## Brief Outline

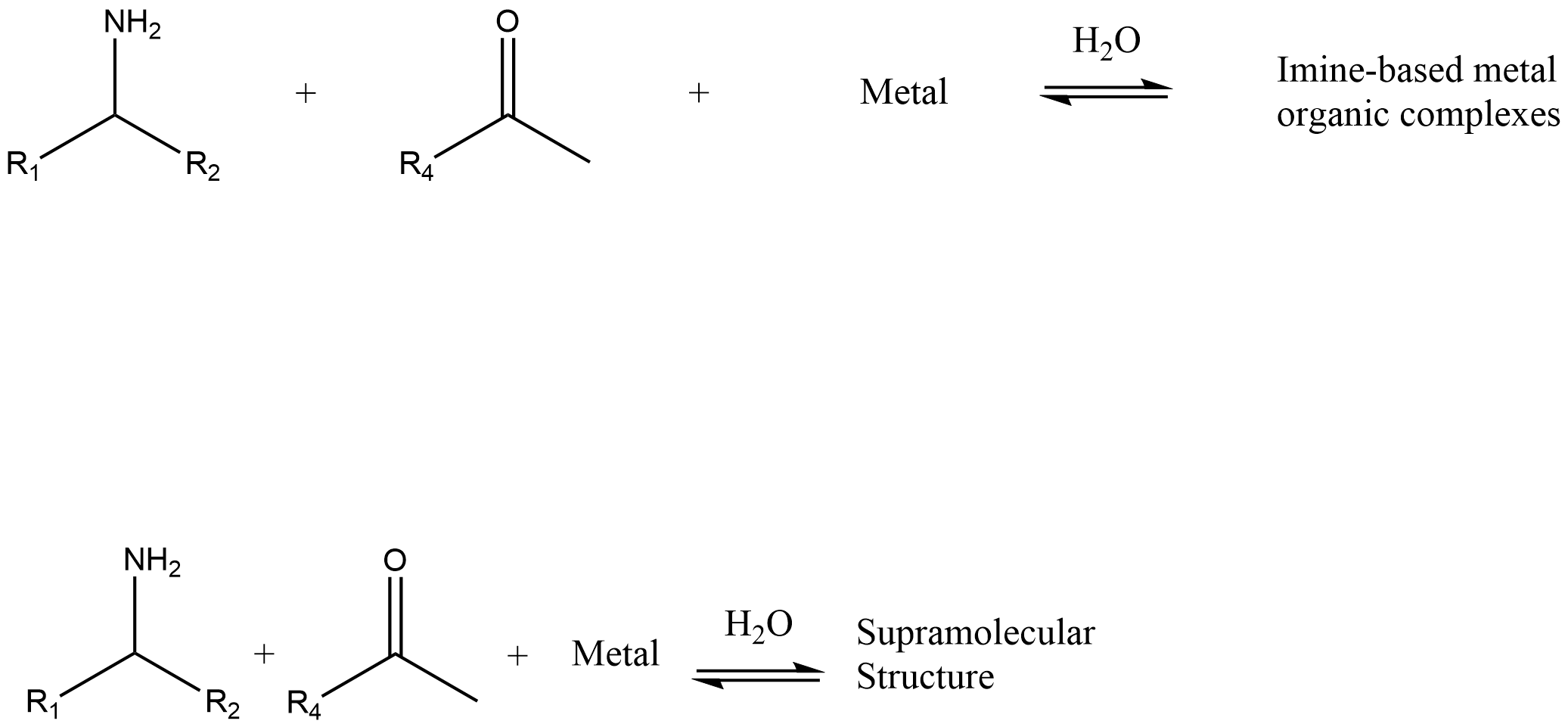


Figure 1: The reaction being carried out in the study.

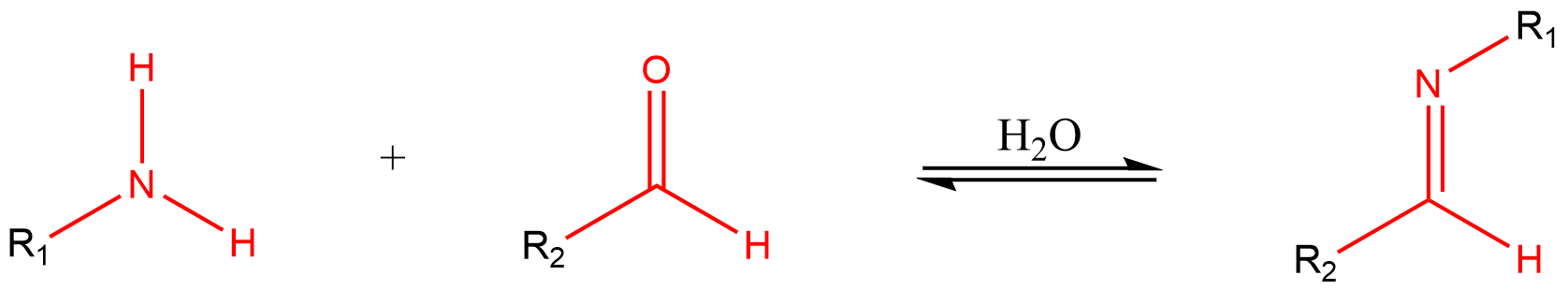
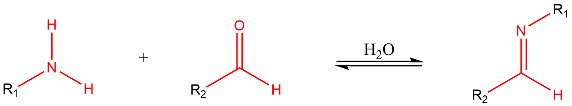
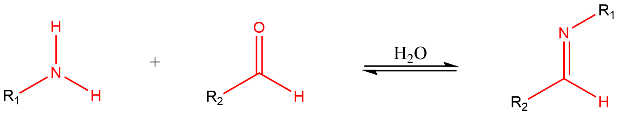
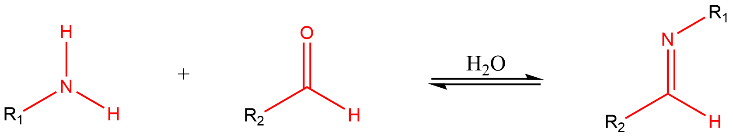


Figure 2: An typical condensation reaction where an amine and aldehyde form an imine.

In each reaction there is one amine, one aldhdye and one metal being reacted. See Figure 1. The amine and the aldehyde react to form and imine (Figure 2). In a chemical reaction, the key areas of interest are called functional group. These are highlighted in red in Figure 2. Essentially, R1 and R­2 could be anything as long as they have the amine  and aldehyde  functional group in their structure. These two functional groups then form an imine . In reality its more complex and there are additional factors that determine if the amine and aldehyde groups react. These factors are also encompassed in the features (see later section).

In chemistry, there are specific elements that have the ability to coordinate (bind to) metals. One of these is the nitrogen. Since there is a nitrogen present in our imine, this can then bind to our metal. The iminie-metal bound product is called an Imine based metal organic complex. (Figure 1).

Figure 3 shows an example formation of a complex. Can you spot the amine, aldhdye, imine functional groups? Can you spot the bond formed between the nitrogen and the metal?

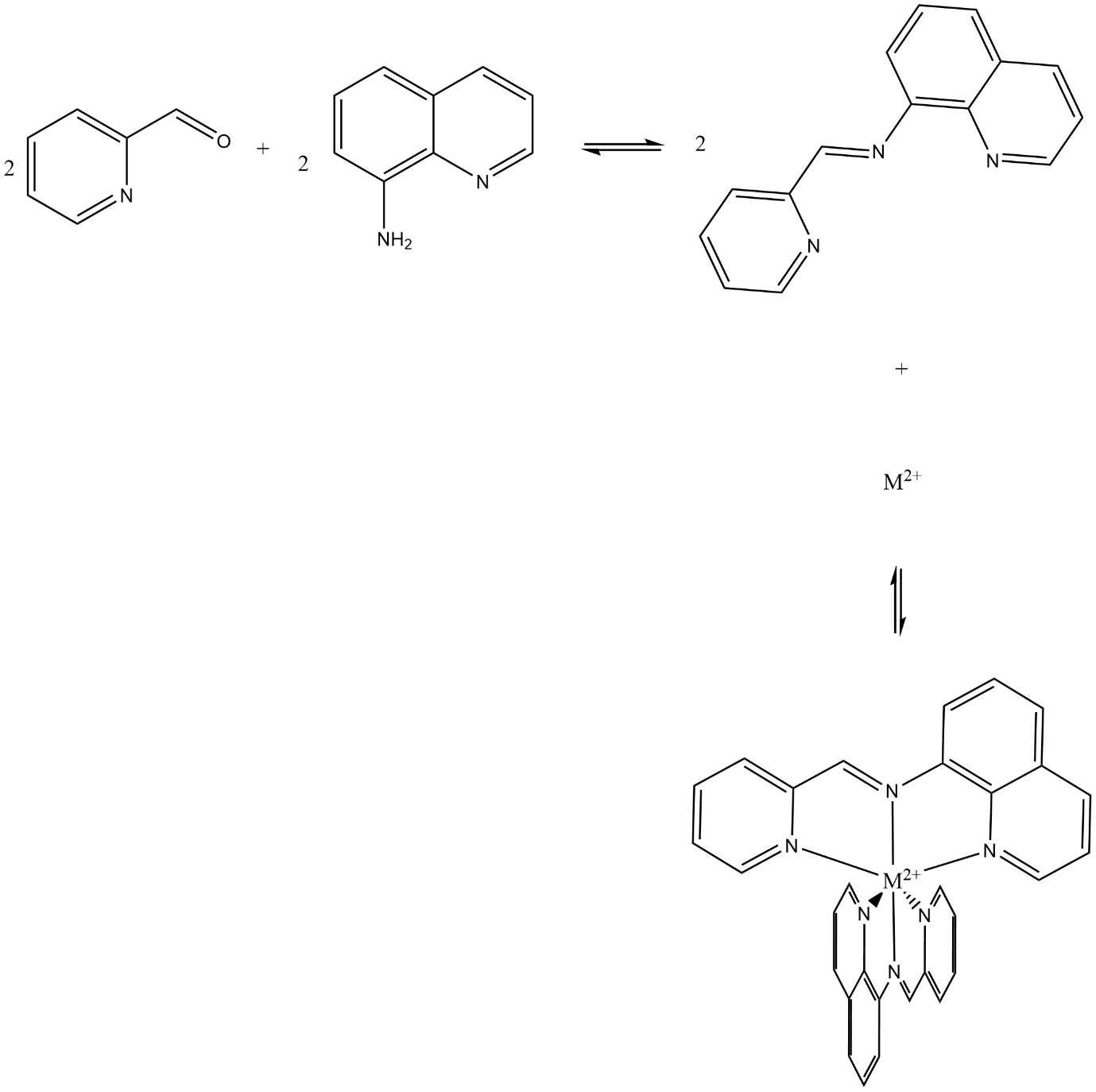


Figure 3: The formation of a Imine based metal organic complex.

The example in Figure 3 is simple, and the reaction actually uses something called a diamine. This is a compound with two amine functional groups. Can you spot them in Figure 4. The use of a diamine allows for the formation of more crazy complexes, and example is shown in Figure 4, where there are two metals in a complex!.

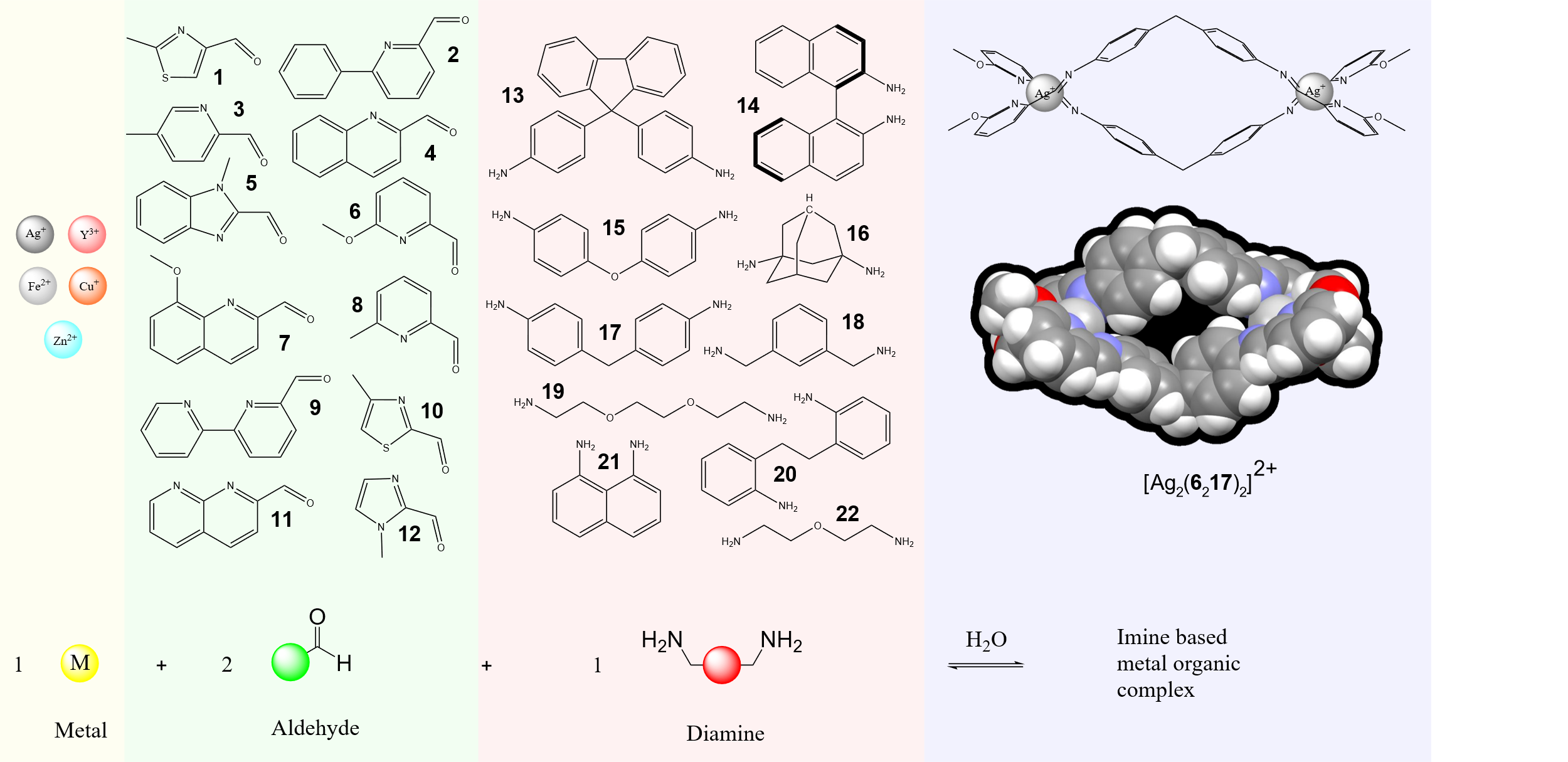


Figure 4: the chemicals and the reaction used in the study.

As Figure 4 shows, you can have a variety of metals, aldehydes, and diamines. There are a few chemical differences that determine the structure of the complex and or if one is even formed. As mentioned previously this can be subdivided into two: factors that affect bonding (i.e. formation of imine, binding of nitrogen from imine to metal) and factors that affect the architecture of the complex (i.e. number of metals in a complex, the geometry of the complex, the size of the complex). Factors that affect architecture are highlighted in red, factors that affect bonding are highlighted in blue, molecular representations or additional information are highlighted in black, any factors in ‘’ are keys in the python compound information dictionary. These factors are shown below and are also used as features for ML. How the factors affect chemistry is also given.

Metal:

* 'metalSize'
* 'metalCharge'
* 'preferredCoordinationGeometry'
* Number of coordination
* 'id'
* 'smiles'
* 'formula'
* 'molecularWeight'
* 'commonName'
* 'chemSpiderObject'
* 'chemicalClass'
* 'morganFingerPrint'

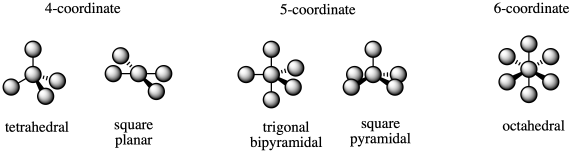
Aldehdye:

* ‘ringSize’
* ‘substituentVolume’
* 'coordinatingNitrogenPartialCharge'
* 'carbonCarbonylPartialCharge'
* 'oxygenCarbonylPartialCharge'
* ‘HomoLumoGap’
* 'id’
* 'smiles'
* 'formula'
* 'molecularWeight'
* 'commonName'
* 'chemSpiderObject'
* 'chemicalClass'
* 'morganFingerPrint'
* 'spectrumIR'
* 'atomCoodinates'

Amine:

* 'amineDistance'
* 'amineRigidity'
* 'numRotatableBonds'
* 'amineDihedralAngle'
* 'numPrimaryAmines'
* 'numAromaticAmimes'
* 'amineClass'
* 'nitrogen1PartialCharge'
* 'nitrogen2PartialCharge'
* ‘HomoLumoGap’
* 'id'
* 'smiles'
* 'formula'
* 'molecularWeight'
* 'commonName'
* 'chemSpiderObject'
* 'chemicalClass'
* 'morganFingerPrint'
* 'spectrumIR'
* 'atomCoodinates'

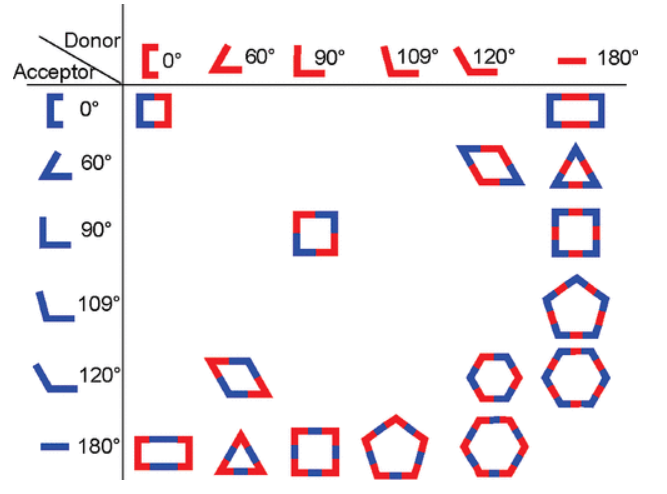
Metal:

* 'preferredCoordinationGeometry': This is the shape the ligand-imine bond forms. There are many types (see picture below for an example). These affect topology of the complex. 
* Number of coordination: Again, metals can have different coordination numbers (see picture above). These affect the topology of the complex. The higher coordination numbers allow for the formation of more complex structures at smaller scales.
* 'metalSize': The smaller the size, the stronger the metal-imine bond.
* 'metalCharge': The greater the positive charge the stronger the metal-imine bond. At a simple level you can think of this as a stronger magnet formed with a larger positive charge, attracting the weak imine magnet. The size of the metal is based on ionic radii.
* 'id': this is the chemspider id to be used in scipy API
* 'smiles': this is a string representation of compound (metal).
* 'formula': this is the composition of the metal compound used in the experiment.
* 'molecularWeight': this is the relative weight of the compound (metal).
* 'commonName': this is the generic name of the compound (metal).
* 'chemSpiderObject': this is a chemspider object, allows user to do further rdkit calculations.
* 'chemicalClass': this is the class of the compound (the possible classes are: Metal, Diamine, Monoaldhyde, Dialdehyde, Monoamine). The class for the metal is: Metal.
* 'morganFingerPrint’: this is a binary representation of the metal.

Aldehdye:

* ‘ringSize’: this is the size of the ring attached to the aldehyde group. The larger ring size makes the aldhdye group less reactive as there are more double bonds available to stabilise the carbonyl (see resonance).
* ‘substituentVolume’: This is the volume of the bulky groups in alpha position (next to the nitrogen) of nitrogen, since the metal links different imines together, there must be enough space to do so, if there is a bulky group in the way, the imine bump into eachother to much making it more difficult to link two imines together.
* 'coordinatingNitrogenPartialCharge': the amount of electrons on the nitrogen affects the formation of the metal-imine bond. Remember this bond is formed via the nitrogens of the imine. The more electrons there are the greater the attraction to the electron poor metal: think of it as a magnet, more electrons = stronger attraction to electron poor areas.
* 'carbonCarbonylPartialCharge': this affects the formation of the imine. The carbon does not have as many electrons as the nitrogen in the amine. Like a magnet, the less electrons there are in the carbon the more strongly attacked the nitrogen is to it.
* 'oxygenCarbonylPartialCharge': To help remove even more electrons from the carbon and making it more reactive to the amine nitrogen, the oxygen helps pull electrons away. Like a magnet, the more electrons on oxygen, the more it pulls the electrons away from the carbon (since they are bonded next to each other). This makes the carbon have less electrons and like a magnet makes it more attractive to the electron rich nitrogen in the amine.
* ‘HomoLumoGap’: In chemistry to break a bond, electrons require a specific amount of energy. The difference between that energy and the energy of the current electrons is the homo lumo gap.
* 'id’: this is the chemspider id to be used in scipy API.
* 'smiles': this is a string representation of compound (aldehdye).
* 'formula': this is the composition of the aldehdye compound used in the experiment.
* 'molecularWeight': this is the relative weight of the compound (aldhdye).
* 'commonName': this is the generic name of the compound (aldhdye).
* 'chemSpiderObject': this is a chemspider object, allows user to do further rdkit calculations.
* 'chemicalClass': this is the class of the compound (the possible classes are: Metal, Diamine, Monoaldhyde, Dialdehyde, Monoamine). The class for the aldehdye is: Monoaldhdye.
* 'morganFingerPrint': this is a binary representation of the metal.
* 'spectrumIR': this is the simulated IR spectrum of the compound (aldhdye). It provides insights into the reactivity of the aldhdye.
* 'atomCoodinates': these are the atoms type and their xyz coodinates (in that order) that make the compound (aldhdye).

Amine:

* 'amineDistance': This is the bond distance between the two functional groups (the number of bonds between the two amine functional groups). This is a proxy for the length between metals, which affects topoly (see picture in 'amineDihedralAngle').
* 'amineRigidity': This is the standard deviation of the angles between the two amine functional groups (see 'amineDihedralAngle') of different conformers (same compound slightly different xyz coordinates of their atoms). This is another proxy for the rigidity of the amine, which affects the topology of the complex.
* 'numRotatableBonds': this is how floppy the compounds are. You can image the more floppy they are the harder it is to get a nice ordered structure. However if its to rigid, you can only get simple topologies.
* 'amineDihedralAngle': this is the angle between the two amine functional groups. This affect topology. See the image below. 
* 'numPrimaryAmines': this is the number of amine functional groups with two hydrogens. Since the hydrogens are small these are more reactive as the compounds areas that don’t react don’t crash into each other as often.
* 'numAromaticAmimes': this the number amine functional groups in a ring. This affects the partial charge of the nitrogen.
* 'amineClass': this is the type of amine present in the compound. The possible options are: alkyl and aromatic amine; alkyl amine; aromatic amine. This affects reactivity as the ring, affects the partial charge of the nitrogen. The ring can remove partial charge, and therefore makes it less reactive towards the electron poor aldehyde. Just like a magnet the positive is attracted to the negative.
* 'nitrogen1PartialCharge' / 'nitrogen2PartialCharge': This is gives an idea of how many electrons are in the nitrogen. Due to the lack of electrons on the carbon of the carbonyl on the aldhdye, the nitrogen of the amine is attracted to it and forms a bond (see formation of imine Figure 2). The more electrons there are on the nitrogen the more reactive with the carbon. You can think of it like a magnet.
* ‘HomoLumoGap’: In chemistry to break a bond, electrons require a specific amount of energy. The difference between that energy and the energy of the current electrons is the homo lumo gap.
* 'id': this is the chemspider id to be used in scipy API.
* 'smiles': this is a string representation of compound (aldehdye).
* 'formula': this is the composition of the amine compound used in the experiment.
* 'molecularWeight': this is the relative weight of the compound (aldhdye).
* 'commonName': this is the generic name of the compound (amine).
* 'chemSpiderObject': this is a chemspider object, allows user to do further rdkit calculations.
* 'chemicalClass': this is the class of the compound (the possible classes are: Metal, Diamine, Monoaldhyde, Dialdehyde, Monoamine). The class for the amine is: Diamine.
* 'morganFingerPrint': this is a binary representation of the metal.
* 'spectrumIR': this is the simulated IR spectrum of the compound (amine). It provides insights into the reactivity of the amine.
* 'atomCoodinates': these are the atoms type and their xyz coodinates (in that order) that make the compound (amine).