Author:
Thayer, George S

Title:
Computational mechanistic analysis of homogeneous copper-catalysed Ullmann type coupling reactions.

Computational mechanistic analysis of homogeneous copper-catalysed Ullmann type coupling reactions

Candidate: George Thayer Supervisor:

Dr Natalie Fey

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Master of Science by Research Degree in Chemistry in the Faculty of Science.

Submission Date: 18th January 2022

Word Count: 33132

Abstract

This project aims to critically review past work exploring the mechanism of homogeneously catalysed coupling reactions whilst exploring the use of amino alcohol ligands in copper-mediated Ullmann-type arylation reactions. The use of Cu(I) and Cu(III) assumes an oxidative addition – reductive elimination mechanism. The coordination sphere of Cu(I) was found to optimise towards 2- and 3- coordinate complexes, with 14-electron 2-coordinate complexes being favourable. Cu(III) can be difficult to isolate experimentally and has also been seen to be problematic to locate computationally due to the inaccessibility of a 5-coordinate geometry for copper in this oxidation state, but calculations showed that it may be viable. The most likely mechanism according to the calculations presented here focuses on a 4-coordinate pathway, which benefits from low steric strain and is accessible as a result of the hemilabile nature of amino alcohols. This pathway shares some similarities such as coordination sphere and solvent with other mechanistic proposals but has a focus on this lower coordination number throughout.

Data mining of the Cambridge Structural Database (CSD) was used to gain a greater understanding of the current abundance of copper species and provided an insight into some of the crystal structure of isolated copper complexes with varying oxidation states.

Solvation, dispersion, basis set and functional effects on the relative free energy and geometry have been assessed to determine a 'best method' for computational analysis in this project, with compromises being made to manage computational costs. The B3LYP density functional was found to be a robust choice for these calculations. For energies, the 6-311+G(d,p) basis set was identified as a suitable basis set for all atoms except Cu and I which were modelled using the SDD basis set with an Effective Core Potential (ECP). Solvation was modelled implicitly using the IEFPCM method and dispersion was corrected for using the GD3 method.

The proposed mechanism is energetically viable compared with proposed mechanisms in literature and shows evidence that copper coordinated with an amino alcohol is in fact catalytically competent, at least from a computational point of view. The use of 2-coordinate Cu(I) complexes requires further investigation as they show promising results in terms of providing a kinetically-favourable route to oxidative addition.

Dedication and Acknowledgements

I dedicate this thesis to all of my friends and family of whom supported me through this year and helped to keep me on track during the pandemic.

I would like to thank my supervisor Dr Natalie Fey for having patience throughout this process and for always supporting me throughout this project. I would also like to thank Mr Derek Durand for his advice and training through this project, as he has also supported, motivated and trained me during this year.

Authors Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: George Thayer DATE: 17/01/2022

Table of Contents

Glossary

6-31G(d) – A double-zeta basis set with polarization.

6-311+G(d,p) – A triple-zeta basis set with polarization and diffuse functions.

B3LYP – The Becke three parameter hybrid functional with the exchange and correlation energies described by Lee, Yang and Parr.

Basis Set – A collection of functions that describe the electronic wavefunction of molecular orbitals.

BP86 – A common GGA density functional.

Density Functional Theory (DFT) – A computational quantum mechanical simulation method used to calculate the structure and energy of molecular systems. (See ch. 1 for review).

DMF – Dimethylformamide.

Dispersion – A correction term used in order to more accurately corrects for the attractive contribution of intramolecular interactions which is not fully captured by standard density functionals

Effective Core Potential (ECP) – A basis set that models the relativistic effects of core electrons in heavier atoms as one simplified potential function.

GGA – Generalized Gradient Approximation.

HF – Hartree Fock (see ch. 4 for computational approaches).

IAT – Iodine Atom Transfer.

LANL2DZ – The Los Alamos National Laboratory 2-double-zeta basis set.

Molecular Mechanics (MM) – A computational methodology that uses classical mechanics for calculation.

PCM – Polarizable Continuum Model

SDD – Stuttgart Dresden basis set, triple zeta quality with effective core potentials for elements beyond Na.

SET – Single Electron Transfer.

SPE – Single Point Energy.

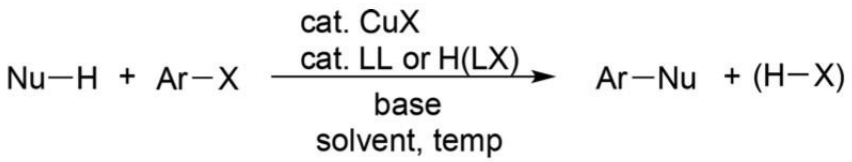
Table of Figures

Table of Tables

1- Introduction

Coupling reactions have been studied extensively and are the basis of many industrial processes such as intermediate synthesis for medicinally relevant compounds , for example enzyme inhibitors and biological screening.[1] Coupling reactions are a useful and practical method for the formation of aryl-N, aryl-O and aryl-C bonds, with industrial uses in the polymer industry.[2] It is common for transition metals such as platinum and palladium to be used in these processes, however, other d-block metals such as copper and nickel have high catalytic activity.[3] The former are often investigated in order to save costs, therefore metals with similar properties are of interest.

The use of phenanthroline and diketone ligands in Ullmann-type copper catalysed arylation reactions to form carbon-heteroatom and carbon-carbon bonds through the coupling of aryl halides with amino alcohols is well reported.[4-8] The general Ullmann-type reaction is shown below in Scheme 1, where the major product is a result of arylation. The nucleophile in this reaction is often an amino alcohol.[2, 5]



Scheme 1: General Ullmann-Type reaction.

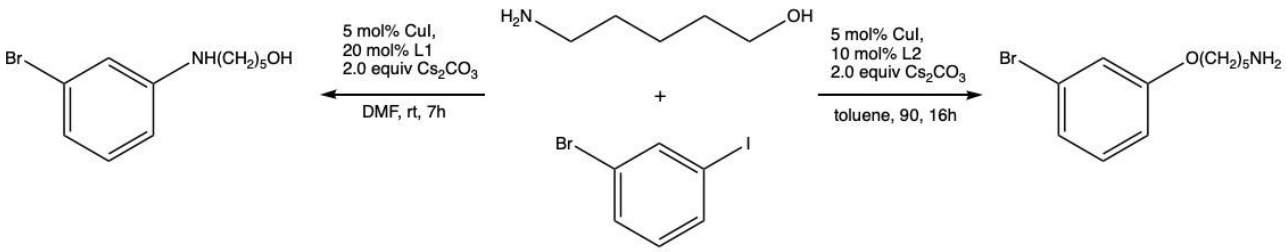
The Ullmann reaction describes the coupling of aryl halides with amines, phenols and amides[9] leading to the formation of biaryl ethers, often paired with a moderate base and a catalyst. Cross coupling traditionally uses an sp^2 hybridized aryl halide electrophile with an organometallic nucleophile and is catalysed by a metal. This is because direct coupling using sp^3 alkyl reagents can lead to undesired HX products.[10]

The use of copper as a substitute for palladium has become more popular in recent decades as a low-cost complementary metal catalyst due to its similar characteristics and shows potential for alternative pathways.[11] These reactions require high temperatures and large amounts of reagents in order to successfully complete, hence the aim to develop more efficient catalysts to aid these processes. The metal catalyst precursor must be introduced to the reaction mixture and is

commonly added using copper salts (CuI) or copper oxides such as CuO_2 that can be reduced from Cu(II) to Cu(I) for this reaction to proceed,[2] which is important as the active catalyst must be able to form experimentally. CuI is another popular salt used to introduce copper cations into solution, particularly in reactions where the substrate involved is ArI, as there is a high concentration of I^- ions in solution.[12] This process is largely about introducing the metal into solution whilst using soluble anions that will not interfere with the reaction, hence why halide salts are preferred as they will already be present. The use of aryl halides in coupling reactions has been explored in detail theoretically and experimentally and some of these will be reviewed below.[13]

There has recently been a large push in the field of copper catalysis, with many attempts made to increase the knowledge base of the mechanism behind copper catalysed coupling reactions whether Ullmann-type or not.[2]

Work undertaken by Yu *et al.*[4] investigated the oxidative addition / reductive elimination pathway for copper mediated arylation using iodobenzene with diketone (L1) and phenanthroline (L2) bidentate ligands, the reaction for this can be seen below in Scheme 2.



Scheme 2: Arylation reaction using 5C chain amino alcohol reproduced from Yu et al.[4]

This was explored in order to determine whether changing the donor atoms in a ligand would have an effect on product selectivity. The B3LYP[14] density functional was used with the 6-31g(d) double zeta quality basis set, using the LanL2DZ[15] basis set as an Effective Core Potential for modelling of Cu and I atoms; computational methodology will be covered in due course (section 2). The resulting energies were corrected with Single Point Energy (SPE) calculations using a larger basis set (6-311+G(d,p)). SPE calculations were also used to explore the solvent effect in this study, where acetonitrile, toluene and DMF were used. However, full optimizations were not carried out in this work as, compared to the gas phase, many solution phase calculations failed to converge.

The amino alcohol substrate used by Yu was 5-amino-1-pentanol to be coupled with an aryl halide, which has been used experimentally[16] as well as theoretically. The ease with which an aryl halide will associate with Cu(I) is influenced by the net charge of the complex, where an aryl iodide will bond to a metal centre readily, allowing for the copper to undergo oxidative addition. However, when neutral phenanthroline is used as a ligand, a positively charged metal centre is unfavourable for oxidative addition,[4] therefore a deprotonated amino alcohol is suspected to bind with the Cu(I) before the aryl halide in order to create a neutral species. The main findings of work conducted by Yu are that in an oxidative addition-based mechanism, activation of the aryl halide is not the main cause for the N or O coupling selectivity influence, instead the nucleophile coordination bears greater significance. This work[4] determined that there were several optimized geometries that justified the theoretical presence of a Cu(III) species involved in the mechanism, whilst also providing lower energy barriers than those postulated by computational studies of the Single Electron Transfer (SET) and Iodine Atom Transfer (IAT) pathways,[5] however, it should be noted that these structures were considered in the gas phase only. This inability to confidently find and isolate structures whilst using implicit solvation raises concern for the ease with which high coordination number copper(III) species can be optimized thus exploration of the range of geometries exhibited by Cu(III) is useful. The mechanisms proposed by Yu *et al*[4] largely represents the basis for this project, as it uses very similar conditions and computational methodology as outlined above and there is a well-defined mechanism for both N and O arylation of amino alcohols, which has been reproduced below in Figure 1.

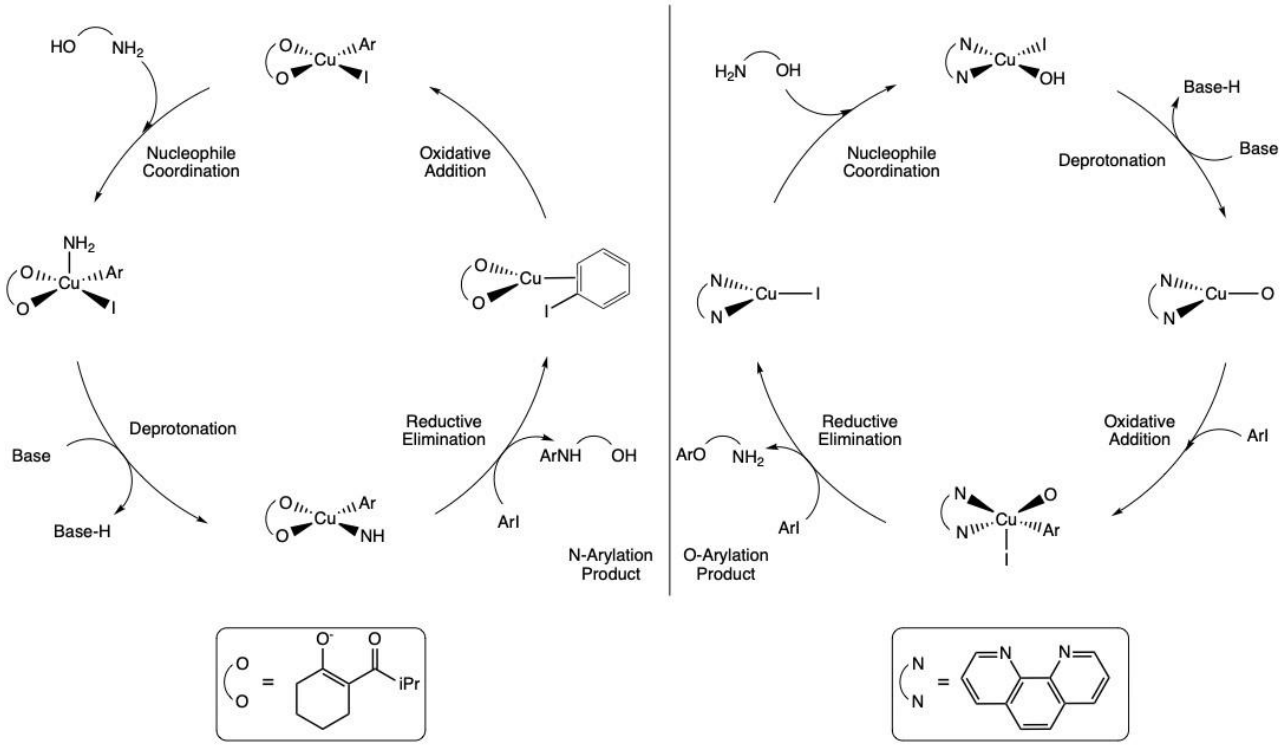


Figure 1: Mechanistic pathway for Cu-mediated arylation of amino alcohols.

The authors note that the alcohol group readily dissociated from the metal centre during optimisations, suggesting that a 4-coordinate Cu(I) species was unlikely to be stable computationally (B3LYP/6-31G(d)/LanL2DZ)[4] whilst using a diketone ligand. This corresponds to an initial 4-coordinate $18e^-$ complex optimizing to a 3-coordinate $16e^-$ complex, which is often seen in Cu(I) geometries when there are steric effects that restrict the structure from being tetrahedral,[17] such as a long chain or bulky substrate, due to the electron repulsion between ligands having a greater effect than the increase in the coordination number.

Optimal pathways for N and O selectivity have been described by Yu *et al.*,[4] yet there is still a lack of detail regarding a comparison of amino alcohols to different ligands in this ar, as well as the ability to perform calculations to challenge or support existing hypotheses on the mechanism. Competition between amino alcohols and other neutral or charged ligands such as phenanthroline or diketone type ligands as outlined in the mechanism shown above in Figure 1 is important for the development of better catalysts. The competition is important to consider, as there is research that alludes to the ability of amino alcohols to be able to act as ligands as well as substrates, evidenced by an amino alcohol deprotonated at the terminal oxygen atom displaying a higher binding energy than several N,N donor ligands.[18] Typically, phenanthroline is a stable ligand and binds in a

bidentate arrangement, however, this may be challenged when using polar solvents and copper(I)[4] where solvent may displace the ligand, favouring monodentate coordination of phenanthroline. Short chain amino alcohols have a similar bite angle and share one of the donor atoms with the previously mentioned ligands of around 85-90°, so might be able to perform the same role in a reaction. There are several mechanistic considerations to be made when determining the catalytic activity of copper-amino alcohol complexes, as in recent literature amino alcohols have been used as a substrate with various O,O$^-$ and N,N coordinating ligands present.[2, 4, 9] This type of competitive binding is likely to be a reaction specific challenge due to the number of different ligands and reagents that are able to be present.

Amino alcohols can coordinate in a mono- or bidentate fashion as ligands, with previous work showing amino alcohols, such as L-proline,[9] to act as excellent ligands for copper catalysed Ullmann type reactions. There are a range of Cu(I)/Cu(III) coordination geometries which are likely dependent on carbon chain length/ligand bite angle and protonation state, as more bulky or sterically demanding[19] ligands will reduce the likelihood of a high coordination number for the metal. Altering the protonation state of the amino alcohol ligand will affect the net charge of a complex, which will cause more electrophilic or nucleophilic species to become involved in bonding. The carbon chain can determine the distance between donor groups and also the bite angle of said ligand. Jones *et al.*[5] suggested the idea of a reaction mechanism analysing the viability of Single Electron Transfer (SET) and Iodine Atom Transfer (IAT) mechanisms, which were explored computationally. The computational methodology used in this work used the B3LYP[19] and MPWB1K[5, 20] functional with the 6-31+G(d,p) basis set, these parameters will be explored in detail. In a similar way to Yu *et al.*[4] the solvation model used was a CPCM implicit method where acetonitrile, DMF and toluene were considered. These calculations did not consider dispersion corrections. Findings showed that for the Ullmann reaction using a diketone type ligand, the activation energy of the SET (+27.2 kcal mol^{-1}) and IAT (+32.9 kcal mol^{-1}) mechanisms were lower than those of an oxidative addition (+64.6 kcal mol^{-1}) or sigma bond metathesis (+57.1 kcal mol^{-1}) mechanism. The results of this also showed that the use of a diketone type ligand promoted a SET mechanism. It should be noted that the use of a phenanthroline ligand caused the activation energies of the SET mechanism (+43.6 kcal mol^{-1}) to be slightly higher than that of oxidative addition (+43.2 kcal mol^{-1}) or sigma bond metathesis (+43.4 kcal mol^{-1}) by 0.4 and 0.2 kcal mol^{-1}

respectively, which is a negligible difference with regard to differentiating between the three mechanisms. For the phenanthroline ligand, the IAT activation energy (+34.0 kcal mol^{-1}) was most favoured. Several of the postulated intermediates were high energy and some were unable to have their geometries optimized, which was observed for Cu(II) and Cu(III) species. The first half of both the SET and IAT mechanisms suggested come with the caveat of a very large barrier for the first 2 steps of the reaction, where a high temperature is required in order to overcome this. As previously noted, there were no dispersion corrections used in this study, therefore the intramolecular interactions favouring higher coordination numbers may be underestimated.

There are several alternative postulated mechanisms for copper-catalysed coupling to proceed, including a single-electron transfer or halogen atom transfer,[5] which involves a Cu(I)/Cu(II) redox couple. These pathways have been suggested to require a lower activation energy than oxidative addition or bond metathesis mechanisms,[4] however, there is compelling evidence that the HAT and SET mechanisms are disfavoured. The studies by Yu *et al.*[4] mentioned previously showed that several Iodine-atom transfer (IAT) transition states were difficult to locate and, when found, displayed large barriers (+42.7 kcal mol^{-1})[4] which suggests that the IAT mechanism is not the lowest energy pathway found at this level of theory. Several SET transition states were compared with those identified using an oxidative addition pathway and were found to have much higher energy barriers, once again giving evidence of the oxidative addition path to be the lowest energy route for the reaction to proceed. There is also experimental evidence against the SET and IAT mechanisms by Tye *et al.*,[21] where the expected products of an aryl halide cyclization through these mechanistic pathways were not formed.[4, 21, 22] Tye *et al.*[21] also suggested that despite being uncommon, alkylcopper(III) and arylcopper(III) species were identified and that a computational study of the free energy of the phenanthroline ligated species (+22.1 kcal mol^{-1}) fell in line with that found in experimental work.

Another study by Giri *et al.*[23] took an experimental kinetic study with additional DFT calculations to study the addition of phenols to aryl halides, which bears resemblance to this research on amino alcohols due to sharing an alcohol functional group. In the work by Giri, the oxidative addition/reductive elimination pathway is targeted and suggested to be most likely. The DFT calculations used B3LYP with 6-311G(d,p) and an effective core potential of LANL2TZ for Cu and I.

Solvation was modelled implicitly using Conductor-like polarizable continuum model (CPCM) but was only used in SPE calculations. The starting complex was a 3-coordinate Cu(I) with a bidentate ligand and a phenol. Most of the initial geometries found in the gas phase were optimized successfully, however, two of the 5-coordinate Cu(III) intermediate attempts were unsuccessful. The Cu(III) geometries that did converge to a minimum were done so in the gas phase also and would benefit from optimization using a solvation model and dispersion corrections in order for the results to be more comparable to the work done using amino alcohols. It was found that computationally, it was more favourable for the halide to dissociate from the Cu(III) intermediate (+21 kcal mol^{-1}) to form a lower energy tetracoordinate Cu(III) intermediate (+7 kcal mol^{-1}) before the reductive elimination step of the reaction, as the transition state for this step was only 5-10 kcal mol^{-1} higher in free energy than the intermediate. This was considerably more achievable than reductive elimination directly from the 5 coordinate Cu(III), which required an increase of +27-29 kcal mol^{-1} in order to proceed.[23] This is useful as it suggests that for a given ligand, the steric bulk of a 5-coordinate Cu(III) intermediate may hinder the reductive elimination step, therefore the removal of iodide ions or exploring the hemi-lability of amino alcohols in order to reduce the coordination number or a complex may be beneficial. This step of a reaction is to be discussed in more detail in Section 9 – Mechanism.

Selectivity for an O- or N-arylation product was achieved whilst using a phenanthroline type ligand for O-arylation and a diketone type ligand for N-arylation.[16] An efficient method for O selectivity without a ligand was postulated, however, not found to be accurate, whereas a ‘ligand-free’ N selective product was able to be formed without additional ligands with a high yield (83%), because a 1,2-amino alcohol is able to act as a supporting ligand as well as a substrate.[16] In this case, O selectivity was only able to be achieved whilst using an amino alcohol with at least 4 methylene groups between the amine and hydroxyl groups. It was hypothesized that two of the leading steps of the mechanism responsible for selectivity were deprotonation and coordination. This implication suggests that a deprotonated substrate may cause a decrease in the electrophilic nature of Cu(I) centre, which would encourage the binding of the amino end of ethanolamine. This suggestion agrees with work from Yu,[4] where it was postulated that a difference in charge would cause the terminal N- donor atom to gain higher selectivity.

The main mechanism considered in this research is thus an oxidative addition/reductive elimination pathway[24] accommodating the coupling of an aryl iodide to an amino alcohol with a C2 backbone. The oxidative addition of Cu(I) is postulated[4] to proceed through the addition of an aryl halide to a 2- or 3-coordinate complex in which the ligands are mono- or bidentate amino alcohols. Copper has been used in many ways to enable organometallic chemists to produce functional catalysts that aid in coupling reactions such as the oxidative Chan-Lam reaction, where boronic acids are coupled with alcohols or amines. Recent work from Bell[25] *et al.* shows that Cu(I) may accelerate aryl transfer with diaryliodoniums, which gives evidence that copper-mediated arylation is likely to benefit from the presence of a transition metal. In literature by Ahn *et al.* [26] this process is able to occur without copper present, however, the addition is significantly more efficient whilst using copper. This example[26] uses the Cu(III) intermediate in a 4-coordinate complex, showing that there are likely to be multiple coordination geometries when the high oxidation state is achieved. The main difference of this reaction to the research undertaken here regarding amino alcohols is the range of primary and secondary amine precursors used as there is no oxygen arylation pathway possible from this work for reference [25] therefore previous work from Yu[4] *et al.* and Jones[5] *et al.* provides a greater insight to both N and O donor ligand pathways.

The formation of C-N bonds from defined aryl-Cu(III) complexes has been studied and squareplanar complexes with several Cu-N bonds have been observed experimentally by ^{1}H NMR.[27] It can be noted that in the mentioned study,[27] the Cu(III) structures were polydentate and macrocyclic, therefore differ in physical properties from amino alcohols. The binding in question contains a Cu(III)-aryl bond amongst 3 N-donor atoms from the same ligand.[27]

The inclusion of a Cu(III) intermediate has been questioned for some time due to the absence of well defined, observed examples.[28] Work from Li and Lan[28] was centred around the involvement of a Cu(III) intermediate and assessed whether it is necessary for this high oxidation state to be used, suggesting alternative mechanistic pathways possible using Cu(I)/Cu(II). The computational methodology used in their work was the commonly used 6-31g(d) double-zeta quality basis set paired with the triple-zeta SDD basis set for the ECP for copper, which is a larger ECP than used by Yu *et al*.[4] Similar to work by Yu *et al*,[4] the SPE calculations used the 6-311+G(d,p) basis set on all other atoms in order to improve the accuracy of energy calculations. The Cu(III) intermediate is

often thermodynamically unfavourable, which suggests that if this higher oxidation state is reached, it will lead to a rapid reductive elimination from what may be the rate determining step.
This study aimed to find a bypass for this intermediate that may allow for the energy of the whole system to be lower. This work gives support for the SET mechanism, as it focusses on the inclusion of radical species as the electron transfer resulting in a Cu(I) to Cu(II) oxidation is more optimal for the formation of new covalent bonds. Binuclear copper species have also been explored as active species that may lead to a low energy reductive elimination; however, these species were slightly thermodynamically unstable, therefore this pathway was less explored.[28]

One study investigated the kinetics of copper catalysed arylation reactions, with a focus on the arylation of perfluoroalkyl copper complexes.[29] The ligands used in this study were bidentate dinitrogen donor ligands. Similarly to previous work,[28] it was also suggested that the formation of a Cu(III) intermediate was rate limiting, with oxidation to the higher 3+ state being facilitated by ancillary ligands, such as bipyridine, providing large electron density around the metal centre. A ligand-free approach was considered, and it was determined that the majority of the hypothesised reaction steps proceeded at a faster rate when ancillary ligands were present, which was determined by ^{19}F NMR.[29] Kinetic studies found that for reactions with aryl halides, less electron donating ligands resulted in a faster rate of reaction compared with more electron donating counterparts. Results from these experimental data were consistent with prior DFT calculations that determined the reactivity of different substituted aryl halides with free energy barriers varying by only ±1 kcal mol^{-1}.[29] There were no trends found in the Gibbs energy calculations, which was thought to be a result of the assumptions made by the implicit solvation model used. This is the same model used for the work in this project as it is simpler than explicit solvation. Despite this, implicit solvation for smaller molecules is useful as the approximations are minimised in comparison to large structures such as proteins and has been suggested to be a good choice for the calculations of binding energies.[30]

In the present study, beginning with only amino alcohols used as ligands removes competition for binding sites on the metal centre, which allows for determination of whether they are viable to be used as ligands as well as substrates. The use of binding energies in separate calculations will help to determine the likely effect of introducing other bidentate ligands into the reaction mixture

alongside amino alcohols, introducing a competition for binding. It would be useful to determine a hierarchy of ligand binding preferences, comparing O,O$^-$ and N,N-bidentate donor ligands with N,O and N,O$^-$ bidentate amino alcohol ligands, as this would provide a better understanding of displacement and hemi-lability for specific arylation reactions.[18] The addition of a base is required for this reaction to occur, with Cs_2CO_3 or KOH often being selected for mild bases.[23]

1.1 - Rationale and Hypotheses

There is a gap in the literature surrounding the use of amino alcohols as ligands as well as substrates,[4, 31] which would be beneficial to fill. Unlike typical O,O and N,N bidentate ligands, amino alcohols can use N,O mixed donor atoms or be monodentate, this may be due to amino alcohols simply being a different type of ligand. The difference in valence and electronic character of these donor atoms can allow for amino alcohols to behave in a different way to phenanthroline and diketone type ligands. The work in this report aims to develop a greater understanding of the binding capabilities of amino alcohols and to determine their suitability as ligands. There are many modes of binding available to these ligands due to their size and character, this work hopes to shed some light on the preferences of amino alcohol binding regarding different oxidation states and coordination numbers of copper.

Studies using the Cambridge Structural Database (CSD)[32, 33] have been used to better understand the coordination of ligands to copper in a wider range of complexes. The CSD is a library of peer reviewed crystallographic data about molecules and compounds that have had their physical properties such as geometry, coordination sphere, bond length and bond angles recorded. The benefit of this type of database is to obtain large quantities of information about specific bonds or atoms with experimental backing.

As a result of past studies,[4] it is hypothesised that ligated copper in solution may display a hierarchy when in competition to bind with different bidentate ligands, where the ligand may be displaced. This hierarchy is likely determined by chemical properties of ligands involved, specifically the donor atoms involved in coordination. It has been seen that N,N donors have a lower metal-ligand bond energy than O,O donors, with amino alcohols placed between.[18] This work shows that there is a

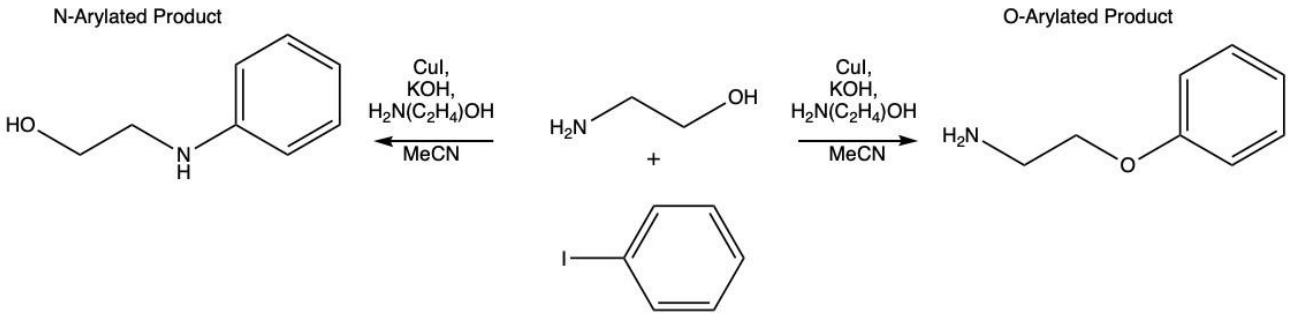
relationship and a possible correlation between the donors and their ability to out-compete each other in a reaction mixture, with consideration given to the charge on each of the ligands.

The goal of computational chemistry work is to improve the ever-growing knowledge base of catalyst design and development. This work is a step towards understanding the nature of amino alcohols in comparison to other ligands and the effect that they can have in a reaction mixture, in order to pursue greater efficiency and accuracy with predicting the catalytic nature of copper. The goal of this project is to explore the use of amino alcohols as ligands as well as substrates, which has yet to be explored in great detail. Many amino alcohols used in literature contain chains of 5 or more carbon atoms separating the NH_2 and OH groups,[4] which encourages the use of amino alcohols as a substrate, due to both the N and O donor atoms being far apart. In this work, the amino alcohol used is ethanolamine, which is able to act as a bidentate ligand due to the small distance between the donor atoms. It has been shown that short chain amino alcohols can act as ligands,[3] therefore further exploration of the conditions and properties that encourage ligand behaviour is important.

Another goal of computational chemistry is prediction of the most likely mechanistic pathway for a reaction to proceed. This type of investigation is useful as it provides an accurate idea of whether a combination of metal and ligand is catalytically competent, whilst beginning to outline a thermodynamically feasible mechanism for a reaction to proceed. One challenge of mechanism prediction is a starting point; therefore, literature was paramount in determining a likely start of the cycle which involved choosing a way to bring copper into solution. A common method for this is the dissolving of copper iodide. The CuI is often formed through a mixture of copper sulphate and potassium iodide, which is ideal due to aryl iodide and KOH being used in the reaction mixture as a nucleophile and base in this reaction. Common solvents for bringing CuI into solution are DMF and DMSO.[34] Experimental work from Sung *et al*[34] suggested that the solubility of Cs_2CO_3 is low in DMSO, whilst CuI solubility was considered suitable, therefore consideration of the conditions required and strength of bases in the arylation of amino alcohols in acetonitrile are important. When assessing the stability of Cu(I) bound with a phenanthroline-type ligand, it was found that an amino alcohol deprotonated and coordinated at the terminal alcohol was the favoured complex.[4]

This is an important detail as it suggests that early deprotonation of amino alcohols may be useful for a low energy pathway.

The main mechanism studied in this work is an oxidative addition/reductive elimination reaction shown below in Scheme 3.



Scheme 3: Scheme for the N- and O- arylation reactions.

The copper species involved are d^{10} Cu(I) and d^{8} Cu(III). The coordination sphere of these species can likely vary; copper(I) to (III) redox catalysis can be observed in several types of mechanism such as Chan-Lam cross coupling.[25] Ligands have been shown to be critical in the progression of copper catalysed reactions, as they make the conditions much milder which is necessary for consideration of these processes in an industrial context.[35] Some low cost ligands that are easily synthesised and therefore at the forefront of study are bipyridine derivatives that generate a high yield experimentally, such as iminopyridine with a 95% yield in an Ullmann type reaction.[35]

The solubility and stability of catalysts is vastly improved with the addition of ligands.[9] Many ligands used in these reactions are bidentate with O,O or N,N donor atoms such as phenanthroline or diketone type ligands, as shown below in Figure 2.

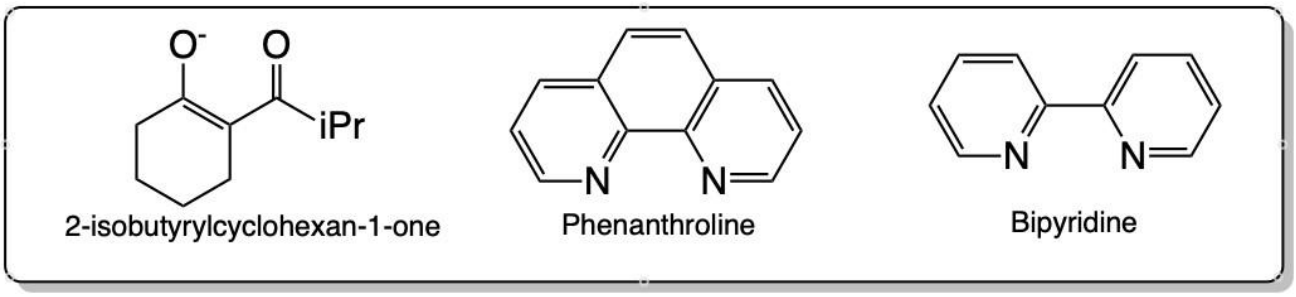


Figure 2: Common O,O and N,N ligands.

When used as a substrate, amino alcohols allow for a tuneable reaction,[4] as they are able to form both a N and an O selective product which can be manipulated to become more likely to yield

either product using ligands, bases and solvents.[9] Literature has displayed the selectivity of several amino alcohol substrates with varying conditions using O,O^- and N,N donor ligands, where there is a strong preference for a N-arylation product when using a charged O,O^- donor ligand and a preference for an O-arylation product with the bidentate N,N type ligands.[4] This selectivity caused by ligand charge suggests that neutral amino alcohols are more likely to bind via the N atom, whereas deprotonated amino alcohols may be more likely to bind via the O atom.[4]

The main mechanistic pathway determined in this research builds on prior structures and pathways suggested by Yu, Jones and many other authors in order to apply this to the addition of amino alcohols as ligands. The pathway considered is shown below in Figure 3 and is explored throughout in more detail, as well as in the final section on mechanism.

It is important to study these reactions as it provides a greater insight of the coordination sphere of copper through different oxidation states whilst being able to determine whether reactions are likely to proceed or not. Ligand exchange can be targeted specifically through computational studies, as one molecule of complex is able to be isolated in order to more accurately assess bond lengths, angles and coordination modes.

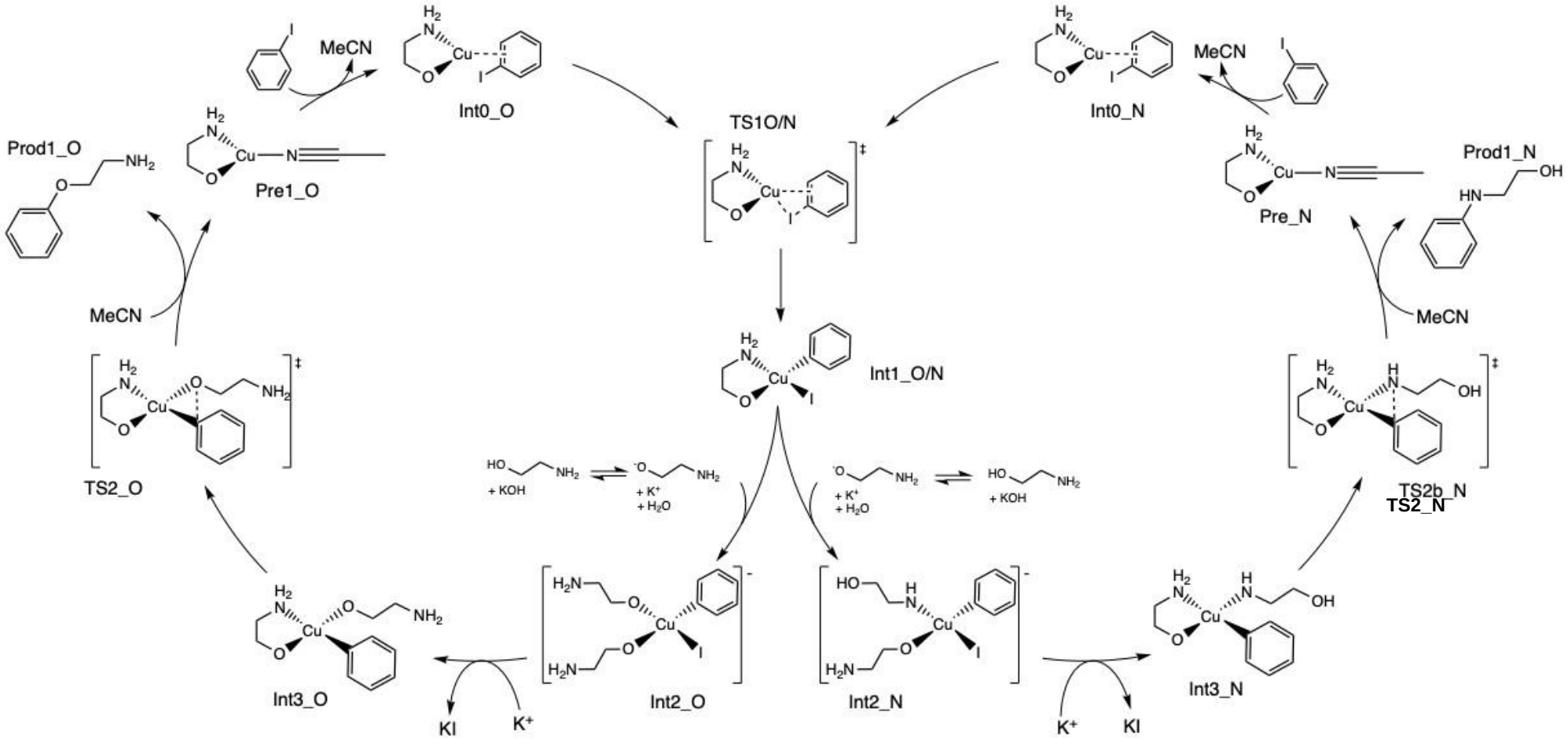


Figure 3: Main catalytic cycle showing N and O selective pathways for arylation of amino alcohols.

As the knowledge base of copper complexes through multiple oxidation states is better mapped out, it has been possible to determine a more accurate prediction for the mechanistic pathways involved. It is important to keep in mind that there is still a gap between the reliability of computational calculations and real-world reactions, as these are only predictions with approximations and do not always prove to be a perfect match for experimental procedures.[34] The main benefit of computational prediction in catalyst design is cutting costs and laboratory time required by determining the most likely result of a reaction mixture. For catalysts made from expensive and rare elements such as rhodium, palladium or platinum, it is essential to minimise the wastage when experimenting. Computational methods that have been used such as B3LYP are far from perfect, however, provide a compromise between accuracy and computational cost.[26]

1.3 - Computational Mechanistic Studies

1.3.1 - Computational methods

Over the last decades, advancements in molecular descriptors through quantum mechanical computation have allowed for reactions to be studied incredibly accurately including the use of catalysts. This planning of reactions is crucial for efficiency in all aspects of lab work.[3, 9] Prediction of reaction pathways enables chemists to narrow down a number of possibilities, calculating whether specific reactions are likely to occur, and almost more importantly, which reactions will not proceed.[36]

There are several derived theoretical methods for the quantum mechanical approximation of the Schrödinger equation, the most rudimentary of which being semi-empirical. These are the lowest cost quantum mechanical methods available and are useful for analysing molecules that are too large for a highly accurate computational method to be applied such as the active site of an enzyme.[37] These methods give the largest compromise in accuracy for a quick calculation, however, the can provide a more useful calculation than molecular mechanics (MM). The use of MM is helpful for treating electrons implicitly and therefore is not quite fit for purpose for polarization or electron transfer, hence MM is not able to be used for modelling chemical reactions in the same way that quantum mechanical (QM) methods are. One benefit of MM in comparison to QM methods is a much lower computer power due to the larger number of assumptions used, which

makes MM a good choice for modelling large structures such as proteins.[38] However, QM methods are useful for protein structures with the downfall of requiring extreme amounts of computer power, therefore smaller discrete molecules are favoured.

The Hartree-Fock approximation is the least computationally expensive, and one of the more inaccurate[39] *ab initio* analytical method used for solving the time-independent Schrodinger equation for open and closed shell systems. This theory relies on describing many one-electron wavefunctions, invoking the Born-Oppenheimer approximation and neglecting electron correlation. This approximation allows for the nuclei of an atom to be treated as a stationary point because the mass of a nuclei is considerably greater than that of an electron, thus a nucleus can be observed to move very slowly relative to an electron. The result of this is the kinetic energy of the nuclei becoming equal to zero. For this reason, and the neglect of previously mentioned electron correlation a more detailed description of the electron density through density functional theory is appropriate.

1.3.2 - Density Functional Theory (DFT)

Early studies by Dirac[40] noted that the exchange energy of a molecule was able to be computed from the charge density, which was further investigated by Kohn, Sham and Hohenberg[41] in order to refine a method for the computation of electronic structures of molecules.[42] Kohn-Sham orbitals were proposed as a model for the charge density of molecular orbitals, and in turn give opportunity to calculate the Mulliken charge of atoms, molecular orbitals of an atom or the electronic structure of a molecule.[41]

DFT can be used for searching for ground and transition state structures of molecules and complexes. Structures can be generated to be optimized with calculated energy values. For small molecules, this can be a quick and efficient method of analysis, however, it requires systematic planning when there are multiple isomers and transition states to locate for one molecule in order to remain a truly efficient process, when a user is able to input accurate geometry guess coordinates, the computational cost is reduced.[43]

In comparison to a Hartree-Fock (HF) level of theory, Density Functional Theory provides a more accurate description of the approximation of the wavefunction, whilst reliably scaling in a similar manner with respect to the number of electrons in a system. The main benefit of DFT methodology is that the functions take into consideration the electron-electron correlation in many-body systems, which is neglected in the HF method. Unlike semiempirical methods, DFT does not calculate the conventional wavefunction, and instead calculates a direct electron density function.[37] There are pure Generalized Gradient Approximation (GGA) and hybrid GGA DFT methods that can be used in order to improve the accuracy of a calculation. These two DFT methods are defined through the functionals used, where pure DFT requires two GGA functionals to describe energy correlation and exchange, whereas the hybrid method uses a mixture of a standard GGA with a part of a Hartree-Fock exchange term.[39] These terms allow for a more comprehensive description of the interaction of wavefunctions between any two given electrons in overlapping orbitals. This overlap of orbitals contributes to the charge density and potential energy of the system.[44] The addition of HF exchange terms is what separates the lower three rungs of the 'Jacobs Ladder'[45] from hybrid DFT methods.

Density Functional Theory has many applications in catalysis, ranging from calculating ionization energies to determining interactions between biological catalysts and proteins. Through the last few decades this has become a powerful tool for prediction and development of kinetic models. DFT brings a useful compromise between accuracy and computational cost, which is sought after in ensuring its use is validated in chemical modelling. Environmental effects such as temperature, concentration and solvents can also be accounted for by using additional input to determine the interaction of solvent molecules with the optimized molecule and monitor the change of kinetic energy in order to modify experiments.[46] The input of a starting geometry can be optimized to a local minimum in order to begin mapping out a reaction; once reactants and products have been identified, it is possible to use these optimized structures to locate transition states between the intermediates which is necessary for mechanistic analysis.

1.3.3 - Functionals

A functional is a function of a mathematical function[37] that is used to describe an initial input function numerically.[47] When categorizing functionals, a Jacobs-ladder approach is used,[45] where climbing a 'rung' denotes an increase in the accuracy of the functionals used, resulting in an improved description of the shape of each orbital in each atom. This ladder analogy defines a system of increasing mathematical descriptors that satisfy conditions of accurately modelling atomic systems, which lowers the number of approximations used.[45] Deciding which functional to use relates to the type of work being undertaken and can be often decided upon based on the time taken and level of accuracy desired. Generalized Gradient Approximation (GGA) DFT functionals such as BP86[48] can be more appropriate for geometry optimizations, and often provide results similar to that of experimental work. GGA functionals are more accurate than Local Spin Density Approximation (LDA) functionals as they reduce energy errors by including terms that are dependent on the spatial variation of the electron density on top of the local density.[45] Hybrid functionals (detailed previously in section 1.3.2) such as B3LYP[14, 19, 49] became a standard for fast computational analysis, and perform well in regard to the description of most atoms, yet can have inaccuracies when modelling large transition metal complexes due to the number of core electrons and commonly under-bind atoms.[14] B3LYP is a functional that is expressed using KohnSham orbitals, which have the same density as the original many-body function.[50] Additional input such as dispersion corrections can be used in order to improve the accuracy of a functional by decreasing the mean absolute error, which are used to more accurately describe the long range electron correlation within a molecule. Dispersion corrections account for part of the van der Waals interaction potential between atoms or molecules.[51] These corrections are able to minimise common under-binding for hybrid GGA functionals and repulsive interactions that are found whilst using LDA and GGA functionals.[52]

1.3.4 - Basis Sets

Basis sets are groups of descriptive functions that are designed to model the wavefunction of molecular orbitals in a way that allows for high levels of flexibility to more accurately model the electron density surrounding an atom. They are designed to have a physical significance, with each

function in a basis set having an origin which can be found within the nuclei of each atom. The functions used in basis sets are a descriptor of the matrix coordinates of an electron and are used to help solve the Schrödinger equation. Common descriptors for orbitals are the Slater-type and Gaussian-type orbital descriptors. The number of basis functions that describe each atomic orbital determines the type of basis set, where minimal basis sets use one basis function (Slater or Gaussian type) per atomic orbital. Some basis sets are referred to as double or triple zeta (ζ), this term is used when there are two or three basis functions used to describe the wavefunction of each atomic orbital. Finally, basis sets can be 'split valence' which use one function for each core atomic orbital, and one or more larger basis function to model the valence atomic orbitals. These basis sets can be altered to use larger numbers of individual functions in order to make optimizations and frequency calculations more accurate, whilst increasing the time taken to perform calculations. Determining an optimal basis set to compromise between time and accuracy is important due to the desire to gain reproducible results in terms of the free energy of copper-amino alcohol complexes

Additional functions can be added to calculation input such as polarization or diffuse functions. Polarization functions increase the flexibility of the orbital description by accounting for the interactions between orbitals as atoms become closer together. These additional polarization functions are denoted with (d) when applied to all atoms excluding hydrogen and (d,p) when being applied to all atoms inclusive of hydrogen. There are additional polarisation functions that build on these foundational d and p functions.

Some basis sets are not ideal for modelling atoms of a certain size due to inaccurate description of the core electrons and because the core electrons do not play a direct role in bonding compared with the valence electrons.[53] A mixed basis set can be applied to a system whereby heavier atoms are modelled using a larger basis set called an Effective Core Potential (ECP), which treats the core electrons together and allows for an accurate description with minimal distortion from the true positions, whilst the lighter atoms are still able to be described using a smaller basis set such as 631G(d). One example of an ECP used for atoms larger than sodium is the Los Alamos National Laboratory 2-Double Zeta (LANL2DZ)[15] ECP, which has been shown to perform well in representing the inner electrons of heavy atoms like gold[54] whilst using a range of different functionals such as

B3LYP and B3P86. Accuracy describing heavy atoms like gold suggests that for the representation of smaller atoms like copper or potassium, LANL2DZ is likely to do an adequate job. Larger basis sets will provide a more accurate description of the wavefunction, therefore the use of the SDD basis set for representation of heavier atoms is likely to perform better than LANL2DZ. The SDD is a triple-zeta basis set with ECPS and therefore, whilst it treats electron density in a very similar manner, accounts for more of the charge density near the nucleus,[55] hence acts as a more accurate ECP.

1.3.5 – Basis Set Superposition Error (BSSE)

When assessing the most commonly used basis sets, the amount of error accumulated from assumptions must be considered. When performing calculations to describe systems with relatively small basis sets such as 6-31G(d), it is common to expect the basis set error (BSE) to be cancelled out as long as the basis set remains the same for comparison of two molecules.[62] The basis set superposition error arises from basis sets, often categorised as Gaussian type orbitals (GTOs), where the wavefunction of a molecule is described from a fixed central nucleus, which suggests that for a finite solution there must be a discrepancy from the true solution due to small basis sets that use a greater number of approximations. Basis functions will overlap in a different manner when using a range of individual basis sets to optimise the geometry of one molecule because of the change in number of descriptors per electron alongside the Hartree-Fock exchange contribution of the functional used. The BSSE varies with interatomic distance, however the separate energy of the atoms does not depend on this distance.[63] One commonly used method for minimising BSSE is to pick a large enough basis set that the complete basis set (CBS) limit is reached.[63] This implies that the number of basis functions per atom is at a maximum for a given computational power, however this is a particularly time-consuming method and is therefore impractical for larger systems. Practically, to work towards the CBS limit is not always computationally efficient or possible, therefore a counterpoise correction (CPC) can be implemented in order to counteract the BSSE.[64] The CPC method requires individual calculation of both reactants together and alone, with then altering the basis set in order to define the energy change caused by the reactant's interactions between molecules.[65] The CPC is larger for transition states and intermediate geometries as two fragments are more likely to influence each other, and therefore is less prominently seen in smaller molecules further apart.[64]

1.3.6 - Solvents

Solvation is an important factor to include in DFT calculations in order to make an accurate model of a real-world reaction. Neglect of solvation in chemical modelling has been referred to as an 'apparent oversimplification.[37] Solvation can be modelled in two ways, explicitly and implicitly. Explicit solvation, also referred to as micro-solvation, physically places molecules of solvent around the molecule being modelled and measures interactions. This method is a large oversimplification as in reality, any molecule larger than a monatomic ion will be surrounded entirely by the reaction medium, not just by 6-10 discrete solvent molecules[37] however, it provides a useful description of the physical solvent-solute interactions on a small scale.

Implicit solvation models as implemented in the software used here (Gaussian)[56] use a self-consistent reaction field (SCRF), which treats solvents as a uniform, continuous dielectric medium that is characterised by each individual solvents' dielectric constant (ε). In this type of solvation modelling, the molecule is placed inside a cavity of spherical/ellipsoidal geometry of which the cavity wall is polarized by the average charge distribution of the solvent within the cavity,[39] hence the name polarizable continuum model (PCM). Figure 4 below shows a physical description of the cavity formed around a molecule reproduced from literature[39] where the van der Waals radii are represented by the spheres centred at the nuclei.

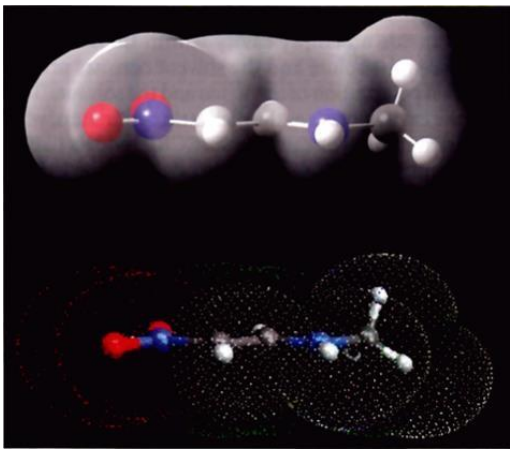


Figure 4: PCM model of a cavity around a molecule.[1]

For larger molecules that are not ideally spherical, this model can prove inaccurate. Several iterations of the PCM model are available to model solvation to a more or less accurate extent, one improved model being the Integral Equation Formalism PCM, or IEFPCM.[39] This model helps to

account for escaped charge from the cavity and has a 'smoothing' effect on the wall of the cavity. This proves relevant as the changing ligand sphere causes a variation in the shape of the complex, and charge must be conserved when deprotonated ligands and substrates are present. The IEFPCM model is of interest in comparison with another solvation model like the Solvent Model Density (SMD). This model has been shown to perform just as well as any PCM model, however, a distinct performance drop was seen whilst analysing aprotic solvents.[57] This gives the IEFPCM appropriate placement for the study of amino alcohols in solvents such as toluene or acetonitrile[4], which are both commonly used.

2 - Methodology

2.1 - Computational Details

Throughout this work, optimizations and frequency calculations were performed using the standard hybrid B3LYP functional[15, 19, 58] alongside the 6-31G(d) basis set for all atoms (C, O, N, H) except for the triple - ζ SDD ECP which was used for Cu, I and K atoms. For some early calculations, the LANL2DZ effective core potential was used instead of SDD, this method will be denoted as BS1, whilst the SDD method will be referred to as BS2.

Single point energy calculations were performed using 6-311+G(d,p) basis set for all atoms (C, O, N, H) and the triple - ζ SDD[59] Effective Core Potential for Cu, I and K atoms (denoted as BS3). The IEFPCM solvation model was used for all calculations in solution. Acetonitrile was the solvent used across all calculations in solution, and has experimental uses with tetradentate ligands to support this choice.[35] Dispersion corrections were added using the empirical GD3 correction method,[60] which uses an increased accuracy in the modelling of intermolecular interactions to prevent the model from becoming too 'loose' as this may cause errors in relative free energies, alongside distorting the bond angles.[46]

All calculations were performed using Gaussian 16W[56], with structures being produced in GaussView 6.1.1.[61] The version of the Cambridge Structural Database used for complex structure searching was ConQuest V2.0.5.[32]

3 - Cambridge Structural Database (CSD) Mining

The CSD is a crystallographic database containing 1.1 million submitted structures from journals and research papers. The search feature of the CSD is largely customisable, allowing you to search broadly or specifically by building queries to fit any task. Structures can be found using their unique 'RefCode', a drawn structure, journal, elements, unit cell or even density. Multiple queries can be layered in order to refine and pinpoint structures. When drawing structures, bond lengths and angles are able to be restricted in order to force the search to remain specific to well determined VSEPR geometries. This is a useful feature as it allows for specific copper geometries to be compiled for a general investigation into the physical properties of copper complexes. It was found initially that copper(I) is highly likely to form 2, 3 and 4 coordinate geometries, which is to be expected as silver and gold, the other coinage metals, both follow similar preferred geometries in their (I) states.[17]

3.1 - Data mining analysis of copper coordination

This investigation was conducted to explore the coordination chemistry of copper, and to gain a greater understanding of N and O bonding to copper complexes, which was compared to Cl, as this is another common ligand. This study sought to gain a more reliable understanding of complexes that may prove more or less favourable when attempting to compare and create catalytic cycles regarding the coupling of aryl iodides to amino alcohols. The Cambridge Structural Database (CSD) was used to mine data bond lengths and bond angles in various copper complexes.[32, 33, 66] This is of interest because it provides crystallographic experimental data on thousands of molecules and compounds, which allows for calculated structures from this project to be compared to crystallographically-determined data averages. With this being said, one of the greatest challenges with comparing CSD data with computationally determined structures is the computational overestimation of metal-ligand bond lengths when using a hybrid functional, which can cause inaccuracies when comparing crystallographic bonds with those optimized with DFT,[67] however, this can depend on the functional used and could be checked using XRD data. Another important factor is the difference caused by the change of phase, as the CSD data is given as solid crystal data whereas many DFT methods solely operate in the gas phase. The hybrid functional overestimation

is counteracted by the use of dispersion corrections. This effect may be further exaggerated when attempting to model a complex with solvation, as many structures from the CSD contain counterions or co-crystallized solvents. Another potential drawback of the CSD being paired with DFT methods is that most DFT functionals do not model dispersion.[67] For this reason, where calculations have been included, dispersion corrections have been employed alongside a B3LYP/6311+G(d,p)/SDD computational method in order to bring these two methods in closer agreement.

The bond lengths of interest were Cu-N, Cu-O and Cu-Cl, the angles looked at were N-Cu-Cl, N-Cu-O and O-Cu-Cl. These bonds and angles are of interest because of the N and O donor atoms found on an amino alcohol, and the presence of halide in the hypothetical rection, this makes these bonds useful for comparing to non-polar and polar ligand donor atoms such as phenanthroline, diketonetype and bipyridine ligands. To understand the trends in bond length and angles in this number of complexes helps to determine more and less favourable geometries and coordination modes of ligands.

The CSD has been used to identify 4-coordinate Cu(I) and Cu(II) structures before, where it was found that Cu(I) will often resort to a tetrahedral crystal structure.[68] A large majority of the Cu(II) complexes found were in a square planar geometry. This is useful as it enables a more in-depth consideration to be given to the tetrahedral geometries when searching for Cu(I) structural characteristics. The CSD only stores oxidation state information when it was provided from the original publication, therefore despite large differences in stereo-electronic preferences between Cu(I)/(II)/(III), a lot of information may not be available for some entries.[68]

3.2 - Methodology

In order to collect the required complexes, several queries had to be set up with restrictions on the constituents and geometry of the specific bonding in the complex. In every structural search, two bonds were always present, in combinations of Cu-N, Cu-O and Cu-Cl to find differences in the geometries and to note the interaction between the donor atoms. The search structures are shown below in Figure 5. The structures largely focus on a Cu(II) centre, especially when considering the

likely oxidation state of an octahedral (6-coordinate) copper complex. This is relevant as there are studies that identify the involvement of Cu(II) in Ullmann-type reactions,[28] therefore being aware of the whole range of oxidation states is useful. However, there are many Cu(I) structures that can comfortably sit in a 4-coordinate geometry, and a more in-depth study of the coordination of complexes with lower coordination numbers is to follow.

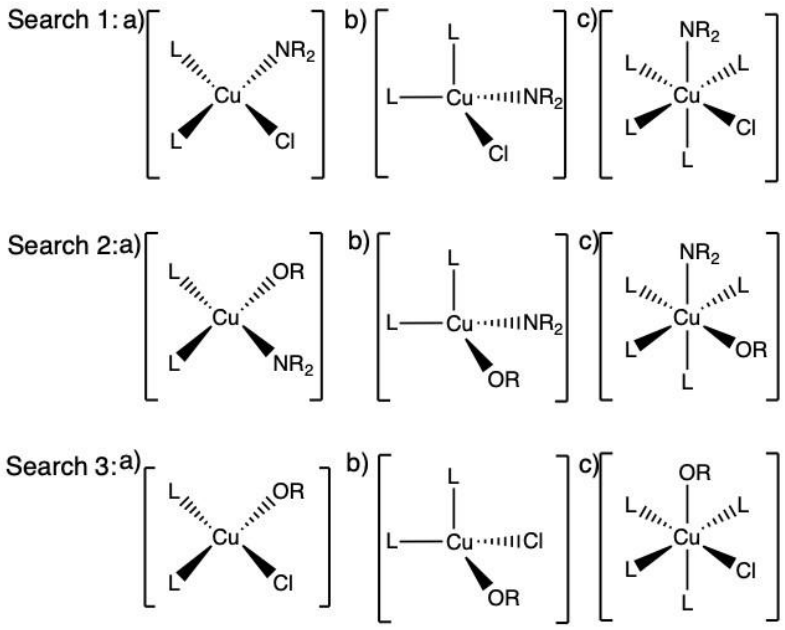


Figure 5: General structures of 4 and 8 coordinate complexes searched for using the CSD with restricted bond angles.

To ensure that only tetrahedral and square planar complexes were found, the coordination of the central copper atom was set to 4. The largest R value of 0.1 was used alongside "organometallic only" in order to get a broad spectrum of complexes. The R-factor describes how well the structure agrees with the experimental model and gives a better refinement for other filters such as excluding disordered or powdered structures. To differentiate between these geometries, the bond angle was restricted to between 88° and 92° for identifying square planar complexes and between 108° and 110° for tetrahedral complexes. These are tight ranges and cause the loss of *trans* isomers in square planar and octahedral complexes, therefore separate searches with restrictions of 170° – 180° were also made to acquire isomers for the six complexes in these geometries.

In order to find octahedral complexes, the bond angle restriction of 88° and 92° was used with a coordination of 6. If this restriction is relaxed then there are many more complexes found, lots of which contain polydentate ligands and heavily distorted octahedral complexes. The bond ranges

were used rather than exact angles of 90° and 109.5° due to the data being reliable yet experimental, which accounted for small discrepancies.

3.3 – Results and Discussion

It is important to note that, whilst the following results display a range of useful information regarding structural properties of crystalized copper complexes, there is initially a large mix of Cu(I) and Cu(II) despite the main focus of this research being Cu(I) based. This is due to the large amount of time required to filter through each search result in order to identify individual oxidation states. Both tetrahedral and square planar geometries were considered because of several copper species that lie somewhere between the two due to steric restrictions.

3.3.1 - Bond Lengths of 4 and 6 coordinate copper

The mean bond length, standard deviation, relative standard deviation (RSD) and interquartile range (IQR) of Cu-N, Cu-Cl and Cu-O are shown below in tables 2-7. A graphical comparison between the Cu-ligand bond lengths of Cl-Cu-N, N-Cu-O and O-Cu-Cl in square planar and octahedral complexes is shown in Appendix Section 12.1.

Table 1:Summary of bond length and deviations Cu-N from search 1.

Geometry	Mean BL / Å	Standard Deviation	RSD / %	IQR
Square Planar Cis	2.001	0.034	1.67	0.04
Square Planar Trans	1.995	0.051	2.54	0.08
Tetrahedral	2.111	0.143	6.75	0.16
Octahedral Cis	2.050	0.131	6.40	0.04
Octahedral Trans	2.063	0.091	4.43	0.07

Table 2: Summary of bond length and deviations Cu-Cl from search 1.

Geometry	Mean BL / Å	Standard Deviation	RSD / %	IQR
Square Planar Cis	2.267	0.061	2.67	0.04
Square Planar Trans	2.236	0.037	1.67	0.03
Tetrahedral	2.362	0.148	6.27	0.16
Octahedral Cis	2.615	0.262	10.04	0.52
Octahedral Trans	2.319	0.150	6.48	0.05

Table 3: Summary of bond length and deviations Cu-N from search 2

Geometry	Mean BL / Å	Standard Deviation	RSD / %	IQR
Square Planar Cis	1.975	0.042	2.84	0.05
Square Planar Trans	1.956	0.046	2.35	0.06
Tetrahedral	2.010	0.111	9.50	0.13
Octahedral Cis	2.026	0.096	11.78	0.04
Octahedral Trans	2.026	0.113	5.60	0.06

Table 4: Summary of bond length and deviations Cu-O from search 2

Geometry	Mean BL / Å	Standard Deviation	RSD / %	IQR
Square Planar Cis	1.923	0.055	5.53	0.05
Square Planar Trans	1.924	0.034	1.78	0.04
Tetrahedral	2.062	0.196	4.75	0.24
Octahedral Cis	2.285	0.269	2.14	0.52
Octahedral Trans	2.009	0.156	7.78	0.05

Table 5: Summary of bond length and deviations Cu-O from search 3

Geometry	Mean BL / Å	Standard Deviation	RSD / %	IQR
Square Planar Cis	1.941	0.062	2.72	0.06
Square Planar Trans	1.969	0.051	2.61	0.08
Tetrahedral	2.008	0.098	4.86	0.07
Octahedral Cis	2.190	0.287	10.23	0.51
Octahedral Trans	2.425	0.325	13.41	0.65

Table 6: Summary of bond length and deviations Cu-Cl from search 3

Geometry	Mean BL / Å	Standard Deviation	RSD / %	IQR
Square Planar Cis	2.255	0.061	2.12	0.03
Square Planar Trans	2.240	0.041	1.82	0.05
Tetrahedral	2.244	0.038	1.73	0.03
Octahedral Cis	2.495	0.255	3.18	0.49
Octahedral Trans	2.623	0.285	10.88	0.56

The average Cu-Cl bond length was greater than the average Cu-N bond length, which was expected due to the large van der Waals radius of a chlorine atom compared to that of a nitrogen atom. However, several outliers were identified, and can be seen in Appendix Section 12.1. These outliers were mostly found in the bond length of Cu-Cl in a square planar geometry and of Cu-N in an octahedral complex. The Cu-Cl outliers may be due to the large variance in the properties of the third and fourth ligands, as this data only concerns a Cu centre that has at least one N and one Cl ligand. A more detailed look at these outliers can be found in section 'Outliers' below.

The results in Tables 2-7 show that for search one and two, square planar geometries lead to the lowest standard deviation, which determines that the Cu-Cl and Cu-N BL in this geometry should have a small range. This is confirmed by a very low IQR in both cases. The IQR of both tetrahedral geometries are larger than SQP, however, the octahedral complexes display a different trend.

Octahedral Cu-Cl and Cu-O each have a large standard deviation and IQR, suggesting that the complexes analysed display a mix of instances where the complexes are distorted by the Jahn-Teller effect. The chlorine ligand is pushed away from the metal centre by repulsion from the high electron density surrounding the equator of the complex as shown below in Figure 6. This tetragonal distortion lowers the energy of the dz^2 orbital and increases the energy of the dx^2-y^2 orbital.[69]

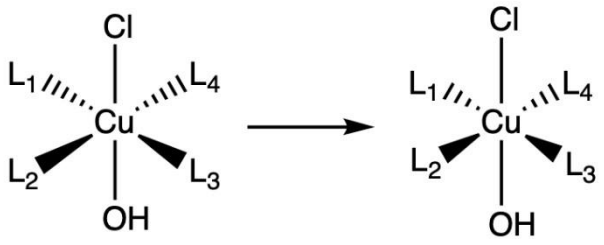


Figure 6:
Jahn-Teller vertical distortion of an octahedral copper complex.

In complexes containing Cu-Cl bonds found by search 3, the tetrahedral complexes had the lowest standard deviation, which suggests that despite a changing ligand sphere, there are not many surrounding atoms that have a large effect on this bond. Being larger, the chlorine atom is more able to distort smaller, less polar atoms surrounding the metal centre. In search 3, the Cu-O bonds in a square planar geometry had the lowest standard deviation, only showing a noticeable difference when O is *trans* to Cl. This demonstrates the *trans* influence of chlorine, causing elongation and sequential weakening of the Cu-O bond.[70] There is a contribution from the polarity of both oxygen and chlorine causing large repulsion when closer together in a square planar or tetrahedral geometry.

The tetrahedral structures from each search can be seen compared in Appendix Section 12.1. This comparison shows that Cl maintains the longest bond length. The average bond length of Cu-N in search 1, where at least one of the surrounding ligands is Cu-Cl was 2.111, which was shortened to 2.010 in search 2, where the Cu-Cl bond was replaced with a Cu-O bond. This is largely due to the smaller size of oxygen allowing nitrogen to sit closer to the metal centre than chlorine.

On average, Cl maintained a greater bond length than either N or O. In octahedral complexes this was more noticeable, due to the Jahn-Teller distortion effect. It is common in octahedral complexes

that two of the larger, more polar ligands sit farther away from the metal centre than the other four, especially in the case of Cu^{2+} which is a d^9 metal centre. This can be seen below in structure FIWFIA in Figure 7 below, where the Cl and O atoms are *trans* towards the top and bottom of the 6coordinate complex.

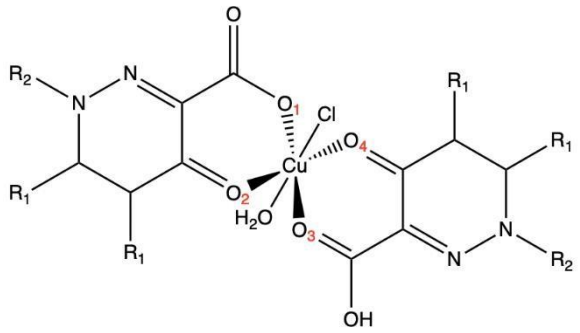


Figure 7: Structure of FIWFIA showing the octahedral central arrangement around a Cu^{2+} centre.

The Cl bond is the longest at 2.604 Å, closely followed by the *trans* Cu-O bond at 2.558 Å, which is a ligated water. The bond lengths are summarised below in Table 7.

Table 7: Bond lengths of Cu-Cl and Cu-O in structure FIWFIA

Bond	**Bond Length (Å)**
Cu-Cl	2.604
Cu-OH_2	2.558
Cu-O (1)	1.998
Cu-O (2)	1.938
Cu-O (3)	2.008
Cu-O (4)	1.932

The average Cu-Cl bond length in octahedral complexes containing oxygen in search 3 was 2.495 Å. Structure FIWFIA shows the effect that highly polar coordination can have on bond length of the trans upper and lower ligands. The average Cu-O bond length in search 3 octahedral complexes was 2.190 Å, which indicates a 0.368 Å increase in bond length when the equatorial ligands are highly polar.

The orientation of many octahedral complexes is such that the Cl or O ligands are situated above and below a planar-like field of the remaining ligands. Jahn-Teller distortion occurs when the d orbitals in a metal complex undergo splitting to reduce the degeneracy and increase the stability of an electronic state[71] as shown below in Figure 8.

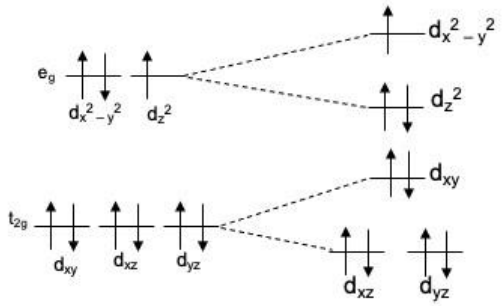


Figure 8: Octahedral d^9 orbital splitting causing Jahn Teller distortion.

One notable observation is that for all ligands except oxygen, in a square planar geometry, the *cis* isomer has a larger average bond length than the *trans* isomer, where the O is *cis* or *trans* to a Cl or N atom. This is due to the consistency in size and electronegativity of the defined ligands in the *trans* position. However, it should be noted that in table 4 the *trans* isomer is only 0.001 Å greater than the *cis* variant which is the smallest difference in bond length average recorded for this geometry.

In complexes containing Cl-Cu-N or O-Cu-N, the average Cu-N bond length remains remarkably similar, however, the bond length between Cu-O is on average shorter than the original Cl-Cu bond, as can be seen in Table 8 and is demonstrated in Table 8 below. The physical change in this structure may be related to the *trans* influence shift caused by the removal of a Cl^- being substituted for an O .[72] The difference in bond length is due to size of the ligand involved in bonding. Alongside this, oxygen is more electronegative than nitrogen, which causes oxygen to form a stronger bond with the central metal ion.

Table 8: Difference between Cu-Cl and Cu-O bond length when Cu is also bound to N.

Geometry	Bond Length Cu-Cl / Å	Standard Deviation	Bond Length Cu-O / Å	Standard Deviation	Difference in Bond Length / Å
Square Planar Cis	2.267	0.061	1.923	0.055	0.344
Square Planar Trans	2.236	0.037	1.924	0.034	0.312
Tetrahedral	2.362	0.148	2.062	0.196	0.301
Octahedral Cis	2.615	0.262	2.285	0.269	0.330
Octahedral Trans	2.319	0.150	2.009	0.156	0.310

There were outliers in the datasets for square planar Cu-O and octahedral Cu-N. These outliers are unlikely to be significant as the IQR is 0.05 and 0.04 respectively, which shows that the majority of datapoints fall within this boundary. The RSD is 5.53% and 11.78% which shows that a large amount of datapoints fall clustered towards the mean value for Cu-O but in the case of Cu-N, there is in fact a larger variability. This is related to the R groups bonded to the N donor. Smaller, less electron withdrawing groups such as hydrogen will allow the nitrogen to form shorter and stronger bonds with the Cu metal centre, whereas larger electron withdrawing groups or groups bonded to the nitrogen will cause longer, weaker bonds. The coordination sphere around the central copper metal is occupied with O, N and Cl atoms with some larger dimer structures containing Cu-Cu bonding.

The IQR of all square planar complexes is the lowest of all the geometries analysed, and frequently shares the lowest standard deviation of BL across all Cu-X bonds taken into consideration. The RSD of all square planar complexes is low, giving reliability to the closeness of the data to the mean value.

The octahedral complexes with Cu-O and Cu-Cl bonds displayed a much larger range of bond lengths than complexes in a square planar or tetrahedral geometry, whereas the Cu-N bonds remained with a small range. This reinforces earlier suggestions of a strong Jahn-Teller effect. However, when the main ligands were Cl and O, the range of Cu-O bond lengths was still large. This may be due to the structural differences, but there are far fewer data points for structures

containing both a Cu-O and Cu-Cl bond in tetrahedral geometry, which is reflected in this difference and the lack of outliers.

The final consideration with regard to octahedral and square planar complexes, is to question at what point can we consider the axial distorted bonds of a tetragonal distorted structure to become dissociated. For large atoms like Cl, the average bond length (2.623 Å) is greater than that of either N (2.026 Å) or O (2.009 Å). Any measured bond lengths above a certain tolerance may be considered as dissociated, or a weak interaction. In O and N atoms, the tolerance is likely to be 2.72.8 Å, whereas with Cl atoms the tolerance for bond association is likely to be 3.1 Å, with bond radii of these atoms suggesting a weak interaction still apparent at 3.2 Å and 3.4 Å respectively.[73] A better understanding of the bond length where a bond can be considered dissociated can be determined by performing energy calculations of a molecule with a set bond length between the two atoms being considered at bond length intervals over a range. This increase in internuclear separation changes the associated bond energy until the bond is broken, where a large decrease in energy could be seen.[74] This was useful to understand as it allows for intermediate structures in the hypothetical mechanism for the arylation of amino alcohols to be analysed, with a greater ability to recognise when a bond is unlikely to be present in an optimized structure.

3.3.3 - Bond Angles of 4 and 6 coordinate Copper

The mean bond angles of several complexes were calculated to determine the average measured angle results for N, O and Cl in experimental crystallographic data of organometallic complexes. The results can be seen below in Table 10.

Table 9: Mean bond angle and deviation of N-Cu-O, N-Cu-Cl and O-Cu-Cl in organometallic complexes

Geometry	Angle	Mean Angle / Degrees	Standard Deviation	RSD / %
Square Planar	N-Cu-Cl	90.30	0.98	1.09
	N-Cu-O	90.28	1.15	1.27
	O-Cu-Cl	90.37	0.99	1.10
Tetrahedral	N-Cu-Cl	109.03	0.58	0.53
	N-Cu-O	108.89	0.60	0.55
	O-Cu-Cl	108.97	0.59	0.54
Octahedral	N-Cu-Cl	90.02	1.14	1.27
	N-Cu-O	90.08	1.11	1.23
	O-Cu-Cl	89.97	1.17	1.30

The tetrahedral complexes show the smallest deviation from the mean of the dataset, yet the furthest deviance from the expected value of 109.5°, however, this is made up for with an RSD of 0.53 – 0.55%, showing that this is an accurate representation of the whole dataset. This, however, stands correct for the small amount of data used and would potentially be larger if the dataset were larger in using a wider range of acceptable bond angles.

The octahedral complexes have the greatest standard deviation; however, they also display the closest to true expected octahedral bond angles of 90°. It is likely that there is some larger deviation from the mean as these complexes have the largest RSD, which is unsurprising as the coordination sphere defined allowed 4 of the 6 ligands to be 'any' ligand. The conclusions that can be drawn from this are that the bond angle is more likely to have a smaller variance when in a tetrahedral arrangement. This is due to the larger bond angle allowing for the ligands to be more spread out. In square planar and octahedral geometries, the ligands are closer together, and when these ligands are large or highly electronegative, electron repulsion causes distortion.

From these calculations, regarding the catalytic nature of copper, it is important to consider the benefits of the *cis/trans* geometries of ligands surrounding the metal centre. The coordination

sphere of Cu(II) is flexible and will comfortably accommodate a 4 or 6-coordinate geometry. Both oxygen and chlorine ligands have a large polar and steric effect on surrounding bonds, especially in the case of non-polar ligands such as nitrogen donors, for example phenanthroline or bipyridine. This is an important finding as it allows for prediction of the number and type of ligands that can be involved in the studied catalytic cycles and helps determine that there may be a preference of the ligands in the coordination sphere when there are polar and non-polar ligands in the reaction mixture, which is shown through the difference in bond length when comparing Cu-N to Cu-O. This is also interesting as it suggests that when considering the selectivity of the product of the reaction, the orientation of ligands and substrates may change based on the polarity. Elongation of bonds may allow for lower energy barriers to overcome when predicting the mechanism of a reaction along with the ability to determine more favourable geometries for amino alcohols to bind with a copper centre, however this elongation must be watched due to the possibility of dissociation.

This work can be extrapolated and allows for the two oxidation states focussed on in literature to be studied to a greater extent. It is relevant to have a good understanding of all the common oxidation states of copper to appreciate the many possibilities when deciding on a mechanism.

3.4 - Bond Lengths of 2 and 5 coordinate Copper

There are suggestions in literature that Cu(I) is most kinetically stable in a 3-coordinate geometry,[4] however, in the CSD it has been found that there are enough 2-coordinate Cu(I) structures to justify further investigation.[17] It is possible that when copper(I) is bound to amino alcohol ligands, there may be low energy structures that contain two amino alcohol units that may be favourable in a monodentate 2-coordinate geometry. This is possible as in literature, optimisations have led to dissociation of one end of an amino alcohol.[4] This may be useful as the use of two ligands bound in a monodentate fashion allows for a more linear starting geometry, which may allow for an aryl halide nucleophile to approach the complex with less steric hindrance. This is appropriate to explore due to the uncertainty surrounding the exact order that the key steps along the catalytic pathway are likely to proceed in, where this study will analyse specifically the Cu(I) and Cu(III) metal centres as they are the redox pair that have been suggested in literature to accommodate the most likely reaction pathway for copper catalysed coupling reactions.[4] The results of this analysis are

discussed below. The 2-coordinate geometry is likely to be more advantageous when using amino alcohols where the nitrogen and oxygen donor atoms are relatively far away from each other, such as long straight chain aliphatic amino alcohols such as 5-aminopentan-1-ol, as used in literature as a substrate.[4]

3.4.1 - Methodology

Three searches were carried out using ConQuest's search builder.[75] In each case, the coordination of the central copper metal was set at 2. The measured 3D parameter for all queries was the average Cu-N and Cu-O bond length. Diagrams of each of the search queries are shown below in Figure 9.

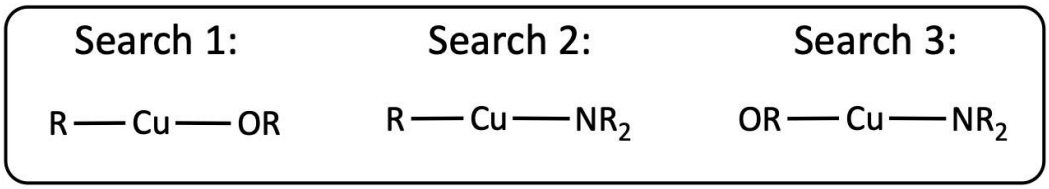


Figure 9: Search queries used in the CSD for 2-coordinate copper complexes.

All three queries restrict the Cu to being bound to two atoms, where searches 1 and 2 are the N or O atom in question, and R is 'any atom'. The final query was used to determine the breadth of experimental submissions having only a N and an O atom bound to the central copper ion. The search criteria for the three queries include filters ensuring that all search results are organometallic, without errors and within below a specified r-factor.[75] For the search needed in this study, a low R-factor was not largely important and was set at R = 0.1 which regards the crystallographic search results as 'fair' quality in comparison to experimental results.

3.4.2 - Results

The average bond lengths from searches in the CSD are shown below in Table 11. The range of data for each search is shown in Figure 10.

Table 10: Average Cu-N and Cu-O bond lengths in 2-coordinate complexes

Bond	Number of search results	Average Bond Length / Angstrom	Standard Deviation
X-**Cu-O**	306	1.849	0.023
X-**Cu-N**	722	1.878	0.038
O-**Cu-N**	25	1.891	0.069
O-Cu-N	25	1.873	0.060

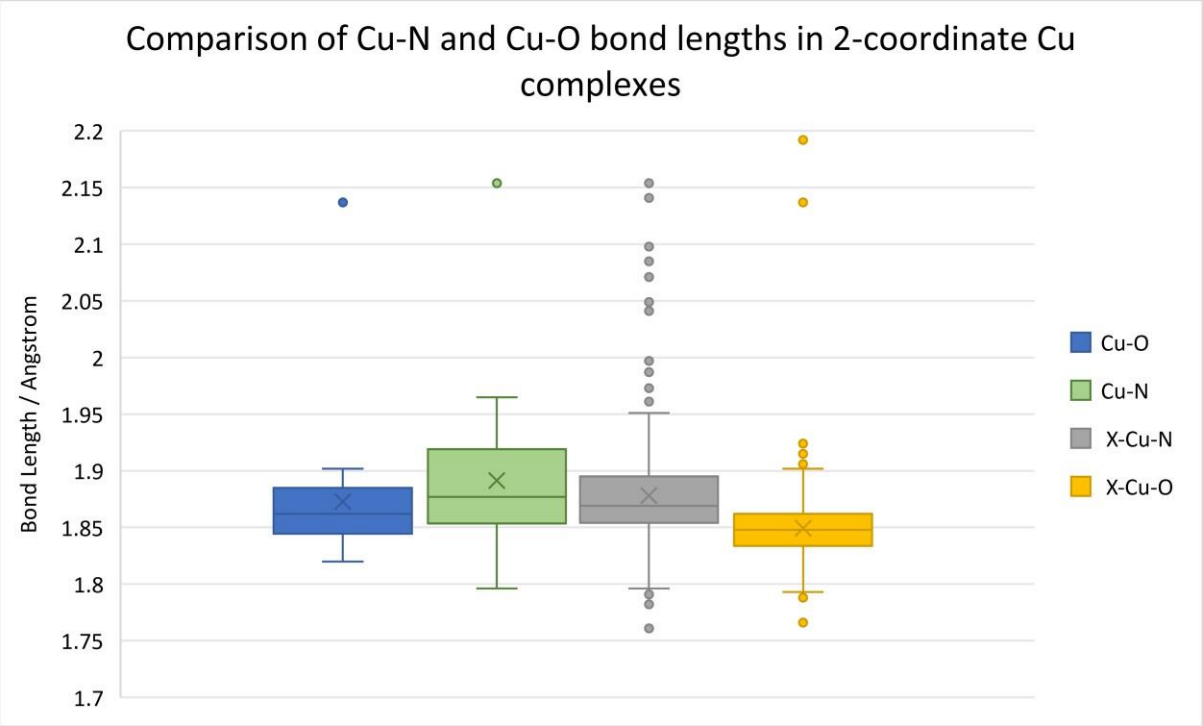


Figure 10: Diagram comparing the entire dataset ranges of Cu-N and Cu-O bond lengths.

The search results show that despite there being hundreds of 2-coordinate Cu structures in the CSD that contain either Cu-N (722) or Cu-O (306) bonds with 'any' atom also bonded with the copper, there is a lack of data surrounding 2-coordinate structures with both bonds present, with only 25 found in the search using the same search specifications. One factor affecting the number of results is the R-factor, which was set to the highest tolerance of 0.1 for using this filter. If this filter is set to a lower value of <0.05, there will be a lower number of results.[75]

The average Cu-N bond length was greater than that of Cu-O in both cases, and in the larger dataset of search 2, the average Cu-N length was greater than that of Cu-N in search 3. The expected

outcome was a longer average Cu-O bond length; however, this was not seen in this dataset. The search 1 Cu-N average bond is greater than the Cu-O bond of search 2 by 0.029 Å, however the search 3 Cu-N average bond length is only 0.009 Å longer than the search 3 Cu-O bond. The standard deviation from both bonds measured in search 3 are twice that of the standard deviation in search 1 or 2, which is likely due to the far fewer datapoints.

There are several outliers in search 3 for the bond length of Cu-N. The interquartile range of the CuN and Cu-O bond lengths from search 3 were lower than that of the IQR in searches 1 and 2. This is due to a much smaller dataset, where a small number of results deviating from the mean bond length will behave as outliers. Several of the search results found were not the same type of structure that is relevant to this work, however, determining a scope for the likelihood of these structures being feasible is important, as only a small amount of literature[76] explores the possibility of 2-coordinate intermediates, even though they are entirely possible with small hemi-labile ligands.

The final oxidation state of copper considered is Cu(III). Literature[76] suggests that for Cu(III) to be feasible, the copper metal centre will likely be 4 or 5-coordinate. While a pentacoordinate Cu(III) can find arrangement in both a square-based pyramidal and trigonal bipyramidal geometry, it has been observed that the 4-coordinate square planar geometry is often preferred,[76] despite a trigonal bipyramidal geometry usually being more stable.[77] The geometry of these can be seen below in Figure 11.[77]

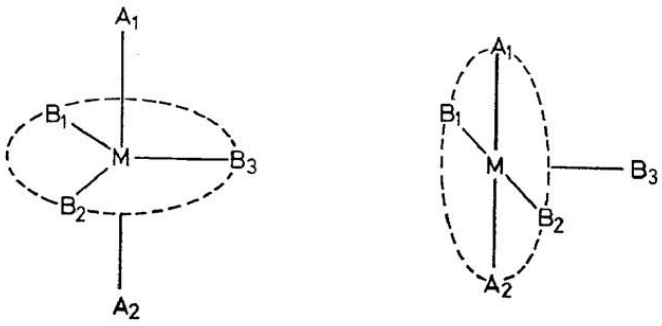


Figure 11: 5-Coordinate geometry showing trigonal bipyramidal (left) and square based pyramidal (right).

Copper in the 3+ oxidation state is difficult to reliably interpret from the CSD, as there are not many structures deposited, however a contribution to this lack of structures may be due to poor

categorization of these structures. A study of the coordination geometries of the d-block element ions showed that at most, there were varying numbers of Cu(III) structures depending on whether 'Cu^{3+}' or 'Copper(iii)' were searched, showing another level of unreliability within the CSD oxidation state search .[78] It was found that of the Cu(III) structures located, a large majority were in a square planar geometry for the 4 coordinate geometries with square-based pyramidal being the next most common for 5 coordinate structures. This contributes to the difficulty in determination and isolation of a Cu(III) complex geometry.[68] Finding several Cu(III) structures in a square planar arrangement is expected due to the known stability of d^8 transition metals in this geometry, which is useful for being able to explore 4-coordinate Cu(III) pathways for a mechanism including amino alcohols, as they may be hemilabile which would allow for greater flexibility in the coordination number that may not have been observed with previous bidentate ligands such as phenanthroline.

The main outcomes of these searches show that there are a large range of available copper structures to identify and that attempting to use postulated Cu(III) intermediates within mechanistic studies is well justified. The exact coordination number of these intermediates will be explored further computationally as the searches did not provide a definitive insight into whether the preferred coordination sphere in this higher oxidation state is likely to be 4- or 5- coordinate. The CSD has helped to identify trends in bond lengths between copper and N/O atoms, which is beneficial moving forwards for determining the arrangement of amino alcohol ligands in the space around the metal centre. These bond lengths can be compared with those in calculations to see how they change with the coordination spheres implemented for the coupling of aryl iodide with an amino alcohol.

4 - Method Validation

It is essential to determine a method of computation that will allow for accurate determination of relevant energy values for this research with the least impact on realistic data values for analysis and interpretation. There are several variables within the computational method that can be altered in order to increase and fine-tune the level of accuracy and efficiency achieved. Basis sets, solvation, dispersion and functionals are the main variables focussed on in this work, as they enable the computational method to be developed to a much higher accuracy, as highlighted in Section 2. There is a need to begin at a less expensive level of theory (BS1) and make changes such as inclusion of larger basis sets or solvents, in order to highlight the need for method validation. Many ideas and progressions in these methods are drawn from literature[4] and the compromise between accuracy and computational cost is considered throughout due to the need for efficiency yet wanting to ensure that research undertaken is comparable to literature through using similar and identical computational methodology.[14, 15] One working cycle studied in this work was chosen in order to compare the energetic and geometric changes that altering the computational method can have. This chosen cycle uses a 3- and 4-coordinate Cu(I)/Cu(III) pathway and is shown below in Figure 12 named Cycle A.

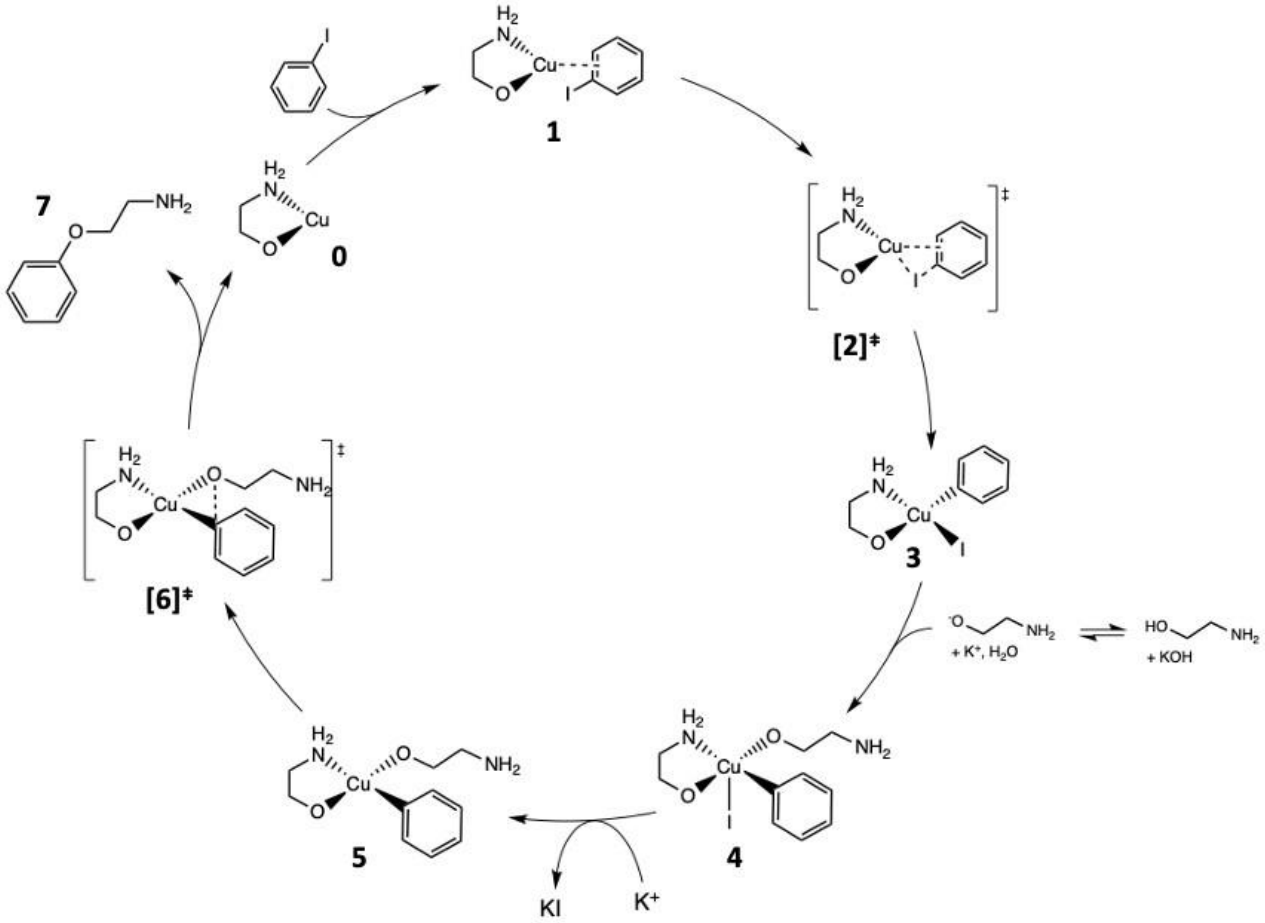


Figure 12:Hypothetical Cycle A.

4.1 - Solvation Effects

Common solvents used for copper catalysed coupling such as toluene, DMF and acetonitrile (MeCN) have been discussed in section 'Solvents'. To provide a greater understanding of the computational method, one catalytic cycle in development (Cycle A) was subjected to the addition of implicit solvation using MeCN to compare to calculations in the gas phase. This is a polar solvent, which is appropriate for use with the KOH base chosen. MeCN is comparable to solvents commonly used in experimental work such as DMF as they have similar dielectric constants (ε) of 36.64 and 36.70, respectively.[5] In these solvents, involvement of charged species is favoured[4] which is appropriate for this system where there is a very high likelihood of the amino alcohol ligands being anionic in solution.

4.1.1 - Energy Effects

To determine the effect of solvation on the energy of each optimized structure in Cycle A, single point energies were carried out using BS2 level of theory with a self-consistent reaction field, using the IEFPCM model for implicit solvation using acetonitrile. Table 12 below shows the comparison of the relative energies of BS2 alone and BS2 with solvation. The optimized geometries were found using Gaussian at the BS2 level of theory.

Table 11: Energy change caused by adding the IEFPCM solvation model to BS2 for Cycle A.

Species from Cycle A	Relative energy of Cycle A BS2 / kcal mol^{-1}	Relative energy of Cycle A BS2 + Solvent / kcal mol^{-1}	Difference in energy caused by adding solvent / kcal mol^{-1}
0	0.00	0.00	0.00
1	-8.32	-6.08	+2.24
[2]‡ *cis*	-1.93	-0.13	+1.80
[2]‡ *trans*	-0.35	0.91	+1.26
3 *cis*	-10.86	-12.50	-1.64
3 *trans*	-17.16	-14.95	+2.21
4	-61.01	-33.69	+27.32
5 *cis*	-45.07	-60.32	-15.25
5 *trans*	-54.00	-56.32	-2.32
[6]‡ *cis*	-36.62	-46.49	-9.87
[6]‡ *trans*	-40.43	-49.14	-8.71
7	-71.24	-77.78	-6.54

As shown above in Table 13, there is an overall energy drop of -6.54 kcal mol^{-1} when adding solvation when calculating the relative energy of the cycle with geometries at the same coordinates. This demonstrates the extent that adding solvent to the reaction is able to thermodynamically stabilize the overall reaction and lead to a reaction that may proceed under less harsh conditions to a lower energy product, despite having a small increase in required input

energy for the first half of the cycle. The *trans* isomer of **3** can be seen to have a lower relative energy (-14.95 kcal mol^{-1}) than the *cis* isomer of **3** (-12.50 kcal mol^{-1}), however, solvation caused a greater stabilization effect on the *cis* isomer, as this structure was seen to drop in energy by -1.64 kcal mol^{-1}. The largest energy drop was that of **5**-*cis*, which is a 4-coordinate Cu(III) structure. This structure is energetically favourable, and the involvement of solvent causes a particularly large stabilizing effect on the amino alcohol groups. Once past **4,** all later intermediates displayed a drop in energy as a result of adding solvation, the large increase in energy of **4** may be due to the decrease in charge from neutral to -1 when the deprotonated ethanolamine substrate is introduced. An interesting calculation to further see the effects of solvation include using explicit solvation to understand more closely the interactions between low coordination number geometries with acetonitrile, and whether there are extra intermediates that bind well with solvent, but that lay outside of the scope of the present project.

4.1.2 - Geometry Effects

Alongside energy effects, the geometry effects of adding implicit solvation to calculations was assessed using re-optimization of the BS2 geometries previously calculated. Table 13 and 14 below show the bond lengths of all Cu-X bonds throughout hypothetical Cycle A, whilst Table 15 summarises the average bond length of each atom in relation to copper, and how this was affected by the addition of solvation.

Table 12: Bond length of each bond to copper for all intermediates in Cycle A at BS2 level of theory.

Species	Ligand Bond Length / Angstrom		Substrate Bond Length / Angstrom		
	Cu-O	Cu-N	Cu-O	Cu-Ar	Cu-I
0	1.885	2.084	-	-	-
1	1.889	2.108	-	2.025	-
[2]‡ *cis*	1.870	2.011	-	1.931	2.502
[2]‡ *trans*	1.858	2.053	-	1.924	2.535
3 *cis*	1.881	2.189	-	1.985	2.680
3 *trans*	1.943	2.083	-	2.014	2.743

4	1.870	-	1.872	1.935	2.605
5 *cis*	1.887	1.975	1.815	1.910	-
5 *trans*	1.840	2.076	1.812	1.908	-
[6]‡ *cis*	1.915	2.012	1.878	1.929	-
[6]‡ *trans*	1.852	2.187	1.871	1.970	-

Table 13: Bond length of each bond to copper for all intermediates in Cycle A at BS2 level of theory with solvation.

Species	**Ligand Bond Length / Angstrom**		**Substrate Bond Length / Angstrom**		
	Cu-O	**Cu-N**	**Cu-O**	**Cu-Ar**	**Cu-I**
0	1.883	2.084	-	-	-
1	1.882	2.103	-	2.011	-
[2]‡ *cis*	1.877	2.168	-	1.978	2.685
[2]‡ *trans*	1.941	2.071	-	1.992	2.732
3 *cis*	1.867	2.015	-	1.923	2.498
3 *trans*	1.855	2.059	-	1.919	2.532
4	1.868	-	1.867	1.928	2.594
5 *cis*	1.886	1.982	1.810	1.920	-
5 *trans*	1.838	2.077	1.815	1.913	-
[6]‡ *cis*	1.911	2.015	1.876	1.927	-
[6]‡ *trans*	1.848	2.176	1.864	1.961	-

Table 14: Average ligand and substrate geometry effect of adding solvation to the BS2 methodology.

Bond length	**Ligand Bond Length / Angstrom**		**Substrate Bond Length / Angstrom**		
	Cu-O	**Cu-N**	**Cu-O**	**Cu-Ar**	**Cu-I**
BS2	1.881	2.078	1.850	1.953	2.613
BS2 + Solvation	1.896	2.051	1.858	1.958	2.619
Difference	+0.015	-0.027	+0.009	+0.005	+0.007

The average bond length of all bonds increased, except for the bidentate Cu-N ligand bond. Solvation had the greatest bond elongation effect overall on the bidentate ligand bonding, which became largely distorted in the presence of acetonitrile. The general increase of each of the bonds may have allowed for the ligand N-donor atom to become closer with the copper centre, as this donor is bound to the ligand O-donor which experienced bond elongation. The Cu-O bonds were expected to be influenced by the addition of acetonitrile due to its polar nature, which was seen more in the ethanolamine ligand than in the monodentate substrate. This suggests that the bite angle and bidentate nature of the ligand also caused for the Cu-O bond to shift. As can be seen in Table 14 and 15, the bond length of the Cu-I in transition states **[2]**‡ *cis* and *trans* increased due to the addition of solvation, however the intermediates **3** *cis* and *trans* formed after this TS experienced a decrease in Cu-I bond length.

Overall, the addition of solvation caused a change in the bond lengths of all atoms. This may be due to the stretching interactions between the solvent cavity and the outer atoms of the molecule, extending the bonds through electrostatic attraction to the solvent. The change in bond length was not large in most cases, with the largest being a decrease of -0.027 A of the bidentate amino alcohol Cu-N bond length. This was likely affected most due to the influence of the solvation model on the deprotonated Cu-O ligand bond, where the continuum model of solvation improves the screening of the negative charge. Structures **0**-**4** experienced an increase in free energy, whilst structures **5**-**7** were seen to experience a decrease in free energy from the addition of implicit solvation. The MeCN caused these structures to become more stable through interactions between the polar solvent and the reducing intermediates. Solvation is an important factor to include with final calculations in this study to have a more accurate model of the real-world reaction, which is performed in solution.

4.2 - Dispersion Corrections

The input of additional parameters such as solvation and dispersion to calculations slightly increases the computational cost of the optimization, however, are necessary to ensure calculation outputs are sufficiently accurate.[31] It should be noted that several of the past studies[4] did not include dispersion corrections, therefore comparison to studies will include some error where comparisons between similar structures may result in a larger relative Gibbs free energy difference than

expected. Dispersion corrections are a necessary feature of any computational study and can lead to relative energy changes of near 10 kcal/mol in the energy surface of organometallic reactions.[79] These corrections describe the weak interactions caused by the variance in the movement of electrons in and around a molecular system. Stand-alone DFT functionals such as B3LYP do not take the fluctuation of charge distribution into account, which gives rise to the need for empirical dispersion corrections.[39] Single point energy calculations of the optimized intermediates in Cycle A were calculated by adding Grimme's empirical dispersion correction D3 to the original BS2 conditions. D3 dispersion is a function that takes into consideration the interatomic distances and interaction energy of electrons.[80]

4.2.1 - Energy Effects

Separately from solvation, the effect of dispersion corrections on the energy of optimized intermediates was also investigated. The BS2 approach previously described was used with empirical dispersion applied to the input and reoptimized. The difference in energy of the intermediates in Cycle A with dispersion corrections is shown below in Table 15.

Table 15: Energy effects of the addition of dispersion corrections to the BS2 methodology.

Species	Relative energy of Cycle A BS2 / kcal mol^{-1}	Relative energy of Cycle A BS2 + Dispersion / kcal mol^{-1}	Difference in energy caused by adding dispersion / kcal mol^{-1}
0	0.00	0.00	0.00
1	-8.32	-26.48	-18.16
[2]‡ *cis*	-1.93	-19.50	-17.57
[2]‡ *trans*	-0.35	-17.92	-17.57
3 *cis*	-10.86	-29.69	-18.83
3 *trans*	-17.16	-35.93	-18.77
4	-61.01	-112.97	-51.96
5 *cis*	-45.07	-74.64	-29.57
5 *trans*	-54.00	-66.67	-12.67

[6]‡ *cis*	-36.62	-58.45	-21.83
[6]‡ *trans*	-40.43	-60.88	-20.45
7	-71.24	-69.80	1.44

The implementation of dispersion corrections allowed for a more accurate measure of long-range electron correlation to be calculated.[51] One notable trend in the data is that dispersion corrections decreased the relative energy of each individual intermediate by 17-21 kcal mol^{-1}. This more accurate description of the atoms in given space is evidence that empirical dispersion is important to increase the chances of studying a more realistic model of a catalytic cycle. Moving forward, empirical dispersion will be added to calculations alongside the previously mentioned solvation model to form another 'best case' methodology. The energy decrease observed for structure **4** was -51.96 kcal mol^{-1}, which represents an extremely large change in free energy of this intermediate, greater than any in the cycle. This energy change may suggest that there is a large instability in this structure and that some of the bonds in the structure are not likely to remain strong interactions because of the variability of the geometry caused by external electrostatic interactions.

4.2.2 – Geometry Effects

Alongside energetic differences, adding dispersion corrections to more accurately model the charge distribution fluctuations is likely to have some effect on the geometry of the reaction intermediates. The BS2 bond lengths are as previously shown in Section 4.2.2 – Table 14. Shown below in Table 16 are the Cu-X bond lengths of ligand and substrates with dispersion corrections implemented.

Table 16:Bond length of each bond to copper for all intermediates in Cycle A at BS2 level of theory with dispersion.

Species	Ligand Bond Length / Angstrom		Substrate Bond Length / Angstrom		
	Cu-O	Cu-N	Cu-O	Cu-Ar	Cu-I
0	1.883	2.084	-	-	-
1	1.882	2.103	-	2.011	-
[2]‡ *cis*	1.877	2.168	-	1.978	2.685
[2]‡ *trans*	1.941	2.071	-	1.992	2.732
3 *cis*	1.867	2.015	-	1.923	2.498
3 *trans*	1.855	2.059	-	1.919	2.532
4	1.868	-	1.867	1.928	2.594
5 *cis*	1.886	1.982	1.810	1.920	-
5 *trans*	1.838	2.077	1.815	1.913	-
[6]‡ *cis*	1.911	2.015	1.876	1.927	-
[6]‡ *trans*	1.848	2.176	1.864	1.961	-

One of the more notable observations from assessing the changes in bond length as a result of incorporating dispersion corrections in to the BS2 calculations of Cycle A, is that with dispersion the copper-iodine bond length is significantly longer in transition states **[2]**‡ *cis* and *trans*, which is the same trend observed with the addition of solvation. These changes were significant, with the average bond length change caused from the addition of dispersion corrections was - 0.004 Å. In the case of structures in Cycle A, this shows that the addition of dispersion corrections causes the structures to tighten up and constrict the bond angles from the overestimation without. This was the expected trend to be noticed.

The average bond length and difference caused by adding dispersion is summarised below in Table 17.

Table 17: Average ligand and substrate geometry effect of adding dispersion to the BS2 methodology.

Bond length	Ligand Bond Length / Angstrom		Substrate Bond Length / Angstrom		
	Cu-O	Cu-N	Cu-O	Cu-Ar	Cu-I
BS2	1.881	2.078	1.850	1.953	2.613
BS2 + Dispersion	1.878	2.075	1.846	1.947	2.608
Difference	-0.003	-0.003	-0.003	-0.006	-0.004

As seen above, there is a small but clear decrease in the average bond length of all ligand and substrate Cu-X bonds across Cycle A which evidences the slight overestimation of bond lengths from the standard B3LYP/BS2 method used. The largest change was the Cu-Ar bond, which may be due to the large amount of electron density found in the aryl group. Compared with solvation, the addition of dispersion has made a smaller change to the bond lengths, and a more even distribution of this bond shortening is apparent.

4.3 – Functional Effects

The effects of changing the density functional were assessed through comparison of two functionals. The originally used B3LYP functional was compared with BP86 in order to observe the difference that a hybrid functional can have on a molecular system compared to a standard generalized gradient approximation (GGA) functional. Both functionals contain the Becke's functional contribution and are differentiated by the addition of either Perdew's 1986 (P86) correlation function or the Lee, Yang and Parr (LYP) correlation function.[52] B3LYP also contains exact exchange functions.

The relative energy pathway of Cycle A was assessed using both functionals at the same geometry. This surface was calculated in order to show that despite changing the DFT functional, the energy level trends through the catalytic cycle remain very similar, as each structure is being described in the same way for each functional. Below in Figure 13 and Table 19 are the energy levels of Cycle A whilst using each functional alongside 6-31g(d)/SDD.

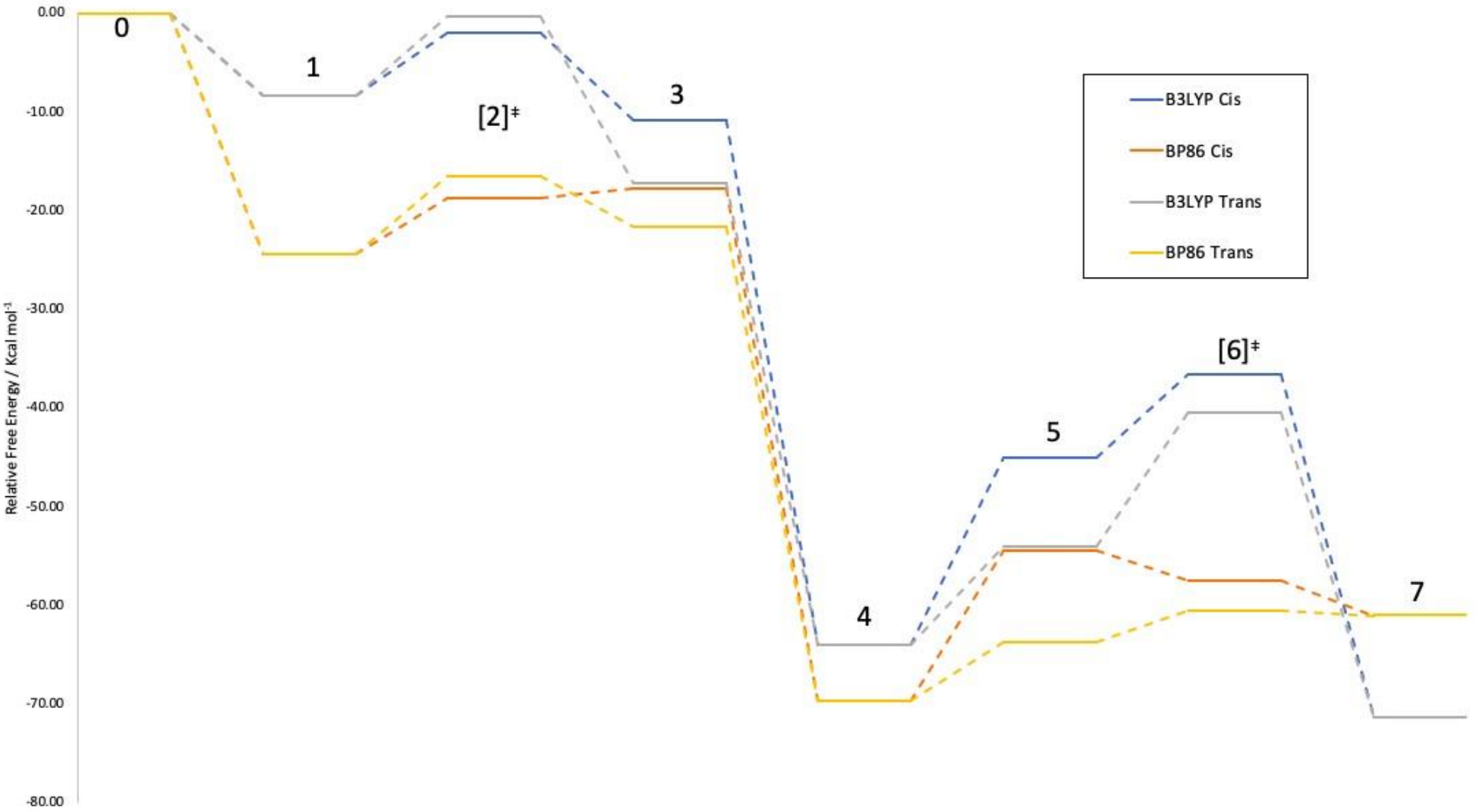


Figure 13: Comparison of Cis and Trans pathways for Cycle A using B3LYP and BP86 functionals.

Table 18: Comparison of free energy of Cycle A using the B3LYP and BP86 functionals in the gas phase – no dispersion corrections.

Species	Relative energy of Cycle A B3LYP / kcal mol^{-1}	Relative energy of Cycle A BP86 / kcal mol^{-1}	Difference in energy / kcal mol^{-1}
0	0.00	0.00	0.00
1	-8.32	-24.33	-16.01
[2]‡ *cis*	-1.93	-18.74	-16.81
[2]‡ *trans*	-0.35	-16.52	-16.17
3 *cis*	-10.86	-17.70	-6.83
3 *trans*	-17.16	-21.66	-4.49
4	-61.01	-69.62	-5.69
5 *cis*	-45.07	-54.48	-9.41
5 *trans*	-54.00	-63.68	-9.68
[6]‡ *cis*	-36.62	-57.47	-20.85
[6]‡ *trans*	-40.43	-60.46	-20.03
7	-71.24	-60.99	10.25

In this cycle, it was seen that overall B3LYP gave results with a less negative relative free energy. The same trend was observed throughout the calculations and cycle, with **[2]**‡ *cis* being the more energetically favoured structure to form, however, after this point onwards where there were isomers, the *trans* isomer was energetically favoured. This is likely due to a change in the coordination sphere, where in **[2]**‡ the aryl iodide is present, and the iodine atom is most likely to be coordinated *cis* to the oxygen donor atom. For both BP86 and B3LYP, energetically the geometries are more favoured when the oxygen donor atoms are *trans* to the iodine ligand, leaving the nitrogen donor atom *trans* to the aryl group. This is likely due to the equal *trans* influence of the oxygen donors paired with the small *trans* influence of the NH_2 donor, causing a stabilizing influence on the aryl group. It can be seen that there is an overall increase of +10.25 kcal mol^{-1} for the overall reaction, despite all B3LYP intermediates having a less negative relative energy throughout. The trends have stayed the same using both functionals, with transition states gaining

energy in both cases regardless to the overall energy drop caused by using BP86 for most intermediates.

4.4 - Changing the Basis Set

One study regarding basis sets was conducted to compare the differences between basis set sizes, to assess what differences in accuracy may be seen from increasing the number of basis functions used to describe each atom. The 6-31G(d) and 6-311G+(d,p) basis sets were chosen for comparison, as they are both well documented methods. The 6-31G(d) basis set is a double-zeta basis with polarization functions applied to all heavy atoms, whereas the valence triple-zeta 6-311G+(d,p) basis set includes polarization applied to all atoms including hydrogen and has a larger number of functions applied to each atom. The latter set also includes diffuse functions applied to all heavy atoms, which is an important property to consider when computationally describing a system with hydrogen bonds or ligands as there is often overestimation of binding energy, which can especially be seen when using charged ligands.[60] In the following work, 6-31G(d)/LanL2DZ is referred to as
BS1, with 6-31G(d)/SDD being referred to as BS2. In cases where 6-311+G(d,p) was used with SDD, BS3 has been used. The cycles explored using basis set alteration were cycles O and N, which have been repeated below for clarity in Figure 14. These cycles are iterations of Cycle A, where the selectivity and most efficient route to product formation has been considered and geometries have been re-optimised using BS2.

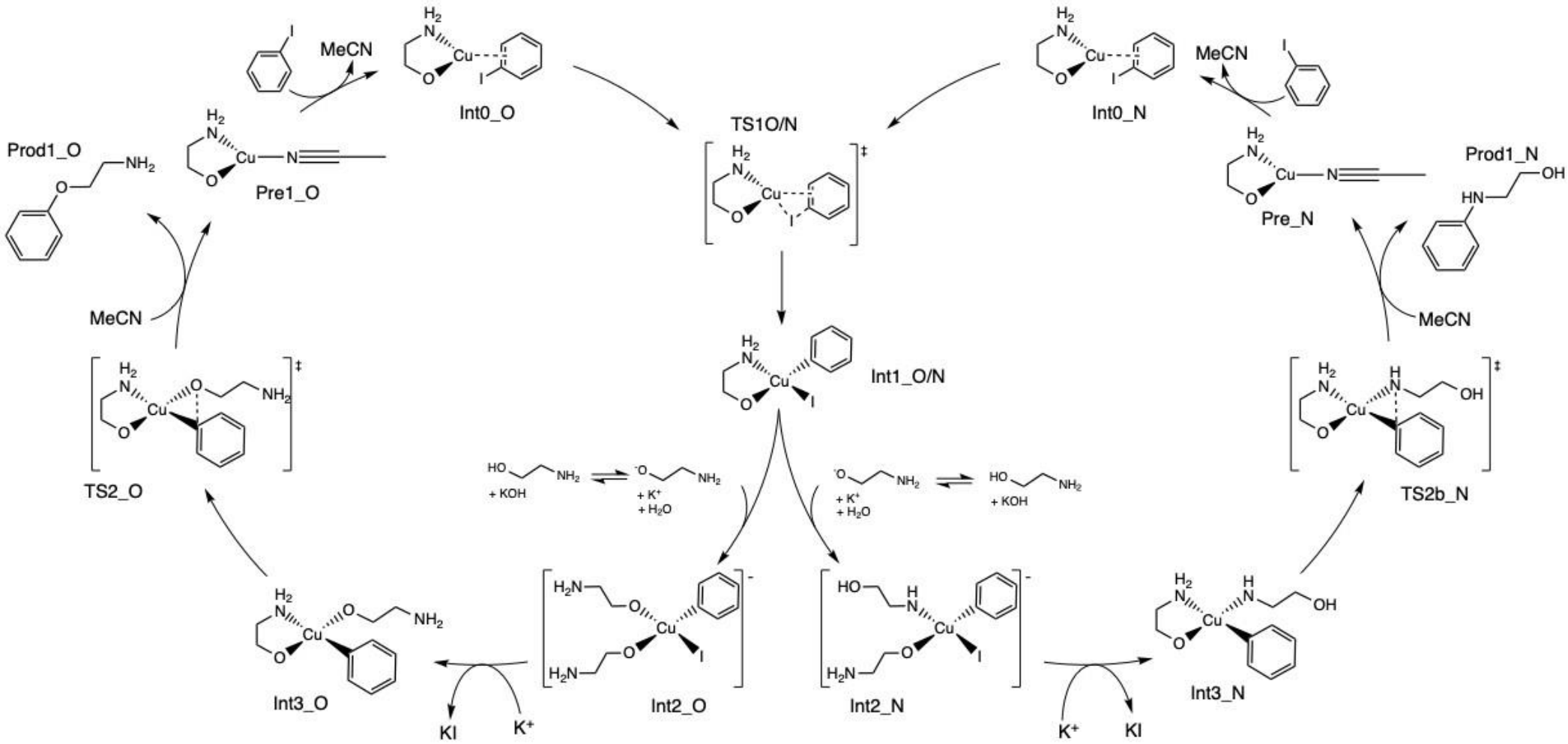


Figure 14: Hypothetical catalytic cycle O and N

4.4.1 - Effective Core Potential

In the original calculations performed, the Los Alamos National Laboratory 2-Double Zeta (LANL2DZ)[15] basis set was used to model the ECP of Cu, I and K atoms, creating a mixed basis set when combined with 6-31G(d). This set up will be referred to as Basis Set 1 (BS1). An alteration made in order to improve the accuracy of the mixed basis set used in this study was to use the SDD[59] basis for ECP modelling. The SDD basis set is triple-zeta quality and therefore larger than the previously mentioned LANL2DZ, which has the compromise of being more flexible, despite being more computationally demanding.[55] The interchange of LANL2DZ for SDD creates another standard which will be called Basis Set 2 (BS2). Yang[15] *et al.* explored the mixed basis set of 6-31+G(d,p) with LanL2DZ used as a basis set for transition metal complexes, which was shown to perform to a greater level of accuracy than 6-31+G(d,p) alone, evidenced by a lower heat of formation energy when compared across multiple functionals,[15] however, in the case of the B3LYP functional, the standalone 6-31+G(d,p) basis set had a lower margin of error[15] than the mixed basis set. Diffuse functions were not initially used in this study to describe the systems in question, increasing basis set size in DFT increases computational cost as between N^3 and N^4 where N is the number of basis functions.[15]

4.4.2 – Results

There were clear trends found when changing the basis set used, and throughout this exploration, the relative energy of each cycle was found to follow the same pathway with altering the basis set causing a shift in the entire pathway. Shown below in Figure 15 is a comparison of Cycle O and Cycle N using BS2 and BS3, where BS3 calculations are single point energies using the same geometries that were optimised in BS2.

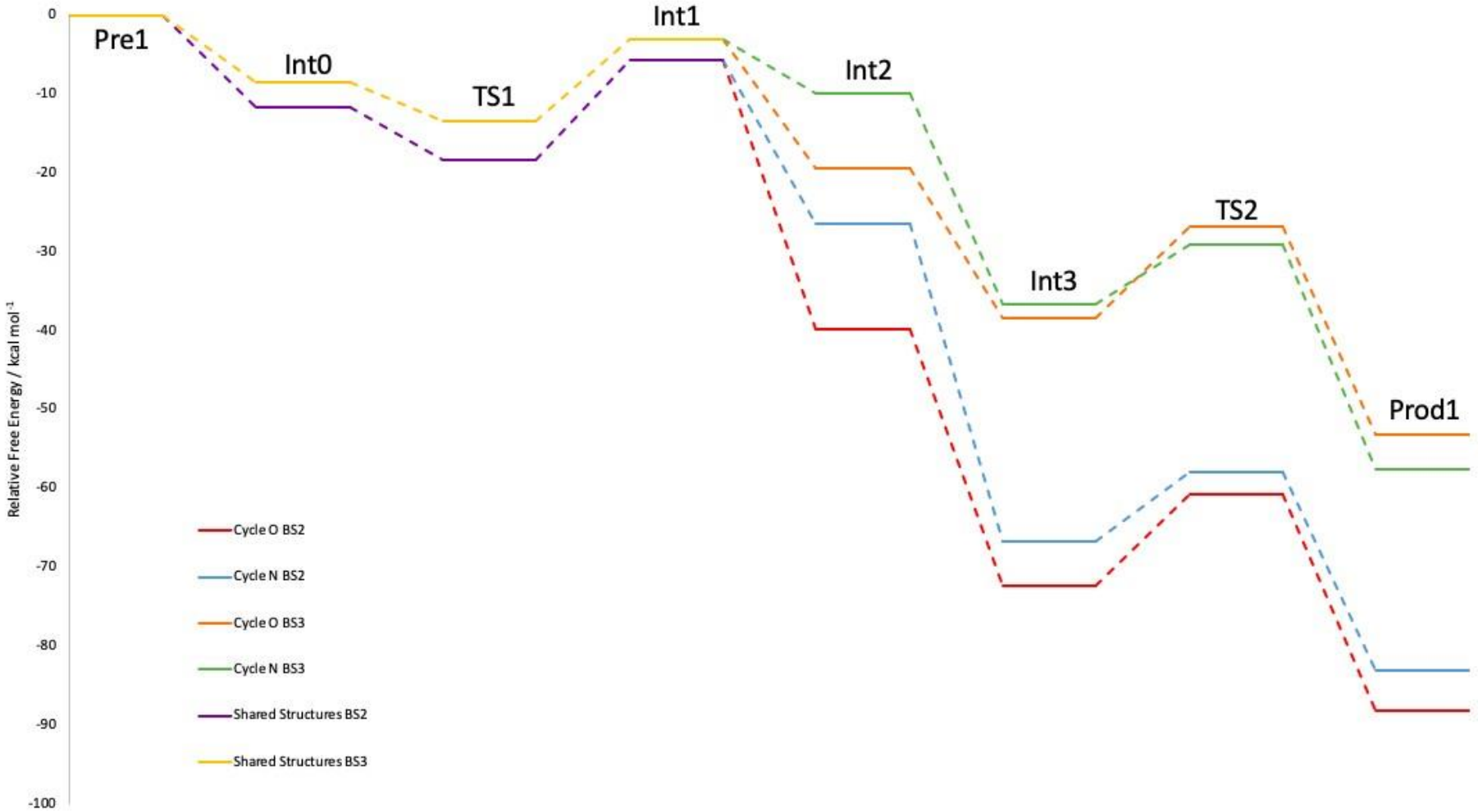


Figure 15: Effects of changing the basis set from 6-31G(d) to 6-311+G(d,p) in comparison with Cycle O and N.

The above figure shows that overall, the larger basis set causes an upward shift in the overall reaction free energy, with each intermediate becoming less negative. The structures that were affected the most were those under TS2. The change of basis set caused TS2_O to become higher in energy than TS2_N, which must be due to the basis set as this was the only variable changed. In all cases other than this, Cycle N was the higher energy pathway, however, was still feasible. The basis set change to BS3 also looks to have brought the O and N pathways closer together in energy compared to BS2, which is likely to be due to the increase in functions used to describe the wave function and shows that BS2 was less accurate. The energy change after Int1 were much larger than the changes prior to Int1, which shows that BS3 had a large impact on these intermediates. The overall slope of the reaction pathway is not as steep using BS3, therefore BS2 may have largely overestimated the relative free energies of the cycle.

Overall, the best method for calculating these energies is the Single Point Energy (SPE) BS3 method of B3LYP with 6311+G(d,p)/SDD using solvation and dispersion corrections, based on the geometries optimised at BS2. These energies tend to be higher than those using the BS2 method, which is likely due to the increased level of accuracy used to describe the wavefunction through the larger basis set with additional overstabilisation caused from the larger basis set. The addition of diffuse functions (+) to the method gives a better description of the atomic orbitals that are positioned farther from the nuclei of the atoms than without, probably improving the modelling of anionic ligands/sites.[60]

5 - Coordination of Cu(I) complexes

Copper(I) throughout literature is often suggested to be 3-coordinate.[2, 4] From the CSD mining previously shown in Section 3, Cu(I) is likely to be found in a 2- or 3-coordinate geometry. As shown in this CSD work, copper(II) has a large range of possible square planar, tetrahedral and octahedral geometries giving rise to geometric isomers. Cu(I) is more electron rich than Cu(II), due to it having a d^{10} electron configuration. When in a 4-coordinate complex, Cu(I) is most stable in a distorted tetrahedral or square planar arrangement, which will be explored in other stages of the catalytic cycle. Two main cycles are to be studied with N and O arylation pathways separated for selectivity.

Copper in complexes **1a** and **1b** is initially expected to be 3-coordinate, which requires the amino alcohol ligand/substrate to be bound once bidentate and a second amino alcohol monodentate to the Cu centre. Literature gives evidence for the ease with which a C5 amino alcohol is able to become monodentate,[4] with less evidence for it acting as a bidentate ligand. The bidentate nature of the amino alcohol is possible due to it containing both O and N donor atoms. Calculations show that the 3-coordinate Cu(I) complexes are likely to readily optimise to a trigonal planar geometry, as shown below in Figure 16. The measured amino alcohol used for this research was selected because the bidentate bite angle between 80-90° is very similar to that of other bidentate phenanthroline and diketone type ligands (~81°),[81] however the tight range gives evidence that a tetrahedral geometry may not be preferred.

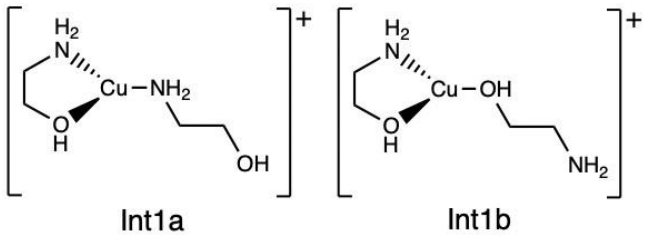


Figure 16: 3-Coordinate Cu(I) complexes with Mono and Bidentate amino alcohols.

Past calculations have given evidence for the possibility of 2- and 3-coordinate Cu(I) complexes.[4] During the process of optimising several initial structures of 3-coordinate Cu(I) complexes, it became common for the bidentate ligand to become dissociated at one end, especially when the

starting geometry had not been refined. The dissociation was recognised as a result of measured bond lengths of over 3 angstroms being present in the optimisations. This suggests that there are low energy structures that contain two monodentate amino alcohol units which may be favourable in a 2-coordinate geometry.

This may be useful as the use of two ligands bound only once each allows for a more linear starting geometry, which allows for an aryl halide nucleophile to approach the complex with less steric hindrance. The 2-coordinate complex may be useful for oxidative addition, as there is a lower electron count in the coordination sphere, increasing this number from 14 to 16 electrons may increase the stability of the complex. There is also more freedom of rotation around the central metal ion, which allows for the nucleophile to distort the starting complex with little repulsion from the linear position to a trigonal planar or square planar structure. This is appropriate to explore due to the uncertainty regarding the geometries that are accessed through this mechanistic pathway. There are possible routes that are more feasible with a 2-coordinate starting geometry as there is a greater chance for aryl groups to bind, however, the 2-coordinate geometry seems elusive due to the opportunity for a stable 3-coordinate complex to be formed. The results of this analysis are discussed below.

The hypothesised mode of coordination to the copper centre is via two monodentate amino alcohol ligands. There are three ways that these amino alcohols can be arranged around the metal ion, as shown below in Figure 17.

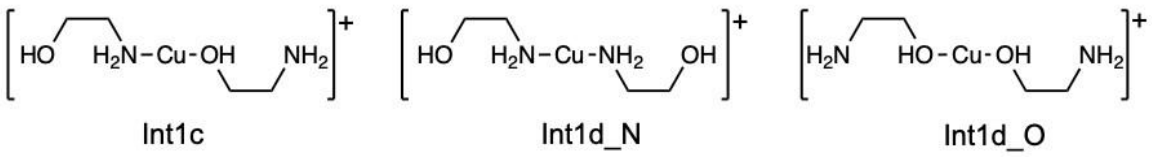


Figure 17:Structures of 2-Coordinate Cu(I) geometries with monodentate amino alcohol ligands.

The bis-monoligation of these ligands allows for donor atom combinations of dinitrogen and dioxygen which is similar to that in literature[4, 5] where the amino alcohol is able to coordinate with copper from either end, whilst allowing to experiment with the effects of only using two donor

atoms. Yu *et al.*[4] focussed on 3-coordinate starting geometries that were Cu(I) complexes bound to one ligand with either the amino alcohol substrate, or an adduct was formed with a dihaptic bond to the aryl ring of the aryl halide. CSD data determined that there were more Cu(I) structures that were 3-coordinate, especially when compared with only 25 structures found that were 2coordinate containing O-Cu-N bonds. Despite this, the 2-coordinate geometries may be difficult to isolate and in a reaction mixture are possibly hemilabile, causing a combination of 2, 3 or even 4 coordinate resting states for these complexes to sit in, which gives some key geometries that can be explored to a greater extent throughout this project.

5.1 - Results and Discussion

Table 20 below shows the relative free energy of fully protonated 2-coordinate and 3-coordinate Cu(I) complexes to compare the coordination mode of amino alcohol ligands to this metal centre, the structures of which have been shown previously in Figure 15 and Figure 16.

Table 19:Free energy of 3-coordinate Cu(I) complexes - bidentate amino alcohol ligand.

Complex	Relative Free energy / kcal/mol
1A	+27.98
1B	+36.72
1C	+2.83
1D_N	**+0.00**
1D_O	+9.30

In the calculations performed in this study, the 3-coordinate Cu(I) structure with a nitrogen bound substrate **1A** was seemingly more stable than the oxygen bound counterpart, however, this was based on gas phase free energy calculations alone using the BS1 level of theory. The alternative to this was achieved using deprotonated complexes as the change in charge may alter the preferred method of coordination. This was partially expected due to previous work by Yu *et al.*[4] that suggests a 3-coordinate oxygen bound amino alcohol has a greater free energy (+15.4 kcal/mol) than that of the 3-coordinate nitrogen bound complex. Despite small inaccuracy, this trend was also seen in this work with the oxygen bound complex **1B** having a free energy +8.74 kcal/mol

greater than Int1a. This is not an ideal comparison as although the ligands used by Yu *et al.* are bidentate, one is β-diketone type, with two oxygen donor atoms and a net charge of -1 and the other frequently ligand used is a neutral phenanthroline type. There is a lack of literature in the region of amino alcohol ambidentate ligands with Cu(I) in comparison with N,N and O,O donor ligands, therefore, to compare amino alcohols with ligands that have similar donor groups and properties is a sufficient place to begin analysis. Another difference is the chain length of the amino alcohol used, as Yu used a 5-carbon amino alcohol substrate, which may cause a difference in charge distribution when deprotonation is induced, which is likely to influence the bond length. Calculations could be done to confirm the effect of increasing the chain length of the amino alcohol ligand on the Gibbs free energy of individual structures, however, they were not performed in this research. When a ligand is used with two nitrogen donor ligands, there is less evidence to suggest that the nitrogen-bound amino alcohol is preferred, as the literature intermediates suggest a 3coordinate structure with a deprotonated oxygen amino alcohol substrate is more favourable.[4]

The 2-coordinate complexes, when compared to others containing a single bidentate ligand, will provide a useful assessment of the copper centre coordination sphere, as the large van der Waals radii of iodine and steric bulk of benzene may be better accommodated when the metal centre is not constrained by a bidentate ligand alongside the second bound amino alcohol substrate, allowing the amino alcohol chain to move freely. The hemi-lability of amino alcohols has not been explored, however, may be regarded as important as a mechanism is developed.

A further calculation to examine the coordination of Cu(I) replaced the amino alcohol ligand with bipyridine, a well-known bidentate ligand. When the starting geometry was arranged to be bidentate in a trigonal planar orientation, with the N donor atoms equidistant from the copper centre, the optimized geometry showed that the bipy ligand became monodentate to the copper centre. This was verified by recording a difference in the Cu-N bond length of each donor atom of 0.30752 Å, caused by the addition of the amino alcohol. Bipy is a particularly good ligand to use as it shares a very similar bite angle and chain length as the short amino alcohol already tested. The bite angle of Bipy was calculated to be 76.8° and the bite angle of ethanolamine was seen to be 80.1°. In the calculations, a bipy ligand was bound with a Cu(I) centre excluding an amino alcohol

and was found to bind in bidentate fashion, with the Cu-N distances remaining equal. More information about monodentate and bidentate ligand identification can be found in section 'ligand exchange'. The structures and bond lengths of the complexes discussed can be seen below in Figure
18.

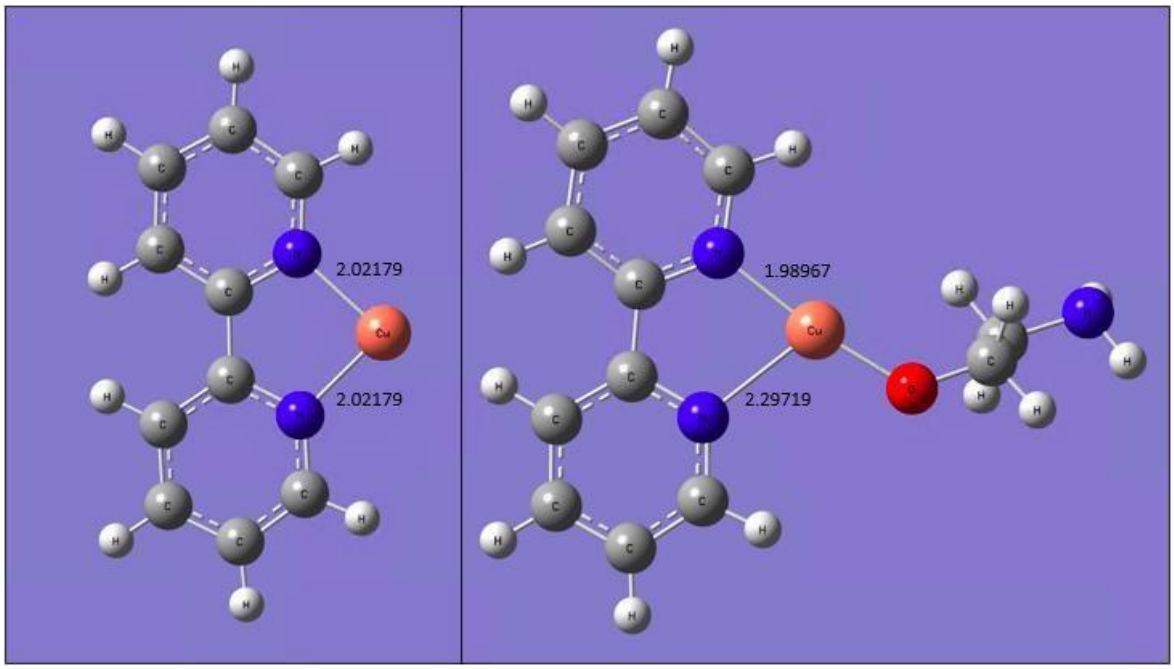


Figure 18: Comparison of the bond length of bipyridine (angstroms) to Cu(I) with and without an amino alcohol substrate.

There is a weak interaction between one of the nitrogen atoms in bipyridine when bound to Cu(I) and an amino alcohol, the optimized geometry shows that the amino alcohol is forcing a distorted trigonal planar geometry as it attempts to become linear to one of the N donor atoms. This shows that neutral bidentate ligands may be undesirable when using amino alcohol substrates, as their bidentate nature is distorted so greatly. This was not seen when fully protonated C2 amino alcohols were used as the bidentate ligand. These geometry optimizations were useful as they present information about the binding nature of Cu(I) and show that common bidentate ligands may not bind in the expected way, and not equally from each donor ligand.

6 - Deprotonation

Calculations were performed to monitor the deprotonation of Cu(I) complexes in the gas phase and in solution using acetonitrile (MeCN) or Toluene (Tol) as the solvent at the previously described 631G(d)/LanL2DZ (BS1) level of theory. These calculations were deemed necessary in order to determine whether a deprotonation step was needed in the cycle, which would be facilitated by the addition of a base such as KOH. The deprotonation step of this reaction is most likely to occur before or after the oxidative addition, but before the reductive elimination.[4] Another important conclusion to draw is that if the amino alcohol is deprotonated, whether it is likely to deprotonate before or after it is bound to the Cu(I) centre. Work from Yu *et al.*[4] suggests that a deprotonation is likely to occur, and that it is most probable to happen once the amino alcohol substrate is already bound to the copper metal centre, however, calculation of the proton affinity of the amino alcohol, alongside the deprotonated complexes will provide evidence for the likelihood of the deprotonation occurring. Optimisations of the amino alcohol were carried out and compared with the complex deprotonation, which in the scope of the overall mechanism is possible and potentially favourable due to this causing the complex to become more electron rich, increasing the nucleophilic characteristics of the deprotonated intermediate[82] which aligns well with previous studies suggesting that this type of arylation reaction is likely to proceed beginning with the coordination of a nucleophile.[9]

Another important reason to determine how likely a deprotonation is to occur is to assess what role a base may play in this reaction. In literature common bases suggested for use in Ullmann coupling reactions are NaOH, KOH, K_2CO_3, Cs_2CO_3.[23] The solvents chosen for these calculations were used in order to compare a polar and a non-polar solvent, which is relevant for mapping the mechanism of a reaction as the solvent may cause a deprotonation step to be more or less favourable depending on the effect on the net charge of the species, however, the solvation model used in this work is implicit, therefore the solvent-molecule interactions are averaged and not explicit in the optimized geometries.

The gas phase calculations were performed first to determine the atom in the Cu(I) complex most likely to be deprotonated. Calculations were performed using several versions of two 3-coordinate Cu(I) complexes to examine whether the most likely deprotonation would change depending on the orientation of the amino alcohol substrate, the two fully protonated complexes in question with labelled protons are shown below in Figure 19.

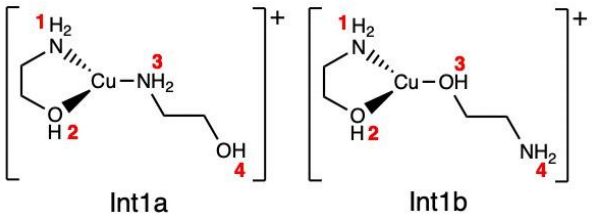


Figure 19: 3-Coordinate Cu(I) complexes with labelled protons

This deprotonation step may be favourable because it allows the complex to gain a neutral overall charge. It was hypothesised that the first single deprotonation of the complex in question was most likely to happen to one of the nitrogen or oxygen atoms of the ligand or substrate. This calculation considers the 3-coordinate complexes and is a comparison of the individual possible deprotonation sites relative to the fully protonated starting complexes in Figure 19. In l iterature,[82] it was suggested that a Cu(I) complexed to an amino alcohol substrate using a N, N donor ligand was considerably more stable when the substrate was bound at the oxygen atom due to the more acidic nature of an -OH group compared with that of an -NH_2 group.

6.1 - Methodology

Calculations were run in the gas phase, with the intention of recalculating the lowest energy, most favourable structures using solvation at theory level BS1. The solvation method used was the Integral Equation Formalism Polarizable Continuum Model (IEFPCM), in which the focus solvents were toluene and acetonitrile as these solvents are well documented in literature and give a comparison between polar and apolar solvent effects on the system.

For calculations of the fully protonated Cu(I) species shown above in Figure 20, additional constraints were necessary in order to prevent the complex from fully dissociating. The

'opt=modredundant' function was used with distance constraints in order to ensure that all three bonds to the metal centre were maintained during the geometry optimization. The impact of introducing the constraints on the optimisation of the 3-coordinate complexes was an easier location of the local minimum, which aided in finding the lowest energy species once the constraints had been removed. Another result of this is that the bond lengths of the amino alcohol ligand are likely to be at their limit, and shows that moving forward, a deprotonated complex may be necessary to facilitate 3-coordinate Cu(I), as the charge is able to remain the same throughout.

6.2 - Results

Copper complexes with coordination numbers of 2 and 3 were assessed in fully protonated, singly deprotonated and doubly deprotonated iterations. The results below summarise the complexes, with singly and doubly deprotonated intermediates denoted with -H and -HH respectively. Table 20 shows a comparison between deprotonated 2-coordinate and 3-coordinate Cu(I) complexes with single and double deprotonations. This represents the energy difference between all measured complexes before undergoing oxidative addition in the gas phase. The structures drawn in Figure 20 show the 2-coordinate monodentate structures optimized. The energies below are relative to the zero point of structure **1D_N** shown previously in Table 19.

Table 20: Relative Free energy 2 and 3-coordinate Cu(I) complexes at BS2 level of theory.

Complex	Relative Free energy / kcal/mol
1A-H1	+10.15
1A-H2	-10.58
1A-H3	-3.81
1A-H4	-3.80
1B-H1	+12.53
1B-H2	-5.86
1B-H3	-11.53
1B-H4	-15.02

1C-H1	+3.59
1C-H2	-14.15
1D_N-H1	-2.72
1D_N-H2	+2.34
1D_O-H1	-9.61
1D_O-H2	-8.29
1C-HH1	+50.00
1C-HH2	+49.98
1C-HH3	+50.76
1D_N-HH1	+67.20
1D_N-HH2	+58.50
1D_N-HH3	+72.80
1D_O-HH1	+41.95
1D_O-HH2	+42.32
1D_O-HH3	+43.72

Table 22 shows a comparison of the average fully protonated and singly deprotonated bond lengths in 3-coordinate complexes.

Table 21: Summary of mean bond lengths of 3-coordinate Cu(I) complexes.

Complex	**Average Bond Length / Angstroms**			
	Cu-O_{Ln}	**Cu-N_{Ln}**	**Cu-O_{Sub}**	**Cu-N_{Sub}**
Int1a	1.799	1.824	-	1.870
Int1b	1.799	1.824	1.870	-
Int1a-H	2.027	1.985	-	1.940
Int1b-H	2.027	1.985	1.940	-

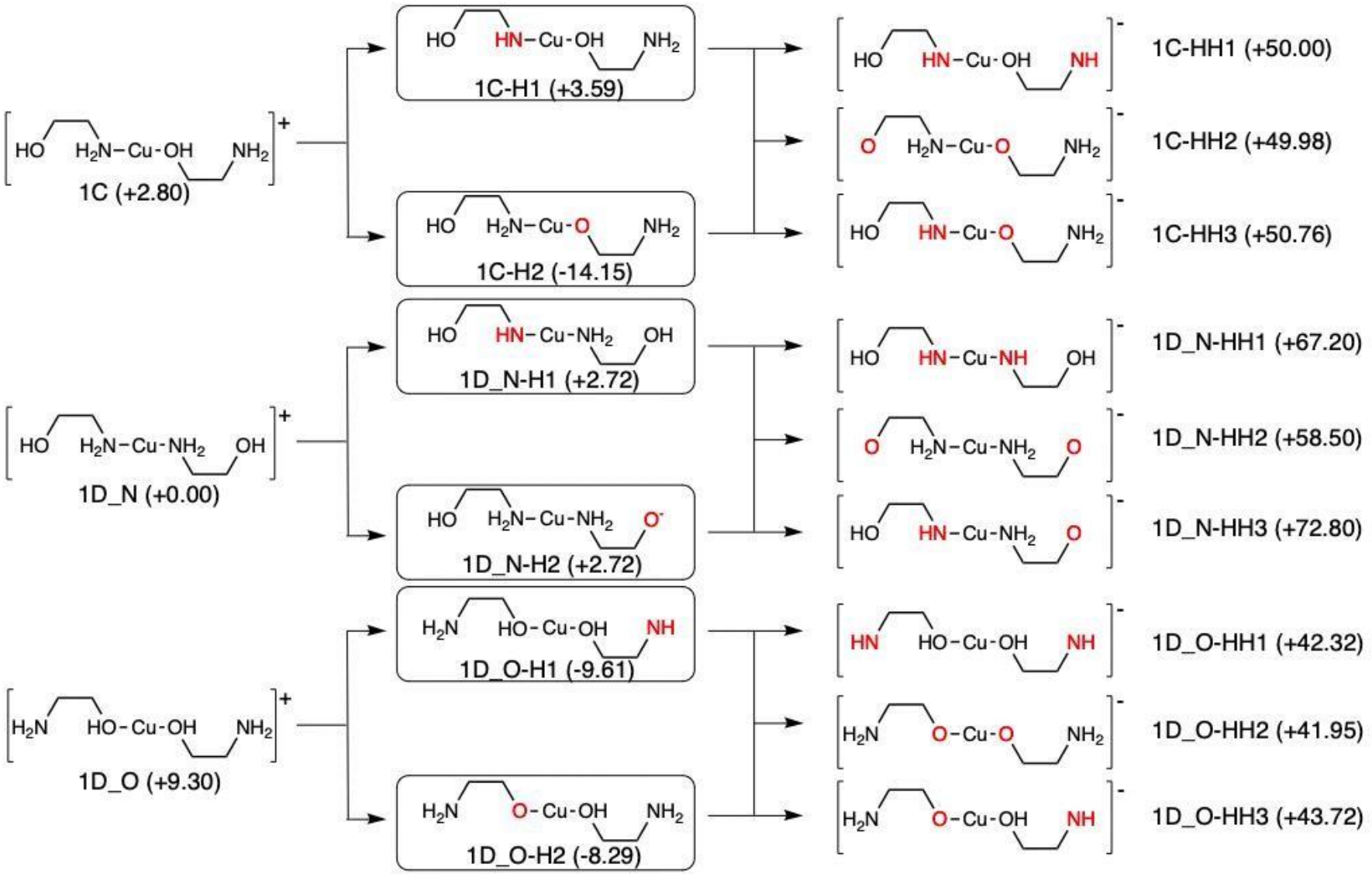


Figure 20: Structures and difference in proton affinities from relative free energies (kcal/mol) of all 2-coordinate monodentate Cu(I) complexes in the gas phase.

When assessing the deprotonated 3-coordinate Cu(I) complexes there was a noticeable change in the trend. In both cases, removing a proton from the donor nitrogen of the ligand caused an energy change of +10.2 and +12.5 kcal/mol from Int1a-H1 and Int1b-H1 respectively (Seen in Table 21), relative to the lowest energy complex. Although a neutral charge for the overall complex is achieved, the energy increase may be due to nitrogen being favoured whilst the amino alcohol is fully protonated, therefore removing a proton from this nitrogen bound structure is likely to be unfavourable/endothermic[69] in comparison to a hydroxyl group. The lowest energy species from all deprotonation calculations were recorded from the removal of a proton from an oxygen, which is to be expected due to this inducing a -1 charge on whichever oxygen atom is deprotonated, which balances the Cu(I) to give an overall neutral complex. Deprotonation of nitrogen also gives the -1 charge; however, this was energetically unfavourable compared with oxygen. Electrostatic attraction between the metal centre and the ligand is more likely achieved through deprotonation of an oxygen donor[83].

The only case where there was a net energy increase through deprotonation of the 3-coordinate Cu(I) geometries was in position 1. This was the deprotonation of the nitrogen atom from the bidentate amino alcohol ligand. Every other position of deprotonation caused a decrease in the energy relative to Int1d, which was the lowest energy 2-coordinate complex. There were particularly large negative energy changes when deprotonating an oxygen atom, which suggests that this deprotonation happens readily, giving evidence that an amino alcohol may be more stable than a fully protonated amino alcohol or a N,N donor ligand when deprotonated.

There is clear evidence that deprotonation has a large effect on the complex that can be further shown by measuring the average M-L bond length before and after deprotonation as shown in Table 22 where the average M-L bond length of both N and O in the bidentate amino alcohol increases significantly upon deprotonation of either of the terminal groups, however, is more significant when deprotonating the O atom. The difference in bond length caused by deprotonation of three coordinate complexes has been calculated and is displayed in Table 23 below where X_{Ln} denotes a bound atom (X) from a ligand, and X_{Sub} denotes a bound atom from the substrate.

Table 22: Change in bond length from fully protonated to singly deprotonated 3-coordinate Cu(I) complexes.

Bond	Change in bond length / Angstrom
Cu-O_{Ln}	-0.228
Cu-N_{Ln}	-0.161
Cu-O_{Sub}	+0.070
Cu-N_{Sub}	+0.070

Measurements of bond length give evidence that the deprotonation of just one of the N or O atoms causes a shift on the charge distribution of the complex, which is similar no matter which atom is deprotonated. Upon deprotonating the amino alcohol at the terminal oxygen donor, the bond length between the ligand and the copper centre (Cu-O_{Ln} /Cu-N_{Ln}) is affected to a much larger extent than the monodentate amino alcohol. The increase in bond length of the monodentate ligand is expected due to the induced negative charge from deprotonation of the oxygen creating a stronger attraction between the negatively charged ligand with the Cu(I) which has low electron density.

As discussed previously in Section 5 with regard to copper coordination, the lack of ligands that are favourably bidentate whilst also allowing the metal centre to stay bound to another amino alcohol unit is suggestive that amino alcohols may be more likely to be stable in a monodentate arrangement. This is further evidenced when the complex is deprotonated as this caused the relative energy to decrease when balanced with base as expected in the gas phase. In order to determine whether the 3-coordinate complexes were truly favourable, the bonds were subjected to the same BS2 conditions used previously on the fully protonated structures of Int1a and Int1b. The results of these calculations show that the average free energy of deprotonated 3-coordinate complexes is more negative than the protonated starting complex. This is useful as it provides insight into potential structures that are useful for constructing a working hypothetical cycle. As seen in literature,[4] the deprotonation of an amino alcohol can cause a drop in the ground state energy of an intermediate relative to the beginning reactants.

6.3.1 - 2-Coordinate Deprotonation

When singly deprotonating 2-coordinate Cu(I) complexes, each intermediate had two possible deprotonation sites due to the ligands only having few combinations for binding modes. The most favourable 2-coordinate structures were more energetically favourable than the lowest energy 3coordinate structures, and more so than the fully protonated monodentate structures of Int1c, Int1d_N and Int1d_O. However, the 3-coordinate structures were used in the next step of the reaction due to a strong suggestion that 3-coordinate systems tend to lead to favourable intermediates after oxidative addition.[4] When deprotonating monodentate complexes singly, there were two complexes, Int1c-H2 and Int1d_O-H1 that showed a decrease of -14.2 and -9.6 kcal mol^{-1} compared to the lowest energy fully protonated structure Int1d_N. These low energy structures may be good choices for the overall mechanism as they may lead to more favourable intermediates that allow for the reaction to proceed.

The highest energy singly deprotonated complexes were Int1c-H1 and Int1d_N-H2 with relative free energies of +2.7 And +3.6 kcal mol^{-1} respectively, relative to Int1d_N. The latter of which was caused by the deprotonation of a terminal oxygen atom, leaving an exposed negative charge on the end of the complex. This would likely be filled with a hydrogen shift from an amino group or form a salt. Evidence for this was potentially seen in optimizations of structures where the crowded NH_2 group was seen close enough to the oxygen due to the short length of the amino alcohol that there is room for transfer of a hydrogen or a hydrogen bond. It is more favourable for the amino alcohol to be bound by a N atom having a coordination number of 3 and to have the induced negatively charged O atom neutralised by a hydrogen shift than for the exposed terminal O atom to remain as O^-. Another option for this charge to be balanced could come from the free K^+ ions in solution from the dissociation of the KOH base.

To summarise, when calculating the free energy of 2-coordinate copper complexes, there are only a few cases where the complexes containing two nitrogen bound amino alcohol ligands has a lower

free energy than complexes with two oxygen bound amino alcohol ligands. This is particularly noticeable in the complexes that have been doubly deprotonated. One instance was found when calculating the energy of fully protonated complexes (Int1d_N), whereas in all cases where there was a single or double deprotonation introduced, the complexes containing two nitrogen bound amino alcohol ligands had a higher free energy. There is more room to explore the nature of this, using the CSD to gain a better scope of the records of 2-coordinate Cu(I) geometries in literature. There are 4 main complexes that show the most promise with being able to provide a low energy mechanistic pathway for this reaction with an aryl halide, which are Int1a, Int1b, Int1d_H and Int1d_O-H1 shown previously in Figure 19 and Figure 20.

Another takeaway from the deprotonation calculations is that there is reason to believe that amino alcohols may be deprotonated in solution, and not after binding to the copper. This can be seen by the lowest energy species already having one amino alcohol deprotonated. A discussion of the pK_a/pK_b values for the relevant bases can be found below in Section 7. Determination of a preference for deprotonation of the amino alcohol will begin to exclude certain complexes, as there may be no logical method of forming these products, which is vital information as a factor of using amino alcohols in this reaction is the selectivity of the product. A change in which end of the amino alcohol is deprotonated may cause a preference for binding that will influence the product being O- or N- arylated.

7 - Bases

In literature reports,[23] the most used bases in this Ullman-type arylation reaction are KOH, KCO_3 and Cs_2CO_3. The purpose of using these bases is generally agreed to be for the dissociation of a halide anion and/or a proton. In cases where amino alcohols have been used as a substrate, Cs_2CO_3 is particularly prevalent as a first-choice base despite the difficulty modelling inorganic bases in organic solvents[4] because it is regarded as a mild base and Cs^+ ions are highly reactive for selective deprotonation.[84] Sodium or potassium inorganic bases have also been documented as successful bases, with NaOH in particular resulting in a N-arylation being highly favoured. Cs-containing bases have been shown to favour O-arylation products, where yields were modest in comparison with the

use of NaOH base.[2] This selectivity may not have been a result of the base used, as it was noted that N/O selectivity was decreased whilst using a milder base such as Cs_2CO_3 or K_3PO_4. The selectivity may have been more of a result of a change of solvent, where DMSO/H_2O were shown to favour N-arylation whereas butyronitrile was shown to favour O-arylation with the Cs_2CO_3 base. The base KOH has been shown to be more effective at assisting in the synthesis of phenols,[2] which is suggesting that O-arylation may be favoured using this, following trends seen with NaOH. The use of a mild base is preferred over a strong base when trying to achieve O-arylation, as stronger bases are more able to deprotonate the largely basic amine group of an amino alcohol, which would encourage N-arylation instead. A mild base is also able to be used instead of a strong one because the substrate is more likely to be deprotonated whilst bound to the Cu(I) centre, due to the increased acidity of the bound functional group in this position.[85]

The study of bases in this reaction is important as to predict the energy cost of the removal of a proton from each amino alcohol in solution requires balancing molecules. The base in this reaction is likely to act as a proton sink, whilst also encouraging the reaction to proceed by providing a cationic K^+ for the iodide to bond with. For the sake of balancing energies, KOH was used in this calculation for balancing the deprotonation, however, a level of uncertainty must be expected as lattice enthalpy of the potassium salt is not considered here. KOH was the initial base used as a replacement for Cs_2CO_3 because it is much more soluble in organic solvents.

The effects of three more common bases were compared over the course of one catalytic cycle, a summary of those results can be seen below in Table 24 where the amino alcohol [$H_2N(C_2H_4)OH$] is noted as **L1**.

Table 23: Testing of bases for deprotonation of amino alcohols.

Deprotonation	**Reaction Energy / kcal mol^{-1}**		
	No Solvent	**Toluene**	**Acetonitrile**
L1 + NaOH → $L1^-$ + H_2O + Na^+	140.14	62.03	10.24
L1 + KOH → $L1^-$ + H_2O + K^+	112.34	42.36	-2.43
L1 + Cs_2CO_3 → $L1^-$ + $CsHCO_3$ + Cs^+	109.28	35.63	5.92

KOH was seen to produce a lower energy pathway for one cycle compared with NaOH and has been reported to be a good mild base to use experimentally. Despite Cs_2CO_3 resulting in the lowest energy deprotonation of an amino alcohol both in the gas phase and in toluene (PCM), KOH was the only base to offer spontaneous deprotonation of an amino alcohol in acetonitrile, where deprotonation is favoured. Thus, KOH was chosen to be the base of choice for each cycle that was calculated. It has also been shown that carbonates interact with transition metals,[4] therefore excluding them may simplify calculations. If not a computational complication, this result may also arise due to the difficulty found with the insoluble nature of Cs_2CO_3 in organic solvents. The energy decreases between using toluene and acetonitrile and is likely due to the nature of acetonitrile being a polar solvent, causing a greater thermodynamic stabilization effect when considering charged species. These calculations do not take into consideration the lattice energy of any of the bases used, which is worth reflecting on, as each of these bases require dissolving into the selected solvent, which has previously been detailed to be a challenge when considering solvation of inorganic bases into an organic solvent.[8, 34] Another consideration of the deprotonation of amino alcohols is the acidity/basicity of the reactants. Bases such as NaOH and KOH have very low basicity (pK_b) of 0.2 and 0.7[86] respectively, which suggest that they will both readily dissociate, decreasing the H^+ concentration in solution. Cs_2CO_3 was used experimentally in acetonitrile, however, the reaction was recorded as 'ligand free' and therefore only truly evidences some solubility of this mild base.[9] Details of the solubility of KOH in acetonitrile are not available, however, laboratory work to determine the solubility limits could be performed in order to determine whether KOH is realistic. It is likely that harsh conditions will have to be used in this reaction such as a high temperature, which would likely aid in dissolving the chosen base. In literature, DMF was used as a solvent, but required temperatures of up to 120°C, this may become an issue as acetonitrile has a low boiling point of 82°C,[87] however, this temperature was reported experimentally[9] to help dissolve Cs_2CO_3.

8 - Ligand Binding and Exchange

There are a range of ligands that can be utilized to have various effects on the coordination of complexes and the progression of a reaction in terms of rate and geometry. Once amino alcohols had been used as ligand and substrate with no competition, the suitability of amino alcohols primarily as a ligand was tested against other ligands that have been the go-to choice for enabling N and O selectivity in coupling of amino alcohols to aryl halides.[4, 5, 23] Alongside the amino alcohol, the selected ligands were phenanthroline, bipyridine and a 2-isobutyrylcyclohexan-1-one (diketone), as shown below in Figure 21.

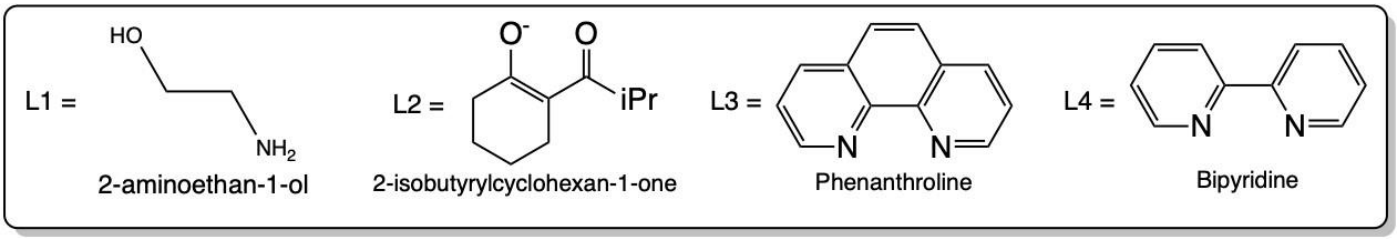


Figure 21: Structure of amino alcohol and common N,N and O,O ligands.

The interest in studying the above ligands is further driven by the comparison of N,N and O,O^- donors to N,O and N,O^- donor ligands. There are studies to suggest that in this specific arylation reaction there may be a hierarchy for the preference of these ligands to bind to a copper metal centre, hypothesising that O,O donors will have the lowest binding energy, where N,O donors have the greatest.[18]

8.1 - Methodology

The binding energy of each ligand versus a fully protonated and singly deprotonated amino alcohol were calculated at the BS2 level of theory with solvation.

The comparison used for this study also shows the addition of each of the ligands in Figure 21 to a 3-coordinate Cu(I) complex in order to determine whether each ligand is likely to displace an amino alcohol in the reaction matrix. The reaction schemes for this can be seen below in equations 1, 2 and 3.

8.2 - Binding Energy

The binding energy of each ligand with a Cu(I) cation is shown below in Table 25 with the deprotonated species denoted with an additional H, as shown below in Equation 1 and 2.

$$Cu^{+} + Ln \rightarrow [CuLn]^{+} \quad \text{Eq.1.}$$

$$Cu^{+} + Ln^{-} \rightarrow CuLn \quad \text{Eq.2.}$$

Table 24: Binding energy of common ligands with Cu(I) using B3LYP with 6-31G(d,p)/SDD and solvation.

Species	Bond Energy kcal mol^{-1}
L1	+189.7
L1H	+218.4
L2	+194.6
L3	+86.9
L4	+86.9

The bond energy of the deprotonated amino alcohol (**L1H**) is greater than that of the fully protonated amino alcohol (**L1**) as shown above in Figure 24. This is due to the negative charge induced by deprotonation in solution. This energy difference would likely be even greater in a solvent such as MeCN, because this highly polar solvent will encourage charged species to partake in the reaction.[4] Both **L3** and **L4** (previously shown in Figure 21) have a very similar bond energy, which is expected due to the structures of phenanthroline and bipyridine being nearly identical, in particular the constituents bound to the N donor sites of these ligands. The diketone type ligand (**L2**) has a much greater bond energy than either of the N,N donor ligands which is evidence of the greater attraction to Cu(I) from a -1 charged ligand. Based on these observations, it is likely that the deprotonated amino alcohol and diketone ligands will form a stronger bond to the Cu(I) than either **L3** or **L4**. It can be seen that **L2** will likely form a stronger bond with Cu(I) than a fully protonated amino alcohol (**L1**), however, deprotonation of this ligand (**L1H**) increases its binding strength significantly. The bond lengths of each of these structures is shown below in Table 26.

Table 25: Bond lengths of fragment of donor atoms with Cu(I) cation

Complex	Bond Length O_1/N_1 / Angstrom	Bond Length O_2/N_2 / Angstrom
Cu-**L1**	2.081 (O)	2.037 (N)
Cu-**L1H**	1.891 (O)	2.114 (N)
Cu-**L2**	2.022	2.022
Cu-**L3**	2.026	2.026
Cu-**L4**	2.024	2.015

The average bond length of Cu-O was seen to decrease when **L1** was deprotonated to **L1H**, which is likely due to the greater attraction between the negatively charged oxygen donor and the copper centre. The Cu-N bond length when using **L1H** is greater than that displayed by any other ligand, which when compared with the Cu-O bond length shows that there is greater interaction between the deprotonated oxygen donor atom and the metal centre than the nitrogen donor atom. The bond length displayed by **L3** and **L4** were extremely similar, which is to be expected because the donor atoms are both N and the general properties of the ligands are the same. The imbalance mentioned from donor atoms in **L1H** gives evidence that if the amino alcohol is to become monodentate, there may be a favoured atom to be dissociated, whereas this may be less predictable in **L1** without the deprotonation.

8.2.1 - Geometry Effects

It is important to know that there are structural changes due to the ligand binding and exchange. When looking at geometries of copper with **L2**, **L3** and **L4**, each of these common bidentate ligands displays monodentate characteristics when an amino alcohol is also bound with the Cu(I) centre. Evidence for this is in the bond lengths shown below in Table 27. This has been seen previously in section 'Cu(I) Coordination' where bipyridine can be seen bonding monodentate with copper with only a weak attraction from the second N donor group. This same trend is apparent with **L2** and **L3**.

Table 26: Bond length comparison of commonly bidentate ligands when bound to Cu(I) with an amino alcohol substrate.

Complex	Bond Length O_1/N_1 / Angstrom	Bond Length O_2/N_2 / Angstrom
Cu-L2-AAL	1.939	2.319
Cu-L3-AAL	1.973	2.435
Cu-L4-AAL	1.990	2.297

In comparison with the bond lengths in Table 19, the O and N bond lengths in Table 27 have increased by ~0.3-0.4 Å. This is a large enough increase to postulate a broken bond. A useful calculation to further examine the effects of changing the ligand is to introduce constraints to force each of the additional ligands to become bidentate. The distances used between donor atoms and the copper centre were elongated by 0.1 Å compared to the monodentate bond length to allow the geometry to come to an equal distance. These calculations are relevant to the competition between ligands and have a large weight in dictating the competitive outcome of multiple ligands in a reaction. There is likely to be ligand exchange when multiple ligands are present, which is outlined in Section 8.3.

8.3 - Ligand Exchange

There is a measure of the effect of supporting ligands on a complex encompassing the ability of an amino alcohol to be encouraged to remain bound or to be removed. This describes the effect of exchanging the ligand bound to copper, whilst not assessing any barriers for interconversion between structures. The dissociation energy of an amino alcohol substrate or the supporting ligand is shown below in Table 28. This was calculated using the formula shown below in Equation 3 describing the energy of the fragments of the substrate and ligated copper with the energy of the complex before dissociation.

$$E_{Diss} = E[Substrate]_{Frag} + E[CuLn]_{Frag} - E[CuLnSub]_{Complex} \quad \text{Eq.3.}$$

Where L1H has been used, this indicates that the amino alcohol ligand has been O-deprotonated. L1 is not present, as the amino alcohols are assumed to be deprotonated throughout.

Table 27: Dissociation energy of Substrate and Ligands from Cu(I) complex.

Complex	Amino Alcohol Dissociation Energy / kcal mol^{-1}	Ligand Dissociation Energy / kcal mol^{-1}
Cu-**L1H**	70.8	70.8
Cu-**L2**	64.3	34.4
Cu-**L3**	146.3	14.8
Cu-**L4**	147.0	15.6

Both **L1H** and **L2** provide a lower energy environment for the removal of the amino alcohol substrate. This is likely due to the electron withdrawing nature of the charged O donor atoms of these ligands. The substrate and ligand dissociation energy of **L1H** are equal due to both of these being amino alcohols. In all cases except for **L1H**, the dissociation of the ligand is energetically more favourable than the dissociation of the amino alcohol. This is particularly true for the uncharged ligands **L3** and **L4**, where the relative free energies of +14.8 and +15.6 kcal mol^{-1} are nearly ten-fold lower than removing the protonated substrate. This is due to the substrate being bound by a deprotonated oxygen atom which has a much higher binding affinity than nitrogen. The dissociation energies of **L2** are more similar than those of the other ligands calculated, showing the electronic influence of the two O donors in **L2**. Due to there being only one of the two O donors bound in this complex, the bond dissociation of the **L2** ligand is less than that of the substrate, as the second O donor is an electron withdrawing group[88] itself and weakens the monodentate Cu-O bond of **L2**.

A further investigation of ligand exchange used the 'opt=modredundant' input command to restrain the structure of the complexes in question to ensure that the ligands were bound bidentate alongside the BS2 methodology. Optimizations were performed with and without restraints in order to understand whether Cu(I) structures were more likely to be 2- or 3- coordinate whilst using typically bidentate ligands. The restraints were used to generate a 3-coordinate Cu(I) complex with the bidentate phenanthroline, or diketone ligand donor atoms restricted to being bound with the copper centre. The amino alcohol substrate was not restrained to a Cu-N bonding distance. This was done to understand the likely reasons for the typically bidentate ligands binding in a monodentate way in these copper(I) complexes. The results of this are shown below in Table 29.

Table 28: Dissociation energy of Ligands and amino alcohol substrate using restrictions at BS2 level of theory.

Complex	Amino alcohol Dissociation Energy / kcal mol^{-1}	Ligand Dissociation Energy / kcal mol^{-1}
Cu-**L2**	46.8	16.8
Cu-**L3**	129.2	-2.3
Cu-**L4**	125.8	-5.6

The results above show that the forced bidentate binding causes an energetic strain on the complex where the energy required for bond dissociation drops for both deprotonated amino alcohol and ligand. This shows that forcing a 3-coordinate structure with these added ligands drops the energy to a point where **L3** and **L4** will readily dissociate under the new strain, whereas the charged diketone ligand is able to remain coordinated. The energy required to dissociate the amino alcohol is greater than any other ligand in all cases, however, the energy has decreased from the values shown in Table 28 by ~20 kcal mol^{-1}. The energy decrease is due to the increased steric demand of a large bidentate ligand having forced bonds, which makes all of the bonds surrounding the copper centre less thermodynamically stable, requiring less energy to be broken. This suggests that the deprotonated amino alcohol is able to outcompete commonly used bidentate ligands in solution and therefore it is valid to begin to appreciate that amino alcohols are able to act as ligands as well as substrates when in a reaction mixture together. These findings are similar to those reported in by Fey *et al*.[18] where the hemilability of ligands has been considered to be useful for protecting binding sites and for altering the selectivity of a reaction. Competition between ligands caused by the similarities in binding energies may be a useful way to selectively 'turn off' ligands through the addition of separate ligands with a greater affinity for the metal centre.[8]

9 - Mechanisms

9.1 – Cycles

The mechanism shown below in Figure 22 is a hypothetical catalytic cycle for this project and describes the coupling of aryl iodide with an amino alcohol through an oxidative addition/reductive elimination pathway, using a similar pathway as investigated by Yu *et al*.[4] This mechanism has two routes, accounting for the N-selective and O-selective products of the reaction. Information regarding the two pathways is useful for being able to favour specific reactions in order to reach desired products faster experimentally. In comparison to the aforementioned 'Cycle A' used for method evaluation, this provides a more in-depth analysis of the most likely and efficient route to product formation. The species analysed are influenced by literature and experimentation of the ideal coordination of amino alcohols and the coordination sphere of Cu(I) and Cu(III) as described above. The 3-coordinate to 4-coordinate pathway is focussed on as a low coordination number pathway that includes both an O-selective and N-selective routes in order to compare the relative energy of each product formation. The relative free energy of each intermediate of the cycle has been collated in Table 29 below for clarity.

Table 29: Relative free energy of final Cycle O and N at BS3 level with solvation and dispersion corrections.

Structure	Relative Free Energy / kcal mol^{-1}	Structure	Relative Free Energy / kcal mol^{-1}
Pre1_O	0.00	Pre1_N	0.00
Int0_O	-8.55	Int0_N	-8.55
TS1_O_*cis*	-3.04	TS1_N_*cis*	-3.04
TS1_O_*trans*	-1.78	TS1_N_*trans*	-1.78
Int1_O_*cis*	-13.48	Int1_N_*cis*	-13.48
Int1_O_*trans*	-15.51	Int1_N_*trans*	-15.51
Int2_O	-19.43	Int2_N	-9.89
Int3_O_*cis*	-38.49	Int3_N_*cis*	-36.64
Int3_O_*trans*	-34.98	Int3_N_*trans*	-34.03
TS2_O_*cis*	-26.81	TS2_N_*cis*	-29.10
TS2_O_*trans*	-26.77	TS2_N_*trans*	-25.46
Prod1_O	-53.27	Prod1_N	-57.59

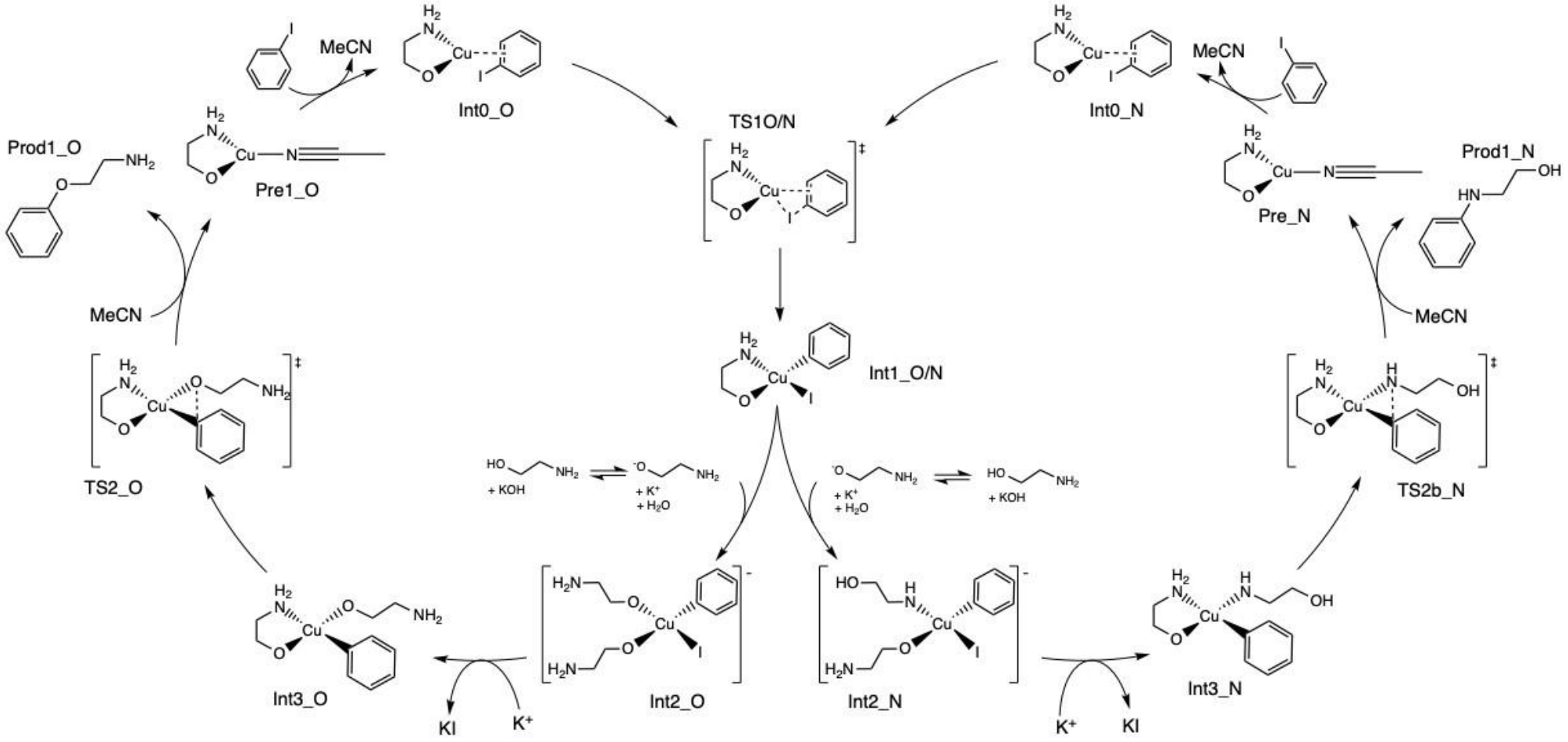


Figure 22: Final catalytic cycle for N- and O- selective coupling of aryl iodide with ethanolamine considered computationally in this thesis.

Upon initial optimisation calculations of the intermediates in proposed Cycle O, Int1a_O, a 4coordinate complex bound to an aryl ring, iodine and bidentate coordination of an amino alcohol could not be found. The solution to this was to use a deprotonated amino alcohol ligand, which readily optimised. The deprotonation of this ligand is not necessarily restricted to one stage of the cycle, unfortunately alone, an easy geometry optimization provides little evidence that amino alcohols may be able to be deprotonated in solution in comparison to proton affinities in MeCN and binding energies previously discussed. To explore this more, the binding energy of protonated vs deprotonated amino alcohols can be compared, as previously shown.

A balanced equation showing the relative energy of the deprotonation of the amino alcohol in the gas phase compared to toluene and acetonitrile was used, as outlined in Section 6 – Deprotonation, which has been repeated below in Equation 3 for clarity.

$$[H_2N(C_2H_4)OH] + KOH \rightarrow [H_2N(C_2H_4)O]^- + K^+ + H_2O \qquad \text{Eq.3}$$

The lowest energy option requires the deprotonation of the -OH group in MeCN (-2.43 kcal/mol) compared to the next lowest deprotonation of $-NH_2$ also in MeCN (+19.42 kcal/mol). The deprotonation of oxygen may lead to greater O-selectivity of the product. In all solvents, deprotonated oxygen is a more favourable amino alcohol structure as previously discussed (Section 7 – Bases), which is not to say that some of these molecules will be deprotonated at the nitrogen or not deprotonated at all whilst in solution. There are a lot of different routes that individual molecules can go down and rearrange to, as displayed by the basicity of the KOH used, where all positive and negative charges can be stabilised, however, full deprotonation at the oxygen atom was preferred and therefore was chosen as the production model amino alcohol.

The second transition state in this cycle is a 4-coordinate complex, which, depending on the binding site of the monodentate amino alcohol, has four isomers. The occurrence of four isomers potentially determines O or N selectivity and whether in these pathways the aryl ring was more energetically favourable to be *trans* to the N or O of the bidentate amino alcohol. The lowest energy transition state was Int1e_TS2b, which has a relative free energy of -40.4 kcal mol^{-1}, which is the lowest energy pathway found by -4.56 kcal/mol compared with Int1e_TS2a. The aryl ring

tends to be lower in energy when *trans* to the N of the amino alcohol which is a trend already seen in past

structures of Int3e, where the aryl ring was significantly more stable in this position *cis* to both O atoms (Section 4). In this mechanistic pathway, N-selectivity for the product is only favoured by 0.0019 kcal mol , therefore a large preference was not observed. However, once the amino alcohol has been bound monodentate to the Cu(III), each individual stage of the hypothesised mechanism is lower in energy when bound to the oxygen, which is further discussed in Section 9.3.1. Cu(III) is likely to be a favourable electronic configuration for beginning the reductive elimination step of the cycle due to it being a d^8 metal centre, which makes it highly susceptible to surrounding ligands causing a strong *trans* influence.[70] This influence makes for a complex with good leaving groups, which was agreed upon by Yu *et al.*[4] despite their suggestion that this Cu(III) would be 5-coordinate, the d^8 electron count was still maintained. Experimental work from Casitas *et al.*[89] suggests that the Cu(I)/(III) redox couple has a low activation barrier and that addition of acid to a Cu(III) complex can easily drive the reaction to proceed at room temperature. In the aforementioned study,[89] the ligand used was 1,10-phenanthroline, therefore direct comparison to amino alcohols is not necessarily appropriate, however, it shows that large nucleophiles can be eliminated from a Cu(III) complex. Research from Jones *et al.*[5] also hypothesises the d^8 Cu(III) centre, despite the use of a SET/IAT mechanism, which gives further evidence that the consensus on this elimination step of the reaction follows with this intermediate.

9.2 - Precursors

The computational method for optimization and frequency calculations for the precursors was the BS2 method previously outlined using implicit IEFPCM acetonitrile solvation and dispersion corrections. The selection of precursors assessed for the cycle are shown below in Figure 23. As previously used, the addition of ‘H’ denotes a deprotonation where the addition of ‘P’ denotes a protonation.

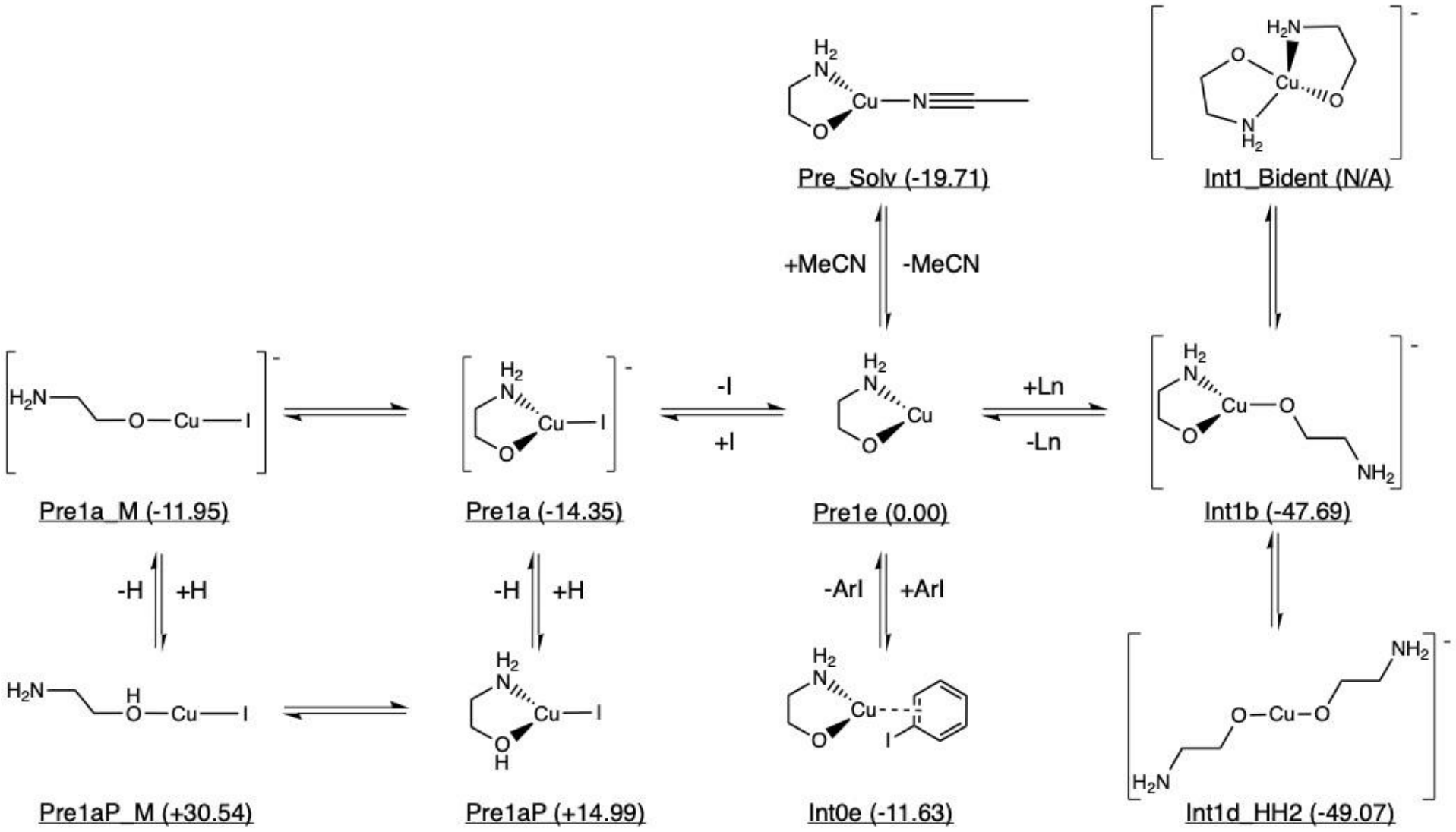


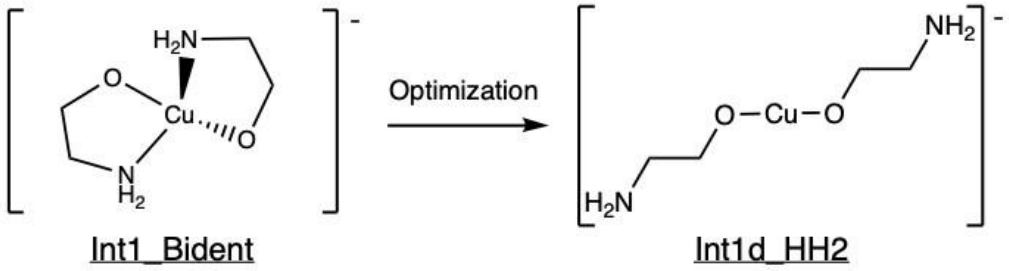
Figure 23: Comparison of precursors to begin the mechanistic pathway using BS2, solvation and dispersion with relative energy in kcal mol^{-1}.

Reports from the groups of Monnier,[9] Yu[4] and Sambiagio[2] show that it is very common for copper salts to be used as precursors to copper catalysed Ullmann reactions. This salt is ideal for an arylation using aryl iodide, as it keeps any free halide ions consistent. One challenge for constructing a reaction mechanism is to choose a suitable starting point, and in particular the complex that will be most likely for allowing turnover of a cycle. Once completion of one cycle has occurred, the arylated amino alcohol product will be formed, along with other fragments from the reductive elimination step of the reaction. There are several complexes with monodentate or bidentate amino alcohols that may be formed in this step, and a further look at the relative Gibbs free energy of these intermediates can be seen in Figure 23 compared to the single amino alcohol bidentate complex of Cu(I), which is likely to be highly unstable due to the large amount of open space that would easily accommodate another amino alcohol, halide ions or solvent. It is likely that structure Int1d_HH2 is a catalyst sink, and may cause issues for turnover.

It can be seen that the only cause of an increase in energy from the starting geometry of Pre1e is caused by the protonation of the amino alcohol, as displayed in Pre1aP_M and Pre1aP. This protonation causes the net charge of the complex to revert to neutral, however, is unlikely to occur as these calculations are run in acetonitrile, which gives amino alcohols a strong preference to remain deprotonated. Without this protonation, it is clear that the addition of a free iodide will stabilize the complex and is a feasible route due to the amount of iodine in solution from ArI and CuI. The other two possible additions to Pre1e are an aryl halide and the coordination of solvent molecules. One molecule of acetonitrile causes a -19.71 kcal mol^{-1} change in relative free energy, whilst addition of aryl iodide causes a -11.63 kcal mol^{-1}. This shows that increasing the coordination number of this Cu(I) complex from 2 to 3 is beneficial, increasing the electron count to be 16 instead of 14. For Cu(I) complexes, the 18-electron rule has many exceptions, one of which following with the coinage group in being largely stable in 2-, 3- or 4- coordinate complexes with only 14 or 16 electrons.

The larger number of electrons does not cause the same level of stabilization for Cu(I) whilst coordinated to two deprotonated amino alcohol units. Especially in the case of Int1d_HH2 where, as predicted by trends of gold and silver, Cu(I) is stable with a linear 14 electron geometry.[54] This

was reinforced by an attempted optimization of a 4-coordinate complex containing two bidentate amino alcohols, which resulted in both amino groups dissociating with interatomic lengths of 3.2 Å. A sketch of this attempted optimization can be seen below in Figure 25.



Int1_Bident Int1d_HH2

Figure 24: Sketch of the optimization of Bidentate amino alcohol ligands using BS2 with solvation and dispersion

For cycles that end with a 3-coordinate geometry, there may already be an iodine atom coordinated, which will make structure Pre1a the most likely intermediate, however, the strength of this Cu-I bond may allow for an amino alcohol or solvent molecule to cause displacement, especially in higher temperature reaction mixtures where the Cu-I bond is under more stress. If a cycle undergoes reductive elimination from a 4-coordinate complex, then coordination of another deprotonated amino alcohol or solvent molecule may be the favoured route to allow for cycle turnover to be successful. In relation to the scheme shown previously in Figure 22 this is why a 4coordinate intermediate route was chosen to proceed whilst not using exclusively bidentate ligands, when possible, as the amino alcohol Cu-O and Cu-N bonds do not present equal affinity towards the metal centre.

To summarise, the relative free energies for the precursor to this reaction have been compared and contrasted. There is a strong preference for the 3-coordinate complexes as well as Int1d_HH2, which will be where a lot of the catalyst is lost, in particular the solvated Pre_Solv and Pre1a showed promise in terms of their relative energies. The energy difference from the solvated intermediate towards the next intermediate in the cycle was lowest, therefore these are the best choices.

9.3 – Oxidative Addition

The oxidative addition stage of the reaction involves a change of oxidation state from copper(I) to copper(III). This is summarised in Figure 22, which shows the change in oxidation state occurring from Int0_O and Int0_N to Int1_O/N, with a transition state that links the two structures from a copper-aryl adduct into a concerted addition of the aryl iodide. This reaction step also requires an increase of coordination number from 3 to 4. One amino alcohol ligand is introduced to Int1_O/N, which is able to form either a 4 or 5 coordinate Cu(III) complex. The 4-coordinate complex is more likely to form as the bond strength of Cu-O and Cu-N are not equal, therefore there is a chance for the bidentate amino alcohol ligand to become temporarily monodentate by dissociation at one end in order to give a less strained complex. The small increase in coordination number may be preferred over an increase from 3 to 5 coordinate, as it allows for more flexibility for the surrounding ligands, which causes a more sterically favourable coordination sphere, explored in more detail below.

The activation barriers to oxidative addition in structures Int1_*cis* and Int1_*trans* are -13.48 and 15.51 kcal mol^{-1} respectively. Compared to Yu *et al*.[4] and Jones *et al*.[5] these barriers are very low. Yu proposed a cycle with an oxidative addition barrier of 16 kcal mol^{-1} which is very similar to the calculated barrier for Int1_*trans*, which shows that this oxidative addition is feasible. Yu also disregarded the barriers from SET and IAT mechanistic considerations from Jones due to their height of over 35 kcal mol^{-1}. One reason for the lower activation barrier may be that the cycle proceeds from a solvent coordinated Cu(I) complex rather than beginning the cycle with a Cu-Aryl adduct. The isomers of the transition state structure TS1_O (-3.04 kcal mol^{-1}) and TS1_N (-1.78 kcal mol^{-1}) are very similar in geometry and thus in relative free energy, with the difference in energy being one of the smallest between isomers at only 1.26 kcal mol^{-1}.

The lowest oxidative addition barrier is for Int1_*trans*, which suggests that there is a preference for the aryl iodide nucleophile to approach the Cu(I) intermediate so that the aryl group is *trans* to the nitrogen donor atom. This is a trend that is predicted to be sustained throughout the mechanism due to the large *trans* influence of aryl functional groups. This arrangement of atoms has a better

fit around the central atom yet does not promote any specific selectivity due to the lack of involvement of a second amino alcohol ligand at this stage of the reaction.

9.3.1 – 3 and 5-Coordinate Geometries

The hypothetical cycles O and N are based on the oxidative addition of Cu(I)/Cu(III) using predominantly 4-coordinate copper complexes. The inclusion of 3- and 5-coordinate copper complexes was investigated, as these structures have been suggested in literature by Yu *et al.*[4] and Giri *et al.*[23] to be feasible in the gas phase. It was thought that a 3-coordinate Cu(I) intermediate would be able to replace Int1_O and Int1_N shown in the final cycle, and that a 5-coordinate Cu(III) complex would be able to replace Int2_O and Int2_N.

The optimizations and free energies for these structures were calculated in the gas phase using B3LYP/6-31G(d)/SDD. This BS2 methodology was followed as both Cycle O and N used this methodology, however, solvation and dispersion were not included because in testing, these conditions caused the structures to dissociate, therefore comparison of energies is not appropriate. It was found that even in the gas phase there were several structures that proved difficult to completely optimise for both 3- and 5-coordinate copper complexes without constraining the distances to remain bonded, as it was common for both the 3- and 5-coordinate structures to reduce to a 2 or 4-coordinate counterpart respectively. In the optimizations, it was common for the NH_2 to become dissociated from the Cu(I) centre, forming a near linear complex. The structures of the intermediates that would not optimize are shown below in Figure 23.

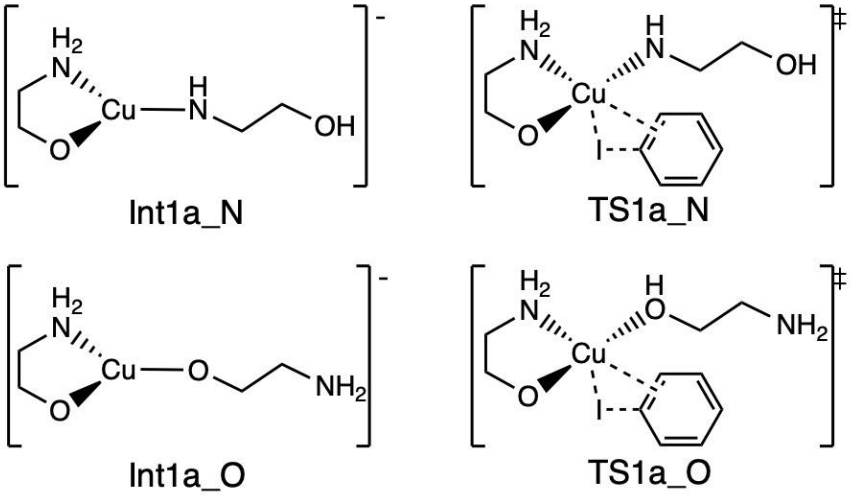


Figure 25: Failed 3 and 5-coordinate geometry optimizations.

When considering Int1_N/O, it is important to consider that even though amino alcohols in acetonitrile are likely to be deprotonated, in the gas phase it was found through several optimizations that a neutral Cu(I) complex with only one of the amino alcohols deprotonated was more favourable due to optimizations completing. This is in line with results using an amino alcohol with a -1 charged diketone type ligand used from Yu *et al.*,[4] where the singly deprotonated amino alcohol in a neutral copper(I) complex resulted in a +19 kcal mol^{-1} relative energy increase.

As stated above, the lack of solvation may have caused the amino alcohols to favour staying protonated, however, the addition of solvation was troublesome for the Cu-I bond, which frequently dissociated to bond lengths over 4 Å.

It can be noted that **Int2a** in this profile has a very large relative free energy drop (-61.0 kcal mol^{-1}) for a structure that is potentially less sterically favourable and may be a result of a dispersive effect. This is likely due to the lack of solvent included in the calculation, which has been previously discussed to provide a stabilizing effect in Section 4.2. As previously mentioned, a report from Yu et al.[4] shows that despite many 5-coordinate structures being optimized, many of these were only found in the gas phase, as the use of solvation caused structures to not converge to a local minimum.

To explore this, the same pathway was reoptimized using BS2 with solvation and dispersion, with Single Point Energy calculations using a larger basis set of 6-311+G(d,p)/SDD (BS3). After calculation, it was clear that **Int2a** was highly unstable and when including implicit solvation in the calculations, the iodine or amino alcohol repeatedly dissociated. As a fix for this, a 4-coordinate Cu(III) intermediate **Int2_O** was constructed, using the hemi-labile nature of the amino alcohol to proceed, as shown below in Figure 26. The N-bound version of this pathway results in **Int2_N** and **Int3_N**.

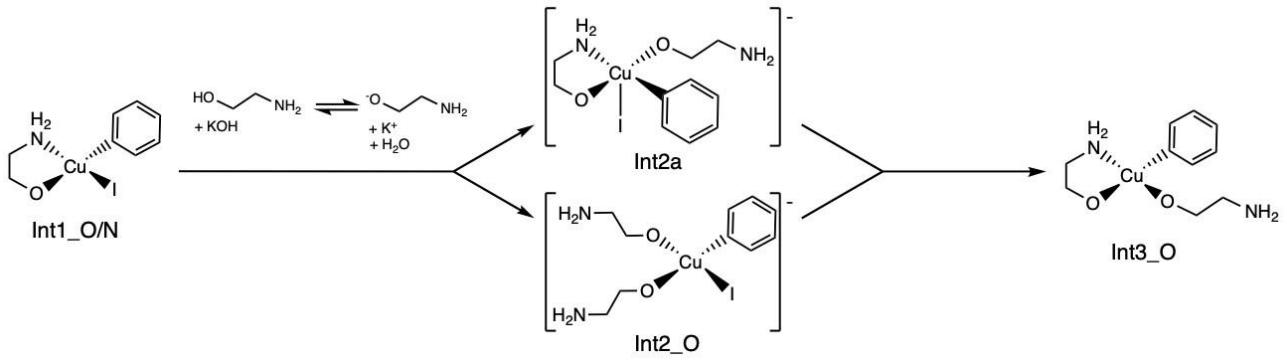


Figure 26: Alternate pathway to reach Int3_O using a lower coordination number.

The change of coordination mode of **Int2a** led to alterations to the initial hypothetical cycle. Whilst using the BS2 level of theory alongside solvation and dispersion corrections, **Int2a** could not be optimized. The 4-coordinate Cu(III) structures have been explored as stable complexes in literature and a few examples of these have been isolated and observed using NMR.[27] Another change made from the initially suggested mechanism is the precursor, which has been changed as a result of the calculations shown in Section 9.1 – Precursors. This section outlined the possible intermediates formed from the reductive elimination product, where the addition of acetonitrile solvent was a highly favourable step (-19.71 kcal mol^{-1}), which would only be easily outcompeted energetically by another deprotonated amino alcohol (-47.69 kcal mol^{-1}). In this case, however, the acetonitrile has been chosen as the precursor step of the reaction, as it is the step that would allow coordination of the following aryl iodide with the lowest barrier, which will eventually allow an amino alcohol to bind and takes into consideration the excess of solvent. This decision may allow for the intermediates nearer the end of the reaction to be formed.

Once the first half of the mechanism was considered and a final mechanism was outlined, a more reliable energetic pathway can be considered, with the energy levels of each pathway for O and N selectivity being compared. Figure 27 below shows the relative free energy of all possible intermediates in the suggested hypothetical catalytic cycle.

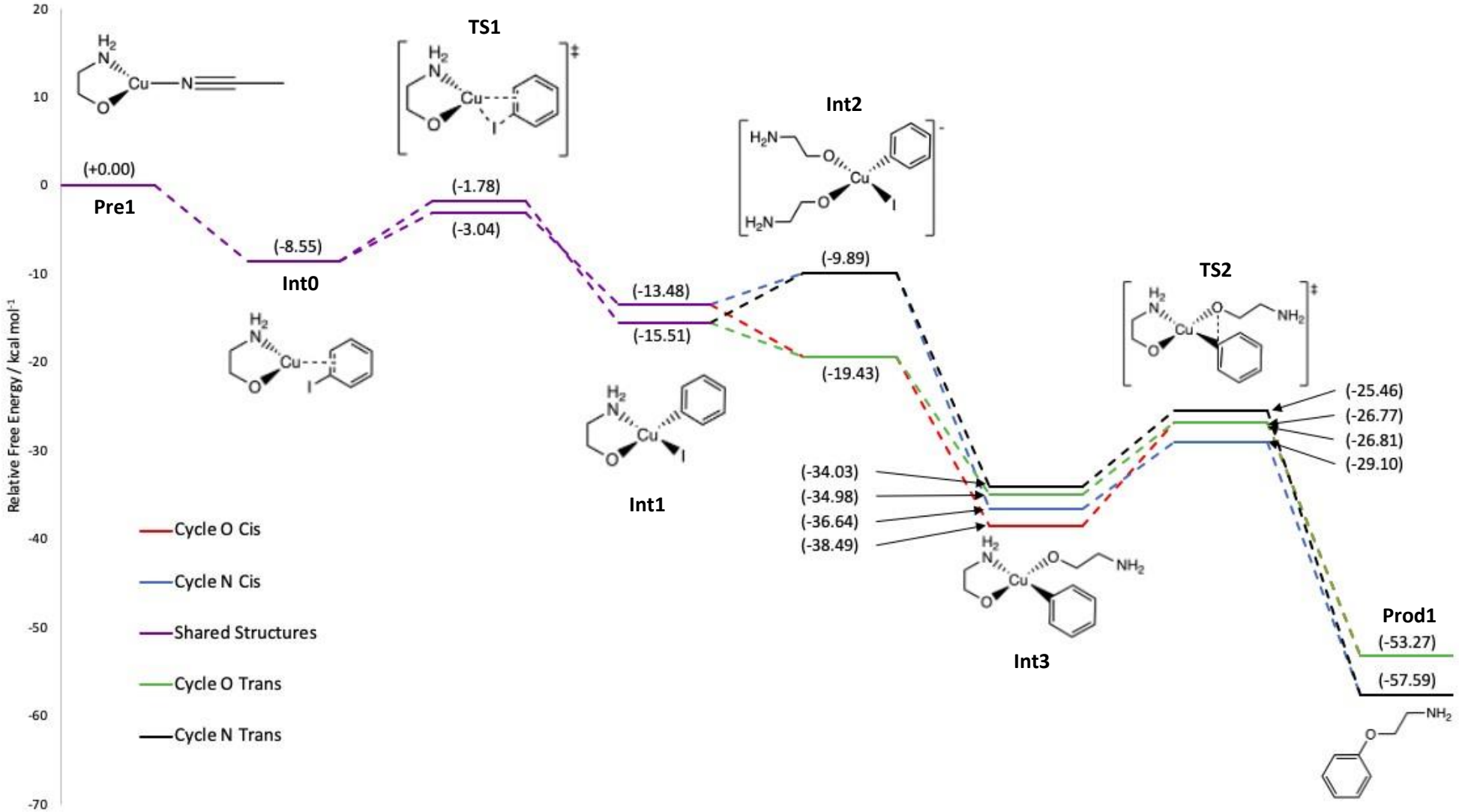


Figure 27: Energy diagram of Cycle O and N comparing isomers and relative free energy using B3LYP with 6-311+G(d,p)/SDD with solvation and dispersion.

The individual isomers of Int3 and TS2 are showed below in Figure 28 for clarity.

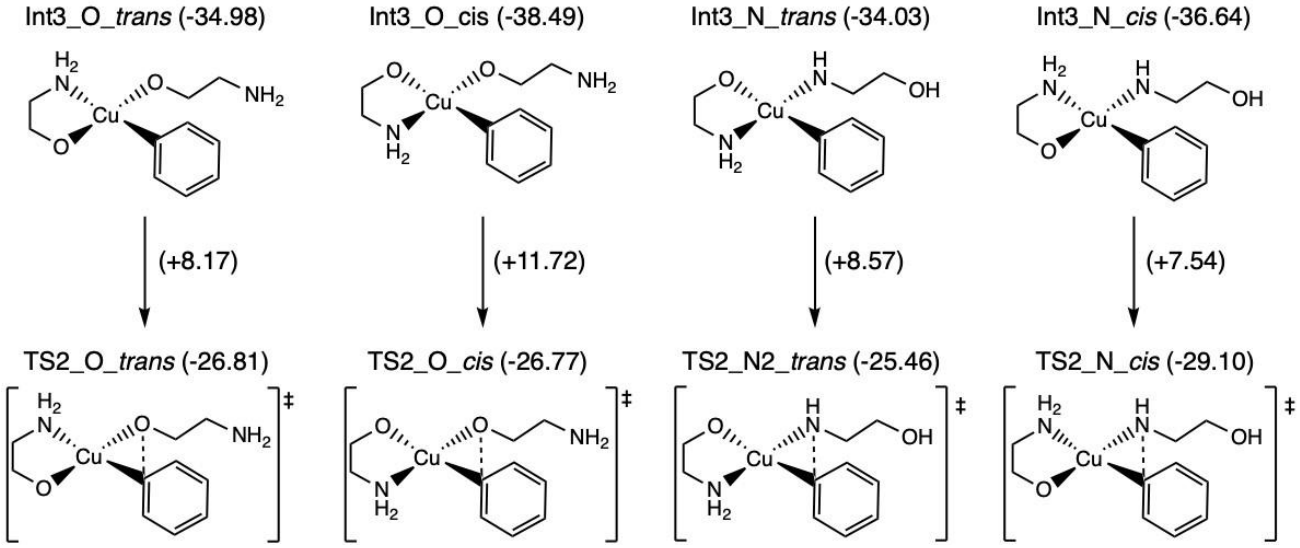


Figure 28:Isomer structures Int3 and TS2 from Cycle O and N comparing relative free energies in kcal mol^{-1}.

Shown above are the four pathways that Int3 can take to form the final transition state of the hypothetical cycles. These options are comparable through steric and energy influences. The main differences between these structures are the orientation of the bidentate amino alcohol and the donor atom of the monodentate amino alcohol bound to the central copper. The lowest barriers are those with a nitrogen donor atom positioned *trans* to the aryl group. This causes Int3_N_*cis* to have the lower energy barrier for the N-arylation pathway as the *cis* is referencing the orientation of the nitrogen groups. In Int3_O_*trans* the orientation results in the same position of the N donor atom, showing that the position of each ligand is vital for ensuring the lowest energy pathway to product formation. A further discussion of this can be seen below in section 9.4.

9.3.2 - Copper – Iodide Bonding

Copper-iodide bonds are the most thermodynamically stable at lower temperatures, where crystallization is favoured, however, this is often due to the highly insoluble nature of CuI in water.[90] The reactions that use copper catalysts are often conducted at between 80-110°C,[34] and are performed in solvents such as Toluene or Tetrahydrofuran (THF) which suggests that Cu-I bonds may be easier to break in these environments as the solubility of CuI increases. Some experiments have been carried out that use acetonitrile at room temperature, however, they were considered to be ligand free.[16] In contrast, some other experiments have recommended the use of acetonitrile at lower temperatures, as it increases the solubility of CuI, with a peak of ~3.5 g CuI/100g at 30°C,[90] however, the solubility does seem to increase with temperatures over 70°C which is the ideal zone for the optimal temperature of this reaction[9, 16] although the DFT calculations in this work did not

take into account the temperature of the reaction. This is suggesting that the lower and higher extremes of the temperature of the reaction, the weaker the interactions between Cu^{+} and I^{-}, which may explain some challenges in the optimization of high coordination complexes containing both copper in the I or III oxidation state, and iodide.

As a result of many optimizations including the bonding between Cu(I)/(III) with aryl iodide and single iodide ions, some trends were observable. It was notable that when 5-coordinate Cu(III) species were attempted to be optimized, they only converged to a global minimum when in the gas phase. Whilst recalculating some of the optimized geometries found in the gas phase using implicit solvation, the iodine atom was seen to dissociate every time in acetonitrile, which gave more evidence for some of the 5-coordinate intermediate containing cycles to become not feasible.

The computational evidence suggests that copper iodide is a suitable salt to rely on for achieving a high concentration of iodide in the reaction mixture. This supply of Cu^{+} and I^{-} is beneficial to the reaction as it suggests that neither of these components will be the limiting factor for an experimental procedure. There will be a large amount of aryl iodide present, therefore the high concentration of iodide ions present will ensure that these I^{-} anions are involved in the reaction, even if the concerted oxidative addition portion of the proposed mechanism does not proceed exactly as detailed.

9.4 – Reductive Elimination

As shown above in Figure 28, structure **Int2_N** is higher in energy than **Int2_O** in all cases, which shows that in polar solvent it is possible that O-binding of ethanolamine may be more favourable. This could be studied with under different conditions in order to investigate the selectivity in solvents such as toluene or DMSO as these have been used in other research.[4, 16]

Overall, the *trans* pathway provides a lower energy route to product formation with few exceptions, for example the *trans* isomer of **TS1_O** is higher in energy than the *cis* isomer, however, this is a very small energy difference of 3.51 kcal mol^{-1}. This can be taken lightly as this geometry is a distorted square planar and truly lies between square planar and tetrahedral, which means that

the only difference between these structures is a slight rotation of the ArI group. A better insight in to this stage of the mechanism could be explored using conformational searches, which would give access to the range of isomers. In contrast to this, there is a clear energy difference between early intermediates of **Int2_O** and **Int2_N**, where the orientation of the amino alcohol ligands is more favourable to be bound by the oxygen donor atom whilst subjected to the conditions used in the BS3 methodology with acetonitrile solvent.

In the reductive elimination stage of the mechanism, there is likely to be a large *trans* influence with regard to the removal of the coupling product, with the 5-coordinate Cu(III) geometries being held under stress through a bulky coordination sphere. This gives a more sterically favourable route to product formation. This is due to an increase in the closeness of the Cu-substrate and aryl groups to be eliminated from the copper(III) complex. Examples of product formation have been seen from Yu,[4] Jones,[5] Sambiagio[2] and more, where, in several of these reactions, the arylated product is formed from a 5-coordinate structure that could not be optimised using solvation in the present study. In the suggested 4-coordinate pathway, all intermediates were able to be optimised in solvation with dispersion corrections and diffuse functions, which suggests that there is an added level of accuracy to these free energy calculations than seen from the aforementioned literature[4] with a more detailed description of the orbitals in each complex.

It is important to assess the barrier connecting structures **Int3** and **TS2**. Despite this barrier having a lower relative free energy than the initial reactants, it still provides a barrier from intermediate Int3 to transition state formation of TS2. This barrier is most likely to be overcome with a reaction temperature above room temperature. A summary of the activation energy from **Int3** to **TS2** for all pathways is shown below in Table 30.

Table 30: Relative barrier from Int3 to TS2 in Cycles O and N.

Pathway	Relative Free Energy of Int3 / kcal mol^{-1}	Relative Free Energy of TS2/ kcal mol^{-1}	TS Activation Barrier / kcal mol^{-1}
Cycle O *Cis*	-38.49	-26.81	+11.68
Cycle O *Trans*	-34.98	-26.77	+8.21
Cycle N *Cis*	-36.64	-29.10	+7.54
Cycle N *Trans*	-34.03	-25.46	+8.57

The greatest barrier to overcome for reductive elimination comes from the *cis* pathway in Cycle O, which is only ~ 3 kcal mol^{-1} greater than the other TS activation barriers in Table 30 shown above, however, this barrier will easily be able to be overcome using the right temperature over an extended period of time, as temperatures between 80-110°C have been shown to be able to aid the reaction. **Int3_O *cis*** has the lowest relative energy of this group, however, requires the greatest amount of energy input to proceed to a transition state. The most favourable route is likely to be from **Int3_N *cis*** to **TS2_N *cis***, as the activation barrier is the lowest to overcome, which may lead to more efficient product formation whilst requiring a lower amount of external input. It is likely that the reaction will be N-selective using the polar solvent and conditions given, with O-selectivity potentially being favoured in a less polar solvent such as toluene. These solvent effects have been seen in previous work by Yu *et al.*[4]

9.5 - Eyring Calculations

One final descriptor for determining how realistic a reaction is to proceed can be shown by estimates of the rate constant (*k*) for a given reaction. The Eyring equation[82] is shown below in equation 4. Where terms used include the Boltzmann constant (k_B), Temperature (T), Planck's constant (*h*), Gas constant (R) and the Gibbs free energy of adjacent intermediates (I_I) and transition states (T_i) of the system ($G^{‡}$ or $I_i - T_i$).

$$k = \frac{k_B T}{h} e^{\left(\frac{-\Delta G^{‡}}{RT}\right)} = \frac{k_B T}{h} e^{\left(\frac{I_i - T_i}{RT}\right)} \quad \text{Eq. 4}$$

The half-life of the reaction can then be determined using the following equation using the calculated rate constant at a certain temperature.

$$t_{1/2} = ln2(k) \qquad \text{Eq. 5}$$

The use of pseudo first order conditions is assumed for these calculations as to make this reaction linearly dependent on the concentration of only one reactant.

Shown in Table 31 below are the rate constants and subsequently half-lives of both *cis* and *trans* barriers to oxidative addition. The values used to calculate the Gibbs free energy of adjacent intermediates and transition states were those that had the greatest energy difference.

Table 31: Rate constant and half-life of one full cycle.

Temperature (K)	Rate Constant (k)	Half-life ($t_{1/2}$) (Hours)
273.15	2.78×10^{3}	6.91×10^{-8}
298.15	1.84×10^{4}	1.05×10^{-8}

At both 273 K and 298 K the single full cycle will proceed at a fraction of an hour. This extremely low half-life is suggestive that the reaction will proceed spontaneously. The barrier analysed is the greatest barrier (11.63 kcal mol^{-1}) and therefore the limiting stage of the reaction.

To fully deduce whether the barrier from Int3 to TS2 is in fact the largest barrier, one cycle may not be appropriate. This was analysed through the use of a second turnover of the cycle. The profile of the second cycle turnover was graphed and is shown below in Figure 29.

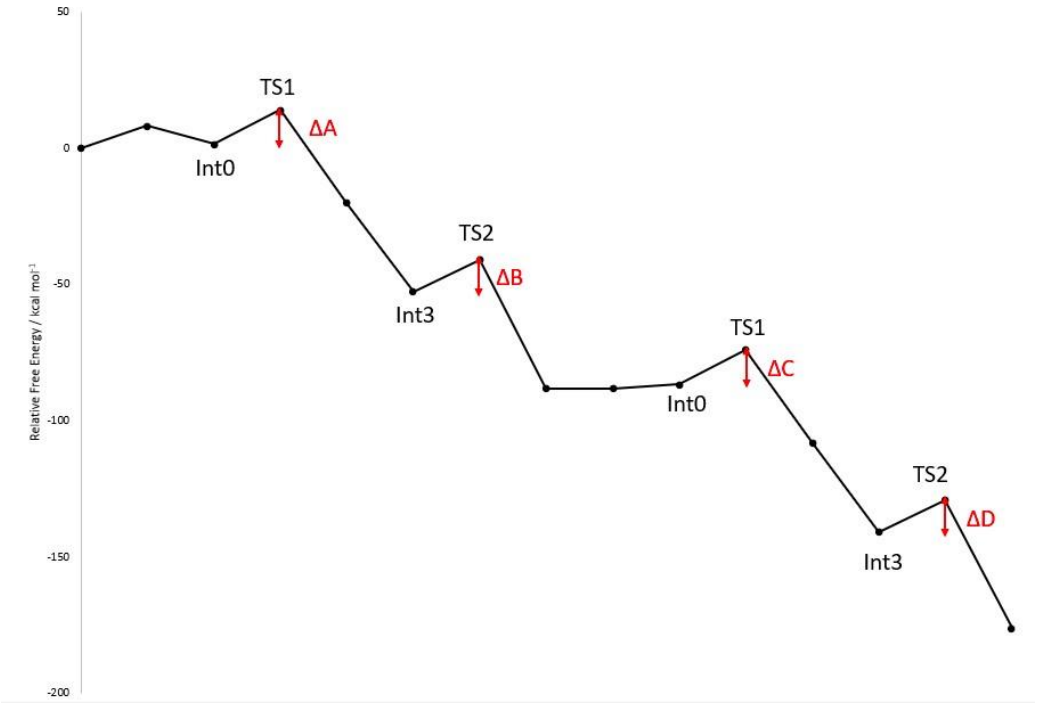


Figure 29: Energy profile of cycle turnover with greatest barriers identified.

This profile shows that the cycle turnover is not the rate determining step for this reaction, therefore it is sensible to presume that this reaction is feasible and if the suggested mechanistic pathway is the preferred route, will turn over without energy input. The greatest barriers are ΔB and ΔD, as previously mentioned as 11.63 kcal mol^{-1}. The barrier of ΔA and ΔC are 6.77 kcal mol^{-1}, which represents the barrier to oxidative addition and is lower than that to reductive elimination.

10 - Conclusions and Future Work

Several goals have been met in this project, the computational method has been examined and refined. There are many relevant Cu(I) and Cu (III) structures that have been analysed from the Cambridge Structural Database, with several outliers being identified and discussed. These structural factors prove invaluable when comparing the geometric effects of different computational methods as they provide citing literature for evidence of bond lengths and angles. Two main mechanistic pathways have been defined and reproduced for both N and O selective products. These pathways showcase the Cu(I)/Cu(III) oxidative addition – reductive elimination route already backed by literature[4] and show that amino alcohols have hemilabile tendencies in the coordination sphere of copper. The most likely cycle also provides a mechanistic pathway which is facilitated by sensible barriers between intermediates, with both N and O selective pathways being energetically feasible. The Narylated product is likely to be the favoured route for this reaction to proceed deduced from a lower reaction energy by -4.32 kcal mol^{-1}.

Method effects were considered, where the addition of implicit solvation caused a stabilizing effect across the board with an average of +2 kcal mol^{-1} increase in free energy in the oxidative addition section of the reaction, whereas an average of -7 kcal mol^{-1} was observed with the reductive elimination section of the reaction. When dispersion corrections were implemented, an average energy change of -18 kcal mol^{-1} was observed. Alteration of the functional used was considered, with a comparison made between B3LYP and BP86 where the potential energy and relative free energy were both examined. The 'best' methodology used throughout was regarded as 'BS3' and used B3LYP/6-311+g(d,p)/SDD with solvation and dispersion included. The B3LYP functional was found to be more accurate than BP86, which agrees with literature from Koch[52] where BP86 and B3LYP functionals tend to have errors of 5 kcal mol^{-1} and 2 kcal mol^{-1} respectively. The use of solvation, dispersion and diffuse functions is necessary to gain the highest level of accuracy for a given system. Using the best computational method shown above provided an insight into the effects of each of these factors, including their benefits. The hybrid B3LYP was the better option for functional choice when paired with the other inputs due to its increased level of accuracy in comparison with pure BP86.

The ongoing work around this project is to continue exploring the possible pathways that amino alcohols of various carbon chain lengths can lead copper catalysed coupling reactions down. For the time being, a realistic pathway centring around 4-coordinate complexes has been hypothesised, however, there is room for improvement, as shown previously with higher coordination numbers. Summarised below are some of the goals and objectives to be considered if this project was continued and are useful general considerations in the field of amino alcohol ligands in copper catalysed coupling reactions.

Another study using the CSD may be useful in order to target some more specific ligands such as phenanthroline or bipyridine, which would show the number of isolated structures that contain these ligands, how they bind to a copper centre and what the effect of surrounding atoms has on these ligands.

One progression to this work would be a more thorough exploration of the similarities between the complexes that have been shown to use a longer carbon chain in the amino alcohols, as this will give more information about product formation and the mono or bidentate nature of amino alcohol ligands. It would also be useful to compare the literature[4] catalytic cycle with dispersion effects added to the geometry input in order to make a more accurate comparison to the data explored with regards to amino alcohol ligands.

It may be beneficial to conduct more research around 2-coordinate copper(I) complexes, as when bound with common bidentate ligands they have been shown to be preferred over 3-coordinate complexes. One area that was left incomplete are cycles only using 2-coordinate complexes as a start and end point due to difficulty coming to a cycle that would turn-over well, despite the possibility of them providing a low energy route to oxidative addition. One of the final main considerations is whether the calculated barriers are low enough for the reaction to proceed experimentally whilst postulating that the catalyst will remain viable whilst incorporating the required coordination of the amino alcohol ligand in order to consider this process ligand assisted. It was shown in Section 9.5 with use of the Eyring equation, where the rate constant and half-life of the reaction at zero degrees was so high that the reaction is highly likely to proceed spontaneously. This is still governed by the choice of starting compound, which is still in question.

11 - Bibliography

1. S. Tasler, J. Mies and M. Lang, *Adv. Synth. Catal.*, 2007, **349**, 2286-2300.
2. C. Sambiagio, S. P. Marsden, A. J. Blacker and P. C. McGowan, *Chem Soc Rev*, 2014, **43**, 3525-3550.
3. L. C. Campeau and N. Hazari, *Organometallics*, 2019, **38**, 3-35.
4. H.-Z. Yu, Y.-Y. Jiang, Y. Fu and L. Liu, *J. Am. Chem. Soc.*, 2010, **132**, 18078-18091.
5. G. O. Jones, P. Liu, K. N. Houk and S. L. Buchwald, *J. Am. Chem. Soc.*, 2010, **132**, 6205-6213.
6. T. Cohen and I. Cristea, *J. Am. Chem. Soc.*, 1976, **98**, 748-753.
7. Q. A. Lo, D. Sale, D. C. Braddock and R. P. Davies, *ACS Catal.*, 2017, **8**, 101-109.
8. G. J. Sherborne, S. Adomeit, R. Menzel, J. Rabeah, A. Bruckner, M. R. Fielding, C. E. Willans and B. N. Nguyen, *Chem Sci*, 2017, **8**, 7203-7210.
9. F. Monnier and M. Taillefer, *Angew. Chem. Int. Ed. Engl.*, 2009, **48**, 6954-6971.
10. N. Kambe, T. Iwasaki and J. Terao, *Chem. Soc. Rev.*, 2011, **40**, 4937-4947.
11. I. P. Beletskaya and A. V. Cheprakov, *Organometallics*, 2012, **31**, 7753-7808.
12. E. R. Strieter, B. Bhayana and S. L. Buchwald, *J. Am. Chem. Soc.*, 2009, **131**, 78-88.
13. A. Klapars and S. L. Buchwald, *J. Am. Chem. Soc.*, 2002, **124**, 14844-14845.
14. J. Tirado-Rives and W. L. Jorgensen, *J Chem Theory Comput*, 2008, **4**, 297-306.
15. Y. Yang, M. N. Weaver and K. M. Merz, Jr., *The journal of physical chemistry. A*, 2009, **113**, 98439851.
16. A. Shafir, P. A. Lichtor and S. L. Buchwald, *J. Am. Chem. Soc.*, 2007, **129**, 3490-3491.
17. M. A. Carvajal, J. J. Novoa and S. Alvarez, *J. Am. Chem. Soc.*, 2004, **126**, 1465-1477.
18. N. Fey, A. Koumi, A. V. Malkov, J. D. Moseley, B. N. Nguyen, S. N. G. Tyler and C. E. Willans, *Dalton Trans*, 2020, **49**, 8169-8178.
19. H. Hirao, *The Journal of Physical Chemistry A*, 2011, **115**, 9308-9313.
20. A. Dkhissi and R. Blossey, *Chem. Phys. Lett.*, 2007, **439**, 35-39.
21. J. W. Tye, Z. Weng, R. Giri and J. F. Hartwig, *Angew. Chem. Int. Ed.*, 2010, **49**, 2185-2189.
22. J. W. Tye, Z. Weng, A. M. Johns, C. D. Incarvito and J. F. Hartwig, *J. Am. Chem. Soc.*, 2008, **130**, 99719983.
23. R. Giri, A. Brusoe, K. Troshin, J. Y. Wang, M. Font and J. F. Hartwig, *J Am Chem Soc*, 2018, **140**, 793806.
24. K. Majid, R. Mushtaq and S. Ahmad, *E-Journal of Chemistry*, 2008, **5**, 680324.
25. N. L. Bell, C. Xu, J. W. B. Fyfe, J. C. Vantourout, J. Brals, S. Chabbra, B. E. Bode, D. B. Cordes, A. M. Z. Slawin, T. M. McGuire and A. J. B. Watson, *Angewandte Chemie International Edition*, 2021, **60**, 7935-7940.
26. S. Ahn, M. Hong, M. Sundararajan, D. H. Ess and M.-H. Baik, *Chemical Reviews*, 2019, **119**, 65096560.
27. L. M. Huffman and S. S. Stahl, *J. Am. Chem. Soc.*, 2008, **130**, 9196-9197.
28. S.-J. Li and Y. Lan, *Chem. Commun.*, 2020, **56**, 6609-6619.
29. E. D. Kalkman, M. G. Mormino and J. F. Hartwig, *J. Am. Chem. Soc.*, 2019, **141**, 19458-19465.
30. D. Chatfield, *Theor. Chem. Acc.*, 2002, **108**, 367-368.
31. J. Jover and N. Fey, *Chemistry – An Asian Journal*, 2014, **9**, 1714-1723.
32. C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, *Acta Crystallographica Section B*, 2016, **72**, 171-179.
33. F. H. Allen and R. Taylor, *Chem. Soc. Rev.*, 2004, **33**, 463-475.
34. S. Sung, D. C. Braddock, A. Armstrong, C. Brennan, D. Sale, A. J. P. White and R. P. Davies, *Chemistry – A European Journal*, 2015, **21**, 7179-7192.
35. A. Ouali, J.-F. Spindler, A. Jutand and M. Taillefer, *Adv. Synth. Catal.*, 2007, **349**, 1906-1916.
36. K. N. Houk and F. Liu, *Acc. Chem. Res.*, 2017, **50**, 539-543.
37. E. G. Lewars, *Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics*, Springer Netherlands, 2010.

38. A. V. Morozov, K. M. S. Misura, K. Tsemekhman and D. Baker, *The Journal of Physical Chemistry B*, 2004, **108**, 8489-8496.
39. J. Foresman and A. Frisch, *Exploring Chemistry With Electronic Structure Methods, 3rd edition*, 2015.
40. P. A. M. Dirac, *Mathematical Proceedings of the Cambridge Philosophical Society*, 1930, **26**, 376-385.
41. P. Hohenberg and W. Kohn, *Physical Review*, 1964, **136**, B864-B871.
42. R. Stowasser and R. Hoffmann, *J. Am. Chem. Soc.*, 1999, **121**, 3414-3420.
43. J. N. Harvey, F. Himo, F. Maseras and L. Perrin, *ACS Catal.*, 2019, **9**, 6803-6813.
44. T. J. Zielinski, E. Harvey, R. Sweeney and D. M. Hanson, *J. Chem. Educ.*, 2005, **82**, 1880.
45. J. P. Perdew and K. Schmidt, *AIP Conference Proceedings*, 2001, **577**, 1-20.
46. C. J. Cramer, *Essentials of computational chemistry : theories and models*, West Sussex, England ; New York : J. Wiley, [2002] ©2002, 2002.
47. K. A. Baseden and J. W. Tye, *J. Chem. Educ.*, 2014, **91**, 2116-2123.
48. N. Emel'yanova, N. Sanina, A. Krivenko, R. Manzhos, K. Bozhenko and S. Aldoshin, *Theor. Chem. Acc.*, 2012, **132**, 1316.
49. Y. Minenkov, Å. Singstad, G. Occhipinti and V. R. Jensen, *Dalton Transactions*, 2012, **41**, 5526-5541.
50. Y. Zhang, X. Xu and W. A. Goddard, *Proceedings of the National Academy of Sciences*, 2009, **106**, 4963.
51. S. Grimme, A. Hansen, J. G. Brandenburg and C. Bannwarth, *Chemical Reviews*, 2016, **116**, 51055154.
52. W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, Wiley, 2015.
53. P. J. Hay and W. R. Wadt, *The Journal of Chemical Physics*, 1985, **82**, 299-310.
54. J. J. Belbruno, *Heteroat. Chem*, 1998, **9**, 651-657.
55. J. Kreutzer, P. Blaha and U. Schubert, *Computational and Theoretical Chemistry*, 2016, **1084**, 162168.
56. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, 2016.
57. E. L. M. Miguel, *How Accurate is the SMD Model for Predicting Free Energy Barriers for Nucleophilic Substitution Reactions in Polar Protic and Dipolar Aprotic Solvents?*, 2016.
58. A. D. Becke, *The Journal of Chemical Physics*, 1993, **98**, 5648-5652.
59. L. Li, J. Hu, X. Shi, W. Ruan, J. Luo and X. Wei, *Int J Mol Sci*, 2016, **17**, 927.
60. A. Bauzá, D. Quiñonero, P. M. Deyà and A. Frontera, *The Journal of Physical Chemistry A*, 2013, **117**, 2651-2655.
61. R. Dennington, T. Keith and J. Milliam, GaussView, ***6.1***, 2019.
62. A. Brakestad, P. Wind, S. R. Jensen, L. Frediani and K. H. Hopmann, *The Journal of Chemical Physics*, 2021, **154**, 214302.
63. R. M. Richard, B. W. Bakr and C. D. Sherrill, *Journal of Chemical Theory and Computation*, 2018, **14**, 2386-2400.
64. Ł. M. Mentel and E. J. Baerends, *Journal of Chemical Theory and Computation*, 2014, **10**, 252-267.
65. S. F. Boys and F. Bernardi, *Mol. Phys.*, 1970, **19**, 553-566.
66. A. Orpen, *Acta Crystallographica Section B*, 2002, **58**, 398-406.
67. N. Fey, S. E. Harris, J. N. Harvey and A. G. Orpen, *Journal of Chemical Information and Modeling*, 2006, **46**, 912-929.

68. P. R. Raithby, G. P. Shields, F. H. Allen and W. D. S. Motherwell, *Acta Crystallographica Section B*, 2000, **56**, 444-454.
69. P. Atkins and T. Overton, *Shriver and Atkins' Inorganic Chemistry*, OUP Oxford, 2010.
70. K. M. Anderson and A. G. Orpen, *Chem. Commun.*, 2001, DOI: 10.1039/B108517B, 2682-2683.
71. T. Onishi, in *Adv. Quantum Chem.*, eds. J. R. Sabin and E. J. Brändas, Academic Press, 2012, vol. 64, pp. 31-81.
72. R. H. Crabtree, *The Organometallic Chemistry of the Transition Metals*, Wiley, 4th edn., 2005.
73. M. Mantina, A. C. Chamberlin, R. Valero, C. J. Cramer and D. G. Truhlar, *The Journal of Physical Chemistry A*, 2009, **113**, 5806-5812.
74. E. N. Golubeva, E. M. Zubanova and G. M. Zhidomirov, *J. Phys. Org. Chem.*, 2013, **26**, 724-729.
75. C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, 2019, DOI: 10.1107/S2052520616003954.
76. S.-L. Zhang, L. Liu, Y. Fu and Q.-X. Guo, *Organometallics*, 2007, **26**, 4546-4554.
77. L. Sacconi, *Pure Appl. Chem.*, 1968, **17**, 95-128.
78. J. Stanley-Gray, Z. Zhang and D. Venkataraman, *J. Chem. Educ.*, 2021, DOI: 10.1021/acs.jchemed.1c00364.
79. C. L. McMullin, J. Jover, J. N. Harvey and N. Fey, *Dalton Transactions*, 2010, **39**, 10833-10836.
80. R. Tonner, *ChemPhysChem*, 2011, **12**, 2352-2352.
81. D. Sasi, V. Ramkumar and N. N. Murthy, *ACS Omega*, 2017, **2**, 2474-2481.
82. V. P. Ananikov, *Understanding Organometallic Reaction Mechanisms and Catalysis: Computational and Experimental Tools*, Wiley, 2014.
83. T. E. Patten, C. Troeltzsch and M. M. Olmstead, *Inorg. Chem.*, 2005, **44**, 9197-9206.
84. R. Rabie, M. M. Hammouda and K. M. Elattar, *Research on Chemical Intermediates*, 2017, **43**, 19792015.
85. G. Franc and A. Jutand, *Dalton Transactions*, 2010, **39**, 7873-7875.
86. K. Popov, L. H. J. Lajunen, A. Popov, H. Rönkkömäki, M. Hannu-Kuure and A. Vendilo, *Inorg. Chem. Commun.*, 2002, **5**, 223-225.
87. E. L. Robert, *The Sigma-Aldrich library of chemical safety data*, Ed. 2. [Milwaukee, Wis., USA] : SigmaAldrich Corp., [1988] ©1988, 1988.
88. Y. D. Wu and D. K. Lai, *J Org Chem*, 1996, **61**, 7904-7910.
89. A. Casitas, M. Canta, M. Solà, M. Costas and X. Ribas, *J. Am. Chem. Soc.*, 2011, **133**, 19386-19392.
90. J. Pan, S. Yang, Y. Li, L. Han, X. Li and Y. Cui, *Crystal Growth & Design*, 2009, **9**, 3825-3827.
91. E. Vega, E. de Julián, G. Borrajo, J. Díez, E. Lastra and M. P. Gamasa, *Polyhedron*, 2015, **94**, 59-66.
92. K. T. Mahmudov, M. Sutradhar, L. M. D. R. S. Martins, M. F. C. Guedes da Silva, A. Ribera, A. V. M. Nunes, S. I. Gahramanova, F. Marchetti and A. J. L. Pombeiro, *RSC Advances*, 2015, **5**, 25979-25987.
93. O. Hietsoi, C. Dubceac, A. S. Filatov and M. A. Petrukhina, *Chem. Commun.*, 2011, **47**, 6939-6941.

12 - Appendix

12.1 – Restrictions and isomers of copper complexes in the CSD.

When using specific angle restrictions, search 1 resulted in 546 square planar complexes with the Cl ligand *cis* to the NR_2 group and 415 square planar complexes with the Cl *trans* to the NR_2, as seen below in Table 32. A CSD search regarding 4 coordinate complexes of c opper with at least one N and one Cl ligand returned 1398 results. This suggests that the restrictions on the bond angle was too tight and did not allow enough steric influence of large ligands and R groups to be accounted for. Upon changing the trans restriction to 160 – 180, the number of *trans* search results increased from 271 to 415. This brings the collective *cis/trans* copper centres to 961, which is a more representative amount of data. There are still structures not accounted for in this search, which is due to using tight restrictions in the queries in order to lower the number of structures seen to be too unlike a ligated complex such as multiple crystals arranged in a cubic lattice, space group restrictions could have been used to further limit the occurrences of these structures. These comparisons were also done for searches 2 and 3, and all the octahedral complexes used as shown below in Table 32.

Table 32: Number of Cis and Trans isomers present in CSD searches.

Search	Complex	*Cis* Isomers	*Trans* Isomers	Total structures	Number of refcodes containing multiple hits
1	Square Planar	546	415	1398	0
	Octahedral	376	118	450	97
2	Square Planar	4334	2686	5592	2096
	Octahedral	5090	2005	5715	1927
3	Square Planar	238	104	389	0
	Octahedral	225	105	311	72

In some query results there were multiple search ‘hits’ found by the CSD. This is indicative that in many of the structures found there is variance in bond lengths, with possible *cis/trans* isomers in the same structure. Whilst this is represented by recording each measured parameter, it results in certain refcodes being repeated. For example, structure AGAWUX[32] was found and is shown below

in Figure 29. This structure contains multiple Cu-N and Cu-O bonds and therefore is picked up twice in the CSD search for each bond parameter. This is one limitation of the CSD as a structure search, as in structures that are mirrored as shown below, the average bond length overall can be skewed. This would be an error that could be fixed by manually checking each structure, however, for many of these searches of hundreds or thousands of structures, this is not feasible. Many of the structures with multiple hits for specific bond lengths were found in search 2 due to the presence of N and O being defined in the search query.

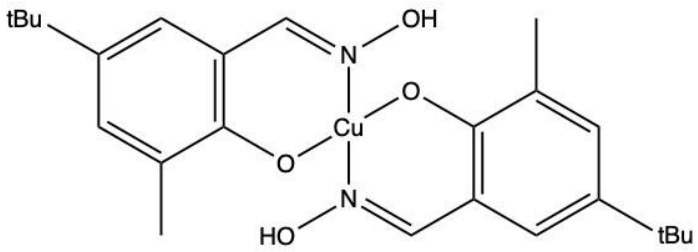


Figure 30: Structure of AGAWUX showing mirrored Cu-N and Cu-O bonds

It is also likely that where the results percentage is greater than 100%, there is an overlap of results due to some octahedral complexes containing multiple *cis* or *trans* bonds, which have therefore contributed to both the cis and trans searches. For example, in search 1 octahedral, the percentage was 109.7%.

12.1 – CSD Graphs

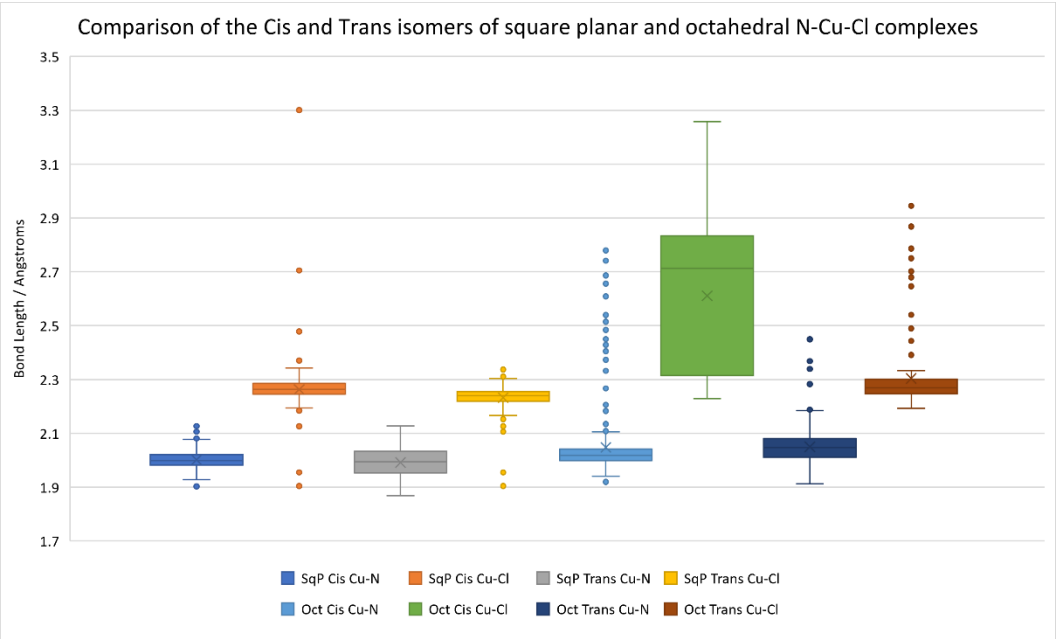


Figure 31: Cis/Trans isomer comparison of square planar and octahedral N-Cu-Cl complexes

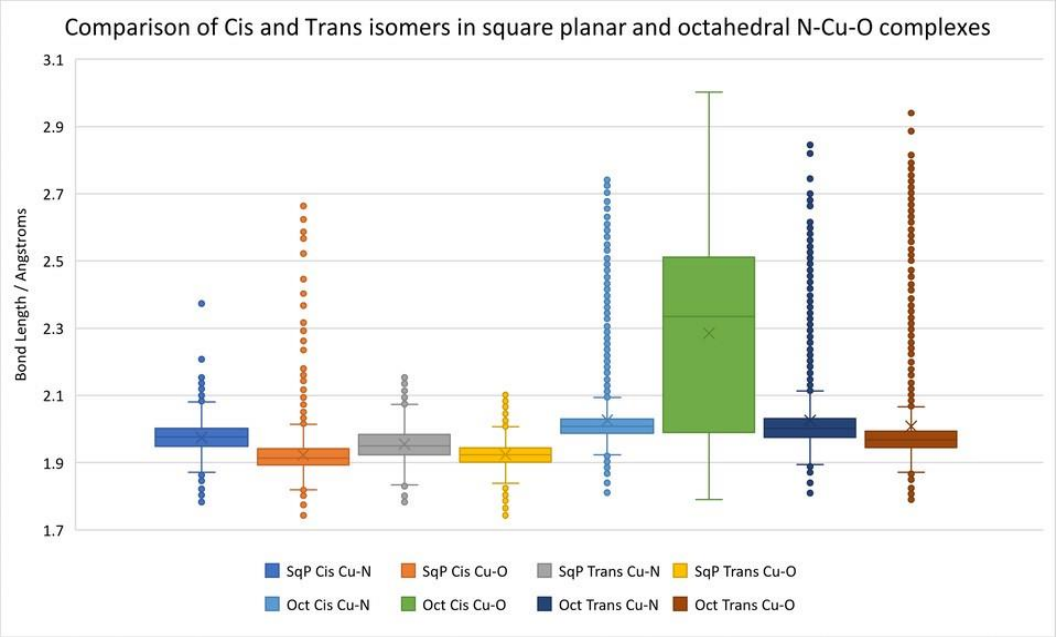


Figure 32: Cis/Trans isomer comparison of square planar and octahedral N-Cu-O complexes

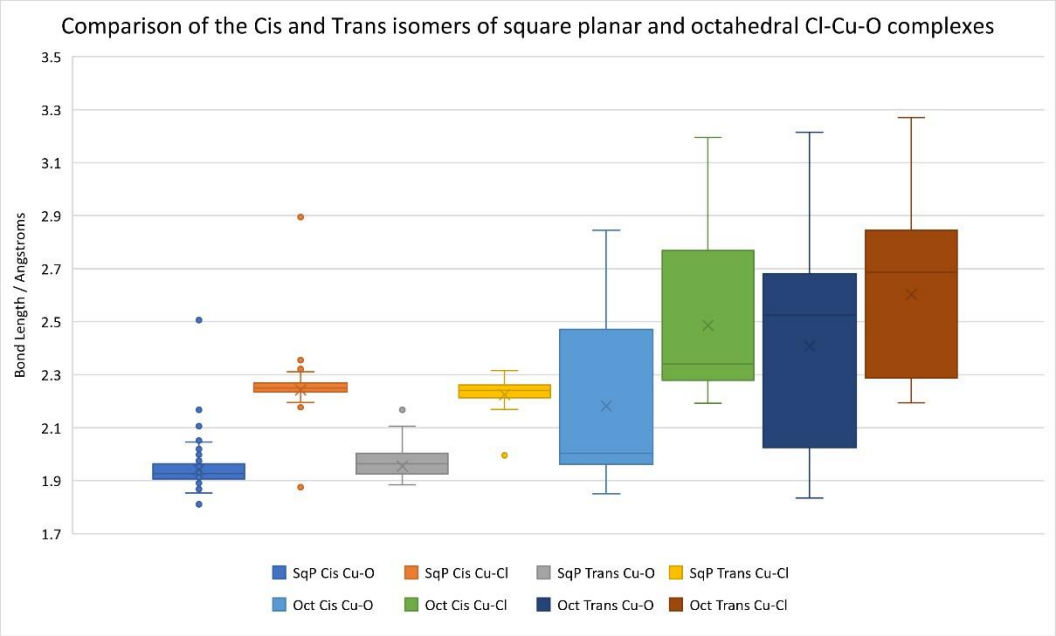


Figure 33: Cis /Trans isomer comparison of square planar and octahedral Cl-Cu-O complexes

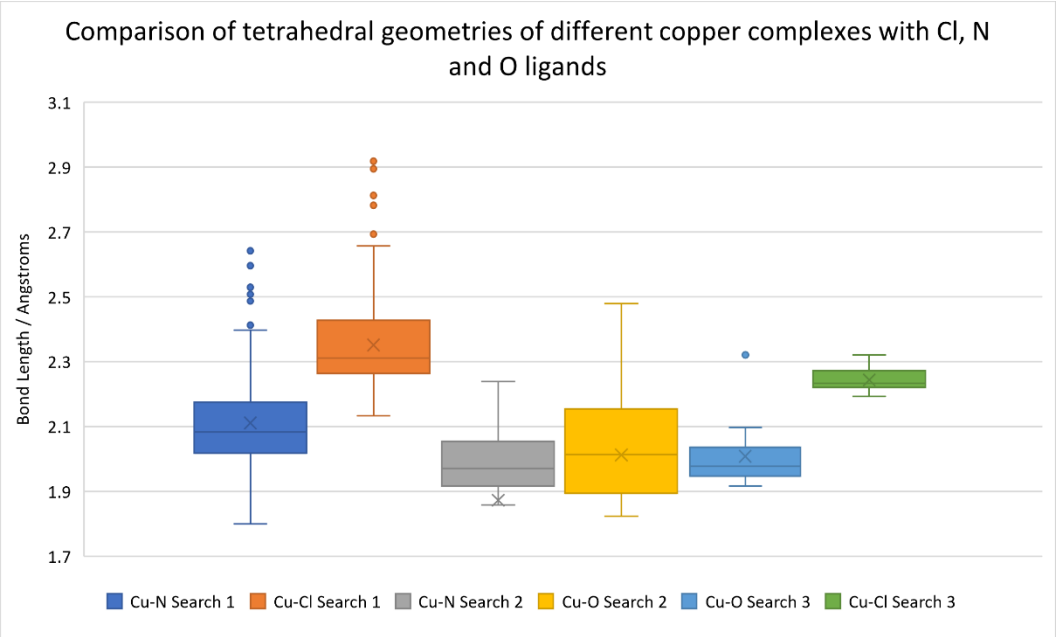


Figure 34: Comparison of tetrahedral Cu complexes with Cl, N and O ligands

12.2 – CSD Outliers

There are many outliers in this data, two extreme outliers from each figure have been chosen for comparison, with data collated below in Table 32.

Table 33: Specific selected outliers for further exploration.

Complex	CSD Ref-Code	Geometry	Outlying bond	Bond Length / A
N-Cu-Cl	NUGFAY	Square Planar	Cu-Cl	3.301
	UKAYIP	Octahedral	Cu-N	2.779
N-Cu-O	KITMEH	Square Planar	Cu-O	2.664
	RUCZIA	Octahedral	Cu-N	2.741
O-Cu-Cl	ITEGAQ	Square Planar	Cu-Cl	2.895
	VEPHIK	Square Planar	Cu-O	2.506

The structure of complex 'NUGFAY'[32] can be seen below in Figure 34 as taken from the CSD alongside a clearer diagram of the central unit structure, which is large and deviates the most from the square planar geometry. As can be seen, the copper centre is closer to a tetrahedral arrangement than a square planar one, which has been forced due to the cubic central arrangement of a combination of copper chlorine bonds. The oxidation state of copper in this structure is +1, however, it belongs in this tetramer cubane-type structure where the Cu(I) are arranged in distorted tetrahedral geometries.[91]

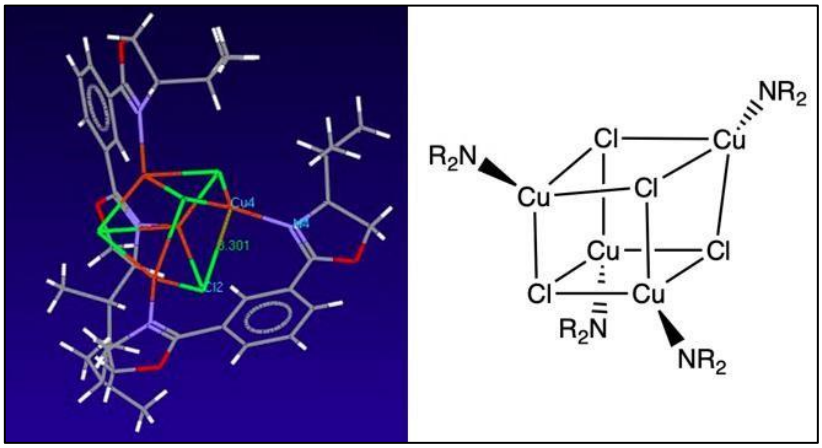


Figure 35: 3D crystal structure and diagram of outlier 'NUGFAY' with labelled relevant Cu and Cl atoms.

The structural geometry of complex 'RUCZIA', shown below in Figure 35,[32] is surrounding the central copper ion is in the (II) state[92] with a coordination of 6 and is bonded to 5 O ligands and one N ligand, creating a very repulsive polar environment. The N donor ligand in question forms a very long bond with the copper as shown below in figure 8 of 2.741 Å, which compared to the mean CuN bond length of 2.026 is extremely long; this may be explained through this structure having uniform

crystal packing, causing symmetry in the molecule. The N donor is also triply bonded to another N atom, which suggests that the Cu-N bond is weak due to the high electron density being held in the triple bond. The complex is unlikely to form a strong N bond due to the O donor atom being doubly bonded elsewhere, which restricts the possible movement. When comparing the *cis* and *trans* isomers of N-Cu-O complexes, there were many datapoints outside of the IQR and maximum data range. These points are not necessarily outliers; however, they do deviate from the mean which is indicative of quite sterically challenging complexes. In these complexes there are several other polar ligands in the coordinate system causing the increase in bond length.

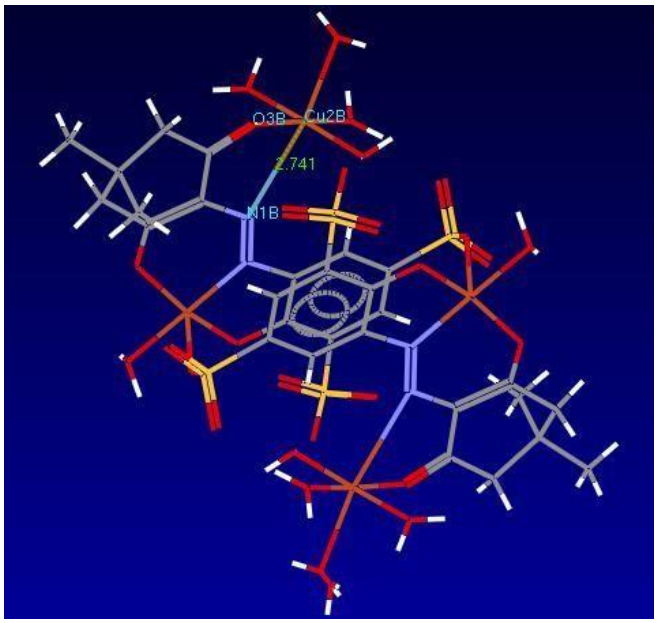


Figure 36: 3D crystal structure of outlier 'RUCZIA' with labelled relevant atoms.

The structure of the outlier 'ITEGAQ'[32] is shown below in Figure 35. This copper(I) centre is bonded to two oxygen donor ligands, one chlorine donor and to another copper metal centre creating a dinuclear unit. The R groups bonded to the oxygen atoms are large and force the O-Cu-Cu bond to be slightly distorted. Alongside this, steric hinderance from the large alkyl R groups causes elongation of the Cl ligand which also has a sterically unfavourable aryl R group. The Cl ligand is present in this structure as it is axially coordinated to the Cu(I) from the solvent.[93] The bond is longer than expected for a square planar copper centre and causes a substantial change in bond angle, forming a geometry like tetrahedral. As referred to regarding structure RUZICA above, this may be largely due to crystal packing rather than an outright bond due to the known presence of

solvent molecules. This provides a further look into the reliability of CSD search queries alone, and that for a more in-depth study of copper geometries, it is important to manually look through some, if not all, structures in order to assess oxidation states and surrounding molecules. This is a process that has been undertaken to critically analyse the accuracy of the database for searching ligands and complexes.[68]

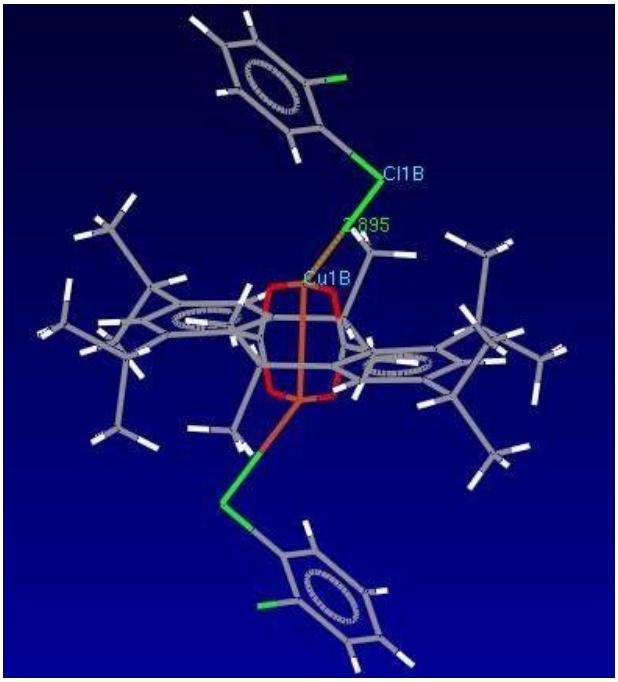


Figure 37: 3D crystal structure of outlier 'ITEGAQ' with labelled relevant atoms.