Effect of Oriented Electric Fields on Biologically Relevant Iron–Sulfur Clusters and Bioinformatics Investigations of Biotin Synthase

Samuel James Howe Gaughan

Acknowledgements

I would like to thank my supervisors, Jonathan Hirst, Anna Croft, and Christof Jäger for their guidance and support throughout my studies; I cannot imagine a better group of people to work with.

I would also like to thank the multitude of people I have had the pleasure of discussing science with, including members of the computational chemistry and the computational biochemical engineering groups, my cohort in the CDT in Sustainable Chemistry and the late Nick Besley who was a fountain of knowledge about the Q-Chem software package in the earlier years of my study. Special thanks to David Robinson who inspired much of my passion in computational biochemistry and guided me onto to the path I am now on.

Finally, many thanks to my parents for their unwavering support not only through my PhD studies, my sister, whose passion for physics has sent her down a similar road to mine and offers me a unique opportunity to always have a like-minded individual to get excited about science with, (and will likely continue to outperform me in her studies!) and the staff at the CDT in Sustainable Chemistry, particularly Pete Licence whose support and advice has been invaluable.

Abstract

Enzymes, as biological catalysts, enjoy several benefits over the more commonly used metal catalysts in chemistry, particularly in terms of sustainability. They can, however, be more complicated to utilise and manipulate and research tends to focus on engineering enzymes for specific tasks where the complexity is reduced and a by-product of this is increased understanding of sequence-structure-function relationship. An alternative approach is to find broader problems whose solutions could be applied to the engineering of many enzymes, or at least a large, multipurpose superfamily of them.

An excellent target for this type of approach is the radical S-adenosylmethionine (rSAM) superfamily, particularly due to its common mechanism of generating a radical species and using careful substrate control to dictate the reaction products across the different enzymes. This common mechanism includes an iron-sulfur cluster which can be influenced by the electrostatic environment, providing a clear path for study and promising powerful engineering opportunities.

The focus of this research is an analysis of the effect of oriented electric fields on several relevant iron-sulfur clusters using a systematic, high throughput density-functional theory (DFT) study to gain both quantitative and qualitative information on the relative energies of spin states, orbitals, vertical electron affinities and spin-coupling constants. In addition, methods are identified for coupling this type of study with bioinformatic information for the purpose of enzyme engineering. Applying this to an exemplar of the rSAM superfamily, biotin synthase (BioB), indicates promising scope for variation at iron-sulfur cluster binding sites, whilst retaining functionality.

Both the DFT results and the bioinformatics analysis represent a promising step towards the potential automation of enzyme engineering and is not limited to biotin synthase or even rSAM enzymes. This could result in improved development of a wide variety of chemical products in sustainable, efficient, and low-carbon syntheses, with the concomitant contributions to mitigating climate change.

List of Original Publications

This thesis is partially based on the following publication that is cited in the text of Chapter 3. The paper is attached in Appendix A3.4 at the end of this thesis. The text also includes unpublished material.

[1] Gaughan, S. J. H.; Hirst, J. D.; Croft, A. K.; Jäger, C. M. Effect of Oriented Electric Fields on Biologically Relevant Iron-Sulfur Clusters: Tuning Redox Reactivity for Catalysis, *J. Chem. Inf. Model.* **2022,** *62,* 591-601.

Chapter 1 – Introduction and Background

1.1 – Motivation

Over millions of years enzymes have evolved into the powerful biological catalysts they are today. Even before their isolation in 1926[1] they were studied and exploited to perform their biological functions *in vitro*. Without the ability to modify the function of enzymes to a significant degree chemists turned to highly versatile metal catalysts to perform a vast array of reactions across an ever-widening chemical space. While useful and almost ubiquitous in their applications, the use of many of these metal catalysts is not sustainable in the long term (and in some cases in the short term) simply based on their supply; some of the most powerful metal catalysts in chemistry are among the rarest in the crust (for example, ruthenium, rhodium, palladium and platinum (Figure 1.1)); and demand, for instance many rare earth metals used in chemistry are also critical for computer processors and other technological staples. This alone is reason enough to motivate the search for a replacement if we are to continue chemistry at the current scale, but there are other concerns such as the toxicity of many metals and metal compounds, the environmental impact of mining for, and processing of, continually depleting rare earth metals and the socio-political implications resulting from the geographical locations of some deposits.

Enzymes on the other hand are principally constructed from amino acids, built from hydrogen, carbon, oxygen, nitrogen, and sulfur, some of the most abundant elements accessible in the Earth's crust. Some enzymes catalyse reactions without the introduction of any non-amino acid molecules (prosthetic groups) while others might create conditions favourable for another species to perform as a catalyst where it otherwise might not. Metalloenzymes are a prime example of this, enhancing the catalytic properties of the more common, and often less toxic, metals to allow reactions that would be difficult or even impossible outside of the enzyme environment. These biological catalysts do not require the creation of artificially extreme conditions to function; they will usually perform optimally at ambient pressure, the standard temperature of the organism they are present in (e.g., 37°C in humans), using water as a solvent, and when degraded they will break down into their constituent amino acid building blocks, requiring little to no waste treatment.

Using a biological catalyst, even if it contains a metal that is less environmentally harmful such as iron, nickel, manganese, or zinc, is clearly superior to the rare earth metal catalysts in terms of sustainability.

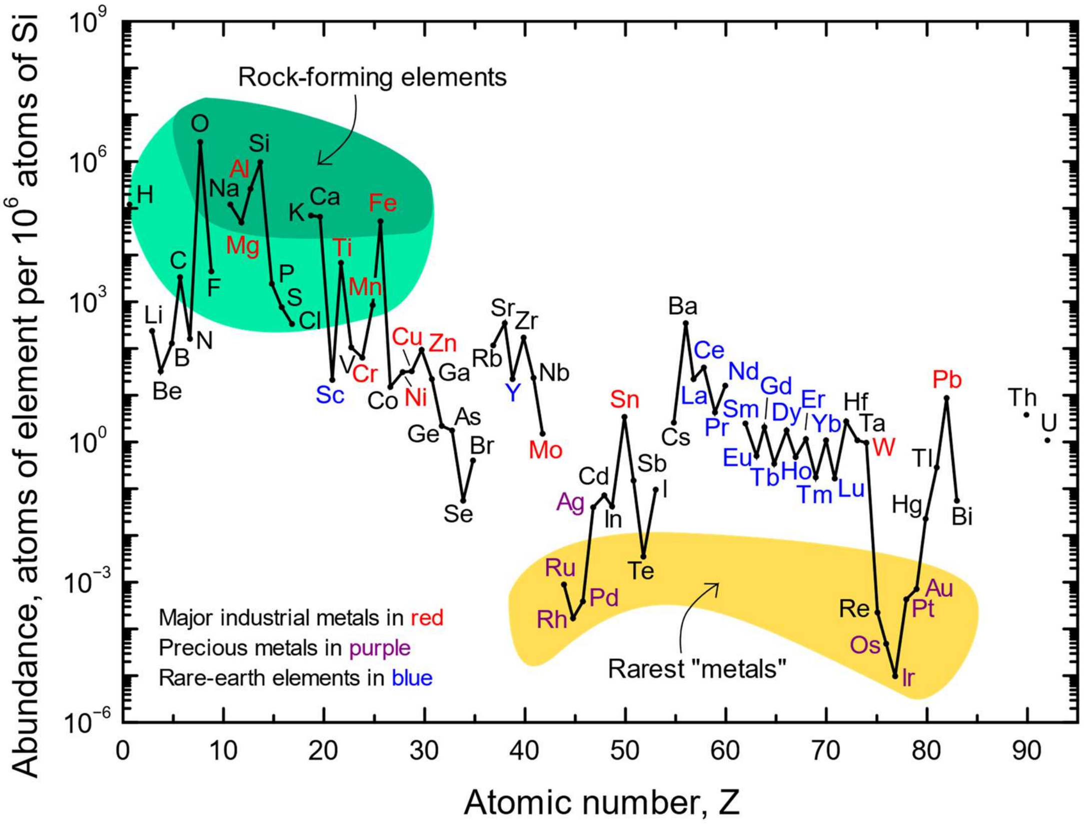


Figure 1.1: Abundance (atom fraction) of the chemical elements in Earth's upper continental crust as a function of atomic number. Many of the elements are classified into (partially overlapping) categories: (1) rock-forming elements (major elements in green field and minor elements in light green field); (2) rare earth elements (lanthanides, La–Lu, Y and Sc; labelled in blue); (3) major industrial metals (global production >~3x10^7 kg/year; labelled in red); (4) precious metals (purple); and (5) the nine rarest "metals"—the six platinum group elements plus Au, Re, and Te (a metalloid) in yellow field. Figure and caption (modified) from USGS Fact Sheet 087-02[2]

In fact, we can quantify this statement more rigorously by applying the 12 principles in green and sustainable chemistry as a critical analysis of enzymes in general.[3] As biological catalysts contain certain inherent properties outlined above, we should consider the principles that seem to be relevant simply by the choice of catalyst:

- *Principle 3: Wherever practicable, synthetic methods should be designed to use and generate substances that possess little or no toxicity to human health* *and the environment.* – Enzymes are present in all forms of life as the foundation of biological processes. They do not possess any inherent toxicity (one should still consider whether the substrates and products of the reaction being catalysed are safe.)

- *Principle 5: The use of auxiliary substances (e.g., solvents, separation agents, etc.) should be made unnecessary wherever possible and, innocuous when used.* – Enzymes predominantly use water as a solvent, arguably the most innocuous solvent one can think of.
- *Principle 6: Energy requirements should be recognized for their environmental and economic impacts and should be minimized. Synthetic methods should be conducted at ambient temperature and pressure.* – Enzymes usually perform optimally at ambient pressure, and may require some heating when used *in vitro*, to simulate a biological environment. Alternatively, we can use organisms such as bacteria as an automated controlled environment for enzymes to function in.
- *Principle 7: A raw material or feedstock should be renewable rather than depleting whenever technically and economically practicable.* – Due to their composition, three-dimensional folds of amino-acid chains, enzymes can be readily constructed without the concerns that metal catalysts raise.
- *Principle 8: Unnecessary derivatization (use of blocking groups, protection/deprotection, temporary modification of physical/chemical processes) should be minimized or avoided, if possible, because such steps require additional reagents and can generate waste.* – An example often cited when discussing this principle is enzymes due to their high specificity compared to other catalysts.
- *Principle 9: Catalytic reagents (as selective as possible) are superior to stoichiometric reagents.* – As biological catalysts, enzymes clearly fit this description.
- *Principle 10: Chemical products should be designed so that at the end of their function they break down into innocuous degradation products and do not persist in the environment.* – As catalysts they should not be considered “products” but even including their disposal we see that enzymes would simply degrade into innocuous amino acids.

- *Principle 12: Substances and the form of a substance used in a chemical process should be chosen to minimize the potential for chemical accidents, including releases, explosions, and fires.* – Similar to principle 3, enzymes tend not to be prone to explosive or combustive reactions due to their presence in nature.

While it would be incorrect to claim that all enzyme-catalysed reactions will therefore be green and sustainable as the substrates and products should still be considered, it is reasonable to state that, in most cases, using a biological catalyst rather than a synthetic (e.g., metal) catalyst for the same reaction is the more sustainable choice. Furthermore, it can be argued that first studying new catalysts *in silico* provides advantages in both sustainability and understanding, in addition to allowing access to enzymes that can be troublesome to use *in vitro* in their current forms.

Why then, if enzymes are the solution to green, sustainable, renewable, and cheap chemical catalysis, do we not simply replace all the metal catalysts with enzymes? The outstanding issue is that of complexity. One of the greatest strengths of enzymes, their highly specific activity, is also one of their greatest weaknesses as this specificity comes from their highly complex structure. The enzymes used in chemistry are almost exclusively taken directly from biological processes with little or no modification, presenting a challenge for reactions that do not have biological analogues. We do not currently have a sufficient understanding of exactly how the amino acid chain links to the structure of an enzyme, and how that structure links to its function to be able to design a protein *de novo*. We can, to a certain extent, engineer enzymes to perform very similar reactions to their wild-type reaction, but the truly ground-breaking discovery would be that of a system that can create an enzyme to catalyse any given reaction from scratch. While this is likely a long way off, we can work towards this goal by gaining a better understanding of an enzyme's amino acid code-structure-function relationships, with a view to re-engineering enzymes along the way.

Individual enzymes are highly specific for a particular substrate and reaction, yet enzymes in general catalyse an enormous range of reactions. The question could then be asked, how many enzymes *could* exist and be chemically active? Certainly, the number of possible unique amino acid sequences is extraordinarily large. A simple calculation can demonstrate this. If we were to arrange 400 amino acids into a single chain, with the choice of the standard 20 amino acids for each position the result would be 20^{400} permutations, approximately 10^{520} (For context the number of atoms in the observable universe is estimated to be between 10^{78} to 10^{82}). The number of unique chains of this length would be half of this value, to eliminate palindromes, which is still of the order 10^{520}. Even if most of these chains would not be catalytically active enzymes, there will be a portion that are. Considering this for every possible chain length results in a prohibitively large space to analyse exhaustively, suggesting that rational design of enzymes is a much more tractable problem, despite its current challenges, and that there may be hypothetical enzymes that can catalyse reactions that seem impossible at our current level of understanding.

1.2 – The rSAM superfamily

With this long-term motivation in mind, we can select a target for study and re-engineering potential which has practical uses, scope for improvement, and would benefit from re-engineering via an *in silico* study to gain insight into the structure-function relationship. While there are advantages associated with *in silico* studies in general, such as fine-tuned control of the system and conditions, along with obtaining precise values for a wide variety of physical properties, systems that are practically difficult to work with *in vitro* benefit even more as more of the preliminary study can be performed *in silico* to reduce the costs associated with practically difficult *in vitro* reactions. The radical *S*-adenosylmethionine (rSAM) superfamily of enzymes is one such candidate, catalysing a diverse and multifaceted set of reactions representative of enzymes' wide reaction scope, in part a result of the core mechanism common to all members of the superfamily, the cleavage of *S*-adenosylmethionine (SAM) into methionine and a radical species which will then react with many different substrates. Remarkably this does not diminish the specificity of these enzymes, with carefully controlled entry and exit structures and highly specific orientations required to trigger the creation of the radical species maintaining specificity despite using the same core reaction. One source of the practical challenges associated with rSAM enzymes is the defining motif holding the iron-sulfur cluster which performs the reductive cleavage of rSAM. When SAM is not bound to the cluster the cluster is susceptible to oxidation in aerobic conditions resulting in the rapid breakdown of this essential component and denaturing the enzyme when used *in vitro*. This further strengthens the argument for applying *in silico* methods to these enzymes to study them without encountering these practical issues.

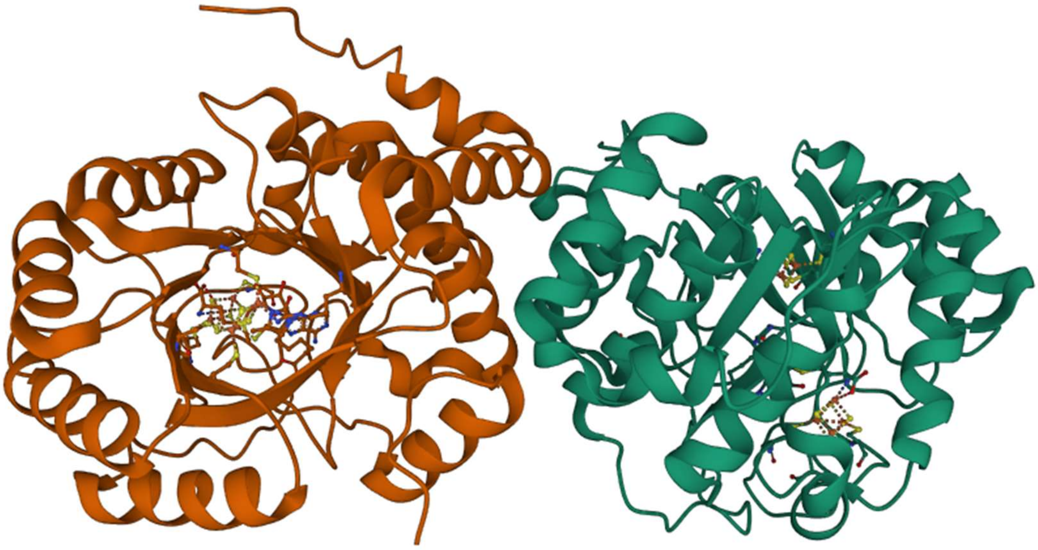


Figure 1.2: The crystal structure of *E.coli* biotin synthase (PDB ID 1R30[4, 5]), a radical SAM enzyme. The enzyme is found as a dimer as shown and the left dimer (orange) is oriented as to look directly down the TIM barrel into the active site.

The rSAM superfamily of enzymes utilises a [4Fe-4S] cluster along with SAM to generate a radical species which then performs the specific reaction that each enzyme is tailored for, although an exception to this rule is tryptophan methyltransferase, which is considered a member of the superfamily but does not require rSAM chemistry as a precursor to its methyl group transfer.[6] While rSAM was first classified as a superfamily in 2001 with an initial membership of 600, performing 30 known functions,[7] the chemistry utilised by this superfamily had been previously observed in members such as lysine 2,3-aminomutase (LAM),[8, 9] pyruvateformate lyase activating enzyme (PFL-AE),[10-13] anaerobic ribonucleotide reductase[14] and biotin synthase.[15, 16] This superfamily's growth has since exploded, with over 110,000 sequences known to be members with confirmed participation in 85 distinct reactions including some analogous to coenzyme B_{12}.[17-19] Structures of these enzymes are clearly lacking in both quantity and quality however, making them an obvious target for computational studies involving structure prediction alongside their wide-ranging uses holding great promise for bio-engineering.[20] There are many structural and mechanistic properties which members of this superfamily possess, but only the cubane iron-sulfur cluster and the cleavage of SAM is common to every single member, signifying their importance. As the superfamily is of great interest and has been reviewed regularly, a summary of the most important discoveries, or those most pertinent to this thesis will follow.

1.3 – Structural Characteristics of the rSAM Superfamily

The [4Fe-4S] cluster present in all rSAM enzymes is most commonly bound by a CX_3CX_2C motif with the three cysteine residues coordinating three of the iron atoms, where the fourth iron atom binds to SAM during the cleavage reaction.[21-23] This fourth iron does not have a protein ligand and as such is labile and air-sensitive in most cases,[22, 24] (with the notable exception of LAM from *B.subtilis*,[25] the understanding of which would be critical to developing other air-stable rSAM enzymes) therefore adding the requirement of anaerobic conditions if the enzyme is to be isolated and remain catalytically active, making it more difficult to study experimentally. In fact this instability has led to very few rSAM enzymes to be shown as catalytic experimentally as they tend to deteriorate into apoenzymes after only a small number of turnover events *in vitro*.[26] SAM binds to this fourth iron in a bidentate manner via the amino and carboxylate moieties and it is this complex which has been described as the “unifying and catalytic feature of rSAM enzymes.”[26-28] In some cases the three cysteine motif can take a slightly different form, or there can be other cysteine rich motifs which potentially bind auxiliary iron-sulfur clusters,[4, 29-33] although these extra clusters tend not to duplicate the reaction of the primary cluster but rather aid it in various ways such as sulfur donation in BioB[34, 35] and MiaB[36] and thioether bond formation in SkfB.[37]

The macrostructure of a rSAM enzyme generally contains a full or partial triose phosphate isomerase (TIM) barrel (Figure 1.2) which holds the active site on an exposed β sheet near the top, the conserved three cysteine motif following the first β strand and the iron-sulfur cluster buried by loop regions such that it is protected from the solvent, although variations on this structural feature are observed in enzymes which still remain members of the rSAM superfamily.[31] As the TIM barrel holds the active site, but there are no common substrate binding motifs (which arguably allows the superfamily to be so varied in the reactions it can catalyse) it has been suggested that the size and shape of the opening of the TIM barrel, along with structural features near the core barrel provide this superfamily's substrate specificity.[38, 39]

1.4 – Iron-sulfur Clusters in Enzymes

Iron-sulfur clusters play a critical role in the reactions catalysed by several families of enzymes, providing a wide variety of functions in each. Their possible role in enabling the emergence of early life[40] and capacity to perform many different roles within enzymatic pathways[41, 42] has led to them being characterised as "one of the most ubiquitous and functionally versatile prosthetic groups in nature".[43] The discovery and purification of ferredoxins in 1962 was an early indication that iron could play roles in enzymes in addition to its well-known presence in hemoproteins.[44] Both iron and iron-sulfur clusters act primarily as mediators for electron transfer with the ability to be either the source or sink for electrons in redox reactions, and iron-sulfur clusters are extremely useful for electron transport due to the delocalisation of electron density across the cluster.[45, 46] The mediator role is one of the most common functions of iron-sulfur clusters, found in a variety of enzymes including those that couple proton transfer to electron transport, such as [FeFe] hydrogenases, which possess a unique version of an iron-sulfur cluster featuring a diiron centre and a bridging dithiolate.[47, 48] Many enzymes containing iron-sulfur clusters use [2Fe-2S], [4Fe-4S] and [3Fe-4S] structures in both redox and non-redox functions. For example, rSAM enzymes make use of a [4Fe-4S] cluster to reductively cleave SAM into methionine and a 5'-deoxyadenosyl radical, the latter of which is used to initiate a variety of radical reactions that have been reviewed previously.[26, 49-51] In some cases, electron transfer may be an intermediate step rather than the complete function of an enzyme, such as the case of the biotin synthase rSAM mechanism where the FeS clusters mediate the donation of a sulfur atom.[52]

1.5 – The 5'-Deoxyadenosyl Radical Mechanism

rSAM enzymes reductively cleave SAM via the input of one electron taken from the iron-sulfur cluster $[4Fe\text{-}4S]^+$ (Figure 1.3) which is accepted to be the catalytically active oxidation state based on observations in both LAM and PFL-AE, particularly helped by the catalytically active 1+ oxidation state being EPR active while the 2+ state is EPR silent.[12, 53-58] The intermediate of this cleavage was hypothesised to be a 5’deoxyadenosyl radical (dAdo$^{\bullet}$) which abstracts a hydrogen atom from the substrate during catalysis,[59-62] and was supported by the isolation of stabilised analogues of dAdo$^{\bullet}$.[63-67]

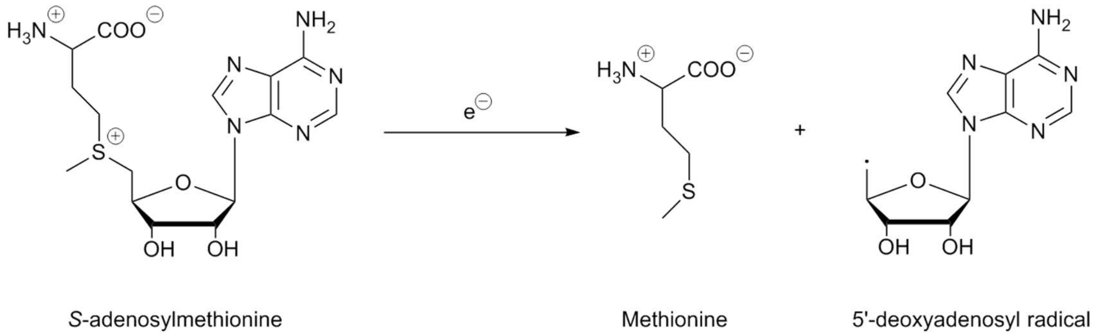


Figure 1.3: Reductive cleavage of SAM to generate methionine and dAdo$^{\bullet}$, mediated by the [4Fe4S] cubane iron-sulfur cluster common to rSAM enzymes. Methionine is a by-product and the dAdo$^{\bullet}$ species will then react with the main substrate of the enzyme.

Controlling the intermediates in rSAM enzymes is particularly interesting as this could be a powerful target for re-engineering and the field has been making progress on solving this puzzle recently.[68-75] Many studies have focused mainly on the chemistry at the iron-sulfur cluster site, but it has become clear that the control of radical intermediates is often the role of surrounding residues which when mutated usually cause the production of different products from SAM cleavage.[68] Particularly important is that the enzymes restrict the environment surrounding the radical such that it is impossible for it to react with anything but the intended molecule, leading to the proposal that the use of van der Waals contacts is common throughout the rSAM superfamily as the means of radical control.[73, 74] The number of reactions that rSAM enzymes catalyse using dAdo$^{\bullet}$ is vast and this subject has been reviewed recently,[26] showing that the potential of the superfamily is almost ubiquitous, allowing a greater scope for re-engineering than other enzyme families.[68, 70]

1.6 – Redox Reactivity and Electrostatics

Redox activity is an important property in enzymes. Reengineering this property has been the focus of experimental studies for many families of enzymes, including the antioxidant peroxiredoxins,[76] thioredoxins, particularly those that act as electron donors for other enzymes,[77] and kinases,[78] amongst other more general studies focused on control via thiol/disulfide exchange,[79, 80] enzyme orientation[81] and substrate specificity.[82] Additionally, redox activity has been used to probe mechanism,[83-87] to understand the effect of changing iron-sulfur cluster ligands on the redox potential,[88] to assess stability and reactivity of the cluster,[89] and to guide improvements in enzyme activity with directed evolution.[90]

Redox reactions, often studied and used via electrochemical methods,[91-93] are sensitive to electric fields. This alone would motivate an analysis of the electrostatic environment generated by an enzyme containing an iron-sulfur cluster. However, electrostatics also play a role in enzyme catalysis in general,[94] including protein-protein interactions,[95] conformational motions,[96] and catalysis.[97, 98] Electrostatic preorganisation in the active site of enzymes, and the electrostatic stabilisation associated with this, are a more recent area of study.[99] This preorganisation has also been quantified computationally.[100, 101] Computational methods such as molecular dynamics simulations, density functional theory (DFT), Valence Bond Theory and Poisson-Boltzmann Equation Solvers have advanced our understanding of the role of electrostatics in enzyme catalysis,[102-104] providing a more complete picture of the function of enzymes such as alpha-amylase,[105] methyltransferases[106] and QueE.[107] These methods have also been used to study specific properties such as the contribution of individual amino acids to the overall electrostatic field of a protein,[108] electrostatic steering and channeling,[109, 110] and the direct effect of the electrostatic field on catalytic rate.[111]

Direct study of electric fields in the context of manipulating catalysis is a growing area of research. A recent review highlighted oriented electric fields as reagents, as well as their effects on enzyme catalysis.[112] DFT as the QM method in a QM/MM approach has been used previously to study the effects of an electric field on biological chromophores[113] and on enzymes.[114] Studies have also been performed on iron-sulfur clusters in the context of enzymes to elucidate information about properties such as coordination, geometry, and electrostatics. These studies have used extended X-ray fine structure spectroscopy (EXAFS) and DFT,[115] including the application of broken-symmetry DFT[116] to iron-sulfur clusters to model antiferromagnetic coupling.

1.7 – Biotin Synthase

An exemplar from the rSAM superfamily is Biotin Synthase (BioB) which catalyses the final step in the synthesis of biotin (vitamin B_7) by performing a radical mediated sulfur insertion reaction, converting dethiobiotin to biotin (Figure 1.4).[34] Each C-S bond formation requires a radical so the cleavage of two SAM molecules by the mechanism shown in Figure 1.3 is performed to fully complete this reaction[26] (although it has been observed that the average number is closer to three due to abortive processes.[117]) There is therefore a significant limitation on total reaction time because the products of the first cleavage must leave the active site before another SAM molecule can enter to form the second C-S bond. A further limitation results from the source of the sulfur inserted into dethiobiotin which appears to come from a [2Fe2S] secondary iron-sulfur cluster which must be repaired before the enzyme can function again.[26, 118, 119] *In vivo* regeneration of this cluster is likely performed by the iron-sulfur cluster assembly system and attempts have been made to artificially reproduce this process.[120, 121] As an rSAM enzyme the classical CX_3CX_2C motif binding the [4Fe4S] cluster is present where the cysteines coordinate three of the iron atoms and the 4th iron is the binding site for SAM as expected. The secondary [2Fe2S] cluster is unusual in that one of the coordinating ligands is an arginine, along with three cysteines. This hints at the possibility of other residue changes that might be useful in enzyme engineering as these iron-sulfur clusters do not bind exclusively to cysteines.

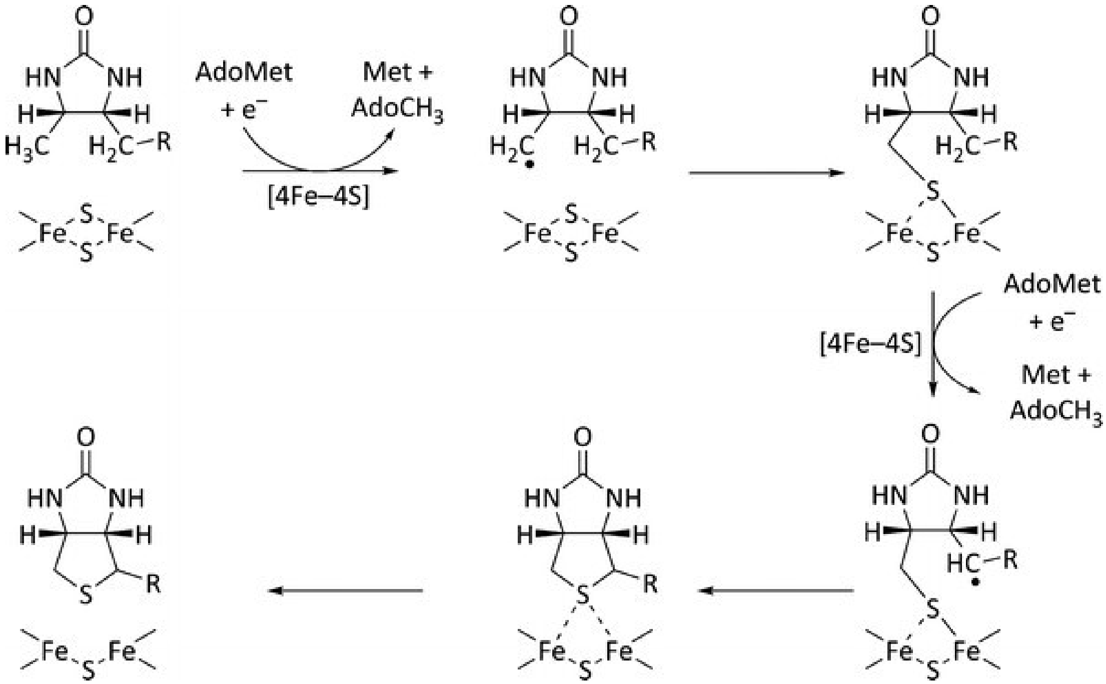


Figure 1.4: Proposed mechanism of sulfur insertion in BioB.[34] Two SAM cleavage events (Figure 1.3) are required to fully complete the process and the sulfur is donated from an auxiliary iron-sulfur cluster in the enzyme. Figure adapted from Fuchs *et al.*[118]

Biotin is classified as a vitamin due to its many essential uses in metabolism and the inability for human biology to synthesise it, although some intestinal bacteria can do this.[122] Biotin is a cofactor for some carboxylases that catalyse parts of fatty acid, glucose and amino acid metabolism, along with having important roles in gene regulation, cell signaling, and histone modification.[123-127] Biotin is present in so many common sources of nutrition that the estimated daily intake exceeds the recommended Adequate Intake and biotin deficiency due to inadequate intake is exceedingly rare, usually caused instead by a genetic disorder.[128] Despite this it is still sold as a dietary supplement and is often present in hair and nail treatment products with the commonly promoted claims of hair and nail strengthening based on very limited literature, some of which only includes people who already have a biotin deficiency,[129-135] and the known effects of biotin deficiency such as hair loss.[136] A much stronger argument for its usefulness comes from biotechnology where biotin is a very common label used in biochemical assays as a bioconjugate, with applications for protein detection, purification and immobilisation.[137]

Considering the challenges associated with rSAM enzymes and specifically BioB, along with the uses of biotin and the scope for improvement of BioB through rational enzyme engineering it is an excellent target for an *in silico* study with re-engineering in mind.

Chapter 2 – Methods

2.1 – The Schrödinger Equation

Ab initio methods use fundamental physical laws to predict the structure and behaviour of molecules or small systems without any empirical parameters. In principle an *ab initio* method aims to solve the time-dependent Schrödinger equation exactly and produce a complete picture of the physical properties of the system. Often a useful approach is to use the time-independent Schrödinger equation instead as the solutions represent the stationary states of the wave function Ψ which in turn represents the quantized energy levels of the system. The time-independent Schrödinger equation can be written as

$$\hat{H}\Psi = E\Psi, \tag{2.1}$$

where Ψ is the wave function, E is the energy and $\hat{H}$ is the Hamiltonian operator, which itself can be written compactly as

$$\hat{H} = \hat{T}_{\mathrm{e}} + \hat{T}_{\mathrm{n}} + \hat{V}_{\mathrm{nn}} + \hat{V}_{\mathrm{ne}} + \hat{V}_{\mathrm{ee}}, \tag{2.2}$$

where $\hat{T}_{\mathrm{e}}$ is the kinetic energy operator for the electrons, $\hat{T}_{\mathrm{n}}$ is the kinetic energy operator for the nuclei, $\hat{V}_{\mathrm{nn}}$ is the nucleus-nucleus repulsion operator, $\hat{V}_{\mathrm{ne}}$ is the electron-nucleus attraction operator, and $\hat{V}_{\mathrm{ee}}$ is the electron-electron repulsion operator.[138-140] This equation has no known analytical solutions for all but the simplest systems so several approximations must be made before methods can be developed to even estimate the solution to the Schrödinger equation.

The Born-Oppenheimer approximation notes that as electrons have an extremely small mass and high velocity relative to nuclei, the nuclei can be viewed as fixed points such that the only particles in motion in the system are electrons. Under this approximation the kinetic energy of the nuclei $\hat{T}_{\mathrm{n}}$ becomes zero and the potential energy between nuclei $\hat{V}_{\mathrm{nn}}$ reduces to a constant. This allows the Hamiltonian in the Schrödinger equation to be written as the electronic Hamiltonian,

$$\hat{H}_{\mathrm{elec}} = \hat{T}_{\mathrm{e}} + \hat{V}_{\mathrm{ne}} + \hat{V}_{\mathrm{ee}}, \tag{2.3}$$

with the time-independent Schrödinger equation becoming

$$\hat{H}_{\text{elec}}\Psi_{\text{elec}} = E_{\text{elec}}\Psi_{\text{elec}}. \tag{2.4}$$

The total energy W under this approximation can then be written as

$$W = E_{\text{elec}} + V_{\text{nn}}, \tag{2.5}$$

where V_{nn} is the constant nucleus-nucleus repulsion. Although this appears easier to solve than the exact form as only the number of electrons and the external potential V_{ext} (the Coulomb potential exerted by the nuclei on the electrons) are needed to define a system completely and uniquely, there is no analytical solution to calculate $\hat{V}_{\text{ee}}$. Due to this the only analytical solutions to the Schrödinger equation are those for single particles, as $\hat{V}_{\text{ee}}$ is not required.[139, 141, 142] The approach to tackle this is to find a numerical solution instead, in this case through the use of the variational principle, which is possible due to a particular property of normalised wave functions

$$\langle\Psi_{\text{trial}}|\hat{H}|\Psi_{\text{trial}}\rangle = E_{\text{trial}} \geq E_0 = \langle\Psi_0|\hat{H}|\Psi_0\rangle, \tag{2.6}$$

where the equality is only true if $\Psi_{\text{trial}} \equiv \Psi_0$. This means that the energy resulting from a trial wave function will always be greater than or equal to the true ground state energy, E_0. If it were possible to conduct an exhaustive search over all possible Ψ_{trial} to minimise E_{trial} the result would be E_0. Unfortunately the set of Ψ_{trial} is infinite, so only a finite subset can be searched and the smallest resultant energy (by the variational principle, the closest to E_0) is taken to be an approximation to E_0, with some convergence criteria used to stop the iterative process when the changes between various E_{trial} results are sufficiently small.[139, 140, 143] While this allows practical numerical calculations using the Schrödinger equation to be feasible for small systems the computational cost required to conduct sufficiently large searches of the set of Ψ_{trial} for most systems is unreasonable. This then highlights one of the most important challenges in computational chemistry, the distinct positive correlation between accuracy and computational cost.

2.2 – *Ab Initio* Methods

Hartee-Fock[144] attempts to approximate the solution to the Schrödinger equation for small systems with more than one electron using a single Slater determinant constructed from one-particle orbitals that satisfies the requirement of the wavefunction to be antisymmetric with respect to electron exchange. This determinant is then used as Ψ_{trial} when applying the variational principle (Eqn 2.6) and as such can be improved upon iteratively. While the initial approach was to use a linear combination of Slater-type[145] atomic orbitals to represent the one-electron wavefunctions the computational cost involved led to the adoption of a linear combination of Gaussian-type[146] orbitals to approximate each Slater-type orbital.

The set of functions used to represent these orbitals is referred to as a basis set. For example, STO-nG basis sets use n Gaussian functions to represent each Slater-type orbital (Figure 2.1), and additional properties can be represented by, for example, polarization or diffuse functions added to the basis set. It is also common for larger systems to use basis sets that put more emphasis on the valence electrons due to their greater importance in chemical reactions. For example, the Pople basis sets[147] of the form x-yzG where x represents the number of Gaussian functions for each core atomic orbital basis function and the numbers after the hyphen indicating that the valence orbitals are represented by two basis functions, one constructed of y Gaussians and the other of z. This is referred to as split-valence double zeta; the valence orbitals are split into a number (zeta) of basis functions, in this case two, while triple-zeta and higher can also be used. There are many other basis sets designed for various purposes and levels of cost/accuracy so only the ones used will be explained further in this thesis.

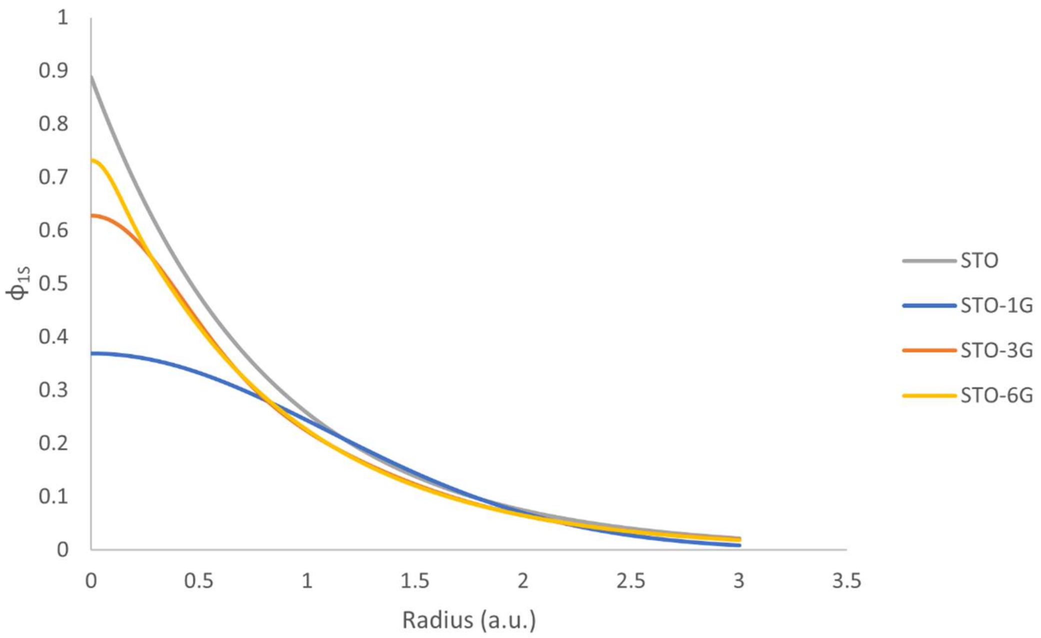


Figure 2.1: Slater-type orbital approximation for a hydrogen atom compared with a single Gaussian (STO-1G) and a linear combination of 3 (STO-3G) and 6 (STO-6G) Gaussian functions to approximate the Slater-type orbital.

One of the major flaws in Hartree-Fock is the lack of consideration for electron correlation, which can be approximated by modifying the process with post-Hartree-Fock methods. Møller-Plesset perturbation theory (MPn) adds correlation to the Hartee-Fock Hamiltonian via perturbations of a particular order n.[148] Higher order MPn methods such as MP5 and above are more accurate in theory but are often difficult to converge, increasing computational cost even further.[149] Another post-Hartree-Fock method to add electron correlation to the system is coupled cluster theory (CC)[150] which adds additional multi-electron wave functions to simulate correlation using a “cluster operator” T which is a linear combination of operators representing each excitation level. CC methods are defined by the levels of excitation taken into account, with parentheses used to indicate that level of excitation was approximated using Many-Body perturbation theory (MBPT) rather than calculated explicitly.[151-153] Configuration interaction (CI) adds correlation via a linear combination of configuration state functions (CSFs) and solves the electronic Schrödinger equation exactly within a certain basis set if all possible CSFs are included.[154] This is known as full configuration interaction (FCI) and is rarely used practically due to the computational time involved and the rapid scaling with system size. CI methods are usually truncated down and the notation for which excitations are used is very similar to that of CC methods.[153] These methods, while generally accurate for small molecules, have computational costs that become prohibitively expensive as the size of the system increases. Each method also has its own difficulties with certain physical properties due to assumptions made in their formulations. Quantum Monte-Carlo (QMC) methods are a relatively new branch of *ab initio* methods with promising results, and tend to be able to tackle larger systems than traditional *ab initio* methods for the same computational power with a comparable level of accuracy, (and are still in the process of being improved[155]) but are still unsuitable for calculations on systems as complex as an enzyme's active site.[156-159]

While *ab initio* methods can produce extremely accurate results for a variety of properties of small molecules and systems, they do not scale well enough to consider using them on such a large system as an enzyme, or even its active site. QMC has been promising in its practical application due to simple parallelisation and reasonable scaling and could be valid for consideration in the future when computational power is greater. If a very specific interaction between two small functional groups needs to be examined then an *ab initio* method could be considered, but otherwise less accurate and cheaper methods are preferable.

2.3 – Density Functional Theory

When using wave function methods on a system with N electrons, $4N$ variables are required to fully describe the state of the system ($3N$ spatial coordinates and N spin variables) which becomes prohibitively difficult to compute in a reasonable amount of time on large systems such as an enzyme's active site. Density functional theory (DFT) is a method to retain the same information using fewer variables, allowing scaling to larger systems with the same level of accuracy.[139, 140] An important property of the wavefunction is that its square is equal to the probability of finding a certain electron in a specific volume element,

$$|\Psi(x)|^2 = p(x), \qquad (2.7)$$

which implies that the integral of this over all space (Eqn 2.8) is the probability of finding a certain electron within the entire volume of the system, which is clearly 1.[138]

$$\int_{-\infty}^{\infty} |\Psi(x)|^2 dx = 1. \tag{2.8}$$

Using this result and the property that electrons are indistinguishable from one another allows us to note that a new variable $\rho(r)$, the electron density, can be defined as the probability of finding any of the N electrons at a particular position. This can be used as a replacement for the wave function as it contains all the information required to completely and uniquely define a Hamiltonian to be able to solve the Schrödinger equation, as shown by the first Hohenberg-Khon theorem.[139, 160] The second Hohenberg-Kohn theorem shows that, like the wave function, the density follows the variational principle and can be used to iterate towards the lowest energy solution to the Schrödinger equation, allowing methods to be created around the density in the same way.[139, 160] If the Hohenberg-Kohn theorems are collected and formulated as a minimisation problem the following is reached;

$$E_0 = \min_{\rho \to N}(F[\rho] + \int \rho(r) V_{\mathrm{ne}} dr), \tag{2.9}$$

where V_{ne} is the electron-nucleus attraction and $F[\rho]$ is the Hohenberg-Kohn functional containing all kinetic, classical Coulombic and non-classical contributions; this is known as the Kohn-Sham method.[139, 160, 161] The kinetic and non-classical contributions do not have known exact forms and therefore have to be approximated, however Kohn and Sham redefined the functional to easily separate the parts without exact forms,

$$F[\rho(r)] = T_{\mathrm{S}}[\rho(r)] + J[\rho(r)] + E_{\mathrm{XC}}[\rho(r)], \tag{2.10}$$

where T_{S} is the non-interacting kinetic energy, J is the Coulomb potential and

$$E_{\mathrm{XC}} \equiv (T[\rho] - T_{\mathrm{S}}[\rho]) + (E_{\mathrm{ee}}[\rho] - J[\rho]) = T_{\mathrm{C}}[\rho] + E_{\mathrm{ncl}}[\rho], \tag{2.11}$$

where T is the total kinetic energy and E_{ee} is the total electron-electron interaction energy. This results in T_{C}, the difference between the total kinetic energy and its non-interacting component, and E_{ncl}, the non-classical component of the electron-electron interaction energy. To generate a Hamiltonian an effective local potential V_{S} is required, which also has an unknown exact form, but can be related to the exchange-correlation energy E_{XC} since V_{S} can be written as

$$V_{\mathrm{S}}(\vec{r}) = \int \frac{\rho(\vec{r}_2)}{r_{12}} d\vec{r}_2 + V_{\mathrm{XC}}(\vec{r}_1) - \sum_{A}^{M} \frac{Z_A}{r_{1A}} \tag{2.12}$$

where $\vec{r}_{\mathrm{p}}$ are the spatial coordinates of particle p, r_{pq} is the distance between the particles p and q, the summation runs from nucleus A to M and V_{XC} can be written as

$$V_{\mathrm{XC}} \equiv \frac{\delta E_{\mathrm{XC}}}{\delta \rho}. \tag{2.13}$$

As neither the exchange-correlation energy E_{XC} or potential V_{XC} are known, this is where DFT is no longer exact; approximations must be made to these variables through the use of functionals which produce expected properties, often through the inclusion of empirical parameters. Therefore DFT, while exact in its theory and formulation (in particular in its inclusion of correlation, which is notably absent from HF, requiring post-HF methods as previously mentioned) and therefore an *ab initio* method in this way, is usually a semi-empirical, or at least approximate, method in practice.[139, 140, 160, 161] However, the improvement in computational cost of DFT over wave function methods allows it to be used on systems of much larger size with acceptable accuracy, and such studies have been undertaken in the context of radical SAM enzymes among others.[162-165]

A vast array of DFT functionals have been constructed for a wide variety of applications, from those made to apply to as many different systems as possible, to those designed with very specific sets of molecules in mind, with a view to accurately predicting physical observations through, in some cases, the use of highly empirically parameterised functionals. Generally functionals can be categorised by their approach to using the electron density, popularised by the introduction of 'Jacob's Ladder'.[166] This description is a reference to a Biblical story about a ladder ascending to Heaven, with the comparison that upon reaching the top of this ladder we would find the perfect 'divine' functional which can perfectly represent all systems and produce the theoretically exact DFT formulation discussed earlier. It can then be envisaged that there are 'rungs' on this ladder representing functionals that achieve results closer to that of the 'divine' functional the further up the ladder we travel. A brief overview of the categories of functional that create these rungs will follow.

The most basic approach is to have a functional that only depends on the value of the electron density at a certain point, often by assuming the electron density surrounding the considered coordinate can be approximated as a homogeneous electron gas with the same density as the coordinate. This is generally given the name Local Density Approximation (LDA) and was a common method in the early days of practical application of DFT.[167-171] However the homogeneous electron gas approximation is not accurate for anything other than a homogeneous electron gas, especially when considering increasingly non-local effects. This leads to a fairly intuitive development in which the first derivative of the density is also used to generate a much more realistic picture of the wider area around the density coordinate in question. These functionals are termed Generalized Gradient Approximations (GGA)[172-174] (examples include PBE[174], BLYP (Becke exchange[175] with Lee, Yang and Parr correlation[176]) OLYP (OPTX exchange[177] with LYP correlation[176]) and BP86 (Becke exchange[175] with Perdew 86 correlation[178])) and variations of them have been built upon to generate most of the functionals we use today. The most obvious extension is to use higher-order derivatives of the density, most commonly the second derivative, which are termed meta-GGAs (examples include TPSS[179], M06-L[180], revTPSS[181] and SCAN[182]). Additionally including a percentage of Hartree-Fock exchange improves the accuracy of GGAs and these are known as hybrid GGAs (for example, the widely used B3LYP (Becke's 3-parameter exchange[183] with LYP correlation[176]) and PBE0 (PBE exchange and correlation[174] with 25% HF exchange) functionals). Hybrid and meta variations are not mutually exclusive so can be combined to form meta hybrid-GGAs (such as the M06 family of functionals[180, 184-187]).

There are also additional considerations for systems being approached from a computational perspective in relation to the electronic configuration of the system. Systems where all electrons are paired, such that the spin multiplicity of the system is 1, are known as "closed-shell" as all molecular orbitals are fully occupied. The converse, where there are one or more unpaired electrons is termed "open-shell", and this can be further categorised by either forcing both spins α and β to occupy the same spatial orbitals where they can be paired ("restricted") or allowing the spins to occupy different spatial orbitals ("unrestricted").

$$\chi_i(\boldsymbol{x}) = \begin{cases} \psi_j(\boldsymbol{r})\alpha(\omega) \\ \psi_j(\boldsymbol{r})\beta(\omega) \end{cases}. \tag{2.14}$$

$$\chi_i(\boldsymbol{x}) = \begin{cases} \psi_j^{\alpha}(\boldsymbol{r})\alpha(\omega) \\ \psi_j^{\beta}(\boldsymbol{r})\beta(\omega) \end{cases}. \tag{2.15}$$

Equations 2.14 and 2.15 show the forms of the Hartree-Fock spin orbitals for restricted and unrestricted systems[188] which can analogously be represented in DFT by rewriting the HK functional[189] (Eqn 2.10) to include spin-resolved densities (Eqn 2.16),

$$F[\rho^{\alpha},\rho^{\beta}] = T_{\mathrm{S}}[\rho^{\alpha},\rho^{\beta}] + J[\rho^{\alpha},\rho^{\beta}] + E_{\mathrm{XC}}[\rho^{\alpha},\rho^{\beta}], \tag{2.16}$$

where the density is also replaced with two spin-resolved densities in the definitions of the kinetic and exchange-correlation energies. In Chapter 3 of this thesis the majority of the calculations performed are unrestricted open-shell calculations, even if the multiplicity of the system is 1, as spin-restricted DFT does not necessarily result in the correct spin-density. This is an important consideration for the systems we will be studying where there are multiple spin-states, and where unrestricted DFT provides an improved spin-density.[190, 191] However unrestricted calculations suffer from "spin contamination", where the separate treatment of α and β electrons that would otherwise occupy the same molecular orbital creates an apparent higher spin, hence "contaminating" the lower spin system with higher spin configurations. This can be remedied in some ways, for example restricted open-shell[192] (ROS) calculations will prevent spin contamination at a significant computational cost, in addition to producing higher energy results, and is therefore generally reserved for systems where the degree of spin contamination is very high. Generally, a simple check to assess how much spin contamination has occurred is to compare the expectation value of the total spin, $< S^2 >$, which most computational chemistry software will produce by default, with the theoretical total spin multiplicity $s(s+1)$, where s is half the number of unpaired electrons.

The situation becomes more complicated when considering unpaired electrons in a molecular system that is spatially disparate. Consider, for example, a dissociating H_2 molecule constructed from two hydrogen atoms A and B such that the α ($ms = +1/2$) and β ($ms = -1/2$) electrons are completely spatially separated.

We can describe this as two “broken symmetry” states where the α electron is located on atom A ($ms_a = +1/2$) and the β electron is located on atom B ($ms_b = -1/2$), and the opposite state where α is at B ($ms_b = +1/2$) and β is at A ($ms_a = -1/2$). For a symmetric system, such as this one, both states would be energetically degenerate, but for many systems, including those we will consider later, there may be other factors influencing the energy or even the reactivity of these states, and as such they will need to be considered. A similar situation arises when considering metal complexes with intermediate bridges linking metal atoms, such as iron-sulfur clusters.[46, 116, 193, 194] These states are not eigenvalues of the $\boldsymbol{S}^2$ operator and are therefore not considered “pure” spin states, but can be related to them with some work.[195]

In the Heisenberg model[196] for describing magnetic systems such as iron-sulfur clusters, the Heisenberg exchange Hamiltonian is

$$H = -\sum_{i \neq j} J_{ij} \hat{e}_i \hat{e}_j, \tag{2.17}$$

where $\hat{e}_i$ is the normalised local spin vector and J_{ij} is the Heisenberg exchange coupling constant.[197, 198] The Heisenberg Double-Exchange (HDE) model[46, 116, 193, 194, 199] includes a description of the direct relationship between a broken symmetry state and the Heisenberg spin ladder[116] (Eqn 3.3) and is elaborated upon in the context of iron-sulfur clusters on page 47 of this thesis.

2.4 –QM/MM, other Hybrid Methods, and Molecular Dynamics

Molecular dynamics is a very intuitive physical method for time-dependent system evolution based on traditional Newtonian physics. It calculates energy using an empirical function rather than quantum mechanics and does not have the level of accuracy that *ab initio* methods and DFT have, but a system can be observed evolving over time with far less computational cost, and therefore larger systems can be studied in this way. Molecular dynamics (MD) begins with a snapshot of the system and calculates the forces on each particle due to the other particles in the system, with the possible addition of an external force field. The acceleration of each particle is calculated based on these forces and the system is allowed to evolve based on the resultant velocities for the length of the chosen time-step. The forces are then recalculated for the next time-step, repeating until the desired total time for the simulation has completed.[200-202] MD simulations have been useful in the modelling of proteins for over four decades, starting with the simulation of the 58 residue bovine pancreatic trypsin inhibitor in 1977 over about 9ps in a vacuum,[203] while advances in computational power have now allowed simulations over 100ns on a fully solvated protein of over 300 residues in length to become a routine calculation.[204] The two major ways in which MD simulations can be used to study proteins are the protein folding problem and the dynamics of enzyme catalysis, although MD is most useful in conjunction with other methods for the latter.[205]

When considering which method to use when modelling a system and weighing up the computational cost against accuracy, an obvious question to ask would be why not use more than one method to combine the best of both worlds? An answer was first formulated with respect to enzyme modelling in 1976 through a hybrid method combining quantum and classical mechanics via the separation of different parts of the system based on how accurate they needed to be modelled.[206] Even at this early stage it was clear to the authors that the classical and quantum potential energies could not simply be summed; there needed to be some form of correction made to account for the unusual separation of the system:

$$V = V_{\text{classical}} + V_{\text{quantum}} + V_{\text{quantum/classical}}. \tag{2.18}$$

This study laid out a general method for using both quantum mechanics and classical mechanics (often now termed molecular mechanics in computational chemistry) in the same simulation to allow for accuracy where it is needed and speed where it is not. Methods based on this idea are now known as hybrid quantum-mechanics/molecular-mechanics (QM/MM) methods[206] and have been used to study proteins and enzymes since their inception,[207-211] including members of the radical SAM superfamily.[212, 213]

The choice of which QM and MM method to use generally follows the same rules as a choice for any other calculation – as long as it can perform the self-consistent-field (SCF) procedure under the influence of the external field caused by the MM model.[214-216] Sources of error occur from the QM and MM methods themselves as in all computational methods, but the coupling of two methods adds an additional source of error: the coupling term must be calculated based on a choice for the nature of the QM/MM boundary. A boundary between the QM and MM regions of the system is clearly non-physical as it is an invention to make modelling of the system easier. As such there is not a perfect solution to treat this boundary region in a satisfying way; only various options from which the most reasonable should be chosen for the system at hand. For QM/MM studies there are three classes of boundary treatment: link atom schemes which introduce a new atom at the point that a covalent bond is cut to close the system; boundary atom schemes where the atom on the MM side of the bond is replaced by an atom which mimics the behaviour of the cut bond for the purposes of the QM simulation; and localised orbital schemes which replace the cut bond by hybrid orbitals which are mostly frozen to cap the QM region.[214] A set of detailed comparisons have been made assessing these three families of boundary treatments but there is no one solution, mirroring the computational cost against accuracy problem of most choices in computational chemistry.[217-224] Adaptive partitioning is a particularly interesting idea within QM/MM which allows for "on-the-fly" reclassification of atoms in order to move the boundary region in systems where the required location of high level QM calculations changes over the course of the simulation. This area of QM/MM is still in development and holds great promise for the increased flexibility of QM/MM simulations while retaining accuracy.[225-230]

Coarse-grained methods were originally proposed as a possible solution to the difficulty in modelling entire proteins and including all the atoms individually in the simulation. The premise was to collect sections of the protein together such that there are fewer interactions to consider while maintaining the overall structure.[206, 231] Using these coarse grains and including empirical interaction parameters between the grains or beads led to simplified ways of modelling protein dynamics and folding.[232, 233] Common implementations of these methods include single-bead,[234] two-bead[235, 236] and hybrid methods,[237, 238] where two-bead and hybrid methods include additional beads to account for the effects of the side chains on amino acids.

All-Atom/Coarse-Grained methods are built upon the same reasoning as QM/MM by reducing the majority of the system to a simpler model in order to perform more accurate calculations on the areas of interest,[239] and have also been used extensively in relation to protein dynamics[240-244] and folding.[245, 246] These methods have very similar challenges, applications and theory to QM/MM and therefore will not be detailed extensively here, although it is worth considering AA/CG methods as an alternative or complementary addition to QM/MM methods.

Quantum chemistry composite methods (QCCMs) are hybrid methods with a different goal in mind. Rather than reducing the computational cost by modelling some regions of the system with molecular dynamics, QCCMs combine several *ab initio* calculations in order to increase the accuracy of the final results, often with a view to obtaining more precise thermochemical and kinetic properties. A recent review makes it clear that the cost of these methods is such that it would be unreasonable to model anything much larger than a single amino acid, although they could be considered for use on small active sites in enzymes in the future as the QM part of a QM/MM calculation.[247] Some QCCMs have been used to study potential mechanisms within radical SAM enzymes, but they have to neglect the surrounding residues and solvent, with the system that remains being studied in the gas phase,[248, 249] reducing the scope of the calculation compared to QM/MM methods which have also been utilised on radical SAM enzymes.[212, 213]

Consideration of solvation can increase the accuracy of a computational model, particularly in the case of enzymes as both their initial folding and their reaction occur in the presence of solvent and should therefore not be dismissed as irrelevant without examination. Rather than explicitly modelling every solvent molecule, which becomes prohibitively expensive when using *ab initio* or other highly computationally intensive methods, we can approximate the solvent to be a polarizable medium giving the same (or sufficiently similar) properties as an explicit solvent model would give.[250] When applying this to DFT it takes the form of a perturbation to the Hamiltonian representing the electrostatic interaction between the solute and the solvent which can be recalculated at every step as the solute has an effect on the polarizable continuum, which in turn alters the effect of the continuum on the solute, becoming an iterative process.[251] In general, the continuum is defined primarily by its polarizability, given by the dielectric constant ε, and the size and shape of the cavity in the continuum in which the solute is embedded. While simple methods with basic cavity shapes such as spheres or ellipses can be envisioned, a powerful alternative is to use the Conductor-like Screening Model (COSMO)[252] which, in addition to a different, conductor-like boundary condition and treatment of solute electron density that reaches out of the cavity (termed "outlying charge"),[253] defines a solvent accessible surface of arbitrary shape such that the imagined centres of solvent molecules that would generate the continuum cannot exist within spheres of radius $R_{\propto}$, where

$$R_{\propto} = R_{\propto}^{\mathrm{vdW}} + R^{\mathrm{solv}}, \tag{2.19}$$

and where $R_{\propto}^{\mathrm{vdW}}$ is the van der Waals radius of atom $\propto$ and R^{solv} is the effective radius of the solvent molecules. The charges generated by these solvent molecules will be located away from the centre of the molecules at some distance δ^{sc}, dependent on the solvent being considered, resulting in the minimum distance of a solvent charge from a solute atom $\propto$ being given by

$$R_{\propto}^{*} = R_{\propto} - \delta^{\mathrm{sc}}. \tag{2.20}$$

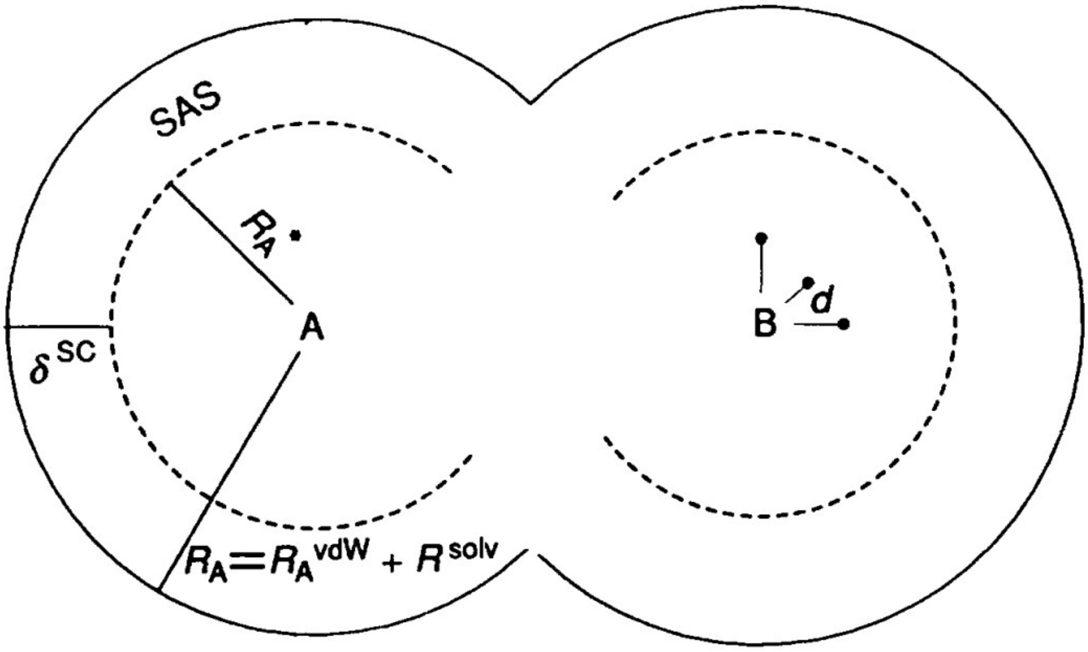


Figure 2.2: Construction of the SAS (Solvent Accessible Surface) (schematic); Solid circles indicate the surface accessible to the centres of solvent molecules, dashed lines the surface accessible to solvent charges. For atom B the three dipole representation points are indicated in addition. Figure and caption adapted from Klamt and Schüürmann.[252]

COSMO is a widespread method of modelling solvation effects in quantum chemical software packages and continues to be expanded upon due to this success.[254]

Chapter 3 – Effect of Oriented Electric Fields on Biologically Relevant Iron–Sulfur Clusters: Tuning Redox Reactivity for Catalysis

3.1 – Introduction and Summary

Enzyme-based iron-sulfur clusters, exemplified in families such as hydrogenases, nitrogenases and rSAM enzymes, feature in many essential biological processes. The functionality of biological iron-sulfur clusters extends beyond simple electron transfer, relying primarily on the redox activity of the clusters, with a remarkable diversity for different enzymes. The active site structure and the electrostatic environment in which the cluster resides direct this redox reactivity. Orientated electric fields in enzymatic active sites can be significantly strong and to understand the extent of their effect on iron-sulfur cluster reactivity can inform first steps towards rationally engineering their reactivity. An extensive systematic density functional theory based screening approach using OPBE/TZP has afforded a simple electric field effect representation. The results demonstrate that the orientation of an external electric field of strength 28.778 MV cm^{-1} at the centre of the cluster can have a significant effect on its relative stability in the order of 35 kJ mol^{-1}. This shows clear implications for the reactivity of iron-sulfur clusters in enzymes. The results also demonstrate that the orientation of the electric field can alter the most stable broken symmetry state, which further has implications on the directionality of initiated electron transfer reactions. These insights open the path for manipulating enzymatic redox reactivity of iron-sulfur cluster containing enzymes by rationally engineering orientated electric fields within the enzymes.

In this study, we directly and systematically investigate the impact of an external electric field on biologically relevant iron-sulfur clusters. We examine the fundamental effect of applying and reorienting a simple electric field on the stability and reactivity of selected model iron-sulfur clusters. The information on how an oriented electric field influences the reactivity of the clusters can later be combined with the knowledge and understanding of the contribution of individual amino acids to the electric field in an enzyme. This will allow bespoke control of iron-sulfur cluster reactivity and stability through mutations of the surrounding residues.[108] Consequences of these adaptations would include either the ability to improve existing reactions by increasing the rate of reaction, integrating oxygen tolerance, or exploiting enzyme selectivity for reactions that were previously only accessible through synthetic approaches.

3.2 – Computational Methods

Unrestricted geometry optimisation calculations using DFT were performed using the Q-Chem software package[255] and compared with literature values[256] using the same model systems, featuring the iron-sulfur clusters [2Fe2S] or [4Fe4S] and four methanethiolate ligands bound to the iron atoms; two per iron or one per iron respectively (Figure 3.1). The geometry of a third cluster, identical to the [4Fe4S] cluster but with one fewer methanethiolate ligands, was also optimised. The initial geometries were based on a crystal structure (PDB ID 1ZOY.) The hybrid OPBE functional,[257, 258] which consists of Handy's optimised exchange (OPTX)[177] and PBE[174] correlation, was used for both the geometry optimisations and further single point calculations. The use of this functional allows comparison with the work of Swart et al.,[256] and the spin states of the iron complexes predicted are consistent with literature studies.[257, 258] The TZP basis set[259-261] was chosen for both geometry optimisations and single point calculations to facilitate this comparison. Methods used by other groups include B(5%HF)P86 and a triple-ζ basis set with polarization functions for accurate bond covalency,[262, 263] and B(5%HF)P86/6-311+G(d) for the QM region of a QM/MM study into the protein environmental effects around iron-sulfur clusters.[264] While those levels of theory would be suitable for optimisations, we chose OPBE/TZP because it demonstrates both the correct structural predictions and accuracy in the relative ranking of spin state energies which is particularly important for this study.[265, 266] The geometry optimisations were also repeated using the polarizable continuum model COSMO[252] with a dielectric constant of 4.0 to simulate a protein environment.[267, 268]

We investigate the effect of an oriented external electric field on the vertical electron affinity (VEA) and the most stable state of the model systems by performing 1296 single point calculations in the presence of two external, equal but oppositely signed point charges, equidistant from the centre of mass of the system, whose bisector intersects this centre. The redox potential is an important property to consider when investigating the reactivity of iron-sulfur clusters. It is dependent on the stability difference of the oxidised and reduced state of the system and the reorganisation energy. The latter is influenced by the relaxation of the system upon electron transfer and the restructuring of the environment, often dominated by solvent reorganisation. The VEA accounts for the energy difference between the oxidised and reduced state without any relaxation (Eqn 3.1) and can be taken as a first approximation for the reactivity difference of the two oxidation states.

$$E_{\mathrm{VEA}} = E_{\mathrm{ox}} - E_{\mathrm{red}}, \tag{3.1}$$

where E_{VEA} is the VEA and is positive if the reduction of the oxidised state is energetically favourable, E_{ox} is the total electronic energy of the oxidised state and E_{red} is the corresponding energy of the reduced cluster with the oxidised cluster's geometry.

The effect of the direction of the external electric field was investigated by rotating point charges defining the field about two axes in the system and mapping them onto a sphere around the model systems. The axes of rotation used to orient the field around the clusters are shown and labelled in Figure 3.1. The notation is described further in the Appendix Section A3.1.

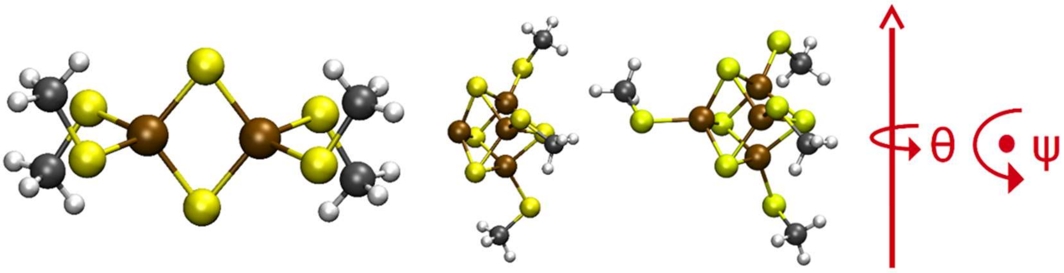


Figure 3.1: Definition of axes for the 2Fe2S, (left) labile 4Fe4S, (middle) and 4Fe4S (right) clusters. All geometries shown are optimised structures for the oxidised clusters at the OPBE/TZP level of theory in a vacuum.

Systems containing atoms with multiple possible oxidation and spin states can be described in several ways. The reduced state of the 2Fe2S cluster introduced earlier could have an oxidation state of +2.5 assigned to both iron atoms, a symmetrical and simplified description. In many cases however it is necessary to define broken symmetry oxidation or spin states, where the additional electron might be fully localised on one of the iron atoms, for systems with ferromagnetic and antiferromagnetic coupling, for example.[193, 269] This would be particularly important if a specific iron atom is involved in a reaction, as the oxidation state would affect the reactivity, or govern whether the reaction would progress at all.

However, it should be noted that Kohn-Sham DFT methods in general have limitations in describing multireference electronic systems such as our model iron-sulfur clusters, and we cannot easily assign formal oxidation states directly for them. To qualitatively verify the results obtained from the broken symmetry approach (see page 28), we can however conduct modified calculations to elucidate useful information and obtain electronic states that are as close as possible in energy to the true pure spin ground state. Here, restricted Open-Shell (ROS)[192] DFT enforces a specific spin state in situations that would otherwise require an unrestricted calculation, such as the reduced [2Fe2S] cluster which has one unpaired electron. This will result in a pure spin state that, however, will be higher in energy than the result from an unrestricted calculation on the same system. Broken-Symmetry DFT (BS-DFT)[116] can also be used to investigate the different broken symmetry states for a single spin state. Using the broken symmetry states and the maximum spin pure spin state we can also calculate the Heisenberg coupling constant J which can provide a direct comparison to experimental results.

There are two broken symmetry states for the 1/2 spin system where the electron density of the unpaired electron is localised to one of the iron atoms, analogous to the system where the formal oxidation states of the iron atoms of the 2Fe2S cluster are +2/+3 and +3/+2. A single point calculation using a default guess wavefunction for the 1/2 mixed spin reduced cluster using one of the point charge positions will give one of the two broken symmetry states. By taking the final wavefunction from each of these situations and using them as guess wavefunctions in two calculations without point charges but preserving the orbital occupancies, we calculated two wavefunctions, one for each broken symmetry state, which do not have orbitals influenced by an external field. We used these wavefunctions as initial guesses when we added point charges, allowing us to see the full range of field effects for both broken symmetry states. Geometry optimisations and subsequent frequency calculations were performed in vacuum and using COSMO with a dielectric of 4.0 for the model 2Fe2S, 4Fe4S and labile 4Fe4S iron-sulfur clusters for the four reported mixed spin states in each case.[256]

3.3 – Effect of an Oriented Electric Field on the Energies and Vertical Electron Affinities of Model Iron-Sulfur Clusters

The energy relative to the calculated ground state for each system is reported in Table 3.1. A brief discussion of the considered spin states is presented in Appendix Section A3.2. All optimised geometries that were used in later calculations were confirmed as minima by frequency calculations. The only case where the ground state is not the low spin state is from the reference value for the reduced 2Fe2S cluster; this is not surprising as the 1/2 and 9/2 spin states are isoenergetic. The geometry and frequency calculations for each spin state agree with the literature values,[256] within 5 kJ mol^{-1} (Table 3.1) and are consistent in the relative stability of each state and can therefore be justified for use in the subsequent investigations for the purpose of determining relative stability and reactivity when a directed external electric field is applied. The absolute energy and $\langle S^2 \rangle$ values for these optimisations are reported in Table 3.2. These values for the optimised geometries are high, as expected for mixed spin states – the theoretical $\langle S^2 \rangle$ values for pure spin states would be 0 for the oxidized clusters and 0.75 for the reduced clusters. Notably the mixed spin $\langle S^2 \rangle$ values display a smaller magnitude of difference between the oxidized and reduced states than the pure spin states, perhaps indicating that the properties of the mixed spin states are more like each other than their pure spin counterparts.

The optimisations in a COSMO environment were generally lower in energy but showed the same trends in stability as the vacuum environment. This is to be expected as prior work found that both PCM (Polarizable Continuum Model) and COSMO approaches converge to very similar structures as the gas phase optimisation, although there were some exceptions.[270] COSMO has also been used in the successful prediction of Mössbauer spectral parameters,[271] reduction potentials[272] and other properties of iron-sulfur clusters and similar molecules.[273] Furthermore, experimental work has shown that the redox potential of biological iron-sulfur clusters is significantly dependent on the environment, when considering different enzymes,[274] or modifying the ligands of the cluster.[89] Optimisations on the labile

4Fe4S cluster can also be justified using the same reasoning and could be used for further study of a similar nature when investigating the behaviour of the cubane cluster where one iron atom is not bound to a ligand (for example SAM in the radical SAM enzyme superfamily).

Table 3.1: Relative energies for OPBE/TZP geometry optimized model iron-sulfur clusters in different mixed spin states, compared with literature values[256] in both vacuum and within a COSMO environment using a dielectric of 4.0. Values presented in kJ mol^{-1}. All values correspond to geometry optimized structures at the given oxidation and mixed spin state. Absolute energy values and <S^2> are presented in Table 3.2.

2Fe2S	**Mixed Spin, mS**			
Oxidised	**mS=0**	**mS=1**	**mS=2**	**mS=5**
Reference	0.00	54.34	104.21	72.64
Vacuum	0.00	65.93	110.44	84.06
COSMO	0.00	57.44	100.68	83.69
Reduced	**mS=1/2**	**mS=3/2**	**mS=5/2**	**mS=9/2**
Reference	0.67	84.11	100.57	0.00
Vacuum	0.00	92.64	98.74	7.47
COSMO	0.00	84.24	91.71	4.16
4Fe4S	**Mixed Spin, mS**			
Oxidised	**mS=0**	**mS=1**	**mS=2**	**mS=9**
Reference	0.00	48.61	82.94	156.38
Vacuum	0.00	45.99	81.33	177.29
COSMO	0.00	44.85	79.85	181.25
Reduced	**mS=1/2**	**mS=3/2**	**mS=5/2**	**mS=17/2**
Reference	0.00	62.09	66.15	63.97
Vacuum	0.00	59.64	66.13	74.39
COSMO	0.00	50.65*	61.11	77.45
Labile 4Fe4S	**Mixed Spin, mS**			
Oxidised	**mS=0**	**mS=1**	**mS=2**	**mS=9**
Vacuum	0.00	58.45	53.92	160.14
COSMO	0.00	56.53	56.28	163.19
Reduced	**mS=1/2**	**mS=3/2**	**mS=5/2**	**mS=17/2**
Vacuum	0.00	64.99	45.80	66.36
COSMO	0.00	56.28	46.79	58.52

*) optimized state with one imaginary frequency. This state was not used in any further calculations.

Table 3.2: Calculated absolute energy and $\langle S^2 \rangle$ for each of the values defined as 0 in Table 3.1 reported in hartree.

Energy ($\langle S^2 \rangle$)	2Fe2S	4Fe4S	Labile 4Fe4S
Oxidised			
Vacuum	-5077.04022 (4.22)	-8401.72576 (6.98)	-7963.61182 (6.85)
COSMO	-5077.22719 (4.23)	-8401.89927 (6.93)	-7963.66590 (6.86)
Reduced			
Vacuum	-5076.83985 (4.28)	-8401.55849 (7.09)	-7963.56871 (7.35)
COSMO	-5077.24384 (4.25)	-8401.93079 (7.04)	-7963.74428 (6.91)

Using VEA as a surrogate for the redox potential is not a new approach but offers an approximation with the potential for high throughput, needed to systematically study an orientated electric field effect. Previous work has gone further to approximate redox potentials from DFT calculations by considering electronegativity,[275] electrophilicity[276-278] and combinations of properties[279, 280]. While these methods would provide more experimentally comparable values, the VEA provides a sufficient description of the change in these properties and is straightforward to compute. This makes it feasible to investigate the effect of a rotating electric field via thousands of individual single point DFT calculations.

The relative energies are in good agreement with literature values obtained for these systems with the order of spin state stability preserved. Slight differences in the energies compared to literature are likely to originate from different implementations of the implicit solvent method (COSMO) and different convergence criteria in the used programmes but do not influence the qualitative agreement.

Single point calculations were performed on the model iron-sulfur clusters in the presence of an electric field of varying orientation for the oxidised and reduced states. The optimised geometry of the ground state spin oxidised cluster was used in all cases to allow the calculation of the VEA. Figure 3.2 displays the effect of the rotation of the electric field on the energy of the two most stable 9/2 and 1/2 spin states of the reduced 2Fe2S model cluster about the principal axes presented in Figure 3.1.

The electric field applied to the cluster consists of two external, equal but oppositely signed point charges, equidistant from the centre of mass of the system, whose bisector intersects this system. The electric field at this centre is a vector with only one component direction since the charges lie on a straight line passing through the origin, where both charges are equidistant from the origin with charges of the same magnitude and opposite signs. This allows us to simply calculate the electric field strength caused by one of the charges and double it. The field strength generated by a point charge with a radial field is given by

$$Q/4\pi\varepsilon_0 r^2, \tag{3.2}$$

where Q is the magnitude of the charge, ε_0 is the vacuum permittivity and r is the separation from the charge to the point we want to measure the field strength at. For this case, Q is the charge on the electron e ($1.60\text{x}10^{-19}$ C) and r is 10 angstroms. Substituting these values into the equation and resolving gives 14.389 MV cm^{-1} for a single charge and therefore 28.778 MV cm^{-1} for the whole field. This field strength is representative of a weak field generated by a protein based on the range of values used in simulations[281] and obtained by experiment.[111, 282, 283]

As can be seen from Figure 3.2, reorienting a directed electric field can influence the electronic stability of the 2Fe2S cluster significantly. Depending on the axis of rotation, the effect on the stability can vary up to 23 kJ mol^{-1} (in the case of rotating around the normal to the plane of the cluster (Normal), Figure 3.2a, d and down to only 1.8 kJ mol^{-1} (in the case of rotating around the FeFe axis, Figure 3.2c). Each spin state is influenced differently in relation to the field orientation, which leads to a varying energy gap between the spin states. However, this variation is never large enough to lead to the 9/2 spin state being the ground state (Figures 3.2a-c). When comparing the electric field effect between the gas phase (Figure 3.2a) and an implicit solvation with a low dielectric constant of 4.0 (Figure 3.3a), the observed effect is maintained, but the energy gap between the two spin states decreases slightly.

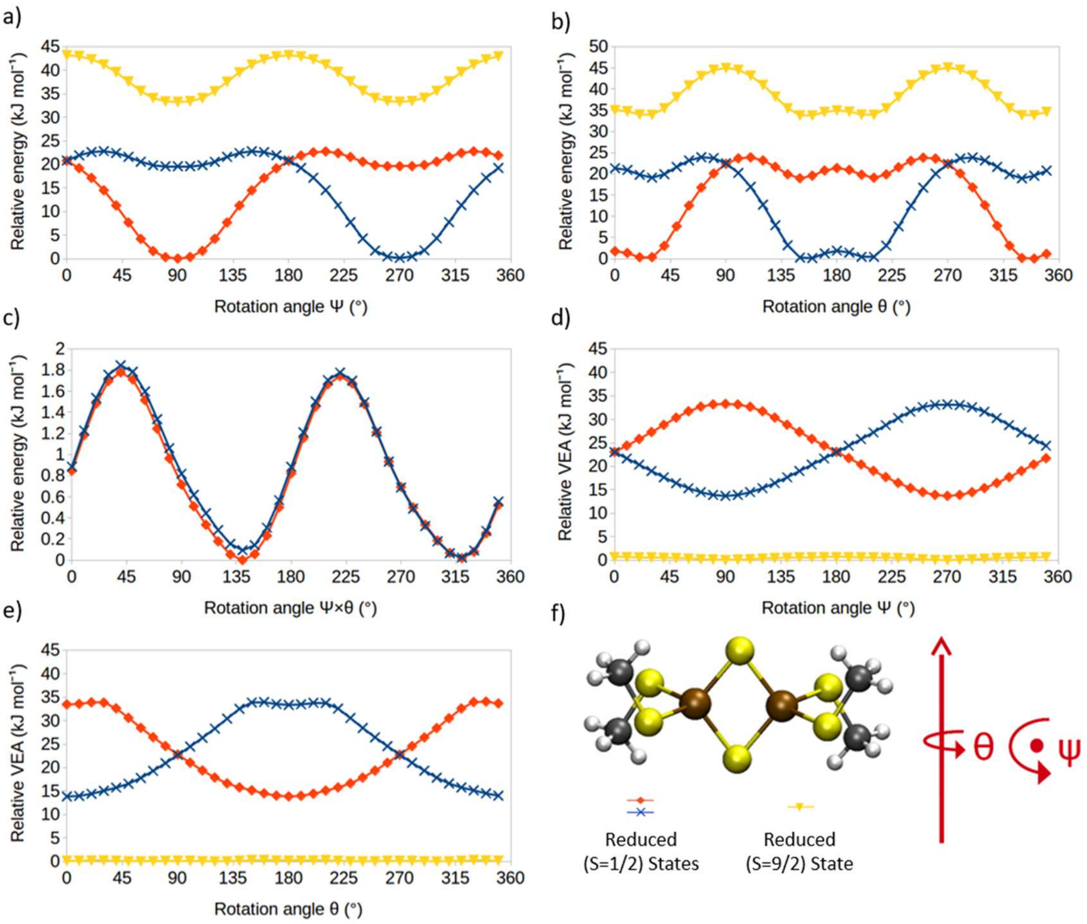


Figure 3.2: Relative energy profiles of the reduced 2Fe2S model cluster (a-c) and the VEA (d, e) of the same cluster with the orientation of an external electric field in the gas phase. Rotation angle indicates the rotation from the initial position of the field about the Normal (a, d), the S-S axis (b, e) or the Fe-Fe axis (c) (ψ x θ represents the rotation around the axis perpendicular to those shown in panel f).

The differences between Figures 3.2a, b, c and 3.3a, b, c are generally in magnitude rather than nature. For both a and b, the addition of a COSMO environment widens the energy gap between the two broken symmetry 1/2 spin states, where the maximum relative energy of these states is increased by approximately 5 kJ mol^{-1}. The 9/2 state also appears to slightly stabilise, closing the gap between the 9/2 and the higher of the 1/2 spin states. However, this still does not result in the 9/2 state being energetically favourable. The rotation about the Fe-Fe axis c shows a decrease in relative energy and a change in the nature of the curve with some perturbations between 90 and 135 degrees. This could be a small effect that was previously hidden by the much larger magnitudes of the other energy profiles but is visible when the relative energy is much smaller. These small differences compared to an increase in computational cost contributed to our choice to continue gas phase calculations only for the full three dimensional analysis.

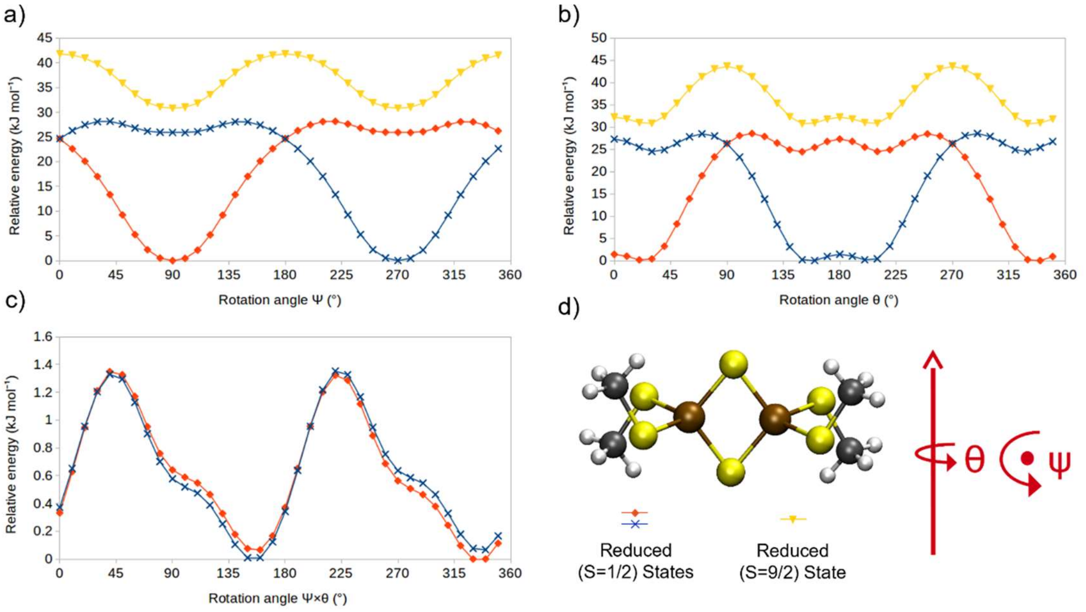


Figure 3.3: Relative energy profiles of the reduced 2Fe2S model cluster with the orientation of an external electric field with the implicit solvation model COSMO at a dielectric of 4.0. Rotation axes are analogous to a) Figure 3.2a, Normal, b) Figure 3.2b, S-S axis, c) Figure 3.2c, Fe-Fe axis. Figure 3.2f is repeated here in panel d for clarity.

Next, it was investigated how the electric field influences the stability of different broken symmetry states which showed a significant effect on the relative stability of those. When rotating the field around the Normal (Figure 3.2a), the most stable spin state flips when the field is exactly aligned along the S-S axis. This means that, depending on the orientation of an external electric field, the cluster adapts by adopting a different broken symmetry state. As these different states will show different reactivity in directed electron transfer, notably in relation to redox potential,[267] they will also show different reactivity patterns with e.g., different substrates reacting with the cluster in an enzymatic active site. The rotation about the Fe-Fe axis causes very little variation in VEA, with the value for the 1/2 spin states remaining between about 22 and 23 kJ mol^{-1} and the 9/2 spin state between 0 and 1 kJ mol^{-1}.

The relationship between the stability of the oxidised and reduced clusters and the changes to the redox reactivity of the clusters are inferred from the VEA calculations for the cluster from the data presented above. Figure 3.2d depicts how the VEA varies with rotation about the Normal. Depending on the rotation axis, ΔE_{VEA} varies by up to 20 kJ mol^{-1}. The results also suggest that the 9/2 spin state of the reduced cluster is unlikely to be involved in redox reactions as it is both less stable and has a lower electron affinity than both 1/2 spin states.

To extend this analysis of the principal axes of the 2Fe2S cluster, we systematically scanned the orientation of the electric field about multiple axes to create a 3-dimensional visualisation. Figure 3.4 depicts the analysis for the 2Fe2S cluster. To understand why a differently orientated electric field influences the stability and reactivity of the 2Fe2S, we then matched these observations to the symmetry of the frontier orbitals of the reactive species. The stability pattern observed for the principal axes' rotations described above is reflected in the analysis of the full rotations. When looking at the VEA, some electric field orientations show significantly higher reactivity compared to others. Considering the individual stabilities of the oxidised and reduced states one can further see the relation with the frontier orbital occupancies, as also depicted in Figure 3.4. Looking at the differences for the two reduced broken symmetry spin states, the effect of the directed electric field is mirror symmetric for each state. The cluster is stabilised when the side with higher electron density of the HOMO is close to the positive charge defining the electric field. Due to the different orbital occupancy of the oxidised state (fully occupied HOMO) this species shows a higher symmetry in reference to the effect of the electric field.

The model iron-sulfur clusters considered in this study are examples of multireference systems that can only be described by DFT with limitations, resulting in describing mixed spin states rather than pure spin states. Broken symmetry DFT provides electronic states that are as close as possible in energy to the pure spin ground state but will allow the introduction of spin contamination. This, however, is the only reasonable choice for a high-throughput study. With this in mind, we can assess the divergence of the computed $<S^2>$ values from the expected $<S^2>$ of a pure spin state for the system to understand the magnitude of spin contamination. While the magnitude of spin contamination is large, it remains somewhat consistent through rotations of the electric field. For example, the reduced [2Fe2S] cluster would have a pure spin state of 0.5, resulting in an $<S^2>$ of 0.75, while $<S^2>$ for the broken symmetry states varies from about 4.09 to 4.17 (Figure 3.5). Similar effects are seen with the other model clusters, consistently small deviations caused by the electric field, but a large overall spin contamination.

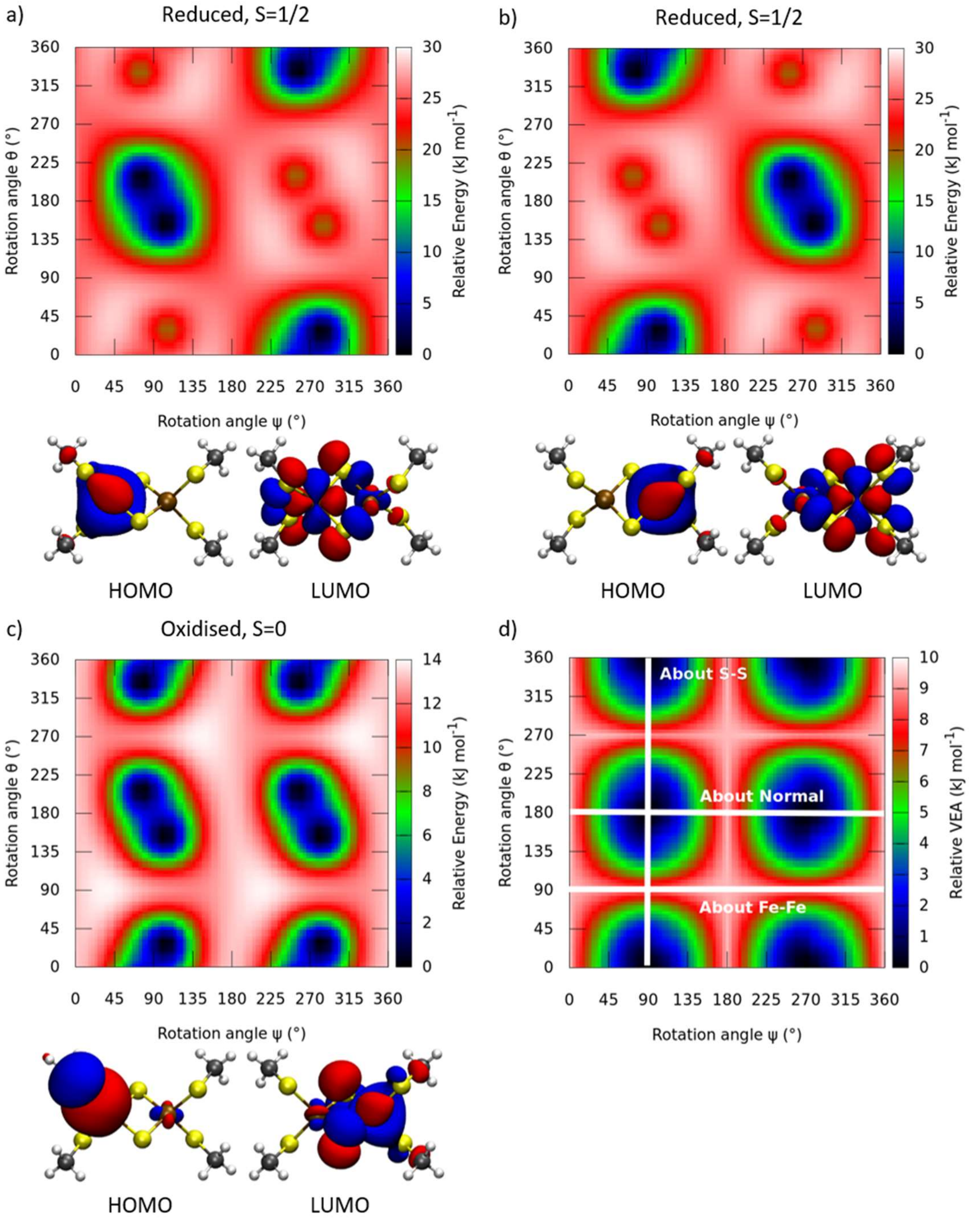


Figure 3.4: Frontier orbitals and effect of the three-dimensional rotation of an orientated electric field around the 2Fe2S cluster. a) Reduced, S=1/2, Broken symmetry state 1; b) Reduced, S=1/2, Broken symmetry state 2; c) Oxidised, mS=0; d) Relative vertical electron affinity with most stable reduced broken symmetry state.

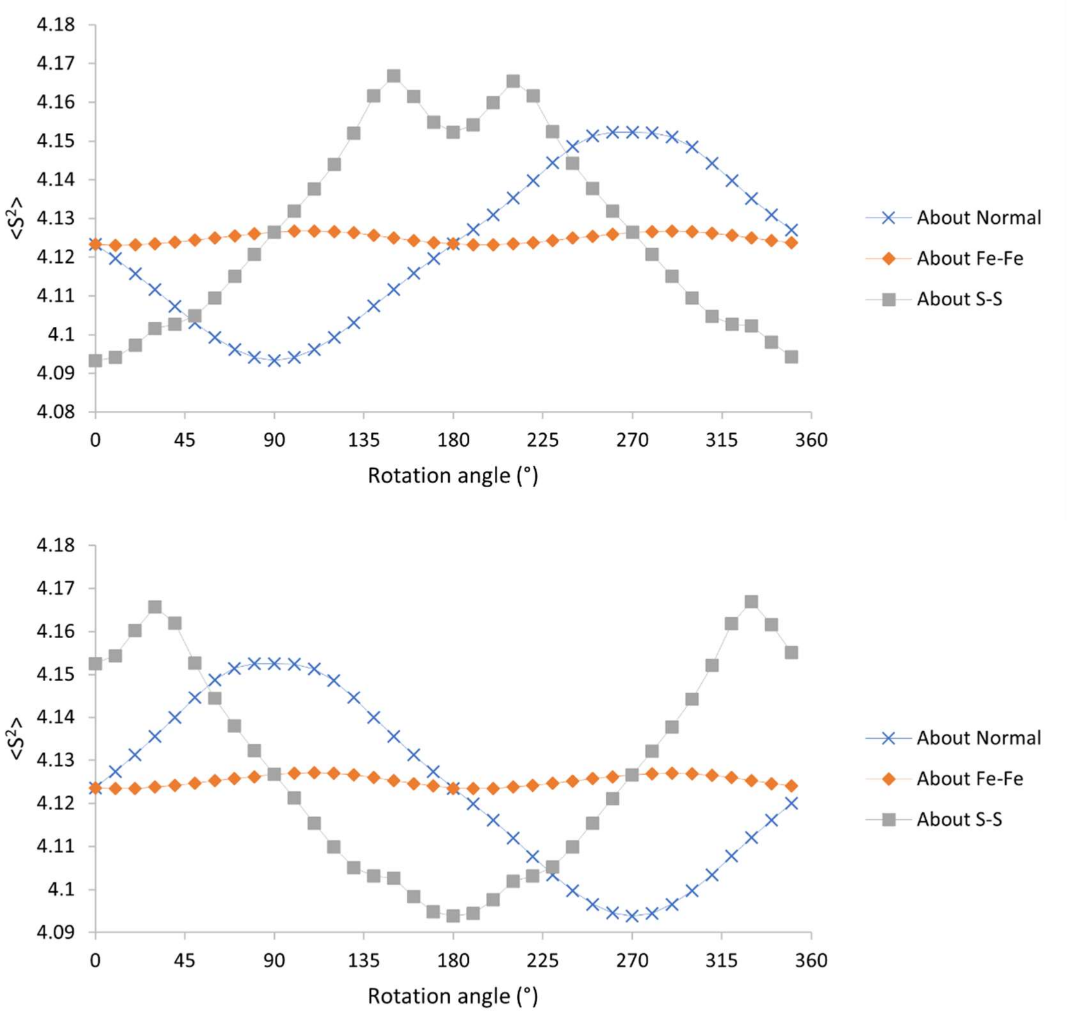


Figure 3.5: $<S^2>$ as a function of rotation angle about the three principal axes of the reduced [2Fe2S] cluster, for both mS = 1/2 broken symmetry states.

To validate the applicability and quality of the broken symmetry approach for our investigations we have also derived the Heisenberg Coupling Constant J (Eqn 2.17) of this system when rotating an electric field around the principal axes detailed above, and as such find the effect of the electric field on the spin coupling constant J (Figure 3.6).

Overall, this analysis confirms that the profile of the energy change with a rotating electric field is the same for a pure spin state and a broken symmetry state for these systems, with the pure spin state higher in energy. While the effect on the oxidised [2Fe2S] cluster only has a range of about 1.5 cm^{-1} it is noticeable that the magnitude of the effect is similar for all rotation axes, while the energy changes for the Fe-Fe axis rotation was an order of magnitude smaller than the other axes. As such, the changes in energy around this axis would imply very little about the system is changing compared to other axes of rotation, but there are other quantities that will still be affected by the field direction around this axis such as J.

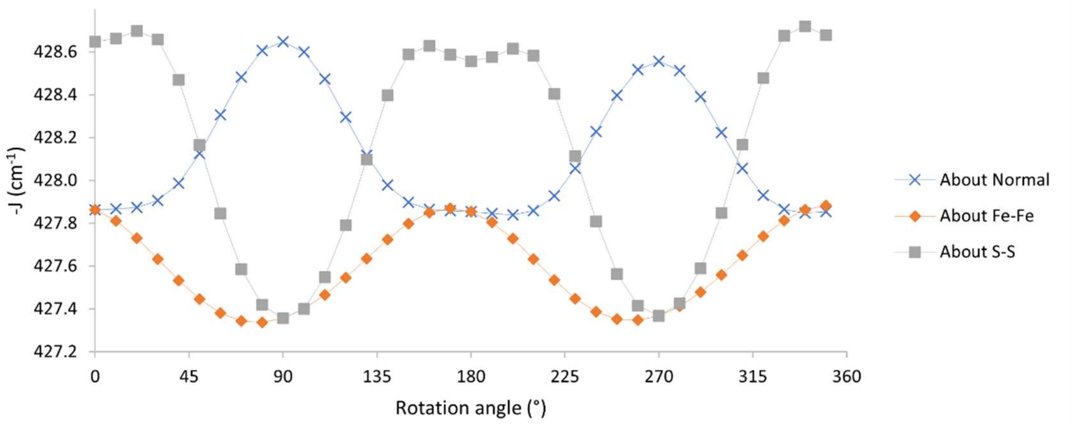


Figure 3.6: Effect of a rotation electric field around the three principal axes of the oxidised [2Fe2S] model cluster on the Heisenberg Coupling Constant J.

The value of J extracted using this method lies within the range of values reported in the literature for the same or very similar systems, noting that there seems to be a common overestimate of J from theoretical calculations compared to experimentally derived values.[46, 193, 199, 284]

We can also further compare the broken symmetry state energetics with the high spin situation which will produce a pure spin state and, using this along with the Heisenberg Double-Exchange (HDE) model,[46, 116, 193, 194, 199] we can reconstruct the spin ladder and provide values for the spin-coupling constant, J, and for the reduced cluster, the double-exchange constant, B. The dependence of these values on the electric field direction can then be shown which allows us to directly connect these computed properties to experimental observables. While this method has been cautioned against,[284] it is still valuable to derive estimates of these values.

Following the HDE model, for the oxidized cluster there is a simple relationship between the energy of a broken symmetry state and the energy of the highest pure spin state,

$$E(S_{\mathrm{max}}) - E_{\mathrm{B}} = -S_{\mathrm{max}}^{2}J, \tag{3.3}$$

where $E(S_{\max})$ is the energy of the highest possible pure spin state of the system, E_{B} is the energy of the broken symmetry state, $S_{\max}$ is the spin of the highest spin state and J is the Heisenberg coupling constant (noting that S^2 in Eqn 3.3 is not $S(S+1)$ but $S \times S$ as per Eqns 32-34 in Noodleman's 1981 formulation[116]). As we have already computed E_{B} for various broken symmetry states under the influence of multiple field directions, we can demonstrate the equivalent effect for the highest pure spin state and therefore find the effect of the field direction on J.

The reduced cluster is slightly more complex, requiring not only $E(S)$ for the highest pure spin state, but for all other pure spin states as well. This then creates a "spin ladder" which has levels,

$$E(S) = JS(S+1) \pm B(S + \tfrac{1}{2}), \tag{3.4}$$

where $E(S)$ is the energy of the pure spin state S and B is the double-exchange constant. When rearranged it becomes clear that this relationship will produce a quadratic equation in S,

$$E(S) = JS^2 + (J \pm B)S \pm \frac{B}{2}, \tag{3.5}$$

such that we can plot each energy value and derive J and B from the equation of the resulting quadratic. To find the energy of the required pure spin states for both the oxidized and reduced [2Fe2S] clusters we performed Restricted Open-Shell DFT calculations to prevent spin contamination.

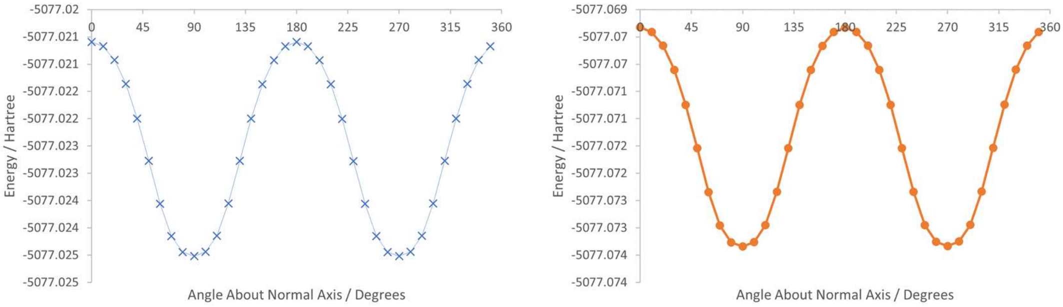


Figure 3.7: Energy of the oxidized [2Fe2S] cluster with changing field direction for a pure spin (left) and a broken symmetry (right) situation. While the energy profiles appear identical the mixed spin state is always lower in energy than the pure spin state.

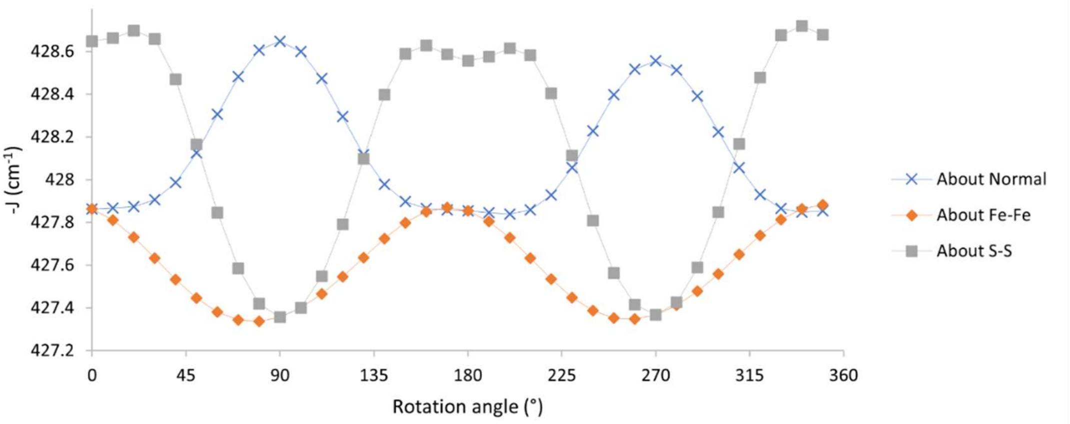


Figure 3.8: -J as a function of changing field direction about the three principal axes of the oxidized [2Fe2S] cluster.

As expected, the pure spin state describes a higher energy system with no spin contamination. The similarity of the energy profiles (Figure 3.7) is encouraging, as it suggests that the effect of the electric field on the energetics of the system is independent of the precise nature of the spin state of the system. Figure 3.8 is interesting, as although we have shown that the magnitude of the variation in energy with changing field direction about the Fe-Fe axis is, unsurprisingly, small compared to other rotational axes as the iron atoms are experiencing very similar fields, the magnitude of the dependence of J on field direction is similar regardless of rotational axis.

When considering the reduced [2Fe2S] cluster we must reconstruct the spin ladder for each field direction by performing calculations at each spin state and studying the resulting quadratic. We have four spin states to calculate for the reduced cluster, S=1/2, S=3/2, S=5/2 and S=9/2, resulting in, for example, Figure 3.9, and using Eqn 3.5 we can extract J and B from the equation of this quadratic in S. For Figure 3.9 the resulting coefficients are displayed in Table 3.3. Due to the performance of the HDE model, the quality of our calculations, or both, the resulting values do not follow this expectation as they are not internally consistent and therefore depend on the order in which we calculate them. We can then find a range of possible values that this model has produced.

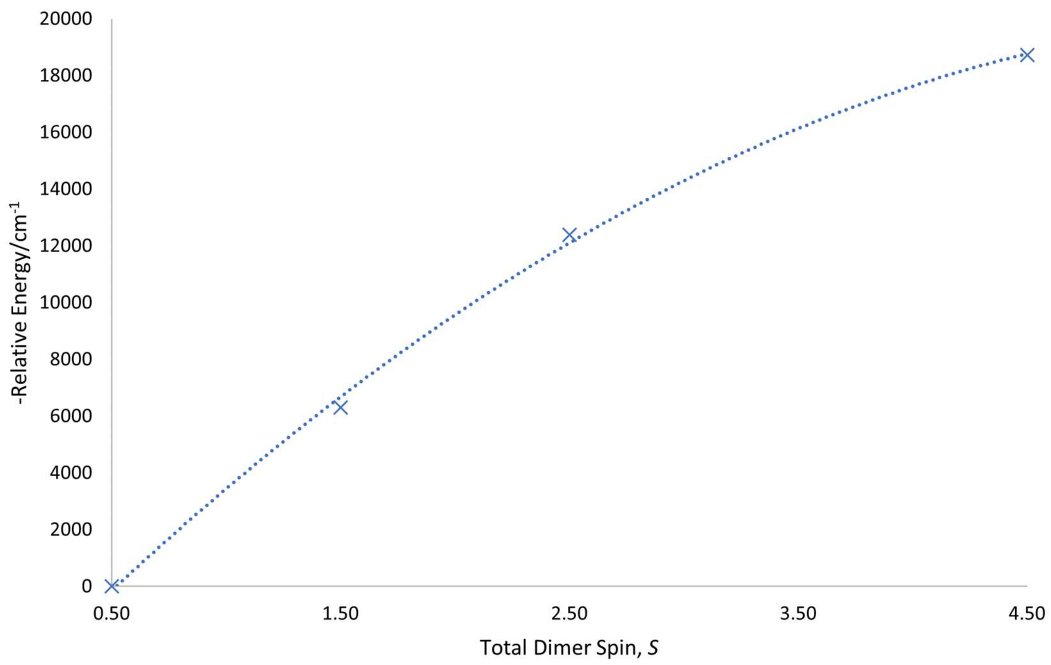


Figure 3.9: Example of the resulting spin ladder displaying a quadratic when -Relative Energy is plotted versus Dimer Spin, S, as in Eqn 3.5.

Table 3.3: Quadratic coefficients and resulting values for J and B corresponding to Figure 3.9.

Quadratic Coefficients	**a**	**b**	**c**
	-698.36	8219.41	-4077.12
Resulting Constants	**J from a=J**	**B_+ (B if c=B/2)**	**B_- (B if c=-B/2)**
	-698.36	-8154.24	8154.24
$J+B_+$	**$J+B_-$**	**$J-B_+$**	**$J-B_-$**
-8852.6	7455.88	7455.88	-8852.6

In Table 3.3 we show the method producing the most internally consistent values (we expect coefficient b to be equal to $J \pm B$) and as such will be the chosen way to calculate these constants. The resulting uncertainty in J becomes more apparent when we plot J as a function of field direction for the reduced cluster, like Figure 3.8 for the oxidised cluster, although for clarity we separate the major axes into individual plots.

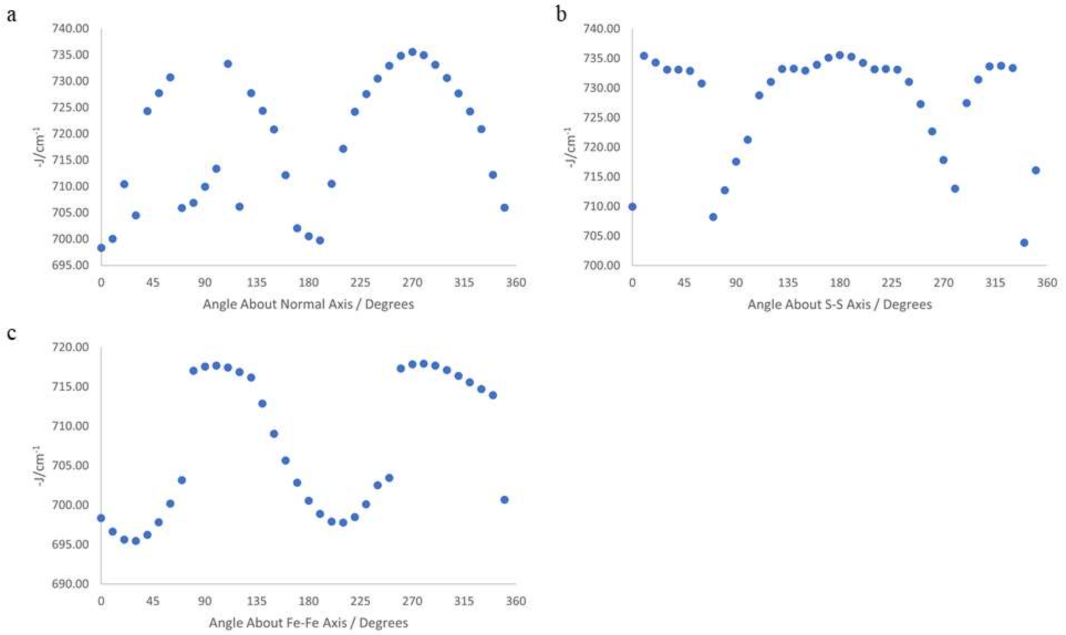


Figure 3.10: $-J$ with changing field direction rotating around a) the Normal, b) the S-S axis and c) the Fe-Fe axis.

While the general trends can be implied, and these trends follow similar patterns to energy as a function of field direction, there are what appear to be anomalous values, particularly in the 0-130 degrees region of Figure 3.10a. It seems that we can conclude that this method might be useful for gaining a general qualitative feel for the dependence of J on electric field direction, it might not be suited to deriving precise values for J or B. We again point to arguments made by Neese and Chan[284] about the accuracy of this model, and we stress that it is useful for understanding relative effects rather than deriving absolute values.

The cubane cluster displays more complex features. While the 2Fe2S cluster showed clear symmetry for the oxidised state, and a clear asymmetry related to the positioning of the unpaired electron density in the reduced states, the 4Fe4S cluster appears to display similar patterns of stability regardless of oxidation or broken symmetry state. The magnitude of these patterns varies, however. The relative energy of the cubane clusters with respect to the rotation of the field is influenced by the location of the HOMO seen in Figures 3.11 and 3.12. The significant change in the location of electron density between these two broken symmetry states results in the quadrants containing the maxima and minima being reversed. The directions of the field that result in the greatest stability remain broadly the same for the reduced and the oxidised states. This suggests that the external electric field has a lower polarizing effect on the 4Fe4S cluster compared to the 2Fe2S cluster and therefore a much lower effect on stability. Any of the four rotations where ψ, θ = 90°, 270° show the largest degree of variation in relative energy. These rotations lie along two perpendicular internal quadrilateral planes of the cluster that are coincident with its edges. The maximum VEA difference observed for the electric field is ΔE_{VEA} = 25.43 kJ mol^{-1}.

The labile cubane cluster also shows only a difference in magnitude rather than location of the lowest energy orientation. There is a much less complex pattern for this molecule than for the [4Fe4S] cluster The points at which the field is aligned along a vector passing through the edge created by the labile iron and a sulfur atom have the highest VEA, and a perpendicular vector has the lowest VEA, suggesting that the effect of this labile iron atom is dominant. Here, the maximal influence of the electric field on the VEA is not along an axis involving the unique iron. The maximum electric field effect on the VEA observed for this cluster is ΔE_{VEA} = 19.97 kJ mol^{-1}. Previous work on similar[285] and related[286] structures shows that the values for the changes in VEA are both significant relative to the absolute VEA but not unreasonably high in magnitude. Notably these VEA values fall within the 0-0.51 eV (0-49 kJ mol^{-1}) range of magnitudes predicted when varying surrounding and connected amino acids around iron-sulfur clusters.[264] It is also of interest to note that the field vector resulting in the maximum and minimum values of relative VEA are the same for both the 4Fe4S and Labile 4Fe4S clusters, approximately passing through a methyl group, the bonded sulfur and iron, and finally through the centre of the cube and the opposite sulfur atom. Since the maxima and minima observed for the VEA also influence electron transfer reactions it is not surprising that the observed field vectors (for maxima and minima) also follow the direction of the initial reductive electron transfer necessary for the formation of the 5'-deoxyadenosyl radical in radical SAM enzymes.[287] As rSAM enzymes share the common feature of SAM bound to the cluster this demonstrates how those enzymes may have evolved to arrange the active site to follow the minimum energy pathway for this initial activation step for the enzymes. Additional electrostatic field effects initiated by other charged residues nearby can either support or reduce this effect.

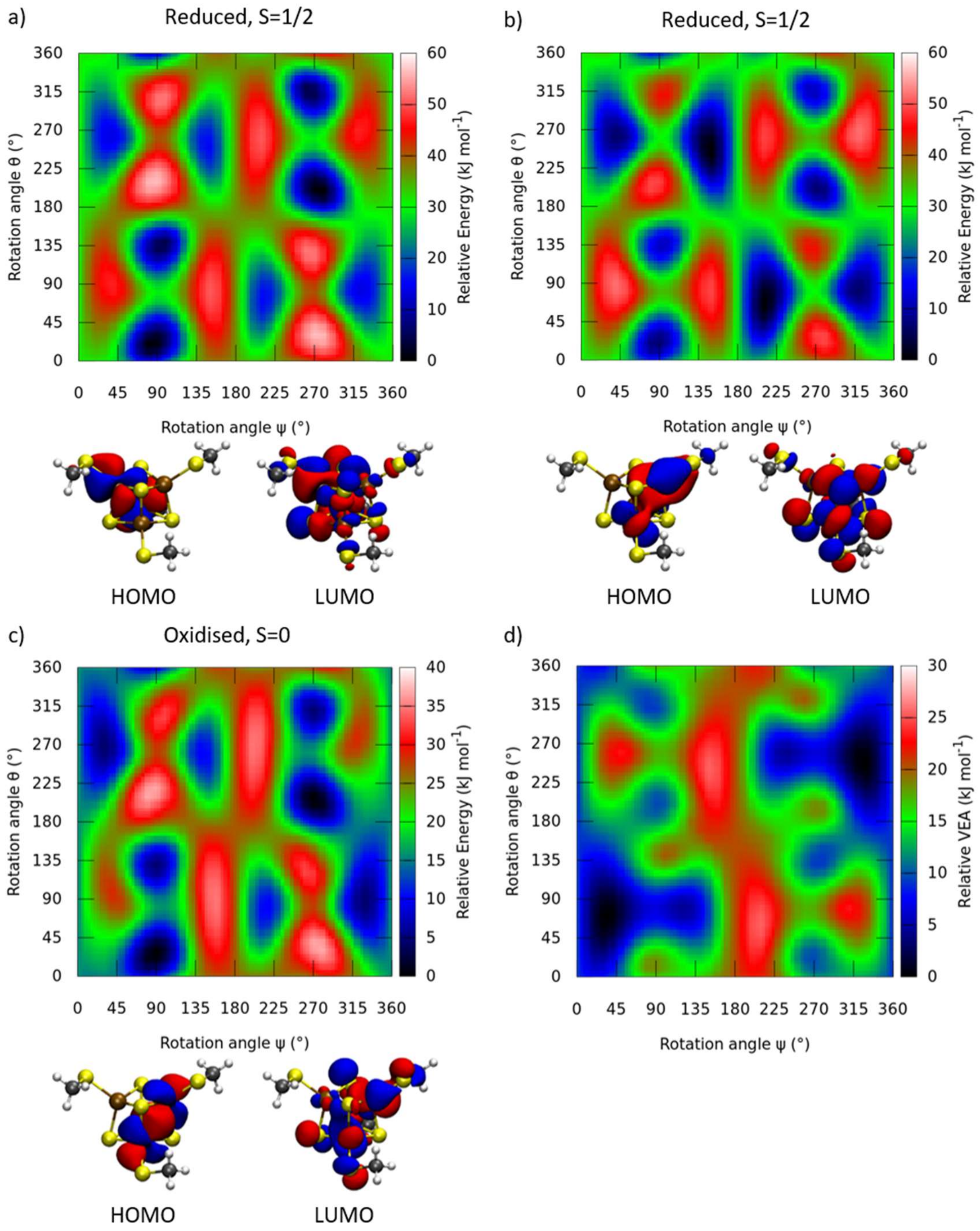


Figure 3.11: Frontier orbitals and effect of the three-dimensional rotation of an orientated electric field around the 4Fe4S cluster. a) Reduced, S=1/2, Broken symmetry state 3; b) Reduced, S=1/2, Broken symmetry state 4; c) Oxidised, mS=0, Broken symmetry state 4; d) Relative vertical electron affinity with most stable reduced broken symmetry state.

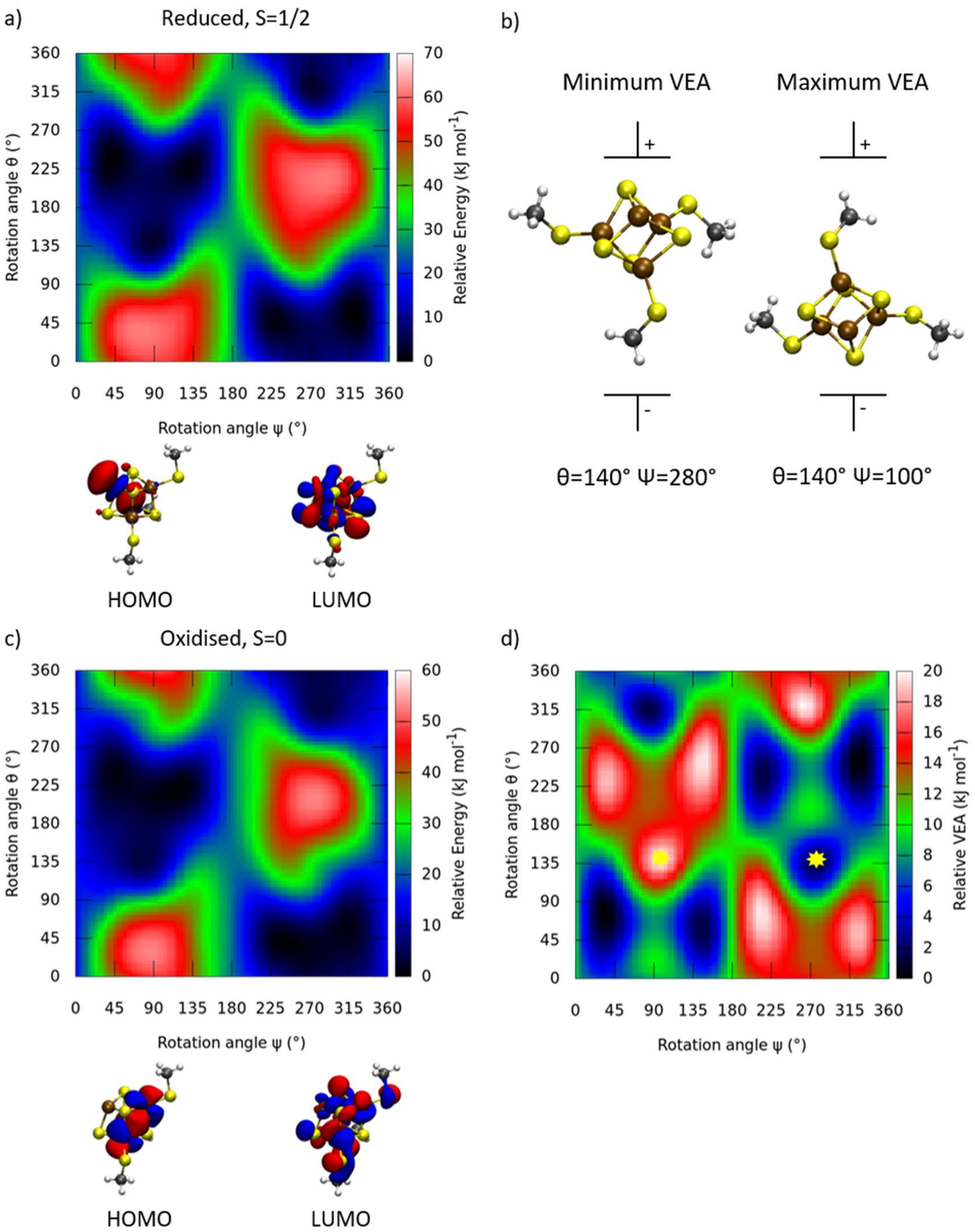


Figure 3.12: Frontier orbitals and effect of the three-dimensional rotation of an orientated electric field around the labile 4Fe4S cluster. a) Reduced, S=1/2; b) Exemplary orientation of the field for the maximum and minimum values of the VEA, indicated in d) by a yellow asterisk; c) Oxidised, mS=0; d) Vertical electron affinity.

While applying a simple electric field allows for a more intuitive analysis of the effect on the iron-sulfur clusters, the field generated by an enzyme would be much more complex, particularly with regards to homogeneity. A systematic study of the same depth as presented in this chapter would be prohibitively expensive when applied to a large range of fields and due to the highly complex nature of the fields in question the results may also be too complex to interpret. The above work demonstrates that a directed electric field of a strength reasonably similar to the net field one would find in an enzyme has a significant effect on the electronic properties of iron-sulfur clusters, so while there are differences between these directed fields and enzyme environments, if there is a correlation between the net field and the properties of the clusters this could be exploited by implying property changes based on the calculated net field generated by an enzyme rather than fully evaluating the effects with systematic DFT calculations.

A first step towards this possible outcome would be to identify effects likely to be caused by enzymatic fields; for example, the relative ranking of the mixed and broken symmetry spin states. To briefly investigate the most extreme changes in the enzyme environment we carried out a small number of calculations with regards to the labile form of the [4Fe4S] cluster using two snapshots of Pyruvate formate-lyase activating enzyme, one with SAM and one without. All but the iron-sulfur cluster atoms were represented by point charges with magnitudes extracted from the .pqr files for these snapshots. Initial geometries taken from the .pqr files were optimised in vacuum at the same level of theory as previous geometry optimisations and mixed and broken symmetry spin states were induced in the same way as previous calculations.

Table 3.4: Relative energies for the mixed spin states of the OPBE/TZP geometry optimized labile [4Fe4S] clusters with initial geometries taken from two snapshots of pfl-AE with (Snapshot 1) and without (Snapshot 2) including SAM. Values presented in kJ mol^{-1} and shown relative to the lowest energy state within a geometry, and globally across both snapshot geometries and the model cluster used previously.

Relative within geometry	**Mixed Spin, mS**			
Oxidised	**mS=0**	**mS=1**	**mS=2**	**mS=9**
Model	0.00	58.42	53.89	160.03
Snapshot 1	0.00	43.58	59.74	180.87
Snapshot 2	0.00	38.74	64.19	183.41
Reduced	**mS=1/2**	**mS=3/2**	**mS=5/2**	**mS=17/2**
Model	0.00	64.95	45.77	66.31
Snapshot 1	0.00	41.31	27.48	83.65
Snapshot 2	0.00	49.51	43.67	99.44
Relative globally	**Mixed Spin, mS**			
Oxidised	**mS=0**	**mS=1**	**mS=2**	**mS=9**
Model	0.00	58.42	53.89	160.03
Snapshot 1	32.46	76.04	92.20	213.34
Snapshot 2	31.26	70.00	95.45	214.67
Reduced	**mS=1/2**	**mS=3/2**	**mS=5/2**	**mS=17/2**
Model	0.00	64.95	45.77	66.31
Snapshot 1	51.68	92.99	79.16	135.34
Snapshot 2	32.74	82.24	76.40	132.18

The relative energies of these optimised clusters (Table 3.4) show that the differences in initial geometry have a significant impact on the vacuum optimised clusters, where there is at least a 10 kJ mol^{-1} increase in energy compared to our model cluster, whose initial geometry was extracted from a crystal structure (PDB ID 1ZOY). This suggests that the choice of initial geometry is a critical factor for the study of the energetics of these iron-sulfur clusters but the general preservation of the relative ranking of spin states implies that while the absolute energies are changing, the ground mixed spin state is unlikely to be replaced when exposed to realistic internal enzymatic fields. The only exception is that the ranking of the mS=1 and mS=2 states of the oxidised cluster is reversed between the model and the snapshot clusters. However, as the ground state remains the same, and by more than 38 kJ mol^{-1} in all cases, it is justifiable to continue to only consider this ground state when analysing the broken symmetry spin states for this cluster.

While the relative ranking of mixed spin states was mostly unchanged by using initial geometries that were influenced by different electric fields, the same was not the case for the broken symmetry states (Table 3.5). Here we see that Snapshot 1 not only has a different ground state and broken symmetry state ranking, but it also appears to be more stabilised compared to the other clusters. This is the case even though no charges are present in these calculations; the only influence SAM had was on the initial geometry chosen for the geometry optimisation. Further work should be done to establish if this dramatic change is caused by SAM or is common to any geometry changes induced by the local electric field.

Table 3.5: Relative energies for the broken symmetry states of the ground state labile [4Fe4S] clusters presented in Table 3.4. Values presented in kJ mol^{-1} and shown relative to the lowest energy state within a geometry, and globally across both snapshot geometries and the model cluster used previously. The broken symmetry states 2 and 5 for the Oxidised cluster from Snapshot 2 could not be found.

Relative within geometry	**Broken Symmetry State**					
Oxidised	**1**	**2**	**3**	**4**	**5**	**6**
Model	0.00	25.14	25.26	25.26	24.14	0.00
Snapshot 1	37.84	0.00	36.96	36.96	0.00	37.84
Snapshot 2	0.00	-	32.44	32.44	-	0.00
Reduced	**1**	**2**	**3**	**4**	**5**	**6**
Model	12.50	18.72	18.86	33.53	26.54	0.00
Snapshot 1	42.35	0.00	28.21	33.88	17.97	29.11
Snapshot 2	11.64	20.77	24.94	30.33	36.07	0.00
Relative globally	**Broken Symmetry State**					
Oxidised	**1**	**2**	**3**	**4**	**5**	**6**
Model	5.25	30.39	30.51	30.51	30.39	5.25
Snapshot 1	37.84	0.00	36.96	36.96	0.00	37.84
Snapshot 2	4.19	-	36.63	36.63	-	4.19
Reduced	**1**	**2**	**3**	**4**	**5**	**6**
Model	15.92	22.14	22.28	36.95	29.96	3.42
Snapshot 1	42.35	0.00	28.21	33.88	17.97	29.11
Snapshot 2	14.27	23.40	27.57	32.96	38.70	2.63

These results suggest that while small changes in the initial geometry can have a large impact on the optimised geometry of iron-sulfur clusters and therefore the energy of the broken symmetry spin states, the mixed spin state appears to be preserved and the broken symmetry spin states are within the same order of magnitude, even in the most extreme case of considering the presence (or lack) of point charges representing SAM in a position bound to the labile iron atom. It should be feasible therefore to consider various enzyme-generated force fields and study if there is a correlation between the nature of the force field (for example, the direction and magnitude of the net force) and the structural or sequence similarities of a family of radical SAM enzymes.

3.4 – Conclusion

We systematically investigated the effect of a rotating orientated electric field on the stability and reactivity of biologically relevant iron-sulfur clusters. Applying DFT calculations on the model clusters, we used the VEA – represented as the energy difference between oxidised and reduced state – as an indicator for the redox reactivity of the clusters. In both cases of pure gas phase and calculations in implicit solvent, we could show that a directed electric field induced by two distant point charges significantly influences the stability and reactivity of the clusters. In agreement with molecular frontier orbital theory, the effects are significantly different for different oxidation and broken symmetry states. This leads to the observation that the most stable broken symmetry state changes when re-orientating an electric field around the 2Fe2S cluster, with significant implication for the reactivity of such clusters in heterogeneous environments like enzyme active sites. Different orientations of the electric field had significantly different influences on the VEA and thus reactivity of the clusters with a maximum effect of up to 25.5 kJ mol^{-1}.

The cubane clusters display more complex behaviour with it still being clear that the external electric field has a dramatic effect on the stability and reactivity of these clusters. The introduction of the COSMO environment alters the stability of the reduced cluster, stabilising the favoured broken symmetry 1/2 spin state relative to the other broken symmetry state, but also stabilising the 9/2 state relative to the most favoured state, suggesting that in a protein environment this stabilisation effect is strengthened. The effects on the VEA are similar and in both cases dependent largely on whether the iron atoms are experiencing unique force field environments or similar ones, shown by Figure 3.2c, where the field is rotated about the Fe-Fe axis and the changes in orientation are therefore the same for both iron atoms. Comparing the field with the orbital occupation of the HOMO and LUMO of the clusters reinforces the hypothesis that the spatial location of electron density contributes significantly to the observed effects.

These strong anisotropic effects of the electric fields demonstrate their relevance for iron-sulfur clusters embedded in heterogeneous enzymatic environments. The impact of the field might explain how enzymes influence the redox reactivity of such clusters in a significant way. This systematic study reveals how these effects might be rationally explored and used as a powerful tool for enzyme engineering where the external electric field would be generated by surrounding residues, which can be mutated for direct control of the properties of the iron-sulfur clusters.

Whilst the model systems presented cannot reflect the complexity of the anisotropic charge distributions inside enzymatic active sites, the presented field strengths and strong directionality of the electric fields are comparable.[112] Thus the model systems offer a good reference on how changing the directionality of electric fields inside the enzymatic active sites can influence catalysis. In the next step this information can be used for rational mutation studies with possible applications that might alter catalytic rates, substrate scope and cluster stability (for example in oxygen sensitive enzymes). Further work should focus on the nature and magnitude of these possible applications and how generalizable this approach is to other similar molecular groups in enzymes.

Chapter 4 – Bioinformatics of BioB

4.1 – Motivation

With a method for analysing the effect of an oriented electric field on iron-sulfur clusters modelled on those present in biotin synthases the next logical step is to look at examples of these enzymes and particularly the environment around the iron-sulfur clusters. Here we present a variety of options to tackle this problem, from specifically mapping important residues around the iron-sulfur clusters to sequence and structure similarity analysis. This type of analysis can further elucidate information about the most important residues and how influential they are in altering the overall electric field of an enzyme and indeed if this effect is a primary driver for the evolution of the sequence and structure of biotin synthases or if it is just one of a plethora of similarly important features.

The goal of this chapter is to identify methods which may be useful in finding patterns in sequence and/or structure that may shed light on the hypothesis that electric fields are a major driver of iron-sulfur cluster containing enzyme chemistry and those patterns that can inform rational engineering of biotin synthases by identifying conserved, co-dependent, or highly variable regions or residues.

4.2 – Identifying Useful Methods

An example of a method to identify key residues of interest with respect to iron-sulfur cluster containing enzymes *a priori* for inclusion in computational calculations or other methods is presented by Harris and Szilagyi.[264] Here the goal was to identify significant interactions to include in a quantum chemical model (for example the QM region of a QM/MM calculation) without the need to run prerequisite calculations to find these interactions each time. To do this, they performed a systematic mapping of the common interactions between iron-sulfur clusters and their surrounding residues and other features of the environment using DFT calculations to create an effective "importance" ranking for each interaction and thus create general selection rules for capturing significant interactions in these environments. The DFT calculations consisted of the 6-311+G(d) basis set[288-290], a gradient-corrected, hybrid functional with 5% Hartree-Fock exchange, 95% Becke88 exchange[175] and 100% Perdew86 correlation.[178] They also included some calculations in a PCM with a dielectric constant of 4.0 and antiferromagnetic coupling was treated using broken-symmetry DFT, similar to Chapter 3.

This system of “importance” for inclusion in quantum chemical models should also provide insight into the “importance” of interactions with iron-sulfur clusters in general, as the more essential an interaction’s inclusion in more accurate computational methods, the more likely it is to have a large impact on the system when altered (by mutating a residue for example). At the very least this gives us an extremely quick and inexpensive first pass for finding residues of interest with respect to mutation for engineering.

The resulting set of *a priori* rules can be simplified for our purposes to include the whole residue rather than the functional group specific definition used in the original paper. We chose to define coordination spheres based on this method as follows:

- 1st Coordination Sphere: Residues covalently or otherwise coordinated directly to the iron-sulfur cluster as a ligand.
- 2nd Coordination Sphere: Residues (including water molecules) covalently bound to the 1st sphere or involved in hydrogen-bonding with the 1st sphere.
- 3rd Coordination Sphere: Residues (including water molecules) involved in hydrogen-bonding to other residues that are themselves covalently bound to the iron-sulfur cluster ligands. Residues (including water molecules) less than 8 Å from an Fe atom in the cluster.

This simplification should capture at least the same residues as the complete method presented in the paper but as we are not using this to select residues for inclusion in quantum chemical calculations, we do not need to limit selection on the basis of computational cost. As such we will usually capture more residues using this method but it is simpler to understand and implement. This method should also be reasonably straightforward to automate and include in enzyme engineering or analysis pipelines as it only needs the nature, volume, and location of each residue rather than the orientation of specific functional groups.

With a method to find potential residues of interest in iron-sulfur cluster containing enzymes it would also be useful to know how nature has treated these residues, namely are they highly conserved, variable or co-dependent? Are there specific patterns we can find that might inform our engineering choices? A clear initial step to take is to collate the set of enzymes we are interested in, in this case biotin synthase-like enzymes, and map out their sequence and structural similarity. We will approach the former problem first.

The Structure-Function Linkage Database (SFLD)[291] is an archive consisting of "a hierarchical classification of enzymes that relates specific sequence-structure features to specific chemical capabilities." This is a powerful tool as it has already identified the structural characteristics of enzymes based on their sequences and we can therefore extract a set of sequences that should have similar structures (and therefore functions) to analyse further. Before becoming an archive, the SFLD had a more interactive set of options for downloading sequence similarity maps that included choosing different parameters affecting the behaviour of a network of related sequences.

To form a network, nodes are initially formed by collecting all the sequences that share 50% pairwise sequence identity with a seed sequence. Edges between these nodes are created with a distance based on the average similarity of the members of these nodes using the Basic Local Alignment Search Tool[292] (BLAST). The BLAST E-value is a measure of significance for the tool; given a random database, it is the number of similarity "hits" expected to be found by chance. The lower this number the more rigorous the algorithm is in deciding what constitutes similarity, giving fewer but more similar results. The SFLD uses a modified parameter,

$$-log_{10}E \geq pE, \qquad (4.1)$$

where pE is the user input for the accuracy parameter. The higher this value is, the stricter BLAST will be in its measure of similarity. The SFLD archive has a pE value of 20 which results in highly interconnected nodes that provide very little information. Previously it was possible to select your own pE value and we chose 85 to create our mapping as the requirements for similarity needed to be stricter due to a generally higher similarity across the data set. Choosing a stricter cut off value allows us to view relationships between nodes at a finer resolution even within a group of very similar nodes. Cytoscape[293] is software that facilitates the visualisation of highly interconnected data sets such as this, allowing for a wide variety of display formats and custom schemes to highlight various properties of interest of the network. This is an extremely powerful analytical tool and the examples presented later form only a small part of Cytoscape's toolbox.

With this method for analysing sequence similarity in hand we then move to incorporating structural similarity into our pipeline, where we can ask such questions as: are the conserved structural features always formed by the same conserved sequences or is there a variety of sequence options to form the same structural motif? If we were to construct a network of structural similarity analogous to the sequence similarity network, would they map onto each other exactly, would the nodes collapse or expand, and are there any common patterns or does it appear random? Answering these questions in full would be an extreme undertaking as predicting protein structures from amino acid sequences is not a solved problem. However, we do have tools to tackle this issue at varying degrees of accuracy, with a similar situation holding for structural similarity analysis. We chose to use a small subset of exemplar sequences from a variety of nodes in order to both establish an initial picture of the relationship between sequence and structure similarity, and if there are any obvious patterns we can exploit to further streamline this process.

Before looking at structural similarity we must first predict the structures of these proteins from their amino acid sequences. The primary method we chose to approach this problem was I-TASSER[294, 295] (Iterative Threading Assembly Refinement) which is a template-based structure prediction and structure-based function annotation software. I-TASSER only requires an amino acid sequence as an input and will establish the most likely resulting structural folds based on databases of known sequence-structure pairing and includes the ability to construct novel structures formed from fragments of known proteins. It has had historic success in the CASP (Critical Assessment of Techniques for Protein Structure Prediction) experiments with first place overall rankings in CASP7-14 and can be considered a reliable tool for protein structure prediction, especially for proteins whose adjacent sequences have confirmed structures listed in databases.

Another method for protein structure prediction is trRosetta,[296-298] which, while capable of including template-based homology modelling like I-TASSER to accelerate the process, has a core functionality focused on *de novo* structure prediction powered by a deep learning neural network and direct energy minimisation. We did not use any of the structures produced by trRosetta in further analysis but we found that it was an excellent tool for confirming the validity of the predictions made by I-TASSER and note that we did not see any significant differences in accuracy for the set of proteins we applied these software to.

After predicting the structures of these proteins we need to perform structural similarity analysis on them, analogous to the sequence similarity analysis from the SFLD. Initially this was done using DeepAlign[299, 300] which is a promising method that includes evolutionary similarities and hydrogen bonding localisation on top of the standard method of comparing the spatial proximity of equivalent residues. The output also breaks down these measures to gather a finer understanding of not just how similar the structures are, but also which measures performed the best and possibly infer the reasons for similarity, such as being either close evolutionary neighbours or an example of convergent evolution. While this is an extraordinarily useful tool for these reasons, we also used another geometry-based method that has more automated visual results analysis for the purpose of presenting it here.

The Dali server[301] is a web-based protein structure comparison tool with a variety of visualisation methods that can elucidate important features of a set of compared structures that may otherwise be lost using a simple scoring system. Dali primarily utilises distance matrix comparison[302] to assess similarity and does not optimise a rigid-body superimposition of structures, although this does form part of the output. This tool was our primary source of results, with DeepAlign acting as a check to ensure there were not extreme divergences in results caused by the choice of software.

While most bioinformatics methodologies will deal with very large sets of data, for smaller sets or single enzymes of interest a manual method for visualising potentially important residues around features of interest is useful to have available. We took the superimposed structures of each protein and the 'most similar' protein overall and manually identified and categorised regions of divergence to establish how well this translates into the scores from the above-mentioned software and if there is a pattern relating to sequence similarity based on not just overall structural similarity but also the nature of the divergences.

4.3 – Application of Methods to a Set of Biotin Synthase-Like Sequences

Applying the ruleset constructed by Harris and Szilagyi[264] to the only available crystal structure of BioB on the Protein Data Bank (PDB ID: 1R30)[4] we generated a map of the approximate coordination spheres around the iron-sulfur clusters in this enzyme. The resolution of this crystal structure is quite poor, at 3.4 Å, so the resulting map will not be a perfect representation of the set of residues that are the most influential with respect to the iron-sulfur clusters. However, this is still a useful methodology to follow. It provides a first approximation for the most critical residues surrounding the iron-sulfur clusters in BioB to consider for re-engineering. Additionally, it can demonstrate the generalisability of the method for any iron-sulfur cluster containing enzyme. Identifying the critical residues in one structure allows us to compare with other sequences and predicted structures to highlight highly conserved, co-dependent, or other patterns of residues that elucidate important information to consider when re-engineering an enzyme.

The highly conserved rSAM CX_3CX_2C motif binding the cubane iron-sulfur cluster is clearly seen, located on residues 53, 57 and 60 in this enzyme, also bound to SAM as this model system includes bound versions of SAM and dethiobiotin (DTB). (Figure 4.1) Similarly, we see three cysteines and one arginine binding the [2Fe2S] cluster, but we see these four residues are spread out along a much longer section of the sequence than for the [4Fe4S] cluster. (Figure 4.2) The implied motif based on this enzyme would be $CX_{30}CX_{59}CX_{71}R$. However, it may be reasonable to allow all four of these residues to be either cysteine or arginine, and that the intervening residue chains could differ in length when searching for this motif in other enzymes.

While not surprising, it should be noted that a large proportion of the residues identified using this method also form part of a key structural characteristic of rSAM enzymes, the TIM barrel, specifically the inner ‘wall’ of this region that is surrounded by eight alpha helices. (Figures 4.3-5) This has significant implications for the complexity of re-engineering the enzyme as mutations that may be beneficial for the reactivity could also be a hinderance to substrate/product entry/exit, or even initial folding, highlighting the critical importance of combining methods of re-engineering that are focused on the reaction itself with structural prediction methods to ensure the conservation of key structural characteristics.

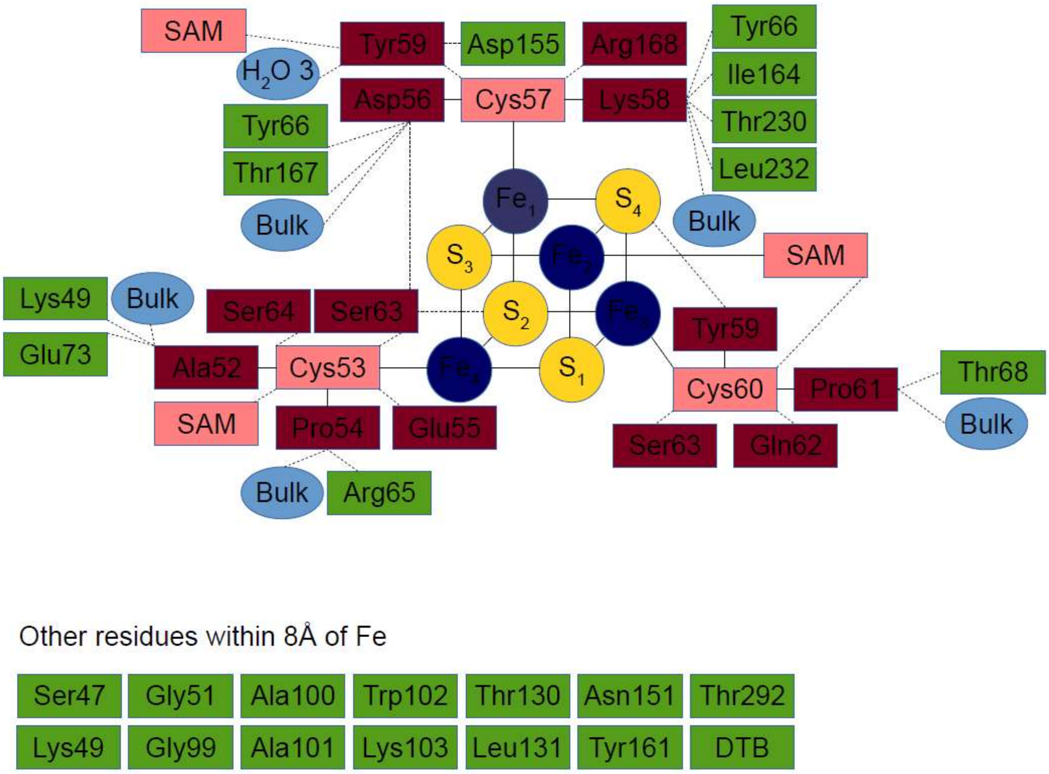


Figure 4.1: Coordination spheres around the [4Fe4S] iron-sulfur cluster in the crystal structure of BioB 1R30. Pink = 1st Coordination sphere, Magenta = 2nd Coordination sphere, Green = 3rd Coordination sphere.

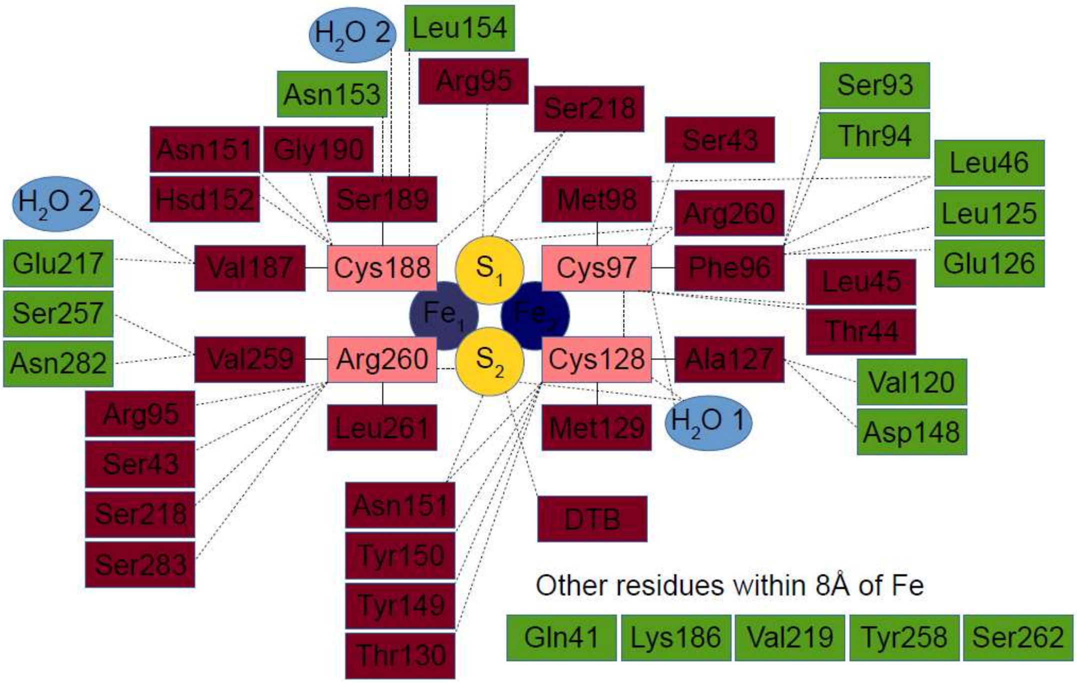


Figure 4.2: Coordination spheres around the [2Fe2S] iron-sulfur cluster in the crystal structure of BioB 1R30. Pink = 1st Coordination sphere, Magenta = 2nd Coordination sphere, Green = 3rd Coordination sphere.

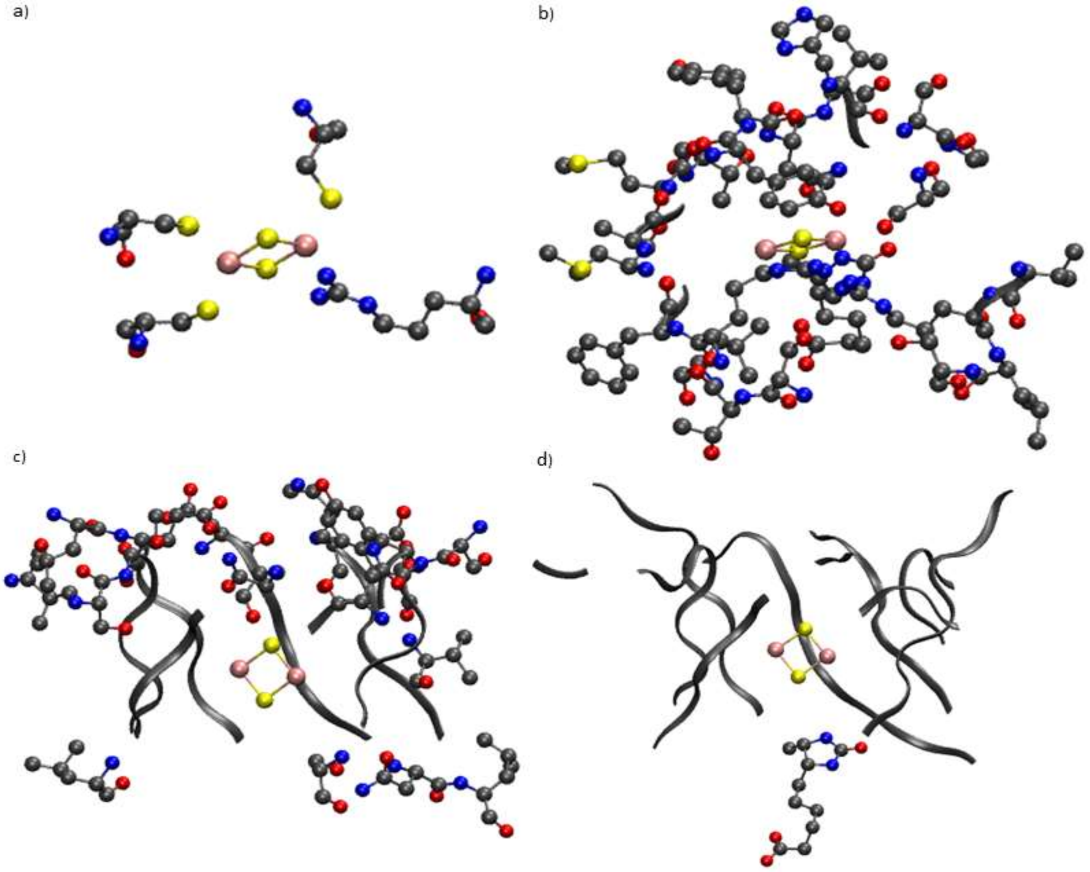


Figure 4.3: 1st (a), 2nd (b) and 3rd (c) coordination spheres around the [2Fe2S] iron-sulfur cluster. Previous coordination spheres are displayed as ribbons in panels b) and c). All three spheres are displayed as ribbons in panel d) along with the [2Fe2S] cluster and DTB.

While this method can be used to identify residues of importance in a specific enzyme it requires a complete three-dimensional structure either from a crystal structure or a structure prediction method, the former being the preferred option. If we want to look at the family of biotin synthases, we first need to apply methods to help us identify which specific proteins we want to study further. Software such as BLAST can be utilised to form parsable networks from large data sets such as protein families and can be used to find common characteristics, co-dependencies, possible evolutionary neighbours, and convergent evolutionary traits.

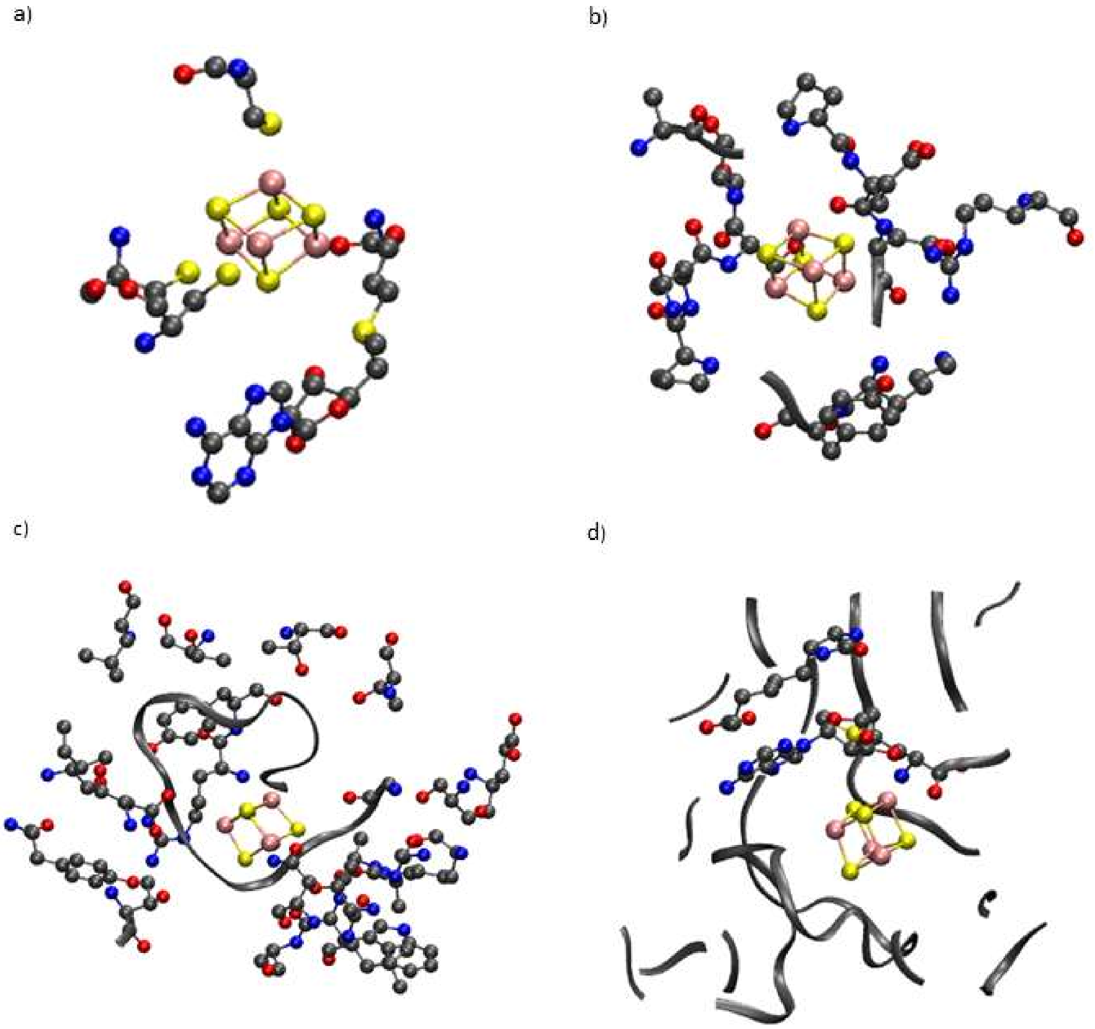


Figure 4.4: 1st (a), 2nd (b) and 3rd (c) coordination spheres around the [4Fe4S] iron-sulfur cluster. Previous coordination spheres are displayed as ribbons in panels b) and c). All three spheres are displayed as ribbons in panel d) along with the [4Fe4S] cluster, SAM and DTB.

We constructed such a sequence similarity network graph using data from the (now archived) Structure-Function Linkage Database.[291, 303] The version of this data that can be accessed in the archived version is for a representative network of biotin synthase-like enzymes with a node similarity threshold of 50 (percentage pairwise sequence identity) and a BLAST E-value cutoff of 20. This creates highly interconnected nodes that provide very little information due to the very high sequence similarity in this subgroup; instead we used data with an E-value of 85 acquired prior to archival which allows us to create a much more defined network graph due to the strict similarity requirements to connect nodes.

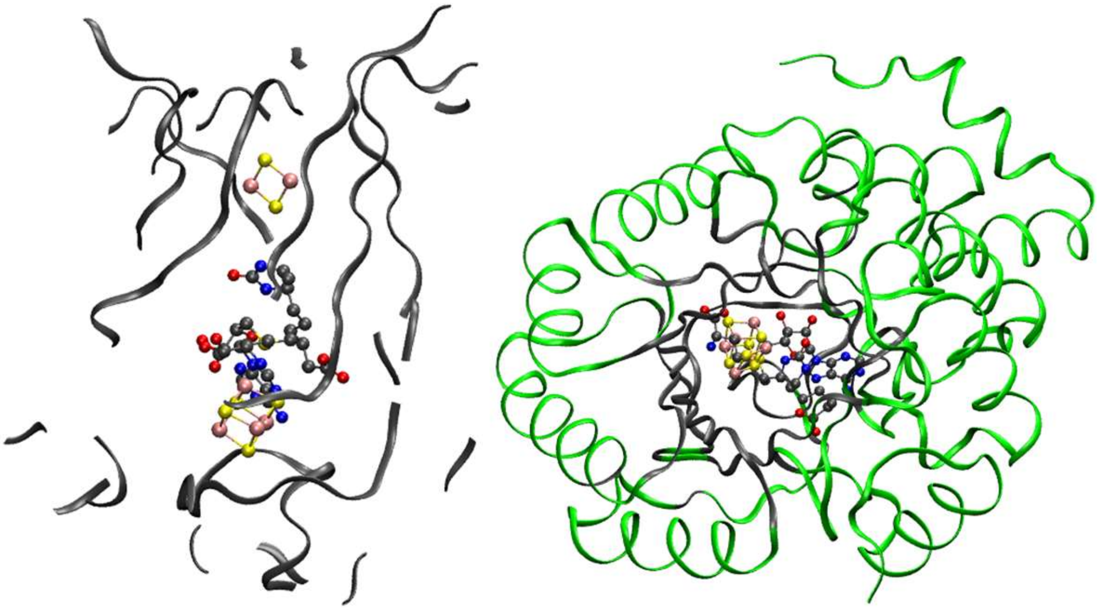


Figure 4.5: The coordination spheres around both iron-sulfur clusters are displayed as grey ribbons, with the clusters themselves, SAM and DTB highlighted. The remainder of the protein is displayed as green ribbons (right).

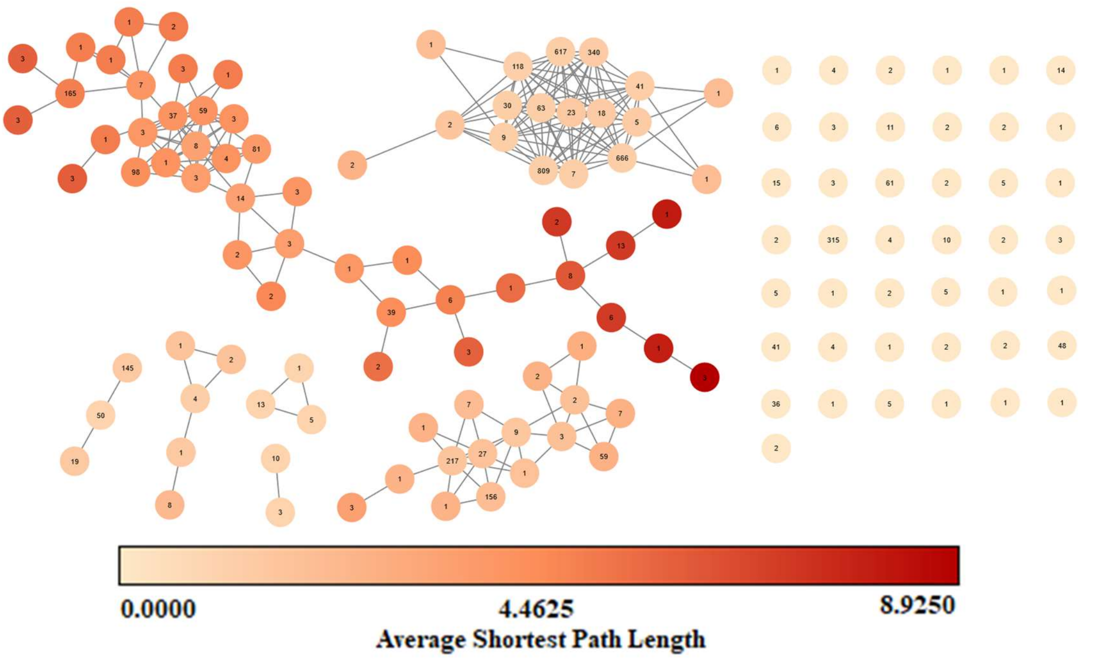


Figure 4.6: Cytoscape sequence similarity map indicating the average shortest path length for each node. A shortest path between two nodes is defined as the minimum number of edges needed to traverse the network from one node to the other. Each node contains a number indicating node size.

We constructed a small number of examples of colour coded networks to demonstrate the usefulness of analysing this data as a network in Cytoscape. Before any shading is applied, we can already note several features of the network such as the large number of 'solitary' nodes on the right hand side of the network. Some of these only contain one sequence meaning they may have unique features reducing their pairwise sequence identity to the rest of the set, or they may have been erroneously assigned to the SFLD set. Others are large sets of sequences with at least 50% pairwise sequence identity but low similarity to other nodes (although a caveat is that we explicitly increased the threshold for creating edges between nodes, so while they appear to be unrelated they are actually strongly related to the other nodes, just not as much as those that *are* connected), which may indicate that they all share a feature no other nodes have, or have, for example, an additional region dedicated to a specific secondary function. In the top middle section of the network is a highly interconnected set of well populated nodes indicating extremely similar sequences and high node similarity due to the short edge distances. Finally, we see several other chains of networks with a varying degree of interconnectedness. It is possible that these chains indicate a path of evolution as each edge would suggest an incremental change but the difference between sequences at opposite ends of this chain might appear to be very divergent.

Figure 4.6 shows the network colour coded by average shortest path length, defined as the average of the shortest path to each connected node on a given node’s subgraph. This gives a qualitative indication of how varied a subgraph is. Highly interconnected subgraphs will have a low average shortest path length while those displaying a possible evolutionary path, such as the one discussed at the end of the last paragraph, will have high values at the ends of the chain, and lower in the middle regions (as intermediate ‘species’). This can clearly be seen with the highly interconnected subgraph in the top middle of the network having a very light colour for all its nodes while the long chain subgraph below it has darker colouration at the extreme ends of the chain.

Node degree is a measure of how many edges a node has. In our case, as each unique node pair can only be connected by at most one edge, this is also the number of other nodes a given node connects to. A node with a high degree will likely contain a large amount of the ‘core’ sequence of its linked nodes and might be some form of common ancestor for them. The multiple nodes of high degree in Figure 4.7 could be produced by variation in a specific feature, such as a secondary function or structural characteristic that is changing while the rest of the sequence is conserved. An enzyme performing the same reaction in many different organisms that have slightly different binding sites and therefore have some surface level structural mutations, while the rest of the enzyme is identical, would be an example of this.

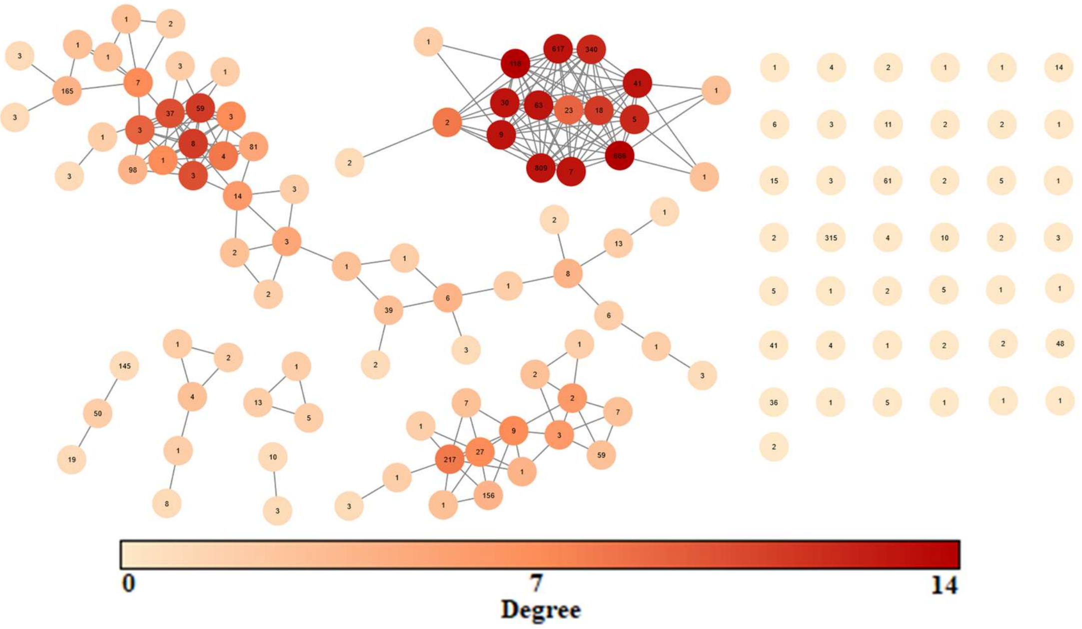


Figure 4.7: Cytoscape sequence similarity map indicating the degree for each node. Degree is defined as the number of edges a node has. Each pair of nodes can only have one edge connecting them. Each node contains a number indicating node size.

The last example we show is that of node size, simply the number of sequences contained within a given node. Using a linear gradient results in Figure 4.8, where it simply highlights the nodes with extremely high population. This obfuscates the finer detail in node size of the rest of the graph as the majority of nodes are much smaller than these large nodes. Figure 4.9 is an attempt to correct for this where 50% of the colour gradient occurs between node sizes 1 and 5, leaving the upper half of the colour gradient for 6 to 809. This provides a much clearer picture of the distribution of sequence population in the nodes and particularly highlights how many of the ‘solitary’ nodes on the right hand side of the network have a significant population, and the number of single sequence nodes contained in the subgraphs, in some cases even connecting two halves of the subgraph that would otherwise be individual subgraphs.

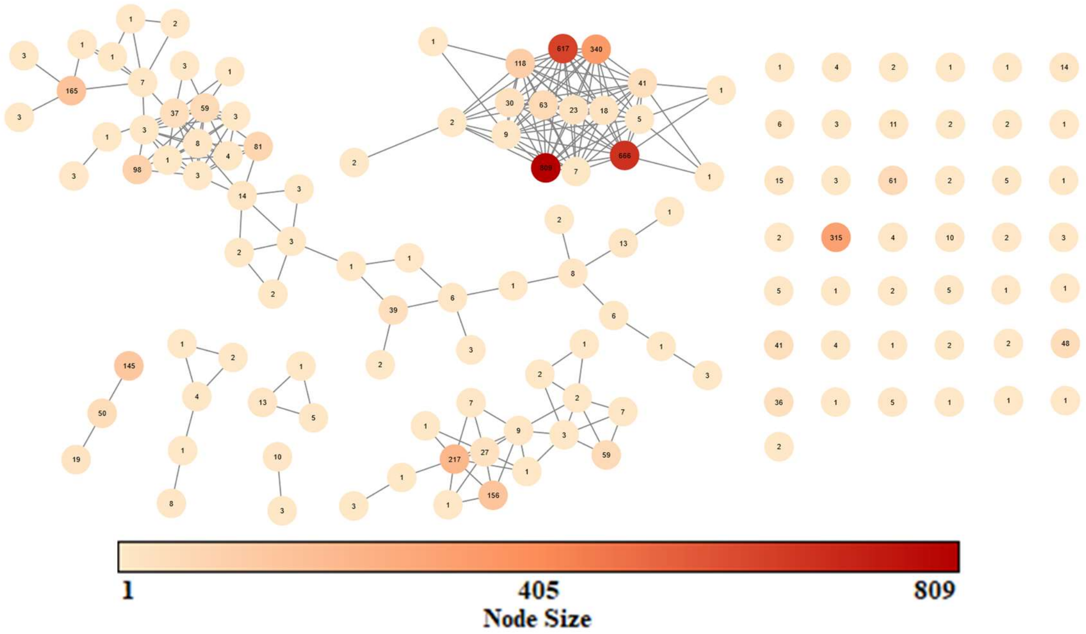


Figure 4.8: Cytoscape sequence similarity map indicating the node size for each node with the midpoint of the colour gradient at 405 sequences. Each node contains a number indicating node size.

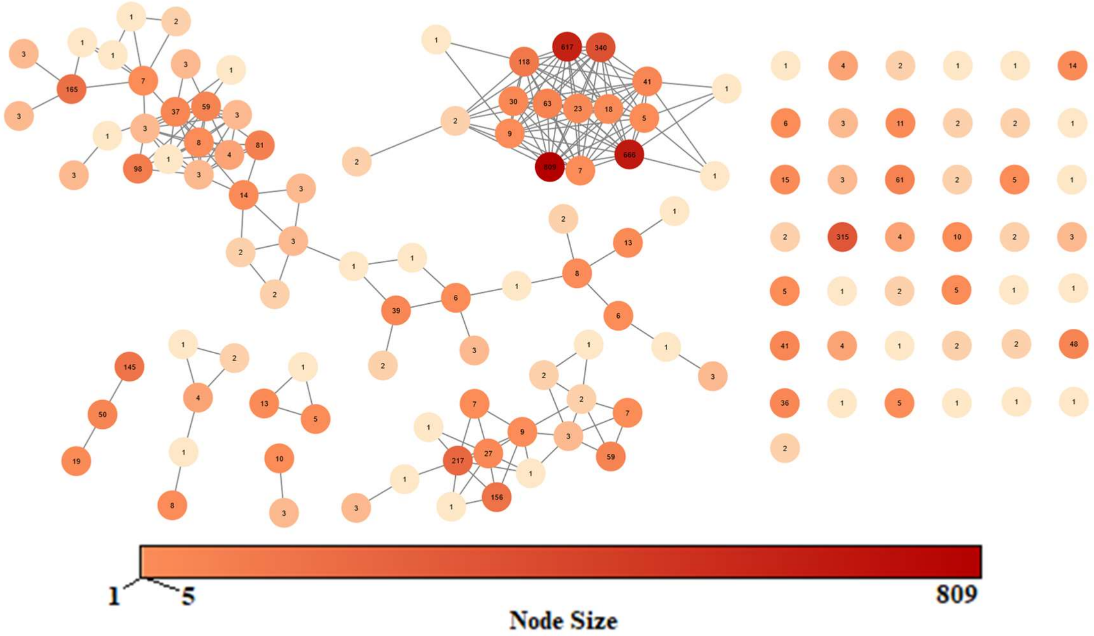


Figure 4.9: Cytoscape sequence similarity map indicating the node size for each node with the midpoint of the colour gradient at 5 sequences. Each node contains a number indicating node size.

We randomly selected a subset of sequences (no more than two per node) from the sequence similarity network and used I-TASSER to generate predicted folded protein structures. These structures were also corroborated by using trRosetta to ensure there were no major discrepancies caused by the choice of structure prediction software. We then used Dali Server to perform an all-to-all sequence and structure similarity comparison for these structures along with a similar pairwise comparison using DeepAlign.

There are several individual metrics that DeepAlign reports that can highlight specific types of similarity but the overall score for general similarity collects all these metrics into a single "DeepScore" which ranks the pairwise comparisons between all the structures. We then calculated a mean DeepScore for each structure to obtain a measure of the mean similarity to all other input structures (Figure 4.10).

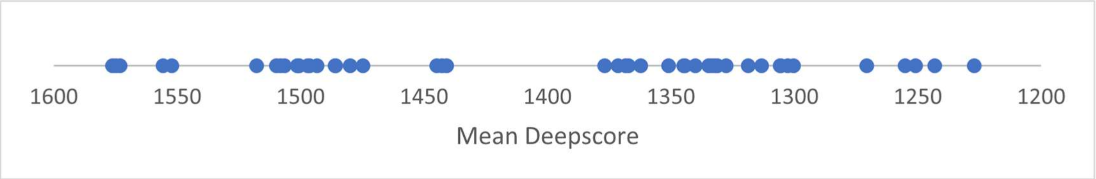


Figure 4.10: Mean DeepScore of each structure compared pairwise with each other structure. A higher score indicates greater average similarity.

As we took two structures from some of the nodes in the sequence similarity network the initial assumption was that the paired clusters shown in Figure 4.10 would be those drawn from the same node, as they have more similar sequences, but this is in fact not the case for some of these pairings. For example, the highest scoring structure with a mean DeepScore of 1628 shares its sequence similarity node with a structure scoring only 1369 whereas the structure ranking second at 1612 does share a sequence similarity node with a closely ranked structure with a DeepScore of 1607. This already suggests, with only a small sample, that a sequence similarity network will not perfectly map onto an analogous structure similarity map.

The next question that arises is if the two structures with high mean DeepScores do not come from the same sequence similarity node, how closely related are they? These two examples are from adjacent nodes highlighted in Figure 4.11.

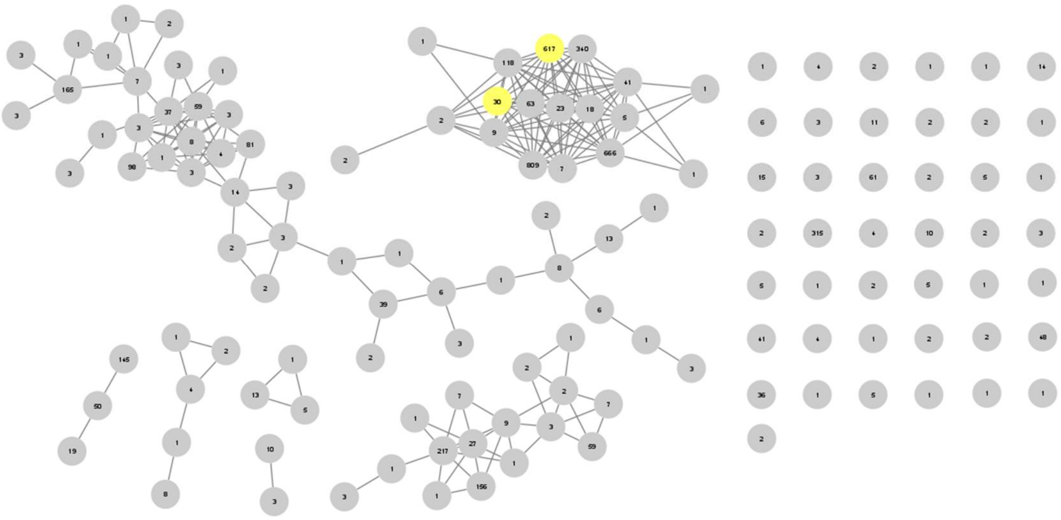


Figure 4.11: Sequence similarity network with two nodes containing high mean DeepScore structures highlighted in yellow. Each node contains a number indicating node size.

How does this look if we generalise this to all structures with mean DeepScore greater than 1400 for example? This seems like a reasonable cutoff point as there is a clear gap shown in Figure 4.10 in this region. Figure 4.12 highlights all the nodes that fit this parameter.

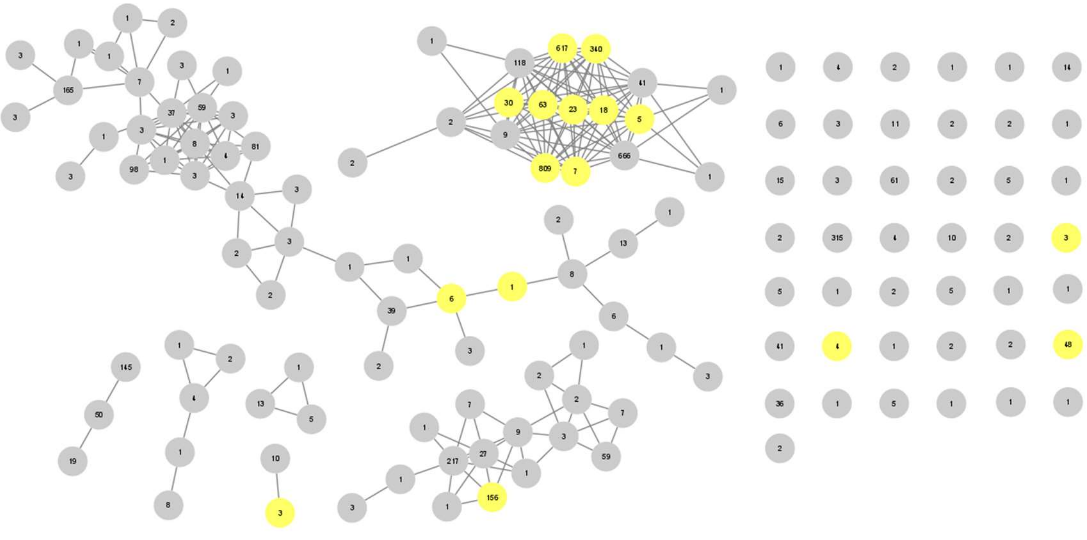


Figure 4.12: Sequence similarity network with nodes containing structures with mean DeepScore >1400 highlighted in yellow. Each node contains a number indicating node size.

We can immediately see that most of these nodes are collected in the top middle subgraph, but as this subgraph contains a large portion of all the sequences in the network, including the three most populous nodes, (Figures 4.8 and 4.9) this seems statistically likely even if we chose sequences to highlight at random. Indeed, several of the nodes contain only one sequence, and three of them have zero edges connecting them to other nodes. This suggests that, at least in this very strict case, sequence similarity networks cannot be used as the only predictor for structural similarity.

We performed an all-to-all structural similarity analysis using Dali and found comparable results to those from DeepAlign. Both the output dendrogram and correspondence analysis provide useful information for our purposes. The dendrogram (Figure 4.14) displays similar structures close to each other and less similar structures further away, much like a phylogenetic tree, and upon closer inspection we can see pairings of closely similar structures that sometimes come from the same node such as structure 47 and 48 both coming from node 1966284 while other pairings such as 26 and 36 are from separate nodes while structure 27, from the same node as 26, pairs with structure 21, from another node. In fact, structure 36 comes from an unconnected node containing only this sequence and is highly similar to structure 26 whose sequence is contained in the highly connected subgraph in the sequence similarity network. (Figure 4.13)

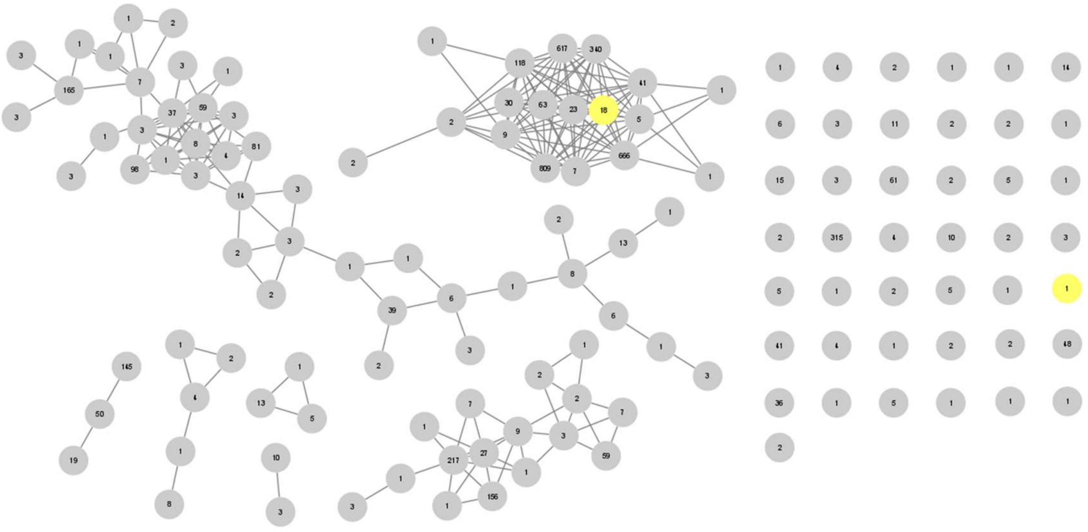


Figure 4.13: Sequence similarity map with highlighted nodes containing sequences corresponding to structure 26 (top) and structure 36 (right). Each node contains a number indicating node size.

While we must accept we have only taken a small subset of structures to perform this analysis on, we have once again seen that highly similar structures will not necessarily come from highly similar sequences. They may even appear, as is the case presented in Figure 4.13, to be entirely unrelated.

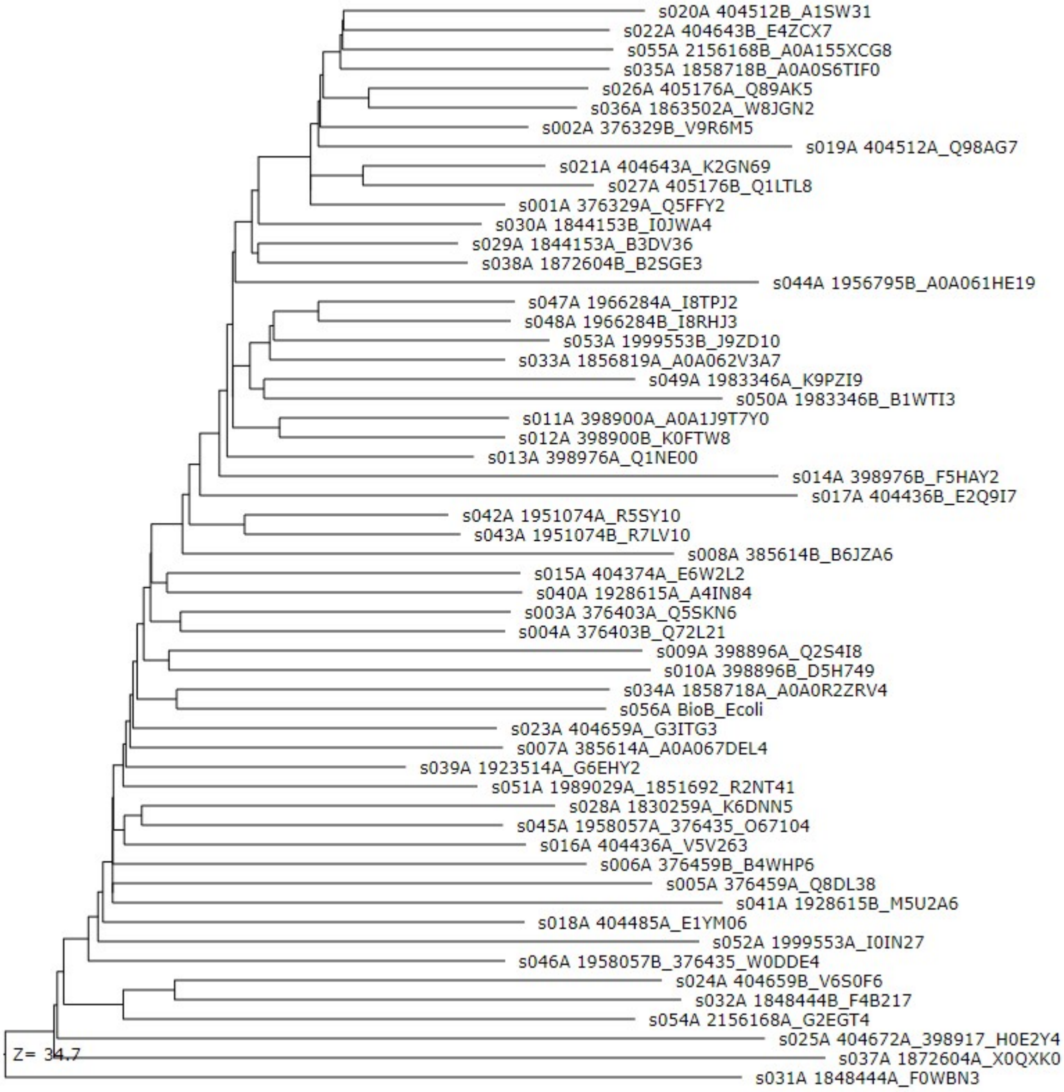


Figure 4.14: Structural similarity dendrogram for a selected subset of structures of biotin synthase-like sequences. Highly similar structures are located near each other analogous to a phylogenetic tree. The nodes are labelled with a structure identifier, the labels of each sequence similarity node the corresponding sequence is located in, and the Uniprot ID for this sequence.

The correspondence analysis shows a similar relationship in a two-dimensional way where similar structures are positioned closer to each other. (Figure 4.15) Figure 4.16 shows the locations of the circled structures' corresponding sequence in the sequence similarity network, highlighted in the same colour. In particular, the nodes highlighted in yellow show a promising result, many of them are connected in the sequence similarity network and are located far from those highlighted in red and green, as would be implied by their distances in the correspondence analysis. (Figure 4.15)

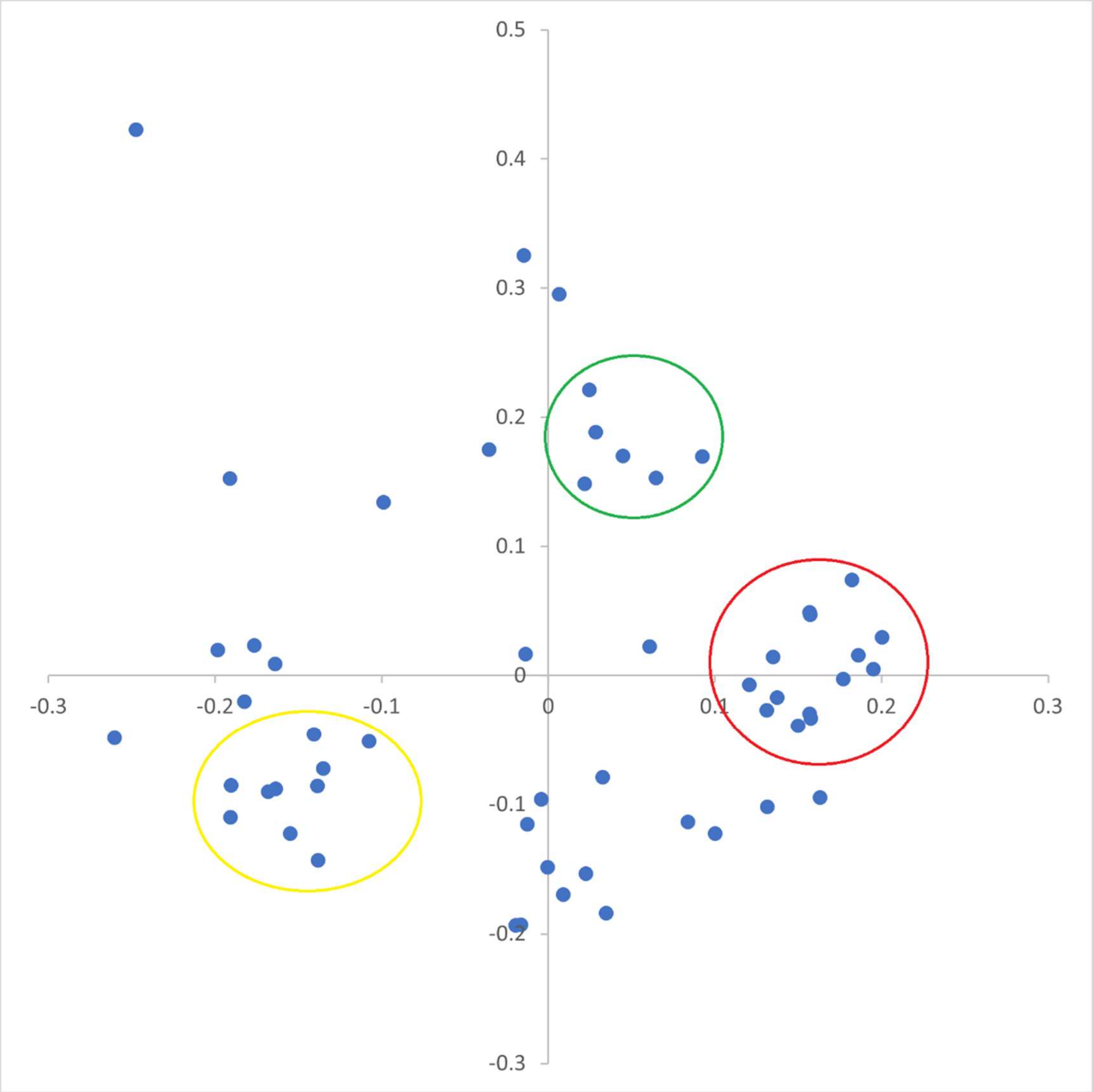


Figure 4.15: Dali structural similarity correspondence analysis of the 56 selected structures

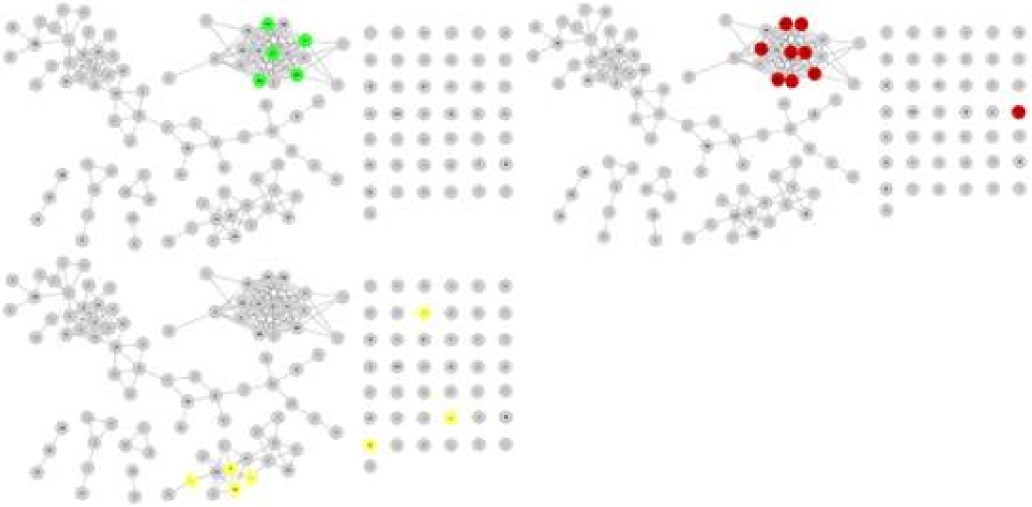


Figure 4.16: Sequence similarity maps with nodes highlighted with the same colours as those structures circled in Figure 4.15. Each node contains a number indicating node size.

When visualising the structures in VMD to assess the effectiveness of the structure comparisons we noted that the differences between the structures could be qualitatively categorised based on the nature of the changes and the severity of those changes. To simplify this task, we compared structure 20 (Uniprot ID A1SW31) to each of the other structures, with particular attention paid to the regions close to the iron-sulfur cluster binding sites. All structures had some global changes as expected for flexible molecules so any small translations in the precise location of a feature was ignored; however changes that were more significant such as a helix being out of phase or features appearing at different points in the chain or being different lengths were noted as important.

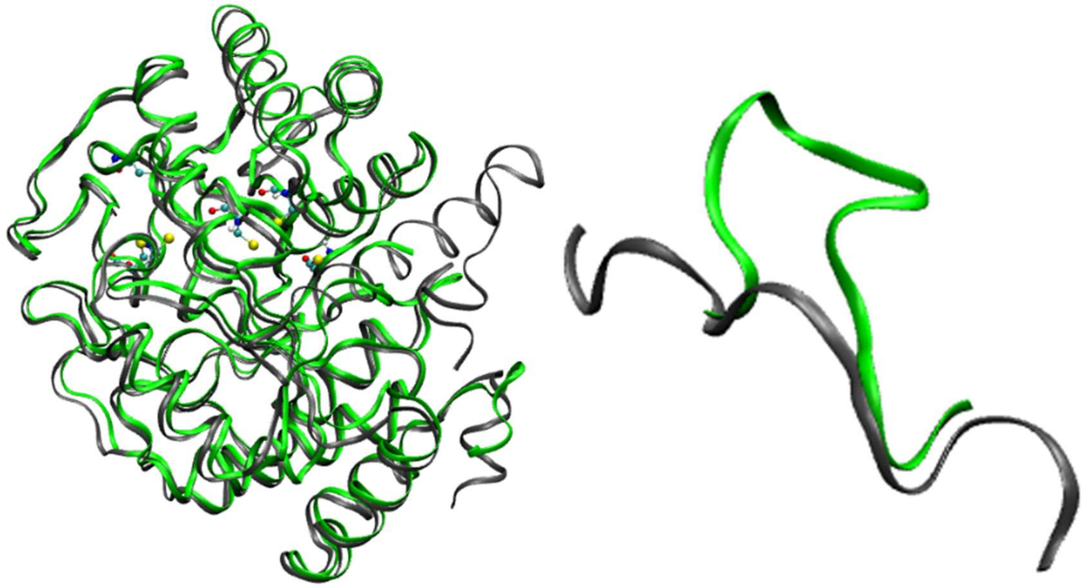


Figure 4.17: Example of structures characterised as having major differences in the tail sections only (left) and a minor inserted loop region that does not affect the surrounding chain (right).

The most populous categories were those where the differences were almost exclusively confined to the tail regions of the enzymes (Figure 4.17), to be expected as these regions are highly flexible. These fell into two broad categories, those which had similar tail sections generally, but their flexibility was likely the cause of differences, and those which had significantly different, but themselves well-defined, (for example, a complete helix) tail sections but maintained core alignment. For these two categories there also exists some instances where a small loop region is inserted that only affects the surface of the protein and does not cause any variation in the rest of the chain, and other cases where the inserted loop does cause some flexibility in the chain, but the distortion is no more significant than minor translations of the chain.

Those structures with greater variability within the core of the enzyme but still followed the same fold, with the same secondary structures at roughly the same points in the chain were categorised based on whether the variability directly impacted the iron-sulfur cluster binding sites or not (Figure 4.18).

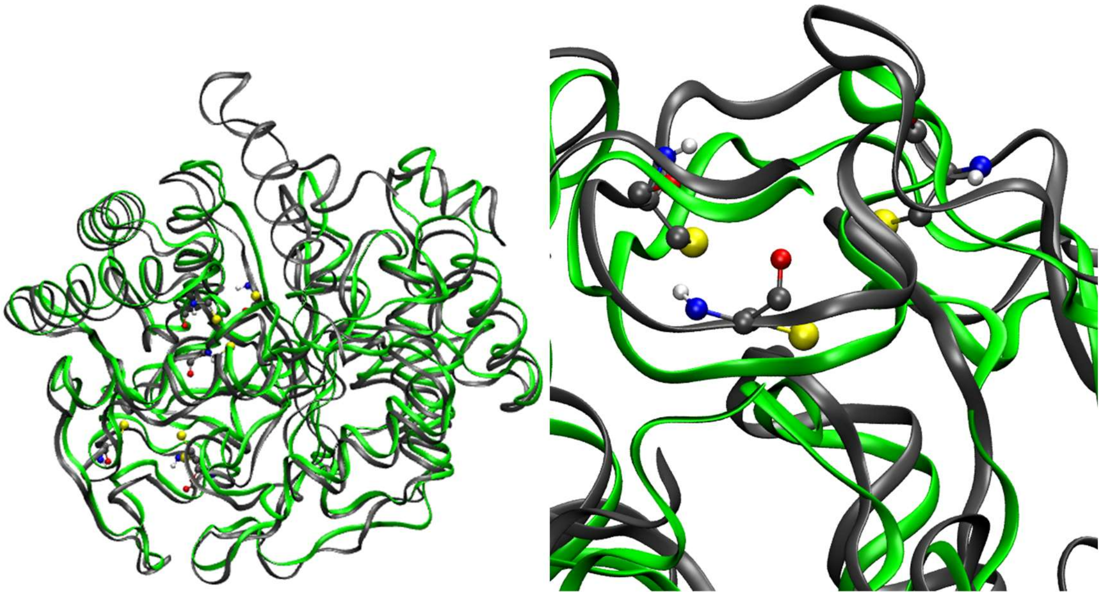


Figure 4.18: Example of a structure with large variability (left) and variation that would be considered impactful on the iron-sulfur cluster binding sites (right).

Table 4.1: Qualitative categorisation of structures based on their differences from structure 20.

Category	Sequence/Structure Identifier
Flexibility in tail section only	1-12,21-22,26-27,29,30,34-37,54-55 (13 with loop 213-219, 50 with loop 156-162)
Additional, well-defined structures in tail region	14,17,19,33,38,41-43,44-49,51,53 (39 with loop 241-248)
Same fold with variability in secondary structures not at Fe-S binding sites	24,31-32,40
Same fold with variability in secondary structures including at Fe-S binding sites	23,25,28,52,56

The final category consists of the enzymes that have such great variability that it appears they have a different fold to that of structure 20, where there are entire regions that do not map onto structure 20 at all. There were no structures in this category for the final analysis, but it should be retained as a possible categorisation.

To identify if the complexity of the all-to-all comparison was influencing the structural mapping we repeated the comparisons of the structures in all but the first two categories in Table 4.1 on a 1-to-1 basis with structure 20. This resulted in either no changes or a small global translation but nothing significant enough to suggest the structure should be in a different category.

4.4 – Conclusion

In this chapter we present a variety of methods for capturing the relationships between sequence and structure similarity and applied them to the dataset of biotin synthase-like enzymes acquired from the SFLD. By examining a limited range of BioB structures, predicted using I-TASSER and verified with trRosetta, we identify that structural similarity is not solely predicted by sequence similarity. This observation is important as sequence similarity is often used to infer structural and functional relations. The main limitation, however, is that this analysis has relied on prediction methods, and experimental structure data would be needed to verify these predictions to solidify these observations.

Closer examination of the structures has identified subclasses of structural change that may have implications for function, narrowing down the scope for initial engineering modifications. If a combination of the prediction, network, and alignment methods were included in an automated system we can imagine a system for generating at the very least patterns in sequence-structure relationship, conserved sequence and/or structural features. Focussing on those important residues for iron-sulfur cluster chemistry, suggested mutations for improving or otherwise editing enzymes that retain functional and structural integrity by uncovering these new structural elements may act as a starting point for further studies, limited by the fact that these are annotated, rather than characterised structures.

Chapter 5 – Conclusion

In this thesis we presented methods for understanding, manipulating and eventually automating the structural and functional analysis of iron-sulfur cluster containing enzymes, with a particular focus on the engineering of the rSAM enzyme biotin synthase. While automation may not be achieved for some time, it is important to lay the groundwork for such an ambitious goal, and with a combination of quantum chemistry, molecular dynamics and modelling, and bioinformatics this does appear to be achievable.

A large portion of this work was dedicated to mapping the effect of an oriented electric field on iron-sulfur cluster chemistry and working towards understanding how these effects manifest and the possible ways in which we might manipulate this. We performed a systematic, high throughput study on model iron-sulfur cluster systems within an oriented electric field and presented the resulting effects on the relative energies of spin states, orbitals, vertical electron affinities and spin-coupling constants. If this could be generalised for more iron-sulfur clusters and with more complex electrostatic environments there is a clear indication that this could be automated to tune the environments to optimise reactivity when paired with the structural and sequential similarity methods displayed later in the thesis. Electric field effects are not solely limited to enzymes, and the methods and approaches developed here could equally be applied to other molecular synthetic systems beyond biochemical catalysis.[112]

Subsequent analysis of the variation in structure across the BioB sequence space indicated that there are some BioB structures that may have scope for variation at the iron-cluster binding sites, whilst potentially retaining functionality. If this is the case, then this variation can be interrogated to further investigate how these structural changes might impact the electric field, and potentially reactivity, at these sites. This may provide an important lead into understanding the limits of the functioning of biotin synthase and how nature might vary this reactivity. Further it could give a lead to protein engineering efforts to improve throughput of this enzyme industrially and circumvent the bottle-neck in current biochemical biotin production, unleashing economic, societal, and environmental advantages.

This work is not limited to biotin synthase, but could in principle impact the large number of radical SAM and other iron-sulfur containing proteins that are of interest to engineers. This could mean improved developments for the production of a large range of biochemical actives, including antivirals, anti-cancer agents and antibiotics, for example from the huge class of RiPP enzymes that are currently a hot topic.[304, 305] The very recent step-change in growth of machine-learning methods applied to enzyme systems additionally provides future opportunities to integrate with the work presented here.[306-308] Speeding up access to such a range of engineerable enzymes could ensure they become a cornerstone of future efforts to produce low-carbon and energy efficient syntheses, with the concomitant contributions to mitigating climate change.

References

1. Sumner, J. B., The isolation and crystallization of the enzyme urease. Preliminary paper. *J. Biol. Chem.* **1926,** *69* (2), 435-441.

2. Haxel, G. B.; Boore, S.; Mayfield, S., Rare Earth Elements—Critical Resources for High Technology: Figure 4 (Modified). In *USGS Fact Sheet 087-02*, Stauffer, P. H.; Hendley, J. W. I., Eds. 2002.

3. Anastas, P. T.; Warner, J. C., *Green Chemistry, Theory and Practice*. Oxford University Press: New York, 1998.

4. Berkovitch, F.; Nicolet, Y.; Wan, J. T.; Jarrett, J. T.; Drennan, C. L., Crystal structure of biotin synthase, an S-adenosylmethionine-dependent radical enzyme. *Science* **2004,** *303* (5654), 76-79.

5. PDB ID 1R30, B., F.; Nicolet, Y.; Wan, J. T.; Jarrett, J. T.; Drennan, C. L., Crystal structure of biotin synthase, an S-adenosylmethionine-dependent radical enzyme. *Science* **2004,** *303* (5654), 76-9., Mol* (D. Sehnal, S. Bittrich, M. Deshpande, R. Svobodová, K. Berka, V. Bazgier, S. Velankar, S.K. Burley, J. Koča, A.S. Rose (2021) Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. Nucleic Acids Research. doi: 10.1093/nar/gkab314), rscb.org, H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne. (2000) The Protein Data Bank Nucleic Acids Research, 28: 235-242.

6. Pierre, S.; Guillot, A.; Benjdia, A.; Sandstrom, C.; Langella, P.; Berteau, O., Thiostrepton tryptophan methyltransferase expands the chemistry of radical SAM enzymes. *Nat. Chem. Biol.* **2012,** *8* (12), 957-959.

7. Sofia, H. J.; Chen, G.; Hetzler, B. G.; Reyes-Spindola, J. F.; Miller, N. E., Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic. Acids. Res.* **2001,** *29* (5), 1097-1106.

8. Chirpich, T. P.; Zappia, V.; Costilow, R. N.; Barker, H. A., Lysine 2,3-Aminomutase - Purification and Properties of a Pyridoxal Phosphate and S-Adenosylmethionine-Activated Enzyme. *J. Biol. Chem.* **1970,** *245* (7), 1778-1789.

9. Petrovich, R. M.; Ruzicka, F. J.; Reed, G. H.; Frey, P. A., Metal Cofactors of Lysine-2,3-Aminomutase. *J. Biol. Chem.* **1991,** *266* (12), 7656-7660.

10. Knappe, J.; Schacht, J.; Mockel, W.; Hopner, T.; Vetter, H.; Edenharder, R., Pyruvate Formate-Lyase Reaction in Escherichia-Coli - Enzymatic System Converting an Inactive Form of Lyase into Catalytically Active Enzyme. *Eur. J. Biochem.* **1969,** *11* (2), 316-327.

11. Knappe, J.; Neugebauer, F. A.; Blaschkowski, H. P.; Ganzler, M., Post-Translational Activation Introduces a Free-Radical into Pyruvate Formate-Lyase. *P. Natl. Acad. Sci-Biol.* **1984,** *81* (5), 1332-1335.

12. Broderick, J. B.; Duderstadt, R. E.; Fernandez, D. C.; Wojtuszewski, K.; Henshaw, T. F.; Johnson, M. K., Pyruvate formate-lyase activating enzyme is an iron-sulfur protein. *J. Am. Chem. Soc.* **1997,** *119* (31), 7396-7397.

13. Knappe, J.; Schmitt, T., Novel Reaction of S-Adenosyl-L-Methionine Correlated with Activation of Pyruvate Formate-Lyase. *Biochem. Bioph. Res. Co.* **1976,** *71* (4), 1110-1117.

14. Mulliez, E.; Fontecave, M.; Gaillard, J.; Reichard, P., An Iron-Sulfur Center and a Free-Radical in the Active Anaerobic Ribonucleotide Reductase of Escherichia-Coli. *J. Biol. Chem.* **1993,** *268* (4), 2296-2299.

15. Birch, O. M.; Fuhrmann, M.; Shaw, N. M., Biotin Synthase from Escherichia-Coli, an Investigation of the Low-Molecular-Weight and Protein-Components Required for Activity in-Vitro. *J. Biol. Chem.* **1995,** *270* (32), 19158-19165.

16. Sanyal, I.; Cohen, G.; Flint, D. H., Biotin Synthase - Purification, Characterization as a [2fe-2s]Cluster Protein, and in-Vitro Activity of the Escherichia-Coli Biob Gene-Product. *Biochemistry-US* **1994,** *33* (12), 3625-3631.

17. Marsh, E. N. G.; Melendez, G. D. R., Adenosylcobalamin enzymes: Theory and experiment begin to converge. *BBA-Proteins. Proteom.* **2012,** *1824* (11), 1154-1164.

18. Buckel, W.; Kratky, C.; Golding, B. T., Stabilisation of Methylene Radicals by Cob(II)alamin in Coenzyme B12 Dependent Mutases. *Chem-Eur. J.* **2006,** *12* (2), 352-362.

19. Banerjee, R., Radical carbon skeleton rearrangements: Catalysis by coenzyme B-12-dependent mutases. *Chem. Rev.* **2003,** *103* (6), 2083-2094.

20. Ji, X. J.; Li, Y. Z.; Xie, L. Q.; Lu, H. J.; Ding, W.; Zhang, Q., Expanding Radical SAM Chemistry by Using Radical Addition Reactions and SAM Analogues. *Angew. Chem. Int. Ed.* **2016,** *55* (39), 11845-11848.

21. Petrovich, R. M.; Ruzicka, F. J.; Reed, G. H.; Frey, P. A., Characterization of Iron-Sulfur Clusters in Lysine 2,3-Aminomutase by Electron-Paramagnetic Resonance Spectroscopy. *Biochemistry-US* **1992,** *31* (44), 10774-10781.

22. Krebs, C.; Broderick, W. E.; Henshaw, T. F.; Broderick, J. B.; Huynh, B. H., Coordination of adenosylmethionine to a unique iron site of the [4Fe-4S] of pyruvate formate-lyase activating enzyme: A Mossbauer spectroscopic study. *J. Am. Chem. Soc.* **2002,** *124* (6), 912-913.

23. Eliasson, R.; Fontecave, M.; Jornvall, H.; Krook, M.; Pontis, E.; Reichard, P., The Anaerobic Ribonucleoside Triphosphate Reductase from Escherichia-Coli Requires S-Adenosylmethionine as a Cofactor. *Proc. Natl. Acad. Sci. USA* **1990,** *87* (9), 3314-3318.

24. Fontecave, M.; Eliasson, R.; Reichard, P., Oxygen-sensitive ribonucleoside triphosphate reductase is present in anaerobic Escherichia coli. *Proc. Natl. Acad. Sci. USA* **1989,** *86* (7), 2147-2151.

25. Chen, D. W.; Ruzicka, F. J.; Frey, P. A., A novel lysine 2,3-aminomutase encoded by the yod0 gene of Bacillus subtilis: characterization and the observation of organic radical intermediates. *Biochem. J.* **2000,** *348*, 539-549.

26. Broderick, J. B.; Duffus, B. R.; Duschene, K. S.; Shepard, E. M., Radical S-Adenosylmethionine Enzymes. *Chemical Reviews* **2014,** *114* (8), 4229-4317.

27. Walsby, C. J.; Hong, W.; Broderick, W. E.; Cheek, J.; Ortillo, D.; Broderick, J. B.; Hoffman, B. M., Electron-nuclear double resonance spectroscopic evidence that S-adenosylmethionine binds in contact with the catalytically active [4Fe-4S](+) cluster of pyruvate formate-lyase activating enzyme. *J. Am. Chem. Soc.* **2002,** *124* (12), 3143-3151.

28. Walsby, C. J.; Ortillo, D.; Broderick, W. E.; Broderick, J. B.; Hoffman, B. M., An anchoring role for FeS clusters: Chelation of the amino acid moiety of S-adenosylmethionine to the unique iron site of the [4Fe-4S] cluster of pyruvate formate-lyase activating enzyme. *J. Am. Chem. Soc.* **2002,** *124* (38), 11270-11271.

29. Yu, L.; Blaser, M.; Andrei, P. I.; Pierik, A. J.; Selmer, T., 4-hydroxyphenylacetate decarboxylases: Properties of a novel subclass of glycyl radical enzyme systems. *Biochemistry-US* **2006,** *45* (31), 9584-9592.

30. Duin, E. C.; Lafferty, M. E.; Crouse, B. R.; Allen, R. M.; Sanyal, I.; Flint, D. H.; Johnson, M. K., [2Fe-2S] to [4Fe-4S] cluster conversion in Escherichia coli biotin synthase. *Biochemistry-US* **1997,** *36* (39), 11811-11820.

31. Bandarian, V.; Drennan, C. L., Radical-mediated ring contraction in the biosynthesis of 7-deazapurines. *Curr. Opin. Struc. Biol.* **2015,** *35*, 116-124.

32. Parent, A.; Guillot, A.; Benjdia, A.; Chartier, G.; Leprince, J.; Berteau, O., The B12-Radical SAM Enzyme PoyC Catalyzes Valine Cβ-Methylation during Polytheonamide Biosynthesis. *J. Am. Chem. Soc.* **2016,** *138* (48), 15515-15518.

33. Chatterjee, A.; Li, Y.; Zhang, Y.; Grove, T. L.; Lee, M.; Krebs, C.; Booker, S. J.; Begley, T. P.; Ealick, S. E., Reconstitution of ThiC in thiamine pyrimidine biosynthesis expands the radical SAM superfamily. *Nat. Chem. Biol.* **2008,** *4* (12), 758-765.

34. Ugulava, N. B.; Sacanell, C. J.; Jarrett, J. T., Spectroscopic changes during a single turnover of biotin synthase: Destruction of a [2Fe-2S] cluster accompanies sulfur insertion. *Biochemistry-US* **2001,** *40* (28), 8352-8358.

35. Jameson, G. N. L.; Cosper, M. M.; Hernandez, H. L.; Johnson, M. K.; Huynh, B. H., Role of the [2Fe-2S] cluster in recombinant Escherichia coli biotin synthase. *Biochemistry-US* **2004,** *43* (7), 2022-2031.

36. Hernandez, H. L.; Pierrel, F.; Elleingand, E.; Garcia-Serres, R.; Huynh, B. H.; Johnson, M. K.; Fontecave, M.; Atta, M., MiaB, a bifunctional radical-S-adenosylmethionine enzyme involved in the thiolation and methylation of tRNA, contains two essential [4Fe-4S] clusters. *Biochemistry-US* **2007,** *46* (17), 5140-5147.

37. Fluhe, L.; Burghaus, O.; Wieckowski, B. M.; Giessen, T. W.; Linne, U.; Marahiel, M. A., Two [4Fe-4S] Clusters Containing Radical SAM Enzyme SkfB Catalyze Thioether Bond

Formation during the Maturation of the Sporulation Killing Factor. *J. Am. Chem. Soc.* **2013,** *135* (3), 959-962.

38. Vey, J. L.; Drennan, C. L., Structural Insights into Radical Generation by the Radical SAM Superfamily. *Chem. Rev.* **2011,** *111* (4), 2487-2506.

39. Nicolet, Y.; Drennan, C. L., AdoMet radical proteins - from structure to evolution - alignment of divergent protein sequences reveals strong secondary structure element conservation. *Nucleic. Acids. Res.* **2004,** *32* (13), 4015-4025.

40. Huber, C.; Wächtershäuser, G., Activated acetic acid by carbon fixation on (Fe,Ni)S under primordial conditions. *Science* **1997,** *276* (5310), 245-247.

41. Beinert, H.; Holm, R. H.; Münck, E., Iron-sulfur clusters: Nature's modular, multipurpose structures. *Science* **1997,** *277* (5326), 653-659.

42. Beinert, H., Iron-sulfur proteins: ancient structures, still full of surprises. *J. Biol. Inorg. Chem.* **2000,** *5* (1), 2-15.

43. Johnson, D. C.; Dean, D. R.; Smith, A. D.; Johnson, M. K., Structure, function, and formation of biological iron-sulfur clusters. *Annu. Rev. Biochem.* **2005,** *74*, 247-281.

44. Mortenson, L. E.; Valentine, R. C.; Carnahan, J. E., An electron transport factor from Clostridium pasteurianum. *Biochem. Bioph. Res. Co.* **1962,** *7*, 448-452.

45. Glaser, T.; Hedman, B.; Hodgson, K. O.; Solomon, E. I., Ligand K-edge X-ray absorption spectroscopy: A direct probe of ligand-metal covalency. *Accounts. Chem. Res.* **2000,** *33* (12), 859-868.

46. Noodleman, L.; Case, D. A., Density-Functional Theory of Spin Polarization and Spin Coupling in Iron-Sulfur Clusters. *Adv. Inorg. Chem.* **1992,** *38*, 423-458.

47. Lubitz, W.; Ogata, H.; Rüdiger, O.; Reijerse, E., Hydrogenases. *Chem. Rev.* **2014,** *114* (8), 4081-4148.

48. Peters, J. W.; Stowell, M. H. B.; Soltis, S. M.; Finnegan, M. G.; Johnson, M. K.; Rees, D. C., Redox-dependent structural changes in the nitrogenase P-cluster. *Biochemistry-US* **1997,** *36* (6), 1181-1187.

49. Cheek, J.; Broderick, J. B., Adenosylmethionine-dependent iron-sulfur enzymes: versatile clusters in a radical new role. *J. Biol. Inorg. Chem.* **2001,** *6* (3), 209-226.

50. Bandarian, V., Radical SAM Enzymes. In *Method. Enzymol.*, Bandarian, V., Ed. Academic Press: 2018; Vol. 606.

51. Jäger, C. M.; Croft, A. K., Anaerobic Radical Enzymes for Biotechnology. *Chembioeng. Rev.* **2018,** *5* (3), 143-162.

52. Fugate, C. J.; Jarrett, J. T., Biotin synthase: Insights into radical-mediated carbon-sulfur bond formation. *BBA-Proteins. Proteom.* **2012,** *1824* (11), 1213-1222.

53. Henshaw, T. F.; Cheek, J.; Broderick, J. B., The [4Fe-4S](1+) cluster of pyruvate formate-lyase activating enzyme generates the glycyl radical on pyruvate formate-lyase: EPR-detected single turnover. *J. Am. Chem. Soc.* **2000,** *122* (34), 8331-8332.

54. Lieder, K. W.; Booker, S.; Ruzicka, F. J.; Beinert, H.; Reed, G. H.; Frey, P. A., S-adenosylmethionine-dependent reduction of lysine 2,3-aminomutase and observation of the catalytically functional iron-sulfur centers by electron paramagnetic resonance. *Biochemistry-US* **1998,** *37* (8), 2578-2585.

55. Broderick, J. B.; Henshaw, T. F.; Cheek, J.; Wojtuszewski, K.; Smith, S. R.; Trojan, M. R.; McGhan, R. M.; Kopf, A.; Kibbey, M.; Broderick, W. E., Pyruvate formate-lyase-activating enzyme: Strictly anaerobic isolation yields active enzyme containing a [3Fe-4S](+) cluster. *Biochem. Bioph. Res. Co.* **2000,** *269* (2), 451-456.

56. Walsby, C. J.; Ortillo, D.; Yang, J.; Nnyepi, M. R.; Broderick, W. E.; Hoffman, B. M.; Broderick, J. B., Spectroscopic approaches to elucidating novel iron-sulfur chemistry in the "Radical-SAM" protein superfamily. *Inorg. Chem.* **2005,** *44* (4), 727-741.

57. Ollagnier-de Choudens, S.; Sanakis, Y.; Hewitson, K. S.; Roach, P.; Munck, E.; Fontecave, M., Reductive cleavage of S-adenosylmethionine by biotin synthase from Escherichia coli. *J. Biol. Chem.* **2002,** *277* (16), 13449-13454.

58. Cosper, M. M.; Jameson, G. N. L.; Davydov, R.; Eidsness, M. K.; Hoffman, B. M.; Huynh, B. H.; Johnson, M. K., The [4Fe-4S](2+) cluster in reconstituted biotin synthase binds S-adenosyl-L-methionine. *J. Am. Chem. Soc.* **2002,** *124* (47), 14006-14007.

59. Wang, S. C.; Frey, P. A., S-adenosylmethionine as an oxidant: the radical SAM superfamily. *Trends. Biochem. Sci.* **2007,** *32* (3), 101-110.

60. Frey, P. A., Lysine 2,3-Aminomutase - Is Adenosylmethionine a Poor Mans Adenosylcobalamin. *FASEB J.* **1993,** *7* (8), 662-670.

61. Jordan, A.; Reichard, P., Ribonucleotide reductases. *Annu. Rev. Biochem.* **1998,** *67*, 71-98.

62. Marsh, E. N. G.; Patterson, D. P.; Li, L., Adenosyl Radical: Reagent and Catalyst in Enzyme Reactions. *Chembiochem.* **2010,** *11* (5), 604-621.

63. Magnusson, O. T.; Frey, P. A., Synthesis and characterization of 3 ',4 '-anhydroadenosylcobalamin: A coenzyme B-12 analogue with unusual properties. *J. Am. Chem. Soc.* **2000,** *122* (37), 8807-8813.

64. Magnusson, O. T.; Frey, P. A., Interactions of diol dehydrase and 3 ',4 '-anhydroadenosylcobalamin: Suicide inactivation by electron transfer. *Biochemistry-US* **2002,** *41* (5), 1695-1702.

65. Mansoorabadi, S. O.; Magnusson, O. T.; Poyner, R. R.; Frey, P. A.; Reed, G. H., Analysis of the cob(II)alamin-5 '-deoxy-3 ',4 '-anhydroadenosyl radical triplet spin system in the active site of diol dehydrase. *Biochemistry-US* **2006,** *45* (48), 14362-14370.

66. Magnusson, O. T.; Reed, G. H.; Frey, P. A., Spectroscopic evidence for the participation of an allylic analogue of the 5 '-deoxyadenosyl radical in the reaction of lysine 2,3-aminomutase. *J. Am. Chem. Soc.* **1999,** *121* (41), 9764-9765.

67. Magnusson, O. T.; Reed, G. H.; Frey, P. A., Characterization of an allylic analogue of the 5 '-deoxyadenosyl radical: An intermediate in the reaction of lysine 2,3-aminomutase. *Biochemistry-US* **2001,** *40* (26), 7773-7782.

68. Farrar, C. E.; Jarrett, J. T., Protein Residues That Control the Reaction Trajectory in S-Adenosylmethionine Radical Enzymes: Mutagenesis of Asparagine 153 and Aspartate 155 in Escherichia coli Biotin Synthase. *Biochemistry* **2009,** *48* (11), 2448-2458.

69. Challand, M. R.; Salvadori, E.; Driesener, R. C.; Kay, C. W. M.; Roach, P. L.; Spencer, J., Cysteine Methylation Controls Radical Generation in the Cfr Radical AdoMet rRNA Methyltransferase. *Plos One* **2013,** *8* (7), Article e67979.

70. Jäger, C. M.; Croft, A. K., Radical Reaction Control in the AdoMet Radical Enzyme CDG Synthase (QueE): Consolidate, Destabilize, Accelerate. *Chem-Eur. J.* **2017,** *23* (4), 953-962.

71. Ding, W.; Li, Y. Z.; Zhang, Q., Substrate-Controlled Stereochemistry in Natural Product Biosynthesis. *ACS Chem. Biol.* **2015,** *10* (7), 1590-1598.

72. Duschene, K. S.; Veneziano, S. E.; Silver, S. C.; Broderick, J. B., Control of radical chemistry in the AdoMet radical enzymes. *Curr. Opin. Chem. Biol.* **2009,** *13* (1), 74-83.

73. Horitani, M.; Byer, A. S.; Shisler, K. A.; Chandra, T.; Broderick, J. B.; Hoffman, B. M., Why Nature Uses Radical SAM Enzymes so Widely: Electron Nuclear Double Resonance Studies of Lysine 2,3-Aminomutase Show the 5'-dAdo• "Free Radical" Is Never Free. *J. Am. Chem. Soc.* **2015,** *137* (22), 7111-7121.

74. Lees, N. S.; Chen, D. W.; Walsby, C. J.; Behshad, E.; Frey, P. A.; Hoffman, B. M., How an enzyme tames reactive intermediates: Positioning of the active-site components of lysine 2,3-aminomutase during enzymatic turnover as determined by ENDOR spectroscopy. *J. Am. Chem. Soc.* **2006,** *128* (31), 10145-10154.

75. Stich, T. A.; Myers, W. K.; Britt, R. D., Paramagnetic Intermediates Generated by Radical S-Adenosylmethionine (SAM) Enzymes. *Accounts. Chem. Res.* **2014,** *47* (8), 2235-2243.

76. Wood, Z. A.; Schröder, E.; Harris, J. R.; Poole, L. B., Structure, mechanism and regulation of peroxiredoxins. *Trends. Biochem. Sci.* **2003,** *28* (1), 32-40.

77. Arnér, E. S. J.; Holmgren, A., Physiological functions of thioredoxin and thioredoxin reductase. *Eur. J. Biochem.* **2000,** *267* (20), 6102-6109.

78. Tsukamoto, Y.; Fukushima, Y.; Hara, S.; Hisabori, T., Redox Control of the Activity of Phosphoglycerate Kinase in Synechocystis sp PCC6803. *Plant. Cell. Physiol.* **2013,** *54* (4), 484-491.

79. Gilbert, H. F., Redox control of enzyme activities by thiol/disulfide exchange. In *Method. Enzymol.*, Academic Press: 1984; Vol. 107, pp 330-351.

80. Richter, A. S.; Grimm, B., Thiol-based redox control of enzymes involved in the tetrapyrrole biosynthesis pathway in plants. *Front. Plant. Sci.* **2013,** *4*, Article 371.

81. Hitaishi, V. P.; Clement, R.; Bourassin, N.; Baaden, M.; De Poulpiquet, A.; Sacquin-Mora, S.; Ciaccafava, A.; Lojou, E., Controlling Redox Enzyme Orientation at Planar Electrodes. *Catalysts* **2018,** *8* (5), 192.

82. Caceres, T. B.; Price, O.; Morales, Y.; Hevel, J., Redox Control of PRMT1 Substrate Specificity. *FASEB J.* **2017,** *31* (S1), 765.11-765.11.

83. Bonifacio, A.; Millo, D.; Keizers, P. H. J.; Boegschoten, R.; Commandeur, J. N. M.; Vermeulen, N. P. E.; Gooijer, C.; van der Zwan, G., Active-site structure, binding and redox activity of the heme-thiolate enzyme CYP2D6 immobilized on coated Ag electrodes: a surface-enhanced resonance Raman scattering study. *J. Biol. Inorg. Chem.* **2008,** *13* (1), 85-96.

84. Skryhan, K.; Cuesta-Seijo, J. A.; Nielsen, M. M.; Marri, L.; Mellor, S. B.; Glaring, M. A.; Jensen, P. E.; Palcic, M. M.; Blennow, A., The Role of Cysteine Residues in Redox Regulation and Protein Stability of Arabidopsis thaliana Starch Synthase 1. *Plos One* **2015,** *10* (9), Article e0136997.

85. Leger, C.; Bertrand, P., Direct electrochemistry of redox enzymes as a tool for mechanistic studies. *Chem. Rev.* **2008,** *108* (7), 2379-2438.

86. Gates, A. J.; Kemp, G. L.; To, C. Y.; Mann, J.; Marritt, S. J.; Mayes, A. G.; Richardson, D. J.; Butt, J. N., The relationship between redox enzyme activity and electrochemical potential-cellular and mechanistic implications from protein film electrochemistry. *Phys. Chem. Chem. Phys.* **2011,** *13* (17), 7720-7731.

87. Elliott, S. J.; Léger, C.; Pershad, H. R.; Hirst, J.; Heffron, K.; Ginet, N.; Blasco, F.; Rothery, R. A.; Weiner, J. H.; Armstrong, F. A., Detection and interpretation of redox potential optima in the catalytic activity of enzymes. *BBA-Bioenergetics.* **2002,** *1555* (1-3), 54-59.

88. Tran, K. N.; Niu, S. Q.; Ichiye, T., Reduction potential calculations of the Fe-S clusters in Thermus thermophilus respiratory complex I. *J. Comput. Chem.* **2019,** *40* (12), 1248-1256.

89. Bak, D. W.; Elliott, S. J., Alternative FeS cluster ligands: tuning redox potentials and chemistry. *Curr. Opin. Chem. Biol.* **2014,** *19*, 50-58.

90. Zhang, L. L.; Cui, H. Y.; Zou, Z.; Garakani, T. M.; Novoa-Henriquez, C.; Jooyeh, B.; Schwaneberg, U., Directed Evolution of a Bacterial Laccase (CueO) for Enzymatic Biofuel Cells. *Angew. Chem. Int. Ed.* **2019,** *58* (14), 4562-4565.

91. Sandford, C.; Edwards, M. A.; Klunder, K. J.; Hickey, D. P.; Li, M.; Barman, K.; Sigman, M. S.; White, H. S.; Minteer, S. D., A synthetic chemist's guide to electroanalytical tools for studying reaction mechanisms. *Chem. Sci.* **2019,** *10* (26), 6404-6422.

92. Hannemann, F.; Guyot, A.; Zöllner, A.; Müller, J. J.; Heinemann, U.; Bernhardt, R., The dipole moment of the electron carrier adrenodoxin is not critical for redox partner interaction and electron transfer. *J. Inorg. Biochem.* **2009,** *103* (7), 997-1004.

93. Ikeda, T.; Kano, K., An electrochemical approach to the studies of biological redox reactions and their applications to biosensors, bioreactors, and biofuel cells. *J. Biosci. Bioeng.* **2001,** *92* (1), 9-18.

94. Warshel, A.; Aqvist, J., Electrostatic Energy and Macromolecular Function. *Annu. Rev. Biophys. Bio.* **1991,** *20* (1), 267-298.

95. Sheinerman, F. B.; Norel, R.; Honig, B., Electrostatic aspects of protein-protein interactions. *Curr. Opin. Struc. Biol.* **2000,** *10* (2), 153-159.

96. Hanoian, P.; Liu, C. T.; Hammes-Schiffer, S.; Benkovic, S., Perspectives on Electrostatics and Conformational Motions in Enzyme Catalysis. *Accounts. Chem. Res.* **2015,** *48* (2), 482-489.

97. Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H. B.; Olsson, M. H. M., Electrostatic basis for enzyme catalysis. *Chem. Rev.* **2006,** *106* (8), 3210-3235.

98. Fried, S. D.; Boxer, S. G., Electric Fields and Enzyme Catalysis. *Annu. Rev. Biochem.* **2017,** *86* (1), 387-415.

99. Warshel, A., Electrostatic origin of the catalytic power of enzymes and the role of preorganized active sites. *J. Biol. Chem.* **1998,** *273* (42), 27035-27038.

100. Morgenstern, A.; Jaszai, M.; Eberhart, M. E.; Alexandrova, A. N., Quantified electrostatic preorganization in enzymes using the geometry of the electron charge density. *Chem. Sci.* **2017,** *8* (7), 5010-5018.

101. Fuller, J.; Wilson, T. R.; Eberhart, M. E.; Alexandrova, A. N., Charge Density in Enzyme Active Site as a Descriptor of Electrostatic Preorganization. *J. Chem Inf. Model.* **2019,** *59* (5), 2367-2373.

102. Náray-Szabó, G., Electrostatic catalysis in enzymes. *J. Mol. Catal.* **1988,** *47* (2), 281-287.

103. Warshel, A., Computer simulations of enzyme catalysis: Methods, progress, and insights. *Annu. Rev. Bioph. Biom.* **2003,** *32*, 425-443.

104. Warshel, A.; Sharma, P. K.; Kato, M.; Parson, W. W., Modeling electrostatic effects in proteins. *BBA-Proteins. Proteom.* **2006,** *1764* (11), 1647-1676.

105. Nielsen, J. E.; Beier, L.; Otzen, D.; Borchert, T. V.; Frantzen, H. B.; Andersen, K. V.; Svendsen, A., Electrostatics in the active site of an alpha-amylase. *Eur. J. Biochem.* **1999,** *264* (3), 816-824.

106. Yang, Z. Y.; Liu, F.; Steeves, A. H.; Kulik, H. J., Quantum Mechanical Description of Electrostatics Provides a Unified Picture of Catalytic Action Across Methyltransferases. *J. Phys. Chem. Lett.* **2019,** *10* (13), 3779-3787.

107. Suess, C. J.; Martins, F. L.; Croft, A. K.; Jäger, C. M., Radical Stabilization Energies for Enzyme Engineering: Tackling the Substrate Scope of the Radical Enzyme QueE. *J. Chem Inf. Model.* **2019,** *59* (12), 5111-5125.

108. Gilson, M. K.; Honig, B. H., Calculation of electrostatic potentials in an enzyme active site. *Nature* **1987,** *330* (6143), 84-86.

109. Eun, C.; Kekenes-Huskey, P. M.; Metzger, V. T.; McCammon, J. A., A model study of sequential enzyme reactions and electrostatic channeling. *J. Chem. Phys.* **2014,** *140* (10), 105101.

110. Wade, R. C.; Gabdoulline, R. R.; Lüdemann, S. K.; Lounnas, V., Electrostatic steering and ionic tethering in enzyme-ligand binding: Insights from simulations. *Proc. Natl. Acad. Sci. USA* **1998,** *95* (11), 5942-5949.

111. Fried, S. D.; Bagchi, S.; Boxer, S. G., Extreme electric fields power catalysis in the active site of ketosteroid isomerase. *Science* **2014,** *346* (6216), 1510-1514.

112. Shaik, S.; Mandal, D.; Ramanan, R., Oriented electric fields as future smart reagents in chemistry. *Nat. Chem.* **2016,** *8* (12), 1091-1098.

113. Orozco-Gonzalez, Y.; Kabir, M. P.; Gozem, S., Electrostatic Spectral Tuning Maps for Biological Chromophores. *J. Phys. Chem. B.* **2019,** *123* (23), 4813-4824.

114. Prah, A.; Frančišković, E.; Mavri, J.; Stare, J., Electrostatics as the Driving Force Behind the Catalytic Function of the Monoamine Oxidase A Enzyme Confirmed by Quantum Computations. *ACS Catal.* **2019,** *9* (2), 1231-1240.

115. Bhave, D. P.; Han, W. G.; Pazicni, S.; Penner-Hahn, J. E.; Carroll, K. S.; Noodleman, L., Geometric and Electrostatic Study of the 4Fe-4S Cluster of Adenosine-5 '-Phosphosulfate Reductase from Broken Symmetry Density Functional Calculations and Extended X-ray Absorption Fine Structure Spectroscopy. *Inorg. Chem.* **2011,** *50* (14), 6610-6625.

116. Noodleman, L., Valence bond description of antiferromagnetic coupling in transition metal dimers. *J. Chem. Phys.* **1981,** *74* (10), 5737-5743.

117. Méjean, A.; Bui, B. T.; Florentin, D.; Ploux, O.; Izumi, Y.; Marquet, A., Highly purified biotin synthase can transform dethiobiotin into biotin in the absence of any other protein, in the presence of photoreduced deazaflavin. *Biochem. Bioph. Res. Co.* **1995,** *217* (3), 1231-1237.

118. Fuchs, M. G. G.; Meyer, F.; Ryde, U., A combined computational and experimental investigation of the [2Fe–2S] cluster in biotin synthase. *J. Biol. Inorg. Chem.* **2010,** *15* (2), 203-212.

119. Farrar, C. E.; Siu, K. K. W.; Howell, P. L.; Jarrett, J. T., Biotin Synthase Exhibits Burst Kinetics and Multiple Turnovers in the Absence of Inhibition by Products and Product-Related Biomolecules. *Biochemistry-US* **2010,** *49* (46), 9985-9996.

120. Reyda, M. R.; Dippold, R.; Dotson, M. E.; Jarrett, J. T., Loss of iron-sulfur clusters from biotin synthase as a result of catalysis promotes unfolding and degradation. *Arch. Biochem. Biophys.* **2008,** *471* (1), 32-41.

121. Choi-Rhee, E.; Cronan, J. E., Biotin Synthase Is Catalytic In Vivo, but Catalysis Engenders Destruction of the Protein. *Chem. Biol.* **2005,** *12* (4), 461-468.

122. Said, H. M., Biotin: Biochemical, Physiological and Clinical Aspects. In *Water Soluble Vitamins: Clinical Research and Future Application*, Stanger, O., Ed. Springer Netherlands: Dordrecht, 2012; pp 1-19.

123. Medicine, I. o., *Dietary Reference Intakes for Thiamin, Riboflavin, Niacin, Vitamin B6, Folate, Vitamin B12, Pantothenic Acid, Biotin, and Choline*. The National Academies Press: Washington, DC, 1998; p 592.

124. Ross, A. C.; Caballero, B. H.; Cousins, R. J.; Tucker, K. L.; Ziegler, T. R., *Modern nutrition in health and disease*. Wolters Kluwer Health Adis (ESP): 2012.

125. Zempleni, J.; Wijeratne, S. S. K.; Kuroishi, T., Biotin. In *Present Knowledge in Nutrition*, 2012; pp 359-374.

126. Pacheco-Alvarez, D.; Solórzano-Vargas, R. S.; Del Río, A. L., Biotin in Metabolism and Its Relationship to Human Disease. *Arch. Med. Res.* **2002,** *33* (5), 439-447.

127. Staggs, C. G.; Sealey, W. M.; McCabe, B. J.; Teague, A. M.; Mock, D. M., Determination of the biotin content of select foods using accurate and sensitive HPLC/avidin binding. *J. Food. Compos. Anal.* **2004,** *17* (6), 767-776.

128. Penberthy, W. T.; Sadri, M.; Zempleni, J., Biotin. In *Present Knowledge in Nutrition*, Eleventh ed.; Marriott, B. P.; Birt, D. F.; Stallings, V. A.; Yates, A. A., Eds. Academic Press (Elsevier): London, United Kingdom, 2020; pp 289-304.

129. Cashman, M. W.; Sloan, S. B., Nutrition and nail disease. *Clin. Dermatol.* **2010,** *28* (4), 420-425.

130. Colombo, V. E.; Gerber, F.; Bronhofer, M.; Floersheim, G. L., Treatment of brittle fingernails and onychoschizia with biotin: scanning electron microscopy. *J. Am. Acad. Dermatol.* **1990,** *23* (6 Pt 1), 1127-32.

131. Hochman, L. G.; Scher, R. K.; Meyerson, M. S., Brittle nails: response to daily biotin supplementation. *Cutis* **1993,** *51* (4), 303-305.

132. Boccaletti, V.; Zendri, E.; Giordano, G.; Gnetti, L.; De Panfilis, G., Familial Uncombable Hair Syndrome: Ultrastructural Hair Study and Response to Biotin. *Pediatr. Dermatol.* **2007,** *24* (3), E14-E16.

133. Shelley, W. B.; Shelley, E. D., Uncombable hair syndrome: observations on response to biotin and occurrence in siblings with ectodermal dysplasia. *J. Am. Acad. Dermatol.* **1985,** *13* (1), 97-102.

134. Mock, D. M.; Baswell, D. L.; Baker, H.; Holman, R. T.; Sweetman, L., Biotin deficiency complicating parenteral alimentation: diagnosis, metabolic repercussions, and treatment. *J. Pediatr.* **1985,** *106* (5), 762-769.

135. Fujimoto, W.; Inaoki, M.; Fukui, T.; Inoue, Y.; Kuhara, T., Biotin deficiency in an infant fed with amino acid formula. *J. Dermatol.* **2005,** *32* (4), 256-261.

136. Zempleni, J.; Wijeratne, S. S.; Hassan, Y. I., Biotin. *Biofactors* **2009,** *35* (1), 36-46.

137. Hermanson, G. T., *Bioconjugate Techniques*. Third ed.; Academic Press: 2013.

138. Schrodinger, E., An undulatory theory of the mechanics of atoms and molecules. *Phys. Rev.* **1926,** *28* (6), 1049-1070.

139. Kock, W. H., M. C., *A Chemist's Guide to Density Functional Theory, Second Edition*. WILEY-VCH: Weinheim, Germany, 2001.

140. Parr, R. G. Y., W., *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, Inc: New York, U.S.A., 1989.

141. Born, M.; Oppenheimer, R., Quantum theory of molecules. *Ann. Phys-Berlin.* **1927,** *84* (20), 0457-0484.

142. Heisenberg, W., The spectrums of atomic systems with two electrons. *Z. Phys.* **1926,** *39* (7/8), 499-518.

143. Epstein, S. T., *The Variation Method in Quantum Chemistry*. Academic Press, Inc: New York, U.S.A., 1974.

144. Hartree, D. R.; Hartree, F. R. S.; Hartree, W., Self-consistent field, with exchange, for beryllium. *P. Roy. Soc. Lond. A. Mat.* **1935,** *150* (A869), 0009-0033.

145. Slater, J. C., Atomic Shielding Constants. *Phys. Rev.* **1930,** *36* (1), 57-64.

146. Boys, S. F.; Egerton, A. C., Electronic wave functions - I. A general method of calculation for the stationary states of any molecular system. *P. Roy. Soc. Lond. A. Mat.* **1950,** *200* (1063), 542-554.

147. Ditchfield, R.; Hehre, W. J.; Pople, J. A., Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971,** *54* (2), 724-728.

148. Moller, C.; Plesset, M. S., Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934,** *46* (7), 0618-0622.

149. Leininger, M. L.; Allen, W. D.; Schaefer, H. F.; Sherrill, C. D., Is Moller-Plesset perturbation theory a convergent ab initio method? *J. Chem. Phys.* **2000,** *112* (21), 9213-9222.

150. Coester, F.; Kummel, H., Short-Range Correlations in Nuclear Wave Functions. *Nucl. Phys.* **1960,** *17* (3), 477-485.

151. Cizek, J., On Correlation Problem in Atomic and Molecular Systems . Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods. *J. Chem. Phys.* **1966,** *45* (11), 4256-4266.

152. Shavitt, I. B., R. J., *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*. Cambridge University Press: 2009.

153. Cramer, C. J., *Essentials of Computational Chemistry*. John Wiley and Sons, Ltd: 2002.

154. Craig, D. P., Configurational Interaction in Molecular Orbital Theory - a Higher Approximation in the Non-Empirical Method. *P. Roy. Soc. Lond. A. Mat.* **1950,** *200* (1063), 474-486.

155. Sellier, J. M.; Dimov, I., On a full Monte Carlo approach to quantum mechanics. *Physica. A.* **2016,** *463*, 45-62.

156. Austin, B. M.; Zubarev, D. Y.; Lester, W. A., Quantum Monte Carlo and Related Approaches. *Chem. Rev.* **2012,** *112* (1), 263-288.

157. Dubecky, M., Quantum Monte Carlo for Noncovalent Interactions: A Tutorial Review. *Acta. Phys. Slovaca.* **2014,** *64* (5), 501-575.

158. Carlson, J.; Gandolfi, S.; Pederiva, F.; Pieper, S. C.; Schiavilla, R.; Schmidt, K. E.; Wiringa, R. B., Quantum Monte Carlo methods for nuclear physics. *Rev. Mod. Phys.* **2015,** *87* (3), 1067-1118.

159. Dubecky, M.; Mitas, L.; Jurecka, P., Noncovalent Interactions by Quantum Monte Carlo. *Chem. Rev.* **2016,** *116* (9), 5188-5215.

160. Hohenberg, P.; Kohn, W., Inhomogeneous Electron Gas. *Phys. Rev. B.* **1964,** *136* (3b), B864-B871.

161. Kohn, W.; Sham, L. J., Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965,** *140* (4a), A1133-A1138.

162. Mantel, C.; Chandor, A.; Gasparutto, D.; Douki, T.; Atta, M.; Fontecave, M.; Bayle, P. A.; Mouesca, J. M.; Bardet, M., Combined NMR and DFT Studies for the Absolute Configuration Elucidation of the Spore Photoproduct, a UV-Induced DNA Lesion. *J. Am. Chem. Soc.* **2008,** *130* (50), 16978-16984.

163. Guo, J. D.; Luo, Y.; Himo, F., DNA repair by spore photoproduct lyase: A density functional theory study. *J. Phys. Chem. B.* **2003,** *107* (40), 11188-11192.

164. Himo, F., C-C bond formation and cleavage in radical enzymes, a theoretical perspective. *BBA-Bioenergetics.* **2005,** *1707* (1), 24-33.

165. Kuchenreuther, J. M.; Myers, W. K.; Stich, T. A.; George, S. J.; NejatyJahromy, Y.; Swartz, J. R.; Britt, R. D., A Radical Intermediate in Tyrosine Scission to the CO and CN- Ligands of FeFe Hydrogenase. *Science* **2013,** *342* (6157), 472-475.

166. Perdew, J. P.; Schmidt, K., Jacob's ladder of density functional approximations for the exchange-correlation energy. *AIP Conf. Proc.* **2001,** *577* (1), 1-20.

167. Parr, R. G. In *Density Functional Theory of Atoms and Molecules*, Dordrecht, Springer Netherlands: Dordrecht, 1980; pp 5-15.

168. Levy, M., Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proc. Natl. Acad. Sci. USA* **1979,** *76* (12), 6062-6065.

169. Sanderson, R. T., Partial Charges on Atoms in Organic Compounds. *Science* **1955,** *121* (3137), 207-208.

170. Gilbert, T. L., Hohenberg-Kohn theorem for nonlocal external potentials. *Phys. Rev. B.* **1975,** *12* (6), 2111-2120.

171. Donnelly, R. A.; Parr, R. G., Elementary properties of an energy functional of the first-order reduced density matrix. *J. Chem. Phys.* **2008,** *69* (10), 4431-4439.

172. Langreth, D. C.; Mehl, M. J., Beyond the local-density approximation in calculations of ground-state electronic properties. *Phys. Rev. B.* **1983,** *28* (4), 1809-1834.

173. Perdew, J. P.; Chevary, J. A.; Vosko, S. H.; Jackson, K. A.; Pederson, M. R.; Singh, D. J.; Fiolhais, C., Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Phys. Rev. B.* **1992,** *46* (11), 6671-6687.

174. Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996,** *77* (18), 3865-3868.

175. Becke, A. D., Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A.* **1988,** *38* (6), 3098-3100.

176. Lee, C.; Yang, W.; Parr, R. G., Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B.* **1988,** *37* (2), 785-789.

177. Handy, N. C.; Cohen, A. J., Left-right correlation energy. *Mol. Phys.* **2001,** *99* (5), 403-412.

178. Perdew, J. P., Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B.* **1986,** *33* (12), 8822-8824.

179. Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E., Climbing the Density Functional Ladder: Nonempirical Meta--Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003,** *91* (14), 146401.

180. Zhao, Y.; Truhlar, D. G., A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J. Chem. Phys.* **2006,** *125* (19), Article 194101.

181. Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Constantin, L. A.; Sun, J., Workhorse Semilocal Density Functional for Condensed Matter Physics and Quantum Chemistry. *Phys. Rev. Lett.* **2009,** *103* (2), 026403.

182. Sun, J.; Ruzsinszky, A.; Perdew, J. P., Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Phys. Rev. Lett.* **2015,** *115* (3), 036402.

183. Becke, A. D., Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993,** *98* (7), 5648-5652.

184. Wang, Y.; Jin, X.; Yu, H. S.; Truhlar, D. G.; He, X., Revised M06-L functional for improved accuracy on chemical reaction barrier heights, noncovalent interactions, and solid-state physics. *Proc. Natl. Acad. Sci. USA* **2017,** *114* (32), 8487-8492.

185. Zhao, Y.; Truhlar, D. G., The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008,** *120* (1), 215-241.

186. Wang, Y.; Verma, P.; Jin, X.; Truhlar, D. G.; He, X., Revised M06 density functional for main-group and transition-metal chemistry. *Proc. Natl. Acad. Sci. USA* **2018,** *115* (41), 10257-10262.

187. Zhao, Y.; Truhlar, D. G., Density Functional for Spectroscopy: No Long-Range Self-Interaction Error, Good Performance for Rydberg and Charge-Transfer States, and Better Performance on Average than B3LYP for Ground States. *J. Phys. Chem. A.* **2006,** *110* (49), 13126-13130.

188. Szabo, A.; Szabó, A.; Ostlund, N. S., *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Macmillan: 1982.

189. Gunnarsson, O.; Lundqvist, B. I., Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism. *Phys. Rev. B.* **1976,** *13* (10), 4274-4298.

190. Pople, J. A.; Gill, P. M. W.; Handy, N. C., Spin-unrestricted character of Kohn-Sham orbitals for open-shell systems. *Int. J. Quantum Chem.* **1995,** *56* (4), 303-305.

191. Perdew, J. P.; Ruzsinszky, A.; Constantin, L. A.; Sun, J.; Csonka, G. I., Some Fundamental Issues in Ground-State Density Functional Theory: A Guide for the Perplexed. *J. Chem. Theor. Comput.* **2009,** *5* (4), 902-908.

192. Roothaan, C. C. J., Self-Consistent Field Theory for Open Shells of Electronic Systems. *Rev. Mod. Phys.* **1960,** *32* (2), 179-185.

193. Noodleman, L.; Baerends, E. J., Electronic structure, magnetic properties, ESR, and optical spectra for 2-iron ferredoxin models by LCAO-X.alpha. valence bond theory. *J. Am. Chem. Soc.* **1984,** *106* (8), 2316-2327.

194. Noodleman, L.; Davidson, E. R., Ligand spin polarization and antiferromagnetic coupling in transition metal dimers. *Chem. Phys.* **1986,** *109* (1), 131-143.

195. Mouesca, J.-M., Density Functional Theory–Broken Symmetry (DFT–BS) Methodology Applied to Electronic and Magnetic Properties of Bioinorganic Prosthetic Groups. In *Metalloproteins: Methods and Protocols*, Fontecilla-Camps, J. C.; Nicolet, Y., Eds. Humana Press: Totowa, NJ, 2014; pp 269-296.

196. Heisenberg, W., Zur Theorie des Ferromagnetismus. *Z. Phys.* **1928,** *49* (9), 619-636.

197. Liechtenstein, A. I.; Katsnelson, M. I.; Antropov, V. P.; Gubanov, V. A., Local spin density functional approach to the theory of exchange interactions in ferromagnetic metals and alloys. *J. Magn. Magn. Mater.* **1987,** *67* (1), 65-74.

198. Pajda, M.; Kudrnovský, J.; Turek, I.; Drchal, V.; Bruno, P., Ab initio calculations of exchange interactions, spin-wave stiffness constants, and Curie temperatures of Fe, Co, and Ni. *Phys. Rev. B.* **2001,** *64* (17), 174402.

199. Gillum, W. O.; Frankel, R. B.; Foner, S.; Holm, R. H., Synthetic analogues of the active sites of iron-sulfur proteins. XIII. Further electronic structural relationships between the analogues [Fe2S2(SR)4]2- and the active sites of oxidized 2Fe-2S* proteins. *Inorg. Chem.* **1976,** *15* (5), 1095-1100.

200. Fermi, E.; Pasta, P.; Ulam, S.; Tsingou, M. *Studies of the Nonlinear Problems*; LA-1940; Los Alamos Scientific Lab., N. Mex.: United States, 1955; p 20.

201. Alder, B. J.; Wainwright, T. E., Studies in Molecular Dynamics .1. General Method. *J. Chem. Phys.* **1959,** *31* (2), 459-466.

202. Rahman, A., Correlations in Motion of Atoms in Liquid Argon. *Phys. Rev.* **1964,** *136* (2a), A405-A411.

203. Mccammon, J. A.; Gelin, B. R.; Karplus, M., Dynamics of Folded Proteins. *Nature* **1977,** *267* (5612), 585-590.

204. Sun, X.; Qian, M. D.; Guan, S. S.; Shan, Y. M.; Dong, Y.; Zhang, H.; Wang, S.; Han, W. W., Investigation of an "alternate water supply system" in enzymatic hydrolysis in the processive endocellulase Cel7A from Rasamsonia emersonii by molecular dynamics simulation. *Biopolymers* **2017,** *107* (2), 46-60.

205. Karplus, M.; Kuriyan, J., Molecular dynamics and protein function. *Proc. Natl. Acad. Sci. USA* **2005,** *102* (19), 6679-6685.

206. Warshel, A.; Levitt, M., Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976,** *103* (2), 227-249.

207. Bakowies, D.; Thiel, W., Hybrid models for combined quantum mechanical and molecular mechanical approaches. *J. Phys. Chem.* **1996,** *100* (25), 10580-10594.

208. Field, M. J.; Bash, P. A.; Karplus, M., A Combined Quantum-Mechanical and Molecular Mechanical Potential for Molecular-Dynamics Simulations. *J. Comput. Chem.* **1990,** *11* (6), 700-733.

209. Gao, J. L.; Truhlar, D. G., Quantum mechanical methods for enzyme kinetics. *Annu. Rev. Phys. Chem.* **2002,** *53*, 467-505.

210. Monard, G.; Merz, K. M., Combined quantum mechanical/molecular mechanical methodologies applied to biomolecular systems. *Accounts. Chem. Res.* **1999,** *32* (10), 904-911.

211. Mordasini, T. Z.; Thiel, W., Combined quantum mechanical and molecular mechanical approaches. *Chimia* **1998,** *52* (6), 288-291.

212. Feliks, M.; Martins, B. M.; Ullmann, G. M., Catalytic Mechanism of the Glycyl Radical Enzyme 4-Hydroxyphenylacetate Decarboxylase from Continuum Electrostatic and QC/MM Calculations. *J. Am. Chem. Soc.* **2013,** *135* (39), 14574-14585.

213. Zhu, W. Y.; Liu, Y. J.; Zhang, R., QM/MM study of the conversion mechanism of lysine to methylornithine catalyzed by methylornithine synthase (PylB). *Theor. Chem. Acc.* **2013,** *132* (9), Article 1385.

214. Senn, H. M.; Thiel, W., QM/MM Methods for Biomolecular Systems. *Angew. Chem. Int. Ed.* **2009,** *48* (7), 1198-1229.

215. Duarte, F.; Amrein, B. A.; Blaha-Nelson, D.; Kamerlin, S. C. L., Recent advances in QM/MM free energy calculations using reference potentials. *BBA-Gen. Subjects.* **2015,** *1850* (5), 954-965.

216. Rivail, J. L.; Ruiz-Lopez, M. F.; Assfeld, X., Quantum modeling of complex molecular systems. In *Challenges and Advances in Computational Chemistry and Physics* [Online] Springer Cham: 2015; p. 523. http://public.eblib.com/choice/publicfullrecord.aspx?p=4178417.

217. Konig, P. H.; Hoffmann, M.; Frauenheim, T.; Cui, Q., A critical evaluation of different QM/MM frontier treatments with SCC-DFTB as the QM method. *J. Phys. Chem. B.* **2005,** *109* (18), 9082-9095.

218. Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M., Frontier bonds in QM/MM methods: A comparison of different approaches. *J. Phys. Chem. A.* **2000,** *104* (8), 1720-1735.

219. Lin, H.; Truhlar, D. G., Redistributed charge and dipole schemes for combined quantum mechanical and molecular mechanical calculations. *J. Phys. Chem. A.* **2005,** *109* (17), 3991-4004.

220. Amara, P.; Field, M. J., Evaluation of an ab initio quantum mechanical/molecular mechanical hybrid-potential link-atom method. *Theor. Chem. Acc.* **2003,** *109* (1), 43-52.

221. Hall, R. J.; Hindle, S. A.; Burton, N. A.; Hillier, I. H., Aspects of hybrid QM/MM calculations: The treatment of the QM/MM interface region and geometry optimization with an application to chorismate mutase. *J. Comput. Chem.* **2000,** *21* (16), 1433-1441.

222. Nicoll, R. M.; Hindle, S. A.; MacKenzie, G.; Hillier, I. H.; Burton, N. A., Quantum mechanical/molecular mechanical methods and the study of kinetic isotope effects:

modelling the covalent junction region and application to the enzyme xylose isomerase. *Theor. Chem. Acc.* **2001,** *106* (1-2), 105-112.

223. Lennartz, C.; Schafer, A.; Terstegen, F.; Thiel, W., Enzymatic reactions of triosephosphate isomerase: A theoretical calibration study. *J. Phys. Chem. B.* **2002,** *106* (7), 1758-1767.

224. Pezeshki, S.; Lin, H., Recent developments in QM/MM methods towards open-boundary multi-scale simulations. *Mol. Simulat.* **2015,** *41* (1-3), 168-189.

225. Heyden, A.; Lin, H.; Truhlar, D. G., Adaptive partitioning in combined quantum mechanical and molecular mechanical calculations of potential energy functions for multiscale simulations. *J. Phys. Chem. B.* **2007,** *111* (9), 2231-2241.

226. Park, K.; Gotz, A. W.; Walker, R. C.; Paesani, F., Application of Adaptive QM/MM Methods to Molecular Dynamics Simulations of Aqueous Systems. *J. Chem. Theor. Comput.* **2012,** *8* (8), 2868-2877.

227. Varnai, C.; Bernstein, N.; Mones, L.; Csanyi, G., Tests of an Adaptive QM/MM Calculation on Free Energy Profiles of Chemical Reactions in Solution. *J. Phys. Chem. B.* **2013,** *117* (40), 12202-12211.

228. Pezeshki, S.; Lin, H., Adaptive-Partitioning Redistributed Charge and Dipole Schemes for QM/MM Dynamics Simulations: On-the-fly Relocation of Boundaries that Pass through Covalent Bonds. *J. Chem. Theor. Comput.* **2011,** *7* (11), 3625-3634.

229. Bulo, R. E.; Ensing, B.; Sikkema, J.; Visscher, L., Toward a Practical Method for Adaptive QM/MM Simulations. *J. Chem. Theor. Comput.* **2009,** *5* (9), 2212-2221.

230. Dohm, S.; Spohr, E.; Korth, M., Developing Adaptive QM/MM Computer Simulations for Electrochemistry. *J. Comput. Chem.* **2017,** *38* (1), 51-58.

231. Levitt, M.; Warshel, A., Computer-Simulation of Protein Folding. *Nature* **1975,** *253* (5494), 694-698.

232. Reynwar, B. J.; Illya, G.; Harmandaris, V. A.; Muller, M. M.; Kremer, K.; Deserno, M., Aggregation and vesiculation of membrane proteins by curvature-mediated interactions. *Nature* **2007,** *447* (7143), 461-464.

233. Yao, X. Q.; Kenzaki, H.; Murakami, S.; Takada, S., Drug export and allosteric coupling in a multidrug transporter revealed by molecular simulations. *Nat. Commun.* **2010,** *1*, Article 117.

234. Clementi, C.; Nymeyer, H.; Onuchic, J. N., Topological and energetic factors: What determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? An investigation for small globular proteins. *J. Mol. Biol.* **2000,** *298* (5), 937-953.

235. Cheung, M. S.; Finke, J. M.; Callahan, B.; Onuchic, J. N., Exploring the interplay between topology and secondary structural formation in the protein folding problem. *J. Phys. Chem. B.* **2003,** *107* (40), 11193-11200.

236. Oliveira, L. C.; Schug, A.; Onuchic, J. N., Geometrical features of the protein folding mechanism are a robust property of the energy landscape: A detailed investigation of several reduced models. *J. Phys. Chem. B.* **2008,** *112* (19), 6131-6136.

237. Wang, Y.; Tang, C.; Wang, E. K.; Wang, J., Exploration of Multi-State Conformational Dynamics and Underlying Global Functional Landscape of Maltose Binding Protein. *Plos. Comput. Biol.* **2012,** *8* (4), Article e1002471.

238. Wang, Y.; Gan, L. F.; Wang, E. K.; Wang, J., Exploring the Dynamic Functional Landscape of Adenylate Kinase Modulated by Substrates. *J. Chem. Theor. Comput.* **2013,** *9* (1), 84-95.

239. Gorfe, A. A.; Chang, C. E. A.; Ivanov, I.; McCammon, J. A., Dynamics of the acetylcholinesterase tetramer. *Biophys. J.* **2008,** *94* (4), 1144-1154.

240. Kamerlin, S. C. L.; Vicatos, S.; Dryga, A.; Warshel, A., Coarse-Grained (Multiscale) Simulations in Studies of Biophysical and Chemical Systems. *Annu. Rev. Phys. Chem.* **2011,** *62*, 41-64.

241. Crowet, J. M.; Parton, D. L.; Hall, B. A.; Steinhauer, S.; Brasseur, R.; Lins, L.; Sansom, M. S. P., Multi-Scale Simulation of the Simian Immunodeficiency Virus Fusion Peptide. *J. Phys. Chem. B.* **2012,** *116* (46), 13713-13721.

242. Shi, Q.; Voth, G. A., Multi-scale modeling of phase separation in mixed lipid bilayers. *Biophys. J.* **2005,** *89* (4), 2385-2394.

243. Coluzza, I.; De Simone, A.; Fraternali, F.; Frenkel, D., Multi-Scale Simulations Provide Supporting Evidence for the Hypothesis of Intramolecular Protein Translocation in GroEL/GroES Complexes. *Plos. Comput. Biol.* **2008,** *4* (2), Article e1000006.

244. Shih, A. Y.; Arkhipov, A.; Freddolino, P. L.; Schulten, K., Coarse grained protein-lipid model with application to lipoprotein particles. *J. Phys. Chem. B.* **2006,** *110* (8), 3674-3684.

245. Thirumalai, D.; O'Brien, E. P.; Morrison, G.; Hyeon, C., Theoretical Perspectives on Protein Folding. *Annu. Rev. Biophys.* **2010,** *39*, 159-183.

246. Zheng, W. H.; Schafer, N. P.; Davtyan, A.; Papoian, G. A.; Wolynes, P. G., Predictive energy landscapes for protein-protein association. *Proc. Natl. Acad. Sci. USA* **2012,** *109* (47), 19244-19249.

247. Karton, A., A computational chemist's guide to accurate thermochemistry for organic molecules. *Wires. Comput. Mol. Sci.* **2016,** *6* (3), 292-310.

248. Hioe, J.; Zipse, H., Radicals in enzymatic catalysis-a thermodynamic perspective. *Faraday. Discuss.* **2010,** *145*, 301-313.

249. Hioe, J.; Zipse, H., Hydrogen Transfer in SAM-Mediated Enzymatic Radical Reactions. *Chem-Eur. J.* **2012,** *18* (51), 16463-16472.

250. Cramer, C. J.; Truhlar, D. G., Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics. *Chem. Rev.* **1999,** *99* (8), 2161-2200.

251. Bonaccorsi, R.; Cimiraglia, R.; Tomasi, J., Ab initio evaluation of absorption and emission transitions for molecular solutes, including separate consideration of orientational and inductive solvent effects. *J. Comput. Chem.* **1983,** *4* (4), 567-577.

252. Klamt, A.; Schüürmann, G., COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2* **1993,** (5), 799-805.

253. Klamt, A.; Jonas, V., Treatment of the outlying charge in continuum solvation models. *J. Chem. Phys.* **1996,** *105* (22), 9972-9981.

254. Klamt, A., The COSMO and COSMO-RS solvation models. *Wires. Comput. Mol. Sci.* **2018,** *8* (1), e1338.

255. Shao, Y. H.; Gan, Z. T.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X. T.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kús, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; Woodcock III, H. L.; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C. M.; Chen, Y. Q.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; DiStasio, R. A.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T. C.; Ji, H. J.; Kaduk, B.; Khistyaev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A. D.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S. P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Neuscamman, E.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stuck, D.; Su, Y. C.; Thom, A. J. W.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z. Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K. L.; Chipman, D. M.; Cramer, C. J.; Goddard, W. A.; Gordon, M. S.; Hehre, W. J.; Klamt, A.; Schaefer, H. F.; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J. D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C. P.; Jung, Y. S.; Kong, J.; Lambrecht, D. S.; Liang, W. Z.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Van Voorhis, T.; Herbert, J. M.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M., Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015,** *113* (2), 184-215.

256. Carvalho, A. T. P.; Swart, M., Electronic Structure Investigation and Parametrization of Biologically Relevant Iron-Sulfur Clusters. *J. Chem Inf. Model.* **2014,** *54* (2), 613-620.

257. Swart, M.; Groenhof, A. R.; Ehlers, A. W.; Lammertsma, K., Validation of exchange - Correlation functionals for spin states of iron complexes. *J. Phys. Chem. A.* **2004,** *108* (25), 5479-5483.

258. Swart, M.; Ehlers, A. W.; Lammertsma, K., Performance of the OPBE exchange-correlation functional. *Mol. Phys.* **2004,** *102* (23-24), 2467-2474.

259. Van Lenthe, E.; Baerends, E. J., Optimized slater-type basis sets for the elements 1-118. *J. Comput. Chem.* **2003,** *24* (9), 1142-1156.

260. Chong, D. P.; Van Lenthe, E.; Van Gisbergen, S.; Baerends, E. J., Even-tempered slater-type orbitals revisited: From hydrogen to krypton. *J. Comput. Chem.* **2004,** *25* (8), 1030-1036.

261. Chong, D. P., Augmenting basis set for time-dependent density functional theory calculation of excitation energies: Slater-type orbitals for hydrogen to krypton. *Mol. Phys.* **2005,** *103* (6-8), 749-761.

262. Harris, T. V.; Szilagyi, R. K., Iron-Sulfur Bond Covalency from Electronic Structure Calculations for Classical Iron-Sulfur Clusters. *J. Comput. Chem.* **2014,** *35* (7), 540-552.

263. Szilagyi, R. K.; Winslow, M. A., On the accuracy of density functional theory for iron - Sulfur clusters. *J. Comput. Chem.* **2006,** *27* (12), 1385-1397.

264. Harris, T. V.; Szilagyi, R. K., Protein Environmental Effects on Iron-Sulfur Clusters: A Set of Rules for Constructing Computational Models for Inner and Outer Coordination Spheres. *J. Comput. Chem.* **2016,** *37* (18), 1681-1696.

265. Swart, M., Accurate Spin-State Energies for Iron Complexes. *J. Chem. Theor. Comput.* **2008,** *4* (12), 2057-2066.

266. Güell, M.; Luis, J. M.; Solà, M.; Swart, M., Importance of the basis set for the spin-state energetics of iron complexes. *J. Phys. Chem. A.* **2008,** *112* (28), 6384-6391.

267. Gilson, M. K.; Honig, B. H., The dielectric constant of a folded protein. *Biopolymers* **1986,** *25* (11), 2097-2119.

268. Sharp, K. A.; Honig, B., Electrostatic Interactions in Macromolecules: Theory and Applications. *Annu. Rev. Biophys. Bio.* **1990,** *19* (1), 301-332.

269. Noodleman, L.; Peng, C. Y.; Case, D. A.; Mouesca, J. M., Orbital interactions, electron delocalization and spin coupling in iron-sulfur clusters. *Coord. Chem. Rev.* **1995,** *144*, 199-244.

270. Emel'yanova, N. S.; Poleshchuk, O. K.; Sanina, N. A.; Aldoshin, S. M., Quantum chemical modeling of ligand substitution in cationic nitrosyl iron complexes. *Russ. Chem. Bull.* **2014,** *63* (5), 1088-1094.

271. Sandala, G. M.; Hopmann, K. H.; Ghosh, A.; Noodleman, L., Calibration of DFT Functionals for the Prediction of Fe-57 Mossbauer Spectral Parameters in Iron-Nitrosyl and Iron-Sulfur Complexes: Accurate Geometries Prove Essential. *J. Chem. Theor. Comput.* **2011,** *7* (10), 3232-3247.

272. Jensen, K. P.; Ooi, B.-L.; Christensen, H. E. M., Accurate computation of reduction potentials of 4Fe-4S clusters indicates a carboxylate shift in Pyrococcus furiosus ferredoxin. *Inorg. Chem.* **2007,** *46* (21), 8710-8716.

273. Han, W.-G.; Liu, T. Q.; Lovell, T.; Noodleman, L., Active site structure of class I ribonucleotide reductase intermediate X: A density functional theory analysis of structure, energetics, and spectroscopy. *J. Am. Chem. Soc.* **2005,** *127* (45), 15778-15790.

274. Maiocco, S. J.; Walker, L. M.; Elliott, S. J., Determining Redox Potentials of the Iron-Sulfur Clusters of the AdoMet Radical Enzyme Superfamily. In *Radical Sam Enzymes*, Bandarian, V., Ed. 2018; Vol. 606, pp 319-339.

275. Miranda-Quintana, R. A.; González, M. M.; Ayers, P. W., Electronegativity and redox reactions. *Phys. Chem. Chem. Phys.* **2016,** *18* (32), 22235-22243.

276. Parr, R. G.; Von Szentpály, L. v.; Liu, S. B., Electrophilicity index. *J. Am. Chem. Soc.* **1999,** *121* (9), 1922-1924.

277. Moens, J.; Roos, G.; Jaque, P.; Proft, F.; Geerlings, P., Can electrophilicity act as a measure of the redox potential of first-row transition metal ions? *Chem-Eur. J.* **2007,** *13* (33), 9331-9343.

278. Chattaraj, P. K.; Giri, S., Electrophilicity index within a conceptual DFT framework. In *Annual Reports on the Progress of Chemistry 2009, Vol 105, Section C: Physical Chemistry*, Webb, G. A., Ed. 2009; Vol. 105, pp 13-39.

279. Moens, J.; Jaque, P.; De Proft, F.; Geerlings, P., The study of redox reactions on the basis of conceptual DFT principles: EEM and vertical quantities. *J. Phys. Chem. A.* **2008,** *112* (26), 6023-6031.

280. Roos, G.; Geerlings, P.; Messens, J., Enzymatic Catalysis: The Emerging Role of Conceptual Density Functional Theory. *J. Phys. Chem. B.* **2009,** *113* (41), 13465-13475.

281. Bradshaw, R. T.; Dziedzic, J.; Skylaris, C. K.; Essex, J. W., The Role of Electrostatics in Enzymes: Do Biomolecular Force Fields Reflect Protein Electric Fields? *J. Chem Inf. Model.* **2020,** *60* (6), 3131-3144.

282. Fried, S. D.; Bagchi, S.; Boxer, S. G., Measuring Electrostatic Fields in Both Hydrogen-Bonding and Non-Hydrogen-Bonding Environments Using Carbonyl Vibrational Probes. *J. Am. Chem. Soc.* **2013,** *135* (30), 11181-11192.

283. Welborn, V. V.; Head-Gordon, T., Fluctuations of Electric Fields in the Active Site of the Enzyme Ketosteroid Isomerase. *J. Am. Chem. Soc.* **2019,** *141* (32), 12487-12492.

284. Sharma, S.; Sivalingam, K.; Neese, F.; Chan, G. K.-L., Low-energy spectrum of iron–sulfur clusters directly from many-particle quantum mechanics. *Nat. Chem.* **2014,** *6* (10), 927-933.

285. Teo, R. D.; Rousseau, B. J. G.; Smithwick, E. R.; Di Felice, R.; Beratan, D. N.; Migliore, A., Charge Transfer between 4Fe4S Proteins and DNA Is Unidirectional: Implications for Biomolecular Signaling. *Chem* **2019,** *5* (1), 122-137.

286. Cervantes-Salguero, K.; Seminario, J. M., Structure and energetics of small iron clusters. *J. Mol. Model.* **2012,** *18* (9), 4043-4052.

287. Nicolet, Y.; Amara, P.; Mouesca, J.-M.; Fontecilla-Camps, J. C., Unexpected electron transfer mechanism upon AdoMet cleavage in radical SAM proteins. *Proc. Natl. Acad. Sci. USA* **2009,** *106* (35), 14867-14871.

288. Wachters, A. J. H., Gaussian Basis Set for Molecular Wavefunctions Containing Third-Row Atoms. *J. Chem. Phys.* **1970,** *52* (3), 1033-1036.

289. Hay, P. J., Gaussian basis sets for molecular calculations. The representation of 3d orbitals in transition-metal atoms. *J. Chem. Phys.* **1977,** *66* (10), 4377-4384.

290. Frisch, M. J.; Pople, J. A.; Binkley, J. S., Self-consistent molecular orbital methods 25. Supplementary functions for Gaussian basis sets. *J. Chem. Phys.* **1984,** *80* (7), 3265-3269.

291. Akiva, E.; Brown, S.; Almonacid, D. E.; Barber, A. E., 2nd; Custer, A. F.; Hicks, M. A.; Huang, C. C.; Lauck, F.; Mashiyama, S. T.; Meng, E. C.; Mischel, D.; Morris, J. H.; Ojha, S.; Schnoes, A. M.; Stryke, D.; Yunes, J. M.; Ferrin, T. E.; Holliday, G. L.; Babbitt, P. C., The Structure-Function Linkage Database. *Nucleic. Acids. Res.* **2014,** *42* (Database issue), D521-30.

292. Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J. H.; Zhang, Z.; Miller, W.; Lipman, D. J., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic. Acids. Res.* **1997,** *25* (17), 3389-3402.

293. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T., Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome. Res.* **2003,** *13* (11), 2498-504.

294. Yang, J.; Zhang, Y., I-TASSER server: new development for protein structure and function predictions. *Nucleic. Acids. Res.* **2015,** *43* (W1), W174-W181.

295. Zheng, W.; Zhang, C.; Li, Y.; Pearce, R.; Bell, E. W.; Zhang, Y., Folding non-homologous proteins by coupling deep-learning contact maps with I-TASSER assembly simulations. *Cell. Rep. Meth.* **2021,** *1* (3), 100014.

296. Du, Z.; Su, H.; Wang, W.; Ye, L.; Wei, H.; Peng, Z.; Anishchenko, I.; Baker, D.; Yang, J., The trRosetta server for fast and accurate protein structure prediction. *Nat. Protoc.* **2021,** *16* (12), 5634-5651.

297. Su, H.; Wang, W.; Du, Z.; Peng, Z.; Gao, S.-H.; Cheng, M.-M.; Yang, J., Improved Protein Structure Prediction Using a New Multi-Scale Network and Homologous Templates. *Adv. Sci.* **2021,** *8* (24), 2102592.

298. Yang, J.; Anishchenko, I.; Park, H.; Peng, Z.; Ovchinnikov, S.; Baker, D., Improved protein structure prediction using predicted interresidue orientations. *Proc. Natl. Acad. Sci. USA* **2020,** *117* (3), 1496-1503.

299. Wang, S.; Ma, J.; Peng, J.; Xu, J., Protein structure alignment beyond spatial proximity. *Sci. Rep-UK.* **2013,** *3* (1), Article 1448.

300. Wang, S.; Peng, J.; Xu, J., Alignment of distantly related protein structures: algorithm, bound and implications to homology modeling. *Bioinformatics* **2011,** *27* (18), 2537-2545.

301. Holm, L., Dali server: structural unification of protein families. *Nucleic. Acids. Res.* **2022,** *50* (W1), W210-W215.

302. Holm, L., DALI and the persistence of protein shape. *Protein. Sci.* **2020,** *29* (1), 128-140.

303. Structure-Function Linkage Database. http://sfld.rbvi.ucsf.edu/archive/django/index.html.

304. Walker, J. A.; Hamlish, N.; Tytla, A.; Brauer, D. D.; Francis, M. B.; Schepartz, A., Redirecting RiPP Biosynthetic Enzymes to Proteins and Backbone-Modified Substrates. *ACS Cent. Sci.* **2022,** *8* (4), 473-482.

305. Glassey, E.; King, A. M.; Anderson, D. A.; Zhang, Z.; Voigt, C. A., Functional expression of diverse post-translational peptide-modifying enzymes in Escherichia coli under uniform expression and purification conditions. *PLOS ONE* **2022,** *17* (9), Article e0266488.

306. Mazurenko, S.; Prokop, Z.; Damborsky, J., Machine Learning in Enzyme Engineering. *ACS Catal.* **2020,** *10* (2), 1210-1223.

307. Feehan, R.; Montezano, D.; Slusky, J. S. G., Machine learning for enzyme engineering, selection and design. *Protein. Eng. Des. Sel.* **2021,** *34*, 1-10.

308. Goldman, S.; Das, R.; Yang, K. K.; Coley, C. W., Machine learning modeling of family wide enzyme-substrate specificity screens. *Plos. Comput. Biol.* **2022,** *18* (2), Article e1009853.

Appendix

Section A3.1 – Axes of rotation definitions and initial positions

The axes about which we rotate the field are defined by features of the cluster, which are synonymous with the three cartesian axes where x is horizontal, z is vertical and y is perpendicular to the plane of the page. The x axis will be defined as the ‘Fe-Fe axis’. Similarly, the z axis becomes the ‘S-S axis’ and the y axis becomes ‘the normal to the plane of the cluster’. For simplicity, these axes can be named FeFe, SS and Normal. The rotation of the simple electric field can then be defined by rotations about these axes (Figure A3.1).

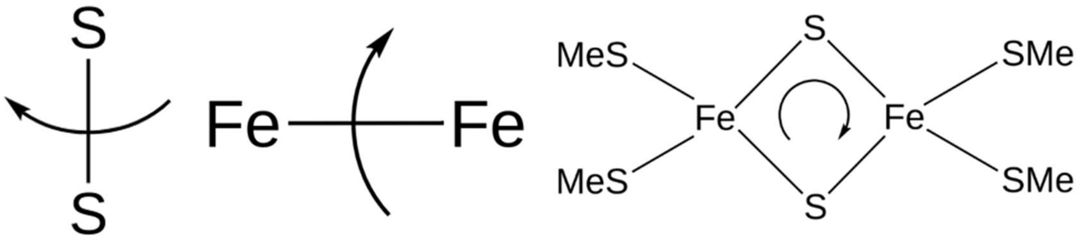


Figure A3.1: Rotations about the SS (left), FeFe (middle) and Normal (right) axes.

Figure A3.2d shows the definition of the angles θ and ψ. If we take the plane of the cluster, $\theta=0^\circ$, increasing ψ moves the field direction around this plane. θ can then be described as the rotation of this plane about the S-S axis and ψ is the angle incremented from the initial position around the normal to this rotated plane. The 2D rotations around the 2Fe2S cluster can then be specified further:

- Rotation about the Normal. The rotation begins in line with the S-S axis. θ is fixed at 0 and ψ is increasing in increments of 10, starting from 0. (Figure A3.2a)
- Rotation about the S-S axis. The rotation begins in line with the Fe-Fe axis. θ is increasing in increments of 10, starting from 0, and ψ is fixed at 90. (Figure A3.2b)
- Rotation about the Fe-Fe axis. The rotation begins in line with the S-S axis. θ is fixed at 90 and ψ is increasing in increments of 10, starting from 0. (Figure A3.2c)

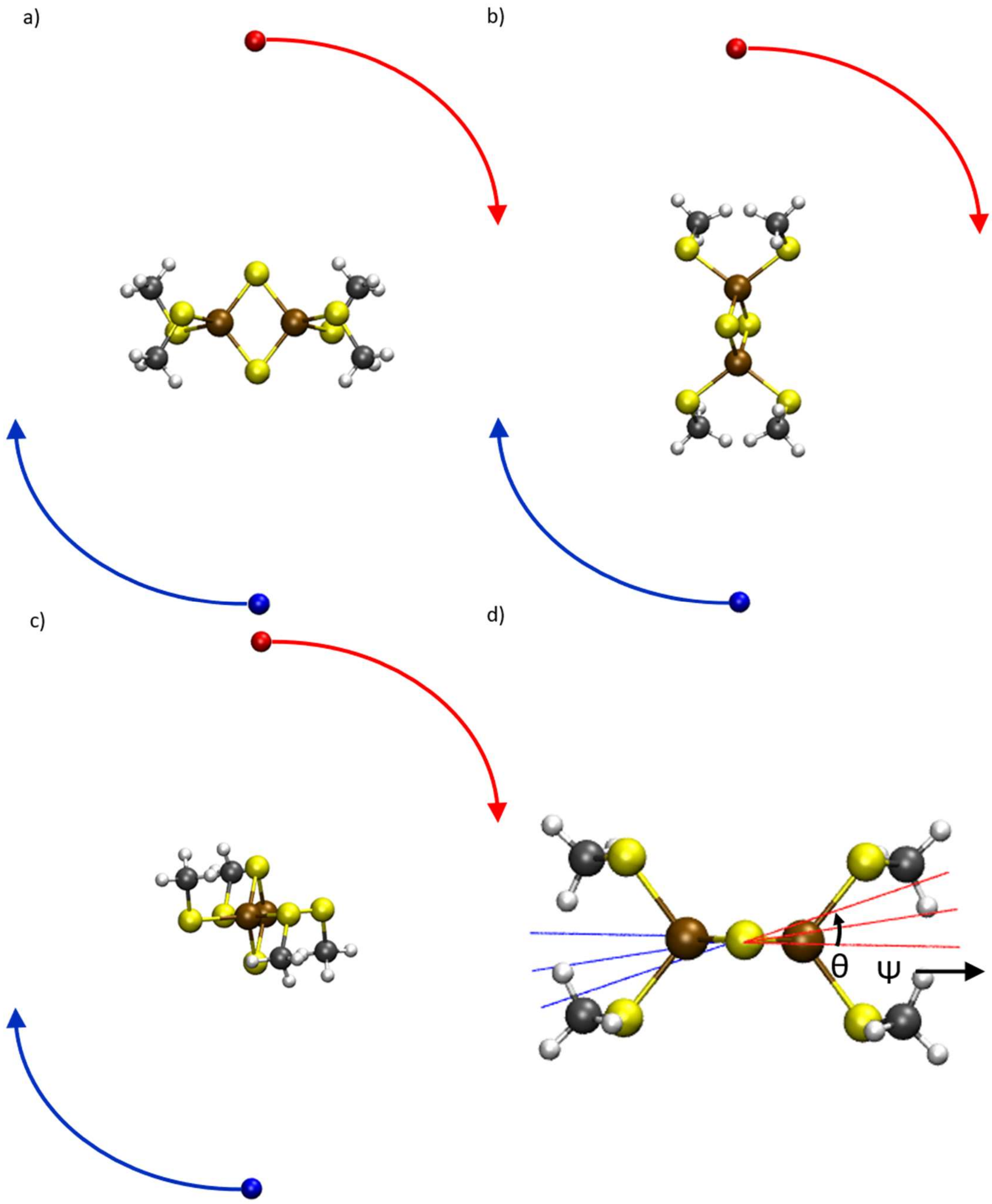


Figure A3.2: Initial position and direction of rotation for the rotation about the Normal (a), the S-S axis (b) and the Fe-Fe axis (c). Effect of changes in the angles θ and ψ shown in d.

Section A3.2 – Iron-Sulfur cluster spin states

Fe has an electronic structure of [Ar]$3d^6 4s^2$ such that Fe(II) would be [Ar]$3d^6$ and Fe(III) would be [Ar]$3d^5$. The iron atoms in our iron-sulfur clusters are in a tetrahedral complex (two sulfur ligands in the cluster and two protein ligands, usually the sulfur in a cysteine) and, as such, the d orbitals will split into two lower energy and three higher energy orbitals. While it is not discussed in the referenced material,[256] it would be prudent to justify the spin states considered in the preceding calculations and to do so requires the knowledge that there is an additional condition affecting the energetic favourability of the spin states in iron sulfur clusters (Figure A3.3) known as antiferromagnetic (AFM) coupling.

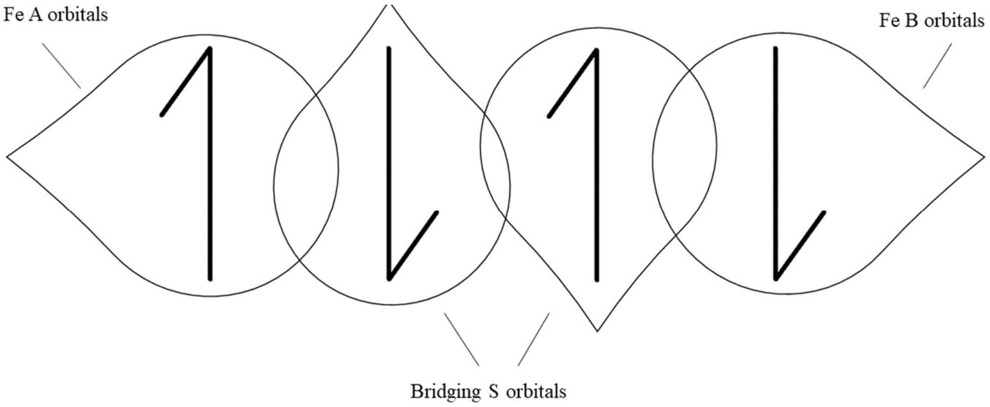


Figure A3.3: Demonstrative sketch of antiferromagnetic coupling where an iron orbital's (Fe A, left) spin induces the opposite spin on the bridging sulfur orbitals and therefore induces the spin of a second iron orbital (Fe B, right) to be the opposite of the first iron orbital.

The spin of the electrons in the orbitals of Fe A (Figure A3.3) can induce the opposite spin in Fe B via influencing the spins of the electrons in the orbitals of the bridging sulfur atom. This does not occur as a rule however, as in some cases it appears that the energy barrier to induce this flip in spins may be too high. In such a case we will refer to this as ferromagnetic (FM) coupling to distinguish it from AFM. Consider the oxidised [2Fe2S] cluster which contains two Fe(III) atoms. These atoms can have their five valence electrons in either a low spin (LS) or high spin (HS) configuration (Figure A3.4).

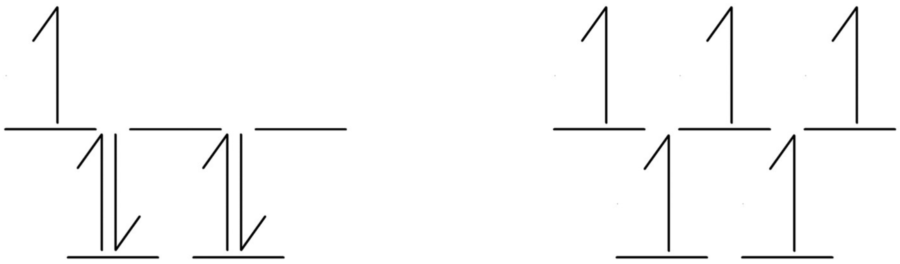


Figure A3.4: Fe(III) low spin (left) and high spin (right) d-orbital configurations.

As we have considered both LS and HS states, we need to consider all possible combinations of these states as well. For the oxidised [2Fe2S] cluster this gives three combinations: LS/LS, LS/HS, and HS/HS. For each of these pairs we also must consider the spin of the unpaired electrons, as this will affect the total mixed cluster spin (mS) reported in the literature[256] and Table 3.1.

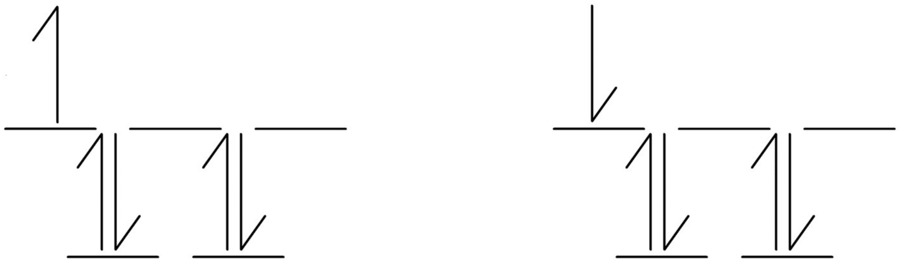


Figure A3.5: Fe(III) low spin ↿ (left) and ⇂ (right) configurations.

Figure A3.5 shows the two possible Fe(III) LS orbital configurations. As such, the LS/LS pair could be constructed from two of the same configurations, resulting in a mS of 1, or one of each configuration, giving a mS of 0. For convenience we can define the left "up" spin configuration as the symbol ↿, and the right "down" configuration as ⇂. Considering the analogous HS configurations (Figure A3.6) and cluster pairings results in a set of paired configurations (Table A3.1).

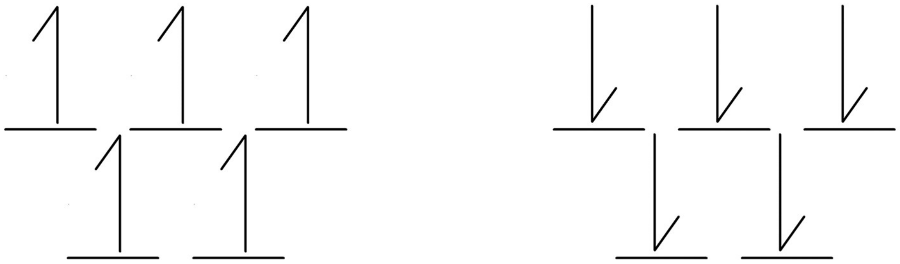


Figure A3.6: Fe(III) high spin ↿ (left) and ⇂ (right) configurations.

Table A3.1: Oxidised [2Fe2S] cluster pair configurations, total cluster mixed spin and a Boolean representing if this spin was considered in the reference material.[256]

Pair Configuration	Total mixed spin, mS	Spin considered?
LS↑/LS↑	1	Y
LS↑/LS↓	0	Y
LS↑/HS↑	3	N
LS↑/HS↓	2	Y
HS↑/HS↑	5	Y
HS↑/HS↓	0	Y

The implication of Table A3.1 is that the energetic favourability of AFM only manifests in the LS/HS pairing in this cluster and not in the LS/LS or HS/HS pairs as we see both mS=1 and mS=5 considered as "possible spin states" despite them being FM pairs, while mS=3, the FM LS/HS pair is completely absent. One possible conclusion is that AFM only occurs as a significant enough energy reduction if there is a different number of unpaired electrons on each iron atom, as is the case when we have a LS/HS pair. To further investigate this, we can consider the same [2Fe2S] in a reduced state. Following the same method as with the oxidised state, while additionally considering the LS and HS reduced configurations (Figure A3.7) gives us an analogous set of paired configurations to that in Table A3.1 (Table A3.2).

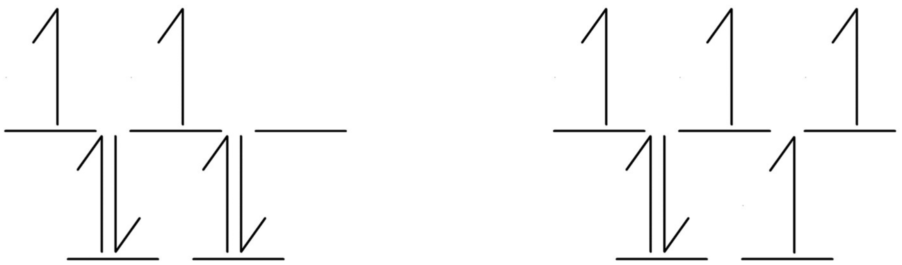


Figure A3.7: Fe(II) low spin (left) and high spin (right) d-orbital configurations.

Table A3.2: Reduced [2Fe2S] cluster pair configurations, total cluster mixed spin and a Boolean representing if this spin was considered in the reference material.[256]

Pair Configuration	Total mixed spin, mS	Spin considered?
$LS_{ox}\uparrow/LS_{red}\uparrow$	3/2	Y
$LS_{ox}\uparrow/LS_{red}\downarrow$	1/2	Y
$LS_{ox}\uparrow/HS_{red}\uparrow$	5/2	Y
$LS_{ox}\uparrow/HS_{red}\downarrow$	3/2	Y
$HS_{ox}\uparrow/ LS_{red}\uparrow$	7/2	N
$HS_{ox}\uparrow/ LS_{red}\downarrow$	3/2	Y
$HS_{ox}\uparrow/ HS_{red}\uparrow$	9/2	Y
$HS_{ox}\uparrow/ HS_{red}\downarrow$	1/2	Y

Note that while we omit pairs of HS/LS in Table A3.1 as they are identical to LS/HS, the situation is somewhat more complex for the reduced cluster as we are considering one Fe(III) (oxidised) and one Fe(II) (reduced) iron atom. However, we do omit any pairings that are simply the "inverse" of an existing pairing, as we did in Table A3.1 for all pairings starting with ↓. For example, $LS_{ox}\uparrow/LS_{red}\downarrow$ and $LS_{ox}\downarrow/LS_{red}\uparrow$ appear to be unique but result in mS=-1/2 and 1/2 respectively. As these clusters would then be indistinguishable, we simply consider them to be identical.

Table A3.2 once again only shows one disallowed pair resulting in a mS of 7/2. However, in the reduced cluster all the pairings have a different number of electrons, so we must refine our previous hypothesis. Consider the pairing which was not present in the literature (Figure A3.8).

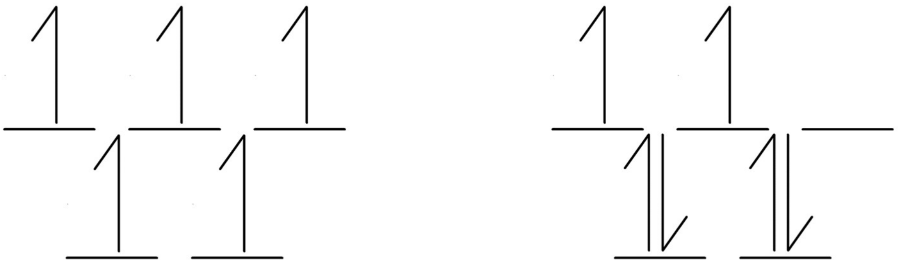


Figure A3.8: $HS_{ox}\uparrow/ LS_{red}\uparrow$ pairing resulting in mS=7/2.

We see in Figure A3.8 that there is a different number of unpaired electrons so we might conclude that this will induce the $LS_{red}\uparrow$ (right) configuration to become $LS_{red}\downarrow$. However only one configuration results in mS=9/2 and this also has a similar phenomenon, a different number of unpaired electrons (Figure A3.9).

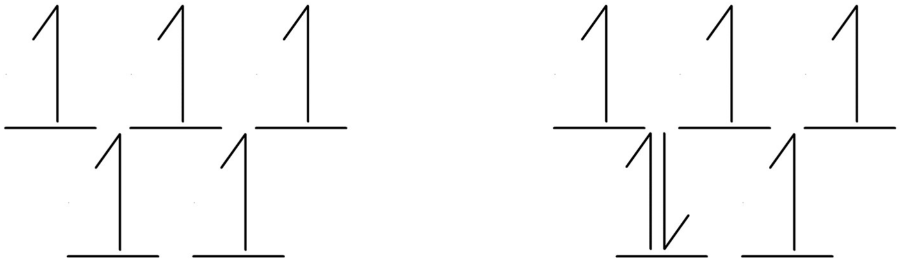


Figure A3.9: $HS_{ox}\uparrow$/ $HS_{red}\uparrow$ pairing resulting in mS=9/2.

As these pairings are the only instances of their respective mS values we know there must be a reason for the first to be disallowed but the second to be allowed. Both pairs contain a $HS_{ox}\uparrow$ configuration containing five unpaired electrons and it appears that this is enough to induce AFM when the paired configuration has two unpaired electrons, but not when it has four. We could therefore conclude that AFM does occur as an energetically favourable condition, but only if the difference in the number of unpaired electrons is at least three (we do not have an example of a difference of two unpaired electrons, this would only occur if the cluster consisted of two Fe(II) atoms). While this is still an unsatisfyingly imprecise justification, it does seem to explain the considered “possible spin states” present in the literature.[256]

Section A3.3 – Nature of the field near the cluster

Single point energy calculations were performed on the system while varying the distance and magnitude of the point charges to verify that the field being generated produces a potential which has a linear dependence on the distance and is unidirectional when near the cluster. The expectation for an electric field generated by two point charges would be that the change in energy of the system caused by the changes to the point charges would be inversely proportional to the distance of the charges, and proportional to the charge squared (if the charges have the same absolute magnitude as they do in our model system) as shown in the relevant equation for electrostatic potential energy U_E, (Eqn A3.1)

$$U_E = k_e\left(\frac{qQ}{r}\right), \qquad \text{(Eqn A3.1)}$$

where k_e is Coulomb's constant, r is the separation between the charges and q and Q are the signed magnitudes of the charges. The calculations were performed on both the reduced and oxidised 2Fe2S cluster where the charges were doubled from (1/8)e to 8e, where e is the charge on the electron in Coulombs, and the distances from the centre of the cluster were incrementally doubled from 10 Å to 640 Å.

Figure A3.10 shows the deviation of the relative energy changes of the cluster from those of the ideal field generated by two point charges. We define the magnitude of the point charges to be Q and the magnitude we used in the main study (1e) to be Q_{REF}. Similarly, the distance of the charges from the centre of mass of the cluster is defined as R, where R_{REF} is the value used in the study (10 Å). For Figures A3.10c and d we take Q_{REF} to be (1/8)e for greater clarity. Based on Eqn A3.1, plotting the ratio of energies at two given values of R against the ratio of the same two values of R should give a straight line at y=1. This is also the case for energy squared when considering the magnitude of the charges. Therefore, plotting one minus this value would be a straight line at y=0 for the ideal case and deviations from this ideal case are readily identifiable. These deviations are very small, at most of the order of 10^{-4}. We can use this result to suggest that this model system creates a faithful representation of a unidirectional field linearly dependent on charge distance to 1 part in 10^4 near to the cluster.

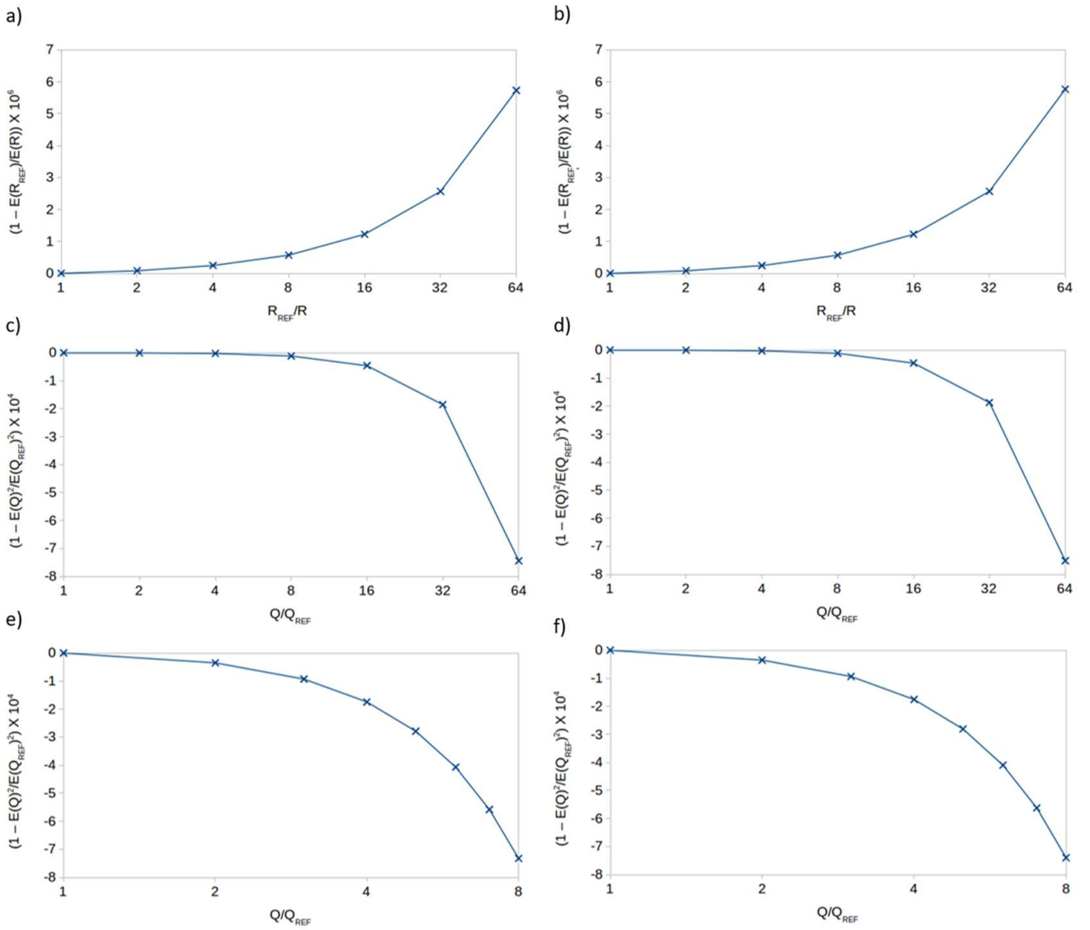


Figure A3.10: a and b show deviation from the expected behaviour given an ideal field generated by two point charges with changing separation, *R*, of the charges. c and d show this for charge, where Q_{REF} is 1/8*e* and e and f show this for Q_{REF} = 1. The left column shows results for the oxidised cluster while the right column shows results for the reduced cluster.

Note that a protein environment will not be as simple as this, but when considering directed changes in the electric field caused by point mutations, a unidirectional field linearly dependent on charge distance should provide a useful analogy.

Section A3.4 – Original Publications

JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

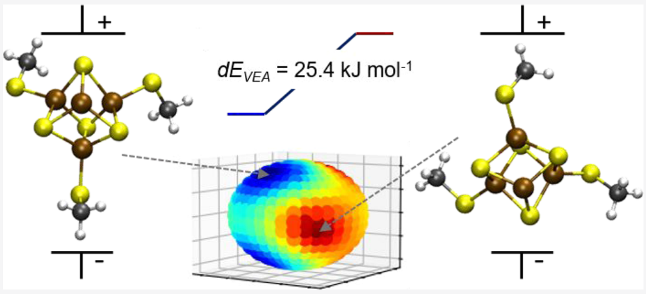
Effect of Oriented Electric Fields on Biologically Relevant Iron–Sulfur Clusters: Tuning Redox Reactivity for Catalysis

Samuel J. H. Gaughan, Jonathan D. Hirst,* Anna K. Croft,* and Christof M. Jäger*

Cite This: *J. Chem. Inf. Model.* 2022, 62, 591−601

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



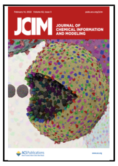
ABSTRACT: Enzyme-based iron−sulfur clusters, exemplified in families such as hydrogenases, nitrogenases, and radical *S*-adenosylmethionine enzymes, feature in many essential biological processes. The functionality of biological iron−sulfur clusters extends beyond simple electron transfer, relying primarily on the redox activity of the clusters, with a remarkable diversity for different enzymes. The active-site structure and the electrostatic environment in which the cluster resides direct this redox reactivity. Oriented electric fields in enzymatic active sites can be significantly strong, and understanding the extent of their effect on iron−sulfur cluster reactivity can inform first steps toward rationally engineering their reactivity. An extensive systematic density functional theory-based screening approach using OPBE/TZP has afforded a simple electric field-effect representation. The results demonstrate that the orientation of an external electric field of strength 28.8 MV cm^{-1} at the center of the cluster can have a significant effect on its relative stability in the order of 35 kJ mol^{-1}. This shows clear implications for the reactivity of iron−sulfur clusters in enzymes. The results also demonstrate that the orientation of the electric field can alter the most stable broken-symmetry state, which further has implications on the directionality of initiated electron-transfer reactions. These insights open the path for manipulating the enzymatic redox reactivity of iron−sulfur cluster-containing enzymes by rationally engineering oriented electric fields within the enzymes.

■ INTRODUCTION

Iron−sulfur clusters play a critical role in reactions catalyzed by several families of enzymes, providing a wide variety of functions in each. Their possible role in enabling the emergence of early life[1] and capacity to perform many different roles within enzymatic pathways[2,3] has led to them being characterized as "one of the most ubiquitous and functionally versatile prosthetic groups in nature".[4] The discovery and purification of ferredoxins in 1962 were an early indication that iron could play roles in enzymes in addition to its well-known presence in hemoproteins.[5] Both iron and iron−sulfur clusters act primarily as mediators for electron transfer with the ability to be either the source or the sink for electrons in redox reactions, and iron−sulfur clusters are extremely useful for electron transport due to the delocalization of the electron density across the cluster.[6,7] The mediator role is one of the most common functions of iron−sulfur clusters found in a variety of enzymes including those that couple proton transfer to electron transport, such as [FeFe] hydrogenases, which possess a unique version of an iron−sulfur cluster featuring a diiron center and a bridging dithiolate.[8,9] Many enzymes containing iron−sulfur clusters use [2Fe2S], [4Fe4S], and [3Fe4S] structures in both redox and nonredox functions. For example, radical *S*-adenosylmethionine (rSAM) enzymes make use of a [4Fe4S] cluster to reductively cleave *S*-adenosylmethionine (SAM) into methionine and the 5′-deoxyadenosyl radical, the latter of which is used to initiate a variety of radical reactions that have been reviewed previously.[10−13] In some cases, electron transfer may be an intermediate step rather than the complete function of an enzyme, such as the case of the biotin synthase rSAM mechanism in which the FeS clusters mediate the donation of a sulfur atom.[14]

Redox activity is an important property in enzymes. Reengineering this property has been the focus of experimental studies for many families of enzymes, including the antioxidant peroxiredoxin,[15] thioredoxins, particularly those that act as electron donors for other enzymes,[16] and kinases,[17] among other more general studies focused on control via thiol/disulfide exchange,[18,19] enzyme orientation,[20] and substrate specificity.[21] Additionally, redox activity has been used to probe the mechanism,[22−26] to understand the effect of changing iron−sulfur cluster ligands on the redox potential,[27] to assess the stability and reactivity of the cluster,[28] and to

Received: July 6, 2021
Published: January 19, 2022



https://doi.org/10.1021/acs.jcim.1c00791
J. Chem. Inf. Model. 2022, 62, 591−601

guide improvements in enzyme activity with directed evolution.[29]

Redox reactions, often studied and used via electrochemical methods,[30−32] are sensitive to electric fields. This alone would motivate an analysis of the electrostatic environment generated by an enzyme containing an iron−sulfur cluster. However, electrostatics also play a role in enzyme catalysis in general,[33] including protein−protein interactions,[34] conformational motions,[35] and catalysis.[36,37] Electrostatic preorganization in the active site of enzymes and the electrostatic stabilization associated with this are a more recent area of study.[38] This preorganization has also been quantified computationally.[39,40] Computational methods such as molecular dynamics simulations, density functional theory (DFT), valence bond theory, and Poisson−Boltzmann equation solvers have advanced our understanding of the role of electrostatics in enzyme catalysis,[41−43] providing a more complete picture of the function of enzymes such as α-amylase,[44] methyltransferases,[45] and QueE.[46] These methods have also been used to study specific properties such as the contribution of individual amino acids to the overall electrostatic field of a protein,[47] electrostatic steering and channeling,[48,49] and the direct effect of the electrostatic field on the catalytic rate.[50]

The direct study of electric fields in the context of manipulating catalysis is a growing area of research. A recent review highlighted oriented electric fields as reagents, as well as their effects on enzyme catalysis.[51] DFT as the quantum mechanics (QM) method in a QM/molecular mechanics (MM) approach has been used previously to study the effects of an electric field on biological chromophores[52] and on enzymes.[53] Studies have also been performed on iron−sulfur clusters in the context of enzymes to elucidate information about properties such as coordination, geometry, and electrostatics. These studies have used extended X-ray fine structure spectroscopy and DFT,[54] including the application of broken-symmetry DFT[55] to iron−sulfur clusters to model antiferromagnetic coupling.

In this study, we directly and systematically investigate the impact of an external electric field on biologically relevant iron−sulfur clusters. We examine the fundamental effect of applying and reorienting a simple electric field on the stability and reactivity of selected model iron−sulfur clusters. The information on how an oriented electric field influences the reactivity of the clusters can later be combined with the knowledge and understanding of the contribution of individual amino acids to the electric field in an enzyme. This will allow the bespoke control of iron−sulfur cluster reactivity and stability through mutations of the surrounding residues.[47] Consequences of these adaptations would include either the ability to improve the existing reactions by increasing the rate of reaction, integrating oxygen tolerance, or exploiting enzyme selectivity for reactions that were previously only accessible through synthetic approaches.

■ METHODS

Unrestricted geometry optimization calculations using DFT were performed using the Q-Chem software package[56] and compared with literature values[57] using the same model systems featuring the iron−sulfur clusters [2Fe2S] or [4Fe4S] and four methanethiolate ligands bound to the iron atoms, two per iron or one per iron, respectively (Figure 1). The geometry of a third cluster, identical to the [4Fe4S] cluster but with one fewer methanethiolate ligands, was also optimized. The initial geometries were based on the crystal structure (PDB ID 1ZOY.) The hybrid optimized Perdew-Burke-Ernzerhof (OPBE) functional,[58,59] which consists of Handy's optimized exchange (OPTX)[60] and PBE[61] correlation, was used for both the geometry optimizations and further single-point calculations. The use of this functional allows comparison with the work of Carvalho and Swart,[57] and the spin states of the iron complexes predicted are consistent with the literature studies.[58,59] The TZP basis set[62−64] was chosen for both geometry optimizations and single-point calculations to facilitate this comparison. The methods used by other groups include B(5%HF)P86 and a triple-ζ basis set with polarization functions for accurate bond covalency[65,66] and B(5%HF)P86/6-311+G(d) for the QM region of a QM/MM study into the protein environmental effects around the iron−sulfur clusters.[67] While these levels of theory would be suitable for optimizations, we chose OPBE/TZP because it demonstrates both the correct structural predictions and accuracy in the relative ranking of spin-state energies, which is particularly important for this study.[68,69] The geometry optimizations were also repeated using the polarizable continuum model COSMO[70] (conductor-like screening model) with a dielectric constant of 4.0 to simulate a protein environment.[71,72]

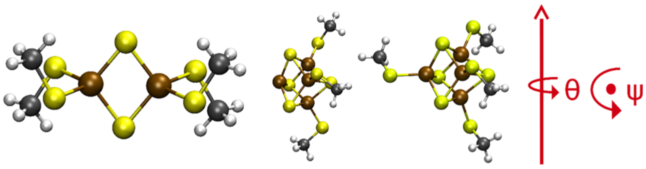


Figure 1. Definition of axes for the [2Fe2S] (left), labile [4Fe4S] (middle), and [4Fe4S] (right) clusters. All geometries shown are optimized structures for the oxidized clusters at the OPBE/TZP level of theory in a vacuum.

We investigate the effect of an oriented external electric field on the vertical electron affinity (VEA) and the most stable state of the model systems by performing 1296 single-point calculations in the presence of two external, equal but oppositely signed point charges, equidistant from the center of mass of the system, whose bisector intersects this center. The redox potential is an important property to consider when investigating the reactivity of iron−sulfur clusters. It is dependent on the stability difference of the oxidized and reduced states of the system and the reorganization energy. The latter is influenced by the relaxation of the system upon electron transfer and the restructuring of the environment, often dominated by solvent reorganization. The VEA accounts for the energy difference between the oxidized and reduced states without any relaxation (eq 1) and can be taken as a first approximation for the reactivity difference of the two oxidation states.

$$E_{\mathrm{VEA}} = E_{\mathrm{ox}} - E_{\mathrm{red}} \tag{1}$$

where E_{VEA} is the VEA and is positive if the reduction of the oxidized state is energetically favorable, E_{ox} is the total electronic energy of the oxidized state, and E_{red} is the corresponding energy of the reduced cluster with the oxidized cluster's geometry.

The effect of the direction of the external electric field was investigated by rotating point charges defining the field about two axes in the system and mapping them onto a sphere around the model systems. The axes of rotation used to orient the field around the clusters are shown and labeled in Figure 1.

The notation is described further in Supporting Information Section S2.

The rotation of the electric field can be defined by the rotations about these axes. Supporting Information Section S2 provides examples of how the point charges are positioned initially and their directions of movement for all rotations applied.

Systems containing atoms with multiple possible oxidation and spin states can be described in several ways. The reduced state of the [2Fe2S] cluster introduced earlier could have an oxidation state of +2.5 assigned to both iron atoms, a symmetrical and simplified description. In many cases, however, it is necessary to define broken-symmetry oxidation or spin states, where the additional electron might be fully localized on one of the iron atoms, for systems with ferromagnetic and antiferromagnetic coupling, for example.[73,74] This would be particularly important if a specific iron atom is involved in a reaction as the oxidation state would affect the reactivity or govern whether the reaction would progress at all.

However, it needs to be noted that Kohn−Sham DFT methods in general have limitations in describing multireference electronic systems such as the presented model iron−sulfur clusters, and we cannot easily assign formal oxidation states directly for them. To qualitatively verify the results obtained from the broken-symmetry approach, we can however conduct modified calculations to elucidate useful information and obtain electronic states that are as close as possible in energy to the true pure spin ground state. Here, restricted open-shell DFT enforces a specific spin state in situations that would otherwise require an unrestricted calculation, such as the reduced [2Fe2S] cluster which has one unpaired electron. This will result in a pure spin state that, however, will be higher in energy than the result from an unrestricted calculation on the same system. Broken-symmetry DFT (BS-DFT)[55] can also be used to investigate the different broken-symmetry states for a single spin state. Using the broken-symmetry states and the maximum-spin pure spin state, we can also calculate the Heisenberg coupling constant J which can provide a direct comparison to experimental results. This method is detailed in Section S6 of the Supporting Information.

There are two broken-symmetry states for the 1/2 spin system, where the electron density of the unpaired electron is localized to one of the iron atoms, analogous to the system where the oxidation states of the iron atoms of the [2Fe2S] cluster are +2/+3 and +3/+2. A single-point calculation using a default guess wave function for the 1/2 mixed spin reduced cluster using one of the point charge positions will give one of the two broken-symmetry states. By taking the final wave function from each of these situations and using them as guess wave functions in two calculations without point charges but preserving the orbital occupancies, we calculated two wave functions, one for each broken-symmetry state, which do not have orbitals influenced by an external field. We used these wave functions as initial guesses when we added point charges, allowing us to see the full range of field effects for both broken-symmetry states. Geometry optimizations and subsequent frequency calculations were performed in vacuum and using COSMO with a dielectric of 4.0 for the model [2Fe2S], [4Fe4S], and labile 4Fe4S iron−sulfur clusters for the four reported mixed spin states in each case.[57]

■ RESULTS AND DISCUSSION

The energy relative to the calculated ground state for each system is reported in Table 1. All optimized geometries that were used in later calculations were confirmed as minima by frequency calculations. The only case where the ground state is not the low spin state is from the reference value for the reduced [2Fe2S] cluster; this is not surprising as the 1/2 and 9/2 spin states are isoenergetic. The geometry and frequency calculations for each spin state agree with the literature values[57] within 5 kJ mol^{-1} (Table 1) and are consistent in the relative stability of each state and can therefore be justified for use in the subsequent investigations for the purpose of determining relative stability and reactivity when a directed external electric field is applied. The $\langle S^2 \rangle$ values for these optimizations are reported in Table S1 of the Supporting Information. The optimizations in a COSMO environment were generally lower in energy but showed the same trends in stability as the vacuum environment. This is to be expected as prior work found that both PCM and COSMO approaches converge to very similar structures as the gas-phase optimization, although there were some exceptions.[75] COSMO has also been used in the successful prediction of Mössbauer spectral parameters,[76] reduction potentials,[77] and other properties of iron−sulfur clusters and similar molecules.[78] Furthermore, experimental work has shown that the

Table 1. Relative Energies for OPBE/TZP Geometry-Optimized-Model Iron−Sulfur Clusters in Different Spin States, Compared with the Literature Values[57] in Both Vacuum and within a COSMO Environment Using a Dielectric of 4.0[a,b]

2Fe2S	**mixed spin, mS**			
oxidized	**mS = 0**	**mS = 1**	**mS = 2**	**mS = 5**
reference	0.00	54.34	104.21	72.64
vacuum	0.00	65.93	110.44	84.06
COSMO	0.00	57.44	100.68	83.69
reduced	**mS = 1/2**	**mS = 3/2**	**mS = 5/2**	**mS = 9/2**
reference	0.67	84.11	100.57	0.00
vacuum	0.00	92.64	98.74	7.47
COSMO	0.00	84.24	91.71	4.16
4Fe4S	**mixed spin, mS**			
oxidized	**mS = 0**	**mS = 1**	**mS = 2**	**mS = 9**
reference	0.00	48.61	82.94	156.38
vacuum	0.00	45.99	81.33	177.29
COSMO	0.00	44.85	79.85	181.25
reduced	**mS = 1/2**	**mS = 3/2**	**mS = 5/2**	**mS = 17/2**
reference	0.00	62.09	66.15	63.97
vacuum	0.00	59.64	66.13	74.39
COSMO	0.00	50.65*	61.11	77.45
labile 4Fe4S	**mixed spin, mS**			
oxidized	**mS = 0**	**mS = 1**	**mS = 2**	**mS = 9**
vacuum	0.00	58.45	53.92	160.14
COSMO	0.00	56.53	56.28	163.19
reduced	**mS = 1/2**	**mS = 3/2**	**mS = 5/2**	**mS = 17/2**
vacuum	0.00	64.99	45.80	66.36
COSMO	0.00	56.28	46.79	58.52

[a] Values presented in kJ mol^{-1}. All values correspond to geometry-optimized structures at the given oxidation and mixed spin state. Absolute energy values are presented in the Supporting Information (Table S1). [b] Optimized state with one imaginary frequency. This state was not used in any further calculations.

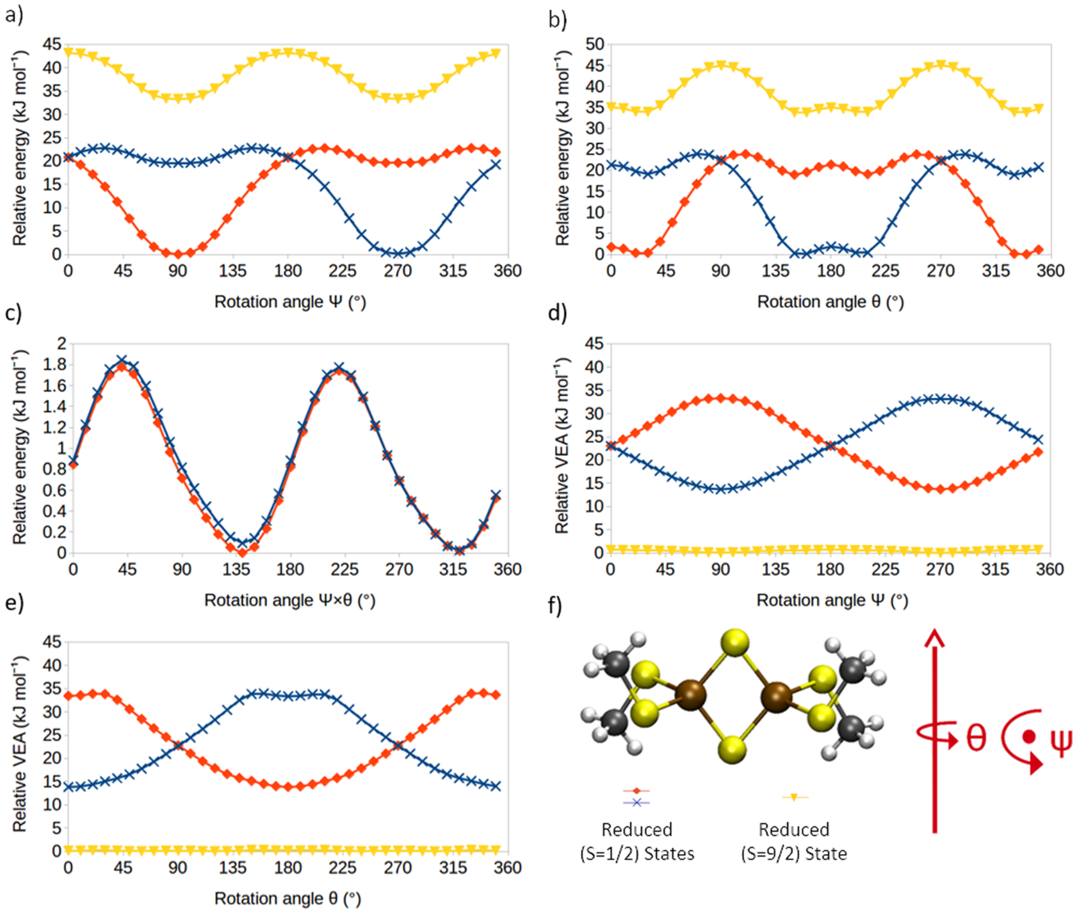


Figure 2. Relative energy profiles of the reduced [2Fe2S] model cluster (a−c) and the VEA (d,e) of the same cluster with the orientation of an external electric field in the gas phase. Rotation angle indicates the rotation from the initial position of the field about the axis normal to the plane of the cluster (a,d), the S−S axis (b,e), or the Fe−Fe axis (c) [$\psi \times \theta$ represents the rotation around the axis perpendicular to those shown in (f)].

redox potential of biological iron−sulfur clusters is significantly dependent on the environment when considering different enzymes[79] or modifying the ligands of the cluster.[28] Optimizations on the labile [4Fe4S] cluster can also be justified using the same reasoning and could be used for further study of a similar nature when investigating the behavior of the cubane cluster where one iron atom is not bound to a ligand (e.g., SAM in the radical SAM enzyme superfamily). Using VEA as a surrogate for the redox potential is not a new approach but offers an approximation with the potential for high throughput needed to systematically study an oriented electric field effect. Previous work has gone further to approximate redox potentials from DFT calculations by considering electronegativity,[80] electrophilicity,[81−83] and combinations of properties.[84,85] While these methods would provide more experimentally comparable values, the VEA provides a sufficient description of the change in these properties and is straightforward to compute. This makes it feasible to investigate the effect of a rotating electric field via thousands of individual single-point DFT calculations.

The relative energies are in good agreement with the literature values obtained for these systems with the order of spin-state stability preserved. Slight differences in the energies compared to the literature are likely to originate from different implementations of the implicit solvent method (COSMO) and different convergence criteria in the used programs but do not influence the qualitative agreement. Single-point calculations were performed on the model iron−sulfur clusters in the presence of an electric field of varying orientation for the oxidized and reduced states. The optimized geometry of the ground-state spin oxidized cluster was used in all cases to allow the calculation of the VEA. Figure 2 displays the effect of the rotation of the electric field on the energy of the two most stable 9/2 and 1/2 spin states of the reduced [2Fe2S] model cluster about the principal axes presented in Figure 1.

As can be seen in Figure 2, reorienting a directed electric field can influence the electronic stability of the [2Fe2S] cluster significantly. Depending on the axis of rotation, the effect on the stability can vary up to 23 kJ mol^{-1} (in case of rotating around the normal to the plane of the cluster, Figure 2a,b,e) and down to only 1.8 kJ mol^{-1} (in case of rotating around the Fe-Fe axis, Figure 2d). Each spin state is influenced differently in relation to the field orientation, which leads to a varying energy gap between the spin states. However, this variation is never large enough to lead to the 9/2 spin state being the ground state (Figure 2a−c). When comparing the electric field effect between the gas phase (Figure 2a) and an implicit solvation with a low dielectric constant of 4.0 (Figure 2b), the observed effect is maintained, but the energy gap between the two spin states decreases slightly. Further results with implicit solvation are presented in Supporting Information Section S5. The $\langle S^2 \rangle$ values for the reduced [2Fe2S] cluster mS = 1/2 broken-symmetry states are also discussed in Section S1 and Figure S1 of the Supporting Information,

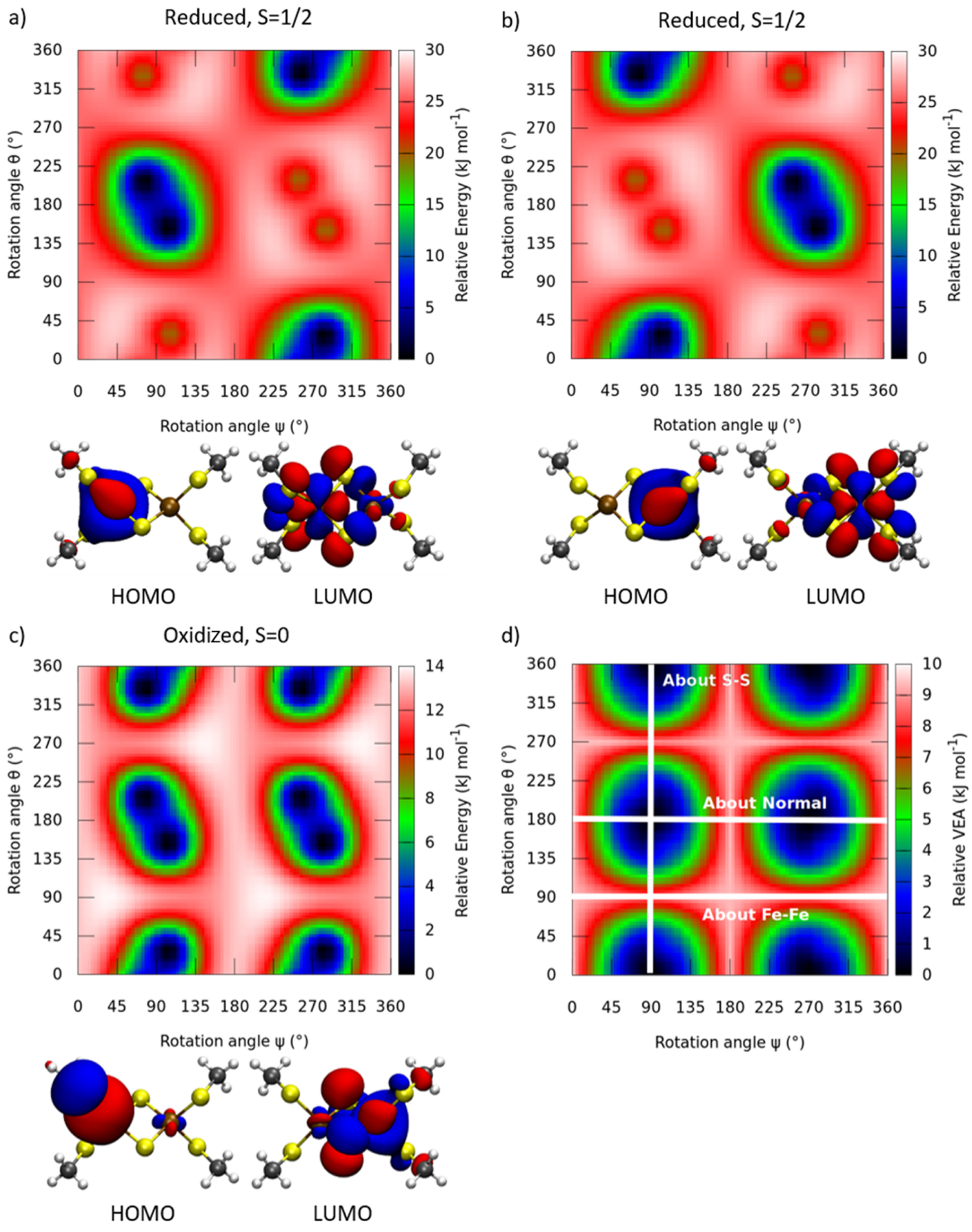


Figure 3. Frontier orbitals and effect of the three-dimensional rotation of an oriented electric field around the [2Fe2S] cluster. (a) Reduced, $S = 1/2$, broken-symmetry state 1; (b) reduced, $S = 1/2$, broken-symmetry state 2; (c) oxidized, mS = 0; (d) relative VEA with the most stable reduced broken-symmetry state.

demonstrating only minor changes in the $\langle S^2 \rangle$ upon the rotation of the electric field around the cluster.

Next, it was investigated how the electric field influences the stability of different broken-symmetry states, which showed a significant effect on the relative stability of those. When rotating the field around the normal to the plane of the [2Fe2S] cluster (Figure 2a,b), the most stable spin state flips when the field is exactly aligned along the S−S axis. This means that depending on the orientation of an external electric field, the cluster adapts by adopting a different broken-symmetry state. As these different states will show different reactivities in the directed electron transfer, notably in relation to redox potential,[71] they will also show different reactivity patterns with, for example, different substrates reacting with the cluster in an enzymatic active site. The rotation about the Fe−Fe axis is diagrammatically presented in the Supporting Information (Figure S1). There is very little variation in VEA, with the value for the 1/2 spin states remaining between about 22 and 23 kJ mol^{-1} and the 9/2 spin state between 0 and 1 kJ mol^{-1}.

The relationship between the stability of the oxidized and reduced clusters and the changes to the redox reactivity of the clusters are inferred from the VEA calculations for the cluster from the data presented above. Figure 2e depicts how the VEA varies with rotation about the normal axis. Depending on the rotation axis, ΔE_{VEA} varies by up to 20 kJ mol^{-1}. The results also suggest that the 9/2 spin state of the reduced cluster is unlikely to be involved in redox reactions as it is both less stable and has a lower electron affinity than both 1/2 spin states.

To extend this analysis of the principal axes of the [2Fe2S] cluster, we systematically scanned the orientation of the

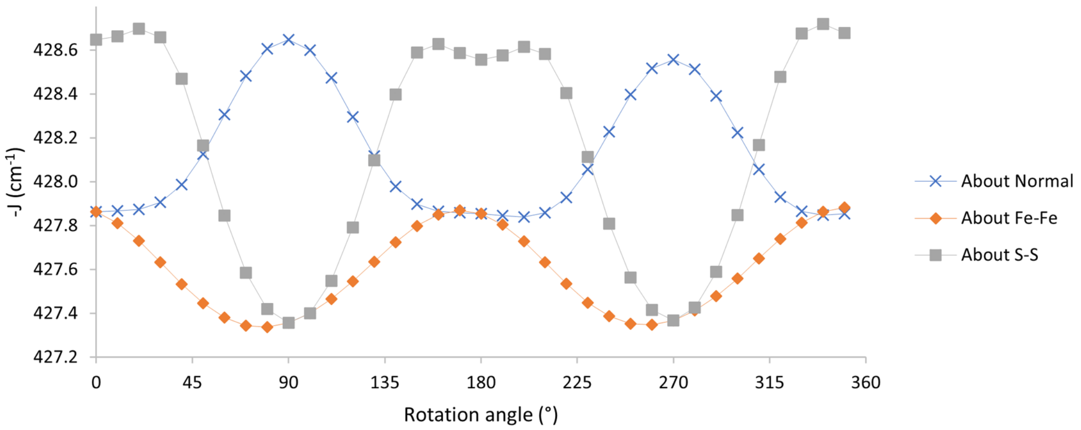


Figure 4. Effect of a rotation electric field around the three principal axes of the oxidized [2Fe2S] model cluster on the Heisenberg coupling constant *J*.

electric field about multiple axes to create a three-dimensional visualization. Figure 3 depicts the analysis for the [2Fe2S] cluster. To understand why a differently oriented electric field influences the stability and reactivity of the [2Fe2S], we then matched these observations to the symmetry of the frontier orbitals of the reactive species.

The stability pattern observed for the principal axes' rotations described above is reflected in the analysis of the full rotations. When looking at the VEA, some electric field orientations show significantly higher reactivity compared to others. Considering the individual stabilities of the oxidized and reduced states, one can further see the relation with the frontier orbital occupancies, as also depicted in Figure 3.

Looking at the differences for the two reduced broken-symmetry spin states, the effect of the directed electric field is mirror-symmetric for each state. The cluster is stabilized when the side with the higher electron density of the highest occupied molecular orbital (HOMO) is close to the positive charge defining the electric field. Due to the different orbital occupancies of the oxidized state (fully occupied HOMO), this species shows a higher symmetry in reference to the effect of the electric field.

To validate the applicability and quality of the applied broken-symmetry approach for our investigations, we have also calculated the pure spin state energies and have derived the Heisenberg coupling constant *J* of this system when rotating an electric field around the principal axes detailed above and as such found the effect of the electric field on the spin coupling constant *J* (Figure 4). A detailed description of the methodology is presented in Section S6 of the Supporting Information. Overall, this analysis confirms that the profile of the energy change with a rotating electric field is the same for a pure spin state and a broken-symmetry state for these systems, with the pure spin state higher in energy. While the effect on the oxidized [2Fe2S] cluster (see Figure 4) only has a range of about 1.5 cm^{-1}, it is noticeable that the magnitude of the effect is similar for all rotation axes, while the energy changes for the Fe−Fe axis rotation were an order of magnitude smaller than the other axes. As such, the changes in energy around this axis would imply that very little about the system changes compared to other axes of rotation, but there are other quantities that will still be affected by the field direction around this axis.

The value of *J* extracted using this method lies within the range of values reported in the literature for the same or very similar systems, noting that there seems to be a common overestimate of *J* from the theoretical calculations compared to experimentally derived values.[7,73,86,87]

The cubane cluster displays more complex features. While the [2Fe2S] cluster showed clear symmetry for the oxidized state and a clear asymmetry related to the positioning of the unpaired electron density in the reduced states, the [4Fe4S] cluster appears to display similar patterns of stability regardless of oxidation or the broken-symmetry state. The magnitude of these patterns varies, however. The relative energy of the cubane clusters with respect to the rotation of the field is influenced by the location of the HOMO seen in Figures 5 and 6. The significant change in the location of the electron density between these two broken-symmetry states results in the quadrants containing the maxima and minima being reversed. The directions of the field that result in the greatest stability remain broadly the same between these and the oxidized state. This suggests that the external electric field has a lower polarizing effect on the [4Fe4S] cluster compared to the [2Fe2S] cluster and therefore a much lower effect on stability. Any of the four rotations where ψ, θ = 90°, 270° show the largest degree of variation in relative energy. These rotations lie along two perpendicular internal quadrilateral planes of the cluster that are coincident with its edges. The maximum VEA difference observed for the electric field is ΔE_{VEA} = 25.4 kJ mol^{-1}.

The labile cubane cluster also shows only a difference in the magnitude rather than the location of the lowest energy orientation. There is a much less complex pattern for this molecule, and noting that the points at which the field is aligned along a vector passing through the edge created by the labile iron and a sulfur atom have the highest VEA and a perpendicular vector has the lowest VEA suggests that the effect of this labile iron atom is dominant. Here, the maximal influence of the electric field on the VEA is not along an axis involving the unique iron. The maximum electric field effect on the VEA observed for this cluster is ΔE_{VEA} = 20.90 kJ mol^{-1}. Previous work on similar[88] and related[89] structures shows that the values for the changes in VEA are both significant relative to the absolute VEA but not unreasonably high in magnitude. Notably, these VEA values fall within the 0−0.51 eV (0−49 kJ mol^{-1}) range of magnitudes predicted when varying the surrounding and connected amino acids around the iron−sulfur clusters.[67] It is also of interest to note that the field vectors resulting in the maximum and minimum values of

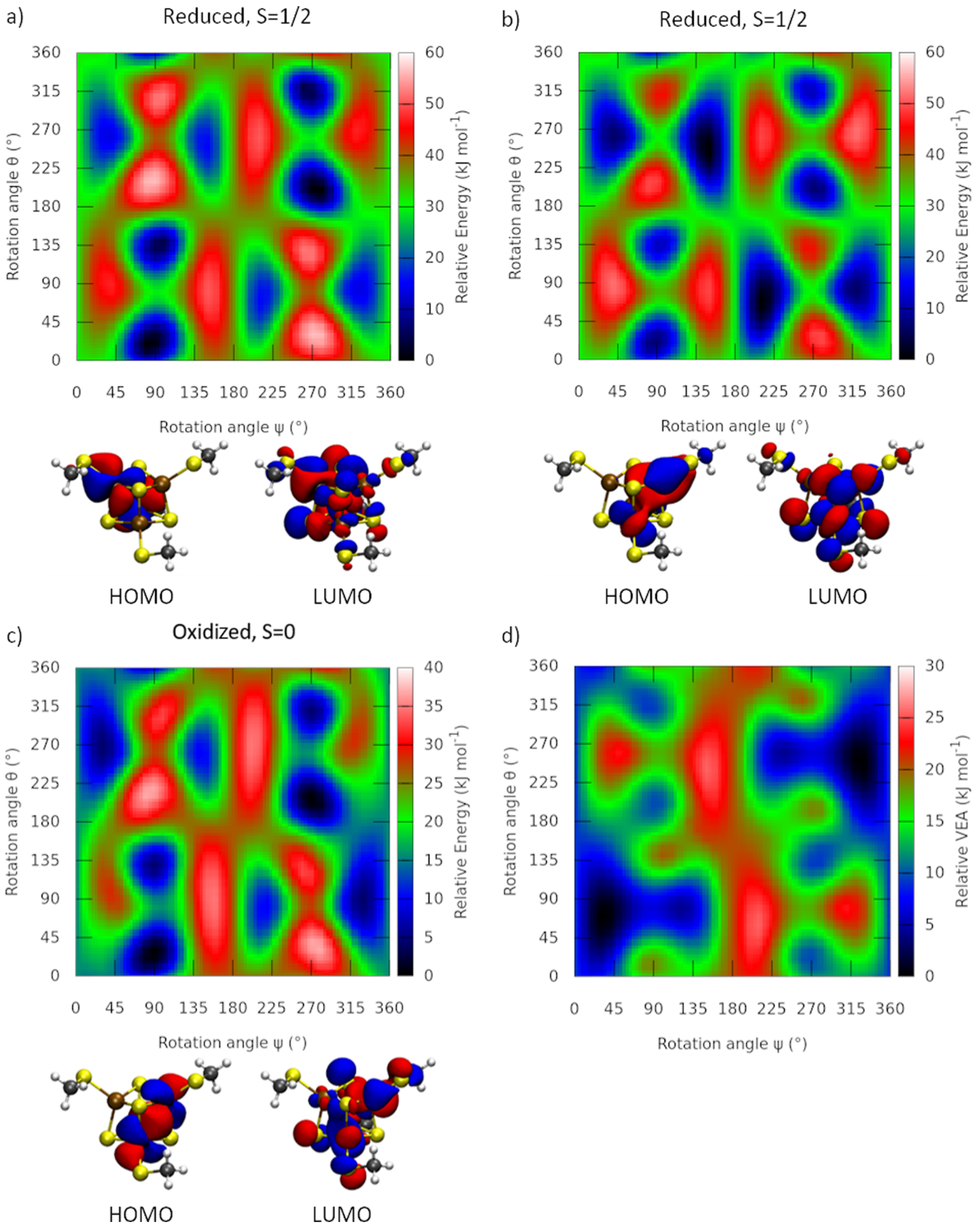


Figure 5. Frontier orbitals and effect of the three-dimensional rotation of an oriented electric field around the [4Fe4S] cluster. (a) Reduced, $S = 1/2$, broken-symmetry state 3; (b) reduced, $S = 1/2$, broken-symmetry state 4; (c) oxidized, mS = 0, broken-symmetry state 4; (d) relative VEA with the most stable reduced broken-symmetry state.

relative VEA are the same for both the [4Fe4S] and labile [4Fe4S] clusters, specifically approximately passing through a methyl group, the bonded sulfur and iron, and finally through the center of the cube and through the opposite sulfur atom. Since the maxima and minima observed for the VEA also influence electron-transfer reactions, it is not surprising that the observed field vectors (for maxima and minima) also follow the direction of the initial reductive electron transfer necessary for the formation of the 5′-deoxyadenosyl radical in radical SAM enzymes.[90] As rSAM enzymes share the common feature of SAM bound to the cluster, this demonstrates how those enzymes may have evolved to arrange the active site to follow the minimum-energy pathway for this initial activation step for the enzymes. Additional electrostatic field effects initiated by other charged residues nearby can either support or reduce this effect.

■ CONCLUSIONS

We systematically investigated the effect of a rotating directed electric field on the stability and reactivity of biologically relevant iron−sulfur clusters. Applying DFT calculations on the model clusters, we used the VEA—represented as the energy difference between the oxidized and reduced states—as an indicator for the redox reactivity of the clusters. In both cases of the pure gas phase and calculations in implicit solvent, we could show that a directed electric field induced by two distant point charges significantly influences the stability and reactivity of the clusters. In agreement with molecular frontier orbital theory, the effects are significantly different for different

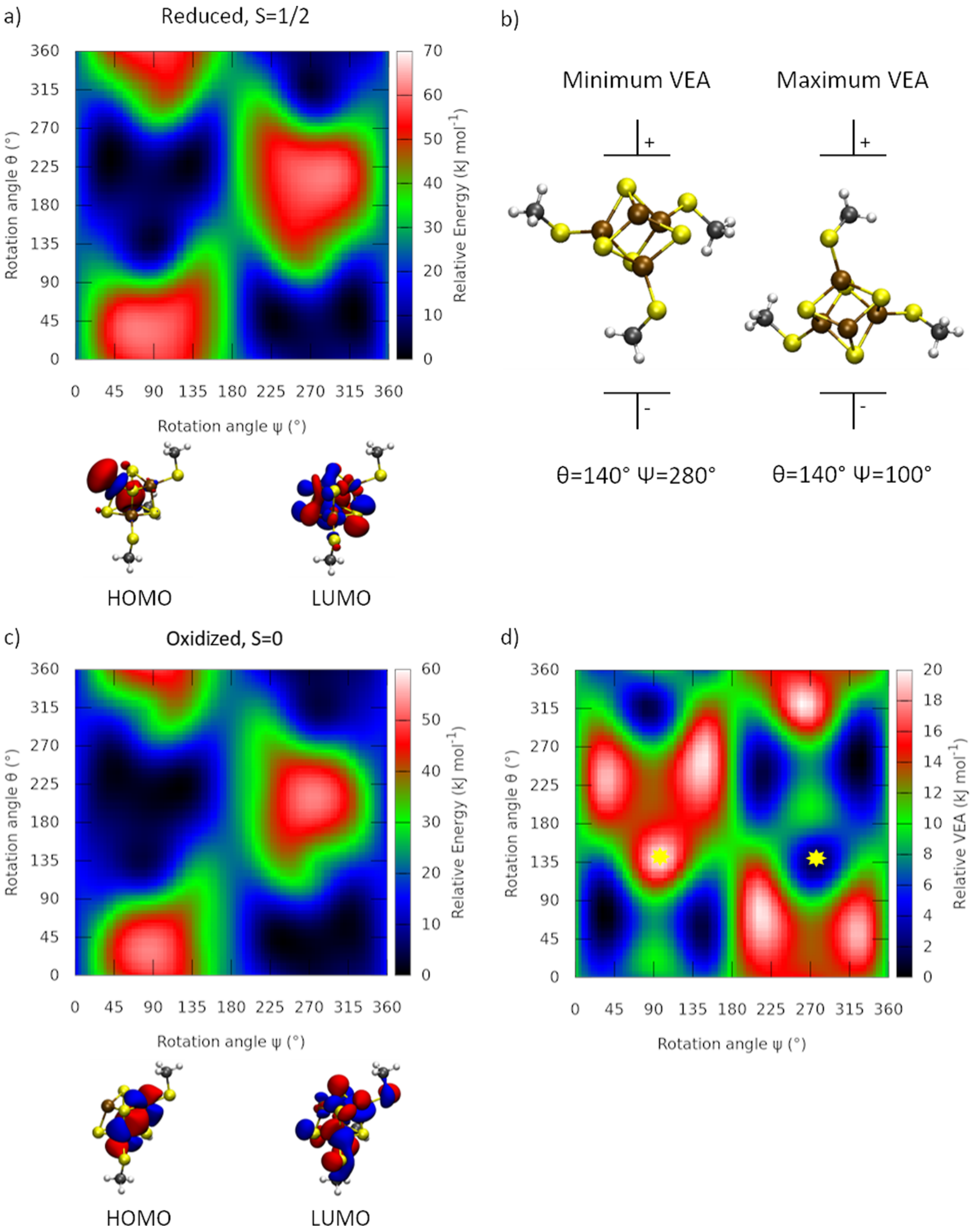


Figure 6. Frontier orbitals and effect of the three-dimensional rotation of an oriented electric field around the labile [4Fe4S] cluster. (a) Reduced, *S* = 1/2; (b) exemplary orientation of the field for the maximum and minimum values of the VEA, indicated in (d) by a yellow asterisk; (c) oxidized, mS = 0; (d) vertical electron affinity.

oxidation and broken-symmetry states. This leads to the observation that the most stable broken-symmetry state changes when reorientating an electric field around the [2Fe2S] cluster with significant implication for the reactivity of such clusters in heterogeneous environments such as enzyme active sites. Different orientations of the electric field had significantly different influences on the VEA and thus the reactivity of the clusters with a maximum effect of up to 25.5 kJ mol.$^{-1}$

The cubane clusters display more complex behavior, with it still being clear that the external electric field has a dramatic effect on the stability and reactivity of these clusters. The introduction of the COSMO environment alters the stability of the reduced cluster, stabilizing the favored broken-symmetry 1/2 spin state relative to the other broken-symmetry state but also stabilizing the 9/2 state relative to the most favored state, suggesting that in a protein environment, this stabilization effect is strengthened. The effects on the VEA are similar and in both cases dependent largely on whether the iron atoms experience unique force-field environments or similar ones, as shown in Figure 2D, where the field is rotated about the Fe—Fe axis and the changes in orientation are therefore the same for both iron atoms. Comparing the field with the orbital occupancy of the HOMO and LUMO of the clusters reinforces the hypothesis that the spatial location of electron density contributes significantly to the observed effects.

These strong anisotropic effects of the electric fields demonstrate their relevance for iron−sulfur clusters embedded in heterogeneous enzymatic environments. The impact of the field might explain how enzymes influence the redox reactivity of such clusters in a significant way. This systematic study reveals how these effects might be rationally explored and used as a powerful tool for enzyme engineering, where the external electric field would be generated by the surrounding residues, which can be mutated for the direct control of the properties of the iron−sulfur clusters.

While the model systems presented cannot reflect the complexity of the anisotropic charge distributions inside enzymatic active sites, the presented field strengths and strong directionality of the electric fields are comparable.[51] Thus, the model systems offer a good reference on how changing the directionality of electric fields inside the enzymatic active sites can influence catalysis. In the next step, this information can be used for rational mutation studies with possible applications that might alter catalytic rates, substrate scope, and cluster stability (e.g., in oxygen-sensitive enzymes). Further work will focus on the nature and magnitude of these possible applications and how generalizable this approach is to other similar molecular groups in enzymes.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.1c00791.

> Further details on methods applied, detailed setup of the electric field calculations, evaluation of the electric field uniformity, and additional details of the results (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Jonathan D. Hirst − *School of Chemistry, University of Nottingham, Nottingham NG7 2RD, U.K.*; orcid.org/0000-0002-2726-0983; Email: jonathan.hirst@nottingham.ac.uk

Anna K. Croft − *Department of Chemical and Environmental Engineering, Faculty of Engineering, University of Nottingham, Nottingham NG7 2RD, U.K.*; orcid.org/0000-0001-5330-150X; Email: anna.croft@nottingham.ac.uk

Christof M. Jäger − *Department of Chemical and Environmental Engineering, Faculty of Engineering, University of Nottingham, Nottingham NG7 2RD, U.K.*; orcid.org/0000-0002-1802-1892; Email: christof.jaeger@nottingham.ac.uk

Author

Samuel J. H. Gaughan − *School of Chemistry and Department of Chemical and Environmental Engineering, Faculty of Engineering, University of Nottingham, Nottingham NG7 2RD, U.K.*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.1c00791

Notes

The authors declare no competing financial interest.

Structural coordinates, QM calculation input and truncated output files, and python scripts used for the automated procedures and analysis can be found in the online repository figshare (https://doi.org/10.6084/m9.figshare.14883615.v2).

■ ACKNOWLEDGMENTS

The authors thank the EPSRC Centre for Doctoral Training in Sustainable Chemistry (EP/L015633/1) for support, and they are grateful for access to the University of Nottingham High Performance Computing Facility. JH is supported by the Royal Academy of Engineering under the Chairs in Emerging Technologies scheme.

■ REFERENCES

(1) Huber, C.; Wächtershäuser, G. Activated Acetic Acid by carbon Fixation on (Fe,Ni)S Under Primordial Conditions. *Science* **1997**, *276*, 245−247.

(2) Beinert, H.; Holm, R. H.; Münck, E. Iron-Sulfur Clusters: Nature's Modular, Multipurpose Structures. *Science* **1997**, *277*, 653−659.

(3) Beinert, H. Iron-Sulfur Proteins: Ancient Structures, Still Full of Surprises. *J. Biol. Inorg Chem.* **2000**, *5*, 2−15.

(4) Johnson, D. C.; Dean, D. R.; Smith, A. D.; Johnson, M. K. Structure, Function, and Formation of Biological Iron-Sulfur Clusters. *Annu. Rev. Biochem.* **2005**, *74*, 247−281.

(5) Mortenson, L. E.; Valentine, R. C.; Carnahan, J. E. An Electron Transport Factor from Clostridiumpasteurianum. *Biochem. Biophys. Res. Commun.* **1962**, *7*, 448−452.

(6) Glaser, T.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. Ligand K-Edge X-ray Absorption Spectroscopy: A Direct Probe of Ligand−Metal Covalency. *Acc. Chem. Res.* **2000**, *33*, 859−868.

(7) Noodleman, L.; Case, D. A. Density-Functional Theory of Spin Polarization and Spin Coupling in Iron-Sulfur Clusters. In *Iron-Sulfur Proteins*; Advances in Inorganic Chemistry; Academic Press, Inc.: San Diego, CA, 1992; Vol. 38, pp 424−467.

(8) Lubitz, W.; Ogata, H.; Rüdiger, O.; Reijerse, E. Hydrogenases. *Chem. Rev.* **2014**, *114*, 4081−4148.

(9) Peters, J. W.; Stowell, M. H. B.; Soltis, S. M.; Finnegan, M. G.; Johnson, M. K.; Rees, D. C. Redox-Dependent Structural Changes in the Nitrogenase P-Cluster,. *Biochemistry* **1997**, *36*, 1181−1187.

(10) Cheek, J.; Broderick, J. B. Adenosylmethionine-Dependent Iron-Sulfur Enzymes: Versatile Clusters in a Radical New Role. *J. Biol. Inorg Chem.* **2001**, *6*, 209−226.

(11) Broderick, J. B.; Duffus, B. R.; Duschene, K. S.; Shepard, E. M. Radical S-Adenosylmethionine Enzymes. *Chem. Rev.* **2014**, *114*, 4229−4317.

(12) Radical SAM Enzymes. In *Methods in Enzymology*; Bandarian, V., Ed.; Academic Press, Elsevier, 2018; Vol. *606*.

(13) Jäger, C. M.; Croft, A. K. Anaerobic Radical Enzymes for Biotechnology. *ChemBioEng Rev.* **2018**, *5*, 143−162.

(14) Fugate, C. J.; Jarrett, J. T. Biotin synthase: Insights Into Radical-Mediated Carbon−Sulfur Bond Formation. *Biochim. Biophys. Acta, Proteins Proteomics* **2012**, *1824*, 1213−1222.

(15) Wood, Z. A.; Schröder, E.; Robin Harris, J.; Poole, L. B. Structure, Mechanism and Regulation of Peroxiredoxins. *Trends Biochem. Sci.* **2003**, *28*, 32−40.

(16) Arnér, E. S.; Holmgren, A. A Physiological Functions of Thioredoxin and Thioredoxin Reductase. *Eur. J. Biochem.* **2000**, *267*, 6102−6109.

(17) Tsukamoto, Y.; Fukushima, Y.; Hara, S.; Hisabori, T. Redox Control of the Activity of Phosphoglycerate Kinase in Synechocystis sp. PCC6803. *Plant Cell Physiol.* **2013**, *54*, 484−491.

(18) Gilbert, H. F. Redox Control of Enzyme Activities by Thiol/Disulfide Exchange. In *Methods in Enzymology*; Academic Press, Elsevier, 1984; Vol. *107*, pp 330−351. DOI: 10.1016/0076-6879(84)07022-1

(19) Richter, A. S.; Grimm, B. Thiol-Based Redox Control of Enzymes Involved in the Tetrapyrrole Biosynthesis Pathway in Plants. *Front. Plant Sci.* **2013**, *4*, 371.

(20) Hitaishi, V.; Clement, R.; Bourassin, N.; Baaden, M.; De Poulpiquet, A.; Sacquin-Mora, S.; Ciaccafava, A.; Lojou, E. Controlling Redox Enzyme Orientation at Planar Electrodes. *Catalysts* **2018**, *8*, 192.

(21) Caceres, T. B.; Price, O.; Morales, Y.; Hevel, J. Redox Control of PRMT1 Substrate Specificity. *FASEB J.* **2017**, *31*, 765−811.

(22) Bonifacio, A.; Millo, D.; Keizers, P. H. J.; Boegschoten, R.; Commandeur, J. N. M.; Vermeulen, N. P. E.; Gooijer, C.; van der Zwan, G. Active-Site Structure, Binding and Redox Activity of the Heme-Thiolate Enzyme CYP2D6 Immobilized on Coated Ag Electrodes: a Surface-Enhanced Resonance Raman Scattering Study. *J. Biol. Inorg Chem.* **2008**, *13*, 85−96.

(23) Skryhan, K.; Cuesta-Seijo, J. A.; Nielsen, M. M.; Marri, L.; Mellor, S. B.; Glaring, M. A.; Jensen, P. E.; Palcic, M. M.; Blennow, A. The Role of Cysteine Residues in Redox Regulation and Protein Stability of Arabidopsis thaliana Starch Synthase 1. *PLoS One* **2015**, *10*, No. e0136997.

(24) Léger, C.; Bertrand, P. Direct Electrochemistry of Redox Enzymes as a Tool for Mechanistic Studies. *Chem. Rev.* **2008**, *108*, 2379−2438.

(25) Gates, A. J.; Kemp, G. L.; To, C. Y.; Mann, J.; Marritt, S. J.; Mayes, A. G.; Richardson, D. J.; Butt, J. N. The Relationship Between Redox Enzyme Activity and Electrochemical Potential-Cellular and Mechanistic Implications from Protein Film Electrochemistry. *Phys. Chem. Chem. Phys.* **2011**, *13*, 7720−7731.

(26) Elliott, S. J.; Léger, C.; Pershad, H. R.; Hirst, J.; Heffron, K.; Ginet, N.; Blasco, F.; Rothery, R. A.; Weiner, J. H.; Armstrong, F. A. Detection and Interpretation of Redox Potential Optima in the Catalytic Activity of Enzymes. *Biochim. Biophys. Acta* **2002**, *1555*, 54−59.

(27) Tran, K. N.; Niu, S.; Ichiye, T. Reduction potential calculations of the Fe-S clusters in Thermus thermophilus respiratory complex I. *J. Comput. Biol.* **2019**, *40*, 1248−1256.

(28) Bak, D. W.; Elliott, S. J. Alternative FeS Cluster Ligands: Tuning Redox Potentials and Chemistry. *Curr. Opin. Chem. Biol.* **2014**, *19*, 50−58.

(29) Zhang, L.; Cui, H.; Zou, Z.; Garakani, T. M.; Novoa-Henriquez, C.; Jooyeh, B.; Schwaneberg, U. Directed Evolution of a Bacterial Laccase (CueO) for Enzymatic Biofuel Cells. *Angew. Chem., Int. Ed.* **2019**, *58*, 4562−4565.

(30) Sandford, C.; Edwards, M. A.; Klunder, K. J.; Hickey, D. P.; Li, M.; Barman, K.; Sigman, M. S.; White, H. S.; Minteer, S. D. A Synthetic Chemist's Guide to Electroanalytical Tools for Studying Reaction Mechanisms. *Chem. Sci.* **2019**, *10*, 6404−6422.

(31) Hannemann, F.; Guyot, A.; Zöllner, A.; Müller, J. J.; Heinemann, U.; Bernhardt, R. The Dipole Moment of the Electron Carrier Adrenodoxin is not Critical for Redox Partner Interaction and Electron Transfer. *J. Inorg. Biochem.* **2009**, *103*, 997−1004.

(32) Ikeda, T.; Kano, K. An Electrochemical Approach to the Studies of Biological Redox Reactions and their Applications to Biosensors, Bioreactors, and Biofuel cells. *J. Biosci. Bioeng.* **2001**, *92*, 9−18.

(33) Warshel, A.; Aqvist, J. Electrostatic Energy and Macromolecular Function. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 267−298.

(34) Sheinerman, F.; Norel, R.; Honig, B. Electrostatic aspects of protein-protein interactions. *Curr. Opin. Struct. Biol.* **2000**, *10*, 153−159.

(35) Hanoian, P.; Liu, C. T.; Hammes-Schiffer, S.; Benkovic, S. Perspectives on Electrostatics and Conformational Motions in Enzyme Catalysis. *Acc. Chem. Res.* **2015**, *48*, 482−489.

(36) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M. Electrostatic Basis for Enzyme Catalysis. *Chem. Rev.* **2006**, *106*, 3210−3235.

(37) Fried, S. D.; Boxer, S. G. Electric Fields and Enzyme Catalysis. *Annu. Rev. Biochem.* **2017**, *86*, 387−415.

(38) Warshel, A. Electrostatic Origin of the Catalytic Power of Enzymes and the Role of Preorganized Active Sites. *J. Biol. Chem.* **1998**, *273*, 27035−27038.

(39) Morgenstern, A.; Jaszai, M.; Eberhart, M. E.; Alexandrova, A. N. Quantified Electrostatic Preorganization in Enzymes Using the Geometry of the Electron Charge Density. *Chem. Sci.* **2017**, *8*, 5010−5018.

(40) Fuller, J.; Wilson, T. R.; Eberhart, M. E.; Alexandrova, A. N. Charge Density in Enzyme Active Site as a Descriptor of Electrostatic Preorganization. *J. Chem. Inf. Model.* **2019**, *59*, 2367−2373.

(41) Náray-Szabó, G. Electrostatic Catalysis in Enzymes. *J. Mol. Catal.* **1988**, *47*, 281−287.

(42) Warshel, A. Computer Simulations of Enzyme Catalysis: Methods, Progress, and Insights. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 425−443.

(43) Warshel, A.; Sharma, P. K.; Kato, M.; Parson, W. W. Modeling Electrostatic Effects in Proteins. *Biochim. Biophys. Acta* **2006**, *1764*, 1647−1676.

(44) Nielsen, J. E.; Beier, L.; Otzen, D.; Borchert, T. V.; Frantzen, H. B.; Andersen, K. V.; Svendsen, A. Electrostatics in the active site of an alpha-amylase. *Eur. J. Biochem.* **1999**, *264*, 816−824.

(45) Yang, Z.; Liu, F.; Steeves, A. H.; Kulik, H. J. Quantum Mechanical Description of Electrostatics Provides a Unified Picture of Catalytic Action Across Methyltransferases. *J. Phys. Chem. Lett.* **2019**, *10*, 3779−3787.

(46) Suess, C. J.; Martins, F. L.; Croft, A. K.; Jäger, C. M. Radical Stabilization Energies for Enzyme Engineering: Tackling the Substrate Scope of the Radical Enzyme QueE. *J. Chem. Inf. Model.* **2019**, *59*, 5111−5125.

(47) Gilson, M. K.; Honig, B. H. Calculation of Electrostatic Potentials in an Enzyme Active Site. *Nature* **1987**, *330*, 84−86.

(48) Eun, C.; Kekenes-Huskey, P. M.; Metzger, V. T.; McCammon, J. A. A model Study of Sequential Enzyme Reactions and Electrostatic Channeling. *J. Chem. Phys.* **2014**, *140*, 105101.

(49) Wade, R. C.; Gabdoulline, R. R.; Ludemann, S. K.; Lounnas, V. Electrostatic steering and ionic tethering in enzyme-ligand binding: Insights from simulations. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5942−5949.

(50) Fried, S. D.; Bagchi, S.; Boxer, S. G. Extreme Electric Fields Power Catalysis in the Active Site of Ketosteroid Isomerase. *Science* **2014**, *346*, 1510−1514.

(51) Shaik, S.; Mandal, D.; Ramanan, R. Oriented Electric Fields as Future Smart Reagents in Chemistry. *Nat. Chem.* **2016**, *8*, 1091−1098.

(52) Orozco-Gonzalez, Y.; Kabir, M. P.; Gozem, S. Electrostatic Spectral Tuning Maps for Biological Chromophores. *J. Phys. Chem. B* **2019**, *123*, 4813−4824.

(53) Prah, A.; Frančišković, E.; Mavri, J.; Stare, J. Electrostatics as the Driving Force Behind the Catalytic Function of the Monoamine Oxidase A Enzyme Confirmed by Quantum Computations. *ACS Catal.* **2019**, *9*, 1231−1240.

(54) Bhave, D. P.; Han, W.-G.; Pazicni, S.; Penner-Hahn, J. E.; Carroll, K. S.; Noodleman, L. Geometric and Electrostatic Study of the [4Fe-4S] Cluster of Adenosine-5′-Phosphosulfate Reductase from Broken Symmetry Density Functional Calculations and Extended X-ray Absorption Fine Structure Spectroscopy. *Inorg. Chem.* **2011**, *50*, 6610−6625.

(55) Noodleman, L. Valence Bond Description of Antiferromagnetic Coupling in Transition Metal Dimers. *J. Chem. Phys.* **1981**, *74*, 5737−5743.

(56) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kuś, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J., III; Woodcock, H. L.; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.,

Jr.; DiStasio, R. A.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyaev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A. D.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Neuscamman, E.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y.-C.; Thom, A. J. W.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K. L.; Chipman, D. M.; Cramer, C. J., III; Goddard, W. A.; Gordon, M. S.; Hehre, W. J.; Klamt, A., III; Schaefer, H. F.; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Van Voorhis, T.; Herbert, J. M.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. Advances in Molecular Quantum Chemistry Contained in the Q-Chem 4 Program Package. *Mol. Phys.* **2015**, *113*, 184−215.

(57) Carvalho, A. T. P.; Swart, M. Electronic Structure Investigation and Parametrization of Biologically Relevant Iron-Sulfur Clusters. *J. Chem. Inf. Model.* **2014**, *54*, 613−620.

(58) Swart, M.; Groenhof, A. R.; Ehlers, A. W.; Lammertsma, K. Validation of Exchange−Correlation Functionals for Spin States of Iron Complexes. *J. Phys. Chem. A* **2004**, *108*, 5479−5483.

(59) Swart, M.; Ehlers, A. W.; Lammertsma, K. Performance of the OPBE Exchange-Correlation Functional. *Mol. Phys.* **2004**, *102*, 2467−2474.

(60) Handy, N. C.; Cohen, A. J. Left-Right Correlation Energy. *Mol. Phys.* **2001**, *99*, 403−412.

(61) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(62) Van Lenthe, E.; Baerends, E. J. Optimized Slater-type basis sets for the elements 1-118. *J. Comput. Chem.* **2003**, *24*, 1142−1156.

(63) Chong, D. P.; Van Lenthe, E.; Van Gisbergen, S.; Baerends, E. J. Even-Tempered Slater-Type Orbitals Revisited: From Hydrogen to Krypton. *J. Comput. Chem.* **2004**, *25*, 1030−1036.

(64) Chong, D. P. Augmenting Basis Set for Time-Dependent Density Functional Theory Calculation of Excitation Energies: Slater-Type Orbitals for Hydrogen to Krypton. *Mol. Phys.* **2005**, *103*, 749−761.

(65) Harris, T. V.; Szilagyi, R. K. Iron-sulfur bond covalency from electronic structure calculations for classical iron-sulfur clusters. *J. Comput. Chem.* **2014**, *35*, 540−552.

(66) Szilagyi, R. K.; Winslow, M. A. On the accuracy of density functional theory for iron-sulfur clusters. *J. Comput. Chem.* **2006**, *27*, 1385−1397.

(67) Harris, T. V.; Szilagyi, R. K. Protein Environmental Effects on Iron-Sulfur Clusters: A Set of Rules for Constructing Computational Models for Inner and Outer Coordination Spheres. *J. Comput. Chem.* **2016**, *37*, 1681−1696.

(68) Swart, M. Accurate Spin-State Energies for Iron Complexes. *J. Chem. Theory Comput.* **2008**, *4*, 2057−2066.

(69) Güell, M.; Luis, J. M.; Solà, M.; Swart, M. Importance of the Basis Set for the Spin-State Energetics of Iron Complexes. *J. Phys. Chem. A* **2008**, *112*, 6384−6391.

(70) Klamt, A.; Schüürmann, G. COSMO A New Approach to Dielectric Screening in Solvents With Explicit Expressions for the Screening Energy and its Gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, *2*, 799−805.

(71) Gilson, M. K.; Honig, B. H. The Dielectric Constant of a Folded Protein. *Biopolymers* **1986**, *25*, 2097−2119.

(72) Sharp, K. A.; Honig, B. Electrostatic Interactions in Macromolecules: Theory and Applications. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 301−332.

(73) Noodleman, L.; Baerends, E. J. Electronic structure, magnetic properties, ESR, and optical spectra for 2-iron ferredoxin models by LCAO-X.alpha. valence bond theory. *J. Am. Chem. Soc.* **1984**, *106*, 2316−2327.

(74) Noodleman, L.; Peng, C. Y.; Case, D. A.; Mouesca, J.-M. Orbital Interactions, Electron Delocalization and Spin Coupling in Iron-Sulfur Clusters. *Coord. Chem. Rev.* **1995**, *144*, 199−244.

(75) Emelyanova, N. S.; Poleshchuk, O. Kh.; Sanina, N. A.; Aldoshin, S. M. Quantum Chemical Modelling of Ligand Substitution in Cationic Nitrosyl Complexes. *Russ. Chem. Bull.* **2014**, *63*, 1088−1094.

(76) Sandala, G. M.; Hopmann, K. H.; Ghosh, A.; Noodleman, L. Calibration of DFT Functionals for the Prediction of 57Fe Mössbauer Spectral Parameters in Iron-Nitrosyl and Iron-Sulfur Complexes: Accurate Geometries Prove Essential. *J. Chem. Theory Comput.* **2011**, *7*, 3232−3247.

(77) Jensen, K. P.; Ooi, B.-L.; Christensen, H. E. M. Accurate Computation of Reduction Potentials of 4Fe−4S Clusters Indicates a Carboxylate Shift in Pyrococcus furiosus Ferredoxin. *Inorg. Chem.* **2007**, *46*, 8710−8716.

(78) Han, W.-G.; Liu, T.; Lovell, T.; Noodleman, L. Active Site Structure of Class I Ribonucleotide Reductase Intermediate X: A Density Functional Theory Analysis of Structure, Energetics, and Spectroscopy. *J. Am. Chem. Soc.* **2005**, *127*, 15778−15790.

(79) Maiocco, S. J.; Walker, L. M.; Elliott, S. J. Determining Redox Potentials of the Iron-Sulfur Clusters of the AdoMet Radical Enzyme Superfamily. In *Methods in Enzymology*; Bandarian, V., Ed.; Academic Press, 2018, pp 319−339. DOI: 10.1016/bs.mie.2018.06.002

(80) Miranda-Quintana, R. A.; Martínez González, M.; Ayers, P. W. Electronegativity and Redox Reactions. *Phys. Chem. Chem. Phys.* **2016**, *18*, 22235−22243.

(81) Parr, R. G.; Szentpály, L. v.; Liu, S. Electrophilicity Index. *J. Am. Chem. Soc.* **1999**, *121*, 1922−1924.

(82) Moens, J.; Roos, G.; Jaque, P.; De Proft, F.; Geerlings, P. Can Electrophilicity Act as a Measure of the Redox Potential of First-Row Transition Metal Ions? *Chem. Eur J.* **2007**, *13*, 9331−9343.

(83) Chattaraj, P. K.; Giri, S. Electrophilicity index within a conceptual DFT framework. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **2009**, *105*, 13−39.

(84) Moens, J.; Jaque, P.; De Proft, F.; Geerlings, P. The Study of Redox Reactions on the Basis of Conceptual DFT Principles: EEM and Vertical Quantities. *J. Phys. Chem. A* **2008**, *112*, 6023−6031.

(85) Roos, G.; Geerlings, P.; Messens, J. Enzymatic Catalysis: The Emerging Role of Conceptual Density Functional Theory. *J. Phys. Chem. B* **2009**, *113*, 13465−13475.

(86) Sharma, S.; Sivalingam, K.; Neese, F.; Chan, G. K.-L. Low-Energy Spectrum of Iron-Sulfur Clusters Directly from Many-Particle Quantum Mechanics. *Nat. Chem.* **2014**, *6*, 927−933.

(87) Gillum, W. O.; Frankel, R. B.; Foner, S.; Holm, R. H. Synthetic Analogues of the Active Sites of Iron-Sulfur Proteins. XIII. Further Electronic Structural Relationships between the Analogues [Fe2S2-(SR)4]2- and the Active Sites of Oxidised 2Fe-2S Proteins. *Inorg. Chem.* **1976**, *15*, 1095−1100.

(88) Teo, R. D.; Rousseau, B. J. G.; Smithwick, E. R.; Di Felice, R.; Beratan, D. N.; Migliore, A. Charge Transfer between [4Fe4S] Proteins and DNA Is Unidirectional: Implications for Biomolecular Signaling. *Chemistry* **2019**, *5*, 122−137.

(89) Cervantes-Salguero, K.; Seminario, J. M. Structure and Energetics of Small Iron Clusters. *J. Mol. Model.* **2012**, *18*, 4043−4052.

(90) Nicolet, Y.; Amara, P.; Mouesca, J.-M.; Fontecilla-Camps, J. C. Unexpected Electron Transfer Mechanism Upon AdoMet Cleavage in Radical SAM Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 14867−14871.