallography

Suitable single crystals were covered with a perfluorinated polyether oil to minimize exposure to oxygen. The X-ray data were collected on a Nonius Kappa CCD diffractometer equipped with an Oxford Cryostream liquid nitrogen cooling stream. Structure determination and refinements were performed by the author at the University of Texas at Austin. All structures were solved by direct methods and refined by full-matrix least squares on F^2 using the Siemens SHELX PLUS

5.0 (PC) software package.[37] All non-hydrogen atoms were allowed anisotropic thermal motion. When possible, all hydrogen atoms were located *via* the model; however, it was occasionally necessary to place the hydrogen atoms in calculated positions (C—H 0.96 Å). The hydrogen atoms were refined using a riding model and a general isotropic thermal parameter. The total number of reflections, collection ranges, and final R-values for each molecule are listed in the appropriate crystallographic data tables.

Samples were mounted on a fine glass fiber using commercial silicon grease as an adhesive. All data collections were performed at 153(2) K using a graphite monochromated Mo Kα radiation (λ = 0.71073 Å). A correction was applied for Lorentz-polarization.

Synthesis of [(η^5-Me$_5$C$_5$)Ge][B(C$_6$F$_5$)$_4$], [1][B(C$_6$F$_5$)$_4$]

A yellow solution of [(η^5-Me$_5$C$_5$)Sn][B(C$_6$F$_5$)$_4$] (0.636 g, 0.682 mmol) in methylene chloride (40 mL) was treated with a methylene chloride solution (40 mL) of decamethylgermanocene (0.234 g, 0.682 mmol) at 0° C. The pale yellow solution was allowed to warm to room temperature and stirred overnight. Lithium chloride was removed by filtration through Celite®. The solvent and volatiles were removed under low pressure until the volume of the solution was approximately 10 mL, at which point it was transferred to a narrow layering vessel and layered with hexane (15 mL). A crop of pale yellow crystals (~20%

yield) formed at the solvent interface. *Spectroscopic Data*: HRMS (CI, CH_4):

calcd for $C_{10}H_{15}Ge$, ^{1}H NMR (CD_2Cl_2): δ 2.231[s, η^5-$C_5(CH_3)_5$]; ^{11}B NMR

(CD_2Cl_2): δ -12.941 [s,$B(C_6F_5)_4$]; ^{13}C{^{1}H} NMR (CD_2Cl_2): δ 9.265 [s, η^5-

$C_5(CH_3)_5$], 123.660 [s, η^5-$C_5(CH_3)_5$]; ^{19}F NMR (CD_2Cl_2): δ -132.936 [s,m-C_6F_5],

-165.314 [t, p-C_6F_5], -167.888 [s, o–C_6F_5].

Synthesis of [(η^5-Me$_5$C$_5$)Sn][B(C$_6$F$_5$)$_4$], [2][B(C$_6$F$_5$)$_4$]

A pale yellow solution of (η^5-Me$_5$C$_5$)SnCl (0.551 g, 1.46 mmol) in

methylene chloride (40 mL) was treated with a methylene chloride solution (40

mL) of $LiB(C_6F_5)_4$ (1.000 g, 1.46 mmol) at 0° C. The resulting orange solution

was allowed to warm to room temperature and stirred overnight. Lithium

chloride was removed by filtration through Celite[®]. The solvent and volatiles

were removed under low pressure until the volume of the solution was

approximately 10 mL, at which point it was transferred to a narrow layering

vessel and layered with hexane (15 mL). A crop of orange crystals (0.597 g,

40.1% yield) formed at the solvent interface. *Spectroscopic data*: matched that

reported by.[27]

Synthesis of [(η^5-Me$_5$C$_5$)$_3$Sn$_2$][B(C$_6$F$_5$)$_4$], [3][B(C$_6$F$_5$)$_4$]

A pale yellow solution of [(η^5-Me$_5$C$_5$)Sn][$B(C_6F_5)_4$] (0.183 g, 0.196

mmol) in methylene chloride (30 mL) was added to a stirred orange solution of

38

(η-Me₅C₅)₂Sn (0.076g, 0.195 mmol) at 0° C. The bright yellow solution was warmed slowly to room temperature and stirred overnight. Lithium chloride was removed by filtration through Celite[®]. The solvent and volatiles were reduced under low pressure until the volume of the solution was approximately 10 mL, at which point it was transferred to a narrow layering vessel and layered with hexane (15 mL). A crop of bright yellow crystals (0.193g, 74.8% yield) formed at the solvent interface. *Spectroscopic Data*: HRMS (CI, CH$_4$): calcd for C$_{30}$H$_{45}$Sn$_2$, *m/z* 645.1565; found, 645.1575; ^{1}H NMR (CD$_2$Cl$_2$): δ2.123[s, η^5-C$_5$(CH$_3$)$_5$]; ^{11}B NMR (CD$_2$Cl$_2$): δ-20.23 [s,B(C$_6$F$_5$)$_4$]; ^{13}C{^{1}H} NMR (CD$_2$Cl$_2$): δ 10.08 [s, η^5-C$_5$(CH$_3$)$_5$] 121.50 [s, η^5-C$_5$(CH$_3$)$_5$]; ^{19}F NMR (CD$_2$Cl$_2$): δ-133.33 [s,*m*-C$_6$F$_5$], -163.98 [t, *p*-C$_6$F$_5$], -167.80 [s, *o*–C$_6$F$_5$]; ^{119}Sn NMR (CD$_2$Cl$_2$): δ -2223.451.

Synthesis of [(η^5-Me$_5$C$_5$)Pb][B(C$_6$F$_5$)$_4$], [4][B(C$_6$F$_5$)$_4$]

An solution of (η^5-Me$_5$C$_5$)PbCl (0.551 g, 1.46 mmol) in methylene chloride (40 mL) was treated with a methylene chloride solution (40 mL) of LiB(C$_6$F$_5$)$_4$ (1.000 g, 1.46 mmol) at 0° C. The resulting orange solution was warmed to room temperature and stirred overnight. Lithium chloride was removed by filtration through Celite[®]. The solvent and volatiles were reduced under low pressure until the volume of the solution was approximately 10 mL, at which point it was transferred to a narrow layering vessel and layered with hexane

(15 mL). A crop of orange crystals (0.597 g, 40.1% yield) formed at the solvent interface.

Spectroscopic Data: HRMS (CI, CH_4): calcd for $C_{34}H_{15}BF_{20}Pb$ 343.0940 found 343.0936; ^1H NMR (CD_2Cl_2): δ 2.67 [s, η^5-$C_5(CH_3)_5$]; ^{11}B NMR (CD_2Cl_2): δ -20.23 [s, $B(C_6F_5)_4$]; $^{13}C\{^1H\}$ NMR (CD_2Cl_2): δ 8.989 [s, η^5-$C_5(CH_3)_5$], 125.943 [s, η^5-$C_5(CH_3)_5$]; ^{19}F NMR (CD_2Cl_2): δ -133.43 [s,m-C_6F_5], -163.94 [t, p-C_6F_5], -167.79 [s, o–C_6F_5].

Synthesis of [(η^5-Me$_5$C$_5$)$_3$Pb$_2$][B(C$_6$F$_5$)$_4$], [5][B(C$_6$F$_5$)$_4$]

An orange solution of [(η^5-Me$_5$C$_5$)Pb][B(C$_6$F$_5$)$_4$] (0.535 g, 0.524 mmol) in methylene chloride (30 mL) was added to a stirred orange solution of (η^5-C$_5$Me$_5$)$_2$Pb (0.250g, 0.523 mmol) at 0° C. The resulting deep red-orange solution was allowed to reach room temperature and stirred overnight. Lithium chloride was removed by filtration through Celite$^®$. The solvent and volatiles were reduced under low pressure until the volume of the solution was approximately 10 mL, at which point it was transferred to a narrow layering vessel and layered with hexane (15 mL). A crop of ruby red crystals (~65% yield) formed at the solvent interface.

Spectroscopic Data: HRMS (CI, CH_4): calcd for $C_{30}H_{45}Pb_2$ 821.305 found 821.306; ^1H NMR (CD_2Cl_2): δ 2.286 [s, η^5-$C_5(CH_3)_5$]; ^{11}B NMR (CD_2Cl_2): δ-20.23; $^{13}C\{^1H\}$ NMR (CD_2Cl_2): δ 9.0887 [s, η^5-$C_5(CH_3)_5$], 125.7410 [s, η^5-

$C_5(CH_3)_5$]; ^{19}F NMR (CD_2Cl_2): δ -133.37 [s,m-C_6F_5], -163.91 [t, p-C_6F_5], -167.76 ['t', o–C_6F_5];

Synthesis of [($η^5$-$μ$-Me_5C_5)In_2][B(C_6F_5)$_4$], [6][B(C_6F_5)$_4$]

[(Toluene)H][B(C_6F_5)$_4$] was prepared by first treating a frozen toluene solution (150 mL) with a large excess of condensed HCl (g) (3 mL, 82.1 mmol), followed by the slow addition of a solution of LiB(C_6F_5)$_4$ (5.00 g, 7.29 mmol) in toluene (60 mL). The reaction mixture was allowed to equilibrate slowly to room temperature and stirred overnight. By means of a double solvent trap, the volume of the solution was reduced to 25 mL and stored at -20° C, thereby affording a crop of colorless crystals (1.609 g, 28.6%). The crystals were harvested and used without further purification. A solution of ($η^5$-Me_5C_5)In (0.648 g, 2.59 mmol) in toluene (40 mL) was added slowly to a stirred solution of [(toluene)H][B(C_6F_5)$_4$] (1.00 g, 1.30 mmol) at -78 °C. The volatiles were immediately removed under reduced pressure from the resulting brown solution to yield a thick brown oily solid. Upon close observation, colorless needle-shaped crystals could be identified in the oily solid. Unfortunately, however, completely pure crystalline samples were never obtained. *Spectroscopic Data*: HRMS (CI, CH$_4$): calcd for $C_{10}H_{15}In_2$ 364.925 found 364.926; ^{1}H NMR (C_6D_6): δ 1.49 [s, $η^5$-$C_5(CH_3)_5$].

41

Synthesis of [(η^5-μ-C$_5$H$_5$)Tl$_2$][B(C$_6$F$_5$)$_4$], [7][B(C$_6$F$_5$)$_4$]

A solution of (η^5-C$_5$H$_5$)Tl (0.159 g, 0.589 mmol) in toluene (30 mL) was added slowly to a stirred solution of [(toluene)H][B(C$_6$F$_5$)$_4$] (0.227 g, 0.294 mmol) at -78 °C. The volatiles were removed immediately under reduced pressure from the resulting brown solution to yield a brown oily solid. Unfortunately, completely pure crystalline samples were never obtained. *Spectroscopic Data*: HRMS (CI, CH$_4$): calc. for C$_{10}$H$_{15}$Tl$_2$ 474.988 found 474.988.

Synthesis of [(Me$_5$C$_5$)H$_2$][B(C$_6$F$_5$)$_4$], [8][B(C$_6$F$_5$)$_4$]

Approach A:

Addition of a toluene solution of [(toluene)H][B(C$_6$F$_5$)$_4$] (0.273 g, 0.344 mmol) to a solution of [Ga(η^5-Me$_5$C$_5$)]$_6$ (0.145 g, 0.118 mmol) in the same solvent at -78 °C resulted in a red-violet colored reaction mixture. The volatiles were removed immediately from the cool solution under reduced pressure to yield a red solid. The red product was extracted into toluene (10 mL) and transferred to a narrow layering tube where it was layered with dichloromethane (15 mL). A crop of colorless crystals (~ 5% yield) formed at the solvent interface.

Approach B:

[(Toluene)H][B(C$_6$F$_5$)$_4$] (0.141 g, 0.183 mmol) in toluene (40 mL) was added to a toluene solution of (Me$_5$C$_5$)H (0.025 g, 0.184 mmol) at -78 °C. The

volatiles were removed immediately from the cool solution under reduced pressure. The pale yellow product was recrystallized from the slow diffusion of toluene and dichloromethane at room temperature to afford a crop of colorless crystals (0.034 g, 24% yield). *Spectroscopic Data:* HRMS (CI, CH$_4$): calc. for C$_{34}$H$_{17}$ 137.1330 found 137.1334; ^{1}H and ^{13}C NMR data matched those published previously by Campbell, *et al.*[31]

TABLES OF X-RAY CRYSTALLOGRAPHIC DATA
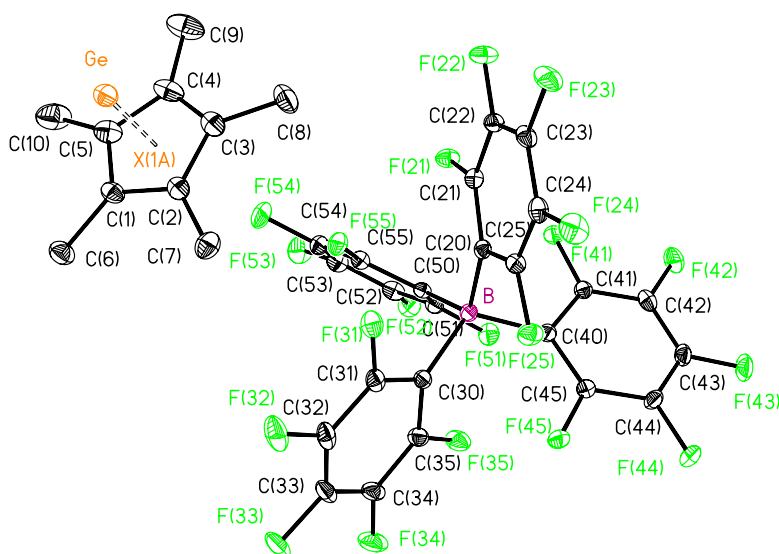
Figure 1.19 Molecular structure of $[(\eta^5\text{-}C_5Me_5)Ge][B(C_6F_5)_4]$, $[1][B(C_6F_5)_4]$, showing the atom number scheme. The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.
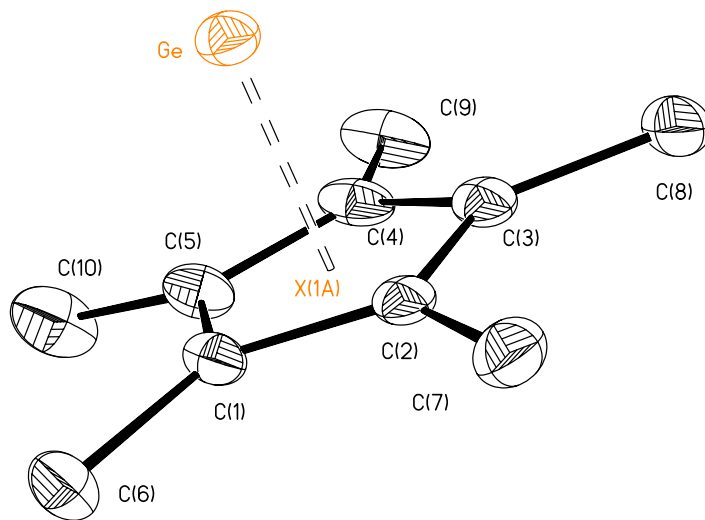
Figure 1.20 Molecular structure of $[(\eta^5\text{-}C_5Me_5)Ge]^+$, $[1]^+$, showing the atom number scheme for the cation. The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.

Table 1.3 Crystal data and structure refinement for [(η^5-C$_5$Me$_5$)Ge][B(C$_6$F$_5$)$_4$], [1][B(C$_6$F$_5$)$_4$].

Identification code	[Cp*Ge][B(C$_6$F$_5$)$_4$]
Empirical formula	C34 H15 B F20 Ge
Formula weight	886.86
Temperature	153(2) K
Wavelength	0.71069 Å
Crystal system	Monoclinic
Space group	P2$_1$/c
Unit cell dimensions	a = 13.456(5) Å α= 90°
	b = 14.052(5) Å β= 90.940(5)°.
	c = 17.494(5) Å γ = 90°
Volume	3307.4(19) Å^3
Z	4
Density (calculated)	1.781 Mg/m^3
Absorption coefficient	1.066 mm^{-1}
F(000)	1744
Crystal size	0.20 x 0.20 x 0.20 mm^3
Theta range for data collection	3.03 to 30.51°.
Index ranges	-19≤h≤19, -19≤k≤18, -24≤l≤24
Reflections collected	16719
Independent reflections	10006 [R(int) = 0.0307]
Completeness to theta = 30.51°	99.0 %
Max. and min. transmission	0.8150 and 0.8150
Refinement method	Full-matrix least-squares on F^2
Goodness-of-fit on F^2	1.015
Final R indices [I>2sigma(I)]	R1 = 0.0519, wR2 = 0.0970
R indices (all data)	R1 = 0.1026, wR2 = 0.1134
Largest diff. peak and hole	1.143 and -1.050 e.Å^{-3}

Table 1.4 Selected bond lengths (Å) for $[(\eta^5\text{-}C_5Me_5)Ge][B(C_6F_5)_4]$, $[1][B(C_6F_5)_4]$).

C(1)-C(5)	1.427(4)	C(3)-C(8)	1.504(4)
C(1)-C(2)	1.433(4)	C(3)-Ge	2.250(3)
C(1)-C(6)	1.497(4)	C(4)-C(5)	1.436(4)
C(1)-Ge	2.237(3)	C(4)-C(9)	1.504(4)
C(2)-C(3)	1.430(4)	C(4)-Ge	2.246(3)
C(2)-C(7)	1.490(4)	C(5)-C(10)	1.497(5)
C(2)-Ge	2.233(3)	C(5)-Ge	2.240(3)
C(3)-C(4)	1.418(4)		

Table 1.5 Selected bond angles (°) for $[(\eta^5\text{-}C_5Me_5)Ge][B(C_6F_5)_4]$, $[1][B(C_6F_5)_4]$.

C(5)-C(1)-C(2)	107.7(2)	C(5)-C(4)-C(9)	125.3(3)
C(5)-C(1)-C(6)	127.1(3)	C(3)-C(4)-Ge	71.78(15)
C(2)-C(1)-C(6)	125.2(3)	C(5)-C(4)-Ge	71.07(16)
C(5)-C(1)-Ge	71.51(15)	C(9)-C(4)-Ge	126.0(2)
C(2)-C(1)-Ge	71.17(14)	C(1)-C(5)-C(4)	107.8(3)
C(6)-C(1)-Ge	124.4(2)	C(1)-C(5)-C(10)	126.6(3)
C(3)-C(2)-C(1)	108.2(2)	C(4)-C(5)-C(10)	125.6(3)
C(3)-C(2)-C(7)	126.4(3)	C(1)-C(5)-Ge	71.30(15)
C(1)-C(2)-C(7)	125.3(3)	C(4)-C(5)-Ge	71.59(15)
C(3)-C(2)-Ge	72.05(15)	C(10)-C(5)-Ge	125.2(3)
C(1)-C(2)-Ge	71.43(15)	C(2)-Ge-C(1)	37.40(10)
C(7)-C(2)-Ge	124.31(19)	C(2)-Ge-C(5)	62.19(10)
C(4)-C(3)-C(2)	107.9(2)	C(1)-Ge-C(5)	37.19(10)
C(4)-C(3)-C(8)	125.8(3)	C(2)-Ge-C(4)	61.84(10)
C(2)-C(3)-C(8)	126.3(3)	C(1)-Ge-C(4)	62.12(10)
C(4)-C(3)-Ge	71.47(16)	C(5)-Ge-C(4)	37.34(11)
C(2)-C(3)-Ge	70.75(14)	C(2)-Ge-C(3)	37.20(10)
C(8)-C(3)-Ge	125.3(2)	C(1)-Ge-C(3)	62.26(10)
C(3)-C(4)-C(5)	108.4(2)	C(5)-Ge-C(3)	62.07(11)
C(3)-C(4)-C(9)	126.2(3)	C(4)-Ge-C(3)	36.75(11)
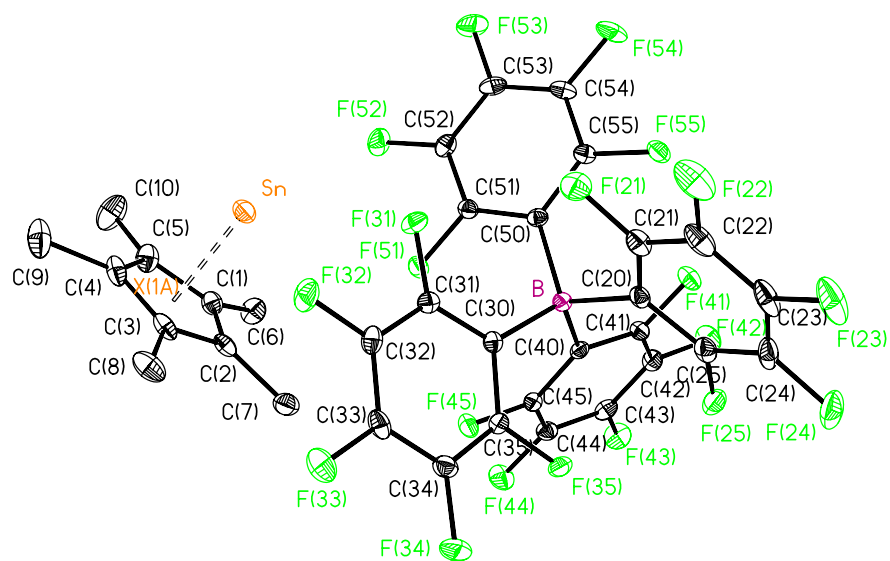
Figure 1.21 Molecular structure of [(η^5-C$_5$Me$_5$)Sn][B(C$_6$F$_5$)$_4$], [2][B(C$_6$F$_5$)$_4$], showing the atom number scheme. The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.

Figure 1.22 Molecular structure of [(η^5-C$_5$Me$_5$)Sn]$^+$, 2$^+$, showing the atom numbering scheme for the cation. The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.

Table 1.6   Crystal   data   and   structure   refinement   for   [(η^5-C$_5$Me$_5$)Sn][B(C$_6$F$_5$)$_4$], [2][B(C$_6$F$_5$)$_4$].

Identification code	p21overc
Empirical formula	C34 H15 B F20 Sn
Formula weight	932.96
Temperature	153(2) K
Wavelength	0.71069 Å
Crystal system	Monoclinic
Space group	P2$_1$/c
Unit cell dimensions	a = 13.442(5) Å α= 90°
	b = 14.160(5) Å β= 91.929(5)°
	c = 17.551(5) Å γ = 90°
Volume	3339(2) Å^3
Z	4
Density (calculated)	1.856 Mg/m^3
Absorption coefficient	0.904 mm^{-1}
F(000)	1816
Crystal size	0.20 x 0.20 x 0.20 mm^3
Theta range for data collection	3.09 to 27.47°.
Index ranges	-17≤h≤17, -18≤k≤18, -22≤l≤22
Reflections collected	13883
Independent reflections	7559 [R(int) = 0.0226]
Completeness to theta = 27.47°	98.7 %
Absorption correction	None
Max. and min. transmission	0.8399 and 0.8399
Refinement method	Full-matrix least-squares on F^2
Goodness-of-fit on F^2	1.057
Final R indices [I>2sigma(I)]	R1 = 0.0317, wR2 = 0.0622
R indices (all data)	R1 = 0.0447, wR2 = 0.0661
Extinction coefficient	0.0040(3)
Largest diff. peak and hole	0.491 and -0.258 e.Å^{-3}

Table 1.7   Selected Bond Lengths (Å) for [(η⁵-C₅Me₅)Sn][B(C₆F₅)₄], [2][B(C₆F₅)₄].

C(1)-C(2)	1.426(4)	C(3)-C(8)	1.498(4)
C(1)-C(5)	1.432(4)	C(3)-Sn	2.444(2)
C(1)-C(6)	1.495(4)	C(4)-C(5)	1.427(4)
C(1)-Sn	2.434(2)	C(4)-C(9)	1.499(4)
C(2)-C(3)	1.432(4)	C(4)-Sn	2.446(3)
C(2)-C(7)	1.498(4)	C(5)-C(10)	1.503(4)
C(2)-Sn	2.428(2)	C(5)-Sn	2.439(3)
C(3)-C(4)	1.428(4)		

Table 1.8 Selected Bond Angles (°) for $[(\eta^5\text{-}C_5Me_5)Sn][B(C_6F_5)_4]$, $[2][B(C_6F_5)_4]$.

C(2)-C(1)-C(5)	107.8(2)	C(5)-C(4)-C(9)	125.5(3)
C(2)-C(1)-C(6)	125.4(2)	C(3)-C(4)-Sn	72.95(14)
C(5)-C(1)-C(6)	126.6(2)	C(5)-C(4)-Sn	72.72(14)
C(2)-C(1)-Sn	72.70(13)	C(9)-C(4)-Sn	125.3(2)
C(5)-C(1)-Sn	73.09(14)	C(4)-C(5)-C(1)	108.2(2)
C(6)-C(1)-Sn	124.14(19)	C(4)-C(5)-C(10)	124.7(3)
C(1)-C(2)-C(3)	108.1(2)	C(1)-C(5)-C(10)	126.9(3)
C(1)-C(2)-C(7)	125.1(2)	C(4)-C(5)-Sn	73.30(14)
C(3)-C(2)-C(7)	126.6(2)	C(1)-C(5)-Sn	72.74(14)
C(1)-C(2)-Sn	73.19(14)	C(10)-C(5)-Sn	124.5(2)
C(3)-C(2)-Sn	73.54(14)	C(2)-Sn-C(1)	34.12(8)
C(7)-C(2)-Sn	123.05(17)	C(2)-Sn-C(5)	56.65(9)
C(4)-C(3)-C(2)	107.9(2)	C(1)-Sn-C(5)	34.17(9)
C(4)-C(3)-C(8)	125.6(3)	C(2)-Sn-C(3)	34.18(8)
C(2)-C(3)-C(8)	126.2(3)	C(1)-Sn-C(3)	56.64(8)
C(4)-C(3)-Sn	73.11(14)	C(5)-Sn-C(3)	56.46(9)
C(2)-C(3)-Sn	72.28(13)	C(2)-Sn-C(4)	56.63(9)
C(8)-C(3)-Sn	125.31(18)	C(1)-Sn-C(4)	56.64(8)
C(3)-C(4)-C(5)	108.0(2)	C(5)-Sn-C(4)	33.98(9)
C(3)-C(4)-C(9)	126.1(3)	C(3)-Sn-C(4)	33.94(9)

51
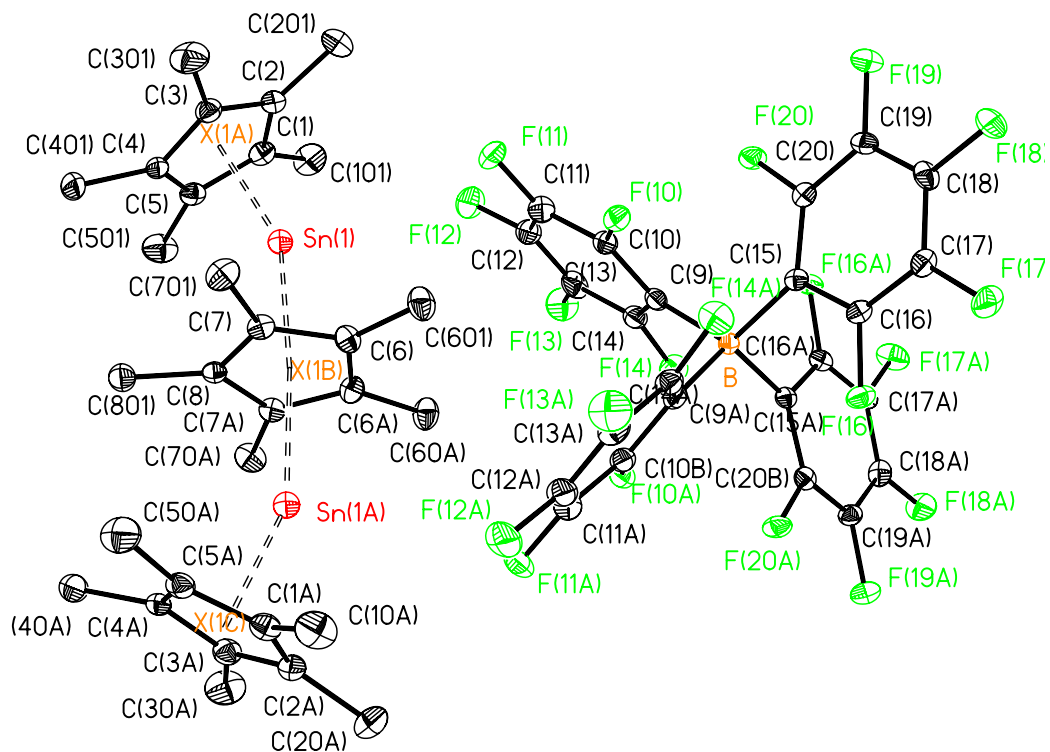
Figure 1.23 Molecular structure of [(η^5-C$_5$Me$_5$)$_3$Sn$_2$][B(C$_6$F$_5$)$_4$], [3][B(C$_6$F$_5$)$_4$], showing the atom numbering scheme. The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.

Figure 1.24 Molecular structure of $[(\eta^5\text{-}C_5Me_5)_3Sn_2]^+$, 3^+, showing the atom numbering scheme for the cation. The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.

Table 1.9 Crystal data and structure refinement for [(η^5-C$_5$Me$_5$)$_3$Sn$_2$][B(C$_6$F$_5$)$_4$], [3][B(C$_6$F$_5$)$_4$].

Identification code	shelx14
Empirical formula	C54 H45 B F20 Sn2
Formula weight	1322.09
Temperature	153(2) K
Wavelength	0.71069 Å
Crystal system	Monoclinic
Space group	C2/c
Unit cell dimensions	a = 22.138(5) Å α= 90°
	b = 14.955(5) Å β= 114.259(5)°
	c = 17.097(5) Å γ = 90°
Volume	5161(3) Å^3
Z	4
Density (calculated)	1.702 Mg/m^3
Absorption coefficient	1.078 mm^{-1}
F(000)	2616
Crystal size	0.20 x 0.20 x 0.20 mm^3
Theta range for data collection	3.02 to 27.44°.
Index ranges	-28≤h≤28, -19≤k≤19, -22≤l≤17
Reflections collected	26112
Independent reflections	5869 [R(int) = 0.0383]
Completeness to theta = 27.44°	99.4 %
Absorption correction	None
Max. and min. transmission	0.8133 and 0.8133
Refinement method	Full-matrix least-squares on F^2
Goodness-of-fit on F^2	1.077
Final R indices [I>2sigma(I)]	R1 = 0.0690, wR2 = 0.1533
R indices (all data)	R1 = 0.0837, wR2 = 0.1597
Largest diff. peak and hole	3.251 and -4.702 e.Å^{-3}

Table 1.10 Selected Bond Lengths (Å) for $[(\eta^5\text{-}C_5Me_5)_3Sn_2][B(C_6F_5)_4]$, $[3][B(C_6F_5)_4]$.

C(1)-C(2)	1.414(9)		C(4)-Sn(1)	2.505(6)
C(1)-C(5)	1.432(9)		C(5)-C(10)	1.514(9)
C(1)-C(6)	1.490(9)		C(5)-Sn(1)	2.522(6)
C(1)-Sn(1)	2.582(6)		C(11)-C(11A)	1.423(12)
C(2)-C(3)	1.430(9)		C(11)-C(12)	1.427(8)
C(2)-C(7)	1.502(9)		C(11)-C(14)	1.506(9)
C(2)-Sn(1)	2.580(6)		C(12)-C(13)	1.430(7)
C(3)-C(4)	1.428(9)		C(12)-C(16A)	1.508(8)
C(3)-C(8)	1.504(9)		C(13)-C(12A)	1.430(7)
C(3)-Sn(1)	2.511(6)		C(13)-C(15)	1.498(12)
C(4)-C(5)	1.419(9)		C(16)-C(12A)	1.508(8)
C(4)-C(9)	1.505(9)			

Table 1.11 Selected Bond Angled (°) for $[(\eta^5\text{-}C_5Me_5)_3Sn_2][B(C_6F_5)_4]$, $[3][B(C_6F_5)_4]$.

C(2)-C(1)-C(5)	108.1(5)		C(4)-C(5)-C(10)	126.5(6)
C(2)-C(1)-C(6)	126.0(6)		C(1)-C(5)-C(10)	125.4(6)
C(5)-C(1)-C(6)	125.8(6)		C(4)-C(5)-Sn(1)	73.0(3)
C(2)-C(1)-Sn(1)	74.1(3)		C(1)-C(5)-Sn(1)	76.0(3)
C(5)-C(1)-Sn(1)	71.4(3)		C(10)-C(5)-Sn(1)	120.7(4)
C(6)-C(1)-Sn(1)	123.1(4)		C(11A)-C(11)-C(12)	107.9(3)
C(1)-C(2)-C(3)	108.0(5)		C(11A)-C(11)-C(14)	126.7(4)
C(1)-C(2)-C(7)	126.1(6)		C(12)-C(11)-C(14)	125.3(5)
C(3)-C(2)-C(7)	125.7(6)		C(11)-C(12)-C(13)	108.4(5)
C(1)-C(2)-Sn(1)	74.1(3)		C(11)-C(12)-C(16A)	126.4(5)
C(3)-C(2)-Sn(1)	71.0(3)		C(13)-C(12)-C(16A)	125.2(6)
C(7)-C(2)-Sn(1)	124.1(5)		C(12)-C(13)-C(12A)	107.2(7)
C(4)-C(3)-C(2)	107.9(5)		C(12)-C(13)-C(15)	126.4(3)
C(4)-C(3)-C(8)	125.3(6)		C(12A)-C(13)-C(15)	126.4(3)
C(2)-C(3)-C(8)	126.3(6)		C(4)-Sn(1)-C(3)	33.1(2)
C(4)-C(3)-Sn(1)	73.3(3)		C(4)-Sn(1)-C(5)	32.8(2)
C(2)-C(3)-Sn(1)	76.4(3)		C(3)-Sn(1)-C(5)	54.4(2)
C(8)-C(3)-Sn(1)	123.0(5)		C(4)-Sn(1)-C(2)	54.00(19)
C(5)-C(4)-C(3)	107.9(5)		C(3)-Sn(1)-C(2)	32.6(2)
C(5)-C(4)-C(9)	126.1(6)		C(5)-Sn(1)-C(2)	53.7(2)
C(3)-C(4)-C(9)	125.3(6)		C(4)-Sn(1)-C(1)	53.89(19)
C(5)-C(4)-Sn(1)	74.2(3)		C(3)-Sn(1)-C(1)	53.7(2)
C(3)-C(4)-Sn(1)	73.7(3)		C(5)-Sn(1)-C(1)	32.6(2)
C(9)-C(4)-Sn(1)	124.9(4)		C(2)-Sn(1)-C(1)	31.8(2)
C(4)-C(5)-C(1)	108.0(5)			

56
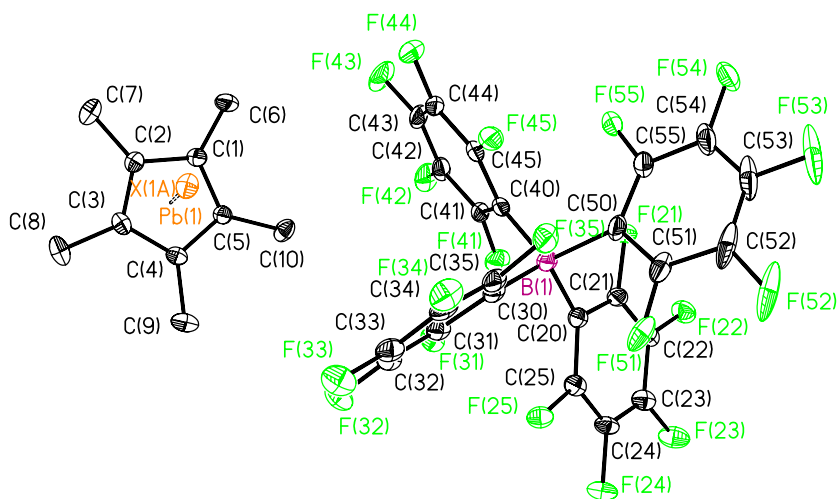
Figure 1.25 Molecular structure of $[(\eta^5\text{-}C_5Me_5)Pb][B(C_6F_5)_4]$, [4][$B(C_6F_5)_4$], showing the atom numbering scheme. The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.

Figure 1.26 Molecular structure of $[(\eta^5\text{-}C_5Me_5)Pb]^+$, 4^+, showing the atom numbering scheme for the cation. The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.

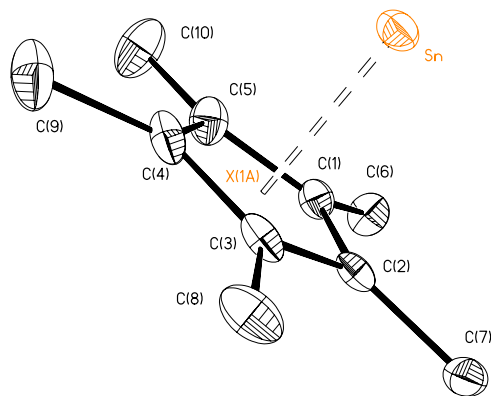**Table 1.12 Crystal data and structure refinement for [(η^5-C$_5$Me$_5$)Pb][B(C$_6$F$_5$)$_4$], [4][B(C$_6$F$_5$)$_4$].**

Identification code	p1bar
Empirical formula	C34 H15 B F20 Pb
Formula weight	1021.46
Temperature	153(2) K
Wavelength	0.71069 Å
Crystal system	Triclinic
Space group	P-1
Unit cell dimensions	a = 11.642(5) Å α= 64.003(5)°
	b = 11.749(5) Å β= 77.837(5)°
	c = 13.392(5) Å γ = 85.223(5)°
Volume	1609.4(11) Å^3
Z	2
Density (calculated)	2.108 Mg/m^3
Absorption coefficient	5.386 mm^{-1}
F(000)	972
Crystal size	0.20 x 0.20 x 0.20 mm^3
Theta range for data collection	3.09 to 27.39°.
Index ranges	-15≤h≤12, -15≤k≤14, -17≤l≤17
Reflections collected	12396
Independent reflections	7270 [R(int) = 0.0171]
Completeness to theta = 27.39°	99.3 %
Absorption correction	None
Max. and min. transmission	0.4122 and 0.4122
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	7270 / 0 / 565
Goodness-of-fit on F^2	1.046
Final R indices [I>2sigma(I)]	R1 = 0.0349, wR2 = 0.0757
R indices (all data)	R1 = 0.0391, wR2 = 0.0780
Largest diff. peak and hole	2.710 and -2.177 e.Å^{-3}

Table 1.13 Selected Bond Lengths (Å) for $[(\eta^5\text{-}C_5Me_5)Pb][B(C_6F_5)_4]$, $[4][B(C_6F_5)_4]$.

C(1)-C(5)	1.427(6)	C(3)-C(8)	1.499(6)
C(1)-C(2)	1.428(6)	C(3)-Pb(1)	2.546(4)
C(1)-C(6)	1.504(6)	C(4)-C(5)	1.424(6)
C(1)-Pb(1)	2.539(4)	C(4)-C(9)	1.508(6)
C(2)-C(3)	1.435(6)	C(4)-Pb(1)	2.557(4)
C(2)-C(7)	1.503(6)	C(5)-C(10)	1.505(6)
C(2)-Pb(1)	2.538(4)	C(5)-Pb(1)	2.559(4)
C(3)-C(4)	1.428(6)		

Table 1.14 Selected Angles (°) for $[(\eta^5\text{-}C_5Me_5)Pb][B(C_6F_5)_4]$, $[4][B(C_6F_5)_4]$.

C(5)-C(1)-C(2)	108.1(3)	C(3)-C(4)-C(9)	126.9(4)
C(5)-C(1)-C(6)	126.8(4)	C(5)-C(4)-Pb(1)	73.9(2)
C(2)-C(1)-C(6)	124.7(4)	C(3)-C(4)-Pb(1)	73.3(2)
C(5)-C(1)-Pb(1)	74.5(2)	C(9)-C(4)-Pb(1)	123.5(3)
C(2)-C(1)-Pb(1)	73.6(2)	C(4)-C(5)-C(1)	108.1(4)
C(6)-C(1)-Pb(1)	123.2(3)	C(4)-C(5)-C(10)	125.8(4)
C(1)-C(2)-C(3)	107.8(4)	C(1)-C(5)-C(10)	125.7(4)
C(1)-C(2)-C(7)	124.4(4)	C(4)-C(5)-Pb(1)	73.8(2)
C(3)-C(2)-C(7)	127.4(4)	C(1)-C(5)-Pb(1)	73.0(2)
C(1)-C(2)-Pb(1)	73.7(2)	C(10)-C(5)-Pb(1)	124.8(3)
C(3)-C(2)-Pb(1)	73.9(2)	C(2)-Pb(1)-C(1)	32.66(13)
C(7)-C(2)-Pb(1)	123.5(3)	C(2)-Pb(1)-C(3)	32.80(14)
C(4)-C(3)-C(2)	107.8(4)	C(1)-Pb(1)-C(3)	54.13(13)
C(4)-C(3)-C(8)	125.8(5)	C(2)-Pb(1)-C(4)	53.99(13)
C(2)-C(3)-C(8)	126.2(5)	C(1)-Pb(1)-C(4)	53.85(13)
C(4)-C(3)-Pb(1)	74.2(2)	C(3)-Pb(1)-C(4)	32.51(14)
C(2)-C(3)-Pb(1)	73.3(2)	C(2)-Pb(1)-C(5)	53.92(13)
C(8)-C(3)-Pb(1)	122.8(4)	C(1)-Pb(1)-C(5)	32.50(12)
C(5)-C(4)-C(3)	108.2(4)	C(3)-Pb(1)-C(5)	53.83(13)
C(5)-C(4)-C(9)	124.6(4)	C(4)-Pb(1)-C(5)	32.32(13)

59
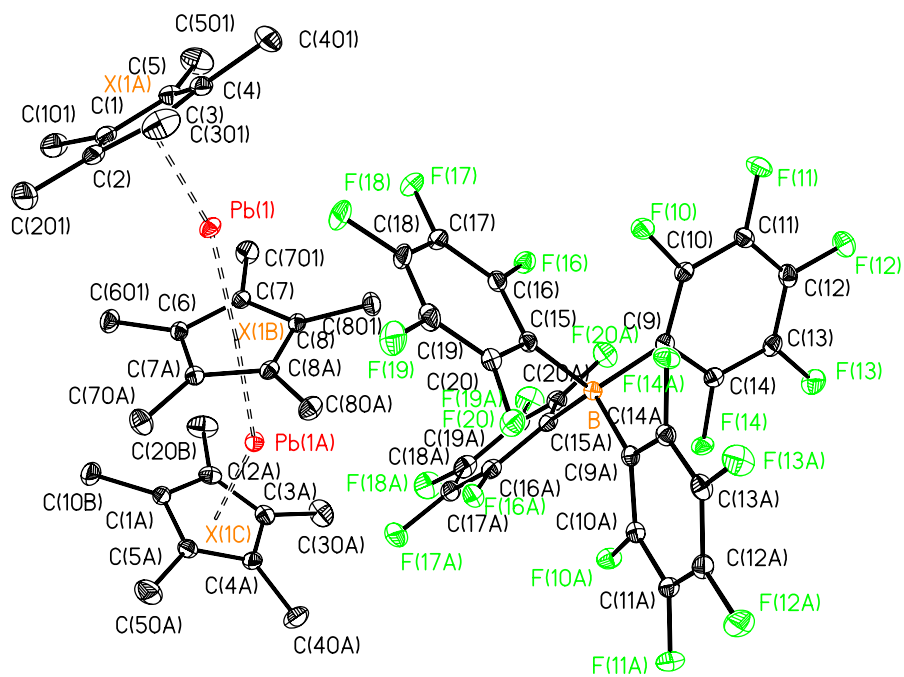
Figure 1.27 Molecular structure of $[(\eta^5\text{-}C_5Me_5)_3Pb_2][B(C_6F_5)_4]$, $[5][B(C_6F_5)_4]$, showing the atom numbering scheme. The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.
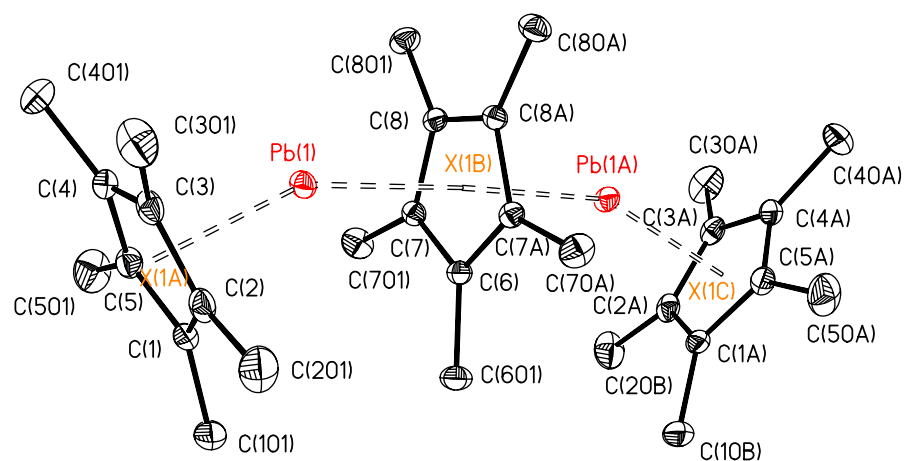
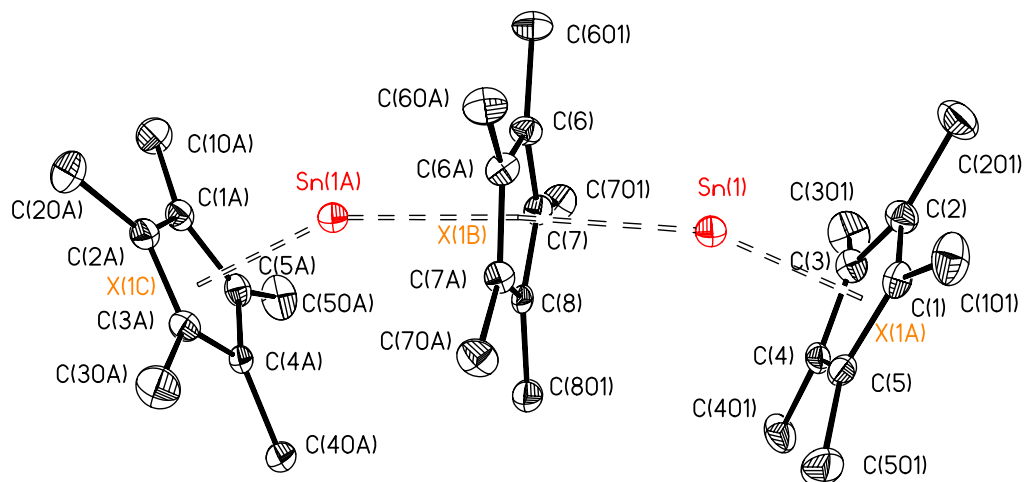**Figure 1.28 Molecular structure of $[(\eta^5\text{-}C_5Me_5)_3Pb_2]^+$, 5^+, showing the atom numbering scheme for the cation. The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.**

Table 1.15 Crystal data and structure refinement for [(η^5-C$_5$Me$_5$)$_3$Pb$_2$][B(C$_6$F$_5$)$_4$], [5][B(C$_6$F$_5$)$_4$].

Identification code	i2overa
Empirical formula	C54 H46 B F20 Pb2
Formula weight	1500.10
Temperature	293(2) K
Wavelength	0.71073 Å
Crystal system	Monoclinic
Space group	I2/a
Unit cell dimensions	a = 17.087(3) Å α = 90°
	b = 14.773(3) Å β = 110.37(3)°
	c = 21.811(4) Å γ = 90°
Volume	5161.6(18) Å^3
Z	4
Density (calculated)	1.930 Mg/m^3
Absorption coefficient	6.625 mm^{-1}
F(000)	2876
Crystal size	0.2 x 0.2 x 0.2 mm^3
Theta range for data collection	3.47 to 27.48°.
Index ranges	-22<=h<=22, -19<=k<=19, -28<=l<=28
Reflections collected	10513
Independent reflections	5852 [R(int) = 0.0148]
Completeness to theta = 27.48°	98.6 %
Absorption correction	None
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	5852 / 0 / 438
Goodness-of-fit on F^2	1.071
Final R indices [I>2sigma(I)]	R1 = 0.0198, wR2 = 0.0454
R indices (all data)	R1 = 0.0225, wR2 = 0.0465
Largest diff. peak and hole	0.706 and -1.188 e.Å^{-3}

Table 1.16 Selected Bond Lengths (Å) for $[(\eta^5\text{-}C_5Me_5)_3Pb_2][B(C_6F_5)_4]$, $[5][B(C_6F_5)_4]$.

C(1)-C(2)	1.423(4)	C(401)-C(4)	1.499(4)
C(2)-C(3)	1.425(4)	C(501)-C(5)	1.504(4)
C(3)-C(4)	1.411(4)	C(6)- C(601)	1.499(5)
C(4)-C(5)	1.422(4)	C(7)- C(701)	1.494(4)
C(1)-C(5)	1.421(4)	C(8)- C(801)	1.508(3)
C(6)-C(7)	1.425(3)	C(1)-Pb(1)	2.592(2)
C(7)- C(8)	1.429(3)	C(2)-Pb(1)	2.620(3)
C(101)-C(1)	1.505(4)	C(3)-Pb(1)	2.673(3)
C(201)-C(2)	1.504(4)	C(4)-Pb(1)	2.672(3)
C(301)-C(3)	1.498(4)	C(5)-Pb(1)	2.605(2)

Table 1.17 Selected Bond Angles (°) for [(η^5-C$_5$Me$_5$)$_3$Pb$_2$][B(C$_6$F$_5$)$_4$], [5][B(C$_6$F$_5$)$_4$].

C(7)-C(6)-C(7A)	108.3(3)	C(201)-C(2)-Pb(1)	120.7(2)
C(7)-C(6)-C(601)	125.86(14)	C(4)-C(3)-C(2)	108.3(2)
C(7A)-C(6)-C(601)	125.86(14)	C(4)-C(3)-C(301)	126.5(3)
C(6)-C(7)-C(8)	107.8(2)	C(2)-C(3)-C(301)	125.0(3)
C(6)-C(7)-C(701)	126.0(2)	C(4)-C(3)-Pb(1)	74.65(15)
C(8)-C(7)-C(701)	126.2(2)	C(2)-C(3)-Pb(1)	72.31(13)
C(8A)-C(8)-C(7)	108.07(13)	C(301)-C(3)-Pb(1)	123.4(2)
C(8A)-C(8)-C(801)	127.11(15)	C(3)-C(4)-C(5)	108.1(2)
C(7)-C(8)-C(801)	124.8(2)	C(3)-C(4)-C(401)	126.1(3)
C(5)-C(1)-C(2)	107.9(2)	C(5)-C(4)-C(401)	125.6(3)
C(5)-C(1)-C(101)	125.0(3)	C(3)-C(4)-Pb(1)	74.75(14)
C(2)-C(1)-C(101)	126.4(3)	C(5)-C(4)-Pb(1)	71.76(14)
C(5)-C(1)-Pb(1)	74.63(14)	C(401)-C(4)-Pb(1)	123.6(2)
C(2)-C(1)-Pb(1)	75.25(14)	C(1)-C(5)-C(4)	108.0(2)
C(101)-C(1)-Pb(1)	123.36(19)	C(1)-C(5)-C(501)	125.5(3)
C(1)-C(2)-C(3)	107.7(2)	C(4)-C(5)-C(501)	126.1(3)
C(1)-C(2)-C(201)	126.3(3)	C(1)-C(5)-Pb(1)	73.62(14)
C(3)-C(2)-C(201)	125.8(3)	C(4)-C(5)-Pb(1)	77.00(14)
C(1)-C(2)-Pb(1)	73.07(14)	C(501)-C(5)-Pb(1)	121.6(2)
C(3)-C(2)-Pb(1)	76.46(14)	C(1)-Pb(1)-C(5)	31.75(8)
C(201)-C(2)-Pb(1)	120.7(2)	C(1)-Pb(1)-C(2)	31.68(8)
C(4)-C(3)-C(2)	108.3(2)	C(5)-Pb(1)-C(2)	52.23(8)
C(4)-C(3)-C(301)	126.5(3)	C(1)-Pb(1)-C(4)	51.79(8)
C(2)-C(3)-C(301)	125.0(3)	C(201)-C(2)-Pb(1)	120.7(2)
C(4)-C(3)-Pb(1)	74.65(15)		
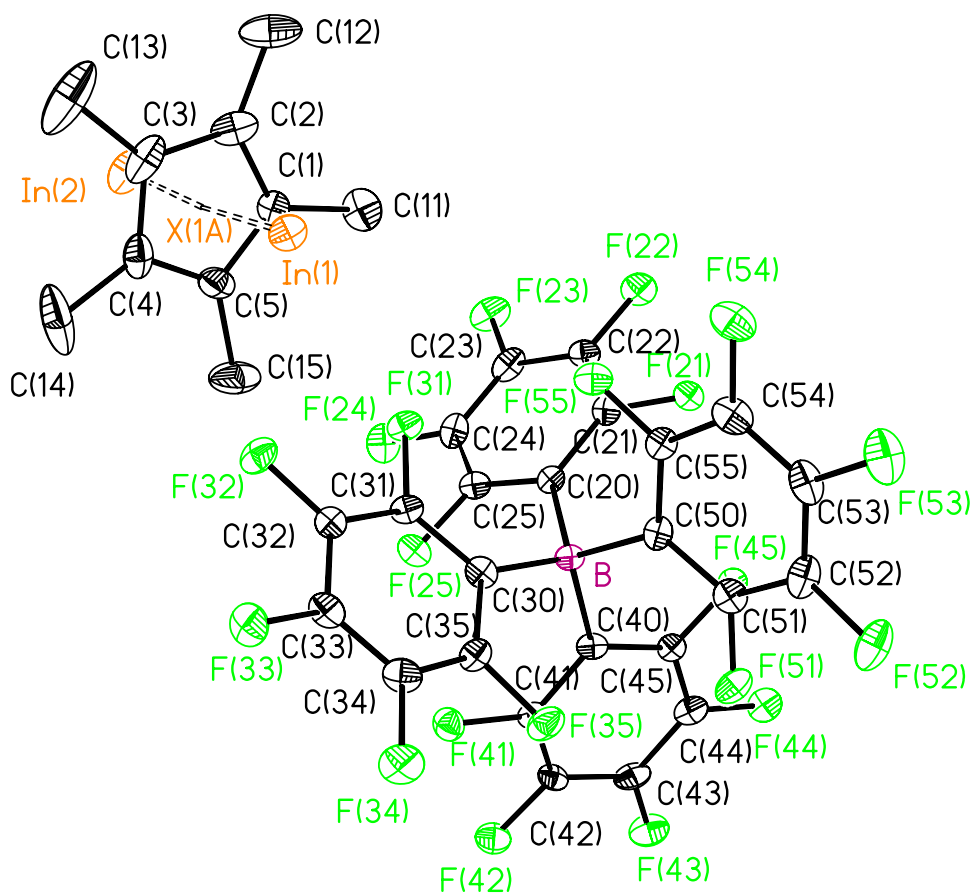
Figure 1.29 Molecular structure of $[In(\eta^5\text{-}\mu\text{-}C_5Me_5)In][B(C_6F_5)_4]$, $[6][B(C_6F_5)_4]$, showing the atom numbering scheme. The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.
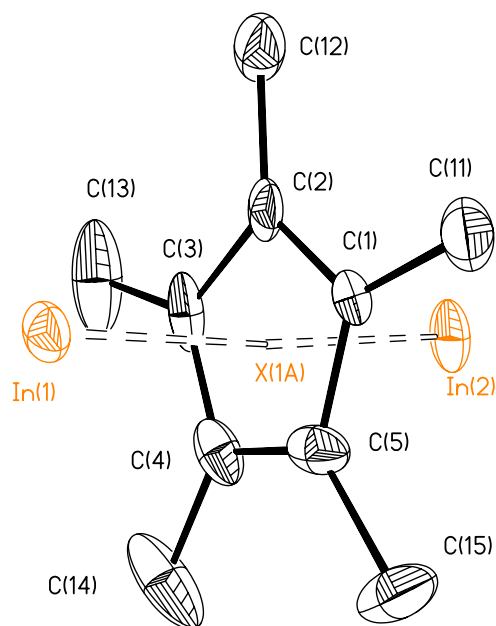
Figure 1.30 Molecular structure of $[In[(\eta^5\text{-}\mu\text{-}C_5Me_5)In]^+$, 6^+, showing the atom numbering scheme. The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.
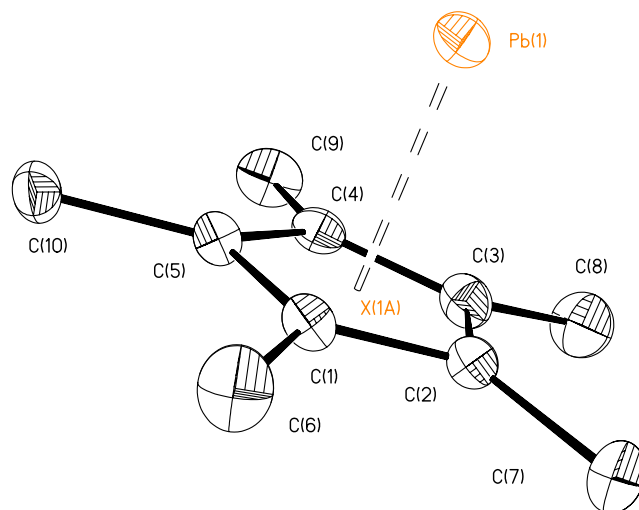
Table 1.18 **Crystal data and structure refinement for [(μ-C₅Me₅)₂In][B(C₆F₅)₄], [6][B(C₆F₅)₄].**

Identification code	p21c
Empirical formula	C34 H15 B F20 In2
Formula weight	1043.91
Temperature	153(2) K
Wavelength	0.71073 Å
Crystal system	Monoclinic
Space group	P2₁/c
Unit cell dimensions	a = 10.065(2) Å α = 90°
	b = 23.157(5) Å β = 103.48(3)°
	c = 16.451(3) Å γ = 90°
Volume	3728.7(13) Å³
Z	4
Density (calculated)	1.860 Mg/m³
Absorption coefficient	1.362 mm⁻¹
F(000)	2008
Crystal size	0.2 x 0.2 x 0.2 mm³
Theta range for data collection	3.65 to 27.50°.
Index ranges	-12≤h≤12, -27≤k≤29, -21≤l≤21
Reflections collected	14935
Independent reflections	8441 [R(int) = 0.0400]
Completeness to theta = 27.50°	98.7 %
Absorption correction	None
Refinement method	Full-matrix least-squares on F²
Data / restraints / parameters	8441 / 0 / 538
Goodness-of-fit on F²	1.032
Final R indices [I>2sigma(I)]	R1 = 0.0495, wR2 = 0.1140
R indices (all data)	R1 = 0.0906, wR2 = 0.1269
Largest diff. peak and hole	0.610 and -0.746 e.Å⁻³

Table 1.19 Selected Bond Lengths (Å) for [(μ-C₅Me₅)₂In][B(C₆F₅)₄], [6][B(C₆F₅)₄].

C(11)-C(1)	1.512(7)	C(3)-C(13)	1.514(8)
C(14)-C(4)	1.543(8)	C(4)-C(5)	1.459(8)
C(1)-C(2)	1.406(7)	C(4)-In(2)	2.675(5)
C(1)-C(5)	1.414(7)	C(5)-C(15)	1.509(9)
C(2)-C(3)	1.388(8)	C(5)-In(2)	2.687(5)
C(2)-C(12)	1.529(8)	C(5)-In(1)	2.701(4)
C(3)-C(4)	1.411(8)		

Table 1.20 Selected Bond Angles (°) for [(μ-C₅Me₅)₂In][B(C₆F₅)₄], [6][B(C₆F₅)₄].

C(2)-C(1)-C(5)	108.1(5)	C(5)-C(4)-In(2)	74.7(3)
C(2)-C(1)-C(11)	126.1(5)	C(14)-C(4)-In(2)	115.0(4)
C(5)-C(1)-C(11)	125.8(5)	C(1)-C(5)-C(4)	105.9(4)
C(3)-C(2)-C(1)	110.4(5)	C(1)-C(5)-C(15)	126.5(6)
C(3)-C(2)-C(12)	124.4(5)	C(4)-C(5)-C(15)	127.6(6)
C(1)-C(2)-C(12)	125.1(5)	C(1)-C(5)-In(2)	76.2(3)
C(2)-C(3)-C(4)	107.1(5)	C(4)-C(5)-In(2)	73.8(3)
C(2)-C(3)-C(13)	124.9(6)	C(15)-C(5)-In(2)	115.4(4)
C(4)-C(3)-C(13)	128.1(6)	C(1)-C(5)-In(1)	76.0(3)
C(3)-C(4)-C(5)	108.5(4)	C(4)-C(5)-In(1)	75.5(3)
C(3)-C(4)-C(14)	125.2(7)	C(15)-C(5)-In(1)	114.2(4)
C(5)-C(4)-C(14)	126.3(7)	In(2)-C(5)-In(1)	130.4(2)
C(3)-C(4)-In(2)	77.8(3)	C(4)-In(2)-C(5)	31.58(18)
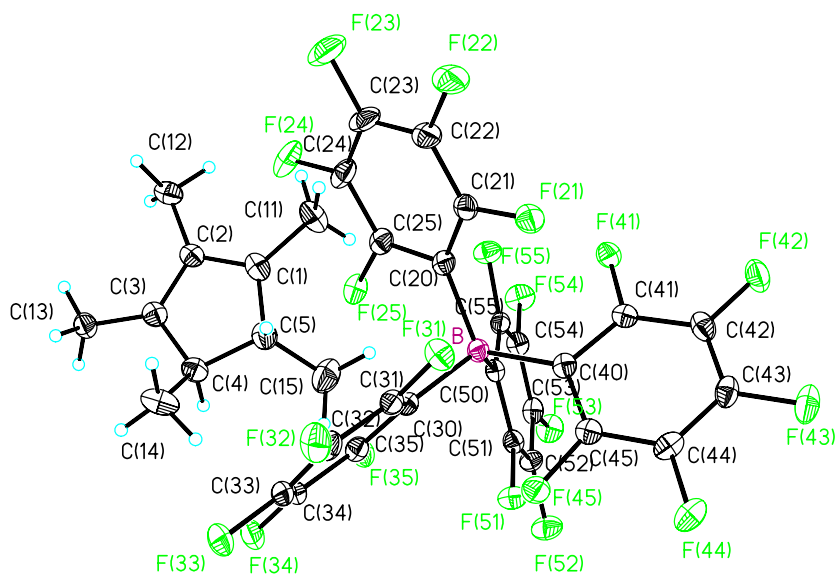
Figure 1.31 Molecular structure of $[(C_5Me_5)H_2][B(C_6F_5)_4]$, $[8][B(C_6F_5)_4]$, showing the atom numbering scheme. The thermal ellipsoids are shown at the 30% probability level.
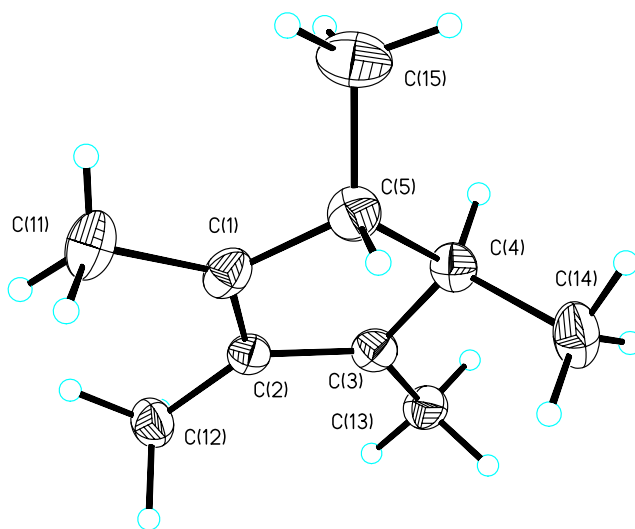
Figure 1.32 Molecular structure of $[(C_5Me_5)H_2]^+$, 8^+, showing the atom numbering scheme. The thermal ellipsoids are shown at the 30% probability level.

69

Table 1.21 Crystal data and structure refinement for [(C$_5$Me$_5$)H$_2$][B(C$_6$F$_5$)$_4$], [8][B(C$_6$F$_5$)$_4$].

Identification code	p21overc
Empirical formula	C24 H17 B F20
Formula weight	696.19
Temperature	153(2) K
Wavelength	0.71073 Å
Crystal system	Monoclinic
Space group	P2$_1$/c
Unit cell dimensions	a = 13.197(3) Å α = 90°
	b = 13.752(3) Å β = 90.11(3)°.
	c = 17.407(4) Å γ = 90°
Volume	3158.9(11) Å^3
Z	4
Density (calculated)	1.464 Mg/m^3
Absorption coefficient	0.168 mm^{-1}
F(000)	1384
Crystal size	0.20 x 0.20 x 0.20 mm^3
Theta range for data collection	2.96 to 27.50°.
Index ranges	-17≤h≤17, -17≤k≤17, -22≤l≤22
Reflections collected	28166
Independent reflections	7256 [R(int) = 0.1108]
Completeness to theta = 27.50°	99.8 %
Absorption correction	None
Max. and min. transmission	0.9673 and 0.9673
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	7256 / 0 / 564
Goodness-of-fit on F^2	1.038
Final R indices [I>2sigma(I)]	R1 = 0.0669, wR2 = 0.1394
R indices (all data)	R1 = 0.1382, wR2 = 0.1650
Largest diff. peak and hole	0.292 and -0.382 e.Å^{-3}

Table 1.22 Selected Bond Lengths (Å) for [(C₅Me₅)H₂][B(C₆F₅)₄], [8][B(C₆F₅)₄].

C(3)-C(2)	1.383(5)		C(4)-C(5)	1.520(5)
C(3)-C(13)	1.470(5)		C(4)-C(14)	1.523(5)
C(3)-C(4)	1.485(5)		C(1)-C(5)	1.463(5)
C(2)-C(1)	1.392(5)		C(1)-C(11)	1.484(5)
C(2)-C(12)	1.494(5)		C(5)-C(15)	1.520(6)

Table 1.23 Selected Bond Angles (°) for [(C₅Me₅)H₂][B(C₆F₅)₄], [8][B(C₆F₅)₄].

C(2)-C(3)-C(13)	125.8(3)		C(5)-C(4)-C(14)	114.0(3)
C(2)-C(3)-C(4)	112.5(3)		C(2)-C(1)-C(5)	111.5(3)
C(13)-C(3)-C(4)	121.6(3)		C(2)-C(1)-C(11)	124.9(4)
C(3)-C(2)-C(1)	107.6(3)		C(5)-C(1)-C(11)	123.6(4)
C(3)-C(2)-C(12)	126.8(3)		C(1)-C(5)-C(15)	116.1(4)
C(1)-C(2)-C(12)	125.5(3)		C(1)-C(5)-C(4)	105.5(3)
C(3)-C(4)-C(5)	102.8(3)		C(15)-C(5)-C(4)	116.1(4)
C(3)-C(4)-C(14)	113.9(4)			

REFERENCES

1. Jutzi, P.; Reumann, G. J., *J. Chem. Soc., Dalt. Trans.* **2000,** 2237.
2. Shapiro, P. J., *Coordin. Chem. Rev.* **1999,** *189,* 1.
3. Jutzi, P.; Burford, N., *Chem. Rev.* **1999,** *99,* 969.
4. Miller, S. A.; Tebboth, J. A.; Tremaine, J. F., *J. Chem. Soc.* **1952,** 632.
5. Kealy, T. J.; Pauson, P. L., *Nature (London)* **1951,** *168,* 1039.
6. Fischer, E. O.; Pfab, W., *Z. Naturforsch. B* **1952,** *7,* 377.
7. Dunitz, J. D.; Orgel, L. E.; Rich, A., *Acta Cryst.* **1956,** *9,* 373.
8. Jutzi, P.; Kohl, F.; Hofmann, P.; Krüger, C.; Tsay, Y.-H., *Chem. Ber.* **1980,** *113,* 757.
9. Atwood, J. L.; Hunter, W. E., *J. Chem. Soc., Chem. Comm.* **1981,** 925.
10. Werner, H.; Salzer, A., *Synthetic Inorganic and Metal Organic Chemistry* **1972,** *2,* 239.
11. Schneider, J. J.; Goddard, R.; Werner, S.; Krüger, C., *Angew. Chem., Int. Ed., Engl.* **1991,** *30,* 1124.
12. Lumme, P. O.; Turpeinen, U.; Kudinov, A. R.; Rybinskaya, M. I., *Acta Cryst., C* **1990,** *46,* 1410.
13. *Metalloorg. Khim (Russ.)* **1992,** *5,* 1102.
14. Armstrong, D. R.; Edwards, A. J.; Moncrieff, D.; Paver, M. A.; Raithby, P. R.; Rennie, M.-A.; Russel, C. A.; Wright, D. S., *J. Chem. Soc., Chem. Comm.* **1995,** 927.
15. Harder, S.; Prosenc, M. H., *Angew. Chem., Int. Ed., Engl.* **1996,** *35,* 97.
16. Wojtczak, W. A.; Fleig, P. F.; Hampden-Smith, M. J., *J. Adv. Organomet. Chem.* **1996,** *40,* 215.
17. Burns, C. T.; Shapiro, P. J.; Budzelaar, P. H. M.; Vij, A., *Organometallics* **2000,** *19,* 3361.
18. Bochmann, M.; Dawson, D. M., *Angew. Chem., Int. Ed., Engl.* **1996,** *35,* 2226.
19. Beswick, M. A.; Gornitzka, H.; Karcher, J.; Mosquera, M. E. G.; Palmer, J. S.; Raithby, P. R.; Russell, C. A.; Stalke, D.; Steiner, A.; Wright, D. S., *Organometallics* **1999,** *18,* 1148.
20. Beswick, M. A.; Palmer, J. S.; Wright, D. S., *Chem. Soc. Rev.* **1998,** *27,* 225.
21. Cowley, A. H.; Macdonald, C. L. B.; Silverman, J. S.; Gorden, J. D.; Voigt, A., *Chem. Comm.* **2001,** 175.
22. Massey, A. G.; Park, A. J.; Stone, F. G. A., *Proc. Chem. Soc.* **1963,** *July,* 212.
23. Piers, W. E.; Chivers, T., *Chem. Soc. Rev.* **1997,** *26,* 345.

24. Doerrer, L. H.; Green, M. L. H., *J. Chem. Soc., Dalt. Trans.* **1999,** 4325.
25. Winter, J. G.; Portius, P.; Kociok-Kohn, G.; Steck, R.; Filippou, A. C., *Organometallics* **1998,** *17,* 4176.
26. Bondi, A., *J. Phys. Chem.* **1964,** *68,* 441.
27. Rhodes, B.; Chien, J. C. W.; Rausch, M. D., *Organometallics* **1998,** *17,* 1931.
28. Jutzi, P.; Dickbreder, R.; Nöth, H., *Chem. Ber.* **1989,** *122,* 865.
29. Beachley Jr., O. T.; Blom, R.; Churchill, M. R.; Faegri Jr., K. J.; Fettinger, C.; Pazik, J. C.; Victoriano, L., *Organometallics* **1989,** *8,* 346.
30. Schluter, R. D.; Cowley, A. H.; Atwood, D. A.; Jones, R. A.; Bond, M. R.; Carrano, C. J., *J. Am. Chem. Soc.* **1993,** *115,* 2070.
31. Campbell, P. H.; Chiu, N. W. K.; Deugau, K.; Miller, I. J.; Sorensen, T. S., *J. Am. Chem. Soc.* **1969,** *91,* 6404.
32. Lambert, J. B.; Lin, L.; Rassolov, V., *Angew. Chem., Int. Ed., Engl.* **2002,** *41,* 1429.
33. Otto, M.; Scheschkewitz, D.; Kato, T.; Midland, M. M.; Lambert, J. B.; Bertrand, G., *Angew. Chem., Int. Ed., Engl.* **2002,** *41,* 2275.
34. Jones, J. N.; Cowley, A. H.; Macdonald, C. L. B., *Chem. Comm.* **2002,** *14,* 1520.
35. Jutzi, P.; Seufert, A., *J. Organomet. Chem.* **1978,** *161,* C5.
36. Jutzi, P.; Kohl, F., *J. Organomet. Chem.* **1979,** *164,* 141.
37. Sheldrick, G. M., *Shelxtl Pc*. 1994, Siemens Analytical X-ray Instruments, Inc.

CHAPTER 2: GROUP 14 INDENYL-CONTAINING

METALLOCENES

INTRODUCTION

Similar to the pentamethylcyclopentadienide ligand discussed in Chapter 1, the indenide ligand, $[C_9H_7]^-$ (Figure 2.1), can rearrange in order to compensate for changes in the total number of electrons in the metal coordination sphere (Figure 2.2).[1] Such a rearrangement has been described as a "ring slippage" or "ring-folding" process[2] or by the more general term of "haptotropic shift".[3] Despite the similarities in the pentamethylcyclopentadienide and indenide ligands, the η^5-to-η^3 ring slippage is significantly more facile for the indenide ligand, resulting in faster reaction rates.[4-10] This increase in reaction rate has been termed the *indenyl effect*.[5] The most widely accepted explanation for the increased reaction rate is the enhanced stability of the η^3-C_9H_7 intermediates as a consequence of the aromaticity gain of the benzene ring formed by the six uncoordinated carbon atoms. However, recent work has shown that the more facile η^5-to-η^3 ring slippage by the indenide ligand might be related to the indenyl—metal bond strength, particularly when the η^5-coordination mode is adopted.[1]
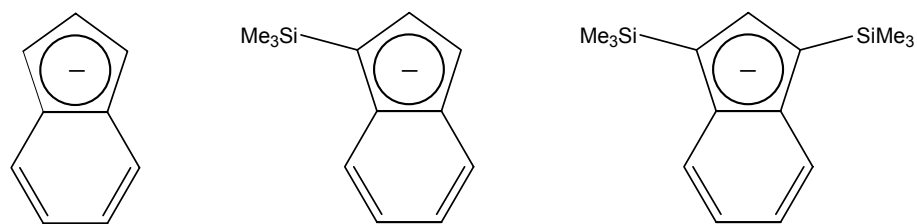
Figure 2.1 General structures of the indenide, 1-trimethylsilylindenide, and 1,3-bis(trimethylsilyl)indenide ligands.
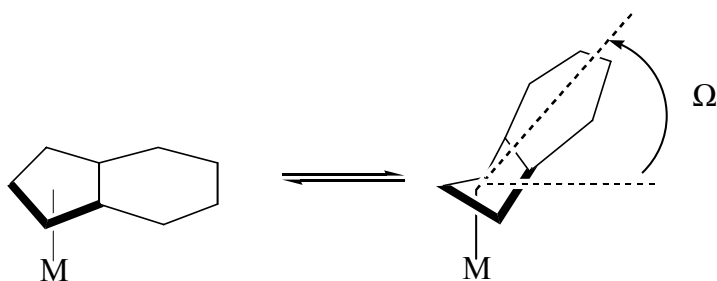
Figure 2.2 The indenyl effect showing the change of coordination from pentahapto to trihapto ligation.

One method for determining the hapticity of indenide complexes is to calculate the slip parameter. The slip parameter, $\Delta_{M\text{-}C}$, is defined as the difference in the average bond length of the metal to the ring junction carbons (C4 and C9) and the average metal—carbon bond length to the adjacent carbons C1 and C3 (Figure 2.3).

$$\Delta_{M-C} = \text{av.}(M - C4, C9) - \text{av.}(M - C1, C3)$$

Equation 2. 1 Slip parameter determination.

For a "true" η^5-coordinated complex, $\Delta_{M-C} = 0.0$ Å, while for η^3-coordinated complexes the range of reported Δ_{M-C} values is 0.69-0.79 Å.
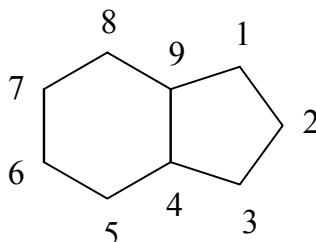
Figure 2.3 Numbering scheme for the indenide ligand.

Mono- or disubstitution of the five-membered ring of the indenide ligand has been shown to impact the bonding flexibility.[11] This is due to the fact that the inductive effect of the trimethylsilyl group ranges from slightly more electron donating than a methyl group to more electron donating than a *tert*-butyl group.[12, 13]

Bis(indenide) complexes can adopt three conformations (Figure 2.4), namely staggered (A), *gauche* (B), and eclipsed (C). In addition to impacting the bonding flexibility of indenide ligands, the use of bulky trimethylsilyl substituents has also been shown to influence these orientation of the ligands in bis(π-indenide) compounds.[14] Brady *et al.*[14] reported the synthesis and X-ray crystal structures of $[1\text{-}(Me_3Si)(\eta^5\text{-}C_9H_5)]_2Cr$ and $[1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)]_2Cr$. The results were surprising in that the indenide ligands are oriented in a staggered fashion in the monosubstituted complex, $[1\text{-}(Me_3Si)(\eta\text{-}C_9H_5)]_2Cr$, while the

disubstituted complex, $[1,3-(Me_3Si)_2(\eta-C_9H_5)]_2Cr$, adopts a *gauche* conformation. The change from a staggered to a *gauche* conformation lowers the symmetry of the bis(indenide)metal complex and allows greater mixing of metal and ligand orbitals. In turn, this changes the magnetic properties of the complexes: $[1-(Me_3Si)(\eta-C_9H_5)]_2Cr$ is a high-spin complex, while $[1,3-(Me_3Si)_2(\eta-C_9H_5)]_2Cr$ is a low-spin complex.[14] The ability to tailor magnetic properties of metal complexes in this fashion opens opportunities for the design of novel materials.
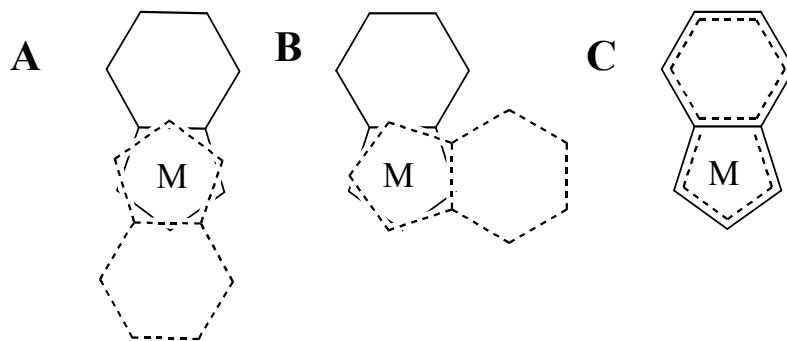
Figure 2.4 Orientations of the indenide ligand in bis(indenyl)metal complexes. A)staggered, B)gauche, and C)eclipsed.

A plethora of bis(π-indenide) transition metal complexes have been reported. The first such complex to be crystallographically charactered was the ferrocene analogue, $(C_9H_7)_2Fe$, in 1953.[15] The rings were predicted to be oriented in a "*trans*" (staggered) configuration; however, a subsequent X-ray diffraction study indicated that in the solid state the ligands are arranged between a *gauche* and an eclipsed conformation. An assortment of ligand arrangements

has been characterized in the solid state structures of transition metal—indenide complexes, including the fully eclipsed bis(indenide)vanadium(II)[16] and bis(indenide)ruthenium(II)[17] complexes and the staggered bis(indenide)nickel(II) complex.[18]

Only two post-transition metal π-indenide compounds have been structurally characterized, namely [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)](η^5-C$_5$Me$_5$)Ge[19] and [1,3-(Me$_3$Si)$_2$(η-C$_9$H$_5$)]$_2$Pb.[20] The germanium atom in the mixed ligand complex, 1,3-(Me$_3$Si)$_2$(C$_9$H$_5$)](C$_5$Me$_5$)Ge, was determined to be η^5-coordinated to both the 1,3-(Me$_3$Si)$_2$(C$_9$H$_5$) and the C$_5$Me$_5$ rings. Interestingly, the germanium—C$_5$Me$_5$ ring centroid distance is significantly shorter than that between the germanium atom and the 1,3-(Me$_3$Si)$_2$(C$_9$H$_5$) ring centroid. This observation implies that the germanium—indenyl bond strength is appreciably weaker than the germanium—C$_5$Me$_5$ bond strength.[19] The lead derivative [1,3-(Me$_3$Si)$_2$(η-C$_9$H$_5$)]$_2$Pb was found to undergo slow decomposition at room temperature; nevertheless, it was possible to collect single crystal X-ray diffraction data. Interesting features of the structure include the rotation of the indenyl ligands by 83.3° and a *gauche* orientation. The ring centroid—lead—ring centroid angle of av. 173.1° is only slightly deviated from linearity.[20]

The present work focuses on the synthesis and structural characterization of group 14 indenide-substituted metallocenes. The study includes both the bis(π-

indenide) and the mixed ligand systems, (π-indenide)(π-C$_5$Me$_5$). These structural data have provided useful insights into the bonding modes of group 14 indenide complexes and shed light on potential nano-materials applications.

RESULTS AND DISCUSSION

Synthesis and Characterization of [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)]Li·2THF, 9

The compound (1,3-trimetylsilane)indenyllithium, **9**, was synthesized in good yield *via* the previously reported[21] method. The reaction of trimethylchlorosilane with [1-(trimethylsilyl)indenyl]lithium was followed by metalation with butyllithium. The ^{1}H NMR spectrum of **9** exhibits four proton signals, indicating C$_{2v}$ symmetry for the indenide ring. The ^{13}C NMR spectrum of **9** evidences six carbon resonances, thus lending further support to the proposed C$_{2v}$ symmetry structure.

Recrystallization of **9** by slow cooling of a concentrated THF solution to -20° C yielded crystals suitable for X-ray diffraction experiments. ^{1}H and ^{13}C NMR analyses of **9** were indicative of the presence of coordinated THF molecules and this was confirmed by X-ray analysis. Analysis of the single crystal X-ray diffraction data confirmed the identity of the compound as [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)]Li·2THF (Figure 2.5). The details of the data collection, structure solution, and refinement are compiled in Table 2.1 and selected metrical parameters are listed in Table 2.2 and Table 2.3.
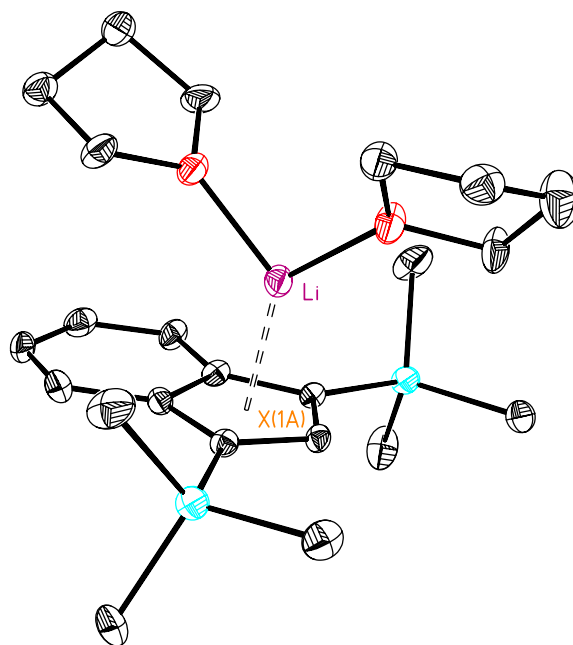
Figure 2.5 Molecular structure of $[1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)]Li\cdot2THF$, **9**, showing the η^5-coordination of the lithium atom to the 1,3-bis(trimethylsilyl)indenide ligand. Atoms are colored as follows: carbon (black), silicon (cyan), lithium (violet), and oxygen (red). Thermal ellipsoids are shown at the 30% probability level.

The $Li\text{—}C_{ring}$ bond distances range from 2.246(9)—2.413(9) Å with an average bond length of 2.344(9) Å. The ring-centroid (X1A)—Li distance is 2.000(4) Å. The close similarity of the $Li\text{—}C_{ring}$ indicated that the lithium atom is pentahapto bonded to the $1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)$ five-membered ring. The slip parameter, $\Delta_{Li\text{-}C}$, is 0.089 Å. The lithium atom is further coordinated to the oxygen atoms of two THF molecules with an average Li—O bond distance of 1.913(8) Å.

Synthesis and Characterization of [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)]$_2$Sn], 10

Bis[1,3-bis(trimethylsilyl)indenyl]tin(II), **10,** was synthesized in good yield *via* the metathetical reaction of tin(II) chloride with two equivalents of [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)]Li in THF solution. The compound decomposes slowly at room temperature over the course of several weeks. The ^{1}H NMR spectrum of **10** exhibited four proton signals thus indicating the equivalence of the two 1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$) ligands. Similarly, the ^{13}C NMR spectrum of **10** shows six carbon resonances, further supporting the argument for the symmetry equivalence of the two 1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$) ligands. The ^{119}Sn NMR spectrum evidences a single peak at δ -2232.

Recrystallization of **10** by slow cooling of a concentrated solution to -20° C yielded crystals suitable for X-ray diffraction experiments. A single-crystal X-ray diffraction study cofirmed the identity of compound **10** as [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)]$_2$Sn]. The molecular structure is illustrated in Figure 2.6. Details of the data collection, structure solution, and refinement are compiled in Table 2.4 and selected metrical parameters are listed in Table 2.5 and Table 2.6.

Figure 2.6   Molecular structure of [1,3-(Me₃Si)₂(η⁵-C₉H₅)]₂Sn, 10, showing the planarity of the 1,3-(Me₃Si)₂(η⁵-C₉H₅) ligands.   Atoms are colored as follows: carbon (black), silicon (cyan), and tin (orange).   Thermal ellipsoids are shown at the 30% probability level.

The Sn—C_{ring} bond distances range from 2.580(3)—2.789(3) Å with an average bond length of 2.432(3) Å. The average ring centroid (X1A/X1B)—Sn bond distance is 2.421(4) Å, which is significantly longer than that reported for Sn(η⁵-C₅Me₅)₂ (2.396 Å).[22]   The bond distances obtained indicate that tin is pentahapto bonded to the five-membered indenyl ring, Δ_{Sn-C} = 0.105 Å.   The $Ind^{2TMS}_{centroid}$ (X1A)—Sn—$Ind^{2TMS}_{centroid}$ (X1B) bond angle is nearly linear

(175.9º) compared to the significantly more acute angle of av. 154.9º for Sn(η^5-C$_5$Me$_5$)$_2$.[22] The linearity of this bond angle is comparable to that determined for

Sn(η^5-C$_5$Ph$_5$)$_2$ (180.0º).[23] The question arises, where is the lone electron pair in

the linear structure?   By means of Fenske-Hall molecular orbital calculations,

Williamson *et al.*[24] determined that the lone pair resides in the tin 5s-like orbital.

The trimethylsilyl groups orient themselves in a *gauche* conformation

which can be seen in Figure 2.7. The two ligands are rotated by approximately

82º relative to each other, which may serve to minimize unfavorable non-bonding

interactions between the indenyl substituents. The *gauche* conformation is similar

to that previous reported for [(Me$_3$Si)$_2$C$_9$H$_5$]$_2$Pb.[20]

ca. 82°
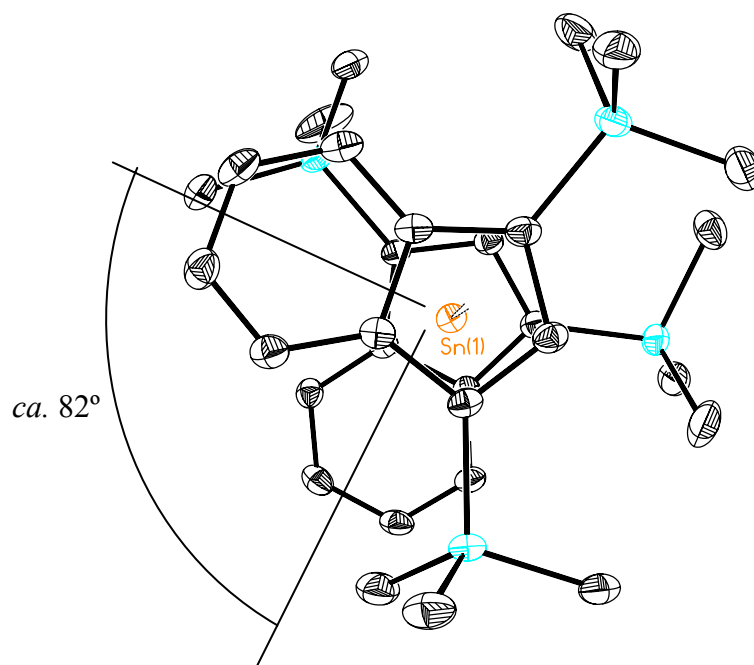
Figure 2.7 Molecular structure of [1,3-(Me₃Si)₂(η⁵-C₉H₅)]₂Sn, 10, showing the *gauche* orientation of the 1,3-(Me₃Si)₂(η⁵-C₉H₅) ligands. Atoms are colored as follows: carbon (black), silicon (cyan), and tin (orange). Thermal ellipsoids are shown at 30% probability.

Synthesis and Characterization of [1,3-(Me₃Si)₂(η⁵-C₉H₅)](η⁵-C₅Me₅)Sn, 11

(Pentamethylcyclopentadienyl)[1,3-bis(trimethylsilyl)indenyl]tin(II), **11,** was synthesized in good yield *via* the metathetical reaction of (pentamethylcyclopentadienyl)tin(II) chloride with an equimolar quantity of [1,3-(Me₃Si)₂(η⁵-C₉H₅)]Li. The product decomposes at room temperature over the course of several days. The ¹H NMR spectrum of **11** exhibits five proton signals. An intense singlet at δ 1.716 indicated the C₅ᵥ symmetry of the

85

pentamethylcyclopentadienide ring, while the other four ^{1}H NMR signals were indicative of the C_{2v} symmetry of the indenide ligand. Similarly, the ^{13}C NMR spectrum of **11** showed eight carbon resonances, two for the $[C_5Me_5]^-$ ligand and six for the $[1,3\text{-}(Me_3Si)_2(\eta\text{-}C_9H_5)]^-$ ligand.

Recrystallization of **11** by slow cooling of a concentrated pentane solution to -20° C yielded crystals suitable for X-ray diffraction experiments. An X-ray diffraction study identified the crystals as the title compound shown in Figure 2.8. Details of the data collection, structure solution, and refinement are compiled in Table 2.7 and selected metrical parameters are listed in Table 2.8 and Table 2.9.
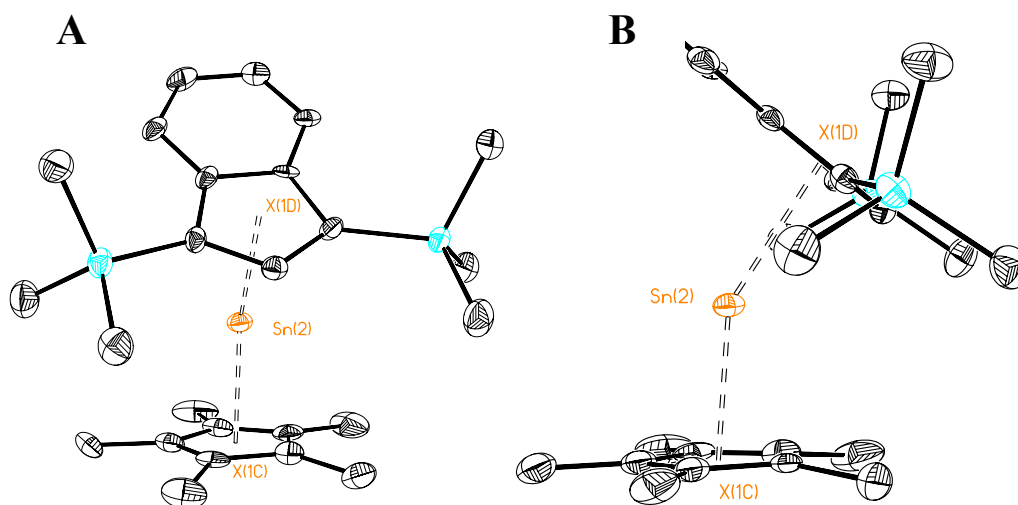
Figure 2.8 (A) Overall geometry of $[1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)](\eta^5\text{-}C_5Me_5)Sn$, **11**, and (B) the angle between $Cp^*_{centroid}(X1C)$—Sn—$Ind^{2TMS}_{centroid}(X1D)$. Atoms are colored as follows: carbon (black), silicon (cyan), and tin (orange). Thermal ellipsoids are shown at the 30% probability level.

The Sn—C_{Cp^*} bond distances range from 2.474(7)—2.708(4) Å with an average bond length of 2.602(4) Å, while the Sn—C_{Ind} bond distances range from 2.692(7) to 2.966(4) Å with an average bond length of 2.835(4) Å. Significantly shorter than the Sn—$Ind^{2TMS}_{centroid}$(X1D) bond distance (2.561 Å), the Sn—$Cp^*_{centroid}$(X1C) distance (2.313 Å) suggests appreciably stronger bonding between the metal center and the C_5Me_5 moiety. The pattern of bond distances suggests that the tin atom is pentahapto bonded to both the $[C_5Me_5]^-$ ring and the 5-membered ring of $[1,3-(Me_3Si)_2(\eta^5-C_9H_5)]^-$, $\Delta_{Sn-C} = 0.155$ Å. The X1C—Sn—X1D bond angle of 151.6º in **11** (Figure 2.8) is similar to that found in $Sn(\eta^5-C_5Me_5)_2$ (154.9º).[22]

Synthesis and Characterization of $[1,3-(Me_3Si)_2(\eta^5-C_9H_5)](\eta^5-C_5Me_5)Pb$, 12

(Pentamethylcyclopentadienyl)[1,3-bis(trimethylsilyl)indenyl]lead(II), **12,** was synthesized in good yield *via* the metathetical reaction of (pentamethylcyclopentadienyl)lead(II) chloride with an equimolar amount of $[1,3-(Me_3Si)_2(\eta^5-C_9H_5)]Li$. The product decomposes at room temperature over the course of a few days. The ^{1}H NMR spectrum of **12** exhibits five proton signals. An intense singlet at δ 2.012 indicated the C_{5v} symmetry of the pentamethylcyclopentadienide ring, while the other four ^{1}H NMR signals were indicative of the C_{2v} symmetry of the indenide ligand. Similarly, the ^{13}C NMR

spectrum of **5** showed eight carbon resonances, two for the $[C_5Me_5]^-$ and six for

the $[1,3-(Me_3Si)_2(\eta^5-C_9H_5)]^-$ ligand.

Recrystallization of **12** by slow cooling of a concentrated pentane solution

to -20° C yielded crystals suitable for X-ray diffraction experiments. An X-ray

diffraction study identified the crystals as the title compound shown in Figure 2.9.

Details of the data collection, structure solution, and refinement are compiled in

Table 2.9 and selected metrical parameters are listed in Table 2.11 and Table

2.12.
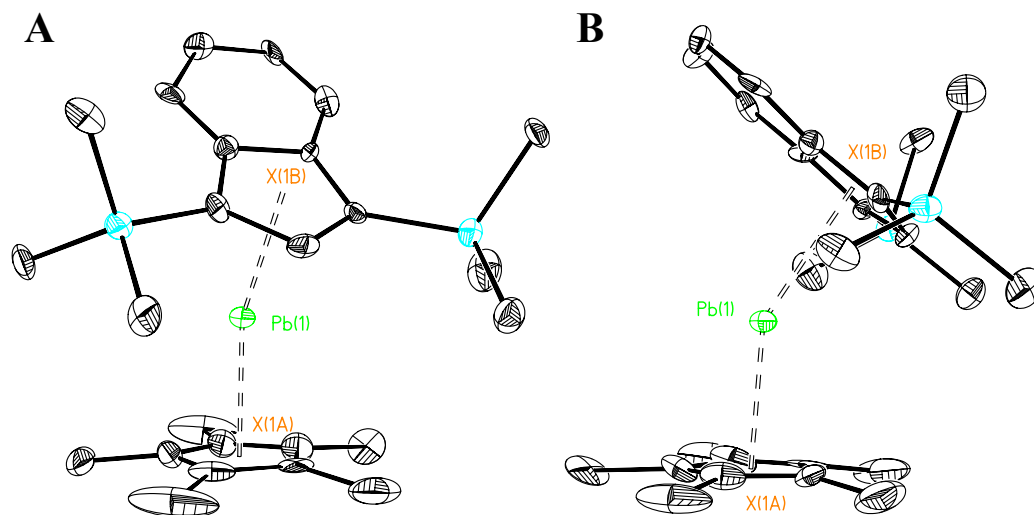
Figure 2.9 (A) Overall geometry of $[1,3-(Me_3Si)_2(\eta^5-C_9H_5)](\eta^5-C_5Me_5)Pb$, 12, and (B) the angle between $Cp^*_{centroid}(X1A)—Pb—Ind^{2TMS}_{centroid}(X1B)$. Atoms are colored as follows: carbon (black), silicon (cyan), and lead (green). Thermal ellipsoids are shown at 30% probability.

The $Pb—C_{Cp^*}$ bond distances range from 2.585(4)—2.783(4) Å with an

average bond length of 2.688 Å, while the $Pb—C_{Ind}$ bond distances range from

2.738(4)—3.055(4) Å with an average bond length of 2.901 Å. Significantly shorter than the Pb—Ind$^{2TMS}_{centroid}$(X1B) bond distance (2.634 Å), the Pb—Cp*$_{centroid}$(X1A) distance (2.406 Å) suggests appreciably stronger bonding between the metal center and the C_5Me_5 moiety. The pattern of bond distances suggests that the lead atom is pentahapto bonded to both the $[C_5Me_5]^-$ ring and the 5-membered ring of $[1,3-(Me_3Si)_2(\eta^5-C_9H_5)]^-$, Δ_{Pb-C} = 0.174 Å. The X1A—Pb—X1B bond angle is 148.42, which is shown in Figure 2.9. This bond angle is similar to that found in Pb(η^5-C_5Me_5)$_2$ (151.3º).[25]

Synthesis and Characterization of [1,3-(Me₃Si)₂(η⁵-C₉H₅)]₂Li₂·Et₂O, 13

In an attempt to synthesize the group 15 bis(π-indenide) complexes, two equivalents of $[1,3-(Me_3Si)_2(\eta-C_9H_5)]$Li (0.399 g, 1.5 mmol) were treated with antimony(III) chloride (0.339 g, 1.5 mmol). Recrystallization of the product by slow cooling of a concentrated solution to -20° C yielded crystals suitable for X-ray diffraction experiments. An X-ray diffraction study revealed that the product is the dilithio derivative $[1,3-(Me_3Si)_2(\eta^5-C_9H_5)]_2Li_2·Et_2O$ illustrated in Figure 2.10. Details of the data collection, structure solution, and refinement are compiled in Table 2.13 and selected metrical parameters are listed in Table 2.14 and Table 2.15.
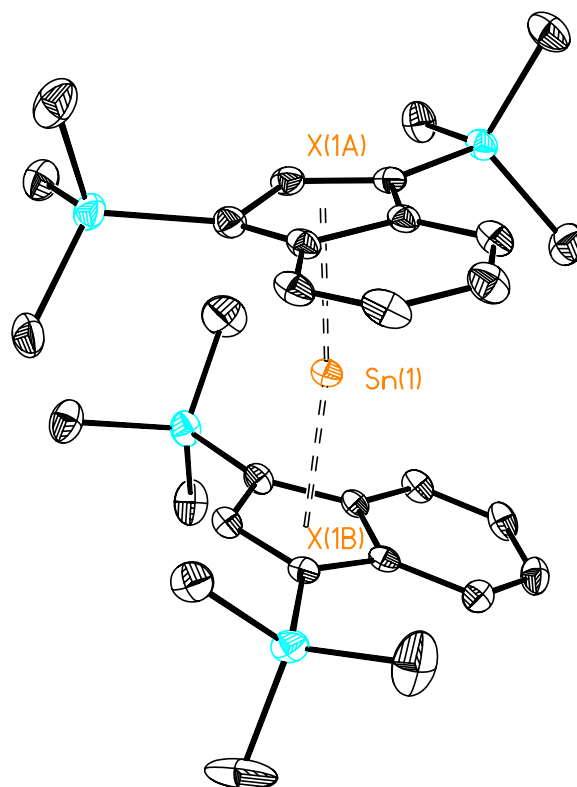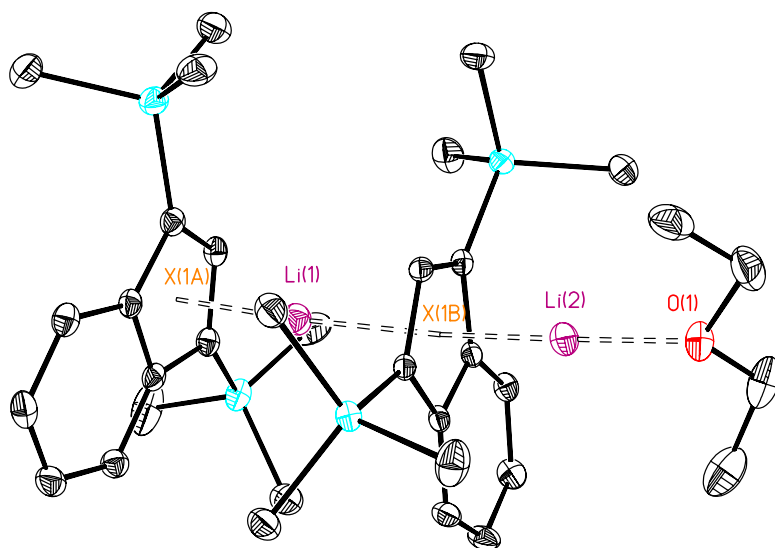
Figure 2.10 **Molecular structure of [1,3-(Me₃Si)₂(η⁵-C₉H₅)]₂Li₂·Et₂O, 13, showing stacked structure. Atoms are colored as follows: carbon (black), silicon (cyan), lithium (violet), and oxygen (red). Thermal ellipsoids are shown at the 30% probability level.**

In general, the structure of base-stabilized **13**, is similar to that reported

for $[Li_2\{C_{24}H_{14}-(SiMe_3)_2\}]\cdot 2THF^{[26]}$ and benzene stabilized $[\{Li(\eta^5-C_5Bz_5)\}_2(C_6D_6)].^{[27]}$ In the case of **13**, one lithium atom is capped by a diethyl

ether moiety, the oxygen atom of which is located 1.877(4) Å from the lithium

atom. The Li(1)—C_{Ind}(external) bond distances range from 2.212(4)—2.433(4) Å

with an average bond length of 2.325 Å. The Li(1)—C_{Ind}(internal) bond distances

are on average slightly longer than the corresponding C_{Ind}(external) distances and

range from 2.301(4)—2.678(4) Å with an average bond length of 2.492 Å. The

Li(2)—C_{Ind}(internal) bond distances range from 2.232(4)—2.280(4) Å and

average 2.333 Å. The external ring centroid (X1A)—Li bond distance of 1.981(4) Å is shorter than that in [{Li(η^5-C$_5$Bz$_5$)}$_2$(C$_6$D$_6$)] (2.015 Å).[27] but longer than that in [Li$_2${C$_{24}$H$_{14}$-(SiMe$_3$)$_2$}]·2THF (1.78 Å).[26] The internal ring centroid (X1B)—Li bond distance of 2.175 Å is longer than that in both [Li$_2${C$_{24}$H$_{14}$-(SiMe$_3$)$_2$}]·2THF (2.094 Å)[26] and [{Li(η^5-C$_5$Bz$_5$)}$_2$(C$_6$D$_6$)] (2.03 Å).[27] The X1B—Li(2) bond distance (1.901 Å) is the shortest of all the Li—X bond distances and identical to the X1B—Li(2) bond distances in [{Li(η^5-C$_5$Bz$_5$)}$_2$(C$_6$D$_6$)] (1.91 Å)[27] within experimental error. For a comparison of bond distances in **6** and [{Li(η^5-C$_5$Bz$_5$)}$_2$(C$_6$D$_6$)], see Figure 2.11. The pattern of bond distances indicates that both lithium atoms are pentahapto bonded to the five-membered rings of the indenide ligands.
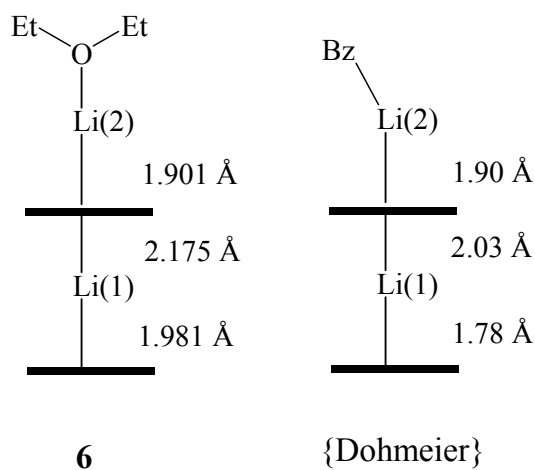
Figure 2.11   Comparison of bond lengths in 13 and [{Li(η^5-C$_5$Bz$_5$)}$_2$(C$_6$D$_6$)].[27]

The X1A—Li(1)—X1B bond angle is nearly linear (173.6º), while the Li(1)—X1B—Li(2) bond angle deviates significantly from linearity (169.0º). As shown in Figure 2.12, the trimethylsilyl groups are oriented in a *gauche* conformation. The two Me$_3$Si ligands are rotated by an angle of 92.6º with respect to each other. Such a conformation may serve to minimize the unfavorable non-bonding interactions between the indenyl substituents. Interestingly, the *gauche* conformation is similar to that reported prviously for [(Me$_3$Si)$_2$C$_9$H$_5$]$_2$Pb[20] and that determined for **10** in the present work.

ca. 93º
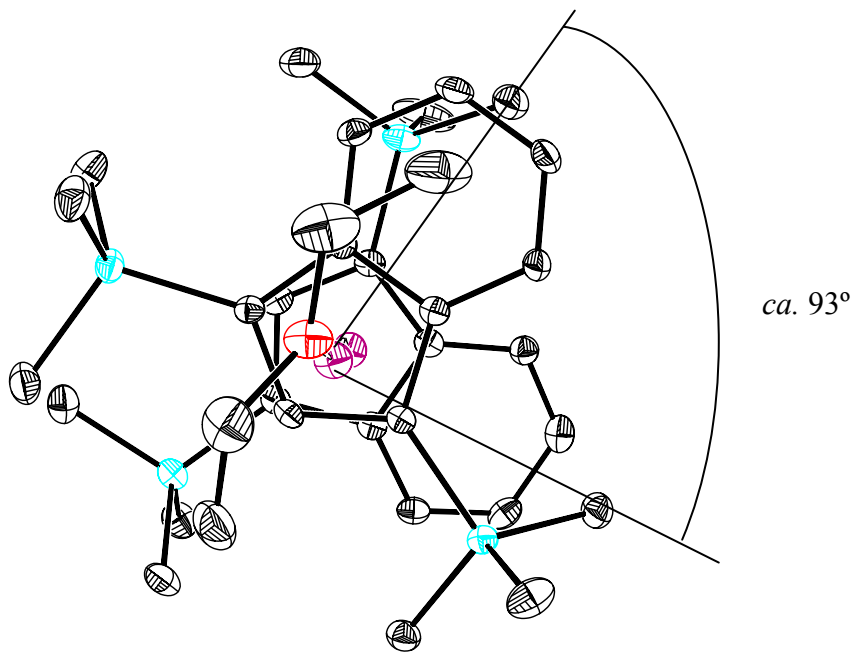
Figure 2.12   Molecular structure of [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)]$_2$Li$_2$·Et$_2$O, 13, showing the *gauche* orientation of the 1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$) ligands. Atoms are colored as follows: carbon (black), silicon (cyan), lithium (violet), and oxygen (red). Thermal ellipsoids are shown at 30% probability.

At this point, the mechanism of formation of the lithium triple-decker metallocene, **13**, is not known. Moreover, it was not possible to reproduce the synthesis of this compound. Unfortunately, no spectroscopic data were acquired for the complex prior to the X-ray crystallographic study.

Reaction of an indenyl cation and a neutral metallocene

In an attempt to synthesize the first example of a post-transition metal, mixed ligand, triple-decker metallocene, the $[B(C_6F_5)4]^-$ salt of the half sandwich cation, $[\{1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)\}Sn]^+$, was allowed to react *in situ* with an equimolar quantity of decamethylstannocene, $Sn(\eta^5\text{-}C_5Me5)_2$. The reaction was expected to yield the mixed ligand, triple-decker cation, $[\{1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)\}Sn(\mu\text{-}\eta^5\text{-}C_5Me_5)Sn(\eta^5\text{-}C_5Me_5)]^+$; however, instead the reaction afforded the crystalline product $[(\eta^5\text{-}C_5Me_5)Sn][B(C_6F_5)4]$ (Figure 2.13). Evidently, the less stable $[\{1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)\}Sn]^+$ cation abstracted a $[C_5Me_5]^-$ moiety from decamethylstannocene to generate the most stable combination of products, namely $[(\eta^5\text{-}C_5Me_5)Sn]^+$ and $[1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)](\eta^5\text{-}C_5Me_5)Sn$.
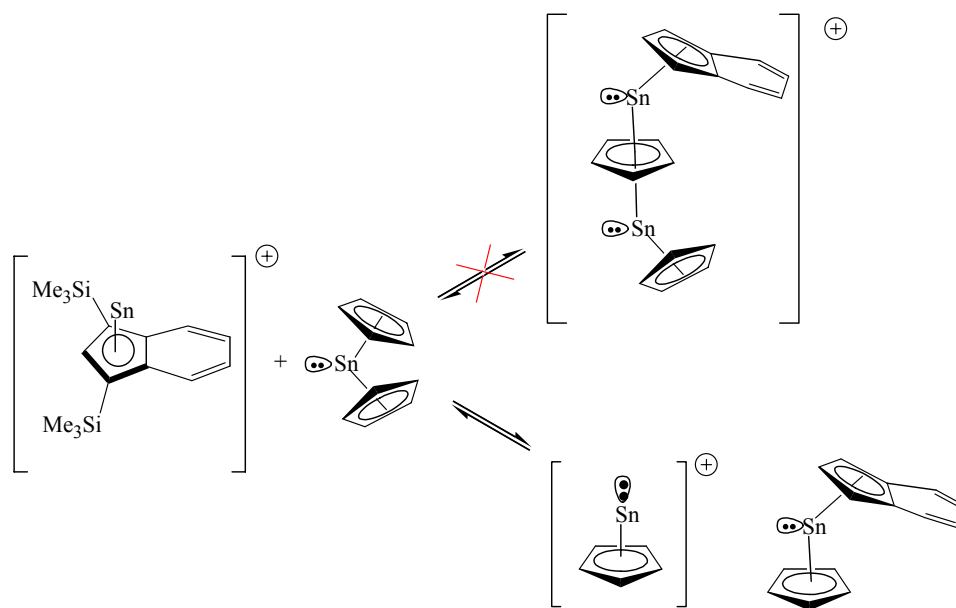
Figure 2.13   Reaction scheme for the intended synthesis of mixed ligand, triple-decker cations. Methyl groups have been omitted for clarity.

The group 14 bis(π-indenyl) and (π-indenyl)(C_5Me_5) metallocenes have been successfully synthesized and structurally characterized. Unlike several of their transition metal analogues, the π-indenide ligand is η^5-coordinated to the group 14 element center in all of the complexes examined. In the bis(π-indenyl) systems, the indenyl ligands orient themselves in a *gauche* fashion in order to minimize methyl group interactions.

The indenide ligand was found to coordinate weakly to the main group element centers. In the mixed ligand systems, and [1,3-(Me_3Si)$_2$(η^5-C_9H_5)](η^5-

$C_5Me_5)E$, the group 14 element center was significantly closer to the C_5Me_5 ring centroid, thus indicating an appreciably stronger bond to the C_5Me_5 ring than compared to the $1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)$ ring. Additional information regarding the bond strengths and stability of the complexes was provided by the reaction of $[\{1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)\}Sn]^+$ with decamethylstannocene. Abstraction of a $[C_5Me_5]^-$ anion by the indenyl cation to yield the most stable products congruently with the observed thermal stabilities and the pattern of bond lengths for the crystalline solids indicates the order of stability for the new complexes is as follows:

$[1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)](\eta^5\text{-}C_5Me_5)Ge >$ $[1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)]_2Ge >$

$[1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)](\eta^5\text{-}C_5Me_5)Sn >$ $[1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)]_2Sn >$

$[1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)](\eta^5\text{-}C_5Me_5)Pb >$ $[1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)]_2Pb >$
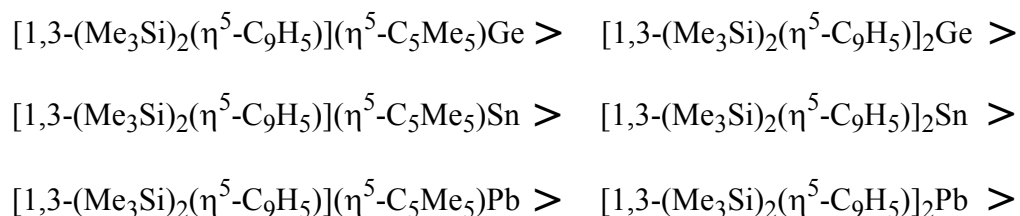
On the basis of the reactions studied, the indenide ligand is highly labile. Such complexes with labile ligands are ideally suited to serve as precursors to nanoparticles. Future work should therefore include the testing of bis(π-indenyl) group 14 metallocene complexes for their use in nanoparticle synthesis.

EXPERIMENTAL SECTION

General Procedures

All solvents were distilled over sodium benzophenone ketyl, except dichloromethane, which was distilled over CaH_2, and degassed prior to use. An M-Braun or VAC Vacuum Atmosphere argon-filled drybox was used for the manipulation of all solid reagents. All reactions were performed under dry, oxygen-free conditions using standard Schlenk or drybox techniques. To ensure the absence of water, all glassware was dried overnight in a 140° C oven before use.

The group 14 dihalides were purchased from a commercial source and used without further purification. The compounds 1,3-$(Me_3Si)_2(\eta-C_9H_5)Li$[21], $(\eta-Me_5C_5)SnCl$[28], and $(\eta-Me_5C_5)PbCl$[29] were prepared according to literature procedures.

Physical Measurements

Low-resolution CI mass spectra were collected on a Finnigan MAT TSQ-700 machine; high resolution mass spectra were measured on a VG Analytical ZAB-VE sector instrument. Mass spectra samples were flame sealed in glass capillaries to prevent exposure to oxygen. Unless otherwise noted, solution phase ^{13}C and ^1H NMR spectra were recorded at 295 K on a GE EQ-300 instrument (^1H,

300 MHz; ^{13}C, 75 MHz). Some ^{13}C and ^1H NMR spectra and all the ^{11}B, ^{19}F, ^{119}Sn, and ^{207}Pb solution-phase NMR spectra were recorded at 295 K on a Varian Inova-500 spectrometer (^1H, 500 MHz; ^{11}B, 160 MHz; ^{13}C, 125 MHz; ^{19}F, 470 MHz; ^{27}Al, 130 MHz). NMR samples were flame-sealed or run immediately following removal from the drybox. The CD_2Cl_2 solvent was obtained in sealed vials from a commercial source and used without further purification. The ^1H and ^{13}C spectra are reported relative to tetramethylsilane (0.00 ppm) and referenced to the solvent. ^{119}Sn chemical shifts are reported relative to Me_4Sn (0.00 ppm), while ^{207}Pb spectra are reported relative to Me_4Pb (0.00 ppm). Unless otherwise noted, the ^{13}C spectra were obtained under conditions of broadband proton decoupling. Melting points were obtained on a Fisher-Johns apparatus after flame-sealing the samples in glass capillaries under argon; the reported values are uncorrected.

X-Ray Crystallography

Structure determination and refinements were performed by the author at the University of Texas at Austin. Suitable single crystals were covered with a perfluorinated polyether oil to minimize oxygen-exposure. The X-ray data were collected on a Nonius Kappa CCD diffractometer equipped with an Oxford Cryostream liquid nitrogen cooling stream. Samples were mounted on a fine glass fiber using commercial silicon grease as an adhesive. All data collections

were performed at 153(2) K using a graphite monochromated Mo Kα radiation (λ = 0.71073 Å). A correction was applied for Lorentz-polarization. All structures were solved by direct methods and refined by full-matrix least squares on F^2 using the Siemens SHELX PLUS 5.0 (PC) software package.[30] All non-hydrogen atoms were allowed anisotropic thermal motion. When possible, all hydrogen atoms were located *via* the model; however, it was occasionally necessary to place the hydrogen atoms in calculated positions (C—H 0.96 Å). The hydrogen atoms were refined using a riding model and a general isotropic thermal parameter. The total number of reflections, collection ranges, and final R-values for each molecule are listed in the appropriate crystallographic data tables.

Synthesis of [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)]$_2$Sn], 10

A yellow solution of SnCl$_2$ (0.405 g, 2.13 mmol) in THF (40 mL) was treated with a two equivalents of a 1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)Li (1.136 g, 4.27 mmol) THF solution (40 mL) of at 0° C. The orange-red solution was stirred at room for 1h. The solvent and volatiles were removed under reduced pressure rendering a red oily solid. The product was extracted into pentane (60 mL) and filtered through Celite® to give an orange-red solution. A thick, red oil was obtained by removing the volatiles. Storage of the oil at -20° C afforded a crop of red crystals of **10** (0.604g , 62.2% yield). *Spectroscopic Data*: HRMS (CI, CH$_4$): calcd for C$_{30}$H$_{46}$Si$_2$Sn, *m/z* 639.177; found, 639.1788; ^{1}H NMR (C$_6$D$_6$): δ 0.348 [s,

(CH$_3$)$_3$Si], 6.36 [s, H(2)], 6.99—7.02 and 7.49—7.52 [AA'BB' system, H(5)—H(8)]; ^{13}C{^{1}H} NMR (C$_6$D$_6$): δ 1.089 [s, (CH$_3$)$_3$Si], 110.405 [s, C(1) and C(3)], 122.759—122.917 [s, C(4), C(9), and C(2)], 124.002 [s, C(6) and C(7)], 141.007 [s, C(5) and C(8)]; ^{119}Sn NMR (C$_6$D$_6$): δ -2168.

Synthesis of [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)](η^5-C$_5$Me$_5$)Sn, 11

A yellow solution of (η^5-C$_5$Me$_5$)SnCl (0.140 g, 0.483 mmol) in THF (40 mL) was treated with a THF solution (40 mL) of 1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)Li (0.128 g, 0.481 mmol) at 0° C. The yellow-orange solution was stirred at room for 1h. The solvent and volatiles were removed under reduced pressure rendering a red oily solid. The product was extracted into pentane (60 mL) and filtered through Celite$^®$ to give a yellow-orange solution. The solution was concentrated under reduced pressure to a volume of approximately 20 mL. Storage of this solution at -20° C afforded a crop of yellow crystals of **11** (~ 45% yield). *Spectroscopic Data*: HRMS (CI, CH$_4$): calcd for C$_{24}$H$_{30}$Si$_2$Sn, *m/z* 499.1299; found, 499.1319; ^{1}H NMR (C$_6$D$_6$): δ 0.499 [s, (CH$_3$)$_3$Si], 1.716 [s, (CH$_3$)$_5$C$_5$], 6.407 [s, H(2)]; 7.081—7.112 and 7.795—7.825 [AA'BB' system, H(5)—H(8)]; ^{13}C{^{1}H} NMR (C$_6$D$_6$): δ 0.97 [s, (CH$_3$)$_3$Si], 9.42 [s, (CH$_3$)$_5$C$_5$], 111.54 [s, C(1) and C(3)], 118.37 [s, (CH$_3$)$_5$C$_5$], 122.22—122.34 [C(4), C(9) and C(2)], 124.06 [s, C(6) and C(7)], 141.03 [s, C(5) and C(8)]; ^{119}Sn NMR (C$_6$D$_6$): δ -2121.

Synthesis of [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)](η^5-C$_5$Me$_5$)Pb, 12

An orange solution of (η^5-C$_5$Me$_5$)PbCl (0.587 g, 1.55 mmol) in THF (40 mL) was treated with a 1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)Li (0.413 g, 1.55 mmol) in THF solution (40 mL) at 0° C. The orange-red solution was stirred at room for 1h. The solvent and volatiles were removed under reduced pressure rendering a red oily solid. The product was extracted into pentane (60 mL) and filtered through Celite® to give an orange-red solution. A thick, red oily solution was obtained by removal of the volatiles under reduced pressure. Storage of the oil at -20° C afforded a crop of red crystals of **12** (0.334 g, 35.7% yield). *Spectroscopic Data*: HRMS (CI, CH$_4$): calcd for C$_{25}$H$_{38}$Si$_2$Pb, *m/z* 603.2356; found, 603.2372; ^{1}H NMR (C$_6$D$_6$): δ 0.493 [s, (CH$_3$)$_3$Si], 1.989 [s, (CH$_3$)$_5$C$_5$], 6.460 [s, H(2)]; 7.118—7.139 and 7.821—7.851 [AA'BB' system, H(5)—H(8)]; ^{13}C{^{1}H} NMR (C$_6$D$_6$): δ 1.20 [s, (CH$_3$)$_3$Si], 9.06 [s, (CH$_3$)$_5$C$_5$], 111.77 [s, C(1) and C(3)], 119.99 [s, (CH$_3$)$_5$C$_5$] 121.85—121.88 [C(4), C(9) and C(2)], 123.66 [s, C(6) and C(7)], 126.94 [s, C(5) and C(8)]; [207]Pb NMR (CDCl$_3$): δ -4402.

Synthesis of [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)]$_2$Li$_2$·Et$_2$O, 13

An orange solution of SbCl$_3$ (0.339 g, 1.5 mmol) in THF (40 mL) was treated with a two equivalents of 1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)Li (0.399 g, 1.5 mmol) in THF solution (40 mL) at 0° C. The orange-red solution was stirred at room for 1h. The solvent and volatiles were removed under reduced pressure rendering a

red oily solid. The product was extracted into pentane (60 mL) and filtered through Celite® to give a yellow-orange solution. The volume of the solution was reduced to ca. 20 mL under reduced pressure. Storage of this solution at -20° C afforded a crop of yellow crystals of **13**.

TABLES OF X-RAY CRYSTALLOGRAPHIC DATA
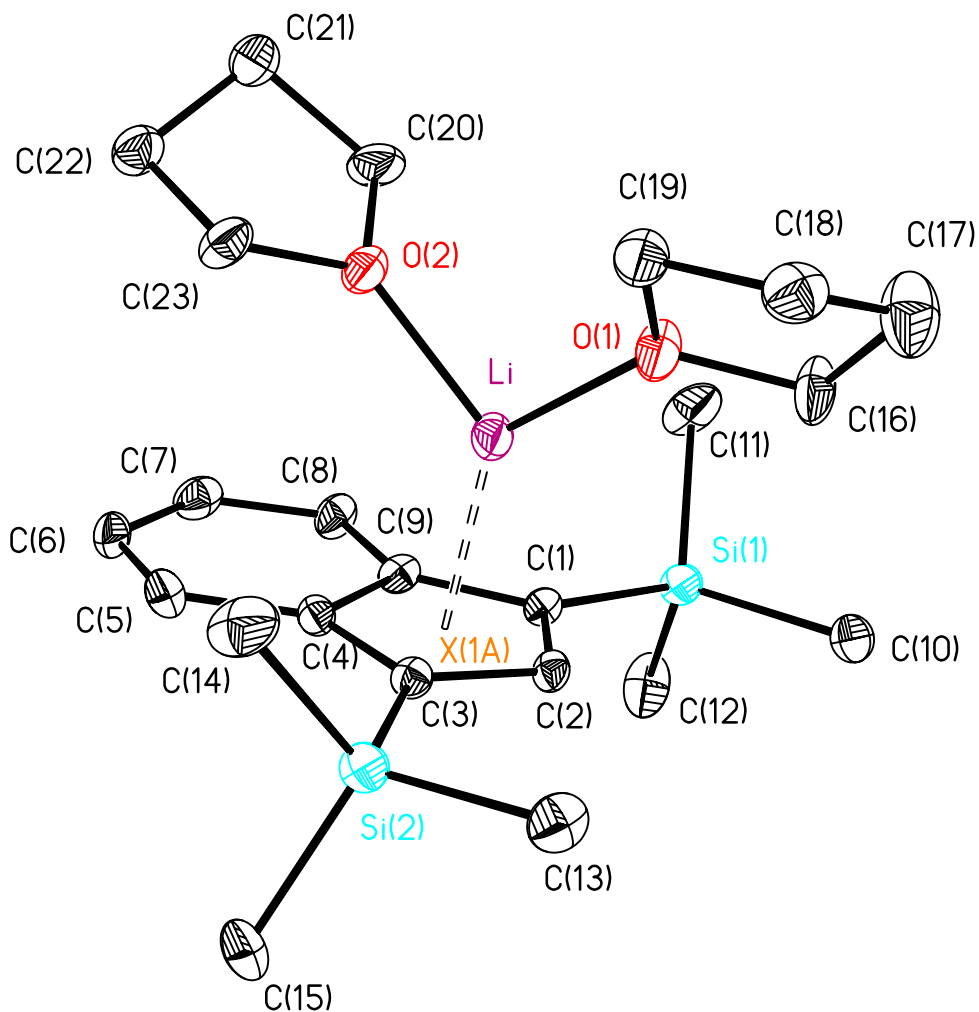
Figure 2.14 Molecular structure of [1,3-$(Me_3Si)_2(\eta^5$-$C_9H_5)$]Li·2THF , **9**, showing the atom number scheme. Atoms are colored as follows: carbon (black), silicon (**cyan**), lithium (**violet**), and oxygen (**red**). The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.

Table 2.1 Crystal Data and Structure Refinement for [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)]Li·2THF, 9.

Identification code	pna21
Empirical formula	C23 H40 Li O2 Si2
Formula weight	411.67
Temperature	153(2) K
Wavelength	0.71069 Å
Crystal system	Orthorhombic
Space group	Pna2$_1$
Unit cell dimensions	a = 15.226(5) Å α= 90.000(5)°.
	b = 10.240(5) Å β= 90.000(5)°.
	c = 16.141(5) Å γ = 90.000(5)°.
Volume	2516.6(17) Å^3
Z	4
Density (calculated)	1.087 Mg/m^3
Absorption coefficient	0.155 mm^{-1}
F(000)	900
Crystal size	0.20 x 0.20 x 0.20 mm^3
Theta range for data collection	2.96 to 30.19°.
Index ranges	-19$\leq$h$\leq$20, -13$\leq$k$\leq$13, -22$\leq$l$\leq$21
Reflections collected	9727
Independent reflections	5794 [R(int) = 0.1359]
Completeness to theta = 30.19°	84.2 %
Absorption correction	None
Max. and min. transmission	0.9696 and 0.9696
Refinement method	Full-matrix least-squares on F^2
Goodness-of-fit on F^2	1.056
Final R indices [I>2sigma(I)]	R1 = 0.0769, wR2 = 0.1084
R indices (all data)	R1 = 0.2120, wR2 = 0.1432
Absolute structure parameter	0.6(2)
Largest diff. peak and hole	0.324 and -0.525 e.Å^{-3}

Table 2.2 Selected Bond Lengths (Å) for [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)]Li·2THF, 9.

C(1)-C(2)	1.416(7)	C(11)-Si(1)	1.859(7)
C(1)-C(9)	1.441(7)	C(12)-Si(1)	1.861(6)
C(1)-Si(1)	1.842(5)	C(13)-Si(2)	1.874(5)
C(1)-Li	2.320(9)	C(14)-Si(2)	1.878(7)
C(2)-C(3)	1.427(7)	C(15)-Si(2)	1.870(6)
C(2)-Li	2.243(9)	C(16)-O(1)	1.445(6)
C(3)-C(4)	1.448(7)	C(16)-C(17)	1.490(8)
C(3)-Si(2)	1.840(5)	C(17)-C(18)	1.490(8)
C(3)-Li	2.327(10)	C(18)-C(19)	1.509(8)
C(4)-C(5)	1.424(7)	C(19)-O(1)	1.465(6)
C(4)-C(9)	1.443(5)	C(20)-O(2)	1.429(7)
C(4)-Li	2.414(9)	C(20)-C(21)	1.499(8)
C(5)-C(6)	1.364(7)	C(21)-C(22)	1.514(8)
C(6)-C(7)	1.411(7)	C(22)-C(23)	1.508(8)
C(7)-C(8)	1.390(7)	C(23)-O(2)	1.462(7)
C(8)-C(9)	1.408(7)	O(1)-Li	1.913(8)
C(9)-Li	2.410(9)	O(2)-Li	1.915(8)
C(10)-Si(1)	1.873(5)	Li-X1A	2.000(4)

Table 2.3 Selected Bond Angles (°) for [1,3-(Me₃Si)₂(η⁵-C₉H₅)]Li·2THF, 9.

C(2)-C(1)-C(9)	105.3(4)	C(1)-Si(1)-C(12)	110.1(3)
C(2)-C(1)-Si(1)	126.9(4)	C(11)-Si(1)-C(12)	109.1(3)
C(9)-C(1)-Si(1)	127.7(4)	C(1)-Si(1)-C(10)	110.9(2)
C(2)-C(1)-Li	69.0(3)	C(11)-Si(1)-C(10)	108.5(3)
C(9)-C(1)-Li	75.7(3)	C(12)-Si(1)-C(10)	107.9(3)
Si(1)-C(1)-Li	122.3(3)	C(3)-Si(2)-C(15)	111.4(3)
C(1)-C(2)-C(3)	113.2(4)	C(3)-Si(2)-C(13)	108.3(2)
C(1)-C(2)-Li	74.9(4)	C(15)-Si(2)-C(13)	109.2(3)
C(3)-C(2)-Li	75.0(4)	C(3)-Si(2)-C(14)	111.7(3)
C(2)-C(3)-C(4)	103.9(4)	C(15)-Si(2)-C(14)	107.6(4)
C(2)-C(3)-Si(2)	126.8(4)	C(13)-Si(2)-C(14)	108.7(4)
C(4)-C(3)-Si(2)	129.2(4)	O(1)-Li-O(2)	98.5(4)
C(2)-C(3)-Li	68.6(3)	O(1)-Li-C(2)	104.7(4)
C(4)-C(3)-Li	75.5(3)	O(2)-Li-C(2)	156.0(5)
Si(2)-C(3)-Li	122.8(3)	O(1)-Li-C(1)	117.3(4)
C(5)-C(4)-C(9)	118.7(5)	O(2)-Li-C(1)	133.8(4)
C(5)-C(4)-C(3)	131.9(5)	C(2)-Li-C(1)	36.1(2)
C(9)-C(4)-C(3)	109.3(4)	O(1)-Li-C(3)	121.6(4)
C(5)-C(4)-Li	124.7(4)	O(2)-Li-C(3)	123.5(4)
C(9)-C(4)-Li	72.4(4)	C(2)-Li-C(3)	36.3(2)
C(3)-C(4)-Li	69.0(3)	C(1)-Li-C(3)	61.4(2)
C(6)-C(5)-C(4)	119.9(5)	O(1)-Li-C(9)	151.7(5)
C(5)-C(6)-C(7)	121.3(5)	O(2)-Li-C(9)	102.5(4)
C(8)-C(7)-C(6)	120.9(5)	C(2)-Li-C(9)	58.3(2)
C(7)-C(8)-C(9)	119.0(5)	C(1)-Li-C(9)	35.4(2)
C(8)-C(9)-C(1)	131.7(5)	C(3)-Li-C(9)	59.7(3)
C(8)-C(9)-C(4)	120.1(5)	O(1)-Li-C(4)	157.1(5)
C(1)-C(9)-C(4)	108.2(4)	O(2)-Li-C(4)	97.9(4)
C(8)-C(9)-Li	124.3(4)	C(2)-Li-C(4)	58.1(3)
C(1)-C(9)-Li	68.9(3)	C(1)-Li-C(4)	59.1(2)
C(4)-C(9)-Li	72.7(4)	C(3)-Li-C(4)	35.5(2)
C(1)-Si(1)-C(11)	110.3(3)	C(9)-Li-C(4)	34.82(17)
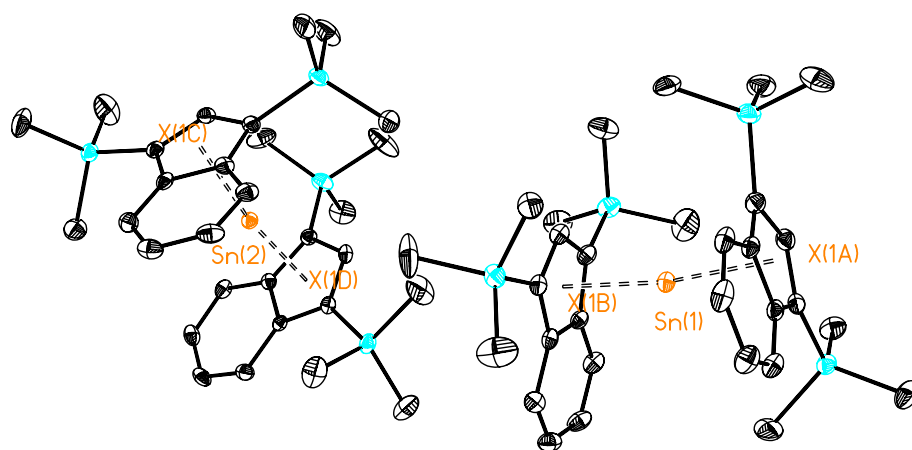
Figure 2.15 Molecular structure of [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)]$_2$Sn, 10, showing the two independent molecules per unit cell. Atoms are colored as follows: carbon (black), silicon (cyan), and tin (orange). The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.
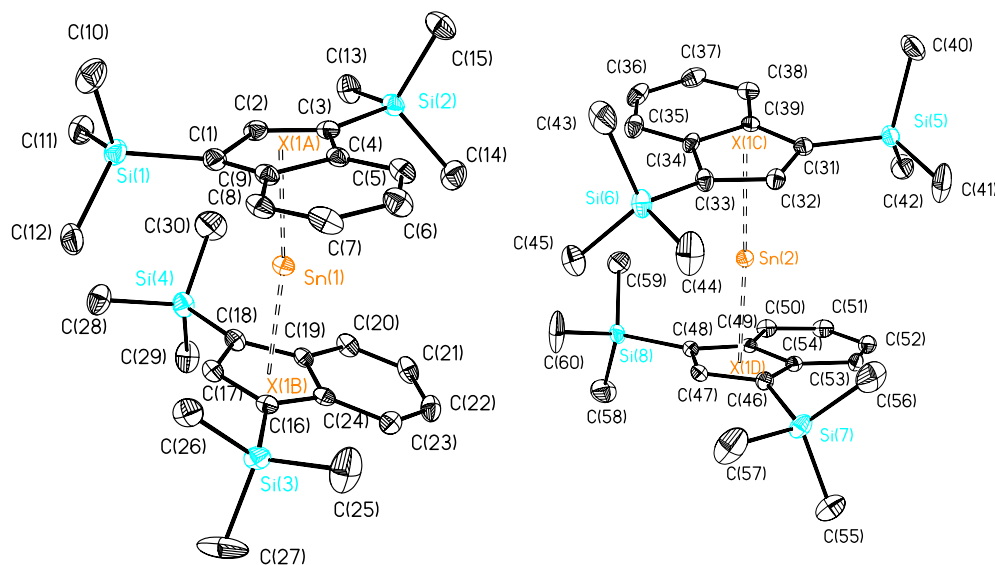
Figure 2.16 Molecular structure of the two independent molecules in the [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)]$_2$Sn, 10, crystal structure. The figure indicates the complete atom numbering scheme. Atoms are colored as follows: carbon (black), silicon (cyan), and tin (orange). The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.

107

Table 2.4 Crystal Data and Structure Refinement for [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)]$_2$Sn, 10.

Identification code	p1bar
Empirical formula	C60 H60 Si8 Sn2
Formula weight	1243.18
Temperature	153(2) K
Wavelength	0.71069 Å
Crystal system	Triclinic
Space group	P-1
Unit cell dimensions	a = 11.440(5) Å α= 103.452(5)°.
	b = 17.300(5) Å β= 105.369(5)°.
	c = 18.587(5) Å γ = 91.091(5)°.
Volume	3437(2) Å^3
Z	2
Density (calculated)	1.201 Mg/m^3
Absorption coefficient	0.898 mm^{-1}
F(000)	1264
Crystal size	0.20 x 0.20 x 0.20 mm^3
Theta range for data collection	3.12 to 27.43°.
Index ranges	-13$\leq$h$\leq$14, -20$\leq$k$\leq$22, -24$\leq$l$\leq$24
Reflections collected	24136
Independent reflections	15537 [R(int) = 0.0289]
Completeness to theta = 27.43°	99.1 %
Absorption correction	None
Max. and min. transmission	0.8409 and 0.8409
Refinement method	Full-matrix least-squares on F^2
Goodness-of-fit on F^2	1.052
Final R indices [I>2sigma(I)]	R1 = 0.0433, wR2 = 0.0798
R indices (all data)	R1 = 0.0702, wR2 = 0.0889
Extinction coefficient	0.00152(17)
Largest diff. peak and hole	0.765 and -0.752 e.Å^{-3}

Table 2.5 Selected Bond Lengths (Å) for [1,3-(Me₃Si)₂(η⁵-C₉H₅)]₂Sn, 10.

C(1)-Sn(1)	2.639(4)	C(32)-C(33)	1.437(4)
C(2)-C(3)	1.411(5)	C(32)-Sn(2)	2.643(3)
C(3)-C(4)	1.453(4)	C(33)-C(34)	1.448(5)
C(3)-Si(2)	1.861(3)	C(33)-Sn(2)	2.610(3)
C(4)-C(5)	1.410(5)	C(34)-C(35)	1.416(5)
C(4)-C(9)	1.446(5)	C(34)-C(39)	1.435(4)
C(5)-C(6)	1.372(5)	C(35)-C(36)	1.356(5)
C(6)-C(7)	1.400(6)	C(36)-C(37)	1.399(5)
C(7)-C(8)	1.359(6)	C(37)-C(38)	1.379(5)
C(8)-C(9)	1.416(5)	C(38)-C(39)	1.409(5)
C(16)-C(17)	1.422(5)	C(46)-C(47)	1.426(4)
C(16)-C(24)	1.441(4)	C(46)-C(54)	1.447(4)
C(17)-C(18)	1.431(4)	C(46)-Sn(2)	2.612(3)
C(17)-Sn(1)	2.580(3)	C(47)-C(48)	1.413(4)
C(18)-C(19)	1.445(5)	C(47)-Sn(2)	2.617(3)
C(18)-Sn(1)	2.600(3)	C(48)-C(49)	1.443(4)
C(19)-C(20)	1.415(4)	C(49)-C(50)	1.411(4)
C(19)-C(24)	1.434(4)	C(49)-C(54)	1.431(4)
C(20)-C(21)	1.361(5)	C(50)-C(51)	1.366(5)
C(21)-C(22)	1.396(5)	C(51)-C(52)	1.404(5)
C(22)-C(23)	1.379(5)	C(52)-C(53)	1.362(5)
C(23)-C(24)	1.410(5)	C(53)-C(54)	1.426(4)
C(31)-C(32)	1.410(4)		

Table 2.6 Selected Bond Angles (°) for [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)]$_2$Sn, 10.

C(2)-C(1)-C(9)	104.3(3)	C(34)-C(33)-Sn(2)	79.23(18)
C(2)-C(1)-Sn(1)	75.5(2)	C(35)-C(34)-C(39)	119.6(3)
C(9)-C(1)-Sn(1)	78.04(19)	C(35)-C(34)-C(33)	131.9(3)
C(3)-C(2)-C(1)	113.4(3)	C(39)-C(34)-C(33)	108.5(3)
C(2)-C(3)-C(4)	104.7(3)	C(36)-C(35)-C(34)	119.5(4)
C(5)-C(4)-C(9)	119.3(3)	C(35)-C(36)-C(37)	121.5(4)
C(5)-C(4)-C(3)	131.8(3)	C(38)-C(37)-C(36)	120.9(3)
C(9)-C(4)-C(3)	108.9(3)	C(37)-C(38)-C(39)	119.6(3)
C(6)-C(5)-C(4)	119.4(4)	C(38)-C(39)-C(34)	118.9(3)
C(5)-C(6)-C(7)	121.0(4)	C(38)-C(39)-C(31)	132.3(3)
C(8)-C(7)-C(6)	122.0(4)	C(34)-C(39)-C(31)	108.9(3)
C(7)-C(8)-C(9)	119.2(4)	C(47)-C(46)-C(54)	104.4(3)
C(8)-C(9)-C(4)	119.2(3)	C(47)-C(46)-Sn(2)	74.36(17)
C(8)-C(9)-C(1)	132.1(3)	C(54)-C(46)-Sn(2)	79.80(17)
C(4)-C(9)-C(1)	108.6(3)	C(48)-C(47)-C(46)	112.8(3)
C(17)-C(16)-C(24)	105.0(3)	C(48)-C(47)-Sn(2)	79.21(17)
C(16)-C(17)-C(18)	112.2(3)	C(46)-C(47)-Sn(2)	73.97(17)
C(16)-C(17)-Sn(1)	79.60(19)	C(47)-C(48)-C(49)	104.9(3)
C(18)-C(17)-Sn(1)	74.73(19)	C(50)-C(49)-C(54)	119.3(3)
C(17)-C(18)-C(19)	104.8(3)	C(50)-C(49)-C(48)	131.8(3)
C(17)-C(18)-Sn(1)	73.18(18)	C(54)-C(49)-C(48)	108.9(3)
C(19)-C(18)-Sn(1)	81.14(18)	C(51)-C(50)-C(49)	119.6(3)
C(20)-C(19)-C(24)	119.3(3)	C(50)-C(51)-C(52)	121.3(3)
C(20)-C(19)-C(18)	132.0(3)	C(53)-C(52)-C(51)	121.2(3)
C(24)-C(19)-C(18)	108.7(3)	C(52)-C(53)-C(54)	119.3(3)
C(21)-C(20)-C(19)	119.8(3)	C(53)-C(54)-C(49)	119.3(3)
C(20)-C(21)-C(22)	121.3(3)	C(53)-C(54)-C(46)	131.9(3)
C(23)-C(22)-C(21)	121.0(4)	C(49)-C(54)-C(46)	108.8(3)
C(22)-C(23)-C(24)	119.5(3)	C(17)-Sn(1)-C(18)	32.09(10)
C(23)-C(24)-C(19)	119.2(3)	C(17)-Sn(1)-C(1)	116.68(11)
C(23)-C(24)-C(16)	131.7(3)	C(18)-Sn(1)-C(1)	124.72(11)
C(19)-C(24)-C(16)	109.0(3)	C(33)-Sn(2)-C(46)	125.62(10)
C(32)-C(31)-C(39)	105.6(3)	C(33)-Sn(2)-C(47)	117.81(10)
C(31)-C(32)-C(33)	112.1(3)	C(46)-Sn(2)-C(47)	31.66(9)
C(31)-C(32)-Sn(2)	79.93(18)	C(33)-Sn(2)-C(32)	31.74(10)
C(33)-C(32)-Sn(2)	72.83(18)	C(46)-Sn(2)-C(32)	119.56(10)

C(32)-C(33)-C(34)	104.9(3)		C(47)-Sn(2)-C(32)	132.11(10)
C(32)-C(33)-Sn(2)	75.43(18)			
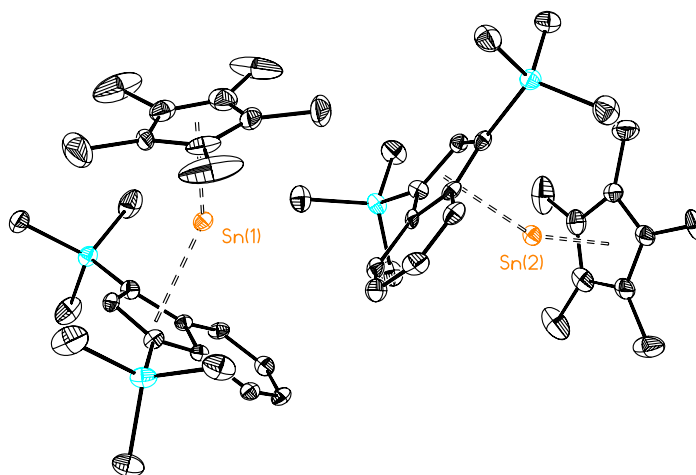
Figure 2.17 Molecular structure of [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)](η^5-C$_5$Me$_5$)Sn, 11, showing the two independent molecules per unit cell. Atoms are colored as follows: carbon (black), silicon (cyan), and tin (orange). The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.
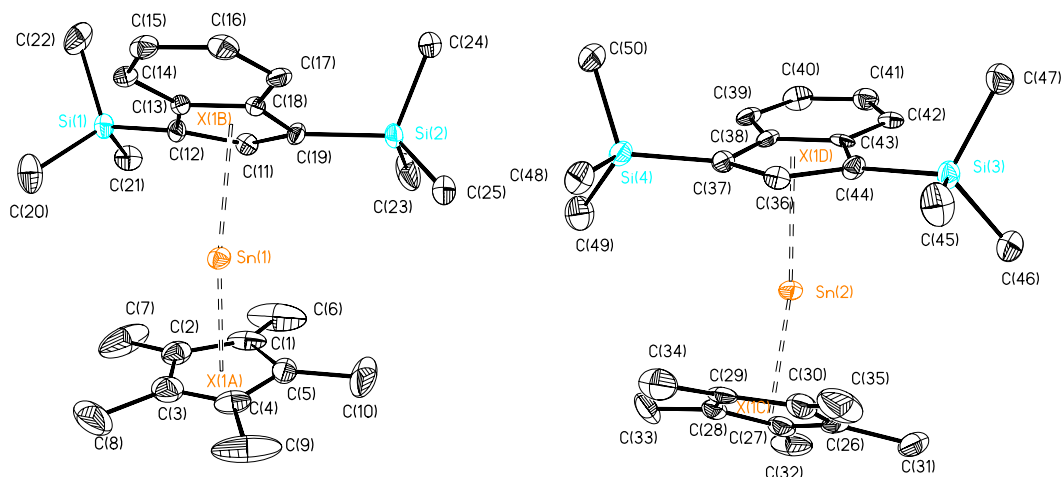
Figure 2.18 Molecular structures and complete atom numbering schemes for the two independent molecules in the crystal structure of [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)](η^5-C$_5$Me$_5$)Sn. Atoms are colored as follows: carbon (black), silicon (cyan), and tin (orange). The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.

112

Table 2.7   Crystal Data and Structure Refinement for [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)](η^5-C$_5$Me$_5$)Sn, 11.

|	
Identification code	p21212
Empirical formula	C50 H76 Si4 Sn2
Formula weight	1026.85
Temperature	153(2) K
Wavelength	0.71069 Å
Crystal system	Orthorhombic
Space group	P2$_1$2$_1$2
Unit cell dimensions	a = 21.232(5) Å α= 90.000(5)°.
	b = 24.238(5) Å β= 90.000(5)°.
	c = 10.398(5) Å γ = 90.000(5)°.
Volume	5351(3) Å^3
Z	4
Density (calculated)	1.275 Mg/m^3
Absorption coefficient	1.053 mm^{-1}
F(000)	2128
Crystal size	0.20 x 0.20 x 0.20 mm^3
Theta range for data collection	3.00 to 27.23°.
Index ranges	-27≤h≤27, -31≤k≤31, -13≤l≤13
Reflections collected	23473
Independent reflections	11903 [R(int) = 0.1077]
Completeness to theta = 27.23°	99.3 %
Absorption correction	None
Max. and min. transmission	0.8171 and 0.8171
Refinement method	Full-matrix least-squares on F^2
Goodness-of-fit on F^2	1.049
Final R indices [I>2sigma(I)]	R1 = 0.0590, wR2 = 0.0850
R indices (all data)	R1 = 0.1275, wR2 = 0.1041
Extinction coefficient	0.00174(12)
Largest diff. peak and hole	0.941 and -1.588 e.Å^{-3}

Table 2.8   Selected Bond Lengths (Å) for [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)](η^5-C$_5$Me$_5$)Sn, 11.

C(1)-C(2)	1.423(11)	C(26)-C(30)	1.404(11)
C(1)-C(5)	1.435(11)	C(26)-C(31)	1.501(11)
C(1)-C(6)	1.499(12)	C(26)-Sn(2)	2.617(6)
C(1)-Sn(1)	2.474(7)	C(27)-C(28)	1.403(10)
C(2)-C(3)	1.337(11)	C(27)-C(32)	1.516(11)
C(2)-C(7)	1.529(12)	C(28)-C(29)	1.440(10)
C(2)-Sn(1)	2.565(7)	C(28)-C(33)	1.495(10)
C(3)-C(4)	1.367(11)	C(28)-Sn(2)	2.649(7)
C(3)-C(8)	1.561(13)	C(29)-C(30)	1.438(11)
C(4)-C(5)	1.410(12)	C(29)-C(34)	1.492(10)
C(4)-C(9)	1.516(11)	C(29)-Sn(2)	2.527(7)
C(5)-C(10)	1.505(12)	C(30)-C(35)	1.520(11)
C(5)-Sn(1)	2.570(8)	C(30)-Sn(2)	2.520(8)
C(11)-C(19)	1.409(9)	C(36)-C(37)	1.415(10)
C(11)-C(12)	1.436(9)	C(36)-C(44)	1.434(9)
C(12)-C(13)	1.438(8)	C(37)-C(38)	1.460(9)
C(13)-C(14)	1.408(9)	C(38)-C(39)	1.411(9)
C(13)-C(18)	1.432(8)	C(38)-C(43)	1.430(9)
C(14)-C(15)	1.358(9)	C(39)-C(40)	1.369(10)
C(15)-C(16)	1.404(10)	C(40)-C(41)	1.384(10)
C(16)-C(17)	1.368(9)	C(41)-C(42)	1.368(10)
C(17)-C(18)	1.410(9)	C(42)-C(43)	1.413(9)
C(18)-C(19)	1.441(9)	C(43)-C(44)	1.447(9)
C(26)-C(27)	1.401(10)		

Table 2.9 Selected Bond Angles (°) for [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)](η^5-C$_5$Me$_5$)Sn, 11.

C(2)-C(1)-C(5)	106.4(7)	C(26)-C(27)-C(28)	110.0(7)
C(2)-C(1)-C(6)	126.9(10)	C(26)-C(27)-C(32)	126.8(8)
C(5)-C(1)-C(6)	125.4(10)	C(28)-C(27)-C(32)	123.2(8)
C(2)-C(1)-Sn(1)	77.2(4)	C(27)-C(28)-C(29)	107.0(7)
C(5)-C(1)-Sn(1)	77.2(4)	C(27)-C(28)-C(33)	126.5(7)
C(6)-C(1)-Sn(1)	121.6(6)	C(29)-C(28)-C(33)	126.4(7)
C(3)-C(2)-C(1)	108.5(7)	C(27)-C(28)-Sn(2)	76.5(4)
C(3)-C(2)-C(7)	126.1(10)	C(29)-C(28)-Sn(2)	69.2(4)
C(1)-C(2)-C(7)	124.4(10)	C(33)-C(28)-Sn(2)	121.6(5)
C(3)-C(2)-Sn(1)	81.0(5)	C(30)-C(29)-C(28)	106.8(6)
C(1)-C(2)-Sn(1)	70.1(4)	C(30)-C(29)-C(34)	125.6(8)
C(7)-C(2)-Sn(1)	123.9(6)	C(28)-C(29)-C(34)	126.7(8)
C(2)-C(3)-C(4)	110.7(8)	C(30)-C(29)-Sn(2)	73.2(4)
C(2)-C(3)-C(8)	124.6(10)	C(28)-C(29)-Sn(2)	78.6(4)
C(4)-C(3)-C(8)	124.7(10)	C(34)-C(29)-Sn(2)	122.4(5)
C(3)-C(4)-C(5)	108.6(7)	C(26)-C(30)-C(29)	108.3(7)
C(3)-C(4)-C(9)	125.3(10)	C(26)-C(30)-C(35)	127.1(8)
C(5)-C(4)-C(9)	126.1(10)	C(29)-C(30)-C(35)	124.3(8)
C(4)-C(5)-C(1)	105.9(7)	C(26)-C(30)-Sn(2)	78.0(4)
C(4)-C(5)-C(10)	126.9(10)	C(29)-C(30)-Sn(2)	73.7(4)
C(1)-C(5)-C(10)	126.8(10)	C(35)-C(30)-Sn(2)	119.7(5)
C(4)-C(5)-Sn(1)	80.0(5)	C(37)-C(36)-C(44)	114.0(6)
C(1)-C(5)-Sn(1)	69.8(4)	C(36)-C(37)-C(38)	103.7(6)
C(10)-C(5)-Sn(1)	120.8(5)	C(39)-C(38)-C(43)	119.2(6)
C(19)-C(11)-C(12)	113.0(6)	C(39)-C(38)-C(37)	131.6(6)
C(11)-C(12)-C(13)	103.8(6)	C(43)-C(38)-C(37)	109.2(6)
C(14)-C(13)-C(18)	118.2(6)	C(40)-C(39)-C(38)	119.0(7)
C(14)-C(13)-C(12)	132.3(6)	C(39)-C(40)-C(41)	122.1(7)
C(18)-C(13)-C(12)	109.5(6)	C(42)-C(41)-C(40)	120.8(7)
C(15)-C(14)-C(13)	120.2(7)	C(41)-C(42)-C(43)	119.5(7)
C(14)-C(15)-C(16)	121.2(7)	C(42)-C(43)-C(38)	119.4(6)
C(17)-C(16)-C(15)	121.2(6)	C(42)-C(43)-C(44)	131.3(6)
C(16)-C(17)-C(18)	118.6(6)	C(38)-C(43)-C(44)	109.3(6)
C(17)-C(18)-C(13)	120.6(6)	C(36)-C(44)-C(43)	103.8(6)
C(17)-C(18)-C(19)	130.7(6)	C(1)-Sn(1)-C(2)	32.7(3)

C(13)-C(18)-C(19)	108.6(5)	C(1)-Sn(1)-C(5)	33.0(3)
C(11)-C(19)-C(18)	104.9(5)	C(2)-Sn(1)-C(5)	52.9(3)
C(27)-C(26)-C(30)	107.8(7)	C(30)-Sn(2)-C(29)	33.1(2)
C(27)-C(26)-C(31)	125.8(9)	C(30)-Sn(2)-C(26)	31.7(2)
C(30)-C(26)-C(31)	126.2(8)	C(29)-Sn(2)-C(26)	53.2(3)
C(27)-C(26)-Sn(2)	77.7(4)	C(30)-Sn(2)-C(28)	53.0(2)
C(30)-C(26)-Sn(2)	70.4(4)	C(29)-Sn(2)-C(28)	32.2(2)
C(31)-C(26)-Sn(2)	121.7(5)	C(26)-Sn(2)-C(28)	51.7(2)
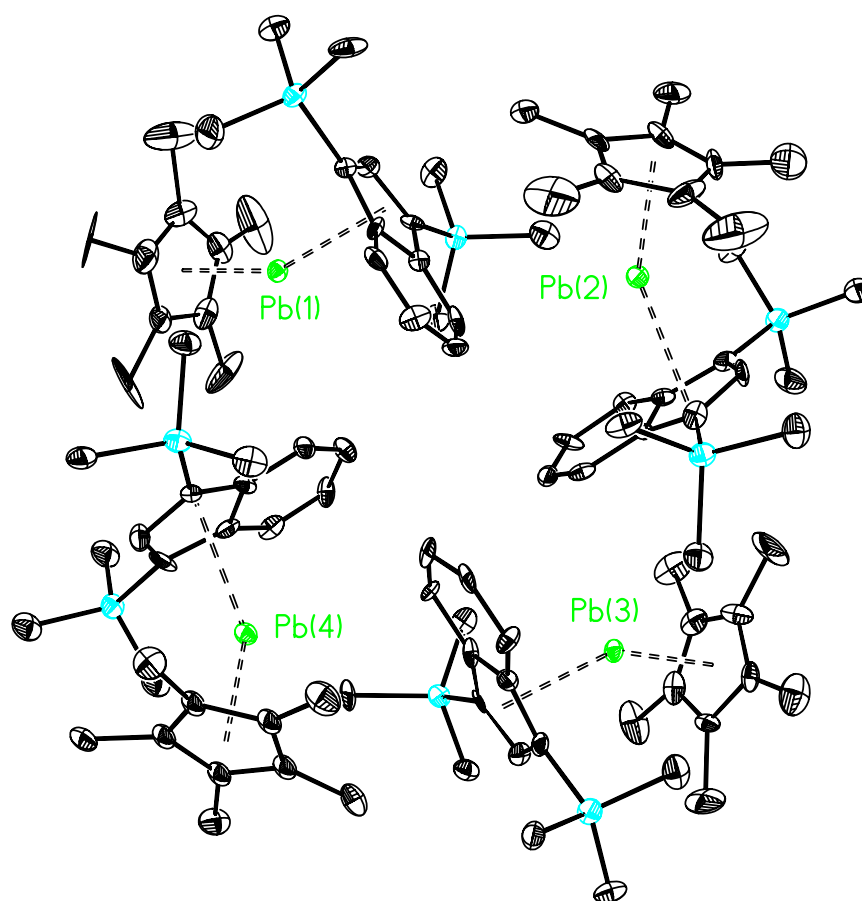
Figure 2.19 Molecular structures of the four independent molecules of [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)](η^5-C$_5$Me$_5$)Pb, 12, per unit cell. Atoms are colored as follows: carbon (black), silicon (cyan), and lead (green). The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.
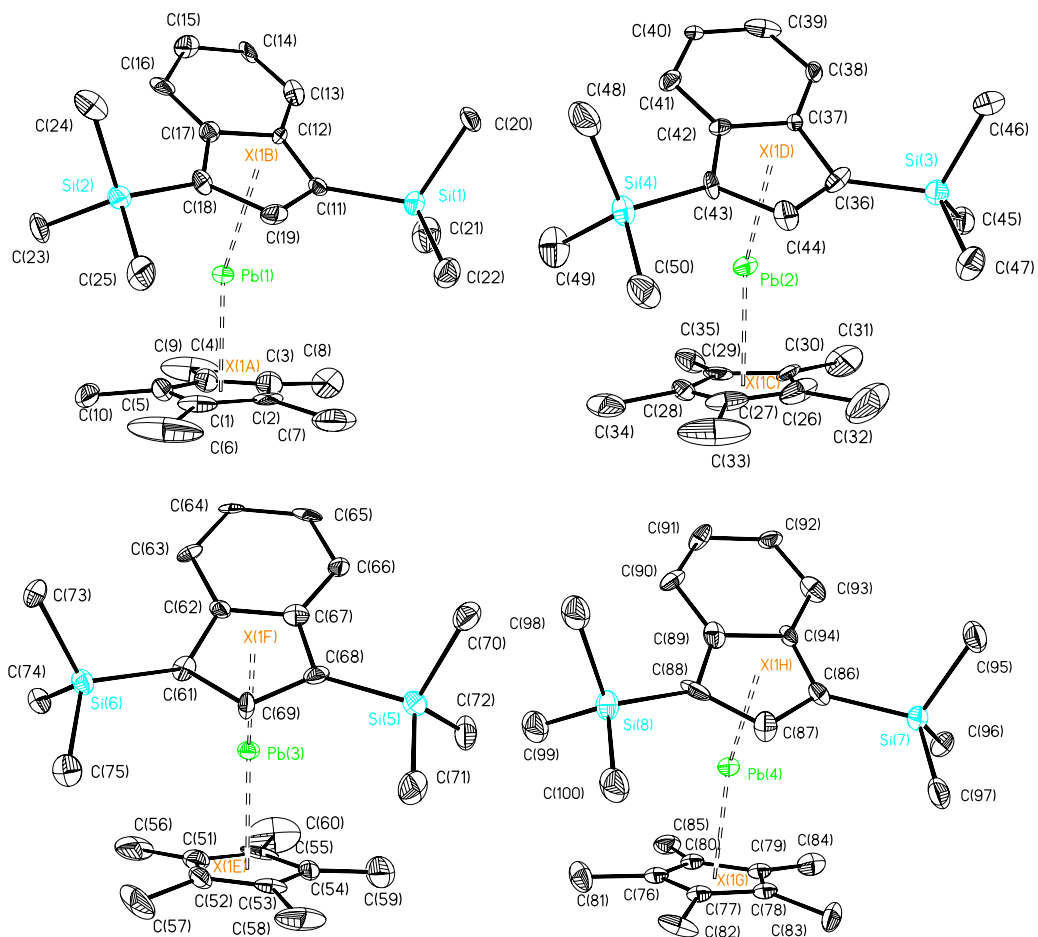
Figure 2.20 Molecular structures and complete atom numbering schemes for the four independent molecules of $[1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)](\eta^5\text{-}C_5Me_5)Pb$, **12**, per unit cell. Atoms are colored as follows: carbon (black), silicon (cyan), and lead (green). The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.

Table 2.10 Crystal data and structure refinement for [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)](η^5-C$_5$Me$_5$)Pb, 12.

Identification code	pna21
Empirical formula	C100 H152 Pb4 Si8
Formula weight	2407.70
Temperature	153(2) K
Wavelength	0.71069 Å
Crystal system	Orthorhombic
Space group	Pna2$_1$
Unit cell dimensions	a = 35.144(5) Å α= 90.000(5)°.
	b = 10.548(5) Å β= 90.000(5)°.
	c = 31.090(5) Å γ = 90.000(5)°.
Volume	11525(6) Å^3
Z	4
Density (calculated)	1.388 Mg/m^3
Absorption coefficient	5.946 mm^{-1}
F(000)	4768
Crystal size	0.20 x 0.20 x 0.20 mm^3
Theta range for data collection	2.91 to 27.50°.
Index ranges	-45$\leq$h$\leq$45, -13$\leq$k$\leq$13, -40$\leq$l$\leq$40
Reflections collected	24183
Independent reflections	24183 [R(int) = 0.0000]
Completeness to theta = 27.50°	98.5 %
Max. and min. transmission	0.3826 and 0.3826
Refinement method	Full-matrix least-squares on F^2
Goodness-of-fit on F^2	1.022
Final R indices [I>2sigma(I)]	R1 = 0.0798, wR2 = 0.1646
R indices (all data)	R1 = 0.1676, wR2 = 0.1983
Absolute structure parameter	0.491(11)
Largest diff. peak and hole	2.895 and -1.182 e.Å^{-3}

Table 2.11 Selected Bond Lengths (Å) for [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)](η^5-C$_5$Me$_5$)Pb, 12.

C(1)-C(5)	1.42(3)	C(51)-C(55)	1.37(3)
C(1)-C(6)	1.44(4)	C(51)-C(52)	1.40(3)
C(1)-C(2)	1.46(4)	C(51)-C(56)	1.54(3)
C(1)-Pb(1)	2.637(18)	C(51)-Pb(3)	2.76(2)
C(2)-C(3)	1.35(4)	C(52)-C(53)	1.38(3)
C(2)-C(7)	1.57(4)	C(52)-C(57)	1.56(3)
C(2)-Pb(1)	2.614(19)	C(52)-Pb(3)	2.595(19)
C(3)-C(4)	1.39(3)	C(53)-C(54)	1.44(3)
C(3)-C(8)	1.54(4)	C(53)-C(58)	1.55(3)
C(3)-Pb(1)	2.71(2)	C(53)-Pb(3)	2.617(18)
C(4)-C(5)	1.38(3)	C(54)-C(59)	1.41(3)
C(4)-C(9)	1.52(4)	C(54)-C(55)	1.41(3)
C(4)-Pb(1)	2.76(2)	C(54)-Pb(3)	2.766(19)
C(5)-C(10)	1.55(3)	C(55)-C(60)	1.50(4)
C(5)-Pb(1)	2.68(2)	C(55)-Pb(3)	2.78(2)
C(11)-C(19)	1.37(2)	C(61)-C(69)	1.40(3)
C(11)-C(12)	1.45(2)	C(61)-C(62)	1.51(3)
C(11)-Pb(1)	2.846(17)	C(61)-Pb(3)	2.838(16)
C(12)-C(13)	1.41(3)	C(62)-C(67)	1.40(3)
C(12)-C(17)	1.48(3)	C(62)-C(63)	1.43(3)
C(13)-C(14)	1.35(3)	C(63)-C(64)	1.36(3)
C(14)-C(15)	1.34(3)	C(64)-C(65)	1.47(3)
C(15)-C(16)	1.41(3)	C(65)-C(66)	1.37(3)
C(16)-C(17)	1.41(3)	C(66)-C(67)	1.43(3)
C(17)-C(18)	1.45(3)	C(67)-C(68)	1.41(3)
C(18)-C(19)	1.45(3)	C(68)-C(69)	1.42(2)
C(18)-Pb(1)	2.87(2)	C(69)-Pb(3)	2.760(19)
C(19)-Pb(1)	2.75(2)	C(76)-C(77)	1.40(3)
C(26)-C(30)	1.46(3)	C(76)-C(80)	1.43(3)
C(26)-C(27)	1.48(4)	C(76)-C(81)	1.52(3)
C(26)-C(32)	1.53(4)	C(76)-Pb(4)	2.704(17)
C(26)-Pb(2)	2.61(2)	C(77)-C(78)	1.41(3)
C(27)-C(28)	1.33(3)	C(77)-C(82)	1.55(3)
C(27)-C(33)	1.42(4)	C(77)-Pb(4)	2.606(16)
C(27)-Pb(2)	2.64(2)	C(78)-C(79)	1.43(3)

C(28)-C(29)	1.40(3)	C(78)-C(83)	1.53(3)
C(28)-C(34)	1.53(3)	C(78)-Pb(4)	2.585(16)
C(28)-Pb(2)	2.73(2)	C(79)-C(80)	1.45(3)
C(29)-C(30)	1.44(3)	C(79)-C(84)	1.48(3)
C(29)-C(35)	1.54(3)	C(79)-Pb(4)	2.702(18)
C(29)-Pb(2)	2.78(2)	C(80)-C(85)	1.48(3)
C(30)-C(31)	1.58(4)	C(80)-Pb(4)	2.783(18)
C(30)-Pb(2)	2.698(19)	C(86)-C(94)	1.42(3)
C(36)-C(44)	1.38(3)	C(86)-C(87)	1.42(3)
C(36)-C(37)	1.47(3)	C(87)-C(88)	1.45(3)
C(37)-C(42)	1.43(2)	C(87)-Pb(4)	2.81(2)
C(37)-C(38)	1.44(3)	C(88)-C(89)	1.44(3)
C(38)-C(39)	1.39(3)	C(88)-Pb(4)	2.87(2)
C(39)-C(40)	1.38(3)	C(89)-C(94)	1.45(2)
C(40)-C(41)	1.38(3)	C(89)-C(90)	1.46(3)
C(41)-C(42)	1.39(2)	C(90)-C(91)	1.32(3)
C(42)-C(43)	1.43(3)	C(91)-C(92)	1.43(3)
C(43)-C(44)	1.43(3)	C(92)-C(93)	1.42(3)
C(43)-Pb(2)	2.86(2)	C(93)-C(94)	1.37(3)
C(44)-Pb(2)	2.74(2)		

Table 2.12 Selected Bond Angles (°) for $[1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)](\eta^5\text{-}C_5Me_5)Pb$, **12**.

C(5)-C(1)-C(6)	125(3)	C(61)-C(69)-Pb(3)	78.7(10)
C(5)-C(1)-C(2)	103(2)	C(68)-C(69)-Pb(3)	80.2(11)
C(6)-C(1)-C(2)	132(3)	C(77)-C(76)-C(80)	111.0(18)
C(5)-C(1)-Pb(1)	76.3(10)	C(77)-C(76)-C(81)	128.2(19)
C(6)-C(1)-Pb(1)	122.4(18)	C(80)-C(76)-C(81)	120.5(19)
C(2)-C(1)-Pb(1)	73.0(10)	C(77)-C(76)-Pb(4)	70.9(8)
C(3)-C(2)-C(1)	111(2)	C(80)-C(76)-Pb(4)	78.0(10)
C(3)-C(2)-C(7)	124(3)	C(81)-C(76)-Pb(4)	123.4(13)
C(1)-C(2)-C(7)	124(3)	C(76)-C(77)-C(78)	106.8(18)
C(3)-C(2)-Pb(1)	79.1(12)	C(76)-C(77)-C(82)	127(2)
C(1)-C(2)-Pb(1)	74.8(10)	C(78)-C(77)-C(82)	125.3(19)
C(7)-C(2)-Pb(1)	120.7(15)	C(76)-C(77)-Pb(4)	78.7(9)
C(2)-C(3)-C(4)	107(2)	C(78)-C(77)-Pb(4)	73.4(8)
C(2)-C(3)-C(8)	127(3)	C(82)-C(77)-Pb(4)	121.3(12)
C(4)-C(3)-C(8)	125(3)	C(77)-C(78)-C(79)	109.4(18)
C(2)-C(3)-Pb(1)	71.5(12)	C(77)-C(78)-C(83)	126.5(19)
C(4)-C(3)-Pb(1)	77.5(12)	C(79)-C(78)-C(83)	123.3(19)
C(8)-C(3)-Pb(1)	123.4(16)	C(77)-C(78)-Pb(4)	75.0(8)
C(5)-C(4)-C(3)	109(2)	C(79)-C(78)-Pb(4)	78.9(10)
C(5)-C(4)-C(9)	127(3)	C(83)-C(78)-Pb(4)	120.8(13)
C(3)-C(4)-C(9)	123(3)	C(78)-C(79)-C(80)	107.1(17)
C(5)-C(4)-Pb(1)	72.1(12)	C(78)-C(79)-C(84)	126.5(19)
C(3)-C(4)-Pb(1)	73.1(12)	C(80)-C(79)-C(84)	126.5(19)
C(9)-C(4)-Pb(1)	121.6(16)	C(78)-C(79)-Pb(4)	69.9(10)
C(4)-C(5)-C(1)	110(2)	C(80)-C(79)-Pb(4)	77.7(10)
C(4)-C(5)-C(10)	121(3)	C(84)-C(79)-Pb(4)	119.0(13)
C(1)-C(5)-C(10)	129(3)	C(76)-C(80)-C(79)	105.6(17)
C(4)-C(5)-Pb(1)	78.6(12)	C(76)-C(80)-C(85)	127.3(19)
C(1)-C(5)-Pb(1)	72.9(11)	C(79)-C(80)-C(85)	127.0(19)
C(10)-C(5)-Pb(1)	123.8(15)	C(76)-C(80)-Pb(4)	71.9(10)
C(19)-C(11)-C(12)	108.8(16)	C(79)-C(80)-Pb(4)	71.6(10)
C(19)-C(11)-Pb(1)	71.9(11)	C(85)-C(80)-Pb(4)	122.5(13)
C(12)-C(11)-Pb(1)	81.6(10)	C(94)-C(86)-C(87)	108.3(17)
C(13)-C(12)-C(11)	136.6(17)	C(86)-C(87)-C(88)	110.8(18)
C(13)-C(12)-C(17)	118.0(17)	C(86)-C(87)-Pb(4)	77.9(12)

122

C(11)-C(12)-C(17)	105.2(16)	C(88)-C(87)-Pb(4)	77.4(12)
C(14)-C(13)-C(12)	121.1(18)	C(89)-C(88)-C(87)	103.4(17)
C(15)-C(14)-C(13)	121.9(18)	C(89)-C(88)-Pb(4)	78.8(12)
C(14)-C(15)-C(16)	122.2(19)	C(87)-C(88)-Pb(4)	73.0(12)
C(15)-C(16)-C(17)	118.6(19)	Si(8)-C(88)-Pb(4)	121.6(9)
C(16)-C(17)-C(18)	132.2(18)	C(88)-C(89)-C(94)	111.1(17)
C(16)-C(17)-C(12)	118.2(17)	C(88)-C(89)-C(90)	130.8(19)
C(18)-C(17)-C(12)	109.6(17)	C(94)-C(89)-C(90)	117.9(19)
C(19)-C(18)-C(17)	103.8(16)	C(91)-C(90)-C(89)	121.2(19)
C(19)-C(18)-Pb(1)	70.5(11)	C(90)-C(91)-C(92)	120.8(17)
C(17)-C(18)-Pb(1)	81.3(11)	C(93)-C(92)-C(91)	119.5(18)
C(11)-C(19)-C(18)	112.5(16)	C(94)-C(93)-C(92)	120.8(19)
C(11)-C(19)-Pb(1)	79.8(11)	C(93)-C(94)-C(86)	134.1(17)
C(18)-C(19)-Pb(1)	79.6(11)	C(93)-C(94)-C(89)	119.7(18)
C(30)-C(26)-C(27)	106.2(19)	C(86)-C(94)-C(89)	106.2(17)
C(30)-C(26)-C(32)	130(3)	C(2)-Pb(1)-C(1)	32.2(8)
C(27)-C(26)-C(32)	123(3)	C(2)-Pb(1)-C(5)	50.1(7)
C(30)-C(26)-Pb(2)	77.3(12)	C(1)-Pb(1)-C(5)	30.9(7)
C(27)-C(26)-Pb(2)	74.5(12)	C(2)-Pb(1)-C(3)	29.3(9)
C(32)-C(26)-Pb(2)	115.3(17)	C(1)-Pb(1)-C(3)	51.3(8)
C(28)-C(27)-C(33)	130(3)	C(5)-Pb(1)-C(3)	49.5(7)
C(28)-C(27)-C(26)	107.8(19)	C(2)-Pb(1)-C(19)	101.6(7)
C(33)-C(27)-C(26)	121(3)	C(1)-Pb(1)-C(19)	104.0(6)
C(28)-C(27)-Pb(2)	79.6(13)	C(5)-Pb(1)-C(19)	133.0(7)
C(33)-C(27)-Pb(2)	121.9(17)	C(3)-Pb(1)-C(19)	125.2(8)
C(26)-C(27)-Pb(2)	72.8(13)	C(2)-Pb(1)-C(4)	48.3(8)
C(27)-C(28)-C(29)	112(2)	C(1)-Pb(1)-C(4)	50.0(7)
C(27)-C(28)-C(34)	122(2)	C(5)-Pb(1)-C(4)	29.3(7)
C(29)-C(28)-C(34)	125(2)	C(3)-Pb(1)-C(4)	29.3(7)
C(27)-C(28)-Pb(2)	71.8(13)	C(19)-Pb(1)-C(4)	149.8(6)
C(29)-C(28)-Pb(2)	77.3(12)	C(2)-Pb(1)-C(11)	107.4(6)
C(34)-C(28)-Pb(2)	125.4(14)	C(1)-Pb(1)-C(11)	123.8(6)
C(28)-C(29)-C(30)	107.8(19)	C(5)-Pb(1)-C(11)	154.5(6)
C(28)-C(29)-C(35)	128(2)	C(3)-Pb(1)-C(11)	118.7(7)
C(30)-C(29)-C(35)	124(2)	C(19)-Pb(1)-C(11)	28.3(5)
C(28)-C(29)-Pb(2)	73.2(13)	C(4)-Pb(1)-C(11)	147.7(6)
C(30)-C(29)-Pb(2)	71.6(12)	C(2)-Pb(1)-C(18)	123.1(8)
C(35)-C(29)-Pb(2)	121.7(13)	C(1)-Pb(1)-C(18)	110.3(6)
C(29)-C(30)-C(26)	106(2)	C(5)-Pb(1)-C(18)	128.2(6)

123

C(29)-C(30)-C(31)	129(2)	C(3)-Pb(1)-C(18)	151.5(8)
C(26)-C(30)-C(31)	125(2)	C(19)-Pb(1)-C(18)	29.9(5)
C(29)-C(30)-Pb(2)	78.1(11)	C(4)-Pb(1)-C(18)	157.4(6)
C(26)-C(30)-Pb(2)	70.9(12)	C(11)-Pb(1)-C(18)	48.5(5)
C(31)-C(30)-Pb(2)	119.3(14)	C(26)-Pb(2)-C(27)	32.7(9)
C(44)-C(36)-C(37)	102.6(16)	C(26)-Pb(2)-C(30)	31.8(7)
C(42)-C(37)-C(38)	119.4(15)	C(27)-Pb(2)-C(30)	52.2(7)
C(42)-C(37)-C(36)	109.4(17)	C(26)-Pb(2)-C(28)	50.2(7)
C(38)-C(37)-C(36)	131.1(16)	C(27)-Pb(2)-C(28)	28.6(7)
C(39)-C(38)-C(37)	117.1(17)	C(30)-Pb(2)-C(28)	50.0(6)
C(40)-C(39)-C(38)	122(2)	C(26)-Pb(2)-C(44)	100.7(7)
C(39)-C(40)-C(41)	122.1(18)	C(27)-Pb(2)-C(44)	98.8(7)
C(40)-C(41)-C(42)	118.4(17)	C(30)-Pb(2)-C(44)	130.0(6)
C(41)-C(42)-C(43)	130.6(17)	C(28)-Pb(2)-C(44)	123.4(6)
C(41)-C(42)-C(37)	120.8(18)	C(26)-Pb(2)-C(29)	50.6(6)
C(43)-C(42)-C(37)	108.7(15)	C(27)-Pb(2)-C(29)	49.4(6)
C(42)-C(43)-C(44)	103.8(15)	C(30)-Pb(2)-C(29)	30.3(6)
C(42)-C(43)-Pb(2)	81.6(11)	C(28)-Pb(2)-C(29)	29.5(6)
C(44)-C(43)-Pb(2)	70.3(11)	C(44)-Pb(2)-C(29)	147.7(6)
C(36)-C(44)-C(43)	115.4(17)	C(26)-Pb(2)-C(43)	122.2(8)
C(36)-C(44)-Pb(2)	82.4(12)	C(27)-Pb(2)-C(43)	106.0(6)
C(43)-C(44)-Pb(2)	80.1(11)	C(30)-Pb(2)-C(43)	154.0(6)
C(55)-C(51)-C(52)	106.9(18)	C(28)-Pb(2)-C(43)	119.0(6)
C(55)-C(51)-C(56)	125(2)	C(44)-Pb(2)-C(43)	29.6(5)
C(52)-C(51)-C(56)	128(2)	C(29)-Pb(2)-C(43)	148.1(6)
C(55)-C(51)-Pb(3)	76.4(12)	C(52)-Pb(3)-C(53)	30.7(6)
C(52)-C(51)-Pb(3)	68.5(11)	C(52)-Pb(3)-C(51)	30.2(7)
C(56)-C(51)-Pb(3)	119.0(15)	C(53)-Pb(3)-C(51)	49.7(6)
C(53)-C(52)-C(51)	108.9(18)	C(52)-Pb(3)-C(69)	104.0(7)
C(53)-C(52)-C(57)	122(2)	C(53)-Pb(3)-C(69)	98.6(6)
C(51)-C(52)-C(57)	128(2)	C(51)-Pb(3)-C(69)	132.8(6)
C(53)-C(52)-Pb(3)	75.5(11)	C(52)-Pb(3)-C(54)	50.5(6)
C(51)-C(52)-Pb(3)	81.3(12)	C(53)-Pb(3)-C(54)	30.9(6)
C(57)-C(52)-Pb(3)	119.2(15)	C(51)-Pb(3)-C(54)	49.2(6)
C(52)-C(53)-C(54)	108.8(18)	C(69)-Pb(3)-C(54)	122.6(6)
C(52)-C(53)-C(58)	123(2)	C(52)-Pb(3)-C(55)	48.9(7)
C(54)-C(53)-C(58)	127(2)	C(53)-Pb(3)-C(55)	49.2(6)
C(52)-C(53)-Pb(3)	73.8(11)	C(51)-Pb(3)-C(55)	28.7(7)
C(54)-C(53)-Pb(3)	80.3(11)	C(69)-Pb(3)-C(55)	147.8(6)

124

C(58)-C(53)-Pb(3)	118.2(14)	C(54)-Pb(3)-C(55)	29.6(6)
C(59)-C(54)-C(55)	131(2)	C(52)-Pb(3)-C(61)	108.7(6)
C(59)-C(54)-C(53)	125(2)	C(53)-Pb(3)-C(61)	117.7(6)
C(55)-C(54)-C(53)	104.2(18)	C(51)-Pb(3)-C(61)	127.9(6)
C(59)-C(54)-Pb(3)	118.1(16)	C(69)-Pb(3)-C(61)	28.9(5)
C(55)-C(54)-Pb(3)	75.6(12)	C(54)-Pb(3)-C(61)	147.3(6)
C(53)-C(54)-Pb(3)	68.9(10)	C(55)-Pb(3)-C(61)	156.2(6)
C(51)-C(55)-C(54)	111.1(19)	C(78)-Pb(4)-C(77)	31.6(6)
C(51)-C(55)-C(60)	124(3)	C(78)-Pb(4)-C(79)	31.3(6)
C(54)-C(55)-C(60)	125(3)	C(77)-Pb(4)-C(79)	51.8(6)
C(51)-C(55)-Pb(3)	74.9(13)	C(78)-Pb(4)-C(76)	50.5(6)
C(54)-C(55)-Pb(3)	74.8(11)	C(77)-Pb(4)-C(76)	30.4(6)
C(60)-C(55)-Pb(3)	121.0(16)	C(79)-Pb(4)-C(76)	50.2(6)
C(69)-C(61)-C(62)	102.1(15)	C(78)-Pb(4)-C(80)	51.0(6)
C(69)-C(61)-Pb(3)	72.4(10)	C(77)-Pb(4)-C(80)	51.0(6)
C(62)-C(61)-Pb(3)	79.9(8)	C(79)-Pb(4)-C(80)	30.7(6)
C(67)-C(62)-C(63)	120.8(18)	C(76)-Pb(4)-C(80)	30.1(6)
C(67)-C(62)-C(61)	108.4(17)	C(78)-Pb(4)-C(87)	101.9(6)
C(63)-C(62)-C(61)	130.6(18)	C(77)-Pb(4)-C(87)	97.2(6)
C(64)-C(63)-C(62)	120.8(19)	C(79)-Pb(4)-C(87)	131.2(6)
C(63)-C(64)-C(65)	118.6(18)	C(76)-Pb(4)-C(87)	122.1(6)
C(66)-C(65)-C(64)	120.7(17)	C(80)-Pb(4)-C(87)	148.2(6)
C(65)-C(66)-C(67)	120.4(18)	C(78)-Pb(4)-C(88)	123.3(6)
C(62)-C(67)-C(68)	110.4(17)	C(77)-Pb(4)-C(88)	105.1(6)
C(62)-C(67)-C(66)	118.5(18)	C(79)-Pb(4)-C(88)	154.4(6)
C(68)-C(67)-C(66)	131.1(19)	C(76)-Pb(4)-C(88)	117.2(6)
C(67)-C(68)-C(69)	104.8(17)	C(80)-Pb(4)-C(88)	146.5(6)
C(61)-C(69)-C(68)	114.3(17)	C(87)-Pb(4)-C(88)	29.6(6)
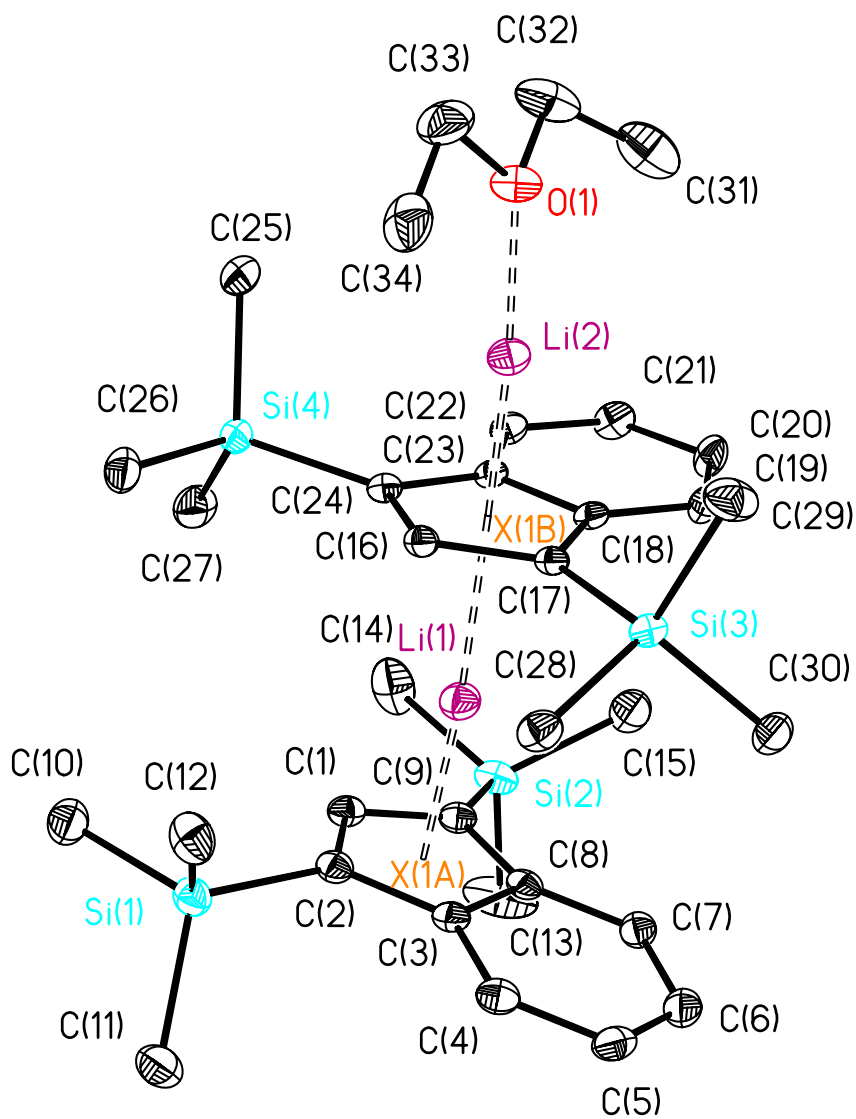
Figure 2.21 Molecular structure and complete atom numbering scheme for one of the two independent molecules in the crystal structure of [1,3-$(Me_3Si)_2(\eta^5-C_9H_5)]_2Li_2 \cdot Et_2O$, **13**. Atoms are colored as follows: carbon (black), silicon (**cyan**), lithium (**violet**), and oxygen (**red**). The thermal ellipsoids are shown at the 30% probability level. All hydrogen atoms have been omitted for clarity.

Table 2.13 Crystal Data and Structure Refinement for [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)]$_2$Li$_2$·Et$_2$O, 13.

Identification code	p21overc
Empirical formula	C34 H56 Cl0 Li2 Mg0 O P0 Si4
Formula weight	607.03
Temperature	153(2) K
Wavelength	0.71069 Å
Crystal system	Monoclinic
Space group	P2(1)/c
Unit cell dimensions	a = 9.959(5) Å α= 90.000(5)°.
	b = 21.193(5) Å β= 93.506(5)°.
	c = 18.068(5) Å γ = 90.000(5)°.
Volume	3806(2) Å^3
Z	4
Density (calculated)	1.059 Mg/m^3
Absorption coefficient	0.179 mm^{-1}
F(000)	1320
Crystal size	0.20 x 0.20 x 0.20 mm^3
Theta range for data collection	2.96 to 27.50°.
Index ranges	-12<=h<=12, -27<=k<=27, -23<=l<=23
Reflections collected	16921
Independent reflections	8721 [R(int) = 0.0951]
Completeness to theta = 27.50°	99.7 %
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	8721 / 0 / 590
Goodness-of-fit on F^2	1.015
Final R indices [I>2sigma(I)]	R1 = 0.0569, wR2 = 0.0929
R indices (all data)	R1 = 0.1489, wR2 = 0.1144
Extinction coefficient	0.0023(4)
Largest diff. peak and hole	0.264 and -0.349 e.Å^{-3}

127

Table 2.14 Selected Bond Lengths (Å) for [1,3-(Me$_3$Si)$_2$(η^5-C$_9$H$_5$)]$_2$Li$_2$·Et$_2$O, 13.

C(1)-C(9)	1.419(3)	C(17)-Li(2)	2.280(5)
C(1)-C(2)	1.422(3)	C(17)-Li(1)	2.352(4)
C(1)-Li(1)	2.212(5)	C(18)-C(19)	1.415(3)
C(2)-C(3)	1.444(3)	C(18)-C(23)	1.439(3)
C(2)-Li(1)	2.342(5)	C(18)-Li(2)	2.263(4)
C(3)-C(4)	1.408(3)	C(18)-Li(1)	2.570(5)
C(3)-C(8)	1.444(3)	C(19)-C(20)	1.369(3)
C(3)-Li(1)	2.433(5)	C(20)-C(21)	1.408(4)
C(4)-C(5)	1.370(4)	C(21)-C(22)	1.363(3)
C(5)-C(6)	1.410(4)	C(22)-C(23)	1.421(3)
C(6)-C(7)	1.368(4)	C(23)-C(24)	1.446(3)
C(7)-C(8)	1.412(3)	C(23)-Li(2)	2.262(4)
C(8)-C(9)	1.440(3)	C(23)-Li(1)	2.678(5)
C(8)-Li(1)	2.382(5)	C(24)-Li(2)	2.267(4)
C(9)-Li(1)	2.256(5)	C(24)-Li(1)	2.557(5)
C(16)-C(24)	1.421(3)	C(31)-C(32)	1.472(5)
C(16)-C(17)	1.436(3)	C(32)-O(1)	1.450(4)
C(16)-Li(2)	2.232(5)	C(33)-O(1)	1.445(3)
C(16)-Li(1)	2.301(5)	C(33)-C(34)	1.479(5)
C(17)-C(18)	1.454(3)	Li(2)-O(1)	1.877(4)
C(1)-C(9)	1.419(3)	C(17)-Li(2)	2.280(5)
C(1)-C(2)	1.422(3)	C(17)-Li(1)	2.352(4)
C(1)-Li(1)	2.212(5)	C(18)-C(19)	1.415(3)
C(2)-C(3)	1.444(3)	C(18)-C(23)	1.439(3)

Table 2.15 Selected Bond Angles (º) for [1,3-(Me₃Si)₂(η⁵-C₉H₅)]₂Li₂·Et₂O, 13.

C(9)-C(1)-C(2)	113.1(2)	Li(2)-C(24)-Li(1)	114.38(16)
C(9)-C(1)-Li(1)	73.17(17)	O(1)-C(32)-C(31)	109.0(3)
C(2)-C(1)-Li(1)	76.88(17)	O(1)-C(33)-C(34)	108.5(2)
C(1)-C(2)-C(3)	104.5(2)	C(1)-Li(1)-C(9)	37.03(10)
C(1)-C(2)-Li(1)	66.87(16)	C(1)-Li(1)-C(16)	132.5(2)
C(3)-C(2)-Li(1)	75.86(16)	C(9)-Li(1)-C(16)	163.2(2)
C(4)-C(3)-C(2)	131.8(2)	C(1)-Li(1)-C(2)	36.25(10)
C(4)-C(3)-C(8)	119.3(2)	C(9)-Li(1)-C(2)	62.05(13)
C(2)-C(3)-C(8)	108.8(2)	C(16)-Li(1)-C(2)	118.16(19)
C(4)-C(3)-Li(1)	129.75(19)	C(1)-Li(1)-C(17)	168.3(2)
C(2)-C(3)-Li(1)	69.01(16)	C(9)-Li(1)-C(17)	154.3(2)
C(8)-C(3)-Li(1)	70.62(16)	C(16)-Li(1)-C(17)	35.93(10)
C(5)-C(4)-C(3)	120.0(3)	C(2)-Li(1)-C(17)	138.9(2)
C(4)-C(5)-C(6)	120.6(3)	C(1)-Li(1)-C(8)	58.96(13)
C(7)-C(6)-C(5)	121.2(3)	C(9)-Li(1)-C(8)	36.04(10)
C(6)-C(7)-C(8)	119.8(2)	C(16)-Li(1)-C(8)	160.4(2)
C(7)-C(8)-C(9)	132.0(2)	C(2)-Li(1)-C(8)	59.60(12)
C(7)-C(8)-C(3)	119.0(2)	C(17)-Li(1)-C(8)	131.17(19)
C(9)-C(8)-C(3)	108.7(2)	C(1)-Li(1)-C(3)	58.18(13)
C(7)-C(8)-Li(1)	128.88(19)	C(9)-Li(1)-C(3)	59.85(13)
C(9)-C(8)-Li(1)	67.21(16)	C(16)-Li(1)-C(3)	131.3(2)
C(3)-C(8)-Li(1)	74.49(16)	C(2)-Li(1)-C(3)	35.13(10)
C(1)-C(9)-C(8)	104.8(2)	C(17)-Li(1)-C(3)	125.3(2)
C(1)-C(9)-Li(1)	69.80(16)	C(8)-Li(1)-C(3)	34.89(9)
C(8)-C(9)-Li(1)	76.75(17)	C(1)-Li(1)-C(24)	113.67(18)
C(24)-C(16)-C(17)	113.0(2)	C(9)-Li(1)-C(24)	130.5(2)
C(24)-C(16)-Li(2)	72.92(17)	C(16)-Li(1)-C(24)	33.49(9)
C(17)-C(16)-Li(2)	73.24(16)	C(2)-Li(1)-C(24)	122.90(18)
C(24)-C(16)-Li(1)	83.18(17)	C(17)-Li(1)-C(24)	57.86(12)
C(17)-C(16)-Li(1)	73.99(16)	C(8)-Li(1)-C(24)	165.8(2)
Li(2)-C(16)-Li(1)	127.00(18)	C(3)-Li(1)-C(24)	154.4(2)
C(16)-C(17)-C(18)	104.27(19)	C(1)-Li(1)-C(18)	150.1(2)
C(16)-C(17)-Li(2)	69.66(16)	C(9)-Li(1)-C(18)	125.12(19)
C(18)-C(17)-Li(2)	70.73(16)	C(16)-Li(1)-C(18)	55.52(11)
C(16)-C(17)-Li(1)	70.09(15)	C(2)-Li(1)-C(18)	172.7(2)
C(18)-C(17)-Li(1)	81.25(16)	C(17)-Li(1)-C(18)	33.99(9)

Li(2)-C(17)-Li(1)	122.30(17)	C(8)-Li(1)-C(18)	124.80(18)
C(19)-C(18)-C(23)	119.4(2)	C(3)-Li(1)-C(18)	144.71(19)
C(19)-C(18)-C(17)	131.7(2)	C(24)-Li(1)-C(18)	54.58(11)
C(23)-C(18)-C(17)	108.81(18)	C(1)-Li(1)-C(23)	122.74(19)
C(19)-C(18)-Li(2)	120.21(19)	C(9)-Li(1)-C(23)	116.53(19)
C(23)-C(18)-Li(2)	71.41(16)	C(16)-Li(1)-C(23)	53.65(11)
C(17)-C(18)-Li(2)	71.95(16)	C(2)-Li(1)-C(23)	149.53(19)
C(19)-C(18)-Li(1)	125.78(18)	C(17)-Li(1)-C(23)	55.31(11)
C(23)-C(18)-Li(1)	78.25(16)	C(8)-Li(1)-C(23)	139.07(19)
C(17)-C(18)-Li(1)	64.76(15)	C(3)-Li(1)-C(23)	173.62(19)
Li(2)-C(18)-Li(1)	114.00(16)	C(24)-Li(1)-C(23)	31.96(8)
C(20)-C(19)-C(18)	119.2(2)	C(18)-Li(1)-C(23)	31.75(8)
C(19)-C(20)-C(21)	121.4(2)	O(1)-Li(2)-C(16)	154.2(2)
C(22)-C(21)-C(20)	121.3(2)	O(1)-Li(2)-C(23)	132.4(2)
C(21)-C(22)-C(23)	119.3(2)	C(16)-Li(2)-C(23)	60.72(13)
C(22)-C(23)-C(18)	119.3(2)	O(1)-Li(2)-C(18)	144.4(2)
C(22)-C(23)-C(24)	131.5(2)	C(16)-Li(2)-C(18)	60.98(13)
C(18)-C(23)-C(24)	109.12(19)	C(23)-Li(2)-C(18)	37.09(10)
C(22)-C(23)-Li(2)	121.24(19)	O(1)-Li(2)-C(24)	135.7(2)
C(18)-C(23)-Li(2)	71.50(16)	C(16)-Li(2)-C(24)	36.80(10)
C(24)-C(23)-Li(2)	71.57(16)	C(23)-Li(2)-C(24)	37.23(10)
C(22)-C(23)-Li(1)	128.20(18)	C(18)-Li(2)-C(24)	62.51(13)
C(18)-C(23)-Li(1)	70.00(15)	O(1)-Li(2)-C(17)	161.0(2)
C(24)-C(23)-Li(1)	69.40(15)	C(16)-Li(2)-C(17)	37.10(10)
Li(2)-C(23)-Li(1)	110.15(17)	C(23)-Li(2)-C(17)	62.39(13)
C(16)-C(24)-C(23)	104.84(19)	C(18)-Li(2)-C(17)	37.32(10)
C(16)-C(24)-Li(2)	70.27(16)	C(24)-Li(2)-C(17)	63.17(13)
C(23)-C(24)-Li(2)	71.20(16)	C(33)-O(1)-C(32)	112.8(2)
C(16)-C(24)-Li(1)	63.33(15)	C(33)-O(1)-Li(2)	117.0(2)
C(23)-C(24)-Li(1)	78.63(15)	C(32)-O(1)-Li(2)	127.3(2)

130

REFERENCES

1. Calhorda, M. J.; Romao, C. C.; Veiros, L. F., *Chemistry A European Journal* **2002,** *8,* 868.
2. O'Connor, J. M.; Casey, C. P., *Chem. Rev.* **1987,** *87,* 307.
3. Anh, N. T.; Elian, M.; Hoffman, R., *J. Am. Chem. Soc.* **1978,** *100,* 110.
4. Rerek, M. E.; Basolo, F., *Organometallics* **1983,** *2,* 372.
5. Rerek, M. E.; Ji, L.-N.; Basolo, F., *J. Chem. Soc., Chem. Comm.* **1983,** 1208.
6. Ji, L.-N.; Rerek, M. E.; Basolo, F., *Organometallics* **1984,** *3,* 740.
7. Jones, D. J.; Mawby, R. J., *Inorg. Chim. Acta* **1972,** *6,* 157.
8. Hart-Davis, A. J.; Mawby, R. J., *J. Chem. Soc.* **1969,** *A,* 2403.
9. Hart-Davis, A. J.; White, C.; Mawby, R. J., *Inorg. Chim. Acta* **1970,** *4,* 441.
10. White, C.; Mawby, R. J., *Inorg. Chim. Acta* **1970,** *4,* 261.
11. Goncalves, I. S.; Ribeiro-Claro, P.; Romao, C. C.; Royo, B.; Taveres, Z. M., *J. Organomet. Chem.* **2002,** *648,* 270.
12. Bursten, B. E.; Callstrom, M. R.; Jolly, C. A.; Paquette, L. A.; Sivik, M. R.; Tucher, R. S.; Wartchow, C. A., *Organometallics* **1994,** *13,* 127.
13. Gassman, P. G.; Deck, P. A.; Winter, C. H.; Dobbs, D. A.; Cao, D. H., *Organometallics* **1992,** *11,* 959.
14. Brady, E. D.; Overby, J. S.; Meredith, M. B.; Mussman, A. B.; Cohn, M. A.; Hanusa, T. P.; Yee, G. T.; Pink, M., *J. Am. Chem. Soc.* **2002,** *124,* 9556.
15. Trotter, J., *Acta Cryst.* **1958,** *11,* 355.
16. Bocarsly, J. R.; Floriani, C.; Chiesi-Villa, A.; Gaustini, C., *Inorg. Chem.* **1987,** *26,* 1871.
17. Webb, N. C.; Marsh, R. E., *Acta Cryst.* **1967,** *22,* 382.
18. Westcott, S. A.; Kakkar, A. K.; Stringer, G.; Taylor, N. J.; Marder, T. B., *J. Organomet. Chem.* **1990,** *394,* 777.
19. Cowley, A. H.; Mardoner, M. G.; Avendano, S.; Roman, E.; Manriquez, J. M.; Carrano, C. J., *Polyhedron* **1993,** *12,* 125.
20. Overby, J. S.; Hanusa, T. P.; Boyle, P. D., *Angew. Chem., Int. Ed., Engl.* **1997,** *36,* 2378.
21. Merola, J. S.; Kacmarcik, R. T., *Organometallics* **1989,** *8,* 778.
22. Jutzi, P.; Kohl, F.; Hofmann, P.; Krüger, C.; Tsay, Y.-H., *Chem. Ber.* **1980,** *113,* 757.
23. Heeg, M. J.; Janiak, C.; Zuckerman, J. J., *J. Am. Chem. Soc.* **1984,** *106,* 4259.

24. Williamson, R. L.; Hall, M. B., *Organometallics* **1986,** *5,* 2142.
25. Atwood, J. L.; Hunter, W. E., *J. Chem. Soc., Chem. Comm.* **1981,** 925.
26. Malaba, D.; Chen, L.; Tessier, C. A.; Youngs, W. J., *Organometallics* **1992,** *11,* 1007.
27. Dohmeier, C.; Baum, E.; Ecker, A.; Koppe, R.; Schnockel, H., *Organometallics* **1996,** *15,* 4702.
28. Bos, K. D.; Bulten, E. J.; Noltes, J. G.; Spek, A. L., *J. Organomet. Chem.* **1975,** *99,* 71.
29. Jutzi, P.; Dickbreder, R.; Nöth, H., *Chem. Ber.* **1989,** *122,* 865.
30. Sheldrick, G. M., *Shelxtl Pc*. 1994, Siemens Analytical X-ray Instruments, Inc.

CHAPTER 3: DHQS AND EPSPS, STRUCTURE-FUNCTION STUDIES ON TWO SHIKIMATE PATHWAY ENZYMES

INTRODUCTION

The seven step shikimate pathway links metabolism of carbohydrates to biosynthesis of aromatic compounds (Figure 3.1). Phospho*enol*pyruvate (PEP) and erythrose 4-phosphate (E4P) are converted to chorismate, the precursor of many aromatic amino acids. The shikimate pathway is found only in microorganisms and plants, leaving animals to obtain most of their aromatic amino acids *via* diet.[1-3]

The sixth step of the shikimate pathway is the target of the inhibitor glyphosate (*N*-phophomethylglycine)[4] the active ingredient in the herbicide Roundup® (Monsanto). Glyphosate is under investigation as a lead compound against parasites which cause diseases like toxoplasmosis and malaria.[5] Other branches from the shikimate pathway lead to the production of vitamins E and K.[3] For such reasons, the shikimate pathway is of great commercial interest.
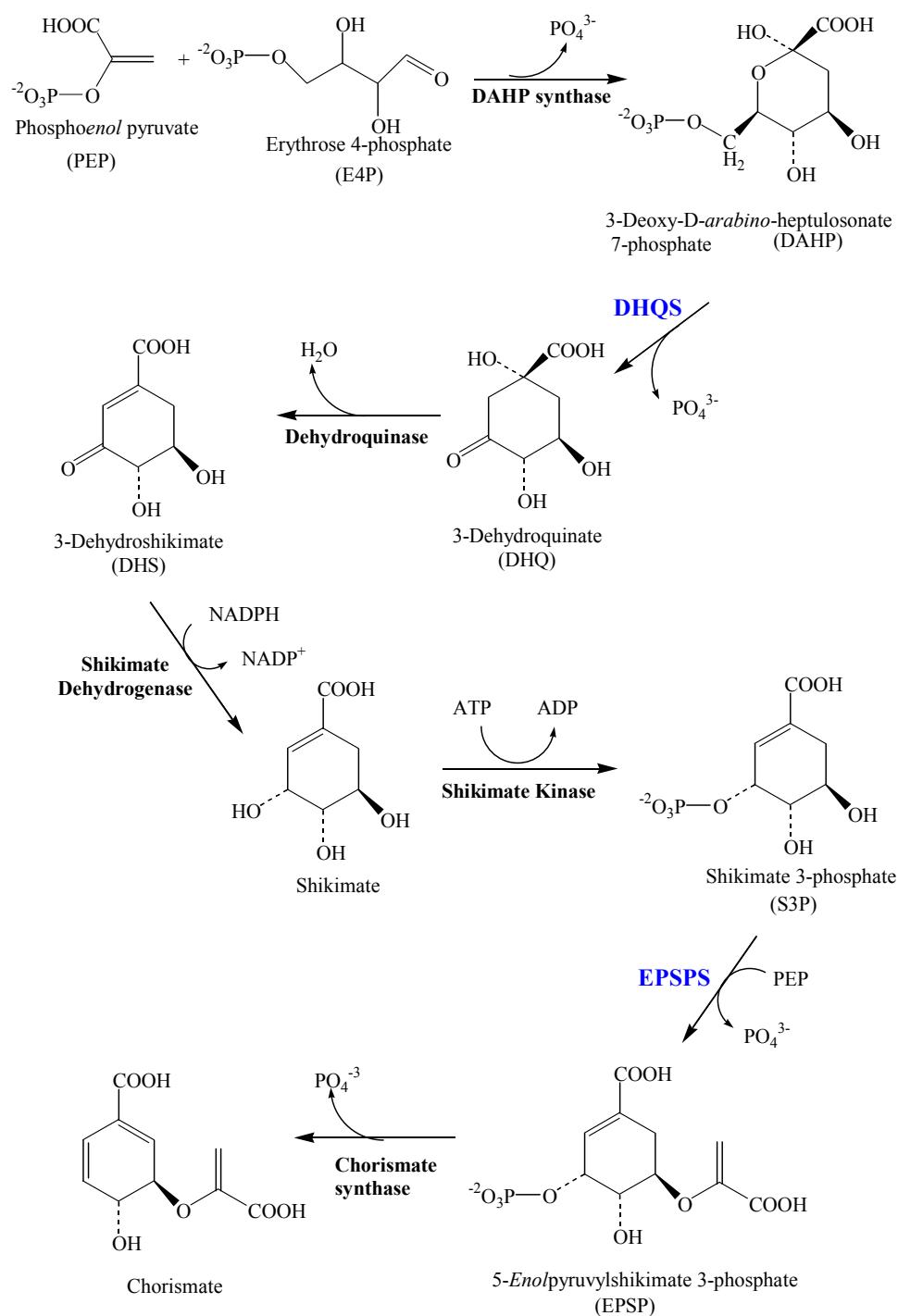
Figure 3.1 The shikimate pathway.

134

Two enzymes in the shikimate pathway, 3-dehydroquinate synthase (DHQS)[6, 7] and 5-*enol*pyruvylshikimate 3-phosphate synthase (EPSPS),[8-10] have been crystallographically observed to undergo conformational changes during catalysis. In the presence of substrate and substrate-analogues, the enzymes exist in a tightly packed structure called the "closed" form, whereas in the absence of substrate or substrate-analogue, the enzyme exists in an "open" conformation. The availability of three-dimensional structural information for both the open and closed conformations of the enzymes presents the unique opportunity to investigate the cause of the conformational changes that occur. A computational software program, GRID, will be employed to investigate the binding recognition sites in the open structure, as well as the binding site of potential substrate and inhibitor molecules in the closed structure of each enzyme.

3-Deoxy-D-Arabino-Heptulosonate 7-Phosphate Synthase

Condensation of PEP and E4P to yield 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAHP) and inorganic phosphate is the first step of the shikimate pathway (Figure 3.1).[3] The synthesis of DAHP is catalyzed by DAHP synthase, which is a metalloprotein inhibited by chelating agents.[11] Several divalent cations including Cu^{2+}, Fe^{2+}, and Zn^{2+} can satisfy the metal requirements of the enzyme.[12, 13]

135

3-Dehydroquinate Synthase

The second reaction of the shikimate pathway is the elimination of phosphate from DAHP to yield 3-dehydroquinate (DHQ) (Figure 3.1).[3] The elimination reaction is catalyzed by DHQ synthase (DHQS), an enzyme requiring the presence of a divalent cation for activity.[14] Catalytic amounts of the cation nicotinamide adenine dinucleotide (NAD^+) are required for DHQS to be active.[15] The reaction involves an oxidation, a β-elimination of phosphate, a reduction, a ring opening, and an intramolecular aldol condensation.[16-18]

3-Dedydroquinate Dehydratase-Shikimate Dehydrogenase

Dehydration of DHQ to give 3-dehydroshikimate (DHS) is the third step of the shikimate pathway (Figure 3.1).[3] The dehydration is catalyzed by DHQ dehydratase and the reaction proceeds by way of a Schiff base mechanism where a lysine residue may serve as an amino donor.[19]

Reduction of DHS to Shikimate

The forth step in the shikimate pathway, the reduction of DHS to shikimate, is catalyzed by a NADP-dependent shikimate dehyrdrogenase (Figure 3.1).[20] Pryrole quinoline is typically used as a cofactor.[21]

Shikimate Kinase

In the fifth step of the shikimate pathway shikimate is phosphorylated to yield shikimate 3-phosphate (S3P) (Figure 3.1). The reaction is catalyzed by shikimate kinase.[3]

5-Enolpyruvylshikimate 3-Phosphate Synthase

The sixth step in the shikimate pathway is the condensation of PEP with shikimate 3-phosphate to produce 5-*enol*pyruvylshikimate 3-phosphate (EPSP) and phosphate (Figure 3.1). The reaction is catalyzed by EPSP synthase (EPSPS).[3] The reaction can be inhibited by glyphosate which does not bind to the apoenzyme, but rather to the enzyme-shikimate 3-phosphate complex.[4] Glyphosate binding is competitive with respect to PEP;[22] however, site-directed mutagenesis studies have indicated that these two molecules do not bind in identical sites on the enzyme.[23-26]

Chorismate Synthase

The final step of the shikimate pathway is the *trans*-1,4 elimination of phosphate from EPSP to yield chorismate, which introduces a second double bond to the ring (Figure 3.1).[27, 28] The reaction is catalyzed by chorismate synthase and requires reduced flavin for activity despite being redox neutral. This behavior makes it similar to DHQS, the second enzyme in the shikimate pathway.

Carbon—oxygen bond cleavage may be initiated by flavin acting as an electron donor to EPSP.

3-Dehydroquinate Synthase Structure and Mechanism

The three-dimensional X-ray crystal structures of *Aspergillus nidulans* dehydroquinate synthase in the open[7, 29] conformation has been determined (Figure 3.2a). DHQS was observed to be a homodimer arranged in a "head-to-tail" orientation. Each monomer consists of an N-terminal *α/β* domain and C-terminal *α*-helical domain. The N-terminal domain is the NAD^+-binding domain and has the characteristic Rossman fold of NAD^+-binding proteins. The NAD^+ is bound in an inverted orientation with respect to other classic Rossman-fold proteins. This orientation of NAD^+ is rare, but not unique to DHQS.[30] The two domains are joined only at the interface running along the bottom of the cleft parallel with the NAD^+-binding site. The majority of contacts between the two monomers occur at the N-terminal domain. The C-terminal domain is primarily composed of an anti-parallel α-helical barrel. The structure has a wide solvent channel located at the junction of the monomers.[7, 29]

The structure of DHQS in the closed conformation was solved containing Zn^{2+}, the co-factor NAD^+ and the slowly-reversible inhibitor carbaphosphonate (CRB).[6] In the presence of carbaphosphonate, DHQS forms a tightly packed structure called the "closed" form,[6, 7, 29] whereas in the absence of substrate or

substrate-analogues the enzyme exists in an "open" conformation.[7, 29] The closing of the structure involves a 14° rotation about an axis, which brings the C-terminal of DHQS into contact with the N-terminal of the other monomer of the dimer (Figure 3.2b). The active site of the enzyme is located in the cleft between the domains. The C-terminal domain contains most of the catalytic residues and the Zn^{2+}- and substrate-binding residues. The conformational change from the open to the closed form of DHQS places residue R130 in the active site region of the closed structure where it plays an integral role in catalysis.[6]
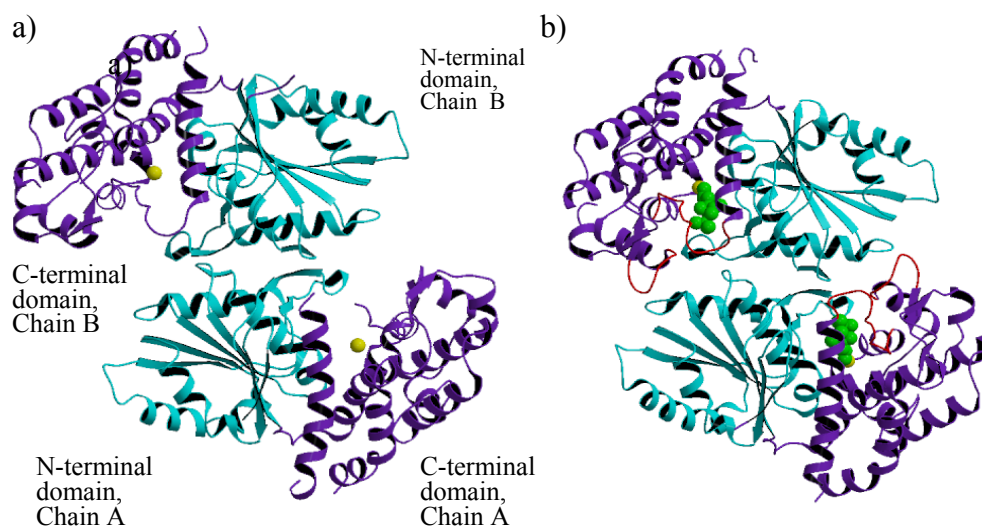
Figure 3.2   a) DHQS "open" structure with Zn^{2+} (yellow) and b) DHQS "closed" structure with Zn^{2+} (yellow) and carbaphosphonate (green). The flexible loops that are important for formation of the active site are shown in red. NAD^+ has been omitted for clarity. The N- and C-terminal domains are indicated in cyan and blueviolet, respectively. Figures were created using MOLSCRIPT[31] and rendered using Render 3D.[32]

Structural studies by Carpenter *et al.*[6] have allowed for elucidation of the residues involved in binding the Zn^{2+} and the substrate-analogue, carbaphosphonate. Analysis of the structure indicated a phosphate binding pockets involving residues N162, K356, R130(B), K152, Wat33, and N268 for the carbaphosphonate phosphonate group. The carboxylate moiety of the carbaphosphonate molecule interacts with residues K250, R264, and K152. Two histidine residues (H271 and H287) and E194 form interactions with the Zn^{2+} (Figure 3.3).

Several hypotheses of the mechanism for the conversion of DAHP to DHQ by DHQS have been proposed. Two early examples of mechanistic work were done by Bartlett *et al.*[18] and Knowles *et al.*[16, 17] Structural information acquired by Carpenter *et al.*[6] has contributed to the understanding of the individual residues participating in the mechanism (Figure 3.3). The substrate is first oxidized at C5 by NAD^+, which involves hydride transfer from C5 of DAHP to C4 of the nicotinamide moiety of NAD^+.[16] Simultaneously, the C5 hydroxyl group of DAHP loses a proton[16] which is relayed to Wat7.[6] An adjacent histidine (H275) could accept the proton.[6] The second step of the reaction is the β-elimination of the phosphate group from intermediate A. The structure indicates the presence of a phosphate binding pocket (K152, N162, N268, H275, and K356) which interacts with the phosphate positioning it so that an oxygen can

remove a proton from C6 of A.[6] The third step of the reaction is the reduction of the C5 ketone of intermediate B by NADH.[16] The hydride ion from C4 of NADH is transferred back to C5 of intermediate B. A ring-opening involving the deprotonation of the hydroxyl at C2 of intermediate C is the fourth step of the reaction mechanism.[17] The structure indicates that the side chain of R260 could act as the base that accepts a proton from Wat20.[6] Rotation about the C5—C6 bond is followed by an intramolecular aldol condensation to reform the ring.[17] No epimerization occurs about the C2 bond in the enzyme-mediated reaction.[18] The absence of epimerization is critical for the continuation of the shikimate pathway as the product, DHQ, must correctly interact with dehydroquinase in the following step in order to produce DHS. The absence of epimerization could be due to hydrogen bond interactions between the carboxylic acid group on C2 of intermediate D and the side chains of residues K152, K250, and R264, as well as by the interactions between the hydroxyl or ketone functional groups on C2 of intermediate D with N268 and Wat20.[6] By binding the intermediates, the protein prevents the formation of unwanted side products.
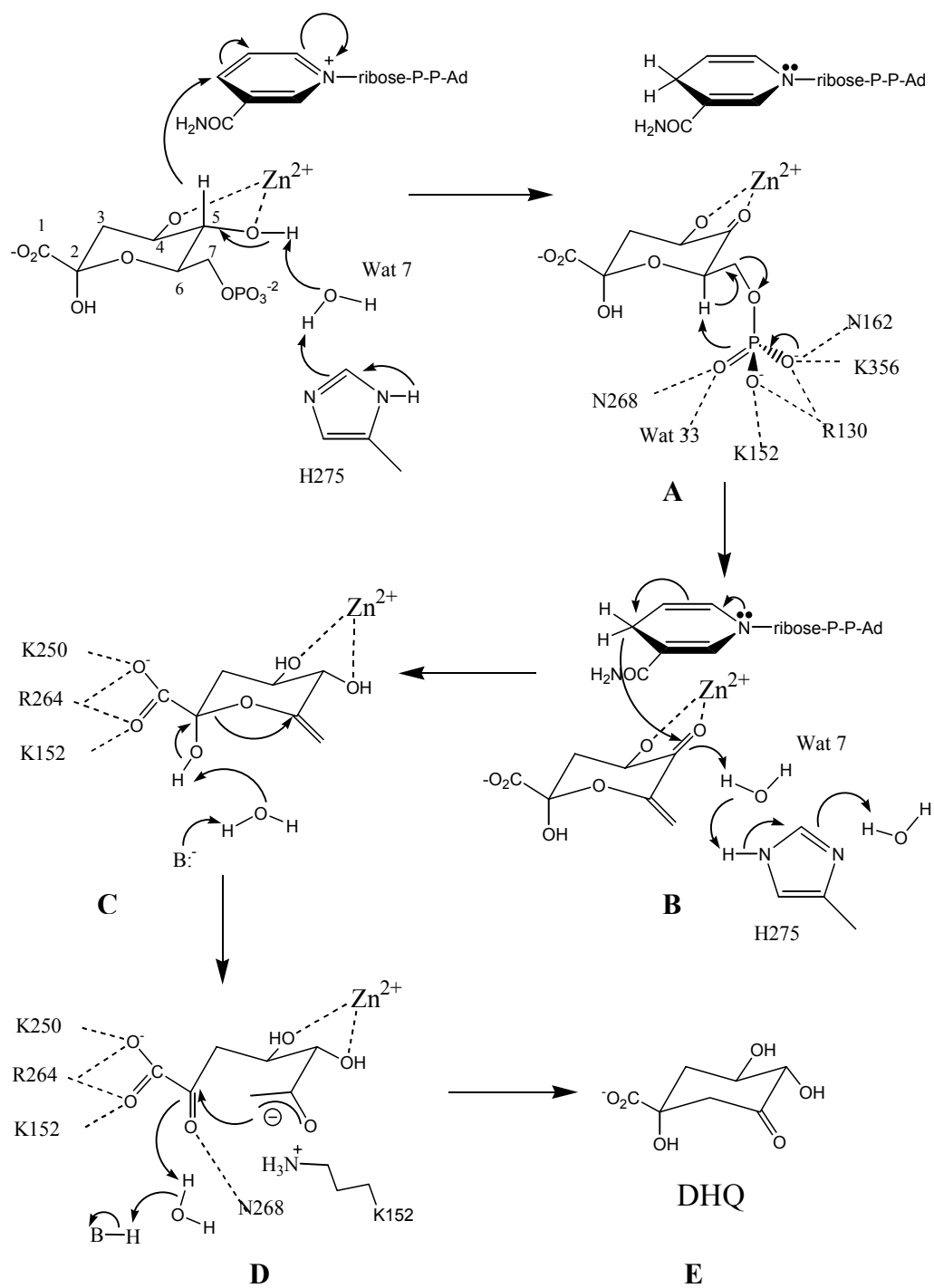
Figure 3.3 DHQS mechanism including the residue interactions as proposed by Carpenter *et al*. based on the acquired structural information.

5-Enolpyruvylshikimate 3-Phophate Synthase Structure, Mechanism, and

Inhibition

Stallings *et al.*[33] initially determined the three-dimensional X-ray crystal structure of *E. coli* EPSPS in 1991; however, since then two higher resolution structures have been determined.[8, 9] The structure consists of two globular domains connected by a double-stranded hinge (Figure 3.4a). The two-domain structure appears to be formed by 6-fold replication of a protein folding unit comprising two parallel helices and a four-stranded β-sheets. Three of these subunits comprise each domain with approximate three-fold symmetry. The two domains of EPSPS are nearly hemispheric and equal in size; however, the amino and carboxyl terminus of the polypeptide backbone are located only in the lower domain. The N-termini point toward the active site, which has been suggested to favor the binding of the negatively charged substrates and inhibitors.[33]

A conformational change has been observed for EPSPS during the catalysis.[8, 10] In the presence of shikimate, EPSPS forms a tightly packed structure called the "closed" form whereas in the absence of substrate it exhibits an "open" conformation (Figure 3.4). The closing of the structure involves a 64° rotation about an axis close to the active site. Several important binding residues including S197 and K411 are rearranged into the active site by the conformational change.[29] Structural studies by Schönbrunn *et al.*[8] and Carpenter *et al.*[9] have

indicated two phosphate binding pockets exist in the closed EPSPS enzyme. Residues S170, K340, S169, N336, and S197 form the shikimate 3-phosphate phosphate moiety binding pocket, while the phosphonate moiety of the glyphosate interacts with residues K22, Q171, R124, G96, N94, and K411 (Figure 3.5). The carboxylate region of the shikimate molecule interacts with residues R27 and S23, while the nearest neighbor residues to the glyphosate's carboxylate moiety are R386, R344, D313, and K22. In addition to possessing phosphate and carboxylate groups, which can be involved in hydrogen bonding, Lennard-Jones and electrostatic interactions, the shikimate 3-phosphate molecule has three hydroxyl groups which interact with residues D313, K22, and K340.[29]

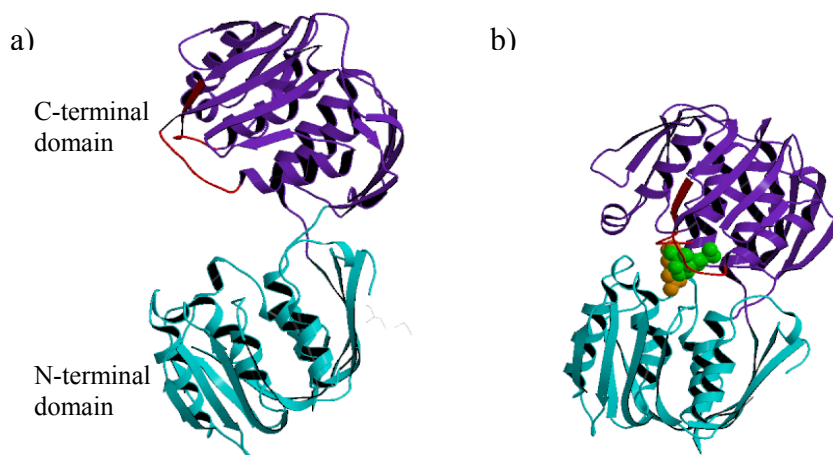**Figure 3.4 a) EPSPS "open" structure and b) EPSPS "closed" structure with glyphosate (orange) and shikimate-3-phosphate (green). The flexible loops that are important for formation of the active site are shown in red. The N- and C-terminal are indicated in cyan and blueviolet, respectively. Figure created using MOLSCRIPT[31] and rendered using Render 3D.[32]**

N-phosphomethylglycine (glyphosate) has been found to act as an inhibitor in EPSPS. Plants do not appear to have the ability to degrade or inactivate glyphosate, which then accumulates in the meristematic areas of the plant producing phytotoxic effects.[4] Glyphosate acts as a competitive inhibitor with respect to PEP and a non-competitive inhibitor with respect to shikimate 3-phosphate in the EPSPS reaction.[34] PEP and glyphosate act as competitive inhibitors; however, their binding sites do not appear to be totally equivalent. Although PEP and glyphosate have similar functional groups, their structures are not identical, which creates variations in their binding site preferences. This difference in binding sites has been demonstrated *via* several site-directed mutagenesis studies.[23-26] A key study by Padgette *et al.*[25] has indicated that EPSPS G96A selectively destabilizes the interaction of enzyme with glyphosate compared to PEP. The difference between interaction of the enzyme with the inhibitor and substrate after single-site mutagenesis suggests that their binding sites are not completely equivalent.

Figure 3.5 The active site of EPSPS indicating the important residue interactions with shikimate 3-phosphate and glyphosate. Residues located in The N- and C-terminal domains are shaded in blue and peach, respectively.

In the reaction catalyzed by EPSPS the intact *enol*pyruvyl group of the PEP is transferred to the 5-hydroxyl group of shikimate 3-phosphate producing EPSP and inorganic phosphate.[3] The reaction is reversible and occurs without the presence of metals or cofactors.[35] There are multiple mechanisms proposed for the reaction. Bondinell, *et al.*[36] suggested the reaction proceeds exclusively with C—O cleavage at the PEP phosphate group, followed by transfer of the intact *enol*pyruvyl group to the acceptor molecule (Figure 3.6). This makes the

reaction unusual as most other enzyme-catalyzed processes utilizing PEP proceed through P—O cleavage. The suggested reversible addition-elimination mechanism proposes the protonation of C3 of PEP is associated with nucleophilic attack on C2 by the 5-hydroxyl group of shikimate 3-phosphate resulting in the formation of the tetrahedral intermediate (Figure 3.6). EPSP is generated following the elimination of phosphate.[36]

Anton *et al.*[37] proposed an alternative to the Bondinell mechanism which involves the stabilization of intermediate I (Figure 3.7) when a negatively charged active site residue of EPSPS forms a covalent bond to PEP. The 5-hydroxyl of shikimate 3-phosphate could then attack C2 of PEP to form the tetrahedral intermediate (Figure 3.7). Conversely, phosphate could be eliminated to form a covalent *enol*pyruvyl enzyme adduct which would react with shikimate 3-phosphate to form EPSP.
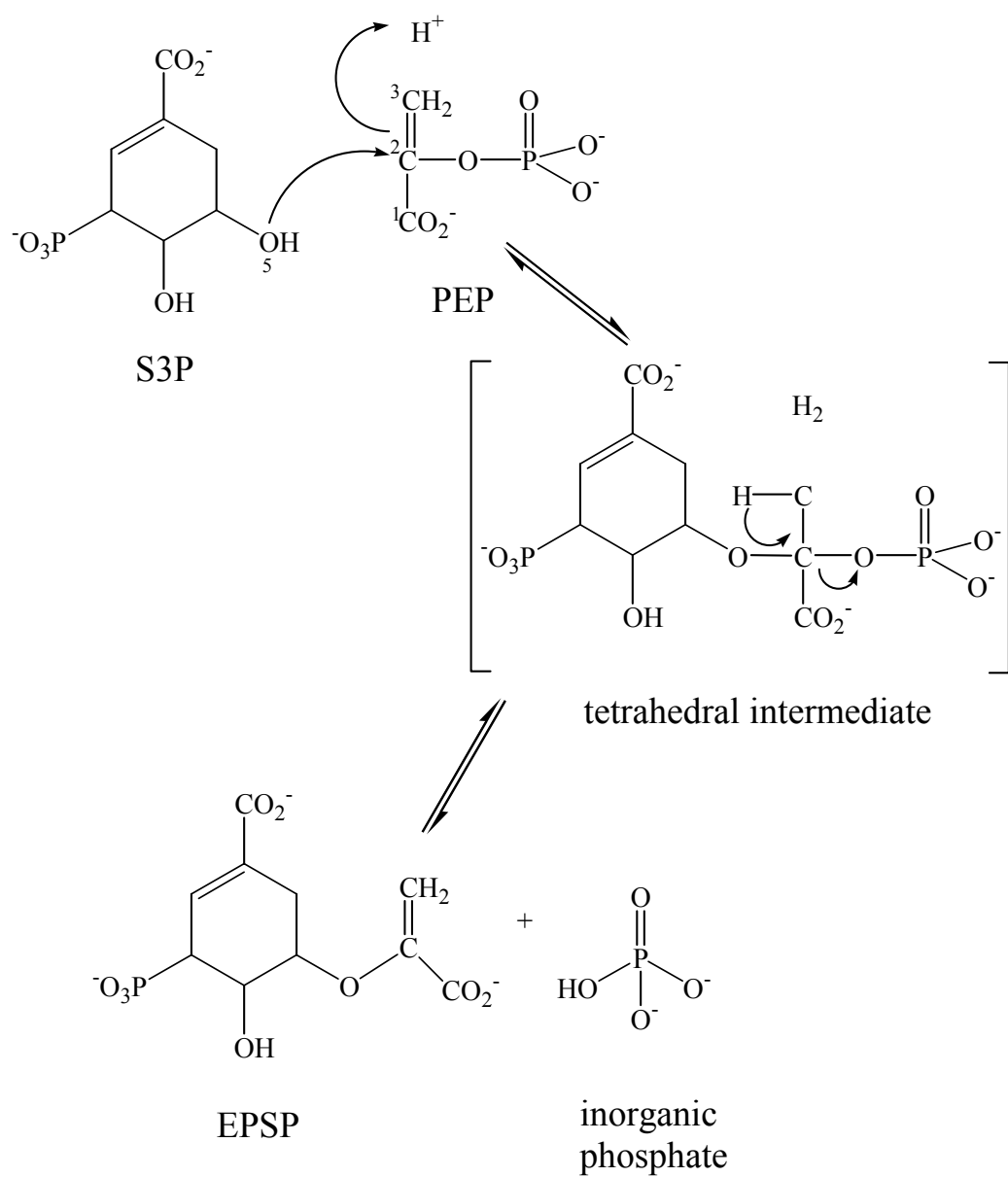
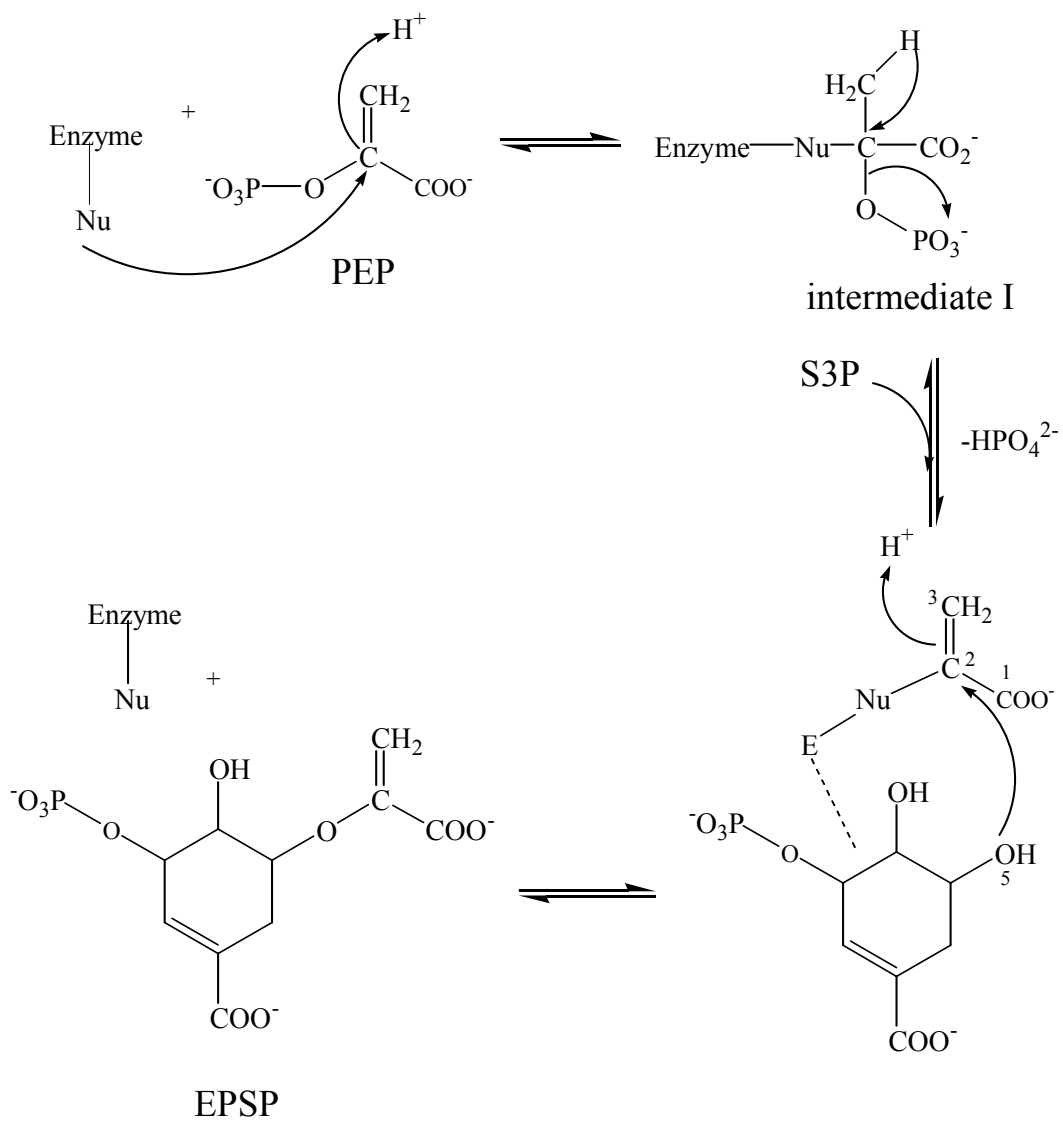**Figure 3.6 Mechanism proposed by Bondinell *et al.* for the EPSP synthase reaction.**

Figure 3.7 Mechanism postulated by Anton *et al.* in 1983 for EPSPS.

149

GRID CALCULATIONS

GRID, a computational program by P. J. Goodford,[38] was designed for detecting energetically favorable binding sites on proteins which have been structurally determined by X-ray crystallography or nuclear magnetic resonance. The target molecule is the species being evaluated. The probe molecule is a small biologically active species such as water, amine nitrogen, carboxy oxygen, or hydroxyl. GRID calculates the interactions between a probe molecule and a target molecule accounting for both the shape and energy of the two species. The Lennard-Jones potential is a function which displays energy minima corresponding to preferred nonbonded atoms by combining the attractive and repulsive forces between atoms (Equation 3.1).[39, 40]

$$E_{\text{lj}} = \frac{A}{d^{12}} - \frac{B}{d^{6}}$$

Equation 3.1 Lennard-Jones energy equation.

In Equation 3.1, d is the distance between a pair of non-bonded atoms and A and B are parameters, which will be discussed later. When d is small, the A/d^{12} term predominates corresponding to a large positive value of E indicating repulsion. The value of d defines the minimum separation between atoms, giving them a radius. However, when d is large the E_{lj} corresponds to the minimum value or bottom of the trough on the Lennard-Jones curve (Figure 3.8). Both

maximum and minimum values can be extended to macromolecular systems and

displayed by computer graphics.
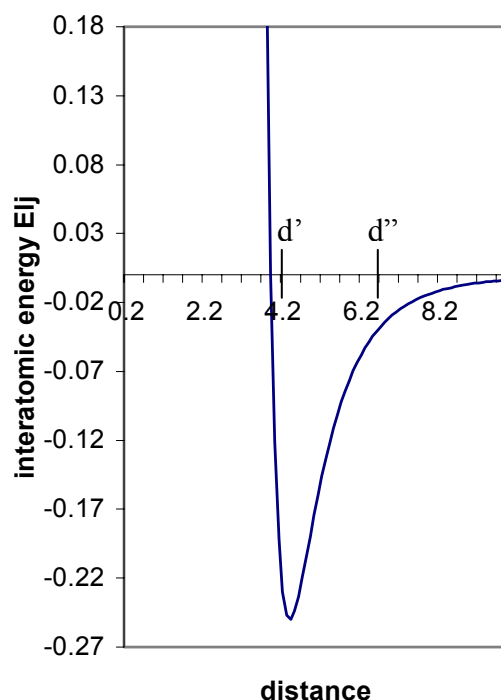
Figure 3.8   A Lennard-Jones curve for the interaction between two nonbonded atoms.   Distances d' and d" define a zone of interatomic attraction.

At distances *d'* and *d"* a zone of mutual attraction exists (Figure 3.8).

GRID uses the *d'* to *d"* zone, as well as electrostatic and hydrogen bond effects to

calculate interactions of a probe group with the target molecule.   An array of

energy values is computed at sample positions of the probe throughout the target

allowing contour plots to be drawn.

The atomic coordinates for the target molecules are acquired from the Protein Data Bank (PDB)[41] and checked for missing atoms. Ligands such as water, chloride ions, and cofactors are removed, unless they are structurally significant. GRUB, a table of parameters included in the GRID software package, is needed to evaluate the Lennard-Jones and other empirical energy functions. GRIN adds the parameters from GRUB to the atomic coordinates file preparing the file for use by GRID. The parameters in GRUB are based on the "extended" atom concept used for the program CHARMM.[42, 43] In the "extended" atom concept, a common functional group is represented as a single-entity rather than as multiple atoms. Significant computing time is saved as the single "extended" group replaces multiple atoms. For example, the amino group is treated as a single entity with a van der Waals radius of 1.75 Å. This value is somewhat larger than that of a nitrogen atom.

Once GRIN has successfully modified the atomic coordinates file, GRID is employed. An array of points is established throughout the target molecule creating the GRID. The potential energy, E_{xyz}, of the probe is calculated at the first point on the first *XY* plane of the GRID. The analysis proceeds until an energy value has been assigned for every point.

Favorable interactions between the probe and the target can be determined by the presence of modest negative energy values. The $-B/d^6$ term of the

Lennard-Jones function, electrostatic effects, and hydrogen-bond interactions all contribute to the slightly negative energy value.

The nonbonded interaction energy, E_{xyz}, of the probe at each location on the GRID is comprised of three components (Equation 3.2). One pairwise interaction between the probe at the GRID point *xyz* and an atom of the target composes each term in the summation. The summation extends over all target atoms.

$$E_{xyz} = \sum E_{lj} + \sum E_{el} + \sum E_{hb}$$

Equation 3.2   The energy equation for nonbonded interactions between atoms where E_{lj}, E_{el}, and E_{hb} is the energy from the Lennard-Jones interactions, electrostatic effects and hydrogen-bond interactions, respectively.

To save computing time, E_{lj} is set to zero if the probe and the pairwise target are more than 8 Å apart. At this distance, E_{lj} is typically −0.01 kcal/mol^{-1} and is diminishing according to the sixth power of the interatomic distance. This assumption does introduce an error on the order of 0.1 kcal/mol^{-1} in ΣE_{lj}. The values of *A* and *B* in Equation 3.1 are calculated from the effective number of electrons, N_{eff}, the polarizability, α_1, and the van der Waals radius, *r*, of the interacting atoms, which was previously reported by Hopfinger.[44]

The electrostatic interaction, E_{el}, does not diminish rapidly with distance. Therefore, it was necessary to evaluate the individual terms pairwise between the

probe group and every target atom. Equation 3.3 is a compromise between procedures based on the algorithm of Warwicker and Watson,[45] which requires extensive computing times and the classical functions which inadequately describe the dielectric effect arising from shielding and coupling effects of the solvent and solute molecules.[44]

$$E_{el} = \frac{pq}{K\zeta} \left[\frac{1}{d} + \frac{(\zeta - \varepsilon)/(\zeta + \varepsilon)}{\sqrt{d^2 + 4s_p s_q}} \right]$$

Equation 3.3   The energy equation for electrostatic interactions between atoms.

The variables p and q are the electrostatic charges on the probe group and the pairwise target atom, respectively. Distance between the probe and the target is d and K is equal to $4\pi\varepsilon_o$, where $\varepsilon_o = 8.854$ F m^{-1}.[44] If a probe has less than seven protein neighbor atoms within a 4 Å distance then it is assumed to be in the solution phase and the $4s_p s_q$ term is set to zero to shorten computational time. The dielectric constants ε and ζ were set to 80 and 4, respectively. Equation 3.3 yields an effective dielectric term if the pairwise groups are embedded in the protein and close enough together that the effect of the solution may be neglected.

Hydrogen bonding effects are determined by a direction-dependent 6-4 function (Equation 3.4), which more accurately predicts the hydrogen-bond energy minimum than the commonly used 12-10 function.[43] Experimentally, the

154

hydrogen-bond energy minimum has been determined to be considerably broader

than the value assigned by the 12-10 function.

$$E_{hb} = \left[\frac{C}{d^6} - \frac{D}{d^4} \right] \cos^m \theta$$

Equation 3.4 The energy equation for hydrogen bond interactions between atoms.

When two identical atoms are interacting, the tabulated values for C and

D, which are specific constants for each donor-acceptor pair, determine their

interatomic separation, d_{min}. If the atoms are of different types, the geometric

mean of their individual D values and the arithmetic mean of their d_{min} separations

are used. The value of C is calculated by setting the differentiated form of

Equation 3.4 to zero (Equation 3.5).

$$C = 0.666 D d_{min}^{2}$$

Equation 3.5 The equation for calculating the value of C in Equation 3.4.

Two possibilities exist for hydrogen bonding between the probe and target

molecules. If the target donates a hydrogen bond, then the bond direction is

determined by the hydrogen position. The angle DHP (Figure 3.9) where D is the

target donor atom, H is the hydrogen, and P is the probe accepting the hydrogen

bond is represented by θ (Equation 3.4). The m term in Equation 3.4 is typically

4, but E_{hb} is zero when $\theta \leq 90°$. If the probe donates the hydrogen bond then it is

155

assumed that it can orient itself in order to form the most effective hydrogen-bond

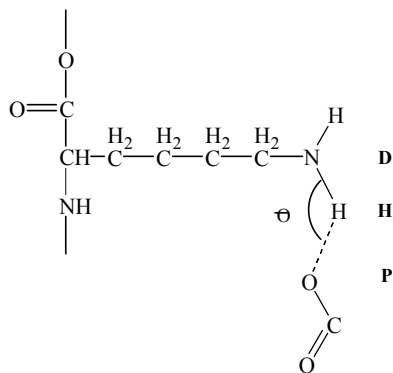interaction with the target atom, and cos θ is set to unity (Equation 3.4).

Figure 3.9   Angle θ is the angle between the protein donor atom, D, the hydrogen atom, H, and the probe accepting the hydrogen bond, P.

Utilizing the ability of GRID to predict energetically favorable binding sites for small molecules, experiments were undertaken to investigate the binding potentials for various probes in DHQS and EPSPS, two enzymes of the shikimate pathway.   As previously discussed, both DHQS and EPSPS undergo conformational changes upon binding of the substrate molecules.   GRID was employed to investigate the differences in the small molecule binding sites of the "open", substrate free, and "closed", substrate bound, forms of each enzyme. Atomic coordinates for DHQS closed (1DQS)[6], and EPSPS closed (1G6S)[8] were obtained from the PDB,[41] while coordinates for the DHQS and EPSPS open structures were obtained from a personal communication.[46] The structures prior to performing GRID calculations are shown (Figure 3.10).
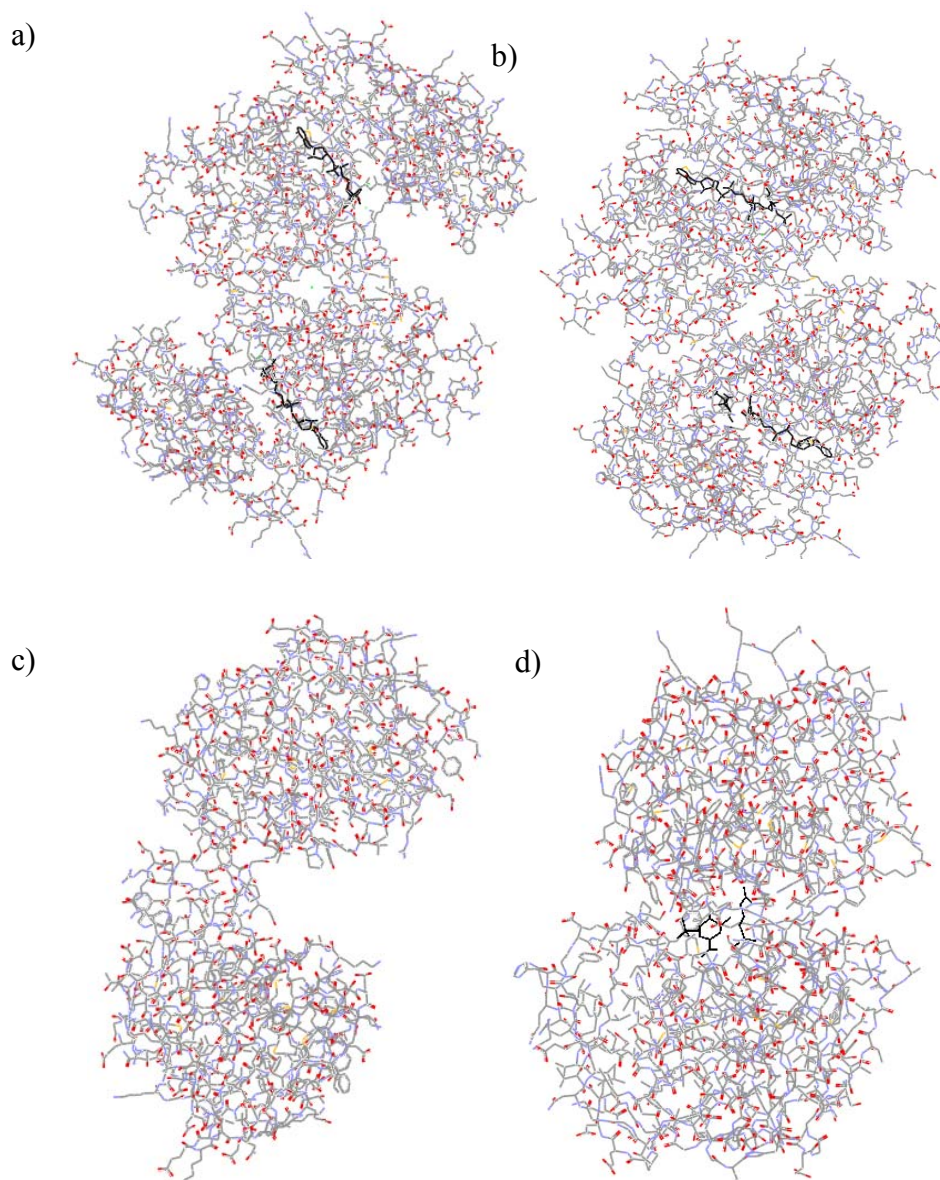
Figure 3.10   a) DHQS "open" with NAD$^+$ and Zn^{2+} all shown in black, b) DHQS "closed" structure with NAD$^+$, Zn^{2+}, and carbaphosphonate all shown in black, c) EPSPS "open" and d) EPSPS "closed" structure with shikimate 3-phosphate, and glyphosate both indicated in black. All structures appear prior to GRID map overlays.

Two experiments were performed using GRID on each protein structure. The first experiment investigated the small molecule binding potential of the protein structure containing only co-factors, where applicable. All substrate-analogues, water molecules, and metal ions were removed from the structure file prior to the initial calculations. GRID was then used to evaluate the energetically favorable binding sites of counterions and appropriate small molecules. The second experiment was performed on the protein structure file containing the co-factors, where applicable, as well as any substrate-analogues and metal ions that were present in the crystal structure. Again, the structure was evaluated with appropriate small molecule probes to locate the most energetically favorable binding sites.

GRID RESULTS

GRID calculations were performed on both the open and closed forms of DHQS. Zn^{2+} and carbaphosphonate (in the closed form) were removed from the structure files and initial calculations were preformed on the NAD^+-containing enzyme only. The metal cation and the substrate-analogue were added back into the coordinate files and the enzymes were reanalyzed using GRID. Unless otherwise stated, all calculations are assumed to be on the enzyme bound only to the cofactor NAD^+.

Several probes were used to investigate potential binding sites on the enzyme. The probes are designed to locate energetically favorable binding sites on the enzyme for the functional moiety they represent. Based on the functional groups of the substrate, DAHP, and a substrate-analogue, carbaphosphonate, probes were chosen (Figure 3.11). The hydroxyl probe was chosen to represent the three OH moieties present on both the DAHP and carbaphosphonate molecules. The carboxy, carbonyl, and multi-atom carboxy probes were selected due to the presence of a carboxy functional group on both DAHP and carbaphosphonate (Figure 3.11). The phosphate or phosphonate moeity on DAHP and carbaphosphonate, respectively, is a primary component of both the substrate and substrate-analogue (Figure 3.11). The hydrogen phosphate monoanion and phosphate dianion probes were chosen to search for binding sites in the enzyme for HPO_4^- and PO_4^{2-}, respectively.

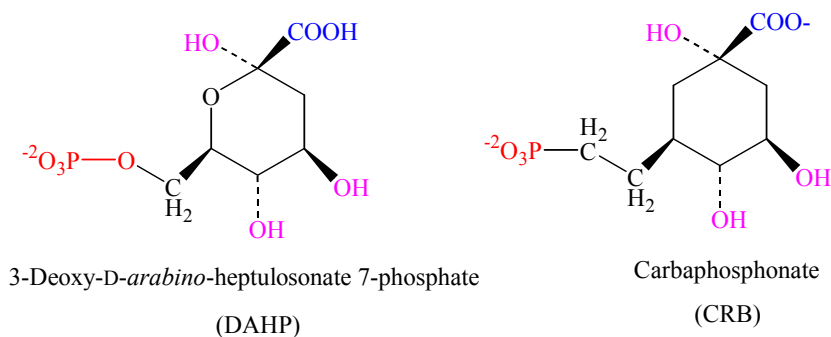3-Deoxy-D-*arabino*-heptulosonate 7-phosphate

(DAHP)

Carbaphosphonate

(CRB)

Figure 3.11 The hydroxyl, carboxy, carbonyl, multi-atom carboxy, hydrogen phosphate monoanion, and phosphate dianion probes will be used to locate the binding pocket for the color coordinated functional groups on the DHQS substrate and substrate-analogue.

159

GRID calculation output was displayed as contour maps using Insight II (Accelrys, Cambridge, UK) and DS ViewerPro 5.0 (Accelrys, Cambridge, UK). By adjusting the visible contour level, the total energy cut off value was set for each calculation. Although qualitative comparisons can be made across systems, quantitative analysis is not recommended as the energy values are probe and target specific.

The presence of energetically favorable sites was determined by visual inspection of the size and shape of the countour within a probe/enzyme system. The user is required to apply knowledge of the biochemistry of the protein being evaluated to determine whether sites identified as energetically favorable are significant to the posed question.

DHQS Calculations

<u>Hydroxyl Probe</u>

The open DHQS structure was explored with the hydroxyl probe to locate regions of the protein that favor interaction with an OH moiety. A number of favorable interactions were located throughout the protein with similar energetic favorability. None of the locations were significantly more favorable when compared to other interactions and, therefore, did not merit being labeled a hydroxyl binding pocket.

The hydroxyl probe was used to investigate the DHQS closed structure for a potential area of interaction for an alkyl hydroxyl group on the enzyme. Energetically favorable interactions were located at residues D146, N162, E194, N268, H271, and H275, corresponding to the binding location of the hydroxyl moieties of the carbaphosphonate molecule (Figure 3.12).

Figure 3.12 The energetically favorable binidng sites for the hydroxyl probe on the DHQS closed structure. Contour level is set at −10 kcal/mol. The carbaphosphonate molecule is overlayed to aid in visualization.

Carboxy Probe

The open DHQS structure was explored with the carboxy probe to located regions of the protein that favor interaction with an anionic carboxy oxygen atom. Although no energetically favorable binding sites were found near the active site,

the probe showed favorable interaction with residues K89(A) and K89(B) at the dimer interface (Figure 3.13).
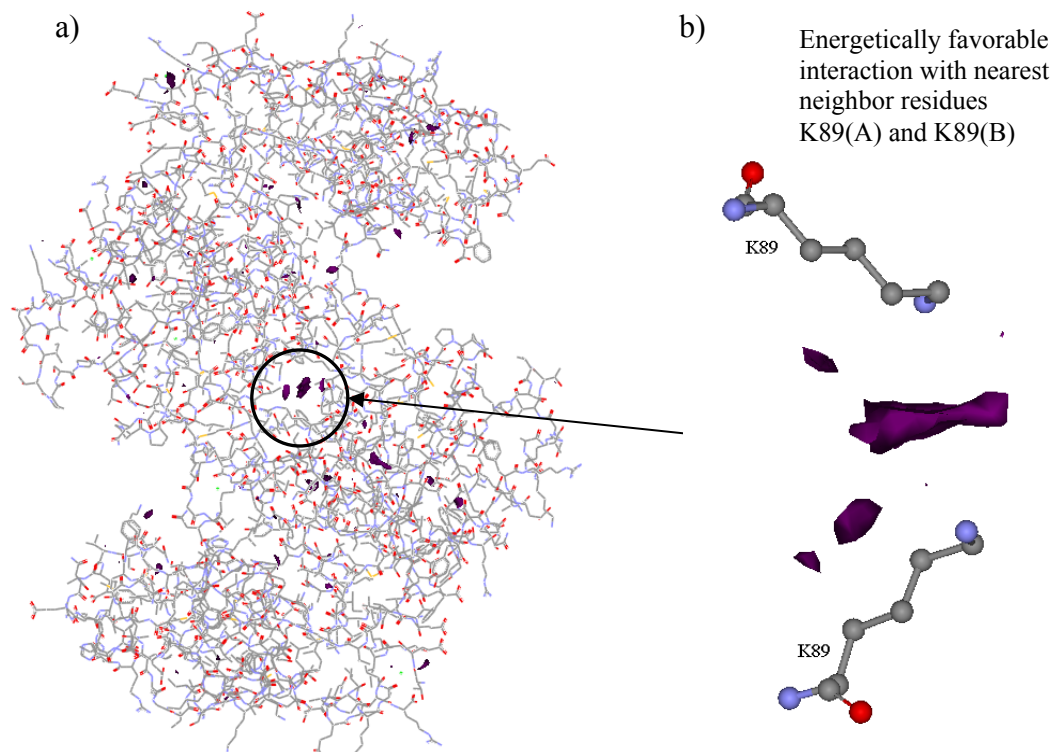
Figure 3.13   a) The energetically favorable binding sites for the carboxy probe on the DHQS open structure and b) the zoomed view of the energetically favorable region.   Contour level set at −10 kcal/mol.

The DHQS closed structure was analyzed with the carboxy probe to locate the most favorable binding site for sp^2 carboxy oxygen atoms. The carboxy probe located an energetically favorable binding site near residues K152, K250, and

R264, which corresponds to the residues associated with the carboxy group of the carbaphosphonate (Figure 3.14).

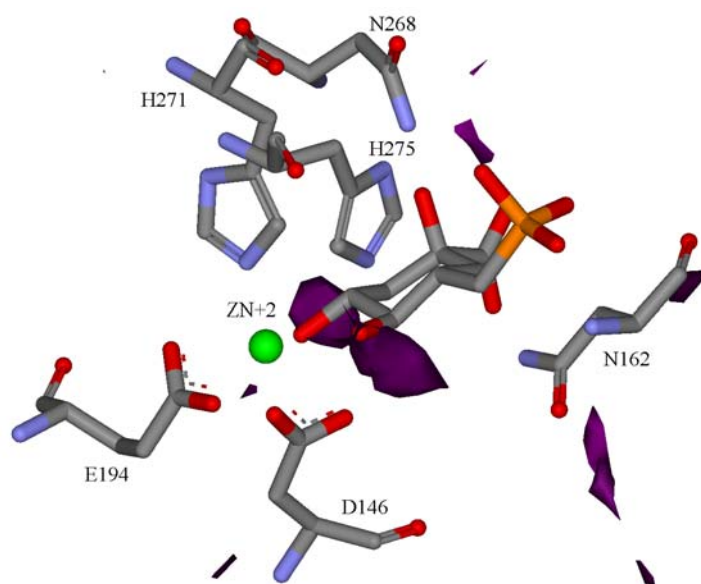**Figure 3.14 Stereoview of the most energetically favorable binding sites for the carboxy probe on the DHQS closed structure. Contour level is set at −10.5 kcal/mol. The carbaphosphonate and Zn^{2+} (green) have been overlayed on the data map to aid in visualization.**

Carbonyl Probe

The carbonyl probe, a small molecule probe designed to locate energetically favorable sites for interaction with a carbonyl oxygen atom, identified residues K152, K250, and R264 as a potential binding pocket in the DHQS open structure. The dimer interface was also determined to be an energetically favorable site for sp^2 carbonyl oxygen binding.

163
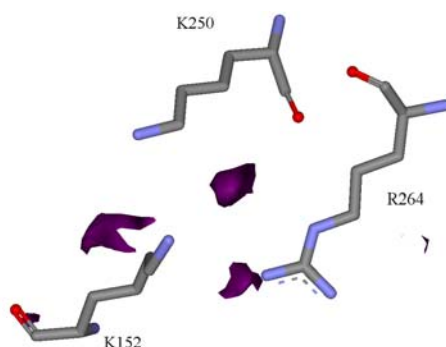
Figure 3. 15 Energetically favorable binding site for the carbonyl probe in the DHQS open structure. Contour is set at −7 kcal/mol.

The same probe was used to investigate the DHQS closed structure. The nearest neighbor residues for the most energetically favorable site were K152, K250, and R264, which are identical to those found by the carboxy probe, (Figure 3.16).

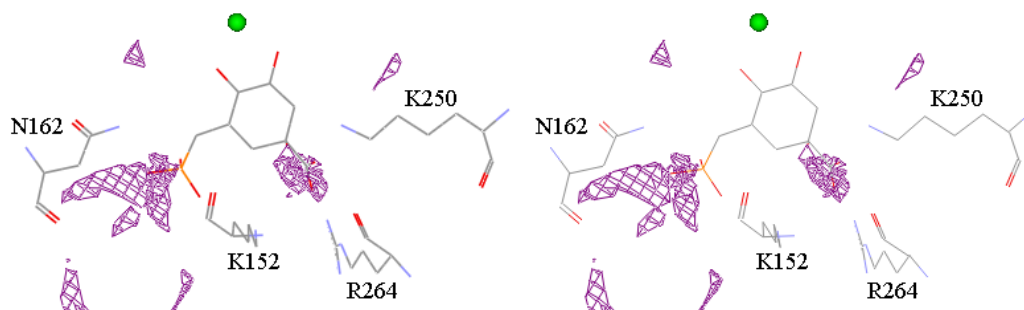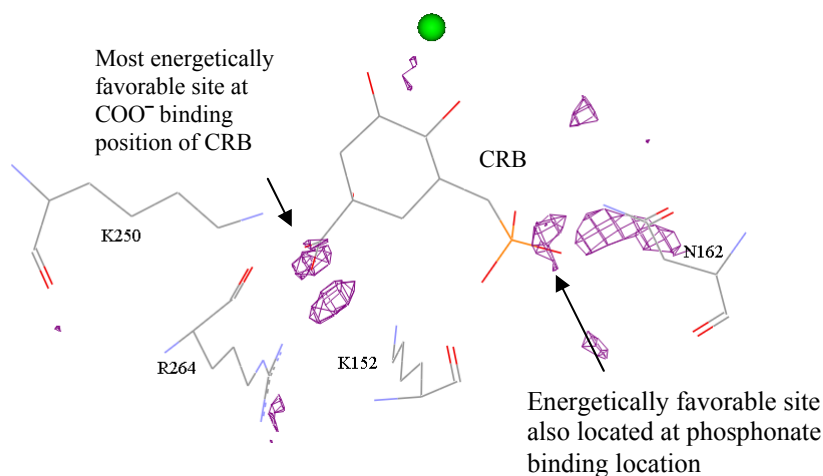**Figure 3.16 The most energetically favorable binding sites for the carbonyl oxygen binding in the DHQS closed structure. Contour level set at −8.5 kcal/mol. The carbaphosphonate (CRB) and Zn^{2+} (green) molecules have been overlayed on the data map to aid in visualization.**

Carboxy Multi-atom Probe

The carboxy multi-atom probe was designed to represent the entire carboxy moiety. The multi-atom probe consists of two sp^2 oxygen atoms bonded to the sp^2 carboxy carbon, which is itself bonded to another β-carbon atom in a methylene group.[38] The probe contains four lone pairs that can accept up to four hydrogen bonds. The multi-atom probe is a more accurate model of a carboxy group than is the single-atom carboxy probe, which represents only an sp^2 oxygen atom bound to a carbon (Figure 3.17).
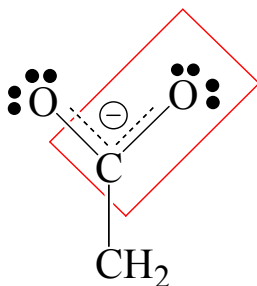
Figure 3.17 The carboxy multi-atom probe is represented with the red box indicating the carboxy single-atom probe.

The DHQS closed structure was analyzed with the carboxy multi-probe to locate the most favorable binding site for a carboxy moiety. The nearest neighbor residues for the first energetically favorable site were determined to be K152, K250, and R264. The size and shape of the contours were found to be nearly identical to those found for the carboxy probe (Figure 3.16).

Phosphate Dianion Probe

Modeling a phosphate dianion with no hydrogens, the phosphate dianion probe was used to investigate the DHQS open structure. The most energetically favorable binding site was identified at the dimer interface with nearest neighbor residues K89(A), K89(B), R86(A), and R86(B) (Figure 3.18). This site is similar to that found to be energetically favorable towards the carbonyl and carboxy probe, but additionally includes residues R86(A) and R86(B). A second binding site for a PO_4^{2-} functional group was determined to have R130(B), N162, and K152 as nearest neighbor residues. These residues correspond to the known binding site of the phosphate unit on the carbaphosphonate in the closed DHQS structure.

a)

b)

R86

K89

R86

K89
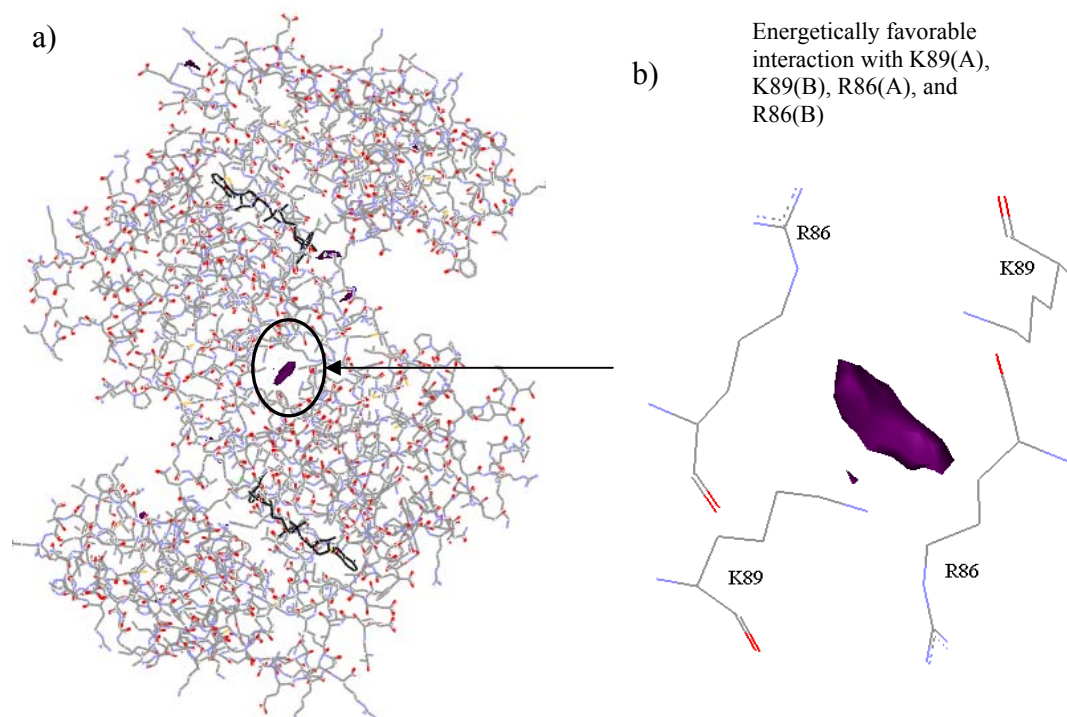
Figure 3.18   a) View of the DHQS open structure indicating the most energetically favorable binding site for the phosphate dianion probe. Zn^{2+} and NAD^+ are indicated in black. b) Detailed view of the energetically favorable site. The contour level is set at -15 kcal/mol in both views.

A closer view of the phosphate binding pocket can be seen in Figure 3.19. Residues N162, K152, and R130(B) interact with the phosphate dianion probe with residue N268 in position to interact upon closure of the structure.

167
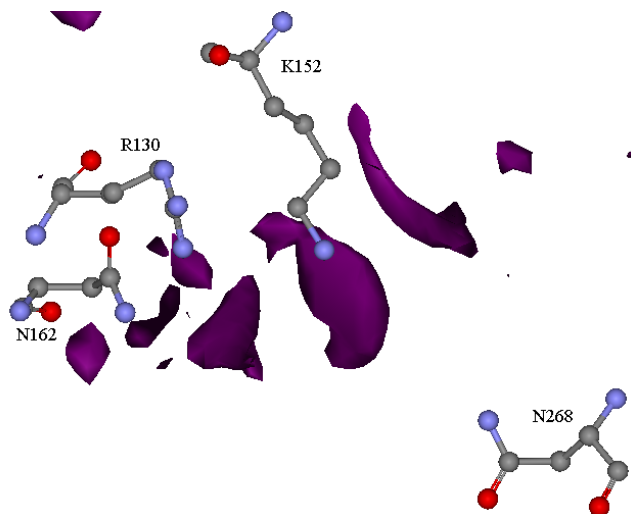
Figure 3.19   View of the phosphate binding pocket in the open DHQS structure. Contour level set at −12 kcal/mol.

The phosphate dianion probe was used to located energetically favorable binding sites in the DHQS closed structure. The most energetically favorable site for PO_4^{2-} binding in the enzyme is located near residues N162, K356, K152, N268, and R130(B).   These residues correspond to the binding site of the phosphonate unit on the carbaphosphonate in the closed DHQS structure.   The second energetically favorable site recognized by the PO_4^{2-} probe is the carboxy binding site of the carbaphosphonate, which has nearest neighbor residues K250, K152, and R264 (Figure 3.20).

168

Energetically favorable site also located at COO$^-$ binding position

K250

R264

K152

CRB

N162

Most energetically favorable site at PO$_4^{2-}$ binding location
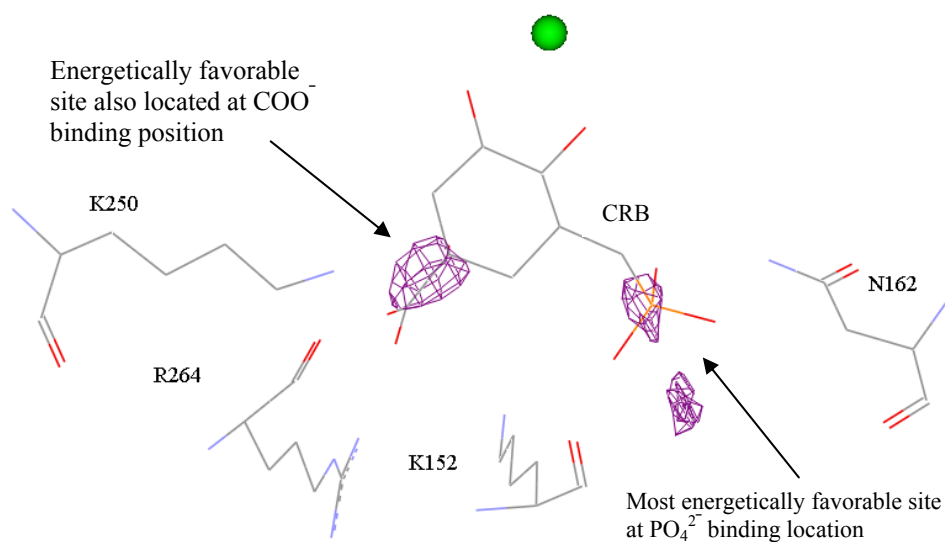
Figure 3.20   The two most energetically favorable binding sites for the phosphate dianion probe in the DHQS closed structure. Contour level set at −17 kcal/mol.   The carbaphosphonate (CRB) and Zn^{2+} (green) have been overlayed on the data map to aid in visualization.

Hydrogen Phosphate Anion Probe

The open DHQS structure was explored with the hydrogen phosphate anion probe to locate favorable sites on the protein for HPO$_4^-$ interaction. Although similar to the phosphate dianion probe, the hydrogen phosphate probe is monoanionic and has one hydrogen atom (Figure 3.21).
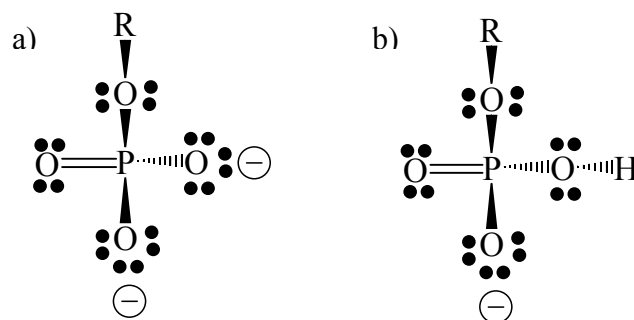
Figure 3.21 The a) phosphate dianion and b) hydrogen phosphate anion probes. The R group indicates the attachement point of the moiety.

Several energetically favorable binding sites were located; however the most favorable site was located at the dimer interface. This location has nearest neighbor residues K89A, K89B, R86A, and R86B. A second binding site was determined to have R130(B), N162, and K152 as nearest neighbor residues. These residues correspond to the known binding site of the phosphonate moiety of the carbaphosphonate in the closed DHQS structure. The size and shape of the energy contours for the HPO_4^- were nearly identical to those found for the phosphate dianion probe (Figure 3.20).

The hydrogen phosphate monoanion probe was employed to determine favorable binding pockets for HPO_4^- in the DHQS closed structure. The most energetically favorable binding site was located at residues K250, R264, and K152, which forms the carboxy binding pocket of the carbaphosphonate molecule. The second site located by the hydrogen phosphate monoanion probe is a surface site. The phosphonate binding pocket of the carbaphosphonate

molecule is identified as the third most energetically favorable binding site with nearest neighbor residues N162, K356, K152, N268, and R130(B).

<u>Zinc Dication Probe</u>

The zinc dication probe was used to analyze the DHQS open structure for potential Zn^{2+}-binding sites. Two energetically favorable binding sites were located on the surface. The first energetically favorable site to occur at the interior of the protein had nearest residue neighbors E194, H287, and H271. The residues involved matched the structurally characterized binding site of Zn^{2+} within a distance of 1.1 Å (Figure 3.22). This can be considered nearly equivalent considering the margin of error for GRID (0.5 Å) and the resolution of the three-dimensional structure (1.8 Å).
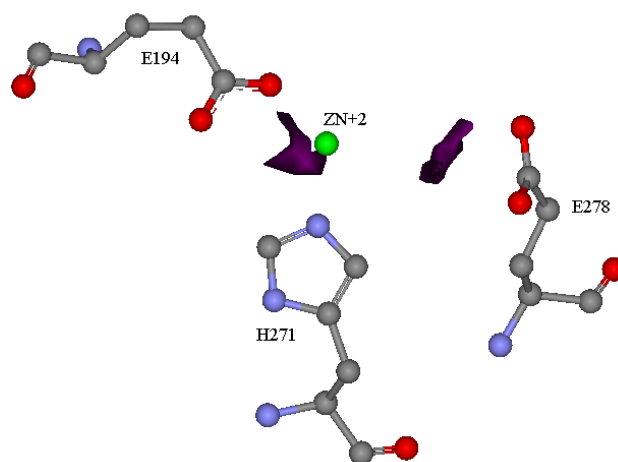
Figure 3.22   An energetically favorable binding sites for Zn^{2+} in the DHQS open structure.   The contour level is set at –50 kcal/mol.   The Zn^{2+} cation (green) has been overlayed on the data map to aid in visualization.

171

The Zn^{2+} was incorporated into the DHQS open structure file and the structure was reanalyzed for potential Zn^{2+}-binding sites. An energetically favorable site was located near residues E81, K84, N162, and D156 (Figure 3.23).
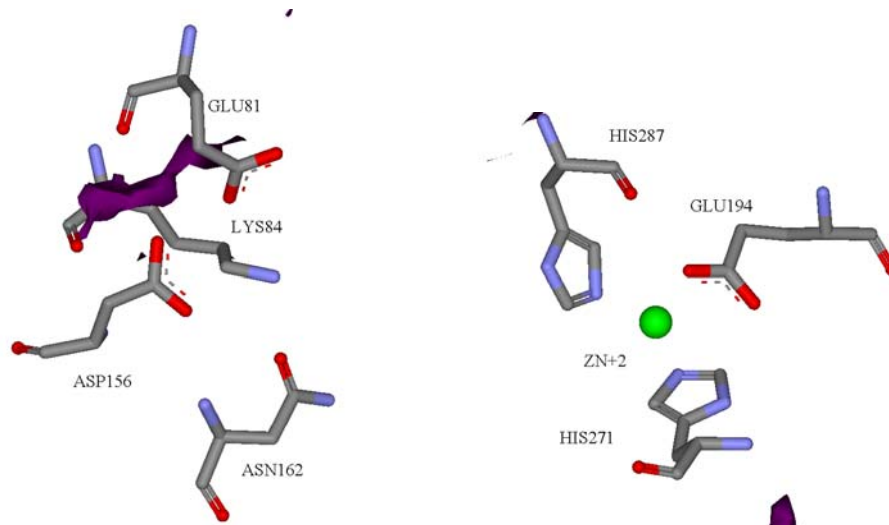
Figure 3.23   View of energetically favorable site for Zn^{2+} binding in the DHQS open structure containing the structurally characterized Zn^{2+} cation. Contour level set at –55 kcal/mol.

The zinc dication probe was employed to investigate the DHQS closed structure for Zn^{2+}-binding sites. The zinc dication probe located two energetically favorable binding sites, the first having nearest neighbor residues E81, K84, N162, and D156, and the second E194, H287, and H271 (Figure 3.24). The point corresponding to the later was 2.1 Å from the structurally characterized binding site of Zn^{2+} (Figure 3.24), indicating a close approximation by GRID of the experimentally determined binding.

172
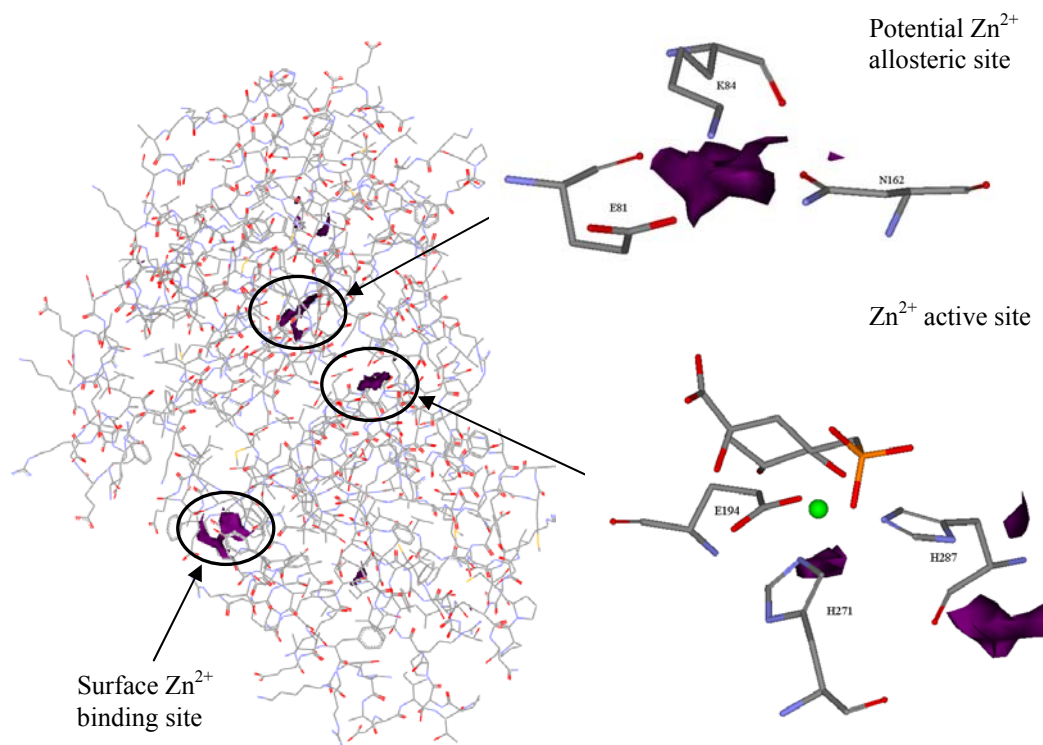
Figure 3.24 One monomer indicating the three most energetically favorable sites for the DHQS closed structure, one at the surface and two near the experimentally known active site. The residues near the two regions of interest have been featured in more detailed figures. The carbaphosphonate (CRB) and Zn^{2+} (green) molecules have been overlayed on the data map to aid in visualization.   The contour level is set at -50 kcal/mol.

Zn^{2+} was incorporated into the DHQS closed structure and it was reanalyzed. Two energetically favorable binding sites were found on the surface. The first energetically favorable site to occur on the interior of the enzyme had nearest neighbor residues E81, K84, N162, and D156, which are identical to the residues found when analyzing the DHQS closed apoenzyme (Figure 3.25). The

173

most energetically favorable binding site for Zn^{2+} in the DHQS closed structure is located 14.8 Å away from the experimentally known Zn^{2+} active site.
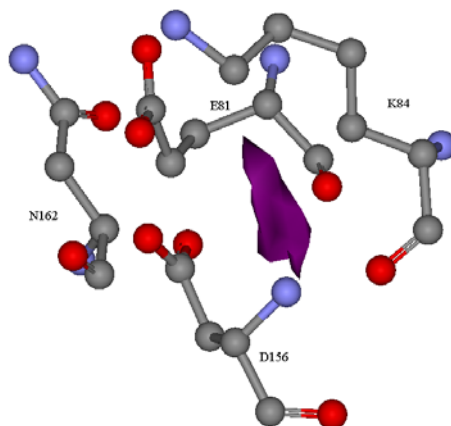
Figure 3.25   The most energetically favorable binding sites for the zinc dication probe for the DHQS closed structure.   Contour level set at –50 kcal/mol.   The interacting residues are indicated as ball-and-stick models.

EPSPS Calculations

GRID calculations were performed on both the open and closed forms of EPSPS. All sulfate and water molecules were removed from the open structure and calculations were preformed on the apoenzyme. In the closed form, the water, formate, glyphosate, and shikimate 3-phosphate molecules were removed prior to the initial calculations. The shikimate 3-phosphate molecule was then placed back into the structure file to simulate binding energies after the first step of the ordered mechanism[47] and the calculations repeated. Unless otherwise stated, all calculations are assumed to be on the apoenzyme.

174

Several probes were used to investigate potential binding sites on the enzyme. The probes are designed to locate energetically favorable binding sites on the enzyme for the functional moiety they represent. Based on the structure of the substrates, shikimate 3-phosphate and PEP, inhibitor glyphosate, probes were chosen to model their functional groups (Figure 3.26). The carboxy, carbonyl, and multi-atom carboxy probes were selected due to the presence of a carboxy functional group on all three molecules (Figure 3.26). The phosphate or phosphonate moeity on shikimate 3-phosphate, PEP, and glyphosate respectively, is a primary binding feature of all three structures (Figure 3.26). The hydrogen phosphate monoanion and phosphate dianion probes were chosen to search for binding sites in the enzyme for HPO_4^- and PO_4^{2-}, respectively. The hydroxyl probe was chosen to simulate the two OH moieties present on the shikimate 3-phosphate molecule. The hydroxyl functional group is one of the only probes, which could be used to distinguish shikimate 3-phosphate binding from PEP and glyphosate binding.

175

Shikimate 3-phosphate
S3P

Glyphosate

Phospho*enol* pyruvate
PEP
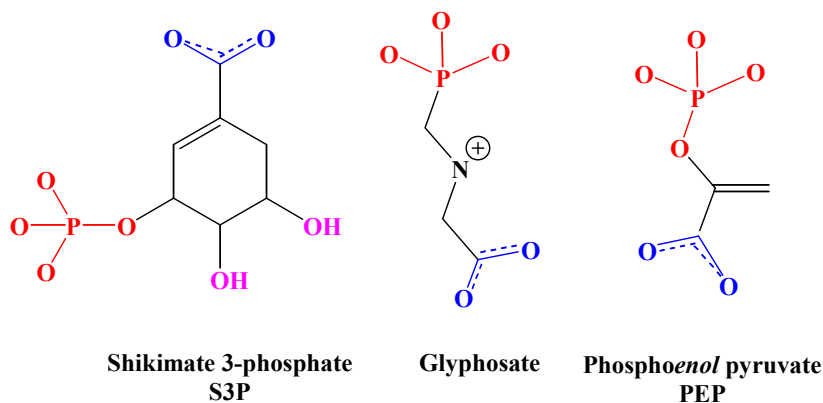
Figure 3.26 The hydroxyl, carboxy, carbonyl, multi-atom carboxy, hydrogen phosphate monoanion, and phosphate dianion probes will be used to locate the binding pocket for the color coordinated functional groups on the DHQS substrate and substrate-analogue.

<u>Hydroxyl Probe</u>

The open EPSPS structure was explored with the hydroxyl probe to locate regions of the protein that favor interaction with an OH moiety. A number of favorable interactions were located throughout the protein with similar energetic favorability. None of the locations were significantly more favorable when compared to other interaction and, therefore, did not merit being labeled as a hydroxyl binding pocket.

The hydroxyl probe was used to investigate the EPSPS closed structure for a potential area of interaction for an alkyl hydroxyl group on the enzyme. Energetically favorable interactions were located at the phosphate binding pocket for the shikimate 3-phosphate molecule and the carboxy binding pocket for the glyphosate molecule (Figure 3.27). Weak interactions were also found between

176

the hydroxyl probe and residues D313 and K22, corresponding to the binding

location of the hydroxyl moieties of the shikimate 3-phosphate molecule (Figure

3.27).
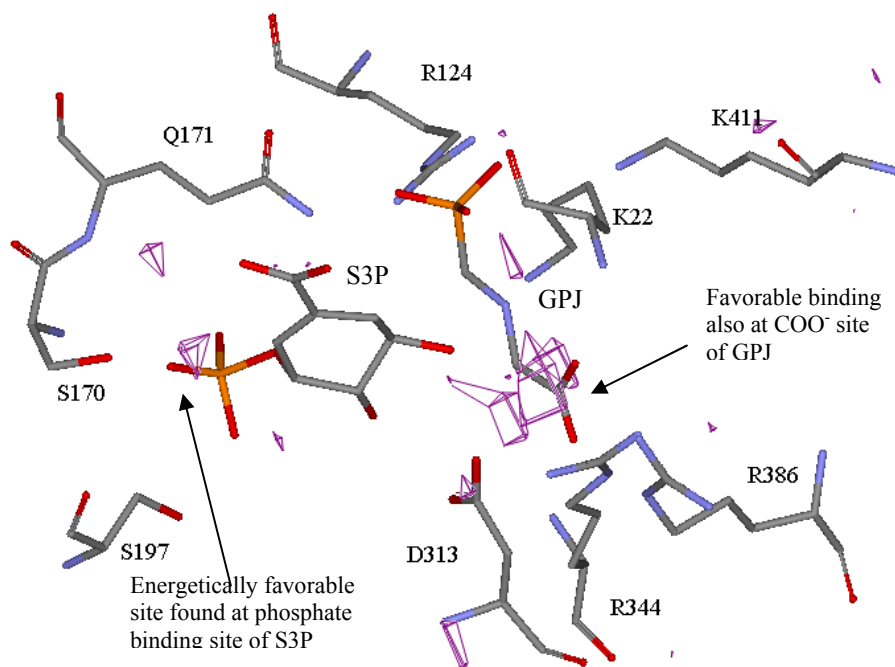
Figure 3.27   The two most energetically favorable binding sites for the hydroxyl probe on the EPSPS closed structure.   Contour level set at −10 kcal/mol.   The shikimate 3-phosphate (S3P) and glyphosate (GPJ) molecules have been overlayed on the data map to aid in visualization.

Carboxy Probe

The open EPSPS structure was investigated with the carboxy probe to

located regions of the protein that favor interactions with an anionic carboxy

oxygen atom.   Several energetically favorable sites were located both on the

surface and the interior of the enzyme.   The carboxy probe located the

experimentally known phosphonate binding pocket of glyphosate in the closed structure, R124, Q171, G96 and N94. At energy values around –8 kcal/mol favorable binding sites are found near residues R27 and S23, which correspond to the structurally known carboxy group binding site of shikimate 3-phosphate in the closed EPSPS structure (Figure 3.28).
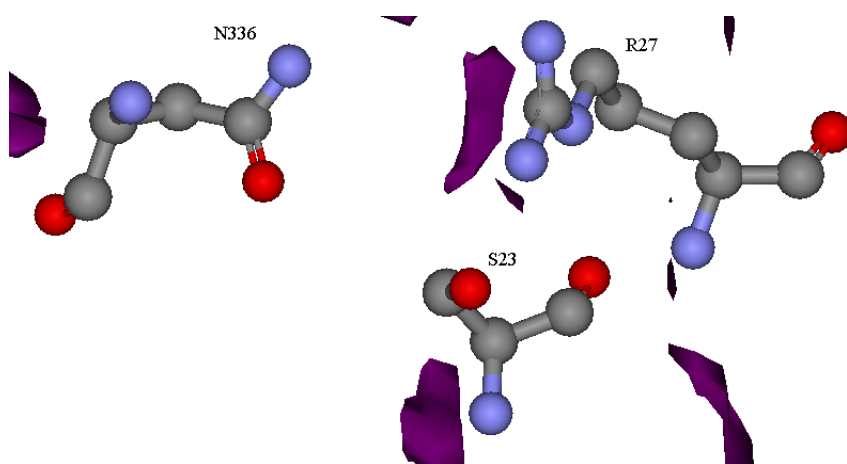
Figure 3.28 View of glyphosate phosphonate binding pocket recognized by the carboxy probe in the EPSPS open structure. Contour set at –9.5 kcal/mol. Interacting residues are indicated as ball-and-stick models.

The EPSPS closed structure was investigated with the carboxy probe to locate the most favorable binding site for sp^2 carbonyl oxygen atoms. Again, several energetically favorable sites were located both on the surface and in the interior of the enzyme. Residues K340 and S170 were identified to form favorable interactions with a carboxy functional group. These residues compose part of the phosphate binding pocket of the shikimate 3-phosphate molecules. A

second site of interaction for a carboxy moiety was determined by GRID to involve residues R344 and R386, which form the carboxy binding pocket of the glyphosate molecule in the EPSPS closed conformation (Figure 3.29).

Figure 3.29 The energetically favorable binding sites for the carboxy probe in the EPSPS closed structure. Contour level set at −11 kcal/mol. The shikimate 3-phosphate (S3P) and glyphosate (GPJ) molecules have been overlayed on the data map to aid in visualization.

Carbonyl Probe

The carbonyl probe, a small molecule probe designed to locate energetically favorable sites for interaction with a carbonyl oxygen atom, was

used to locate favorable sites in the open EPSPS structure. A potential binding site for a sp^2 carbonyl oxygen functional group was found near residues R124, Q171 and G96. The size and shape of the contours were found to be nearly identical to those found for the carboxy probe (Figure 3.28).

The EPSPS closed structure was evaluated with the carbonyl probe. An energetically favorable binding site for a sp^2 carbonyl oxygen functional group was found near residues K340 and R344. The site corresponds to the structurally identified binding pocket of the carboxy functional group on glyphosate. The size and shape of the contours, which are involved in the interaction, were found to be nearly identical to the carboxy probe (Figure 3.29).

Carboxy Multi-atom Probe

The carboxy multi-atom probe (Figure 3.17) was employed to locate the most energetically favorable binding sites for a carboxy functional group in the EPSPS open enzyme. Energetically favorable binding sites for the carboxy multi-atom probe were found to have nearest neighbor residue interactions S23 and R27 (Figure 3.30). These residues are known to form the carboxy binding pocket for the shikimate 3-phosphate molecule in the closed EPSPS structure.

Energetically favorable carboxy
binding pocket before enzyme closes
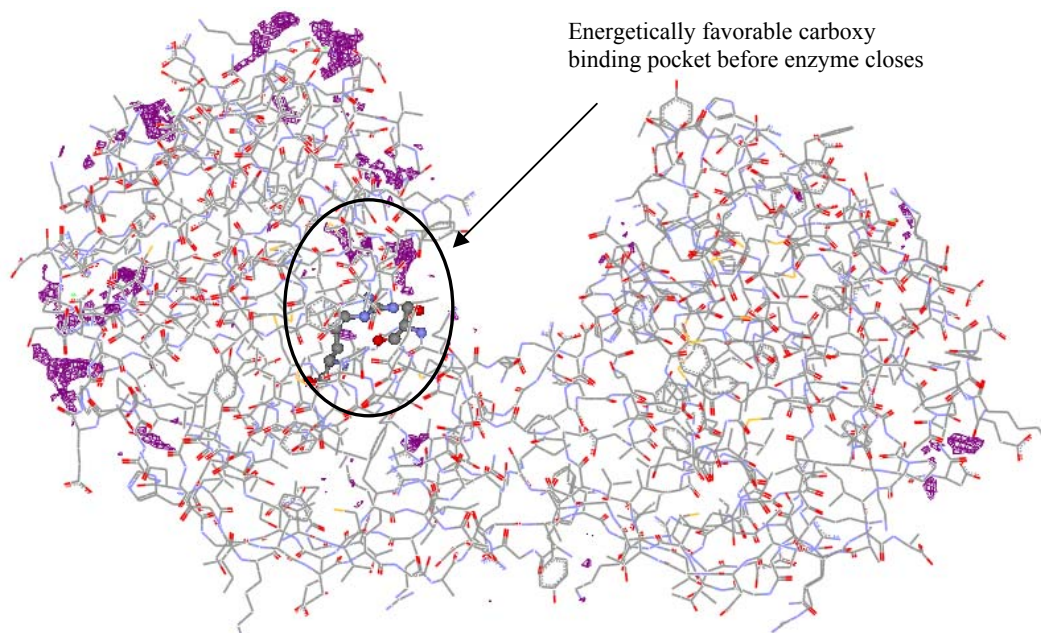
Figure 3.30 View of the full structure of the apoenzyme illustrating the binding pocket located by the carboxy multi-atom probe in the EPSPS open structure. The contour level is set at –12 kcal/mol.

The EPSPS closed structure was evaluated for energetically favorable binding sites for a COO^- functional group using the carboxy multi-atom probe. Energetically favorable sites were located near residues S170, Q171, and S197. These residues correspond to the known binding region of the phosphate unit of the shikimate 3-phosphate molecule. After locating the shikimate 3-phosphate binding region, the phosphonate binding residues of the glyphosate molecule, R124, G96, and Q171, were identified as energetically favorable residues toward a carboxy multi-atom probe (Figure 3.31).
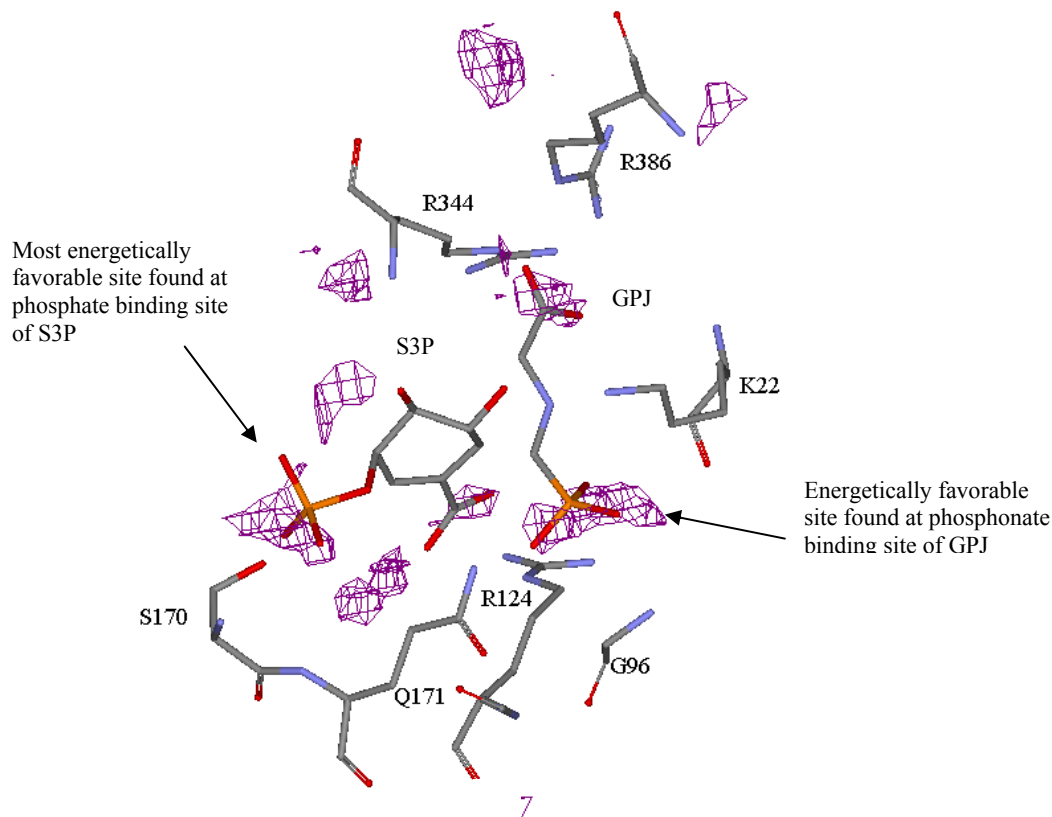
181

Figure 3.31 The energetically favorable binding sites for the carboxy multi-atom probe in the EPSPS closed structure. Contour level set at −15 kcal/mol. The shikimate 3-phosphate (S3P) and glyphosate (GPJ) molecules have been overlayed on the data map to aid in visualization.

Phosphate Dianion Probe

The EPSPS open structure was analyzed with the phosphate dianion probe to simulate a phosphate dianion with no hydrogens. The probe located three important energetically favorable binding sites for a PO_4^{2-} moiety. Residues Q171, R124, N94, K411, and G96 create a binding pocket for phosphate, which is known to be the active site for the phosphonate group of the glyphosate molecule. The residues participating in the binding site of the phosphate functional group on the shikimate 3-phosphate molecule were also identified, S170, S169, and S197. Additionally, residues R344 and R386 were determined to create a binding pocket energetically favorable for PO_4^{2-} bonding; however, the location is the structurally

characterized binding site of the carboxy moiety of the glyphosate molecule (Figure 3.32).
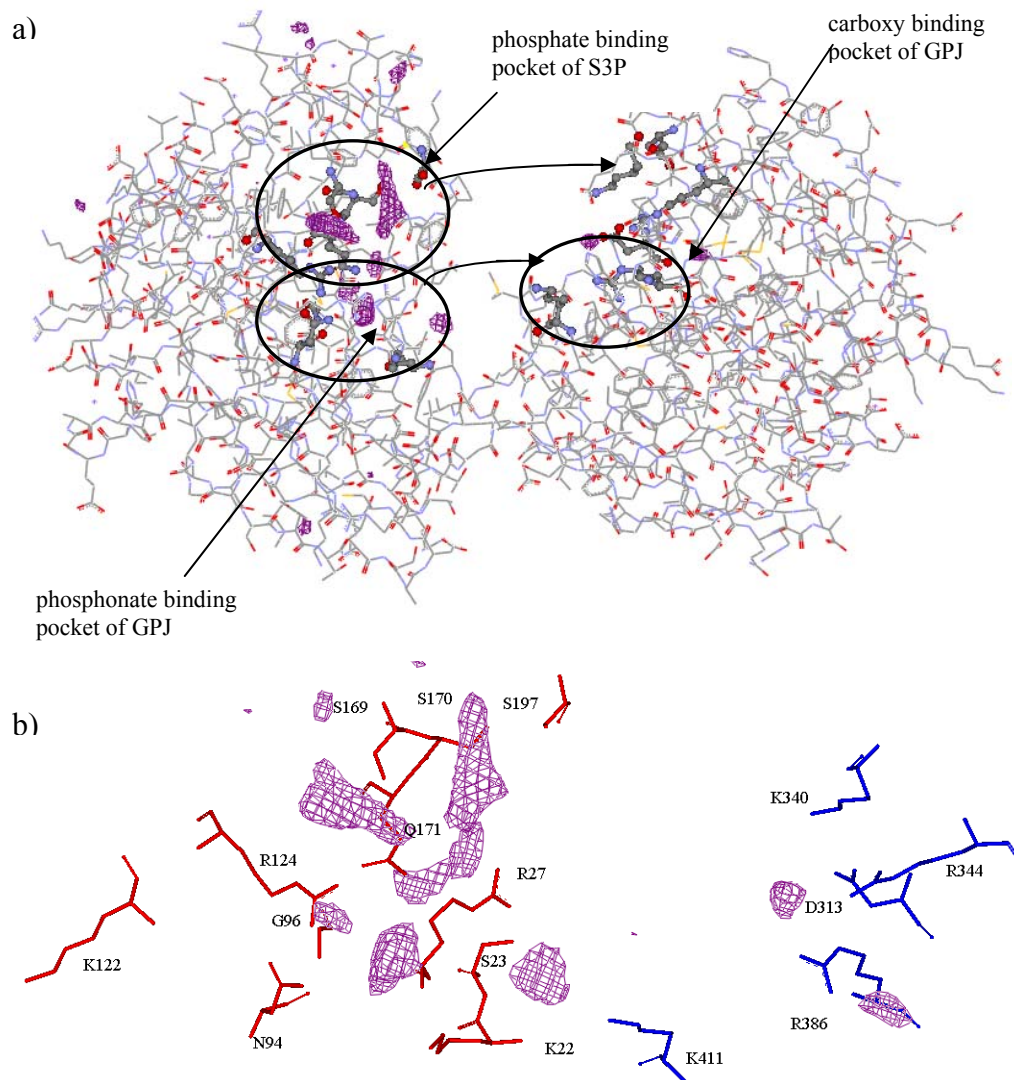
Figure 3.32 a) View of the full structure of the apoenzyme illustrating the three important binding pockets located by the phosphate dianion probe in the EPSPS open structure. S3P and GPJ represent shikimate 3-phosphate and glyphosate, respectively. b) A detailed view of the active site with the N- and C-terminal domains indicated in red and blue respectively. The contour level is set at –12 kcal/mol.

The EPSPS closed structure was evaluated with the phosphate dianion probe to locate the most favorable binding sites for a PO_4^{2-} moiety. Residues interacting with the most energetically favorable binding site were G96, R124, and K411, which are the known nearest neighbor residues of the phosphonate group of the bound gylphosate molecule. A second site favorable for phosphate binding was located with nearest neighbor residue interactions K340, S170, and S197, which is the binding pocket of the shikimate 3-phosphate molecule (Figure 3.33).

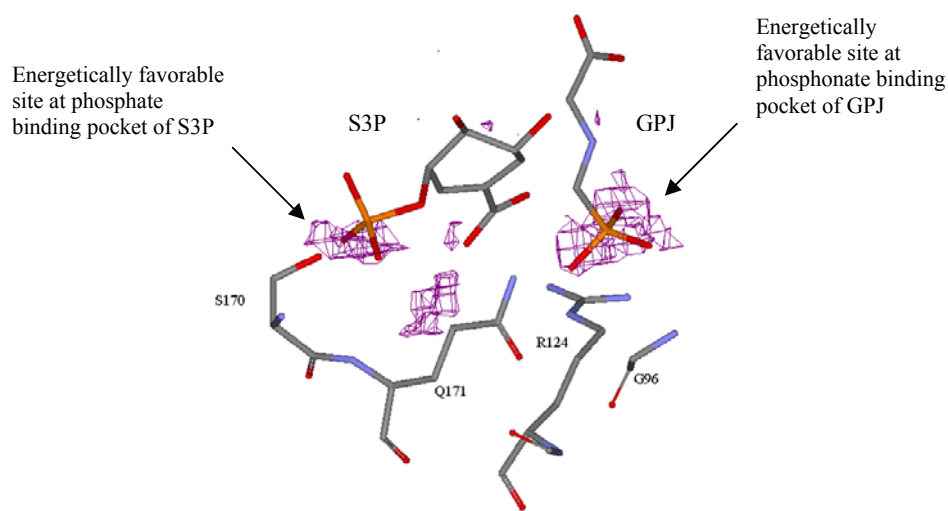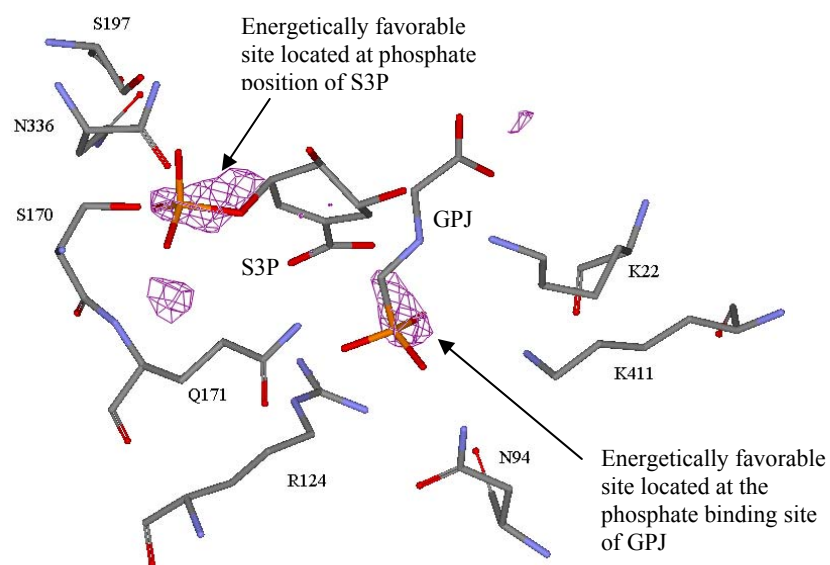**Figure 3.33 Energetically favorable binding sites for the phosphate dianion probe in the EPSPS closed structure. Contour level set at −18 kcal/mol. The shikimate 3-phosphate (S3P) and glyphosate (GPJ) molecules have been overlayed on the data map to aid in visualization.**

184

Hydrogen Phosphate Monoanion Probe

The open EPSPS structure was explored with the hydrogen phosphate monoanion probe to locate energetically favorable sites on the protein for HPO_4^- interactions to occur. Residues K411, R124, G96, Q171, and N94 create a binding pocket for HPO_4^- functional groups. The site is known to be the active site for the phosphonate group of the glyphosate molecule. The residues participating in the binding site of the phosphate functional group on the shikimate 3-phosphate molecule were also identified, S170, S169, and S197. Additionally, residues R344 and R386 were determined to create a binding pocket energetically favorable for HPO_4^- interactions; however, the location is the structurally determined binding site of the carboxy moiety of the glyphosate molecule. The size and shape of the countour maps were nearly identical to those determined using the phosphate dianion probe (Figure 3.32).

Using the hydrogen phosphate anion probe to locate energetically favorable binding pockets for HPO_4^- in the EPSPS closed structure, similar results were observed to those found using the phosphate dianion probe. Residues interacting with the most energetically favorable binding site were G96, R124, and K411, which are the known nearest neighbor residues of the phosphonate group of the bound gylphosate molecule. A second site favorable for phosphate

binding was located with nearest neighbor residue interactions K340, S170, and S197, which is the structurally determined binding pocket of the shikimate 3-phosphate molecule. The size and shape of the contours for the hydrogen phosphate monoanion probe and the phosphate dianion probe on the EPSPS closed structure were nearly identical (Figure 3.33).

DISCUSSION AND CONCLUSIONS

DHQS Discussion

DHQS is a multi-functional enzyme, which is believed to perform five consecutive chemical reactions in a single active site.[15-18, 48-50] Whether the enzyme is actively involved in each of these steps,[15, 18] or rather several steps occur spontaneously[16, 17, 48] has been the topic of considerable debate. The substrate-analogue, carbaphosphonate, has been crystallographically characterized bound to the enzyme and has provided insight into the residues which could play key roles in the binding of the inhibitor and the catalytic mechanism.[6] The structure provided evidence to support DHQS's ability to perform multi-step catalysis without the formation of unwanted by-products.

Another interesting characteristic of DHQS is that it exists in two forms, open[7, 29] and closed.[6, 7, 29] Prior to binding of the substrate, DHQS is loosely packed with a moderately exposed active site; however, in the presence of

substrate or substrate analogue, the structure under goes a conformational change to a more tightly packed form (Figure 3.10 a and b).

In addition to performing multi-step catalysis and interesting conformational changes, DHQS, is of appeal as a potential target for the development of antimicrobials.[3] Because the shikimate pathway is absent from mammals, antimicrobials targeted at DHQS are expected to have limited toxicity in humans.[3]

To aid in understanding the substrate binding site potentials in the active site region, the program GRID was implemented. The use of GRID has provided information regarding the active site recognition features in the open DHQS structure and potential interactions, which may initiate enzyme closure. Residues involved in the catalytic mechanism of DHQS were determined using GRID software. Identification of the key active site binding pockets has allowed for the design of a potential antimicrobial target compound.

In an attempt to identify substrate recognition pockets in the open enzyme, the DHQS open structure based on crystallographic analysis was evaluated using several GRID probes. The substrate, DAHP, has both a carboxy and a phosphonate moiety, which play an important role in binding and orienting the substrate molecule in the enzyme. The known carboxy binding pocket was corroborated using GRID analysis. The carboxy, carbonyl, and carboxy multi-

atom probes located interactions with the C^δ and amino group of the side chain of K250 and the amino group of K152. The phosphate dianion and hydrogen phosphate monoanion probes identified similar interactions, the N^ε of the side chain and the terminal amino group of the side chain of R264, the side chain amino group of K152 and K250 (Figure 3.13 and Figure 3.18). The phosphate binding pocket for the substrate was determined using phosphate dianion and hydrogen phosphate monoanion probes (Figure 3.35). Energetically favorable interactions with the side chain amino groups of R130(B) and the side chain amino groups of K152 and N162 indicate the presence of a phosphate binding pocket prior to the conformational change (Figure 3.19).
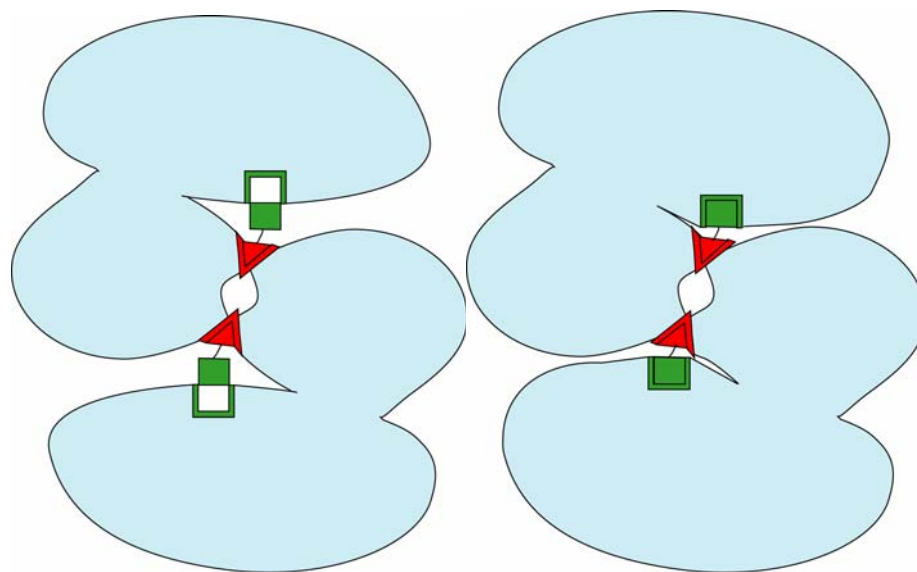
Figure 3.34 Cartoon diagram of binding pockets and carbaphosphonate prior to the closure of DHQS followed by the docking of the substrate and enzyme closure. The phosphate and carboxy binding pockets are indicated in red and green respectively.

The presence of the energetically favorable phosphate binding site in the open structure may indicate a substrate binding recognition feature. The phosphate functional group of the substrate could initially bind in the pocket formed by R130(B), K152, and N162, which would correctly position the molecule before the closure of the enzyme. DHQS undergoes a conformational change in the presence of substrate or substrate-analogue;[6, 7, 29] GRID calculations indicate that open structure binding recognition features may play a vital role in the conformational change which occurs in the enzyme.

Evaluation of the DHQS closed structure with GRID probes has increased our understanding of which residues play an important role in the catalytic mechanism of converting DAHP to DHQ. The carboxy, carbonyl, and multi-atom carboxy probe located similar energetically favorable binding pockets. The most energetically favorable pocket involved residues K152, K250, and R264. These residues are proposed to play an important role in the catalytic mechanism of DHQS by binding the carboxy functional group of the substrate.[6] Based on GRID output, the backbone nitrogen and oxygen of K152, the side chain amino group of K250, as well as the backbone oxygen and the side amino group of R264 are the moieties that create a binding pocket energetically favorable for a carboxy functional group (Figure 3.35). Residues K256, N268, R130(B), N162, and K356 also show energetically favorable interactions with the carboxy, carbonyl, and

multi-atom carboxy probes (Figure 3.35). Based on the X-ray crystal structure, these residues form the phosphate binding pocket of the substrate-analogue molecule. Analysis of the DHQS closed structure with the hydrogen phosphate monoanion and phosphate dianion probes indicated identical atom interactions as those found using the carboxy, carbonyl, and multi-atom carboxy probes (Figure 3.14, Figure 3.16, and Figure 3.20).
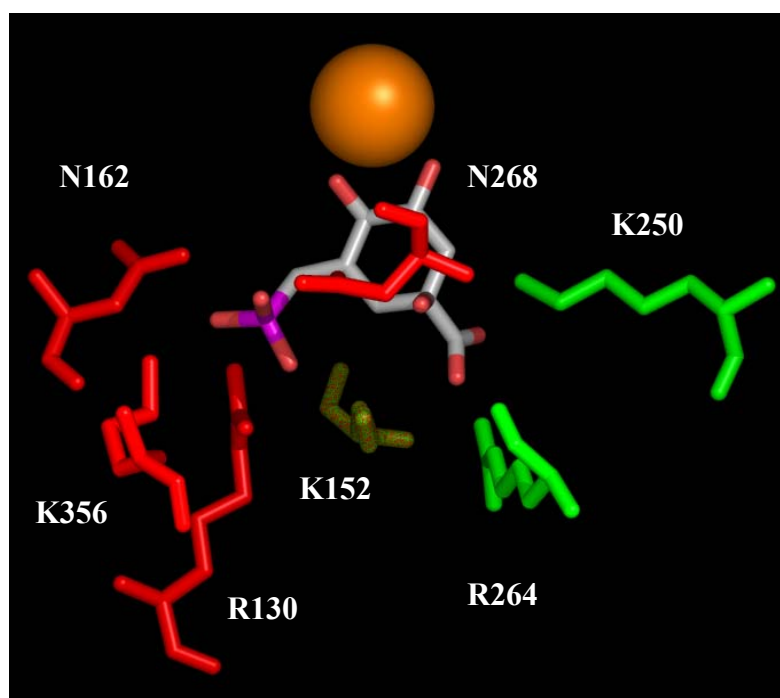
Figure 3.35 Residues in the active site of DHQS with energetically favorable interactions towards the carbonyl, carboxy, carboxy multi-atom probes indicated in green, while residues with energetically favorable interactions towards the phosphate and hydrogen phosphate probes shown in red. The green/red residue interacts with both the carboxy and phosphonate moieties of carbaphosphonate. Zn^{2+} is indicated in orange. Figure created using Pymol (DeLano Scientific, San Carlos, California).

190

Evaluating the residue interactions proposed by Carpenter *et al.*[6] and the new information provided *via* GRID calculations, specific atomic interactions participating in the mechanism can be discussed. The substrate is held in place by energetically favorable binding pockets toward the carboxy and phosphate moieties. During the ring-opening step, the possibility of epimerization at C2 of intermediate D (Figure 3.3) arises. Previous work has indicated that no epimerization occurs about the C2 bond in the enzyme-mediated reaction.[18] GRID calculations indicate that the absence of epimerization is due to strong bonding interactions between the carboxy group on C2 of intermediate D and the backbone nitrogen and oxygen groups of K152, the side chain nitrogen of K250 and the backbone oxygen and side chain amino group of R264.

Interestingly, H275 is not believed to contribute to the binding pocket for the phosphate moiety of the substrate; however, it is recognized by all five GRID probes explored to be energetically favorable toward carboxy-type and phosphate-type probes. The nitrogen atom and CH group of the histidine ring of H275 were all determined to be part of an energetically favorable region for phosphate and carboxy binding. H275 is hypothesized to participate in a proton-shuffling system[6] during formation of intermediates A and C of the mechanism (Figure 3.3). In the formation of intermediate A, H275 acts as a proton acceptor, while it is a proton donor for the formation of intermediate C (Figure 3.3). Considering

the proximity to the active site and the amphoteric nature of this histidine residue, it is not surprising that it acts favorably toward the GRID probes utilized.

Inhibitors designed to target DHQS are promising as antimicrobial agents due to the presence of the shikimate pathway in microorganisms, but not mammals.[51] Carbaphosphonate is a nanomolar-level, slowly reversible inhibitor of DHQS, which was designed to competitively bind with the substrate DAHP.[16, 17] Modifications to the functional groups of carbaphosphonate, such as replacement of the phosphonate moiety with $CH_3(CO_2H)_2$ and mono saturation of the cyclohexanol ring, have also produced DHQS inhibitors.[50] Based on GRID calculations and previous methods used to inhibit metalloenzymes (*e.g.* [52]), potential inhibitors of DHQS could involve coordination to the Zn^{2+} and blocking of the active site. Taking advantage of the high affinity of Zn^{2+} for S^-, an inhibitor, such as that shown in Figure 3.36, has the potential to occupy the substrate binding pocket. The carboxy functional group of the proposed inhibitor would bind to residues K250, R264, and K152, in a manner similar to the substrate, while the sulfide moiety interacts with the Zn^{2+}. The substrate, DAHP, has an overall charge of –3, while that of the proposed inhibitor is –2 making it easier for the inhibitor to cross cell membranes.
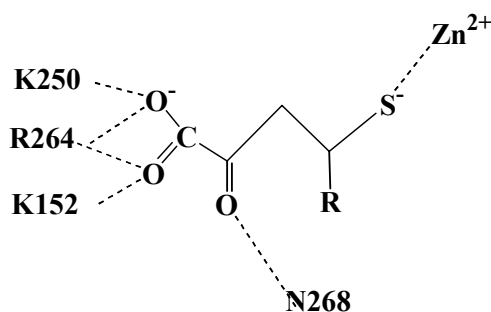
Figure 3.36 Proposed inhibitor for DHQS, which utilizes both the carboxy binding pocket and the Zn^{2+} cation.

DHQS requires the presence of a divalent cation,[14] such as Zn^{2+} for activity. However, *in vitro* trivalent lanthanides,[53] such as Sm^{3+} and Eu^{3+}, have also been shown to activate the enzyme. Interestingly, the presence of large quantities of metal ions in solution has been shown to decrease the activity of the enzyme, suggesting the presence of an allosteric site.[54] Investigating the DHQS open and closed structures for the presence of Zn^{2+} active sites has provided additional evidence to support the hypothesis of an allosteric site. The zinc dication probe located two potential Zn^{2+} binding sites near the active site of both the open and closed DHQS structure (Figure 3.22 and Figure 3.24). In the open structure the most energetically favorable site was determined to interactions with residues H287, E194, and H271. This site matched the known binding site of Zn^{2+} within a distance of approximately 1 Å. In the DHQS closed structure, the zinc dication probe found residues E81, K84, N162, and D156 to provide the most energetically active site for Zn^{2+} binding. This site is approximately 15 Å away

193

from the structurally characterized binding site of Zn^{2+}. The known Zn^{2+} binding site, H287, E194, and H271, was also determined to be energetically favorable in the DHQS closed structure. GRID provides significant evidence for the presence of an allosteric site, which could account for the reduced activity when the enzyme is in the presence of a large excess of metal cations.

During catalysis, DHQS has been shown to undergo a conformational change (Figure 3.34). DHQS forms a tightly packed structure in the presence of substrate called the "closed" form,[6, 7, 29] whereas in the absence of substrate it exists in an "open" conformation.[7, 29] The availability of structural data for both the open and closed forms of DHQS has allowed for a computational study to identify recognition features of the enzyme that induce the substrate-dependent conformational change. Both the carboxy and phosphate binding pockets are present prior to the closure of the structure. If the substrate initially binds to the carboxy site then the charge on the phosphate moiety could potentially attract the binding pocket residues, R130 and N162, initiating the conformational change (Figure 3.37).
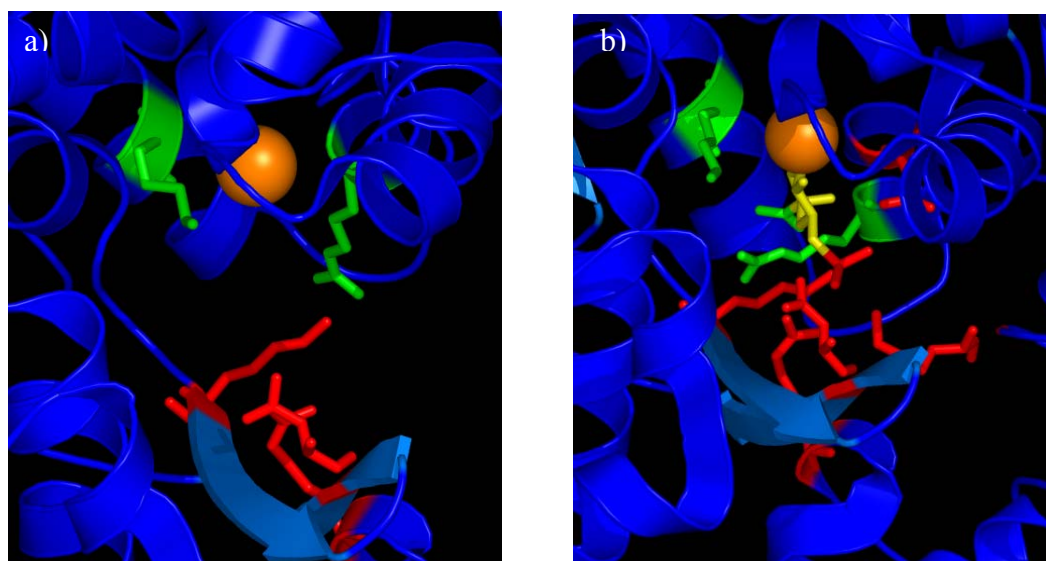
Figure 3.37 Carboxy (green) and phosponate (red) binding pockets present in DHQS a) open and b) closed conformations. Zn^{2+} is indicated in orange, while carbaphosphonate is depicted in yellow with the carboxy and phosphonate moieties highlighted in green and red, respectively. Figure created using Pymol (DeLano Scientific, San Carlos, California).

EPSPS Discussion

EPSPS is one of the most extensively studied enzymes in the shikimate pathway due to the commercial use of the inhibitor glyphosate as an herbicide.[34] Glyphosate does not bind to the apoenzyme, but rather to the enzyme—shikimate 3-phosphate complex.[34] Much work has been devoted to determining the exact binding site of glyphosate and how it competes with PEP.[8, 23-25] GRID calculations were undertaken to investigate the binding site of glyphosate and shikimate 3-phosphate EPSPS.

In addition to locating functional group binding pockets in EPSPS, GRID calculations were utilized to investigate the residues which play a vital role in the catalytic mechanism of EPSPS. Considerable debate over the possibility of a random rather than an ordered mechanism for substrate binding in EPSPS has occurred.[36, 37] The use of GRID will provide insight into which mechanism is most likely to occur based on the presence of binding pockets in the apoenzyme.

High resolution X-ray crystal structures of EPSPS, uncomplexed and two complexed structures (one with shikimate 3-phosphate and glyphosate[9, 29] and the other with shikimate 3-phosphate and formate),[8, 9, 29] have improved our understanding of the mechanism and inhibition of EPSPS. The availability of both forms of EPSPS, open and closed, provides the opportunity to perform computational analysis to further investigate the environment around the active site. GRID calculations were employed to provide insight into the interactions of shikimate 3-phosphate and glyphosate with EPSPS and identify the presence of binding pockets, which may contribute to the conformational change that occurs in the enzyme.

The shikimate pathway has been identified as a potential drug target for the development of new antimicrobial agents.[51] The presence of the shikimate pathway in bacteria, fungi, plants and some apicomplexan parasites, but not in mammals,[5] indicates that antimicrobial agents aimed to disrupt this pathway may

196

have minimal side effects on mammals. A primary use of GRID calculations is the design of lead compounds based on information regarding energetically favorable binding sites in a structure. GRID calculations were performed on the structure of EPSPS to search for potential energetically favorable binding sites for antimicrobial lead compounds.

As discussed previously, EPSPS catalyzes the condensation of PEP with shikimate 3-phosphate to produce EPSP (Figure 3.1).[3] The reaction can be inhibited by glyphosate which does not bind to the apoenzyme, but rather to the enzyme-shikimate 3-phosphate complex.[4] In an attempt to identify substrate and inhibitor recognition pockets in the open enzyme, the EPSPS open structure was evaluated using several GRID probes. Four binding pockets were located, the phosphate/phosphonate binding pockets for shikimate 3-phosphate and glyphosate and the carboxy binding pockets for shikimate 3-phosphate and glyphosate (Figure 3.38).
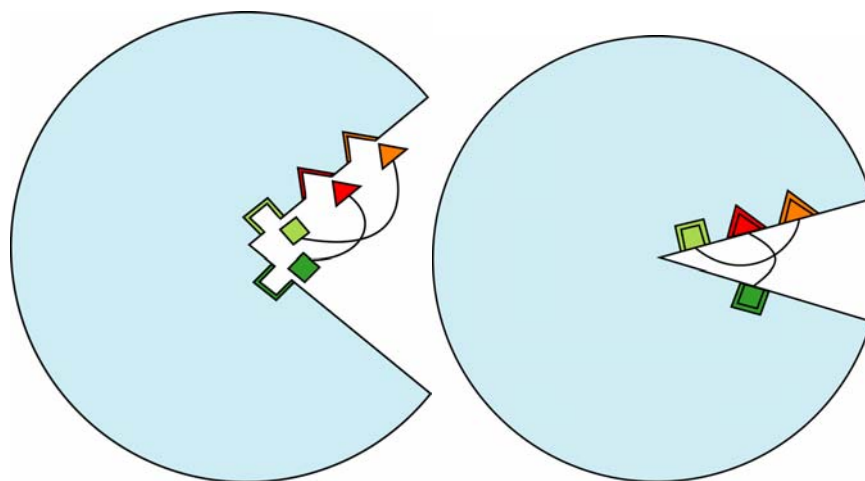
Figure 3.38 Cartoon diagram of phosphate and carboxy binding pockets prior to the closure of EPSPS. The enzyme closes and the shikimate 3-phosphate and glyphosate molecules are shown in their docking sites. The phosphonate and carboxy binding pockets of glyphosate are indicated in red and green, respectively, while the phosphate and carboxy binding pockets of shikimate 3-phosphate are indicated in orange and light green.

The carboxy binding pocket of shikimate 3-phosphate was determined to be present in the open EPSPS structure (Figure 3.38). Analysis with the carbonyl, carboxy, and carboxy multi-atom probes indicated interactions with side chain oxygen of S23 and amino group of R27 (Figure 3.27). The importance of R27 in enzyme activity has been demonstrated by the mutation from arginine to alanine resulting in a complete loss of activity.[55] This mutation, along with the GRID calculations, indicates the vital role played by the presence of the basic guanido group in the binding of the carboxy functional group of shikimate 3-phosphate.

The phosphate functional group of shikimate 3-phosphate is known to bind to S170, S169, and S197; this was corroborated using GRID calculations.

198

The hydrogen phosphate and phosphate probes located interactions with the backbone nitrogen and side chain oxygen of S170, the side chain oxygen of S169 and S197 (Figure 3.33). The presence of this energetically favorable binding site prior to the closure of the enzyme may indicate it plays a critical role in substrate initiated conformational change.

The carboxy binding pocket of shikimate 3-phosphate was determined to be present in the open EPSPS structure (Figure 3.38). Analysis with the carbonyl, carboxy, and carboxy multi-atom probes indicated interactions with the side chain amino groups of R344 and R386 (Figure 3.29 and Figure 3.31). These residues may play an important role in the EPSPS mechanism by positioning PEP within the active site to enable protonation of PEP by the 5-hydroxyl of shikimate 3-phosphate.[56]

GRID calculations using the phosphate and hydrogen phosphate probes located an energetically favorable binding site involving residues G96, R124 and Q171 (Figure 3.33). These residues are known to be involved in forming the phosphonate binding pocket of glyphosate and are potentially involved in the binding of PEP. The phosphate-type probes interacted with the backbone nitrogen group of G96, the amino side chain groups of R124 and the side chain amino group of Q171. Based on site-directed mutagenesis studies, the residue has been shown to play a vital role in glyphosate binding. Mutation of the glycine to

199

an alanine greatly reduces the enzyme's affinity for glyphosate.[25, 57] The mutation and the information acquired from GRID calculations indicate the presence of the hydrogen to be important for the binding of the phosphonate moiety of glyphosate.

The EPSPS reaction involves the transfer of the *enol*pyruvyl group of PEP to the 5-hydroxyl group of shikimate 3-phosphate to produce EPSP and inorganic phosphate (Figure 3.1). The most widely accepted mechanism (Figure 3.6)[36] proposes that the reaction involves protonation of C2 of PEP by the acquisition of a proton from a base in the enzyme, with associated nucleophilic attack on C2 of PEP by the C5 hydroxyl of shikimate 3-phoshpate to form the tetrahedral intermediate (Figure 3.6). The following step is the removal of a proton from the C5 hydroxyl by a proton acceptor in the enzyme, resulting in the formation of EPSP and inorganic phosphate (Figure 3.6).[36]

GRID calculations were utilized to provide additional insight into the mechanism for the catalytic conversion of shikimate 3-phosphate to 5-*enol*pyruvylshikimate 3-phosphate by EPSPS. Both shikimate 3-phosphate and PEP have a carboxy and a phosphate moiety. The GRID probes, carboxy, carbonyl, carboxy multi-atom, hydrogen phosphate monoanion, and phosphate dianion, located identical binding pockets in the closed structure which were the phosphate/phosphonate binding pockets of shikimate 3-phosphate and glyphosate.

Important atomic interactions in the shikimate 3-phosphate binding pocket were identified as the side chain oxygen and the backbone nitrogen group of S170 and the side chain oxygens of S169 and S197. The amino group of K411, the side chain amino groups of R124, the backbone nitrogen of G96, the side chain oxygen and amino group of Q171 and the oxygen of N94 were identified as the energetically favorable binding pocket for glyphosate (Figure 3.39).

The phosphate/phosphonate binding pockets play an important role in positioning the substrates in the enzyme. In order for protonation of C2 of PEP to occur, the C5 hydroxyl of shikimate 3-phosphate must be within approximately 2.7 Å of the C2 of PEP to initiate a nucleophilic attack.[58] Additionally, PEP and the proton acceptor residue, believed to be G341,[8] must be within a distance of approximately 2.8 Å to transfer a proton.[59] After formation of the tetrahedral intermediate, a proton is removed from the C5 hydroxyl by a proton acceptor in the enzyme, G341.[8] Again, the phosphate binding pockets may functions to maintain the correct position and orientation of the small molecules during the deprotonation step.

The similarities in functional groups on PEP and glyphosate make it impossible to distinguish between the binding sites of the two molecules utilizing GRID. Both structures contain a carboxy and phosphate moiety (Figure 3.26); the primary difference between the two structures is the separation between the

carboxy and phosphate groups. The maximum separation between the carboxy

and phosphate groups is 5.2 Å and 4.2 Å for glyphosate and PEP, respectively.

GRID is designed only to identify potential functional group binding sites with no

distinction made for the distance between two moieties. However, based on the

lack of additional energetically favorable sites, PEP and glyphosate binding sites

are thought to be similar.

Figure 3.39   The active site of EPSPS indicating the important residue interactions with shikimate 3-phosphate and glyphosate. Residues located in The N- and C-terminal domains are shaded in blue **and** peach**, respectively. Residues found to be energetically favorable toward all probes used are shown in** red**.**

Because the shikimate pathway is absent from mammals, EPSPS is an attractive target for the development of new antimicrobial agents.[3] Based on the structure and mechanism of the widely used inhibitor glyphosate, several attempts have been made to develop lead compounds in the design of new EPSPS inhibitors (*e.g.* [60, 61]). Most work has been targeted at altering the structure of glyphosate to improve inhibition;[61] however, a different approach is the design of a bisubstrate inhibitor which will simultaneously occupy both the S3P- and PEP-binding sites.[60] Previous attempts at designing an EPSPS bisubstrate inhibitor include compounds **A**[62] and **B**[63] in Figure 3.40. Compound **A** mimics the S3P-PEP intermediate and is a potent EPSPS inhibitor. Compound **B** is also an effective EPSPS inhibitor due to its multiple anionic charge at the five position of the dihydroxylbenzoate. Both inhibitors utilize the more synthetically accessible benzene ring, rather than the chiral shikimate ring.
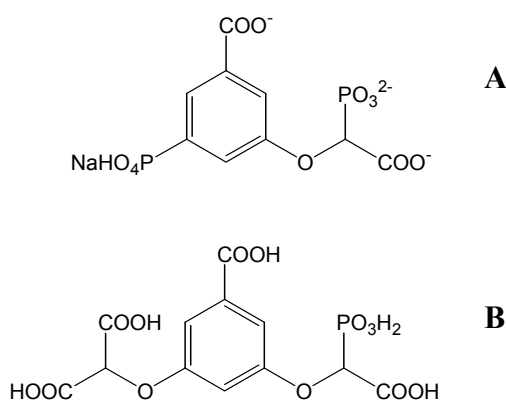
Figure 3.40 Structure of EPSPS inhibitors.

Based on data acquired from GRID calculations, the energetically favorable binding pocket for a phosphate or phosphonate moiety makes the presence of such a group in a potential inhibitor advantageous to assure strong binding. Prior to substrate binding, phosphate binding pockets are present for both the substrate and inhibitor. An energetically favorable binding location for the carboxy moiety of the substrate is also present. Once the structure closes, the two phosphate binding pockets are most energetically favorable. A potential method for inhibiting the functioning of the enzyme it to take advantage of the two phosphate binding pockets by creating a bisubstrate inhibitor which will interact with both pockets. One possibility is the ester shown in Figure 3.41. The phosphate and carboxy functional groups on the shikimate-like structure would bind, correctly orienting the phosphate functional group on the aromatic system to interact with the PEP binding pocket. The flexibility of the ester chain would allow for adjustment of the inhibitor to fit in the binding pocket. Like compounds **A** and **B** (Figure 3.40), the proposed inhibitor also utilizes the synthetically more accessible benzene ring.
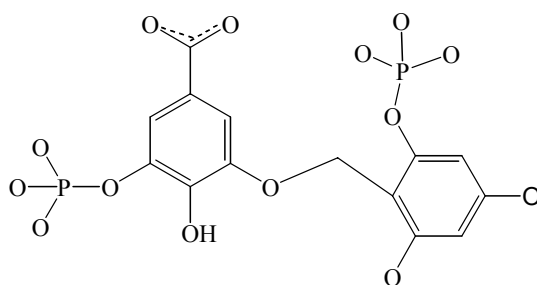
Figure 3.41   Structure of proposed inhibitor for EPSPS based on GRID calculations.

Kinetic evidence by Boocock and Coggins[47] has supported the hypothesis of an ordered sequential mechanism in which binding of shikimate 3-phosphate precedes binding of PEP.   GRID calculations utilizing the carboxy, carbonyl, multi-atom carboxy, hydrogen phosphate and phosphate probes lend additional evidence to strengthen the hypothesis of an ordered binding mechanism.   As previously discussed, the carboxy and phosphate binding pockets for shikimate 3-phosphate proved to be more energetically favorable toward the appropriate probes than were the carboxy and phosphate binding pockets of glyphosate.   The appropriate substrate will be bound to the most energetically favorable binding site before the secondary binding sites will be occupied.   In this case, the shikimate 3-phosphate would first dock into the pocket containing residues S170, S169, and S197 (*i.e.* the most energetically favorable site) prior to the docking of the glyphosate molecule.

It has been shown *via* small-angle X-ray scattering (SAXS) studies that shikimate 3-phosphate can induce a conformational change in EPSPS.[10] The structure is a tightly packed and called the "closed" form whereas in the absence of substrate it exhibits an "open" conformation. The availability of structural data for both the open and closed forms of EPSPS has allowed for a detailed computational study to identify recognition features of the enzyme that induce the substrate-dependent conformational change. GRID calculations indicate that strong phosphate and carboxy binding pockets exist in the open EPSPS structure. These energetically favorable pockets provide a means by which the molecules can bind before the enzyme closes. As indicated by GRID, the most probable mechanism for this is the initial binding of the phosphate moiety into the energetically favorable binding pocket involving residues S170, S169, and S197. The phosphate binding pocket orders the orientation of the shikimate molecule directing the carboxy group toward the carboxy binding pocket located in the opposite domain (Figure 3.32). Due to the proximity of the carboxy moiety on the shikimate molecule to the energetically favorable binding pocket, the attractive forces would be predicted to initiate the closing of the enzyme.
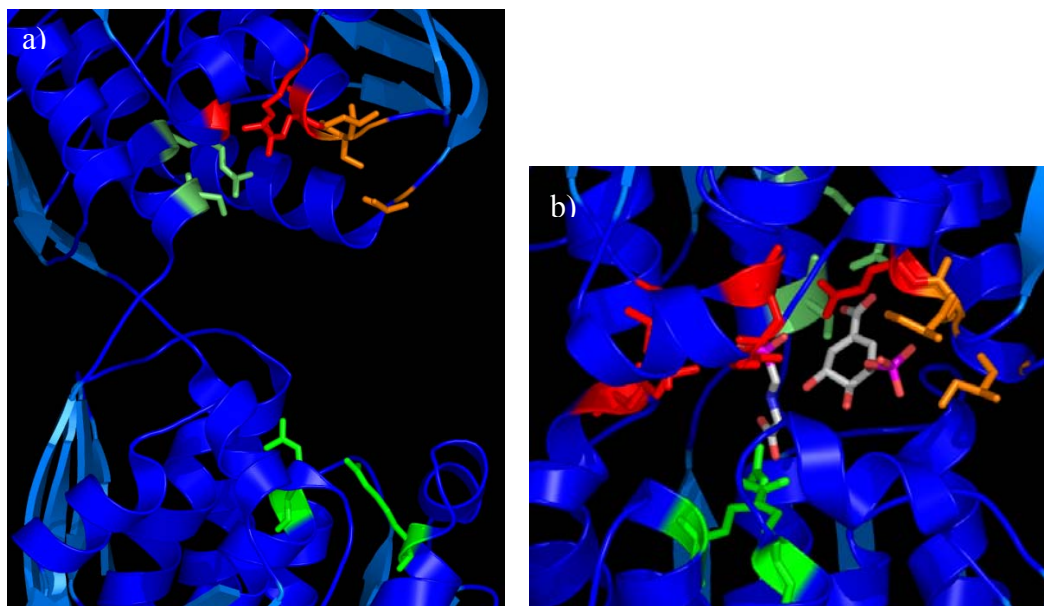
Figure 3.42 **The carboxy (green) and phosponate (red) binding pockets for glyphosate and the carboxy (light green) and phosphate (orange) binding pockets for shikimate 3-phosphate are shown in the EPSPS a) open and b) closed conformations. Glyphosate amd shikimate 3-phosphate molecules are shown with gray carbon atoms, blue nitrogen atoms, red oxygen atoms and pink phosphorous atoms. Figure created using Pymol (DeLano Scientific, San Carlos, California).**

Conclusions and Future Work

GRID calculations were determined to be a valuable method for acquiring information regarding the binding sites for two shikimate pathway enzymes, DHQS and EPSPS. Binding site recognition features, which are present in the substrate free structures, were identified using GRID and their role in orchestrating the conformational change of the enzyme was proposed. GRID

calculations allowed for identification of residues which play a key role in the catalytic mechanism of the enzyme. As both enzymes are potential targets for antimicrobials, suggestions for prospective inhibitors were made based on GRID analysis.

The presence of a phosphate binding pocket in DHQS prior to the conformational change of the enzyme indicated that the carbaphosphonate molecule may initially dock into the phosphate binding pocket and then initiate the structural change by attracting the carboxy binding pocket (Figure 3.34 and Figure 3.38). The residue interactions proposed by Carpenter et $al.$[6] for the catalytic mechanism of DHQS were corroborated by GRID analysis. A plausible inhibitor, based on the strong interactions formed between Zn^{2+} and S^-, as well as the location of the carboxy binding pocket, was proposed (Figure 3.36).

EPSPS showed energetically favorable binding interactions toward GRID probes in the open enzyme. It was concluded that the conformational change occurs when the phosphate moiety of the shikimate 3-phosphate molecule binds to the open enzyme and attracts the carboxy binding pocket located in the opposite domain (Figure 3.38 and Figure 3.42). Based on GRID analysis, a potential inhibitor was proposed which utilized the presence of the strong phosphate binding pockets and the proximity of the shikimate 3-phosphate and glyphosate binding sites (Figure 3.41).

The future direction of this project should include mutations of the phosphate binding pockets in DHQS and EPSPS to determine whether their absence in the open structure will impact the ability of the enzyme to close in the presence of substrate molecules, as predicted by GRID analysis. Additionally, crystallization of DHQS in the presence of excess dications has the potential to provide structural information regarding the location of an allosteric binding site. GRID analysis provided information regarding the design of potential DHQS and EPSPS inhibitors. These inhibitors should be synthesized and their potential inhibitory effects evaluated in the respective enzyme.

EXPERIMENTAL SECTION

GRID Calculations

All calculations were performed using the Linux operating system and the GREAT, GRIN, and GRID software package[38] version 21. Structure files were obtained from the PDB,[41] with the exception of the EPSPS and DHQS open structures which were provided by Dr. E. P. Carpenter. All remarks and water molecules were removed from the protein database file before processing. PDB files consist of two atom-type records, HETATM, which is used to define the properties of any molecule such as solvent molecules, ions, inhibitors, etc., while ATOM records are normally restricted to amino-acids, peptides, proteins, nucleic acids, other biological macromolecules such as sugars. GRID uses a similar format, therefore; all substrates, inhibitors, and metal ions were identified as heteroatoms by changing the ATOM record to HETATM before beginning the calculations.

Proteins of similar nature, *e.g.* DHQS open and DHQS closed, were superpositioned by XtalView software[64] version 3.2.1 to according to their C^{α} position to place them in a similar orientation for future comparison. Protein residues missing from the protein database file were modeled in idealized positions according to the location of the side chain using XtalView software and

refined using CNSsolve,[65] Crystallography and NMR System software version 1.1 at one refinement cycle of two hundred minimization steps.

The files were prepared for use in GRID using program GRIN. The charge of the protein was calculated in GRIN and counterions of Na^+ or Cl^- were added appropriately to create a charge balanced structure. To determine the best placement positions for the counterions, a GRID calculation was performed using the Na^+ or Cl^- probe to locate the most energetically favorable binding site for the counterions. The counterions were then manually entered in suggested positions away from the active site and set to move in response to the probe. As the probe moves around the GRID and into the vicinity of the counterion, the counterion reacts and shifts slightly to avoid being "seen" by the probe. Moving counterions prevent a large charged area from being present during the experiment.

All GRID calculations were performed on the charge balanced protein coordinate files with the following directives: NETA (Number of Extra Target Atoms) equal to the number of HETATM records setting the small molecules to be included as part of the target, NPLA (Number of Planes) equal to 2 thus the GRID points were 0.5 Å apart to improve the accuracy of the contour surfaces, and MOVE equal to -1 to allow the counterions to be flexible with respect to the probe.

Probes used in the course of the experiment included both single-atom multi-atom types. The hydroxyl probe used was modeled to locate energetically favorable areas for phenol or carboxy OH binding. Three closely related probes were the carbonyl, carboxy and multi-atom carboxy. The carbonyl probe was designed to locate regions of favorable binding energy toward sp^2 carbonyl oxygens, while the carboxy probe was modeled as an sp^2 carboxy oxygen atom. The multi-atom carboxy probe is a combination of both the single-atom carboxy and carbonyl probe, which more accurately reflects the binding affinity of a carboxy group. The zinc probe used was modeled as the zinc dication. The hydrogen phosphate monoanion, HPO_4^-, and the phosphate dianion, PO_4^{2-}, probes were investigated but yielded different results only in the presence of a hydrogen-accepting region.

Visualization of Data

The output data generated by GRID were read into MSI Insight II (Accelrys, Cambridge, UK) software version 2000 for visualization of the energy maps. The data map files were output from Insight II (Accelrys, Cambridge, UK) as ascii files and reopened in DS Viewer Pro 5.0 (Accelrys, Cambridge, UK) for generation of the graphics. Alwyn Jones' O software[66] was used to visualize the three-dimensional structure and determine nearest neighbor residues in the protein.

REFERENCES

1. Hawkins, A. R.; Lamb, H. K.; Moore, J. D.; Charles, I. G.; Roberts, C. F., *J. Gen. Microbiol.* **1993,** *139* 2891.
2. Pittard, A. J., *Biosynthesis of the Aromatic Amino Acids*, in *In Eschericia Coli and Salmonella Typhimurium: Cellular and Molecular Biology*, F. C. Niedhardt, Editor. 1996, American Society for Microbiology: Washington, DC. p. 458.
3. Herrmann, K. M.; Weaver, L. M., *Annu. Rev. Plant Mol. Biol.* **1999,** *50* 473.
4. Steinrücken, H. C.; Amrhein, N., *Biochem. Biophys. Res. Comm.* **1980,** *94* 1207.
5. Roberts, F.; Roberts, C. W.; Johnson, J. J.; Kyle, D. E.; Krell, T.; Coggins, J. R.; Coombs, G. H.; Milhous, W. K.; Tzipori, S.; Ferguson, D. J. P.; Chakrabarti, D.; McLeod, R., *Nature* **1998,** *393* 801.
6. Carpenter, E. P.; Hawkins, A. R.; Frost, J. W.; Brown, K. A., *Nature* **1998,** *394* 299.
7. Nichols, C. E.; Ren, J.; Lamb, H. K.; Hawkins, A. R.; Stammers, D. K., *J. Mol. Biol.* **2003,** *327* 129.
8. Schonbrunn, E.; Eschenburg, S.; Shuttleworth, W. A.; Schloss, J. V.; Amrhein, N.; Evans, J. N. S.; Kabsch, N., *Proc. Natl. Acad. Sci.* **2001,** *98* 1376.
9. Carpenter, E. P., **in preparation,**
10. Gallay, O.; Carpenter, E. P., *unpublished work*
11. McCandliss, R. J.; Herrmann, K. M., *Proc. Natl. Acad. Sci.* **1978,** *75* 4810.
12. Baasov, T.; Knowles, J. R., *J. Bacteriol.* **1989,** *171* 6155.
13. Stephens, C. M.; Bauerle, R., *J. Biol. Chem.* **1991,** *266* 20810.
14. Frost, J. W.; Bender, J. L.; Kadonaga, J. T.; Knowles, J. R., *Biochemistry* **1984,** *23* 4470.
15. Srinivasan, P. R.; Rothchild, J.; Sprinson, D. B., *J. Biol. Chem.* **1963,** *238* 3176.
16. Bender, S. L.; Widlanski, T.; Knowles, J. R., *Biochemistry* **1989,** *28* 7560.
17. Widlanski, T.; Bender, S. L.; Knowles, J. R., *Biochemistry* **1989,** *28* 7572.
18. Bartlett, P. A.; McLaren, K. L.; Marx, M. a., *J. Org. Chem.* **1994,** *59* 2082.
19. Chaudhuri, S.; Duncan, K.; Graham, L. D.; Coggins, J. R., *Biochem. J.* **1991,** *275* 1.
20. Anton, P. R.; Coggins, J. R., *Biochem. J.* **1988,** *249* 319.
21. Elsemore, D. A.; Ornston, L. N., *J. Bacteriol.* **1994,** *176* 7659.
22. McDowell, L. M.; Klug, C. A.; Beusen, D. D.; Schaefer, J., *Biochemistry* **1996,** *35* 5395.

23. Shuttleworth, W. A.; Evans, J. N., *Arch. Biochem. Biophys.* **1996,** *334* 37.
24. Shuttleworth, W. A.; Evans, J. N., *Biochemistry* **1994,** *33* 7062.
25. Padgette, S. R.; Re, D. B.; Gasser, C. S.; Eichholtz, D. A.; Frazier, R. B., *J. Biol. Chem.* **1991,** *266* 22364.
26. Majumder, K.; Selvapandiyan, A.; Fattah, F. A.; Arora, N.; Ahmad, S.; Bhatnagar, R. K., *Eur. J. Biochem.* **1995,** *229* 99.
27. Hawkes, T. R.; Lewis, T.; Coggins, J. R.; Mousedale, D. M.; Lowe, D. J.; Thorneley, R. N. F., *Biochem. J.* **1990,** *265* 899.
28. Balasubramanian, S.; Abell, C.; Coggins, J. R., *J. Am. Chem. Soc.* **1990,** *112* 8581.
29. Brown, K. A.; Carpenter, E. P.; Watson, K. A.; Coggins, J. R.; Hawkins, A. R.; Koch, M. H. J.; Svergun, D. I., *Biochem. Soc. T.* **2003,** *June 31* 543.
30. White, S. A.; Peake, S. J.; McSweeny, S.; Leonard, G.; Cotton, N. P. J.; Jackson, J. B., *Structure* **2000,** *8* 1.
31. Kraulis, J., *J. Appl. Crystallogr.* **1991,** *24* 946.
32. Merritt, E. A.; Murphy, M. E. P., *Acta Crystallogr.* **1994,** *D50* 869.
33. Stallings, W. C.; Abdel-Meguid, S.; Lim, L. W.; Huey-Sheng, S.; Dayringer, H. E.; Leimgruber, N. K.; Stegeman, R. A.; Anderson, K. S.; Sikorski, J. A.; Padgette, S. R.; Kishore, G. M., *Proc. Natl. Acad. Sci.* **1991,** *88* 5046.
34. Steinrücken, H. C.; Amrhein, N., *Eur. J. Biochem.* **1984,** *143* 351.
35. Levin, J. G.; Sprinson, D. B., *J. Bio. Chem.* **1964,** *239* 1142.
36. Bondinell, W. E.; Vnek, J.; Knowles, P. F.; Sprecher, M.; Sprinson, D. B., *J. Biol. Chem.* **1971,** *246* 6191.
37. Anton, D. L.; Hedstrom, L.; Fish, S. M.; Abeles, R. H., *Biochemistry* **1983,** *22* 5903.
38. Goodford, P. J., *J. Med. Chem.* **1985,** *28* 849.
39. Buckingham, R. A., *Proc. R. Soc. London, A* **1924,** *168* 264.
40. Jones, J. E., *Proc. R. Soc. London, A* **1924,** *106* 463.
41. Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Bryce, M. D.; Rodgers, J. R.; Kennard, O.; Shikanouchi, T.; Tasumi, M., *J. Mol. Biol.* **1977,** *112* 535.
42. MacKerellJr., A. D.; Brooks, B.; BrooksIII, C. L.; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M., *Charmm: The Energy Function and Its Parameterization with an Overview of the Program*, in *The Encyclopedia of Computational Chemistry*, P. v. R. S. e. *al.*, Editor. 1998, John Wiley & Sons: Chichester. p. 271.
43. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M., *Journal of Computational Chemistry* **1983,** *4* 187.

44. Hopfinger, A. J., *Conformational Properties of Macromolecules*. 1973, New York: Academic Press.

45. Warwicker, J.; Watson, H. C., *J. Mol. Biol.* **1982,** *157* 671.

46. Carpenter, E. P., *Atomic Coordinates of the Three-Dimensional Crystal Structure of Substrate Bound Epsps*, J. N. Jones, Editor. 2003: London.

47. Boocock, M. R.; Coggins, J. R., *FEBS Lett.* **1983,** *154* 127.

48. Bartlett, P. A.; Satake, K., *J. Am. Chem. Soc.* **1988,** *110* 1628.

49. Lambert, J. M.; Boocock, M. R.; Coggins, J. R., *Biochem. J.* **1985,** *226* 817.

50. Montchamp, J.-L.; Frost, J. W., *J. Am. Chem. Soc.* **1997,** *119* 7645.

51. Herrmann, K. M., *Plant Physiol.* **1995,** *107* 7.

52. Cushman, D. W.; Cheung, H. S.; Sabo, E. F.; Ondetti, M. A., *Biochemistry* **1977,** *16* 5484.

53. Moore, J. D.; Skinner, M. A.; Swatman, D. R.; Hawkins, A. R.; Brown, K. A., *J. Am. Chem. Soc.* **1998,** *120* 7105.

54. Bender, S. L.; Mehdi, S.; Knowles, J. R., *Biochemistry* **1989,** *28* 7555.

55. Shuttleworth, W. A.; Pohl, M. E.; Helms, G. L.; Jakeman, D. L.; Evans, J. N., *Biochemistry* **1999,** *38* 296.

56. Saunders, J. E., *Structure Function Studies of Epsp Synthases from Pathogenic Organisms*, in *Department of Biological Sciences*. 2003, Imperial College of Science, Technology and Medicine: London. p. 295.

57. Sost, D.; Amrhein, N., *Arch. Biochem. Biophys.* **1990,** *282* 433.

58. Madura, J. D.; Jorgensen, W. L., *J. Am. Chem. Soc.* **1986,** *108* 2517.

59. Bala, P.; Grochowski, P.; Lesyng, B.; McCammon, J. A., *J. Phys. Chem.* **1996,** *100* 2535.

60. Marzabadi, M. R.; Gruys, K. J.; Pansegrau, P. D.; Walker, M. C.; Yuen, H. K.; Sikorski, J. A., *Biochemistry* **1996,** *35* 4199.

61. Knowles, G., *Journal of Antimicrobial Chemotherapy* **1993,** *32* 172.

62. Miller, M. J.; Ream, J. E.; Walker, M. C.; Sikorski, J. A., *Bioorgan. Med. Chem.* **1994,** *2* 331.

63. Miller, M. J.; Cleary, D. G.; Ream, J. E.; Snyder, K. R.; Sikorski, J. A., *Bioorgan. Med. Chem.* **1995,** *3* 1685.

64. McRee, D. E., *J. Struct. Biol.* **1999,** *125* 156.

65. Brünger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L., *Acta Crystallogr., D* **1998,** *D54* 905.

66. Jones, T. A.; Kjeldgaard, M., *O*. 2000.

CHAPTER 4: STRUCTURAL STUDIES OF MYCOBACTERIUM TUBERCULOSIS CATALASE-PEROXIDASE

INTRODUCTION

In 2001, the World Health Organization reported that 1 in 3 people are infected with latent *Mycobacterium tuberculosis* worldwide. There are 8 million new cases of active tuberculosis (TB) annually and 5000 individuals will die from an infectious pulmonary disease per day.[1] Tuberculosis is a global problem, which requires further research to completely eradicate the disease.

Streptomycin entered the market in 1946 as the first agent used to treat TB.[2] Today, a combination therapy of four drugs: isonicotinic acid hydrazide (isoniazid, INH), rifampicin, pyrazinamide, and ethambutol/streptomycin[3] are used to treat tuberculosis. However, there is a growing resistance to INH and rifampicin among the TB strains. Some drug-resistant strains have been found to be resistant to 7 drugs.[4] The development of new chemotherapeutic agents is hampered by a lack of knowledge regarding the pathogenesis of *M. tuberculosis*.

Catalase-peroxidase (CP), encoded by the *katG* gene, has been shown to restore INH-sensitivity in resistant *Mycobacterium smegmatis* and *M. tuberculosis* strains[5] indicating a clear functional link between the *katG* gene and INH

216

activity.[6] The exact role of CP in INH-sensitivity is unknown despite several attempts at uncovering its function. Investigating the chromosomal DNA from a group of eight INH-resistant strains of *M. tuberculosis* revealed that two-thirds of the highly resistant strains completely lacked the *katG* gene. Cole *et al.*[7] thereby suggested gene deletion as the mechanism behind INH-resistance. In many cases, rather than being deleted from the chronosome, the *katG* gene was found to contain point mutations, deletions, or insertions of a few bases.[8] Still other findings suggest that the *katG* gene is present in 78% of INH-resistant isolates.[9]

M. tuberculosis CP is a dual-function enzyme, which is both catalase and peroxidase active. The catalatic activity is comparable to that of monofunctional catalases, such as Horseradish peroxidase isozyme C (HRPC). *M. tuberculosis* CP shows a broad spectrum of peroxidatic activity and can oxidize a wide range of one electron donors.

Catalase Activity

The CPs show no significant homology to known catalase sequences and it seems that this dissimilarity may imply different active site architecture compared to conventional heme catalases. However, inhibition studies with potassium cyanide, hydroxylamine, and azide have indicated the catalatic activity is mediated at the heme moiety.[10] Mutation of one residue in the putative active site of CP can reduce both catalase and peroxidase functionality.[11, 12] Therefore,

it has been suggested that the catalase activity proceeds *via* reduction of Compound I with H_2O_2 (Figure 4.1). However, addition of H_2O_2 to CP Compound I using excess peroxoacetic acid did not lead to ferric CP; electron donation from H_2O_2 appeared to be blocked. The peroxoacetic acid may sterically impede access to the two electron donor site, indicating that peroxoacetic acid is still bound or catalase activity may not proceed *via* Compound I.[13]
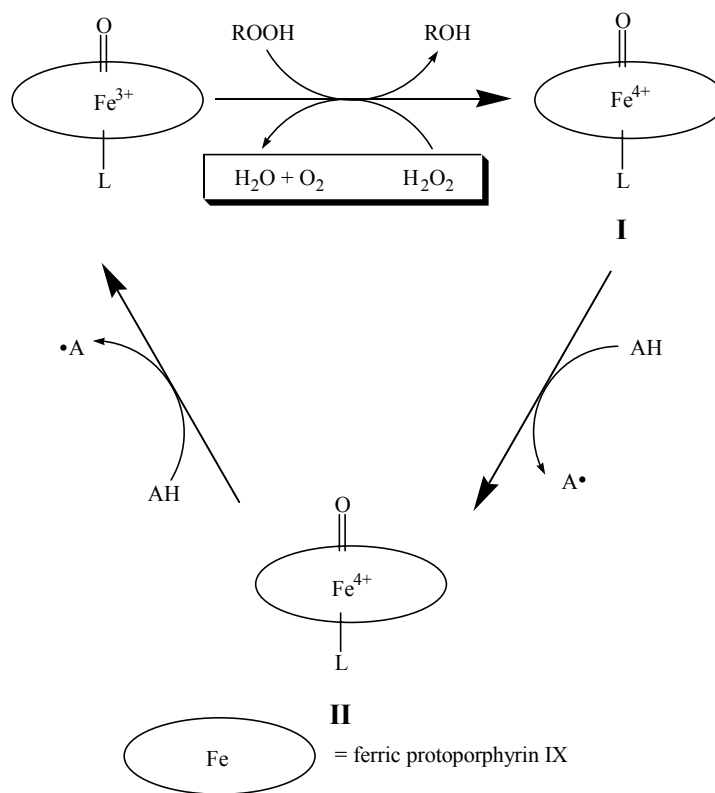
Figure 4.1 Schematic of dual peroxidase and catalase mechanism. Catalase mechanism is highlighted in the shadow box.

218

Peroxidase Activity

Catalase-peroxidases can oxidize a wide range of one-electron donors. The oxidation proceeds in an enzyme like HRPC *via* the peroxidase cycle,[14, 15] which begins with the resting state ferric enzyme reacting with one equivalent of H_2O_2 to give Compound I (Figure 4.2). Compound I is an oxyferryl iron-protoporphyrin IX π-cation radical which is reduced by a single electron from the substrate to yield Compound II. The final step is the reduction of Compound II to the resting state by another single electron typically from a second substrate molecule (Figure 4.2).[16-18] In the case of yeast cytochrome *c* peroxidase (CcP), the cation radical is not localized on the protoporphyrin ring but is stabilized on a tryptophan residue adjacent to the heme.[19-21]

$$Fe(III)R \quad + \quad H_2O_2 \quad \longrightarrow \quad \underset{\text{compound I}}{Fe(IV)OR\bullet} \quad + \quad H_2O$$

$$\underset{\text{compound I}}{Fe(IV)OR\bullet} \quad + \quad substrate \quad \longrightarrow \quad \underset{\text{compound II}}{Fe(IV)OR} \quad + \quad oxidized\ substrate$$

$$\underset{\text{compound II}}{Fe(IV)OR} \quad + \quad substrate \quad \longrightarrow \quad Fe(III)R \quad + \quad H_2O \quad + \quad oxidized\ substrate$$

Figure 4.2 Peroxidase mechanism.

A mechanism for the reduction of the pseudo-stable CP Compound I intermediate has been proposed by Regelsberger *et al.*[13] where both the heme protoporphyrin radical and the Fe(IV) species are reduced simultaneously by two

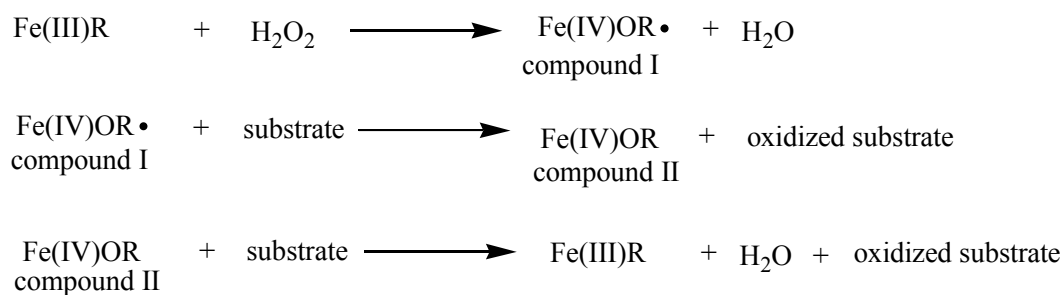electrons. One electron originates from a substrate molecule and one from an amino acid residue near the heme moiety. The neighboring residue could potentially stabilize the radical, similar to W191 in C*c*P Compound I.[19, 22] The oxidized residue is then reduced by the donation of a second electron from another substrate molecule (Figure 4.2).

Until recently, the *M. tuberculosis* CP radical was assigned to the protoporphyrin π-cation radical of Compound I;[23] however, evidence for the presence of a tyrosyl radical in *M. tuberculosis* CP has been presented.[24] Using rapid freeze-quench electron paramagnetic resonance Chouchane *et al.*[24] demonstrated the formation of tyrosyl radical in high yield from the resting enzyme treated with alkyl peroxide.

In addition to showing catalase and peroxidase activity, *M. tuberculosis* CP exhibits activity toward numerous other systems. This flexible enzyme behaves as a haloperoxidase capable of oxidizing chloride, bromide, and iodide in a peroxide- and enzyme-dependent manner.[25] *M. tuberculosis* CP demonstrates effective removal of ONOOH at a rate comparable to that exhibited by other peroxynitrite decomposition catalysts.[26] P450-like mono-oxygenase activity is yet another property exhibited by the versatile CP.[27] Wengenack *et al.* reported that a catalytic quantity of superoxide dismutase (SOD) abolished INH turnover, while addition of a superoxide-generating system to *M. tuberculosis* CP plus INH

increases the oxidation rate of INH 5-fold.[26] Based on these results, it was concluded that *M. tuberculosis* CP exhibits a superoxide-dependent pathway for INH oxidation. However, the SOD-INH oxidation pathway has been criticized due to the presence of a SOD in *M. tuberculosis*, which has the potential to break down any superoxide present prior to its utilization by CP to oxidize INH. Magliozzo *et al.*[28] have reported that *M. tuberculosis* CP demonstrates Mn(II) peroxidase activity; however, this hypothesis has been disputed based on evidence by Bondiguel *et al.*[29] that Mn^{2+} has the ability to catalyze INH oxidation in the absence of enzyme.

Soon after its introduction into clinical use, bacterial resistance to INH was reported.[6, 30] Analysis of *M. tuberculosis* indicated that INH was metabolized to a number of products with 4-pyridylmethanol and isonicotinic acid predominating (Figure 4.3).[31, 32] However, the metabolized products were not observed in INH-resistant organisms.[6] Many INH-resistant isolates lack catalase and peroxidase activity.[30, 33] After Zhang *et al.*[34] restored INH-sensitivity to an INH-resistant catalase-negative strain by introducing the *katG* gene, it is widely accepted that CP oxidizes INH[35, 36] and the *M. tuberculosis katG* gene confers INH sensitivity *in vivo*.[5] However, the activated form of INH and the mechanism by which INH is metabolized remain highly disputed.
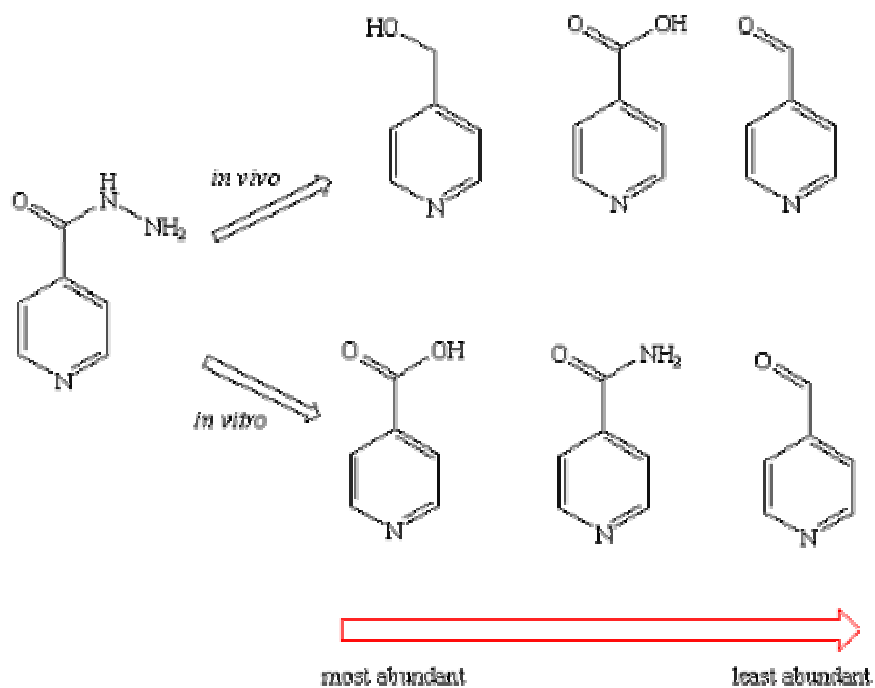
Figure 4.3 **Product profiles for INH oxidation by whole cells of *M. tuberculosis* (*in vivo*) and by purified CP (*in vitro*).**

Mechanistic analysis of INH oxidation has yielded varied results between *in vivo* and *in vitro* studies. Isonicotinic acid, isonicotinamide, and pyridine-4-carboxyaldehyde are the three oxidation products formed during *in vitro* analysis (Figure 4.3).[28, 29, 35, 37] Isonicotic acid and pyridine-4-carboxyaldehyde, as well as isonicotinol, which has not been seen *in vitro*, were determined to be oxidation products of INH *in vivo* (Figure 4.3).[38]

In vivo, the predominant oxidation product is the alcohol, which is not detected *in vitro*. Conversely, the isonicotinamide derivative has not been observed *in vivo*.[38] Pyridine-4-carboxaldehyde may play an important role *in*

vivo since it is rapidly reduced to the alcohol when fed exogenously to whole cells of *M. tuberculosis*.[6] The differences observed *in vitro* versus *in vivo* are most likely due to the difference that arises when measuring the systems. *In vitro* the ability to detect less abundant products exists due to the limited system; where as, *in vivo* the plethora of components may interfere with observing reaction products and/or intermediates.

Using ^{18}O NMR, Johnsson and Schultz determined that there are two simultaneous metabolic pathways operating in INH metabolism to form isonicotinic acid.[35] An acyl radical and an acyldiazonium ion were the postulated intermediates in the pathway.[35] ^{15}N NMR studies led to the development of a mechanism which involves splitting of the N—N bond in INH.[35] Additional HPLC studies involving ^{15}N-labelled INH have led to the proposal of a C—N bond splitting mechanism.[29] Formation of an acyl radical and diazene is followed by its transformation into hydrazine, which in turn becomes ammonia.[29] The initial reduction of INH to a hydrazyl radical by one electron has been determined from square wave and cyclic voltammograms.[39]

Peroxidatic activity of CPs is believed to mediate INH activation. CP-dependent INH-oxidation requires O_2 *in vitro*[28, 40] and *in vivo*.[6, 41] Under aerobic conditions, H_2O_2 is a requirement *in vitro*.[28, 29] Additionally, the observation that INH can be oxidized by HRPC has contributed to the hypothesis

of peroxidatic activity as the activation mediator for INH.[23, 42] However, INH is oxidized both in the presence and absence of peroxide[28, 35, 43] raising questions regarding oxidation *via* a Compound I/II pathway. Although the exact pathway is unknown, cyclic and square wave voltammetry show that the ferric CP midpoint potientials are too small to promote INH oxidation without formation of a high valent intermediate such as Compound I and II or oxyferrous CP.[39]

Two alternative mechanisms to the peroxidatic pathway for INH activation are a cytochrome P450-like pathway[28] and a superoxide-dependent pathway.[26] The P450-like pathway involves reduction of the heme iron followed by binding of O_2 to produce the oxyferrous form of the enzyme, which are analogous to Compound III (Figure 4.1) in peroxidases. The oxyferrous CP then oxidizes INH, which returns it to the ferric resting state (Figure 4.4).[26] Based on generation of superoxide during INH oxidation by HRPC, myeloperoxidase and extracts of *M. tuberculosis* H37Ra (a strain with attenuated virulence),[44-46] the superoxide-dependent pathway for INH oxidation was formulated. Trace metals can autooxidize hydrazines by O_2 generating superoxide.[44, 47, 48] Oxyferrous CP is formed by binding of superoxide to the resting state ferric enzyme, which is followed by subsequent oxidation of INH and CP returning to the resting state (Figure 4.5).
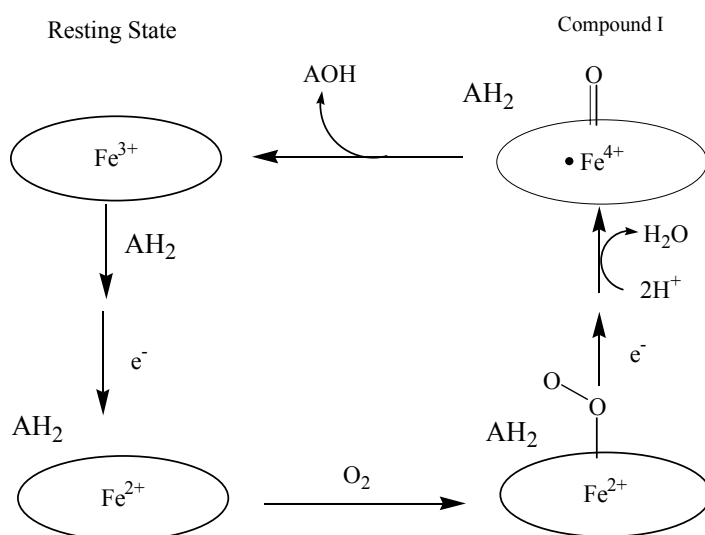
Figure 4.4 Proposed cytochrome P450-like pathway for iron activation and oxidation of INH. AH_2 and AOH represent a substrate molecule and an oxidized substrate molecule, respectively.
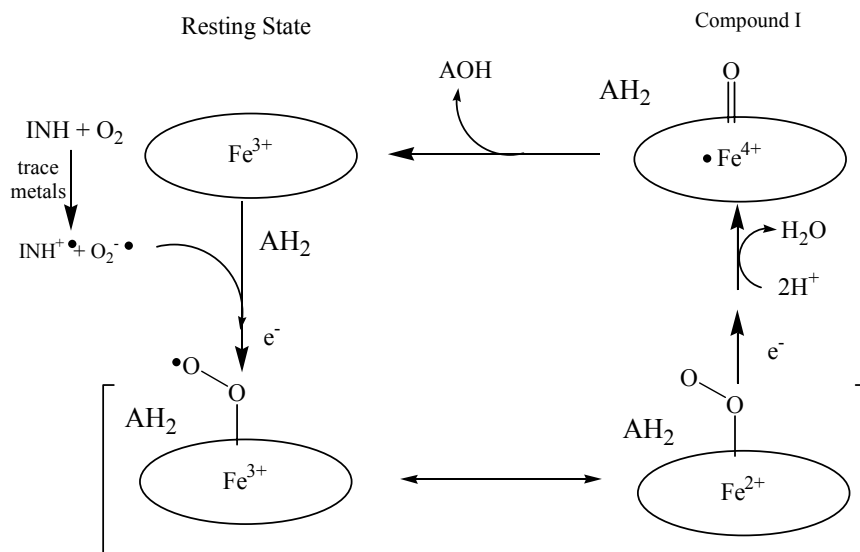

Figure 4.5 Proposed superoxide-dependent pathway for iron activation and oxidation of INH. AH_2 and AOH represent a substrate molecule and an oxidized substrate molecule, respectively.

The presence and importance of the iron protoporphyrin IX moiety in INH oxidation is well documented; however, other metal ions have been reported to play a role in the catalytically competent form of CP. Both Cu^{2+}[48] and Mn^{2+}[28] have been found to stimulate *in vitro* INH oxidation. However, Mn^{2+} has been shown to oxidize INH in the absence of CP[29, 48] spawning much debate regarding the role, if any, of Mn^{2+} in INH oxidation. The proposed Mn^{2+} binding site, determined by sequence alignment of *M. tuberculosis* CP with fungal manganese peroxide, is shown in Figure 4.6.

```
172   SHTVAR          ADKV   181   MnP
169   GHTFGKTHGAGPADLV  284   mtCP
```

Figure 4.6   The key aspartate residue in manganese peroxidase (MnP) is found in M. tuberculosis CP and is predicted to be at the start of a loop (bold) as in MnP.   The positively charged arginine residue of MnP is conserved in *M. tuberculosis* as lysine (underlined).

NMR measurements have proven to be a valuable tool in determining the location and distances of INH-binding in *M. tuberculosis*. Using ^{15}N-labelled INH and longitudinal relaxation measurements, the iron-nitrogen distance from the heme to the hydrazide moiety of INH was determined to be 3.8 Å for wild-type *M. tuberculosis*.[49]   Iron-hydrogen distances from the heme to INH were found to be approximately 12 Å,[50] much longer than the iron-nitrogen distances. Two-dimensional nuclear overhauser effect spectroscopy (NOESY), corroborated

226

these findings.[51] The presented data suggests that the INH hydrazide moiety points towards the heme iron atom.

Based on $^{14}N^{15}N$ and $^{15}N^{14}N$ NMR for HRPC[51] and *M. tuberculosis*,[29] as well as INH incubation studies of the two enzymes,[29, 35, 37] Horseradish Peroxidase C (HRPC) has been suggested as a model compound for the mechanism and activation of INH in *M. tuberculosis*.[51] In an NMR-based three dimensional model of INH bound to HRPC, Banci *et al.*[51] described the residues involved in binding INH. The pyridine moiety of INH is located in a hydrophobic "aromatic binding pocket",[52] near the heme-18 methyl and makes non-bonded interactions with residues G69, A140, P141, and F179, as well as the heme itself.[51] The hydrazide moiety forms hydrogen bond to the backbone carbonyl oxygen of P139 at a distance of 2.7-2.8 Å and to a nitrogen atom of the distal histidine residue, H42, at a distance of 3.0 Å. Considering the similarities in HRPC and *M. tuberculosis* CP, the model of INH bound to HRPC provides a structural framework to further understand the molecular interactions controlling INH binding and turnover.[51]

Structural information about *M. tuberculosis* can be derived from HRPC,[51] a peroxidase belonging to the Class III superfamily of plant, fungal, and bacterial peroxidase.[53] HRPC has similar active site architecture at the N-terminal domain of *M. tuberculosis*.[54] The availability of a three dimensional

crystal structure of HRPC has clarified the residues located in the distal and proximal pockets of the heme molecule, identified the location of Ca^{2+} ions and indicated the presence of a substrate access channel. The 306 residue structure of HRPC is composed primarily of α-helices and the structure is similar to other crystallographically characterized peroxidases (Figure 4.7a).[55] The distal pocket of the heme moiety is composed of key catalytic residues R38, F41, and H42. Residues H170, Q176, S73, S35, and R31 form the proximal heme pocket. Two Ca^{2+} ions are present in the structure. The distal Ca^{2+} site is seven-coordinate with carbonyl and side chain oxygen atoms as ligands, while the proximal Ca^{2+} is located near T171. Interestingly, Ca^{2+} is a requirement for the formation of the structural environment of the heme during *in vitro* folding of the denatured apoprotein.[56] F68, F142, and F179 form a hydrophobic patch near the exposed heme edge, which is the suggested site of the pre-electron transfer complex formed between substrate and high oxidation state intermediates.[54]

Benzhydroxamic acid, a compound structurally similar to INH has been crystallographically characterized bound to HRPC.[52] INH is predicted to occupy the same region of the active site in HRPC and have a similar orientation to that seen in the crystal structure of the HRPC-BHA complex.[51] However, spectroscopic studies have indicated that HRPC has a higher affinity for BHA then INH.[57, 58] It has been suggested that HRPC's reduced affinity for INH is

associated with unfavorable desolvation of the pyrindinyl group predicted to remain non-hydrogen-bonded in the complex.[51, 58]

Cytochrome *c* peroxidase (C*c*P), a class I peroxidase, has been structurally characterized to contain ten helical segments with two clearly defined domains.[59] The heme moiety is located between the domains with only one of the pyrrol ring edges exposed to the external environment. The junction between the domains is lined with aliphatic and aromatic side chains. The heme iron atom is coordinated to a tryptophan, arginine, and histidine on the distal side and a histidine side chain on the proximal side of the heme. The tryptophan residue is thought to be the site of free radical formation in C*c*P comound I.[19-21] A threonine side chain forms a hydrogen bond with one of the heme propionates and a water molecule serves as the sixth heme ligand.[59] The heme iron atom is displaced 0.3 to 0.5 Å below the plane of the heme toward the proximal histidine. The most exposed surface occurs near the guanidinium group of the distal arginine residue and is the most direct route for ligands entering the pocket.[59]

The X-ray crystal structure of another class I peroxidase, ascorbate peroxidase (APX), has been reported.[60] The overall fold of APX is nearly identical to that determined for C*c*P. The active site structure is the same, including the hydrogen bonding interactions between the promixal histidine ligand, a buried aspartic acid residue and a tryptophane residue. Unlike C*c*P, a

cation binding site is present approximately 8 Å from the α-carbon atom of the tryptophan residue in APX.[60] The cation in APX was hyphothesized to be one factor that makes it more difficult to oxidize W179 than the corresponding W191 in C*c*P, allowing for the formation of a prophyrin π-cation radical in compound I rather than a tryptophan radical.[60]

Ascorbate, the natural substrate of APX, was successfully crystallized bound to the enzyme. The overall structure of the enzyme and active site were unaffected by ascorbate binding. Previously, the δ-heme edge has been implicated as the binding site for aromatic substrates[61, 62] and ascorbate;[62, 63] however, ascorbate was crystallographically determined to bind at the γ-*meso* position of the heme.[64] The basic residues, R172, and K30, which are involved in binding of ascorbate, are not present in C*c*P.[64] This difference accounts for the low activity of C*c*P towards ascorbate.[65] A potassium cation was not observed near W179 in the APX-ascorbate structure; therefore, a more concise explanation was presented for formation of the prophyrin π-cation radical rather than a tryptophan radical. The direct coupling of the substrate to the heme 6-propionate bypasses the need for involvement of W179 in ascorbate oxidation and compound I is reduced through a porphyrin π-cation intermediate.[64]

The crystal structures of two CP homologues have been reported, *Haloarcula marismortui*[66] *and Burkholderia pseudomallei*[67] CP, which share 56

% identity / 70 % similarity and 65 % identity / 77 % similarity, respectively, with *M. tuberculosis* CP. Structural information from these enzymes, *H. marismortui* CP and *B. pseudomallei* CP, has provided valuable insight into the architecture of *M. tuberculosis* CP by molecular modeling techniques.

H. marismortui CP is a dimer composed of two identical subunits related by non-crystallographic two-fold symmetry (Figure 4.7b).[66] Two structurally similar domains, the N-terminal domain and C-terminal domain, constitute each subunit. The arrangements of secondary structure in the two domains are identical to that of C*c*P and APX.[66]

The heme active site is buried inside the N-terminal domain and is similar in structure to the active site of C*c*P and APX.[66] The proximal site residues are H259, D372, and W311, while the distal residues are R95, W92, and H96. Unlike APX, a potassium cation is not observed near the proximal tryptophan residue in *H. marismortui* CP. The electronic environment around the proximal side of the heme most closely resembles that of C*c*P, which may support formation of a tryptohan radical rather than a prophyrin π-cation radical as formed by APX.[66]

Similar to HRPC, *H. marismortui* CP has a substrate access channel. Around the heme iron atom, the cylindrical channel is approximately 3.5 Å in diameter at S305 and the conical channel increases in diameter as it approaches the surface. Surface residues G194 and G271 mark the acidic entrance to the

access channel. The size of the substrate access channel prevents large substrates from accessing the heme.[66]

Novel covalent modifications were observed near the active site of *H. marismortui* CP between Cε1 of Y218 and Cη2 of W95, as well as Cε2 of Y218 and Sδ of M244 [66]. The conservation of W95, Y218, M44, and R409 suggests that the covalent bonds may be common among all CPs.[66] Mutation of W95 causes loss of catalase activity,[12, 68] indicating that the covalent bonds may be important for the catalase activity in CPs.

Carpena *et al.*[67] reported the structure of *B. pseudomallei* in 2003 (Figure 4.7c). Similar to *H. marismortui* CP, *B. pseudomallei* CP is a dimer with one heme group per subunit, which is located in the N-terminal domain.   One unidentified metal ion (mostly likely sodium based on its refinement properties) per subunit is coordinated to G122, G124, and S494, as well as two water molecules which interact with the side chains of G128, G198, and D427. The position of this unidentified metal ion is about 7 Å from the position of Ca^{2+} in HRPC.[54, 67]

The distal side of the heme active site is composed of residues R108, W111, and H112, while the proximal site residues are H279 in close association with D389 and W330. The access route to the distal side of the heme active is *via* a funnel shaped channel which is approximately 14 Å from the heme iron atom.

Unlike *H. marismortui* CP, there is an addition of at least two atoms to the vinyl group on ring I of the heme moiety.[67] Carpena *et al.*[67] hypothesized that hydrogen peroxide is the likely source of the modification due to its accessibility to the active site.

The covalent bonding network first identified in *H. marismortui* CP, was also observed in *B. pseudomallei* CP.[67] The presence of the W-Y-M adduct in both *H. marismortui* CP and *B. pseudomallei* CP suggests that it may be a general feature of CPs. The purpose of the modification is unclear; however, one possibility is they stabilize the enzyme by reducing the likelihood of other more damaging oxidations from occurring.[67]

Interestingly, the structure of *B. pseudomallei* CP contained an unidentified mass of electron density prior to the constricted region of the access channel. Carpena *et al.*[67] proposed that the site is potentially the active site for INH activation (Figure 4.7d). Proposed interactions are residues W323 and S324 for the hydrazide functional group and water coordinated to W309. Building on the hypothesis of this being the site for INH activation, a means of electron transfer was devised, which involves the transfer of electrons from the main-chain carbonyl group of A265 to M264.[67]
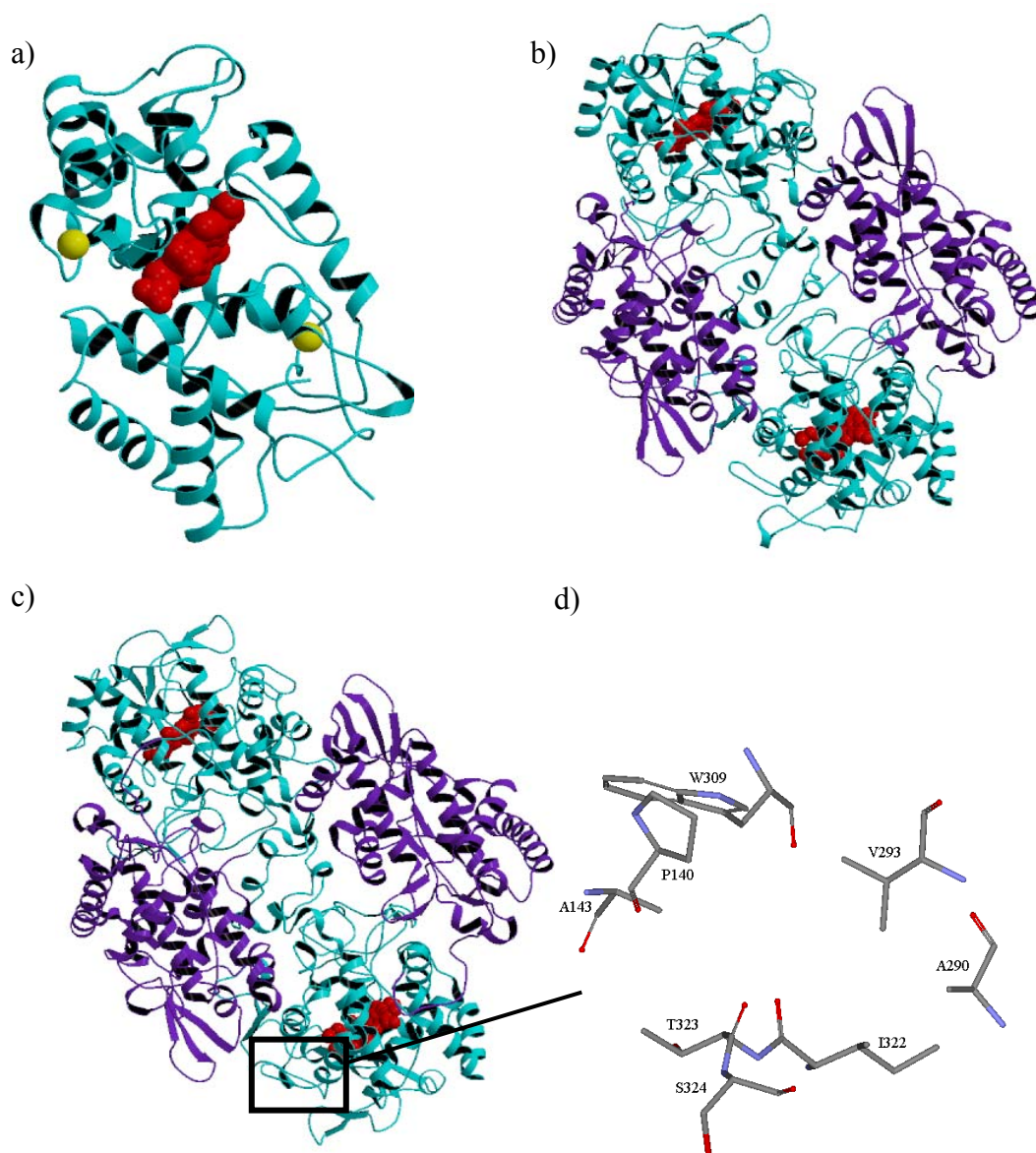
Figure 4.7   Structure of a) HRPC, b)*H. marismortui* **CP, c)***B. pseudomallei* **CP, and d)proposed INH active site. The N-terminal domain is indicated in cyan, while the C-terminal domain is in blueviolet. The heme moiety is shown in red. The Ca^{2+} ions in HRPC are shown in yellow. The figure was created in MOLSCRIPT[69] and rendered in Render 3D.[70]**

Despite the lack of a complete three-dimensional structure, the active site of *M. tuberculosis* CP was partially characterized (Figure 4.8 and Figure 4.9). Using EXAFS, Powers *et al.*[71] generated a three dimensional structure of the *E. coli katG* active site showing a heme group interacting with residues W105, H267, R102, and H106 (Figure 4.9). The iron center has been shown by Raman and optical spectroscopy to be a high spin-state, Fe^{3+}, 5-coordinate complex with some six-coordinate character.[23] The oxidized species appears to be a mixture of both six-coordinate low spin and five-coordinate high spin state heme molecules.
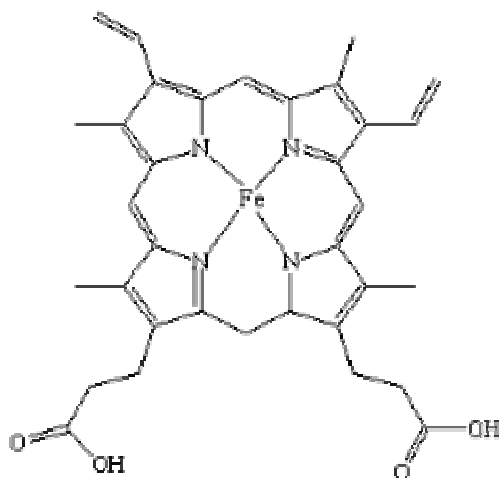
Figure 4.8 Iron protoporphyrin IX prosthetic group.
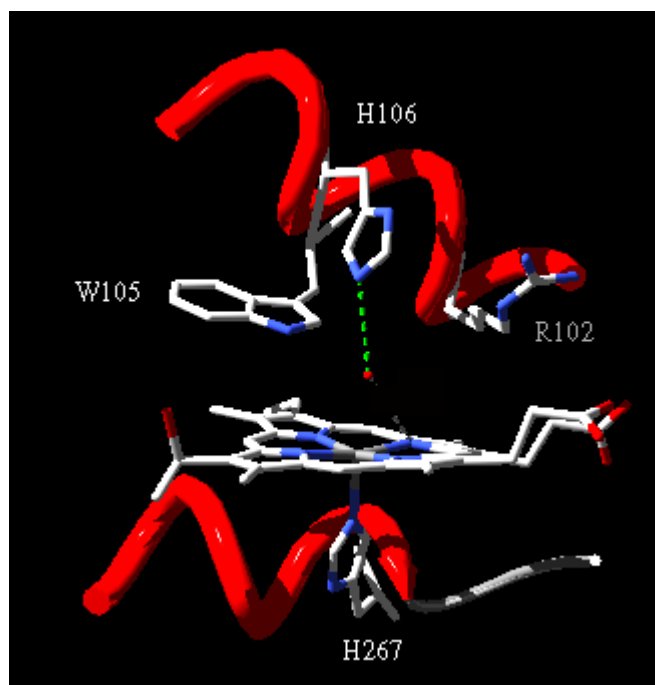
Figure 4.9 Active site of *M. tuberculosis* CP as determined by EXAFS by Powers *et al*. Figure created in Swiss PDB Viewer[72] and rendered in Pov-Ray.[73]

To aid in better understanding the binding and mechanism of INH activation, a homology model for *M. tuberculosis* has been created (Figure 4.10)[74] from the crystal structure of *H. marismortui* CP.[66] The enzyme is a dimer with two identical subunits and is composed primarily of α-helices. The active site is located in the N-terminal domain, which consists of approximately 430 residues, with the heme moiety sequestered within the structure. Substrate access to the active site is through a narrow channel that prevents access of a large substrate.[74] The active site is identical to that described by Powers *et al*.[71] using

EXAFS analysis. A comparison of residues in the active site pockets of *H. marismortui* CP, *B. pseudomallei* CP, HRPC, and the *M. tuberculosis* CP model can be seen in Table 4.1.
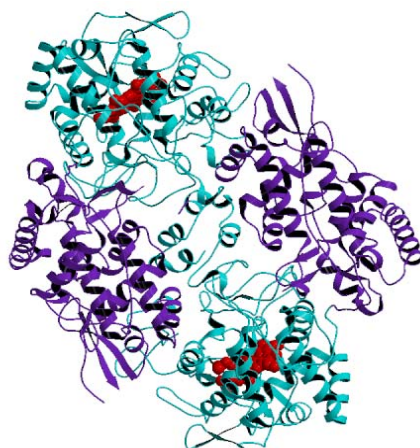
Figure 4.10 *M. tuberculosis* **CP model. The N-terminal domain is indicated in** cyan**, while the C-terminal domain is in** blueviolet**. The heme moiety is shown in** red**. The figure was created in MOLSCRIPT**[69] **and rendered in Render 3D.**[70]

Enzyme	Distal H$_2$O$_2$ binding pocket			Proximal Heme iron pocket			PDB code
*mt*CP	R104	W107	H108	H270	W321	D381	
*ha*CP	R92	W95	H96	H259	W311	D372	1ITK[66]
*bp*CP	R108	W111	H112	H279	W330	D389	1MWV[67]
HRPC	R38	F41	H42	H170	F221	D247	1ATJ[54]

Table 4.1 Comparison of postulated active site residues of *M. tuberculosis* CP with other peroxidase and CP enzymes.

Both the HRPC and *B. pseudomallei* CP crystal structures indicate the presence of metal cations. *M. tuberculosis* CP is expected to have similar sites. To help identify which metal cations may be present in *M. tuberculosis* CP, induced coupled plasma atomic emission spectroscopy study was undertaken.

GRID calculations were conducted to comparatively analyze the similarities in homology around the active site of HRPC, *B. pseudomallei* CP, and *M. tuberculosis* CP. The binding location, as well as residues which participate in the binding of INH, were investigated. Based on information acquired *via* GRID calculations, hypotheses regarding INH binding and activation are proposed.

The availability of a three dimensional X-ray crystal structure of *M. tuberculosis* CP would provide the structure of the heme pocket, as well as provide information regarding potential binding and activation sites for INH. Of particular interest is the presence or absence the access channel proposed for INH activation in *B. pseuodomallei* CP. In an attempt to allucidate structural information for *M. tuberculosis* CP, crystallization studies were performed.

RESULTS

ICP-AES

Results of the ICP-AES analysis of the wild-type, heme-added *M. tuberculosis* CP are summarized in Table 4.2. The high phosphorous content

originates from the KH_2PO_4/K_2HPO_4 buffer used during preparation of the enzyme. The presence of sulfur is expected from the cysteine and methionine residues. The iron concentration indicates a stoichiometry of approximately 2 hemes per *M. tuberculosis* CP dimer based on the sulfur to iron ratio for the known genetic sequence. Interestingly, there was a significant quantity of zinc present, which indicated the need to perform GRID analysis to investigate potential binding sites for a zinc dication in the protein. Other trace elements such as copper and magnesium were observed in the sample.

	Al	Ca	Cr	Fe	K	Mn
Mean concentration (ppm)	8.063	7.267	12.74	71.662	424.103	1.68
	Na	Ni	P	S	Zn	
Mean concentration (ppm)	1.747	8.606	284.18	127.9	28.02	

Table 4.2 ICP-AES results shown for wild-type heme added *M. tuberculosis* CP. Concentrations are shown in ppm with the contribution from the blank subtracted. Only metals which were present in greater than a 1 ppm quantity are shown.

GRID Calculations

Utilizing the ability of GRID to predict energetically favorable binding sites for small molecules, experiments were undertaken to investigate both the binding potential for INH and a probable location for a structural metal cation, such as Zn^{2+}. Since an X-ray crystal structure for *M. tuberculosis* CP was

unavailable, atomic coordinates for the *H. marismortui* homology model created by Eady *et al.*[74] was utilized (Figure 4.10). In addition to the *M. tuberculosis* CP model, the HRPC, APX, C*c*P, and *B. pseudomallei* CP structures were analyzed. The atomic coordinates for *B. pseudomallei* CP and Horseradish Peroxidase (HRP) were acquired from the PDB (1ATG, 1APX, 1CCP, and 1MWV, respectively). The INH-HRPC model developed by Banci *et al.*[51] was used to illustrate the suggested binding location of INH in HRPC.

Water molecules were removed from the structure file prior to the calculations. GRID was then used to evaluate the energetically favorable binding sites of counterions. Counterions were placed in positions to avoid the active site cavity, as well as sites known to contain electron density. Appropriate small molecule probes were then used to analyze the electronically neutral structure.

Several probes were used to investigate potential binding sites for INH in the enzyme. The probes are designed to locate energetically favorable binding sites in the enzyme for the functional moiety they represent. Based on the structure of INH and proposed oxidation products, probes were chosen to model their functional groups (Figure 4.11). The carboxy and carbonyl probes were selected due to the presence of a carbonyl moiety on INH. The N1 and N2 probes represent neutral, flat NH and NH_2 groups, respectively, which are present at the hydrazide moiety of INH. The aliphatic neutral amide multi-atom probe, conhr,

was used to simulate the hydrazinocarbonyl group of INH (Figure 4.11). The closest representation of INH is the multi-atom probe, ar.conhr, which was employed to locate energetically favorable binding positions for aromatic neutral amide functional groups (Figure 4.11). The ar.conhr probe was set to position the NH moiety *trans* with respect to the carbonyl oxygen and closely represents the pyridine-like ring and hydrazinocarbonyl group. Figure 4.12 illustrates the structures of HRPC, APX, C*c*P, *B. pseudomallei* CP, and *M. tuberculosis* CP prior to performing GRID calculations.
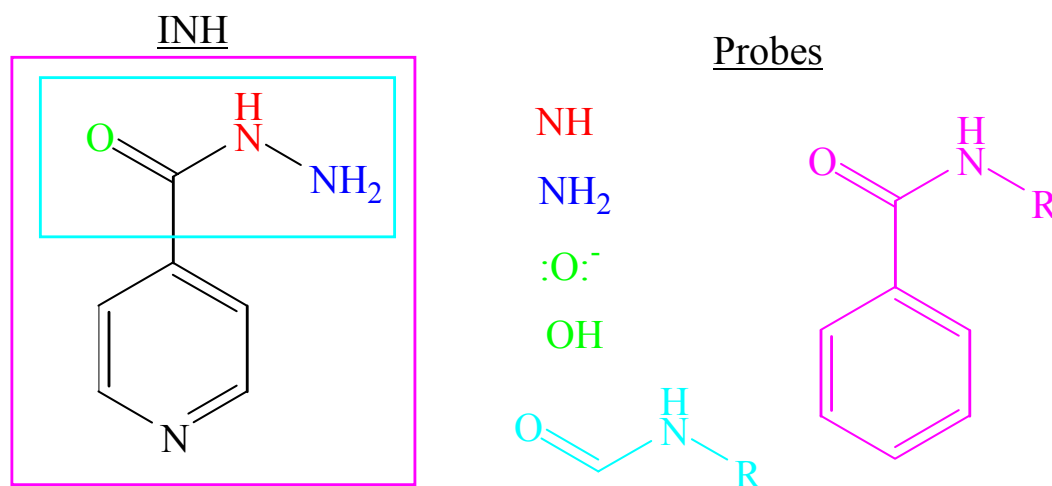
Figure 4.11 Probes used to investigate potential binding locations of INH are indicated on the right and are color coordinated to the diagram of INH on the left.

GRID calculation output was displayed as contour maps using Insight II (Accelrys, Cambridge, UK) and DS ViewerPro v. 5.0 (Accelrys, Cambridge, UK). By adjusting the visible contour level, the total energy cut off value was set

for each calculation. Although qualitative comparisons can be made across systems, quantitative analysis is not recommended as the energy values are probe and target specific.

The presence of energetically favorable sites was determined by visual inspection of the size and shape of the contour within a probe/enzyme system. The user is required to apply knowledge of the biochemistry of the protein being evaluated to determine whether sites identified as energetically favorable are significant to the posed question.

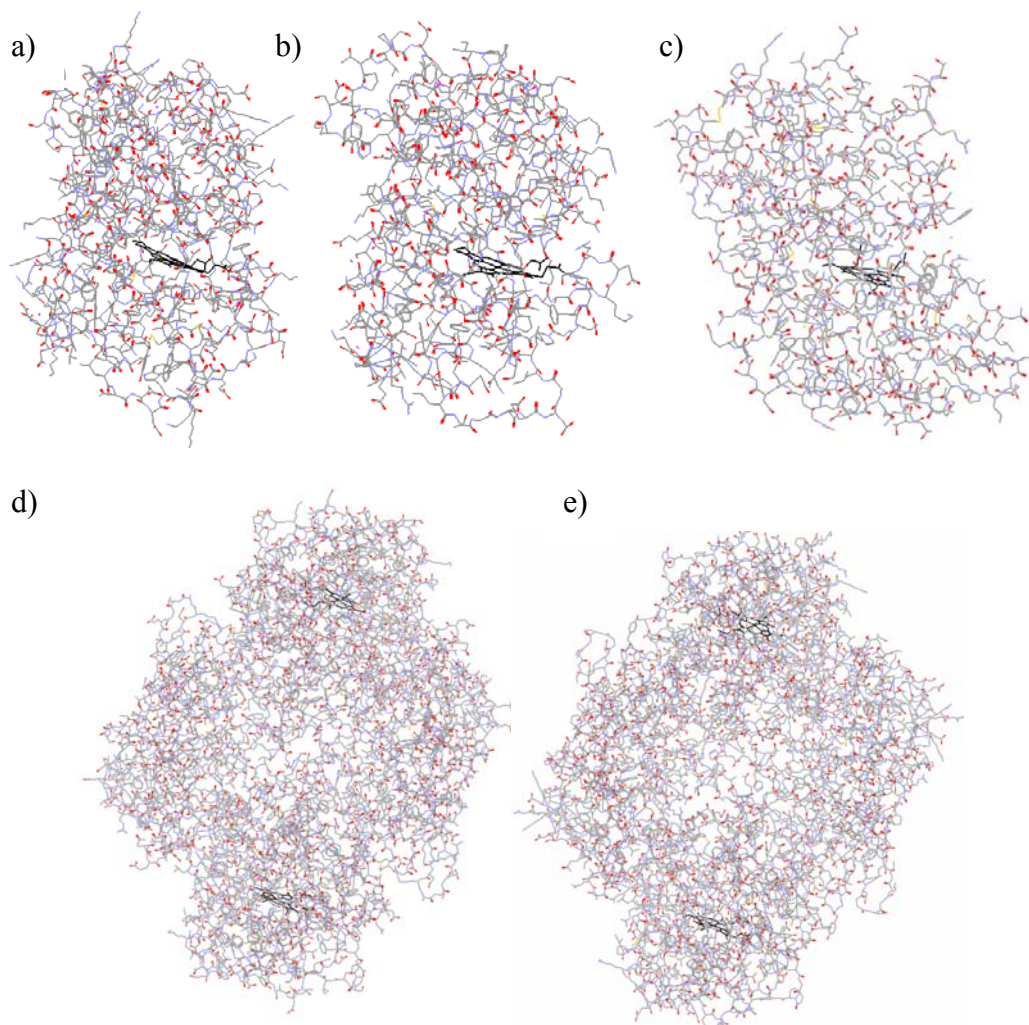**Figure 4.12 The structure of a)C*c*P, b)APX, c)HRPC, d)*B. pseudomallei* CP, and e)*M. tuberculosis* CP prior to GRID calculations. The heme moiety is indicated in black.**

NH Probe

A neutral, flat NH group probe, N1, was implicated to investigate favorable binding sites for INH. In C*c*P, APX, HRPC, *B. pseudomallei* CP, and

M. tuberculosis CP, an energetically favorable site was located near the heme active site. The class I peroxidases, C*c*P and APX, had similar interactions with the NH probe. In C*c*P, energetically favorable interaction included R48, W51, H52, and the iron and pyridinyl nitrogen atoms of the heme moiety (Figure 4.13a), while in APX, residues R38, W41, H42, and the iron and pyridinal nitrogen atoms of the heme moiety were determined to be energetically favorable (Figure 4.13b). The site involved interactions with residues R38, F41, H42, N70, as well as the iron and pyridinyl nitrogen atoms of the heme moiety in HRPC (Figure 4.13c), a class III peroxidase. A similar site was found for *B. pseudomallei* CP involving interactions with residues H112, W111, R108, and the iron and pyridinyl nitrogens of the heme moiety (Figure 4.13d). The NH probe used to analyze *M. tuberculosis* CP located energetically favorable interactions with residues H108, W107, N137, R104, and the iron and pyridinyl nitrogens of the heme moiety (Figure 4.13e).

Figure 4.13 Energetically favorable binding site for the NH probe with a)C*c*P, b)APX, c)HRPC, d)*B. pseudomallei* CP, and e)*M. tuberculosis* CP. Contour levels are set at -8, -7.5, −7, -8, and -8 kcal/mol, respectively. INH has been overlayed in c) to aid in visualization.

The structure of C*c*P was analyzed with the N2 probe, which is designed to locate favorable sites for neutral, flat NH$_2$ functional group binding. An energetically favorable binding site was located involving residues R48, W51, and H52. In APX, residues R38, W41, H42, and the iron and pyridinal nitrogen atoms of the heme moiety were determined to interact with the N2 probe. The HRPC structure was analyzed with the N2 probe. Interactions with residues H42, P139, L138, and G69, as well as the heme moiety were identified for HRPC (Figure 4.14a). Analyzing *B. pseudomallei* CP with the NH$_2$ probe showed energetically favorable interactions with residues H112, W111, N141, as well as the pyridinyl carbon atoms of the heme (Figure 4.14b). In *M. tuberculosis* CP, favorable interactions between the NH$_2$ probe and residues W107 and N137 and the pyridinyl carbon atoms of the heme group (Figure 4.14c) were located.
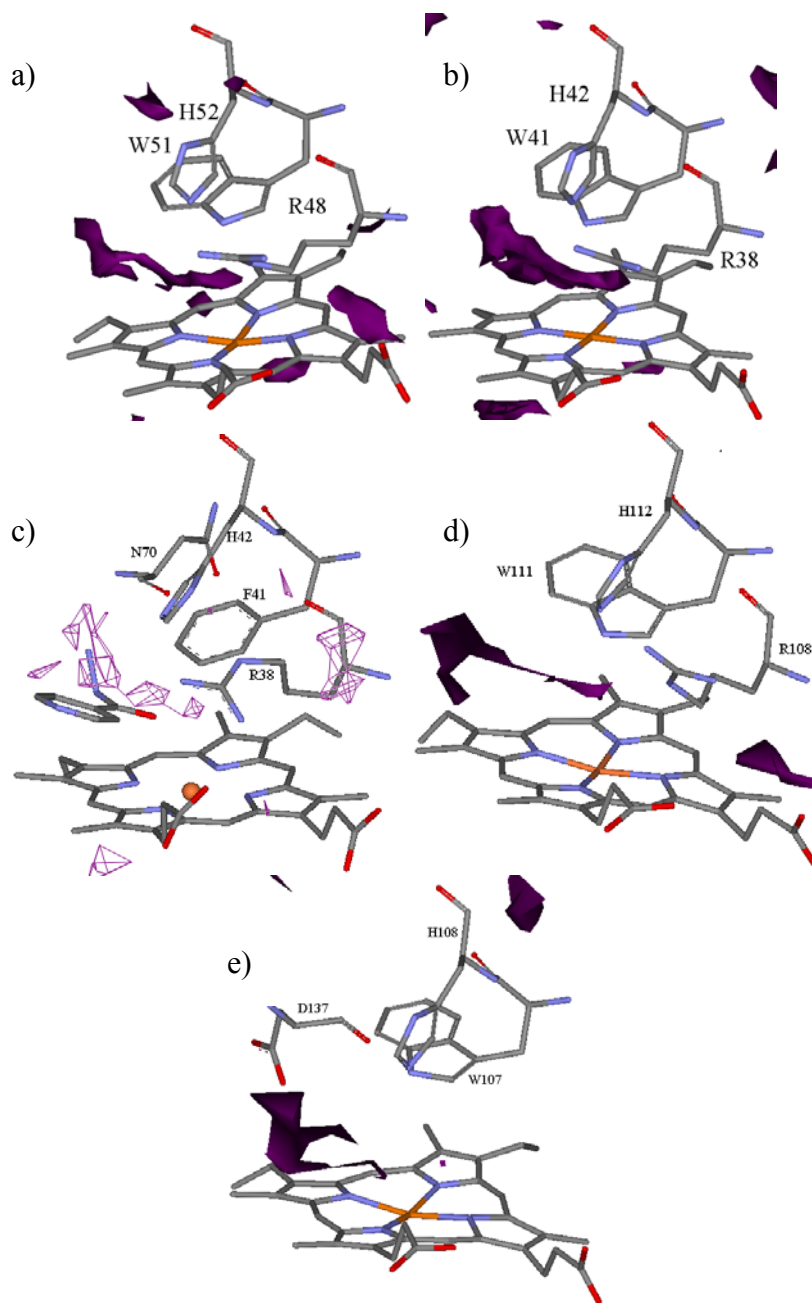
Figure 4.14   Energetically favorable binding site for the NH₂ probe with a)C*c*P, b)APX, c)HRPC, d)*B. pseudomallei*** CP, and e)***M. tuberculosis*** CP. Contour levels are set at -10, -9, –9, -10, and -8 kcal/mol, respectively.   INH has been overlayed in c) to aid in visualization.**

Carbonyl Probe

The class I peroxidase, C*c*P was analyzed with the carbonyl probe to locate favorable sites for *sp²* carbonyl oxygen binding. Residues R48, W51, and H52, as well as the iron and pyridinyl nitrogen atoms of the heme moiety were determined to interact with the carbonyl probe (Figure 4.15a). Energetically favorable interactions between APX and carbonyl probe were found to be similar to C*c*P, R38, W41, and H42, as well as the iron and pyridinyl nitrogen atoms of the heme moiety (Figure 4.15b). The HRPC structure was analyzed with the. Near the heme active site, the carbonyl probe located favorable interactions with R38, F41, H42, and the iron and pyridinyl nitrogen atoms of the heme group in HRPC (Figure 4.16c). In *B. pseudomallei* CP, the carbonyl probe identified favorable interactions with H112, W111, H279, R108, and the iron and pyridinyl nitrogens of the heme group (Figure 4.16d). Analysis of *M.* tuberculosis CP with the carbonyl probe located energetically favorable interactions with residues H108, W107, R104, and the iron and pyridinyl nitrogen atoms of the heme moiety (Figure 4.16e).
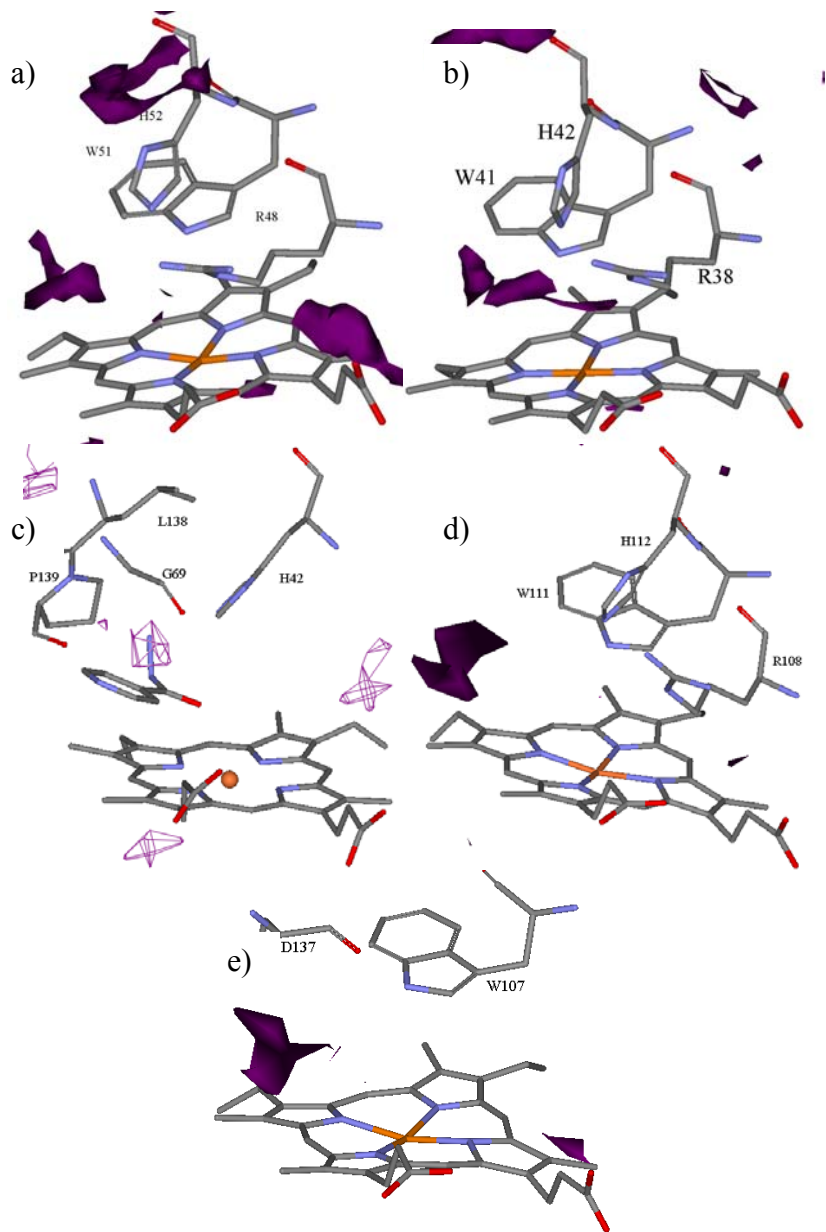
Figure 4.15 Energetically favorable binding site for the carbonyl probe with a)C*c*P, b)APX, c)HRPC, d)*B. pseudomallei* CP, and e)*M. tuberculosis* CP. Contour levels are set at −12 kcal/mol. INH has been overlayed in c) to aid in visualization.

Carboxy Probe

Analysis of C*c*P, APX, HRPC, *B. pseudomallei*, and *M. tuberculosis* with the carboxy probe located energetically favorable regions for anionic carboxy oxygen atom binding near the heme moiety. In the two class I peroxidases, energetically favorable interactions with the carboxy probe were with the arginine, tryptophan, and histidine residues nearest the heme, R48, W51, and H52 for C*c*P (Figure 4.16a) and R38, W41, and H42 for APX (Figure 4.16b). In HRPC, the strongest interactions were found to be with R38, F41, H42, P139, and the iron and pyridinyl nitrogen atoms of the heme group (Figure 4.16c). Analysis of *B. pseudomallei* CP located energetically favorable residues H112, W111, H279, R108, as well as the iron atom and pyridinyl nitrogen atoms of the heme moiety (Figure 4.16d). *M. tuberculosis* CP was evaluated with the carboxy probe and residues H108, W107, and R104, as well as the iron atom of the heme group were identified as being energetically favorable binding sites (Figure 4.16e).
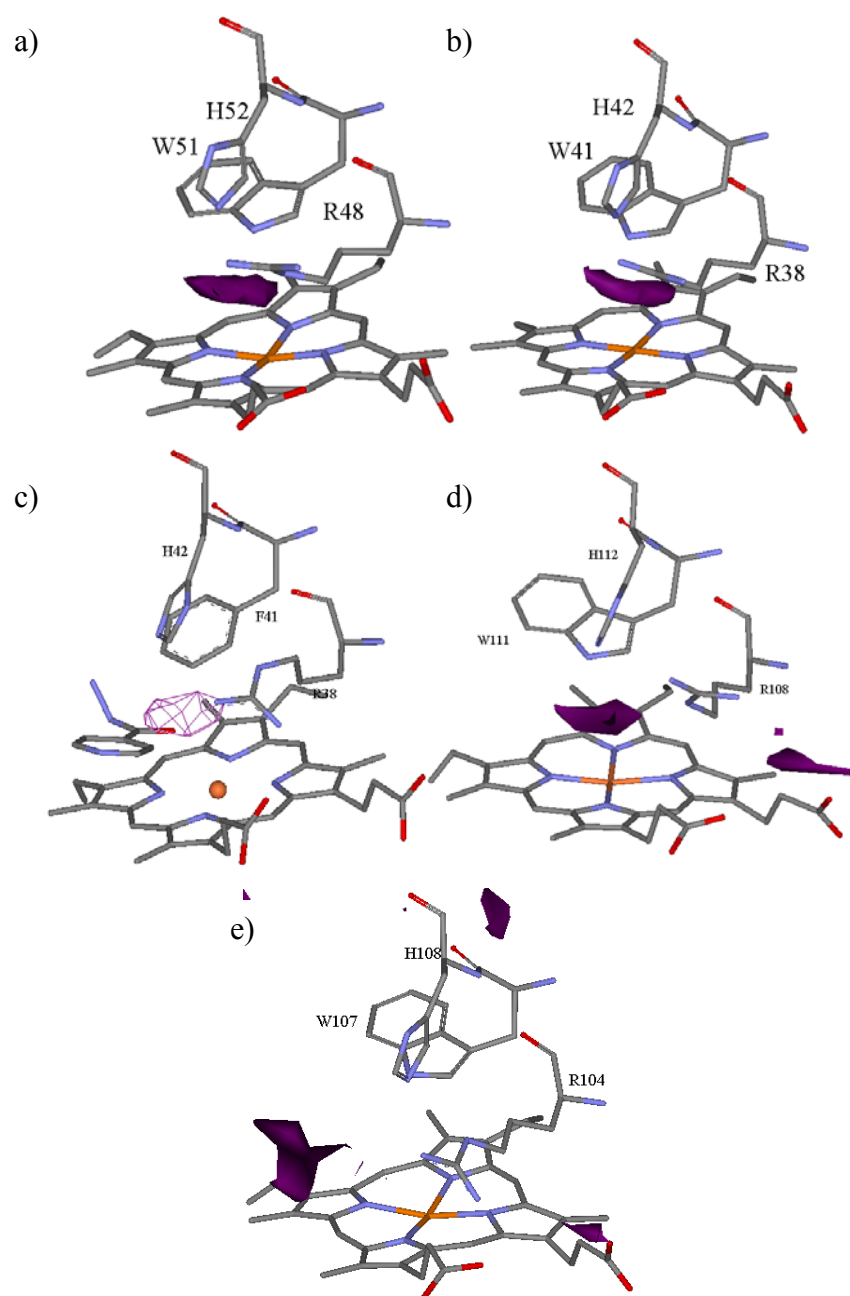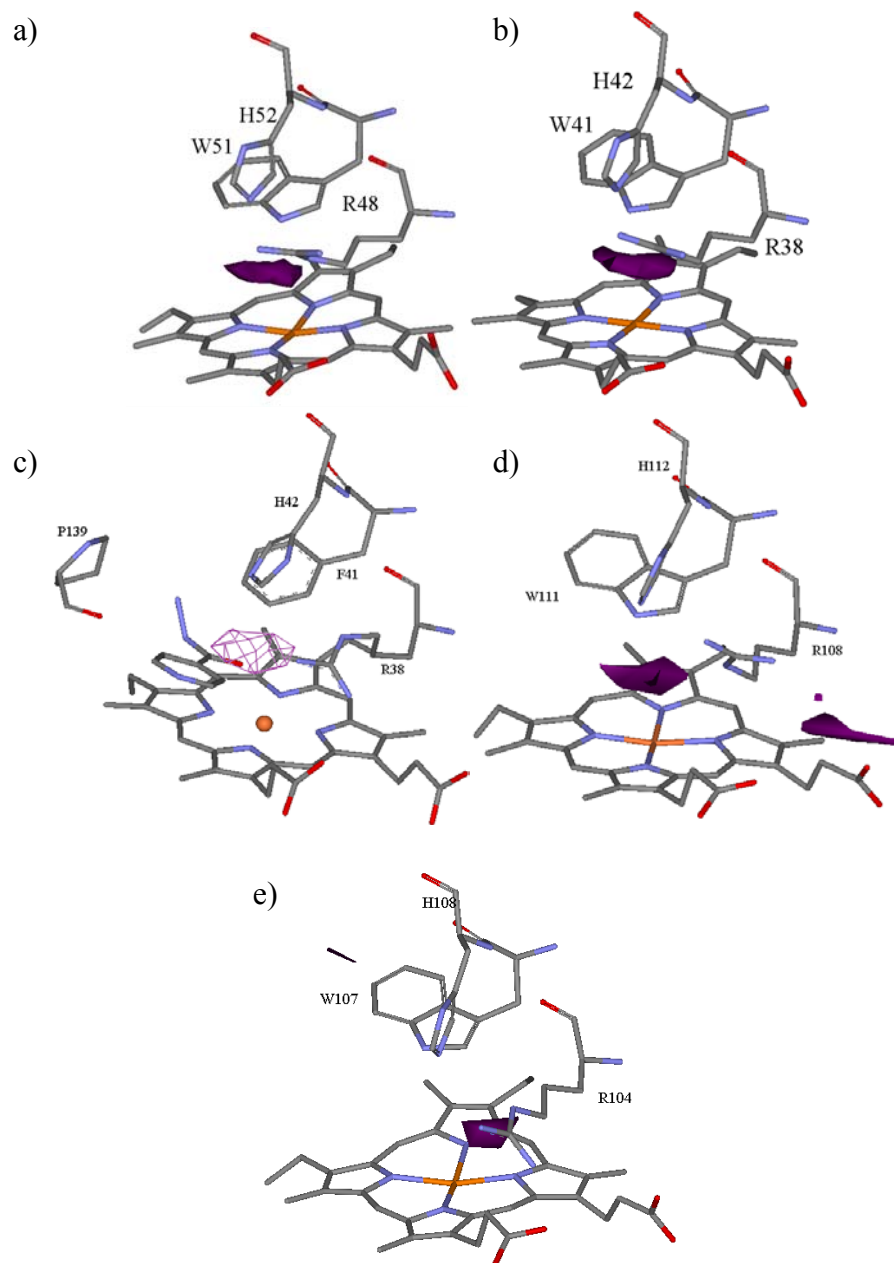
Figure 4.16 Energetically favorable binding site for the carboxy probe with a)C*c***P, b)APX, c)HRPC, d)***B. pseudomallei*** CP, and e)***M. tuberculosis*** CP. Contour levels are set at -20, -20, -14, -20, and -16 kcal/mol, respectively. INH has been overlayed in c) to aid in visualization.**

251

Aromatic Neutral Amide Probe

The multi-atom probe, ar.conhr, was employed to locate energetically favorable binding positions for aromatic neutral amide functional groups. The probe was set in a *trans* orientation, where the NH moiety was *trans* with respect to the carbonyl oxygen. Energetically favorable interactions with residues R48, W51, H52, and P145 were located in C*c*P (Figure 4.17a). Similar interactions were found in APX, R38, W41, H42, and P132 (Figure 4.17b). The aromatic neutral amide probe found a pocket of favorable interactions with residues G69, H42, N70, R38, as well as the iron and pyridinyl nitrogen atoms of the heme moiety of HRPC (Figure 4.17c). In *B. pseudomallei* CP residues H112, W111, H279, R108, as well as the iron and pyridinyl nitrogen atoms of the heme moiety were found to interact favorably (Figure 4.17d). Analysis of *M. tuberculosis* CP with the aromatic neutral amide probe indicated energetically favorable interactions with residues H108, W107, R104, N137, as well as the pyridinyl nitrogen atoms of the heme moiety (Figure 4.17e).

Figure 4.17   Energetically favorable binding site for the aromatic neutral amide probe with a)C*c*P, b)APX, c)HRPC, d)*B. pseudomallei* CP, and e)*M. tuberculosis* CP. Contour levels are set at -8, -8, –7, -8.5, and –8.5 kcal/mol, respectively. INH has been overlayed in c) to aid in visualization.**

<u>Aliphatic Neutral Amide Probe</u>

The multi-atom probe, conhr, was employed to locate energetically favorable binding positions for aliphatic neutral amide functional groups in CcP, APX, HRPC, *B. pseduomallei* CP, and *M. tuberculosis* CP. The probe was set so that the NH moiety was in a *trans* orientation with respect to the carbonyl oxygen. A pocket of energetically favorable interactions was located in CcP involving residues R48, W51, H52, and P145 (Figure 4.18a). Analysis of APX with the aliphatic neutral amide probe indicated the residues R38, W41, H42, and P132 interacted favorably (Figure 4.18b). In HRPC, the aromatic neutral amide probe found a pocket of favorable interactions with residues G69, H42, N70, R38, as well as the pyridinyl carbon atoms of the heme moiety (Figure 4.18c). Residues H112, W111, H279, R108, as well as the iron and pyridinyl nitrogen atoms of the heme moiety in *B. pseduomallei* CP showed energetically favorable interactions with the aromatic neutral amide probe (Figure 4.18d). Analysis of *M. tuberculosis* CP indicated energetically favorable interactions with residues H108, W107, R104, N137, as well as the pyridinyl carbon atoms of the heme moiety (Figure 4.18e).
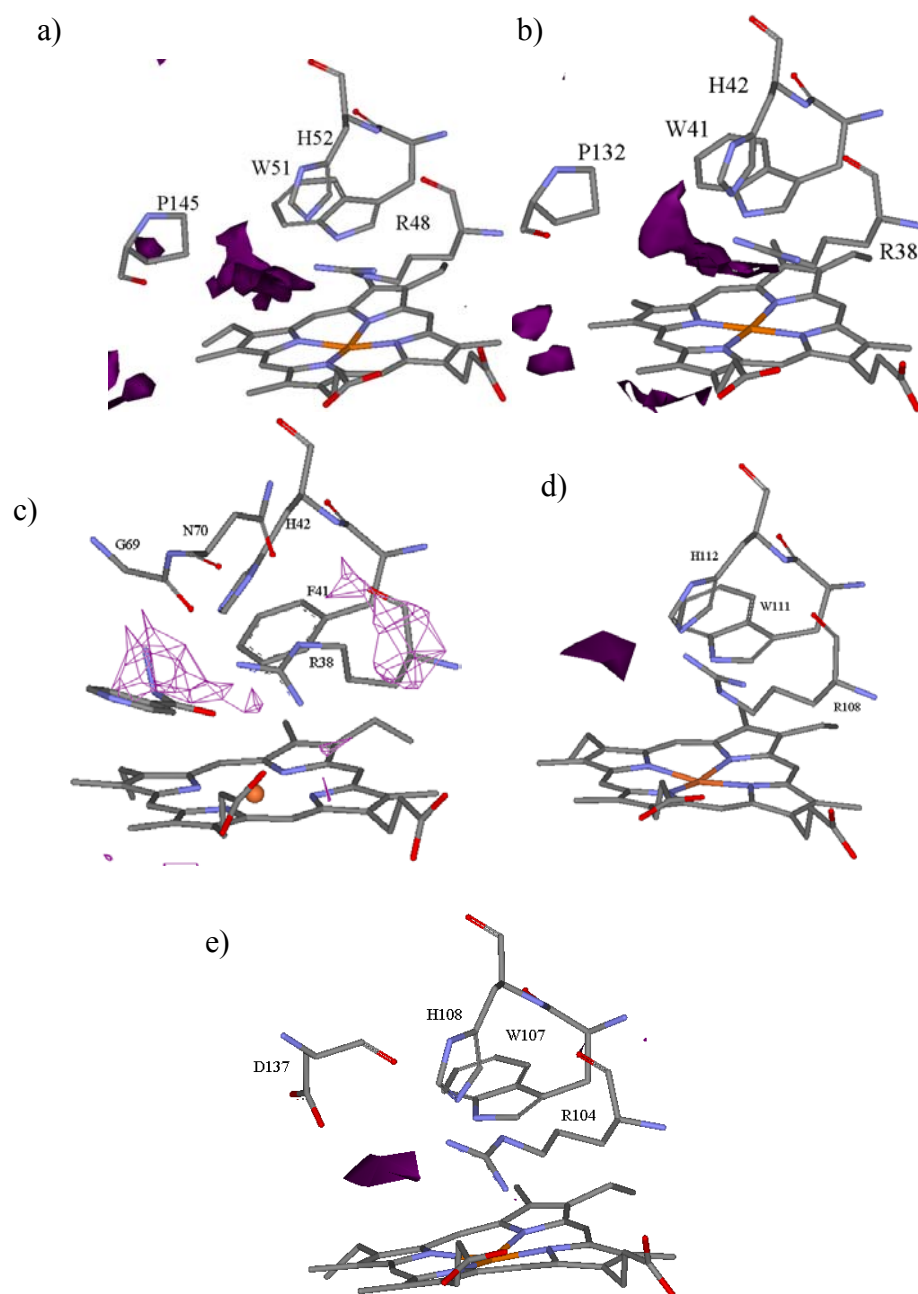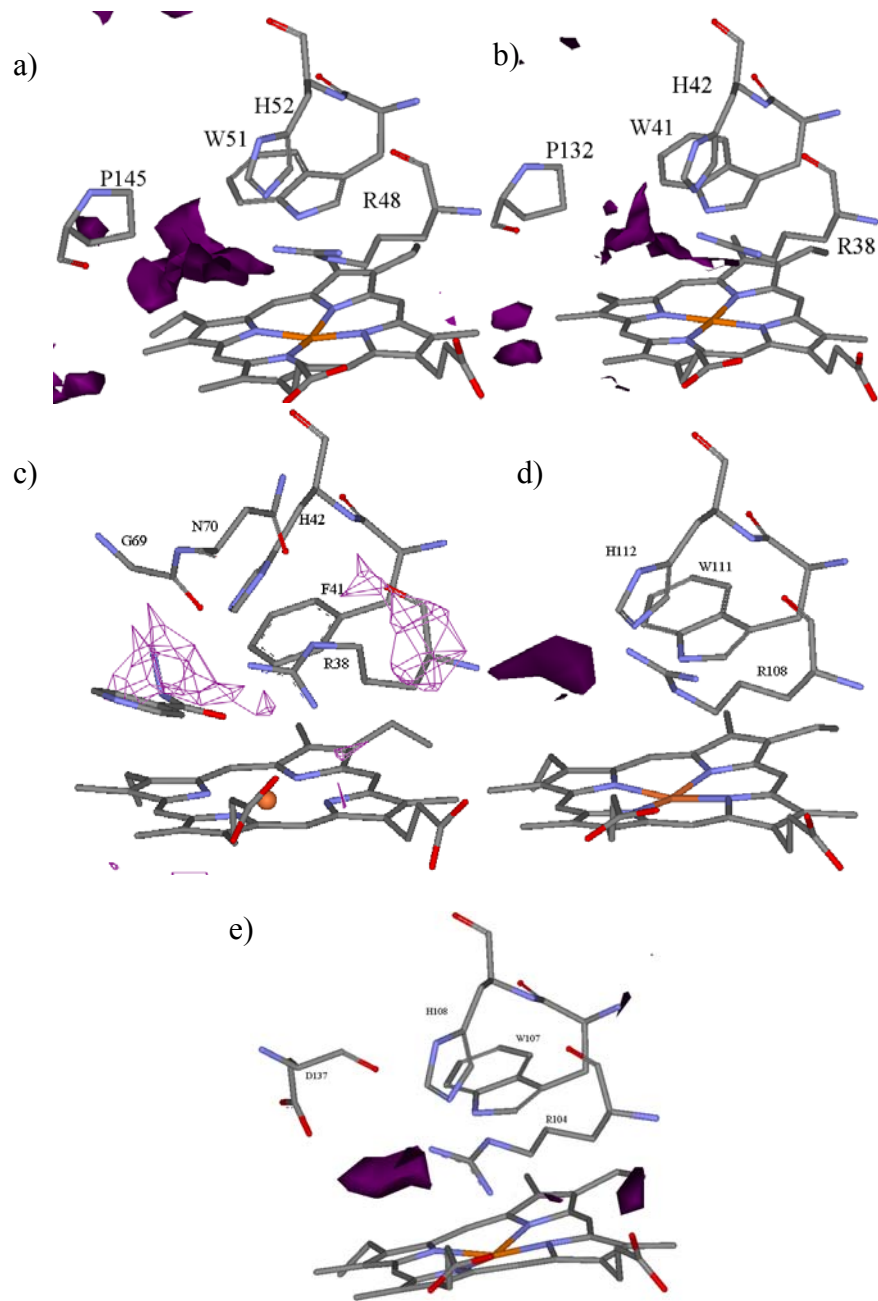
Figure 4.18   Energetically favorable binding site for the aliphatic neutral amide probe with a)C*c*P, b)APX, c)HRPC, d)*B. pseudomallei* CP, and e)*M. tuberculosis* CP.   Contour levels set at -8, -8.5, −7, -8.5, and −9 kcal/mol, respectively. INH has been overlayed in c) to aid in visualization.

Zinc Dication Probe for M. tuberculosis CP

The Zn^{2+} probe was employed to locate energetically favorable binding positions for Zn^{2+} interactions. The Zn^{2+} probe found a pocket of favorable interactions with residues Y95, D94, P100, K274, and H276, as well as the heme moiety (Figure 4.19).
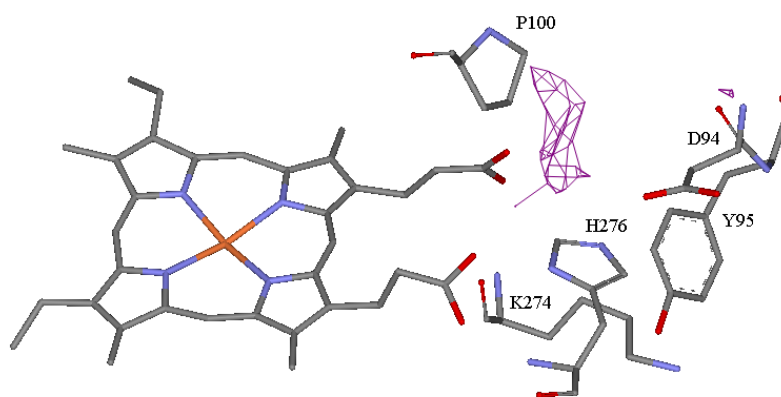
Figure 4.19   Energetically favorable binding site for the zinc dication probe with *M. tuberculosis* CP.   Contour level set at -50 kcal/mol.

Crystallization of M. tuberculosis CP

Crystals of *M. tuberculosis* CP were obtained under a variety of conditions and have been shown to be reproducible (Figure 4.20). Crystallization conditions included conditions similar to those described for *B. pseduomallei* CP. Crystals chosen for X-ray diffraction study were obtained from 4% Polyethylene glycol (PEG) 3,350 and 0.1 M sodium acetate (pH 4.6) after one day at 20 °C. The crystals were red-brown plates typically 0.2 x 0.2 x 0.05 mm in size.

Figure 4.20 Crystals of *M. tuberculosis* CP grown in 4% Polyethylene glycol (PEG) 3,350 and 0.1 M sodium acetate (pH 4.6) after one day at 20 °C.

Data Processing

Diffraction data were collected on beamLine 9.6 at the Synchrotron Radiation Source (SRS), Daresbury, UK using a CCD detector and 0.877 Å radiation. The data set was processed with the HKL suite of programs.[75] A summary of data processing parameters is shown in Table 4.3. The data sets were processed as $P4_22_12$ based on observed systematic absences, ($00l$: $l = 2n$ and $h00$: $h = 2n$).

257

M. tuberculosis CP

Crystal	*M. tuberculosis* CP
Radiation Source	BeamLine 9.6 (SRS, Daresbury, UK)
λ (Å)	1.488
Temperature (K)	100
Space Group	$P4_22_12$
Unit cell parameters	a = 150.33 α = 90.000 b = 150.33 β = 90.000 c = 154.28 γ = 90.000
No. of molecules in asymmetric unit	2
Resolution range (Å)	30 – 2.4
Total observations	1,932,986
Unique reflections	62,065
Completeness (%)	90.4 (89.1)*
R_{sym} (%)	15.9 (39.1)*
$(I/\sigma(I))$	11.2 (4.2)*

Table 4.3 Data collection and processing parameters for *M. tuberculosis* CP crystals. * values in parentheses are for the high-resolution shell.

Sequence Analysis and Molecular Replacement

Sequence alignment of *B. pseudomallei* CP amino acid sequence was performed against that of *M. tuberculosis* CP using the program *Genedoc*. A 64%

sequence identity was determined for the two amino acid sequences (Figure 4.21). Active site residues, H112 (H108), W111 (W107), H279 (H270), and R108 (R104), are conserved between the two CP sequences (*M. tuberculosis* CP numbers are shown in parentheses), further confirming the similarity between the two enzymes. Due to the significant sequence identity between *B. pseudomallei* and *M. tuberculosis* CP, the *B. pseudomallei* CP structure[67] was used as a search model.

```
B. pseudomallei CP   1    MPGSDAGERR RGVHEQRRNR MSNEAKCP-F --HQAAGNGT SNRDWWPNQL DLSILH
M. tuberculosis CP   1    ------MPEQ HPPITETTTG AASNG-CPVV GHMKYPVEGG GNQDWWPNRL NLKVLH

B. pseudomallei CP   58   LSDPMGKDEN YAQAFEKLDL AAVKRDLHAL MTTSQDWWPA DFGHYGGLFI RMAWHS
M. tuberculosis CP   54   VADPMGAAFD YAAEVATIDV DALTRDIEEV MTTSQPWWPA DYGHYGPLFI RMAWHA

B. pseudomallei CP   118  RTADGRGGAG EGQQRFAPLN SWPDNANLDK ARRLLWPIKQ KYGRAISWAD LLILTG
M. tuberculosis CP   114  RIHDGRGGAG GGMQRFAPLN SWPDNASLDK ARRLLWPVKK KYGKKLSWAD LIVFAG

B. pseudomallei CP   178  ESMGFKTFGF AGGRADTWEP EDVYWGSEKI WLELSGGPNS RYSGDRQLEN PLAAVQ
M. tuberculosis CP   174  ESMGFKTFGF GFGRVDQWEP DEVYWGKEAT W--L---GDE RYSGKRDLEN PLAAVQ

B. pseudomallei CP   238  YVNPEGEDGN PDPVAAARDI RDTFARMAMN DEETVALIAG GHTFGKTHGA GPASNV
M. tuberculosis CP   229  YVNPEGPNGN PDPMAAAVDI RETFRRMAMN DVETAALIVG GHTFGKTHGA GPADLV

B. pseudomallei CP   298  EAAGIEAQGL GWKSAYRTGK GADAITSGLE VTWTTTPTQW SHNFFENIFG YEWELT
M. tuberculosis CP   289  EAAPLEQMGL GWKSSYGTGT GKDAITSGIE VVWTNTPTKW DNSFLEILYG YEWELT

B. pseudomallei CP   358  GAHQWVAK-- GADAVIPDAF DPSKKHRPTM LTTDLSLRFD PAYEKISRRF HENPEQ
M. tuberculosis CP   349  GAWQYTAKDG AGAGTIPDPF GGPGR-SPTM LATDLSLRVD PIYERITRRW LEHPEE

B. pseudomallei CP   416  FARAWFKLTH RDMGPRARYL GPEVPAEVLL WQDEIPAVDH PLIDAADAAE LKAKVL
M. tuberculosis CP   408  FAKAWYKLIH RDMGPVARYL GPLVPKQTLL WQDPVPAVSH DLVGEAEIAS LKSQIR

B. pseudomallei CP   476  TVSQLVSTAW AAASTFRGSD KRGGANGARI RLAPQKDWEA NQPE-QLAAV LETLEA
M. tuberculosis CP   468  TVSQLVSTAW AAASSFRGSD KRGGANGGRI RLQPQVGWEV NDPDGDLRKV IRTLEE

B. pseudomallei CP   535  FNGAQRGGKQ VSLADLIVLA GCAGVEQAAK NAGHAVTVPF APGRADASQE QTDVES
M. tuberculosis CP   528  FNSAAPGNIK VSFADLVVLG GCAAIEKAAK AAGHNITVPF TPGRTDASQE QTDVES

B. pseudomallei CP   595  EPVADGFRNY LKGKYRVPAE VLLVDKAQLL TLSAPEMTVL LGGLRVLGAN VGQSRH
M. tuberculosis CP   588  EPKADGFRNY LGKGNPLPAE YMLLDKANLL TLSAPEMTVL VGGLRVLGAN YKRLPL

B. pseudomallei CP   655  AREQALTNDF FVNLLDMGTE WKPTAADADV FEGRDRATGE LKWTGTRVDL VFGSHS
M. tuberculosis CP   648  EASESLTNDF FVNLLDMGIT WEPSPADDGT YQGKD-GSGK VKWTGSRVDL VFGSNS

B. pseudomallei CP   715  LAEVYGSADA QEKFVRDFVA VWNKVMNLDR FDLA 748
M. tuberculosis CP   707  LVEVYGADDA QPKFVQDFVA AWDKVMNLDR FDVR 740
```

Figure 4.21 Sequence alignment of *B. pseudomallei* and *M. tuberculosis* CP. Identical residues are highlighted in blue and similar residues are highlighted in yellow. Sequence was aligned and figure created in BioEdit.[76]

Refined X-ray Crystal Structure of M. tuberculosis CP

The X-ray crystal structure of *M. tuberculosis* CP was successfully refined to 2.4 Å resolution and is shown in Figure 4.22. The structure consists of 716 residues, 703 water molecules, and 3 glycerol molecules. One well-defined heme

moiety is buried in the N-terminal domain. The dimeric structure is generated by crystallographic symmetry (Figure 4.22).
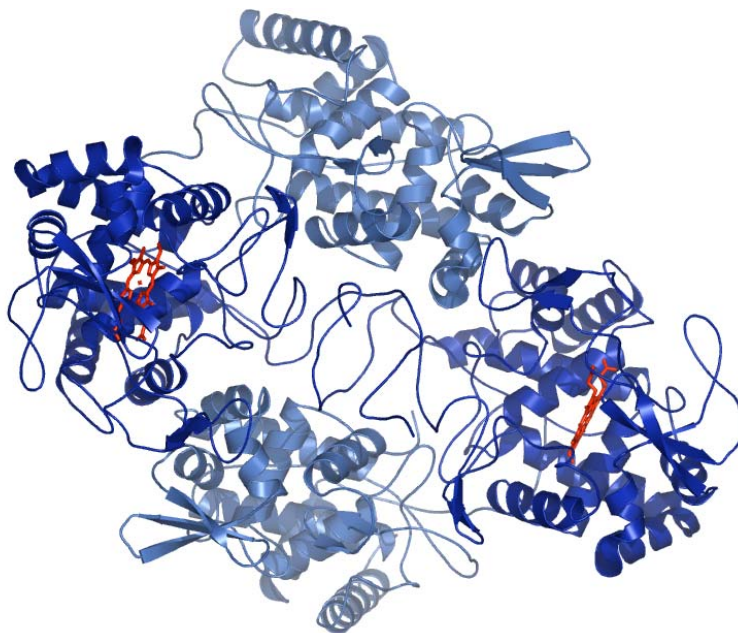
Figure 4.22 Structure of *M. tuberculosis* CP with the heme moieties indicated in red. Figure was produced in PyMOL (DeLano Scientific, San Carlos, California).

The overall fold is similar to that of *B. pseudomallei* CP[67] and *H. marismortui* CP;[66] with two identical subunits related by non-cyrstallographic two-fold symmetry. The first 431 residues compose the N-terminal domain, while the remaining 285 compose the C-terminal domain. Structurally, *M. tuberculosis* CP contains mostly α-helices with few β-sheets, which is similar to both *B. pseudomallei* CP and *H. marismortui* CP (Figure 4.22).

261

The active site of *M. tuberculosis* CP is shown in (Figure 4.23). The organization around the active site is similar in all three CP structures available, as indicated by the number of conserved residues, H108, H270, W107, and R104. Electron density around the heme moiety is well defined (Figure 4.24). The heme moiety in the *B. pseudomallei* CP structure contained a two atom substituent attached to a vinyl group.[67] This modification was not found in the *M. tuberculosis* CP structure. The location of water molecules in the active sites does vary between the catalase-peroxidases, with *M. tuberculosis* CP having only two water molecules present on the distal side of the heme moiety (Figure 4.23). *B. pseudomallei* CP has three water molecules on the distal side of the heme moiety with one molecule directly above the iron atom,[67] while *H. marismortui* CP has both a molecule with two water molecules and a molecule with three water molecules.[66]
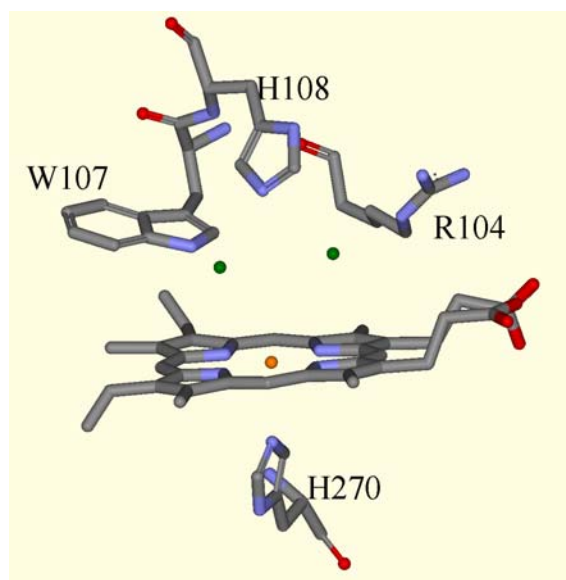
Figure 4.23 Active site of *M. tuberculosis* CP with the water molecules shown in green. Figure was created in DS Viewer Pro 5.0 (Accelrys, Cambridge, UK).
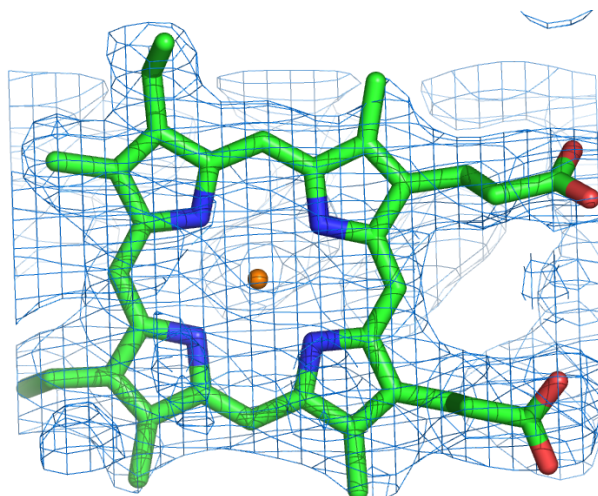
Figure 4.24 $2F_o$-$2F_c$ electron density map (blue) calculated to 2.4 Å resolution contoured at 1 σ. The carbon atoms are shown in green, the nitrogen atoms in blue, the oxygen atoms in red and the Fe^{3+} in orange. Figure was produced in PyMOL (DeLano Scientific, San Carlos, California).

When compared to the other structurally characterized catalase-peroxidases, five additional residues were determined at the N-terminal of *M. tuberculosis* CP. Residues G24, H25, M26, K27, and Y28 are indicated in Figure 4.25 in red. The N-terminal residues appear to interact with residues Y197, K200, and G208 of the C-terminal domain.
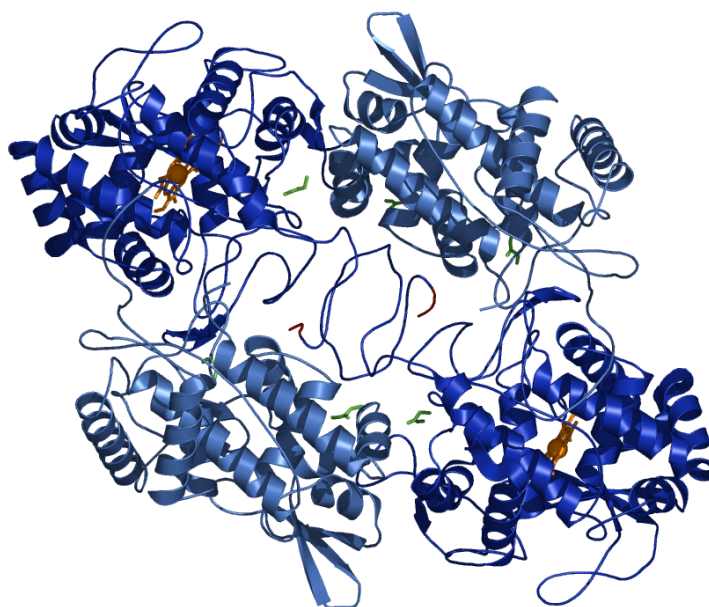
Figure 4.25 Additional residues located at the N-terminal of *M. tuberculosis* CP are indicated in red, while glycerol molecules are shown in green. Figure created in PyMOL (DeLano Scientific, San Carlos, California)

Although the identity is unknown, a potential metal cation was located near the γ-edge of the heme moiety. The region of electron density is coordinated to the oxygen atom of the water molecule 261, the O2α and O1δ atoms of the propionate groups of the heme moiety, and NH1 of R104 in a distorted-

tetrahedral-type geometry. Based on the coordination geometry, the cation is unlikely to be Na^+ or K^+, but could be a dication such as Zn^{2+} or Ca^{2+}.
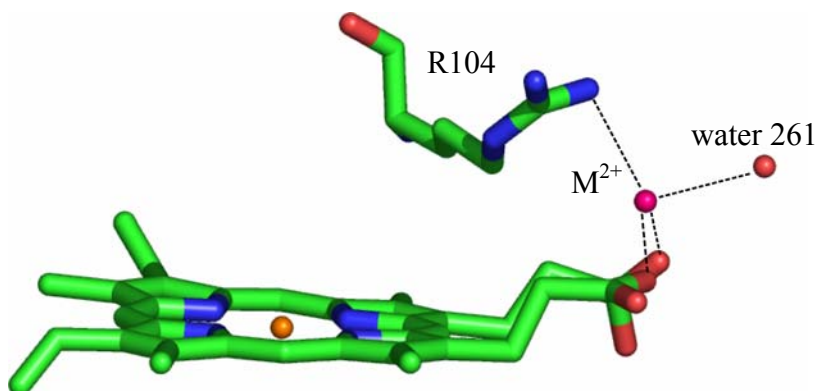
Figure 4.26 Unknown metal cation site in *M. tuberculosis* CP. Figure created in PyMOL (DeLano Scientific, San Carlos, California) and edited in Adobe Illustrator CS (Adobe Systems, San Jose, California).

Three glycerol molecules, which were incorporated into the enzyme from the cyroprotectant solution, were located in the structure (Figure 4.25). One molecule is located near residues G607, R693, N701, and L43, which is approaching the region close to the N-terminal. The second molecule is near residues W191, K157, and K158, while the third molecule is located near residues L587 and K613.

A region of continuous electron density was observed between $C\epsilon 1$ of Y229 and $C\eta 2$ of W107 and between $C2\epsilon$ of Y229 and $S\delta$ of M255 (Figure 4.27). The existence of the covalent bonds between these side chains was confirmed in both subunits. Donation of electrons by the $S\delta$ of M255 to form the extra

covalent bond places a positive charge on the side chain. The guanidine group of R418 forms hydrogen bonds with Oη of Y229 and the amide nitrogen atom of M255, which stabilize the covalent bond between Y229 and M255 by fixing the relative position of the residues. The covalent modification does not disturb the orientation of Y229 and the heme, which has been shown to be important for the reaction of the second peroxide with Compound I.[12]
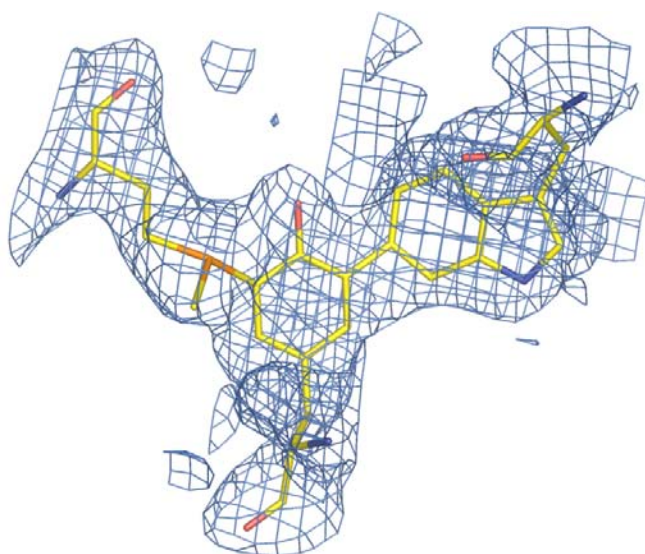
Figure 4.27   Electron density map showing covalent bond modification around W107-Y229-M255.   The figure was created in PyMOL (DeLano Scientific, San Carlos, California).

The most obvious access route to the active site of *M. tuberculosis* CP is *via* a channel positioned similarly to other peroxidases. The narrow channel prevents large substrates access to the heme moiety. S315, a residue which has been shown to be important for INH activation,[43] is situated along the access

266

channel. The funnel-shaped cavity was found to contain a continuum of water

molecules in the X-ray crystal structure (Figure 4.28).
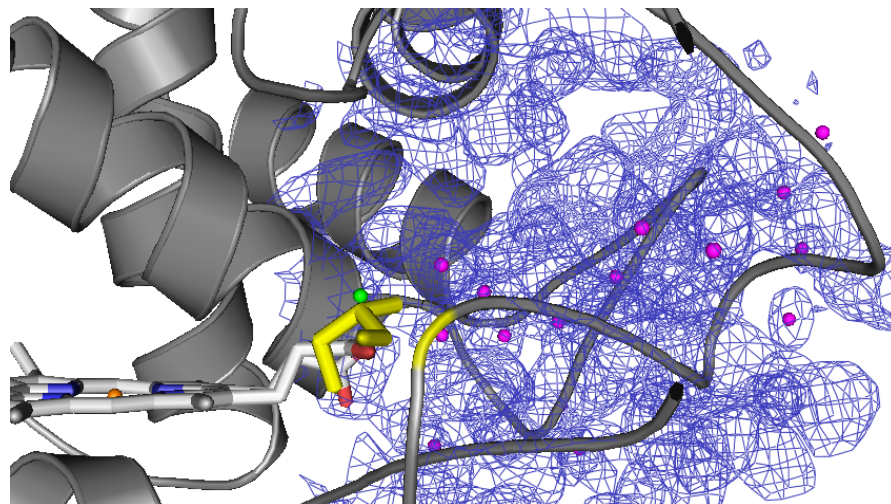
Figure 4.28 Electron density map showing the continuum of water molecules (magenta) located in the access channel. The S315 residue is indicated in yellow and the unidentified metal cation is green. Figure created in PyMOL (DeLano Scientific, San Carlos, California).

DISCUSSION

Improving treatment methods for tuberculosis is hindered by the lack of information regarding the binding/activation site of INH in *M. tuberculosis* CP. In order to aid in understanding where INH might bind, as well as what residues may play a critical role in the activation of the pro-drug, GRID analysis of *M. tuberculosis* CP and two other class III peroxidases, *B. pseudomallei* CP and HRPC, was undertaken. The possibility of a second active site in *B. pseudomallei* CP, as reported by Carpena *et al.*,[67] was scrutinized using GRID analysis.

As previously discussed, *M. tuberculosis* CP is a dual-function enzyme, which is both catalase and peroxidase active. In an attempt to identify any differences between the peroxidases and the catalase-peroxidases, GRID analysis was performed on the crystal structures of APX and C*c*P. The absence of a loop located on the γ-edge of the heme in the catalase-peroxidase, was determined to be a potential difference between the peroxidases and catalase-peroxidases.

The presence of metal cations in the crystal structures of *B. pseudomallei* CP, HRPC, APX, and C*c*P, led to the hypothesis that *M. tuberculosis* CP would also contain a metal cation. ICP-AES analysis indicated a significant quanity of Zn^{2+} present in the enzyme. To elucidate the potential binding location of the Zn^{2+}, GRID analysis was performed on the *M. tuberculosis* CP model. The

information gained from the two studies was then applied to the crystal structure of *M. tuberculosis* CP, once acquired.

Although a model of *M. tuberculosis* CP was readily available, the X-ray crystal structure of the enzyme was desirable. Crystallization and structural analysis of *M. tuberculosis* CP, has provided information regarding the environment around the heme moiety, the size and shape of the access channel and the presence of extra residues at the N-terminal of the enzyme. The covalent bonding system between W-Y-M, found in *B. pseudomallei* CP and *H. marismortui* CP, was determined to be present in *M. tuberculosis* CP as well.

GRID calculations for the binding of INH in *M. tuberculosis* CP, *B. pseudomallei* CP, and HRPC closely follow the HRPC-INH model presented by Banci *et al.*[51] Residues will be presented corresponding to the number assigned to *M. tuberculosis* CP with the residue numbers for *B. pseudomallei* CP and HRPC presented in parentheses. The NH and NH$_2$ probes determined that residues H108 (H112, H42), the W107 (W111, F41), and R104 (R108, R38), as well as the iron and pyridinyl nitrogen atoms of the heme moiety are energetically favorable for interaction with the hydrazide moiety of INH (Figure 4.29). The residues are similar to those reported by Banci *et al.*[51] for hydrazide interaction of INH in HRPC. The carbonyl moiety, modeled by both the carboxy and carbonyl probes, shows clear interactions with heme iron as well as residues H108

(H112, H42), W107 (W111, F41), and R104 (R108, R28). In addition to coordination by the iron, the nitrogen atoms of the heme appear to be involved in interaction with the carbonyl moiety of INH (Figure 4.29) Two multi-atom probes, conhr and ar.conhr, represent a larger portion of the INH structure than their single-atom counterparts.   Although the conhr should only locate the hydrazinocarbonyl of INH the two probes gave similar results. Residues H108 (H112, H42), W107 (W111, F41), R104 (R108, R28), as well as the pyridinyl nitrogen atoms of the heme moiety are predicted by GRID to create an energetically favorable binding pocket for the pyridine functional group of INH. The "aromatic binding pocket" is not identical to that predicted in the HRCP model, heme-18 methyl, G69, A140, P141, and F176. Despite the differences in interactive residues, the location of the binding site appears to be similar (Figure 4.29).

Figure 4.29 GRID maps depicting energetically favorable binding sites for the NH (green), NH$_2$ (blue), carbonyl oxygen (red) and aromatic neutral amide (purple) probes on (A) *M. tuberculosis* CP (B) *B. pseudomallei* CP, and (C) HRPC. GRID map contour levels are set at -8, -9, -12, and -8.5 kcal/mol for the NH, NH$_2$, carbonyl, and aromatic neutral amide probes, respectively. In all panels the heme is in gray and protein mainchains and side chains are depicted in yellow. Interacting amino acid residues are indicated as stick figures. This figure was produced using DS ViewerPro 5.0 (Accelrys, Cambridge, UK).

Data from the GRID calculations suggests that the INH binding site in *B. pseudomallei* CP is similar to that modeled in HRPC[51] and crystallographically determined for BHA in HRPC,[77] not the secondary active site previously reported.[67] Key residues involved in interactions with the functional groups of

INH are H112, W111, R108, and D141, as well as the heme moiety. Interestingly, the putative secondary active site suggested by Carpena *et al.* (Figure 4.7d)[67] did not show any energetically favorable interactions with the probes studied. The idea of INH binding at this position was not supported by GRID calculations.

GRID calculations to determine the most energetically favorable binding site for a pyridinyl substrate, such as INH, were performed on APX and C*c*P. All residues numbers will be indicated for APX with C*c*P in parentheses. The NH, NH_2, and carbonyl probes all indicated interactions with Nϵ of R38 (48), Nϵ1 of W41 (51), Nϵ2 of H42 (52), and the iron and pyridinyl nitrogen atoms of the heme moiety (Figure 4.30). In addition to the R, W, and H residues indicated to be energetically favorable toward the NH, NH_2, and carbonyl probes, the aromatic neutral amide probe located interactions with the O atom of P132 (145) (Figure 4.30). The energetically favorable interactions determined for APX and C*c*P were similar to those found for the catalase-peroxidases, *M. tuberculosis* CP and *B. pseudomallei* CP (Figure 4.29).
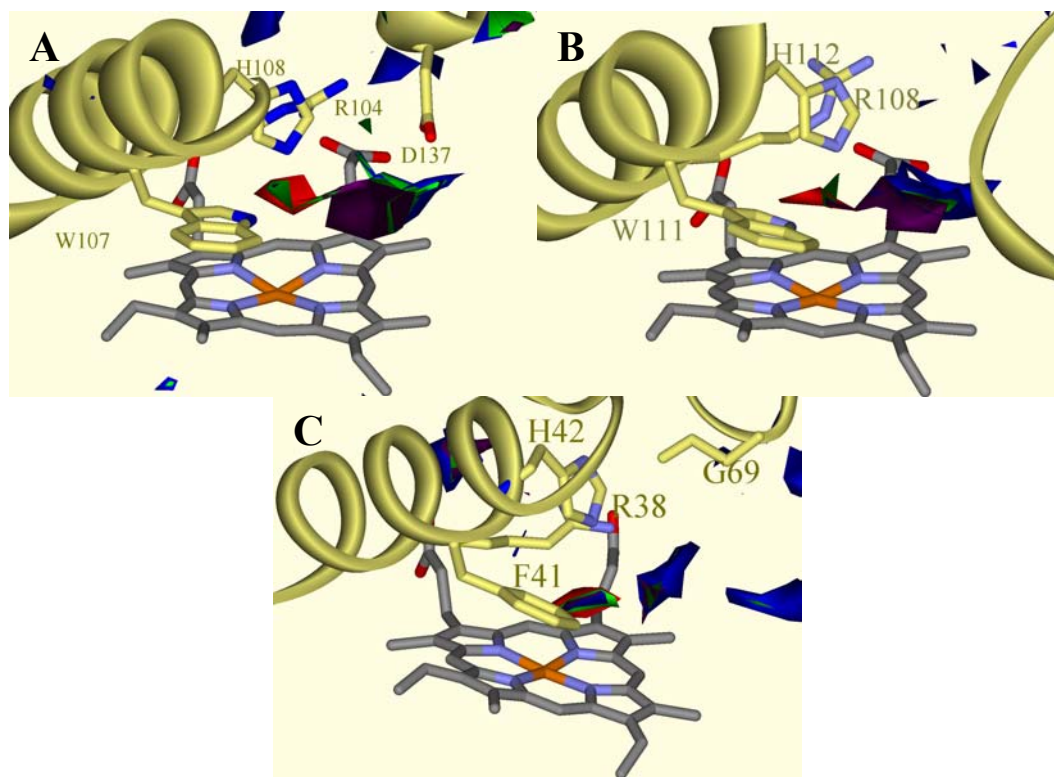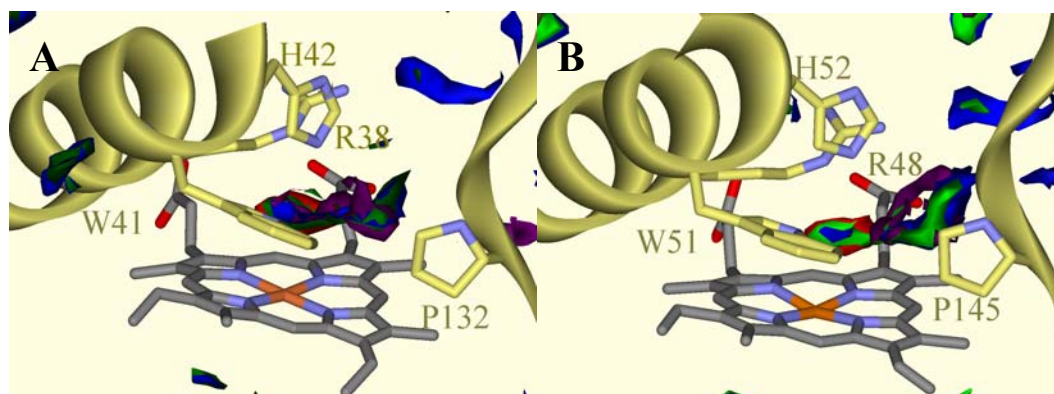
Figure 4.30 GRID maps depicting energetically favorable binding sites for the NH (green), NH$_2$ (blue), carbonyl oxygen (red) and aromatic neutral amide (purple) probes on (A) APX and (B) CcP. GRID map contour levels are set at -8, -9, -12, and -8.5 kcal/mol for the NH, NH$_2$, carbonyl, and aromatic neutral amide probes, respectively. In all panels the heme is in gray and protein mainchains and side chains are depicted in yellow. Interacting amino acid residues are indicated as stick figures. This figure was produced using DS ViewerPro 5.0 (Accelrys, Cambridge, UK).

GRID calculations indicated no obvious differences between the active sites of the classical peroxidases, APX and CcP, and the catalase-peroxidases, *M. tuberculosis* CP and *B. pseudomallei* CP for the probes investigated.   The similarity in the active site as determined by GRID indicates the catalase function is most likely due to differences other then the active site.   The presence of covalently modified tryptophan, methionine, and tyrosine residues in the catalase-peroxidases, but not the classical peroxidases, is one hypothesis for the difference in activity. A second notable difference is the absence of a loop located at the γ-edge of the heme in APX and CcP that is present in all of the catalase-peroxidases

which have been structurally characterized (Figure 4.31). Sequence alignment indicates a clear absence of the 32 (29 for HRPC) residues that constitute the loop present in *M. tubercuolosis* CP, *B. pseudomallei* CP, and *H. marisortui* CP (Figure 4.32).
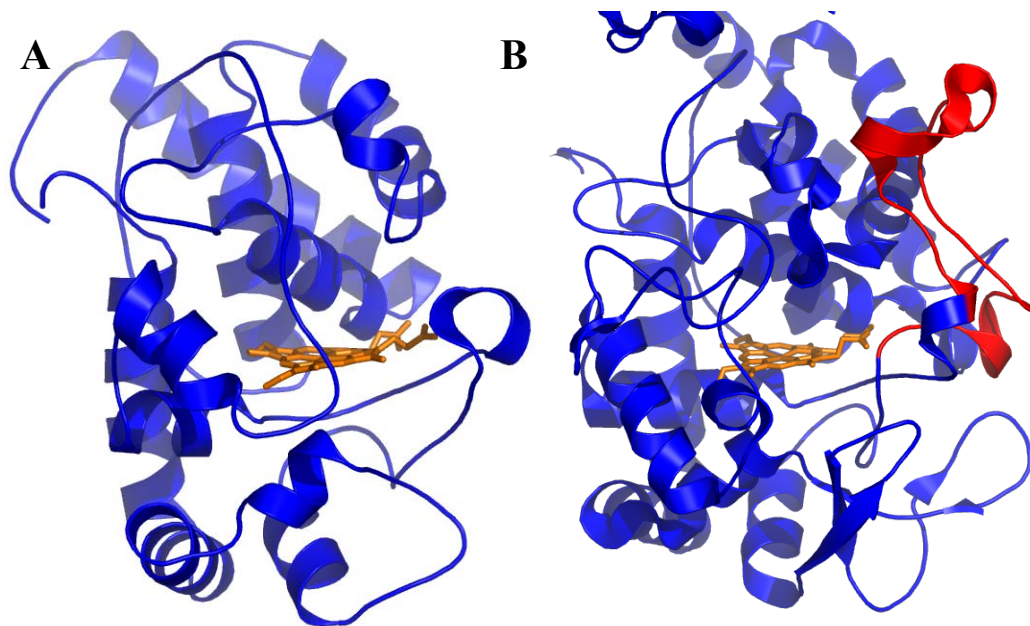
Figure 4.31   Structure of (A) APX and (B) *B. pseudomallei* CP with the "missing" loop indicated in red. Figure was produced in PyMol (DeLano Scientific, San Carlos, California).

```
B. pseduomallei CP    1   MPGSDAGPRRRGVHEQRRNRMSNEAKCPFHQAAGN-GTSNRDWWPNQLDLSILHRH
M. tuberculosis CP    1   -----MPEQHPPITETTTGAASNGCPVVGHMKYPVEGGGNQDWWPNRLNLKVLHQN
H. marismortui CP     1   ----------------MAETPNSDMSGATGGRSKRPKSNQDWWPSKLNLEILDQN
APX                   1   -------------------------------------------------------
CcP                   1   ----------------------------------------------------MIT
1ATJ                  1   -------------------------------------------------------

B. pseduomallei CP   60   DPMGKDFNYAQAFEKLDLAAVKRDLHALMTTSQDWWPADFGHYGGLFIRMAWHSAG
M. tuberculosis CP   56   DPMGAAFDYAAEVATIDVDALTRDIEEVMTTSQPWWPADYGHYGPLFIRMAWHAAG
H. marismortui CP    44   GPVEDDFDYAEEFQKLDLEAVKSDLEELMTSSQDWWPADYGHYGPLFIRMAWHSAG
APX                   1   ---GKSYPTVSPDYQKAIEKAKRKLRGFIAEKK--C-------APLILRIAWHSAG
CcP                   8   HVASVEKGRSYEDFQKVYNAIALKLREDDEYDN--YIG----YGPVLVRLAWHISG
1ATJ                  1   ----QLTPTFYDNSCPNVSNIVRDTIVNELRSDP-------RIAASILRLHFHDCF

B. pseduomallei CP  120   ADGRGGAGEGQQRFAPLNSWPDNAN--IDKARRLLWPIKQKYGRAISWADLLILTG
M. tuberculosis CP  116   HDGRGGAGGGMQRFAPLNSWPDNAS--IDKARRLLWPVKKKYGKKLSWADLIVFAG
H. marismortui CP   104   ADGRGGAAGGRQRFAPINSWPDNAN--IDKARRLLLPIKQKYGQKISWADIMILAG
APX                  49   KTKTGGPFGTIKHQAELAHGANNG---IDIAVRLLEPIKEQFP-IVSYADFYQLAG
CcP                  62   HDNTGGSYGGTYRFKKEFNDPSNAG--LQNGFKFLEPIHKEFP-WISSGDLFSLGG
1ATJ                 50   DASILLDNTTSFRTEKDAFGNANSARGFPVIDRMKAAVESACPRTVSCADILTIAA

B. pseduomallei CP  178   ESMGFKTFGFAGGRADTWEPE-DVYWGSEKIWLELSGGPNSRYSGDRQLENPLAAV
M. tuberculosis CP  174   ESMGFKTFGFGFGRVDQWEPD-EVYWGKEATWLGDE-----RYSGKRDLENPLAAV
H. marismortui CP   162   ESMGFKTFGYAGGREDAFEEDKAVNWGPEDEFETQER-----FDEPGEIQEGLGAS
APX                 105   EITGGPEVPFHPGREDKPEPP--------------------------------
CcP                 119   QEMQGPKIPWRCGRVDTPEDT--------------------------------
1ATJ                110   TLAGGPSWRVPLGRRDSLQAF--------------------------------

B. pseduomallei CP  237   IYVNPEGPDGNPDPVAAARDIRDTFARMAMNDEETVALIAGGHTFGKTHGAG-PAS
M. tuberculosis CP  228   IYVNPEGPNGNPDPMAAAVDIRETFRRMAMNDVETAALIVGGHTFGKTHGAG-PAD
H. marismortui CP   217   IYVNPEGPDGNPDPEASAKNIRQTFDRMAMNDKETAALIAGGHTFGKVHGADDPEE
APX                 125   ----EEG--RLPDATKGSDHLRDVFGKAMGLSDQDIVALSGGHTIGAAHKER-
CcP                 139   ---TLDN-GRLPDADKDAGYVRTFEQRLNMNDRE-VVALMGAHALGKTHLKN-
1ATJ                130   ---LDLANANLPAPFFTLPQLKDSFRNVGLNRSSDLVALSGGHTFGKNQCRFIMDR

B. pseduomallei CP  296   EPEAAGIEAQGLGWKSAYRTGKGADAITSGLEVTWTTTPTQWSHNFFENLFGYEWE
M. tuberculosis CP  287   EPEAAPLEQMGLGWKSSYGTGTGKDAITSGLEVVWTNTPTKWDNSFLEILYGYEWE
H. marismortui CP   277   EPEAAPIEQQGLGWQNKNGNSKGGEMITSGIEGPWTQSPTEWDMGYINNLLDYEWE
APX                 171                            SGFEGPWTSNPLIFDNSYFTELLTGEKD
CcP                 186                            SGYEGPWGAANNVFTNEFYLNLLNEDWK
1ATJ                186                           -FSNTGLPDPTLNTTYLQTLRGLCPLNGNLSA

B. pseduomallei CP  356   PAGAHQWVAKG--ADAVIPDAFDPSKKHRPTMLTTDLSLRFDP---AYEKISRRFH
M. tuberculosis CP  347   PAGAWQYTAKDG-AGAGTIPDPFGGPGRSPTMIATDLSLRVDP---IYERITRRWL
H. marismortui CP   337   PGGAWQWAPKSEELKNSVPDAHDPDEKQTPMMITTDIALKRDP---DYREVMETFQ
APX                 199   --------------------G--------LLQIPSDKALLTDS---VFRPLVEKYA
CcP                 219   DANNEQWDSKSG--------Y---------MMIPTDYSLIQDP---KYLSIVKEYA
1ATJ                222   DLRTPTIFDN---------KYYVNLEEQKGLIQSDQELFSSPNATDTIPLVRSFA

B. pseduomallei CP  411   QFADAFARAWFKLTHRDMGPRARYLGPEVPAEVLLWQDPIPAVDHPLIDAADAAEL
M. tuberculosis CP  403   ELADEFAKAWYKLIHRDMGPVARYLGPLVPKQTLLWQDPVPAVSHDLVGEAEIASL
H. marismortui CP   394   EFGMNLAKAWYKLTHRDMGPPERFLGPEVPDEEMIWQDPLPDADYDLIGDEEIAEL
APX                 229   VFFADYAEAHLKLSELGFAEA----------------------------------
CcP                 259   KFFKDFSKAFEKLLENGITFPKDAPSPFIFKTLEEQGL-----------------
1ATJ                272   TEFNAEVEAMDRMGNITPLTGTQGQIRLNCRVVNS--------------------
```

275

```
B. pseudomallei CP   471 LASGLTVSQLVSTAWAAASTFRGSDKRGGANGARIRLAPQKDWEANQPE-QLAAVL
M. tuberculosis CP   463 RASGLTVSQLVSTAWAAASSFRGSDKRGGANGGRIRLQPQVGWEVNDPDGDLRKVI
H. marismortui CP    454 LDSDLSVSQLVKTAWASASTYRDSDKRGGANGARLRLEPQKNWEVNEPE-QLETVL
APX                  249 -------------------------------------------------------
CcP                  296 -------------------------------------------------------
1ATJ                 306 -------------------------------------------------------

B. pseudomallei CP   530 AIRTAFNGAQRGGKQVSLADLIVLAGCAGVEQAAKNAGHAVTVPFAPGRADASQEQ
M. tuberculosis CP   523 EIQESFNSAAPGNIKVSFADLVVLGGCAAIEKAAKAAGHNITVPFTPGRTDASQEQ
H. marismortui CP    513 NIQTEFNDSRSDGTQVSLADLIVLGGNAAVEQAAANAGYDVEIPFEPGRVDAGPEH
APX                  249 -------------------------------------------------------
CcP                  296 -------------------------------------------------------
1ATJ                 306 -------------------------------------------------------

B. pseudomallei CP   590 SMAVLEPVADGFRNYLKGKYRVPAEVLLVDKAQLLTLSAPEMTVLLGGLRVLGANV
M. tuberculosis CP   583 SFAVLEPKADGFRNYLGKGNPLPAEYMLLDKANLLTLSAPEMTVLVGGLRVLGANY
H. marismortui CP    573 SFDALKPKVDGVRNYIQDDITRPAEEVLVDNADLLNLTASELTALIGGMRSIGANY
APX                  249 -------------------------------------------------------
CcP                  296 -------------------------------------------------------
1ATJ                 306 -------------------------------------------------------

B. pseudomallei CP   650 HGVFTAREQALTNDFFVNLLDMGTEWKPTAADADVFEGRDRATGELKWTGTRVDLV
M. tuberculosis CP   643 LGVFTEASESLTNDFFVNLLDMGITWEPSPADDGTYQGKD-GSGKVKWTGSRVDLV
H. marismortui CP    633 LGVFTDEPETLTNDFFVNLLDMGTEWEPAADSEHRYKGLDRDTGEVKWEATRIDLI
APX                  249 -------------------------------------------------------
CcP                  296 -------------------------------------------------------
1ATJ                 306 -------------------------------------------------------

B. pseudomallei CP   710 SQLRALAEVYGSADAQEKFVRDFVAVWNKVMNLDRFDLA 748
M. tuberculosis CP   702 SELRALVEVYGADDAQPKFVQDFVAAWDKVMNLDRFDVR 740
H. marismortui CP    693 DRLRAISEVYGSADAEKKLVHDFVDTWSKVMKLDRFDLE 731
APX                  249 ------------------------------------- 249
CcP                  296 ------------------------------------- 296
1ATJ                 306 ------------------------------------- 306
```

Figure 4.32 Sequence alignment of *B. pseudomallei* CP, *M. tuberculosis* CP, *H. marismortui* CP, APX, C*c*P, and HRPC. Identical residues are shown in blue, while similar residues are shown in yellow. The loop missing found in the CPs that is not present in APX, C*c*P, and HRPC is highlighted in red.

ICP-AES indicated the presence of Zn^{2+} in *M. tuberculosis* CP and the zinc dication GRID probe predicted an energetically favorable binding site near residues D94, Y95, P100, H274, and H276 (Figure 4.19). The identified potential binding pocket for Zn^{2+} was adjacent to the propionate functional group of the heme moiety (Figure 4.8). Once the crystal structure of *M. tuberculosis* CP was

solved, a potential metal cation site was located in the region predicted by GRID. The distorted tetrahedral geometry of the site suggests that the metal cation might be Zn^{2+}; however other dications such as Ca^{2+} are not precluded. Based on the proximity of the energetically favorable binding site to the heme moiety, the metal cation has the potential to play a structurally important role in maintaining the integrity of the active site.

The three-dimensional X-ray crystal structure of *M. tuberculosis* CP has characterized an additional five residues at the N-terminal of the structure (Figure 4.25). The "hooked" structure of the additional residues interacts with a loop of the N-terminal domain of the other monomer. Interestingly, a hydrophobic pocket is created by the π-stacking of two tyrosine residues, Y28 and Y197. These contacts may play a vital role in the dimerization of the enzyme.

Presence of the continuous electron density between Y229, W107, and M255 in *M. tuberculosis* CP indicates that this may be a common feature of all catalase peroxidases and may contribute to their catalase activity. The catalase activity of the catalase peroxidases is lost upon mutation of the tryptophan residue.[12, 68] The exact function of the covalent modification is unknown; however, one hypothesis is that it stabilizes the enzyme by reducing the likelihood of damaging oxidations.[67]

The electrostatic environment of the active site in the vicinity of the proximal H276, D381, and W328 has a greater similarity to CcP than to APX. Formation of a tryptophan radical as the second oxidizing species of compound I is known for CcP.[78] The electrostatic similarity in the active site may imply that the compound I intermediate of catalase peroxidases are formed *via* a mechanism similar to that in CcP, rather than APX.

The X-ray crystal structure of *M. tuberculosis* CP indicates an access channel to the active site which extends from the γ-edge of the heme moiety to the exterior of the enzyme. A continuum of water molecules is located in the funnel-shaped channel (Figure 4.28). The mutation S315T, which is located along the access channel, results in the loss of the activation of INH with no loss of peroxidase or catalase activity.[43] It seems likely that INH accesses the heme moiety *via* the channel, which is small enough to prevent entry of larger, unwanted substrate molecules.

CONCLUSIONS AND FUTURE WORK

Based on GRID calculations, the INH binding site in HRPC, *B. pseduomalle* CP, and *M. tuberculosis* CP is predicted in a similar location above the δ-edge of the heme moiety. In Class I peroxidases, APX and C*c*P, GRID analysis of INH-like probes yielded results similar to those found for HRPC and the catalase peroxidases.

The successful crystallization and structural characterization of *M. tuberculosis* CP has provided a plethora of useful information. The presence of a loop, which partially blocks entrance into the substrate access channel, is present in the catalase peroxidase, but absent in the classical peroxidases suggesting a possible correlation between the loop and catalase activity. A metal cation binding site was located near the γ-edge of the heme and may play a role in maintaining the structural integrity of the active site. Five additional residues at the N-terminal may play a role in the dimerization of the enzyme. Finally, the region of covalent modification between residues W107, Y229, and M255 appears to be a common feature of all catalase peroxidases and potentially impacts the catalase activity.

A focus for future work involving *M. tuberculosis* CP, should include crystallization studies to trap isoniazid or an isoniazid analogue, such as benzhydroxamic acid, in the active site of the enzyme. Structurally

characterization of such a system will provide valuable information for the design of more potent anti-tuberculosis treatment. Additionally, kinetic studies to evaluate the impact of a metal chelator, such as ethylenediamminetetraacetate (EDTA), on catalase and peroxidase activities should be performed to determine the likelihood of a metal cation binding site near the γ-edge of the heme.

EXPERIMENTAL SECTION

General Synthetic Procedures

Chemicals were purchased from Sigma-Aldrich (Poole, UK). Protein molecular weight standards and Coomassie Brilliant Blue were from Amersham Pharmacia Biotech (Little Chalfont, UK). Media reagents were purchased from Merck (Lutterworth, UK) and were made according to published protocols [79]. Media were sterilized by autoclaving and stored at room temperature. Luria broth (LB) contained 1% (w/v) tryptone, 0.5% (w/v) yeast extract, and 0.5% (w/v) NaCl in distilled water. Luria agar (L-agar) consisted of LB with the addition of 1.5% (w/v) agar. 2xYT containted 1.6% (w/v) tryptone, 1.0% (w/v) yeast extract, and 0.5% (w/v) NaCl in distilled water. Antibiotics were added as required at concentrations of 100 μg/mL ampicillin and 10 μg/mL tetracycline.

Several groups have used *Escherichia coli* to express the *katG* gene of *M. tuberculosis* including Zhang *et al.,*[5] Johnsson *et al.,*[37] Wengenack *et* al.,[43] Hillar and Loewen,[42] Nagy *et al.,*[80] and, most recently, Saint-Joanis *et al.*[81] Using similar methods, *E. coli* UM255 was grown in either LB or 2xYT. Since *E. coli* UM255 contains the tetracycline resistant Tn*10* transposon in its *katE* gene, as well as a mutated *katG* gene it lacks endogenous catalase and peroxidase activity. This *E. coli* strain was used for expression of the enzyme with catalase and peroxidase activity.

pTBCP[80] (Figure 4.33), a gift from Dr. J. M. Nagy, Imperial College London, was constructed by ligation of a PCR-generated fragment from pYZ56[5], a pUC19 derivative containing the *M. tuberculosis katG*, encoding catalase-peroxidase, into the *Eco*RI and *Hind*III sites of p*Trc*99A. The translated gene product, *M. tuberculosis*, utilizes the ATG start codon contained within the *Nco*I restriction site resulting in three additional amino acid residues at the N-terminus upstream of the mycobacterial start codon (GTG) translated as valine. The resulting N-terminal sequence is M-E-F-V.
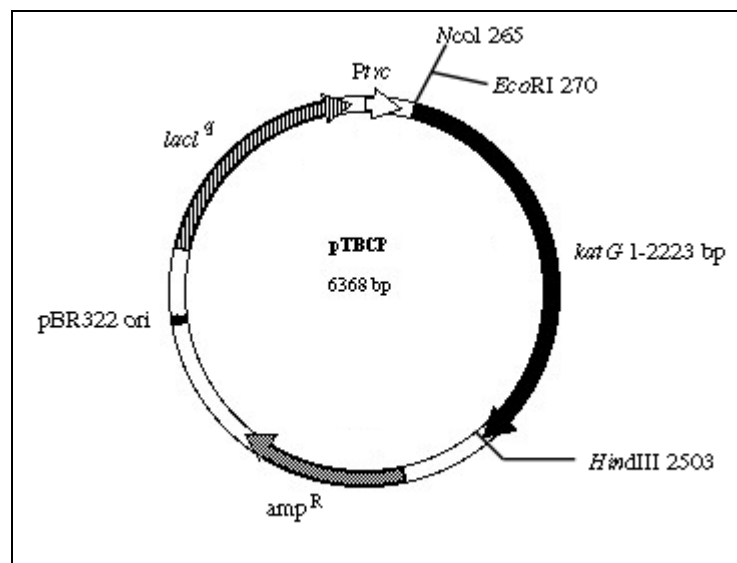

Figure 4.33 Plasmid diagram for pTBCP.

Crude Cell Extract

An L-agar plate containing 100 µg/mL ampicillin and 10 µg/mL tetracycline streaked from a glycerol stock solution of *E. coli* UM255[pTBCP].

282

20 mL of LB containing 100 μg/mL ampicillin and 10 μg/mL tetracycline were inoculated with one colony of *E. coli* UM255. The cultures were grown overnight at 37 °C with shaking; each LB culture was inoculated into 1 L of 2xYT containing 100 μg/mL ampicillin and 10 μg/mL tetracycline. Cells were grown at 37 °C with shaking until a cell density corresponding to an OD_{550} of 0.4-0.6 was reached at which point they were induced with IPTG to a final concentration of 0.2 mg/mL. Growth was continued for an additional 12 h. Cells were harvested by centrifugation at 7500 *g*, at 4 °C for 50 min. Cells were resuspended in 10 mL of KH_2PO_4/K_2HPO_4 phosphate buffer (pH 6.5).

Heme Addition

The level of heme incorporation into CP enzymes has been shown to vary from sub-stoichiometric to stoichiometric. The Reinheitzahl value ($R_z : A_{Sôret\ band}/A_{280\ nm}$), which is typically between 0.2 and 0.7, is used as a measure of purity. The inconsistent incorporation of heme into the CP enzyme is believed to be due to overproduction of the enzyme in *E. coli* resulting in normal levels of heme biosynthesis being insufficient.[23]

Alteration of a previously reported method was used for heme addition [82]. A 24.5 mM solution of hemin chloride solution was prepared using Milli-Q water. In order to completely dissolve the hemin chloride, the pH was raised to around 12 by the addition of 1-2 μL of 1 M NaOH. The final pH of the solution was

approximately 6.0. A 1 to 25 v/v ratio of heme solution was added to the resuspended cells at 4 °C and mixed well by vortexing to increase the stoichiometric amount of holoenzymes present. Cells were lysed by sonication with 3 x 30 sec bursts on power level 8, using a microtip and an XL2020 sonicator (Labcaire Systems, Clevedon, UK). Insoluble material was removed by centrifugation at 10000 g at 4 ˚C for 50 min. The supernatant was treated with 100 μg/mL DNAse and 10 μg/mL RNAse for 1 h at 4 ˚C and then recentrifuged as just described. Free hemin was removed from the solution by subsequent centrifugation and chromatography.

DEAE-Sepharose Anion-Exchange Chromatography

A Pharmacia XK-16/40 column containing 25 mL of DEAE-Sepharose (Fast Flow) attached to a Pharmacia Gradifrac system was equilibrated with 100 mM KH_2PO_4/K_2HPO_4 pH 6.6 buffer at a flow rate of 2.0 mL/min. The supernatant obtained from the cell extraction was passed through a Millipore 0.45-μm filter and loaded onto the column at a flow rate of 2.0 mL/min. The column was washed with 30 mL of 100 mM KH_2PO_4/K_2HPO_4 (pH 6.5) and *M. tuberculosis* CP was eluted with a 150-mL linear gradient of 0.0-0.6 M NaCl in 100 mM KH_2PO_4/K_2HPO_4, pH 6.5 at a flow rate of 3 mL/min. 2.0-mL fractions were collected between 0.25 and 0.40 M NaCl elution and pooled. The samples

were analyzed by SDS polyacrylamide gel electrophoresis for the presence of the enzyme and pooled.

Gel Filtration

Fractions obtained from the DEAE-Sepharose anion-exchange column were further purified using a Pharmacia XK-16/70 column containing 100 mL Superose 6 Prep Grade. The column was equilibrated with phosphate buffered saline [PBS: 0.01 M KH_2PO_4/K_2HPO_4 buffer, 0.0027 M KCl, 0.137 M NaCl (pH 7.5)] overnight at a flow rate of 0.5 mL/min. The fractions were passed through a Millipore 0.45-μm filter and loaded onto the column with a flow rate of 0.5 mL/min. 1.0-mL fractions were collected and fractions from the absorbance maximum region were pooled.

Resource Q Anion-Exchange Chromatography

For further purification, the fractions were loaded onto a Pharmacia Resource Q column (6 mL) at a flow rate of 2.0 mL/min, which had first been equilibrated with 30 mL of 100 mM KH_2PO_4/K_2HPO_4 (pH 6.5). The enzyme, *M. tuberculosis* CP, was eluted with a 30-mL linear gradient of 0-1.0 M NaCl in 100 mM KH_2PO_4/K_2HPO_4 (pH 6.5) at a flow rate of 2.0 mL/min. 1.0-mL fractions were collected and pooled.

Protein Concentration Determination

Concentration of *M. tuberculosis* CP was determined using a Shimadzu UV-1601 UV-visible spectrophotometer and the extinction coefficient values for wild-type heme-added *M. tuberculosis* CP: $\varepsilon_{280} = 1.5 * 10^5$ M^{-1}cm^{-1} and $\varepsilon_{405} = 3.4 * 10^5$ M^{-1}cm^{-1} [82].

HiTrap Desalt Column

To replace the NaCl containg buffer with KH_2PO_4/K_2HPO_4 buffer, a HiTrap Desalt column was used (Amersham Pharmacia Biotech, Little Chalfont, UK). The column was equilibrated with 25 mL of KH_2PO_4/K_2HPO_4 (pH 6.5) at a flow rate of approximately 5 mL/min. The 3.3 mg/mL sample of *M. tuberculosis* CP was loaded onto the column in 1.5 mL aliquots at a flow rate of 5 mL/min. Elution of the protein with the desired buffer yielded 2 mL of $2.0*10^{-4}$ mM (0.016 mg/mL) sample.

SDS Polyacrylamide Gel Electrophoresis

Analysis of the protein was performed using hand-poured 80 mm x 80 mm x 1mm gels consisting of 6 mL resolving gel layered by 2 mL stacking gel. The resolving gel was prepared by 0.1% sodium dodecyl sulphate (SDS) and 12% polyacrylamide,[83] which polymerized in the presence of 0.04% (v/v) tetraethylmethyldiamine and 0.1% (w/v) ammonium persulphate. The

stacking gel was prepared by the polymerisation of 0.1% sodium dodcyl sulphate (SDS) and 5% polyacrylamide in the presence of 0.1% (v/v) tetraethylmethyldiamine and 0.1% (w/v) ammonium persulphate).[83]   The gels were ran on a Novex minigel system (Invitrogen, Groningen, Netherlands). Protein samples were mixed in a 1 in 6 ratio with 5x concentrated protein sample buffer [250 mM Tris-HCl (pH 6.8), 10% (w/v) SDS, 500 mM Dithiothreitol (DTT), 0.5% (w/v) bromophenol blue, 50% (v/v) glycerol]. Before loading the proteins on the gels, samples were boiled for 5 min. The gels were stained in 0.08% (w/v) Coomassie Brilliant Blue, 10% (v/v) acetic acid, and 30% (v/v) methanol for 1 h to aid visualization. Destaining solution, 30% (v/v) ethanol in 10% acetic acid, was used to remove stain not bound to the protein.

Induced Coupled Plasma Atomic Emission Spectroscopy

Purified *M. tuberculosis* CP was analyzed by induced coupled plasma atomic emission spectroscopy (ICP-AES) using a Fisons ARL 3580B to determine Li, Na, K, Be, Mg, Ca, Sr, Ba, Al, La, Ti, V, Cr, Mo, Mn, Fe, Co, Ni, Cu, Ag, Zn, Cd, Pb, P, As, and S content by Dr. G. Fowler (Imperial College, London, UK).   For each measurement, two samples were analyzed, one containing only the acid matrix and the second containing both protein and acid matrix. Samples were prepared in 25 mL round bottom flasks fitted with rubber septa and vented using 2 stainless steal 20 gauge needles.   1-mL 2.5 mg/mL

protein (KH_2PO_4/K_2HPO_4, pH 6.5) in 2 mL concentrated nitric acid and 0.5 mL perchloric acid was heated until a heating mantle for 3 h at 50 ˚C while stiring. The temperature was increased to 150 ˚C and the samples heated for an additional 3 h. Both samples were heated to dryness at 190 ˚C. The remaining solid residue was dissolved in 2 mL of 5 M hydrochloric acid and heated for 1 h at 60 ˚C. The sample was diluted to 10 mL with Mili-Q water followed by filtration through a 0.45-μm filter.

GRID Calculations

All calculations were performed using the Linux operating system and the GREAT, GRIN, and GRID software package [84] version 21. Structure files were obtained from the PDB [85], with the exception of the *M. tuberculosis* model, which was provided by N. A. J. Eady.[74] All remarks and water molecules were removed from the protein database file before processing.

The HRPC PDB file and the INH-docked NMR model of HRPC, obtained from Dr. L. Banci,[51] were superpositioned by Alwyn Jones' O software[86] to according to their C^{α} position. This superpositioning aided in visually comparing the calculated position of INH by GRID to that determined in the NMR model. Protein residues missing from the protein database file were modeled in idealized positions according to the location of the side chain using Swiss PDBViewer.[72]

The files were prepared for use in GRID using program GRIN. The charge of the protein was calculated in GRIN and counterions of Na^+ or Cl^- were added appropriately to create a charge balanced structure. To determine the best placement positions for the counterions, a GRID calculation was performed using the Na^+ or Cl^- probe to locate the most energetically favorable binding site for the counterions. The counterions were then manually entered in suggested positions away from the active site and set to move in response to the probe. When the probe is moving around the GRID and moves into the vicinity of the counterion, the counterion reacts and shifts slightly to avoid being "seen" by the probe. Moving counterions prevent a large charged area from being present during the experiment.

All GRID calculations were performed on the charge balanced protein coordinate files with the following directives: NPLA (Number of Planes) equal to 2 thus the GRID points were 0.5 Å apart to improve the accuracy of the contour surfaces and MOVE equal to -1 to allow the counterions to be flexible with respect to the probe.

Probes used in the course of the experiment included both single-atom multi-atom types. The carbonyl probe was designed to locate regions of favorable binding energy toward sp^2 carbonyl oxygens, while the carboxy probe was modeled as an sp^2 carboxy oxygen atom. The N1 and N2 probes were used to

model NH and NH$_2$ functional groups. Two multi-atom probes, AR.CONHR and CONHR, modeled a larger portion of the INH structure. The aromatic neutral amide probe was set so that the NH group was *trans* with respect to the carbonyl oxygen and was used to locate energetically favorable binding sites for aromatic amides. The aliphatic neutral amide probe, CONHR, was also set in a *trans* orientation and searched for sites which would be energetically favorable for aliphatic amide binding. The zinc and calcium probes were used to model the zinc and calcium dication, respectively.

Visualization of Data

The output data generated by GRID were read into MSI Insight II[87] software version 2000 for visualization of the energy maps. The data map files were output from Insight II as ascii files and reopened in DS Viewer Pro v. 5.0 for generation of the graphics.

Crystallization of M. tuberculosis CP

All buffers were filtered through a 0.45-µl filter prior to use in crystallization trials. *M. tuberculosis* CP sample fractions from the Resource Q purification step were pooled. The protein was buffer exchanged into KH$_2$PO$_4$/K$_2$HPO$_4$ (pH 6.5) using a HiTrap Desalt Column. Using a Centricon-30 protein concentrator, the protein sample was concentrated to 22 mg/mL. The

protein concentration was determined according to the Beer-Lambert law using the extinction coefficient $\varepsilon_{405} = 3.4 * 10^5$ $M^{-1}cm^{-1}$ [82]. The concentrated protein was used to set up hanging-drop[88] crystallization trays with Hampton Research Crystal Screen Kits 1 and 2. Multi-well crystallization plates were used with each well containing 500 µl of reservoir solution. 2-µl protein droplets were placed on siliconized glass cover slips and were mixed with 2 µl reservoir solution. Each cover slip was inverted onto the reservoir solution-filled wells of the multi-well plate, the opening of which had been encircled with vacuum grease to provide a seal. The trays were placed in the dark at 20 °C.

Data Collection and Processing

Diffraction data were collected at the Synchrotron Radiation Source, Daresbury, UK using an Area Detector System Corporation Quantum-4 CCD detector and 0.877 Å radiation. Crystals were prepared for cryocooling by soaking in a cryoprotection buffer containing the components of the reservoir solution for the crystal being tested and 20% glycerol. A single crystal was mounted in a nylon cryoloop and maintained at 100 K using a nitrogen cryostream.

Data sets were processed using the HKL package version 1.97.2,[75] specifically *Denzo* and *XdisplayF*. The peak search option in *XdisplayF* was used to identify and select desired peaks from the oscillation image. The peaks were

autoindexed using *Denzo*[89] and ideal cells for all 14 Bravais lattices were calculated.

The crystal and detector parameters were refined for the displayed oscillation image using a least-squares method, which allows each parameter to be refined in turn. During the refinement process, mosaicity (the smallest angle through which the crystal can rotate, about any axis or combination of axes, while a reflection is still observed),[90] and the spot size and shape are estimated by the user. These refined parameters are then applied to all images of the data set. During this process, *Denzo*[89] looks for spots on all images of the data set and uses the profile-fitting method to estimate the intensities of spots and subsequently calculates the reflection intensities. To calculate the diffraction intensity, the background must be estimated and then subtracted from the reflection profile. The profile-fitting method within *Denzo*[89] is a two-step process. First a profile is predicted based on the profiles of other reflections. Secondly, the diffraction intensity of each spot is calculated by estimating and then subtracting the background from the reflection profile.

Data Scaling and Merging

Scaling and merging of data sets was performed using *Scalepack* from the HKL software package.[75] The program calculates single, isotropic scale and *B* factors for each of the images of the processed data. All images are put on a common

scale, thereby taking into account changes in beam intensity. *Scalepack* also performs postrefinement, the implementation of which enables improved unit cell parameters, mosaicity and crystal orientation of each image, but with the same unit cell value for the whole data set. Postrefinement was run two times, first to write a list of outlying observations and on following rounds these observations were rejected in the refinement of the orientation parameters. Following postrefinement, *Scalepack* merges the data. All related spots, for example, partial reflections, symmetry related spots and Friedel pairs, are merged together and an R_{merge} value is calculated. The quality of the diffraction data can by analyzed by the agreement between symmetry related reflections, the R_{merge}, or the ratio of the intensity to the error of the intensity, $I/\sigma(I)$.

Generation of FreeR Flag

The output reflection file was converted to MTZ format using the *f2mtz* program from the Collaborative Computational Project, Number 4 (CCP4 Program Suite v. 4.1). The resulting MTZ file was processed using *truncate* which calculates a Wilson plot, uses the resulting scale factor to place all reflections on an absolute scale and deal with weak reflections by ensuring no reflection is negative. A best estimate of $|F(hkl)|$ from I, the standard deviation of I and the distribution of intensities in resolution shells are calculated. All negative reflections are forced to be positive and the weakest reflections (less than three

standard deviations) are inflated. The program *freerflag* from the CCP4 Program Suite then designated 5% of the data as a test set for subsequent generation of the *R*-free value.[91]

Molecular Replacement

Molecular replacement was performed using the automated program *molrep*[92] from the CCP4 Program Suite. The program is based upon the fast rotation function[93] and uses the translation function defined as the overlap of the observed and calculated Patterson functions for a given orientation of the model.[92]

M. tuberculosis CP was found to have 66% sequence identity to *B. pseudomallei* CP and was expected to be similar in structure. The crystal structure of *B. pseudomallei* CP[67] was used as a search model for the molecular replacement process.

Refinement

Following molecular replacement, the model was underwent several cycles of refinement in order to obtain closer agreement between the calculated and observed structure factors. The agreement index between calculated and observed structure factors is represented by *R,* where *k* represents a weighting factor (Equation 4. 1).

$$R = \frac{\sum\limits_{hkl} \left\| F_{\mathrm{obs}} \right| - k \left| F_{\mathrm{calc}} \right\|}{\sum\limits_{hkl} \left| F_{\mathrm{obs}} \right|} * 100\%$$

Equation 4. 1 Formula for determining the reliability index of a structure.

The macromolecular structure determination software, Crystallography and NMR System (*CNS*),[91] was used for structure refinement. Refinement within *CNS* utilizes maximum likelihood with cross-validation, in order to reduce over-fitting, which in turn is combined with simulated annealing.[91] Between each refinement step the model was inspected and manually rebuilt where necessary in the molecular graphics program *O*.[86] Observation of the $3F_{\mathrm{o}}$-$2F_{\mathrm{c}}$ Fourier maps indicated when the selection of alternative side chain rotamers and adjustment of torsion angles was required.

Appendix. Compound Numbers

$[(\eta^5\text{-}Me_5C_5)Ge][B(C_6F_5)_4]$ $[1][B(C_6F_5)_4]$

$[(\eta^5\text{-}Me_5C_5)Sn][B(C_6F_5)_4]$ $[2][B(C_6F_5)_4]$

$[(\eta^5\text{-}Me_5C_5)_3Sn_2][B(C_6F_5)_4]$ $[3][B(C_6F_5)_4]$

$[(\eta^5\text{-}Me_5C_5)Pb][B(C_6F_5)_4]$ $[4][B(C_6F_5)_4]$

$[(\eta^5\text{-}Me_5C_5)_3Pb_2][B(C_6F_5)_4]$ $[5][B(C_6F_5)_4]$

$[(\eta^5\text{-}\mu\text{-}Me_5C_5)In_2][B(C_6F_5)_4]$ $[6][B(C_6F_5)_4]$

$[(\eta^5\text{-}\mu\text{-}Me_5C_5)Tl_2][B(C_6F_5)_4]$ $[7][B(C_6F_5)_4]$

$[(Me_5C_5)H_2][B(C_6F_5)_4]$ $[8][B(C_6F_5)_4]$

$[1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)]Li\cdot2THF$ 9

$[1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)]_2Sn]$ 10

$[1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)](\eta^5\text{-}C_5Me_5)Sn$ 11

$[1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)](\eta^5\text{-}C_5Me_5)Pb$ 12

$[1,3\text{-}(Me_3Si)_2(\eta^5\text{-}C_9H_5)]_2Li_2\cdot Et_2O$ 13

REFERENCES

1. Bleed, D.; Watt, C.; Dye, C., *Global Tb Control*. 2001, World Health Organization Report.
2. Schatz, A.; Waksman, S. A., *P. Soc. Exp. Biol. Med.* **1944,** *57,* 244.
3. Maher, D.; al., e., **1997,** *WHO/TB/97.220,*
4. Frieden, T. R.; Sterling, T.; Mendez, A. P.; Kiburn, J. O.; Cauthen, G. M.; Dooley, S. W., *N. Engl. J. Med.* **1993,** *328,* 521.
5. Zhang, Y.; Heym, B.; Allen, B.; Young, D.; Cole, S., *Nature* **1992,** *358,* 591.
6. Youatt, J., *Am. Rev. Respir. Dis.* **1969,** *99,* 729.
7. Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S. V.; Eiglmeier, K.; Gas, S.; Barry, C. E.; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Barrell, B. G., *Nature* **1998,** *393,* 537.
8. Altamirano, M.; Marostenmaki, J.; Wong, A.; FitzGerald, M.; Black, W. A.; Smith, J. A., *J. Infect. Dis.* **1994,** *169,* 1162.
9. Stoeckle, M. Y.; Guan, L.; Riegler, N.; Weitzman, I.; Kreiswirth, B.; Kornblum, J.; Laraque, F.; Riley, L. W., *J. Infect. Dis.* **1993,** *168,* 1063.
10. Hochman, A.; Shemesh, A., *J. Biol. Chem.* **1987,** *262,* 6871.
11. Regelsberger, G.; Jakopitsch, C.; Ruker, F.; Krois, D.; Peschek, G. A.; Obinger, C., *J. Biol. Chem.* **2000,** *275,* 22854.
12. Hillar, A.; Peters, B.; Pauls, R.; Loboda, A.; Zhang, H.; Mauk, A. G.; Loewen, P. C., *Biochemistry* **2000,** *39,* 5868.
13. Regelsberger, G.; Jakopitsch, C.; Engleder, M.; Ruker, F.; Peschek, G. A.; Obinger, C., *Biochemistry* **1999,** *38,* 10480.
14. Kaim, W.; Schwederski, B., *Bioinorganic Chemistry*. 1991, Chichester, UK: Wiley.
15. Dawson, J. H., *Science* **1988,** *240,* 433.
16. Schulz, C. E.; Devaney, P. W.; Winkler, H.; Debrunner, P. G.; Doan, N.; Chiang, R.; Rutter, R.; Hager, L. P., *FEBS Lett.* **1979,** *103,* 102.
17. Moss, T. H.; Ehrenburg, A.; Bearden, A. J., *Biochemistry* **1969,** *8,* 4159.
18. Dunford, H. B., *Peroxidases in Chemistry and Biology*. 1991. 2.
19. Erman, J. E.; Vitello, L. B.; Mauro, J. M.; Kraut, J., *Biochemistry* **1989,** *28,* 7992.
20. English, A. M.; Tsaprailis, T., in *Advances in Inorganic Chemistry*. 1995, Academic Press: New York. p. 79.
21. Yonetani, T.; Anni, H., *J. Biol. Chem.* **1987,** *262,* 9547.

22. Mauro, J. M.; Fishel, L. A.; Hazzard, J. T.; Meyer, T. E.; Tollin, G.; Cusanovich, M. A.; Kraut, J., *Biochemistry* **1988,** *27,* 6243.

23. Chouchane, S.; Lippai, I.; Magliozzo, R. S., *Biochemistry* **2000,** *39,* 9975.

24. Chouchane, S.; Girotto, S.; Yu, S.; Magliozzo, R. S., *J. Biol. Chem.* **2002,** *277,* 42633.

25. Jakopitsch, C.; Regelsberger, G.; Furtmuller, P. G.; Ruker, F.; Peschek, G. A.; Obinger, C., *Biochem. Biophys. Res. Comm.* **2001,** *287,* 682.

26. Wengenack, N. L.; Jensen, M. P.; Rusnak, F.; Stern, M. K., *Biochem. Biophys. Res. Comm.* **1999,** *256,* 485.

27. Magliozzo, R. S.; Marcinkeviciene, J. A., *J. Am. Chem. Soc.* **1996,** *118,* 11303.

28. Magliozzo, R. S.; Marcinkeviciene, J. A., *J. Bio. Chem.* **1997,** *272,* 8867.

29. Bodiguel, J.; Nagy, J. M.; Brown, K. A.; Jamart-Gregoire, B., *J. Am. Chem. Soc.* **2001,** *123,* 3832.

30. Middlebrook, G.; Cohn, M. L.; Scheffer, W. B., *Am. Rev. Tuberc.* **1954,** *70,* 852.

31. Youatt, J., *Aust. J. Exp. Biol. Med.* **1960,** *38,* 245.

32. Youatt, J., *Aust. J. Chem.* **1961,** *14,* 308.

33. Middlebrook, G., *Am. Rev. Tuberc.* **1952,** *69,* 471.

34. Zhang, Y.; Garbe, T.; Young, D., *Mol. Microbiol.* **1993,** *8,* 521.

35. Johnsson, K.; Schultz, P. G., *J. Am. Chem. Soc.* **1994,** *116,* 7425.

36. Shoeb, H. A.; B. U. Bowman, J.; Ottolenghi, A. C.; Merola, A. J., *Antimicrob. Agents Ch.* **1985,** *27,* 399.

37. Johnsson, K.; Froland, W. A.; Schultz, P. G., *J. Biol. Chem.* **1997,** *272,* 2834.

38. Mdluli, K.; Slayden, R. A.; Zhu, Y.; Ramsaswamy, S.; Pan, X.; Mead, D.; Crane, D. D.; Musser, J. M.; Barry, C. E., *Science* **1998,** *280,* 1607.

39. Wengenack, N. L.; Lopes, H.; Kennedy, M. J.; Tavares, P.; Pereira, A. S.; Moura, I.; Moura, J. J. G.; Rusnak, F., *Biochemistry* **2000,** *39,* 11508.

40. Zabinski, R. F.; Blanchard, J. S., *J. Am. Chem. Soc.* **1997,** *119,* 2331.

41. Takayama, K.; Davidson, L. A., *Antibiotics* **1979,** *5,* 98.

42. Hillar, A.; Loewen, P. C., *Arch. Biochem. Biophys.* **1995,** *323,* 438.

43. Wengenack, N. L.; Uhl, J. R.; Amand, S. A. L.; Tomlinson, A. J.; Benson, L. M.; Naylor, S.; Kline, B. C.; Cockerill, F. R.; Rusnak, F., *J. Infect. Dis.* **1997,** *176,* 722.

44. van der Walt, B. J.; van Zyl, J. M.; Kriegler, A., *Biochem. Pharmacol.* **1994,** *47,* 1039.

45. Schoeb, H. A.; Jr., B. U. B.; Ottolenghi, A. C.; Merola, A. J., *Antimicrob. Agents Ch.* **1985,** *27,* 408.

46. Schoeb, H. A.; Jr., B. U. B.; Ottolenghi, A. C.; Merola, A. J., *Antimicrob. Agents Ch.* **1985,** *27,* 399.
47. Hill, H. A.; Thornalley, P. J., *FEBS Lett.* **1981,** *125,* 235.
48. Winder, F. G.; Denneny, J. M., *Biochem. J.* **1959,** *73,* 500.
49. Todorivi, S.; Juranic, N.; Macura, S.; Rusnak, F., *J. Am. Chem. Soc.* **1999,** *121,* 10962.
50. Wengenack, N. L.; Todorovic, S.; Yu, L.; Rusnak, F., *Biochemistry* **1998,** *37,* 15825.
51. Banci, L.; Bodiguel, J.; Brown, K. A.; Eady, N. A. J.; Jamart-Gregoire, B.; Jones, J. N.; Pierattelli, R., **in preparation,**
52. Veitch, N. C., *Biochem. Soc. T.* **1995,** *23,* 232.
53. Welinder, K. G., *Biochim. Biophy. Acta* **1992,** *1080,* 215.
54. Gajhede, M.; Schuller, D. J.; Henricken, A.; Smith, A. T.; Poulos, T. L., *Nat. Struct. Biol.* **1997,** *4,* 1032.
55. Schuller, D. J.; Ban, N.; van Huystee, R. B.; McPherson, M.; Poulos, T. L., *Structure* **1996,** *4,* 311.
56. Smith, A. T.; Santama, N.; Dacey, S.; Edwards, M.; Bray, R. C.; Thorneley, R. N.; Burke, J. F., *J. Biol. Chem.* **1990,** *265,* 13335.
57. Schonbaum, G. R., *J. Biol. Chem.* **1973,** *248,* 502.
58. Aitken, S. M.; Turnbull, J. L.; Percival, M. D.; English, A. M., *Biochemistry* **2001,** *40,* 13980.
59. Poulos, T. L.; Freer, S. T.; Alden, R. A.; Edwards, S. L.; Skogland, U.; Takio, K.; Eriksson, B.; Xuong, N.; Yonetani, T.; Kraut, J., *J. Biol. Chem.* **1980,** *255,* 575.
60. Patterson, W. R.; Poulos, T. L., *Biochemistry* **1995,** *34,* 4331.
61. Mandelman, D.; Jamal, J.; Poulos, T. L., *Biochemistry* **1998,** *37,* 17610.
62. Hill, A. P.; Modi, S.; Sutcliffe, M. J.; Turner, D. D.; Gilfoyle, D. J.; Smith, A. T.; Tam, B. M.; Lloyd, E., *Eur. J. Biochem.* **1997,** *248,* 347.
63. Lad, L.; Mewies, M.; Raven, E. L., *Biochemistry* **2002,** *41,* 13774.
64. Sharp, K. H.; Mewies, M.; Moody, P. C.; Raven, E. L., *Nat. Struct. Biol.* **2003,** *10,* 303.
65. Yonetani, T.; Ray, G. S., *J. Biol. Chem.* **1965,** *240,* 4503.
66. Yamada, Y.; Fujiwara, T.; Sato, T.; Igarashi, N.; Tanaka, N., *Nature* **2002,** *9,* 691.
67. Carpena, X.; Loprasert, S.; Mongkolsuk, S.; Switala, J.; Loewen, P. C.; Fita, I., *J. Mol. Biol.* **2003,** *327,* 475.
68. Zamocky, M.; Regelsberger, G.; Jakopitsch, C.; Obinger, C., *FEBS Lett* **2001,** *492,* 177.
69. Kraulis, J., *J. Appl. Crystallogr.* **1991,** *24,* 946.
70. Merritt, E. A.; Murphy, M. E. P., *Acta Crystallogr.* **1994,** *D50,* 869.

71. Powers, L.; Hillar, A.; Loewen, P. C., *Biochem. Biophys. Acta* **2001,** *1549,* 44.
72. Guex, N.; Peitsch, M. C., *Electrophoresis* **1997,** *18,* 2714.
73. Cason, C., *Persistence of Vision Raytracer (Pov-Ray).* 1991, Persistence of Vision Development Team.
74. Eady, N. A. J., **in preparation.,**
75. Otwinowski, Z.; Minor, W., *Processing of X-Ray Diffraction Data Collected in Oscillation Mode*, in *Methods in Enzymology*, C. W. C. Jr. and R. M. Sweet, Editors. 1997, Academic Press. p. 307.
76. Hall, T. A., *Nucl. Acids. Symp. Ser.* **1999,** 95.
77. Henriksen, A.; Schuller, D. J.; Meno, K.; Welinder, K. G.; Smith, A. T.; Gajhede, H., *Biochemistry* **1998,** *37,* 8054.
78. Sivaraja, M.; Goodin, D. B.; Smith, M.; Hoffman, B. M., *Science* **1989,** *245,* 738.
79. Sambrook, J.; Maniatis, T.; Fritsch, E. F., *Molecular Cloning: A Laboratory Manual.* 2nd ed. 1989: Cold Spring Harbor Laboratory Press.
80. Nagy, J. M.; Cass, A. E.; Brown, K. A., *J. Biol. Chem.* **1997,** *272,* 31265.
81. Saint-Joanis, B.; Souchon, H.; Wilming, M.; Johnsson, K.; Alzari, P. M.; Cole, S. T., *Biochem. J.* **1999,** *338,* 753.
82. Jesmin. 2001, University of London: London.
83. Harlow, E.; Lane, D., *Antibodies: A Laboratory Manual.* 1988, Cold Spring Harbor, NY.
84. Goodford, P. J., *J. Med. Chem.* **1985,** *28,* 849.
85. Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Bryce, M. D.; Rodgers, J. R.; Kennard, O.; Shikanouchi, T.; Tasumi, M., *J. Mol. Biol.* **1977,** *112,* 535.
86. Jones, T. A.; Kjeldgaard, M., *O.* 2000.
87. *Insight Ii.* 2000, Accelrys: San Diego.
88. McPherson, A., *The Preparation and Analysis of Protein Crystals.* 1982, New York: John Wiley and Sons.
89. Otwinowski, Z., *Oscillation Data Reduction Program*, in *Data Collection and Processing*, L. Sawyer, N. Isaac and S. Bailey, Editors. 1993, Proceedings of the CCP4 Study Weekend. p. 63.
90. Gewirth, D., *A Sescription of Programs Denzo, Xdisplayf and Scalepack*, in *The Hkl Manual*. 1994, Yale University: New Haven, CT.
91. Brünger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L., *Acta Crystallogr., D* **1998,** *D54,* 905.
92. Vagin, A. A.; Teplyakov, A., *J. Appl. Crystallogr.* **1997,** *30,* 1022.

93.   Crowther, R. A., *The Molecular Replacement Method*, M. G. Rossmann, Editor. 1972, Gordon & Breach: New York. p. 173.

Vita

Jamie Nicole Jones was born in Knoxville, Tennessee on October 8, 1978, the daughter of James W. and Roberta C. Jones. After graduating as valedictorian from Wartburg Central High School, Wartburg, Tennessee, she entered the University of Tennessee in Knoxville, Tennessee in August 1996. Jamie received a Bachelor of Science degree in Honors Chemistry from the University of Tennessee in May 2000. In August of 2000, she entered the Graduate School of the University of Texas at Austin under the supervision of Professor Alan H. Cowley.

Permanent address: 703 Catoosa Road
                       Wartburg, TN 37887

This dissertation was typed by Jamie N. Jones, the author