CHAPTER 3

Pharmacophore modelling studies:

3.1   Introduction:

According to Langer & Wolber (2004), the key goal of computer-aided molecular design methods in medicinal chemistry is to reduce the time and cost associated with drug discovery and development by identifying the most promising candidates on which further development efforts can be focussed. The amount of compounds that can be tested *in vitro* is limited due to the costly and time-consuming nature of screening thousands of compounds until a promising hit is found. Computational modelling has helped researchers to understand the key interactions between a ligand and a macromolecule and to increase the hit rates in experimental screening subsets that have been subjected to *in silico* screening before they were analyzed *in vitro.*

According to Seidel *et al.* (2010), a pharmacophore is not the representation of a real molecule or the real association between functional groups; it is merely an abstract concept that describes the common steric and electrostatic complimentaries of bio-active molecules with the target of interest. 3D pharmacophore models are the concrete three-dimensional representation of pharmacophore features. In order for the model to have adequate predictive powers, it must be able to describe the nature and location of the functional groups that are involved in the interactions between the ligand and the target receptor, and define the types and characteristics of non-covalent bonds in a uniform and understandable way. This can be achieved by categorizing the interactions into pharmacophoric features such as hydrogen bond donors and acceptors, positively and negatively charged groups and hydrophobic regions. The 3D location of the interactions and the spatial orientation of directed features are also important.
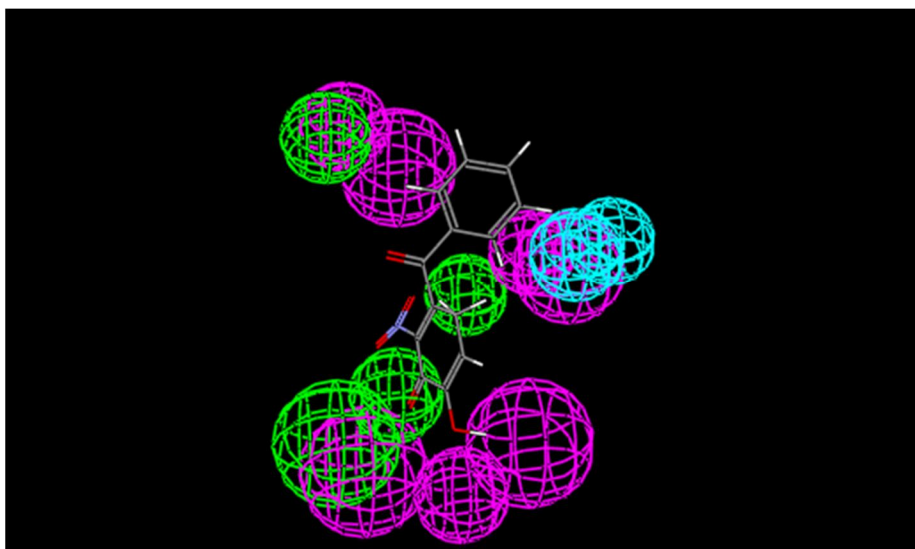
Figure 3.1.1 A *structure-based* pharmacophore model. The green spheres represent hydrogen bond acceptor features, the purple spheres represent hydrogen bond donor features and the cyan spheres represent hydrophobic features.

Various screening platforms can be used to build pharmacophore models. In this study Accelrys Discovery Studio® was used. Due to their abstract nature and simplicity, 3D pharmacophore models are efficient filters for the screening of large drug databases. Lower overall search times are achieved by reducing the complexity of the hit identification process due to the sparse pharmacophoric representation of ligand-target interactions. The pharmacophore based queries identifies novel drug candidates with different scaffolds and functional groups than the original ligands used in the pharmacophore design. This can assist companies to develop novel compounds with better pharmacokinetic and toxicological properties. However, the pharmacophoric representations are greatly simplified and therefore they cannot explain the complete biophysical nature of drug interactions (Seidel *et al.*, 2010).

The workflow of the virtual screening of compounds can be broken down into several well defined steps as illustrated in figure 3.1.5. First a query pharmacophore model is created that specifies the type and geometrical constraints of the chemical features that need to be matched by the screened molecules. Pharmacophore models can be created using structure- or ligand-based approaches. This study used a *structure-based* method wherein chemical features are determined using the complementaries between a ligand and its binding site. This method requires the structural information of the macromolecule and the active conformation of the bound ligand. Because it incorporates the binding site interactions, it often forms a highly restrictive model with orientation constrained features (Seidel *et al.*, 2010).

The next step is the creation of the ligand database to be screened. The database has to be conformationally flexible and this is achieved by recording pre-computed conformations of all the molecules that will be screened. Then the database is searched for conformations that will bind to the macromolecule (represented by the pharmacophore model) using a multistep filtering process. First a fast pre-filtering step is applied, where molecules that do not fit into the macromolecule are eliminated on the basis of their feature-types, feature-counts and quick distance checks. Molecules that might fit the query are closely examined to see whether their conformations are able to match the spatial arrangement of the query features using 3D matching algorithms. The latter process is slower, but more restrictive. The conformers that satisfy the query are put into a hit-list. In order to be on the hit-list, Discovery studio® calculates the root mean square deviation of the interaction and only selects interactions with a value below 1. The distance between each pair of features of the database molecule must be so that it is possible to align them within the specific tolerances of the mapped query features. Fit-values are then computed for the hit-list and this is used to score and rank the screening results (Seidel *et al.,* 2010).

The hit-list can then be analyzed to validate and refine the pharmacophore model. If a pharmacophore model selects *n* molecules from a database with *N* entries, then the selected hit list consists of active compounds (true positive compounds) and decoys (false positive compounds). Active molecules that are not retrieved by the model are defined as false negative compounds while the true negative compounds are the database decoys that were not selected (Kirchmair *et al.,* 2008).
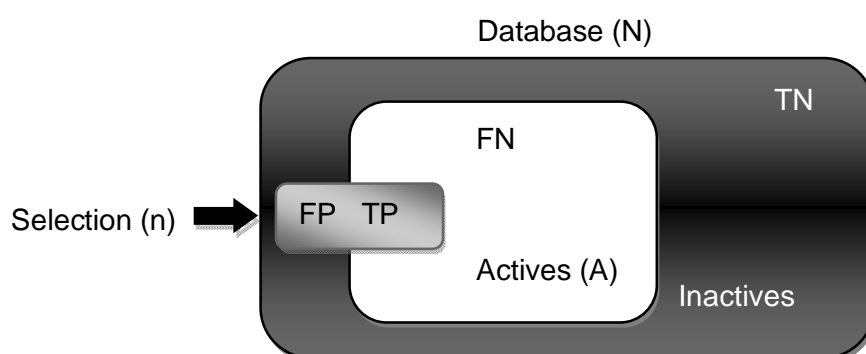

Figure 3.1.2 Selection of *n* molecules from a database containing *N* entries (Adapted from Kirchmair *et al.*, 2008).

According to Kirchmair *et al.* (2008) and Seidel *et al.* (2010), there are several useful calculation measures which can be used including sensitivity (Se), which is the ratio of the retrieved true positive compounds (TP) to all active compounds (A) in the database (the sum

of the TP and false negative (FN) compounds). The values can range between 0 and 1 where Se = 0 means none of the active compounds were found and Se = 1 means all the active compounds were found.

$$Se = \frac{N \; selected \; actives}{N \; total \; actives} = \frac{TP}{TP + FN}$$

Specificity (Sp) is the amount of rejected truly negative compounds (TN) divided by the amount of TN compounds plus the amount of false positive (FP) compounds. This value also ranges from 0-1 and it indicates the percentage of truly inactive compounds that were identified where a value of 0 means none of the negative compounds were identified and a value of 1 means that all of the compounds were identified.

$$Sp = \frac{N \; discarded \; inactives}{N \; total \; inactives} = \frac{TN}{TN + FP}$$

The yield of actives (Ya) quantifies the probability that one of the *n* selected compounds is active. It represents the hit rate that would be achieved if all the molecules selected by the pharmacophore model were tested for activity. However it contains no information about the composition of the database.

$$Ya = \frac{TP}{n}$$

The accuracy (Acc) describes the percentage of molecules that are correctly identified by the model.

$$Acc = \frac{TP + TN}{N} = \frac{A}{N}.Se + \left(1 - \frac{A}{N}\right).Sp$$

The balanced labelling performance (l_{bal}) is a weighted accuracy descriptor. If all the active and inactive molecules are correctly identified by the model, its value should be 1.

$$l_{bal} = \frac{1}{2}.Se + \frac{1}{2}.Sp$$

A more modern method for the assessment of screening results is the Receiver Operating Characteristics (ROC) curves which show the increase of false positives as true positives increases - in other words it describes Se for any possible change of *n* as a function of (1-Sp). The Y-coordinate denotes the true positive value whereas the X-coordinate denotes the false positive rate. If all the molecules scored by a pharmacophore model screening protocol with sufficient discriminatory power are ranked according to score, from high to low, most of the active molecules will have a higher score than the decoys. Since some of the actives will

have a lower score than some of the decoys, an overlap between the distribution of active molecules will occur which will lead to the prediction of false positives and negatives.

An ideal ROC curve representing distributions where no overlap between the scores of active molecules and decoys exist, would rise vertically along the Y-axis until the X-axis (Se) reaches 1 i.e. all the actives are retrieved where after it would continue horizontally to the right until all the decoys are retrieved, which corresponds to Se = 1 and Sp = 0. The ROC curve for a set of actives and decoys with randomly distributed scores tends towards the Se = 1-Sp line asymptotically with an increasing number of actives and decoys. Normally with screening protocols, the ROC curve lies between the ideal curve and the random graph because there is normally an overlapping in the distribution of active molecules and decoys. The ROC curve of a random database is represented by the median. After the hit-list is analyzed the pharmacophore model can be refined to deliver better results (Kirchmair *et al.,* 2008; Seidel *et al.,* 2010).
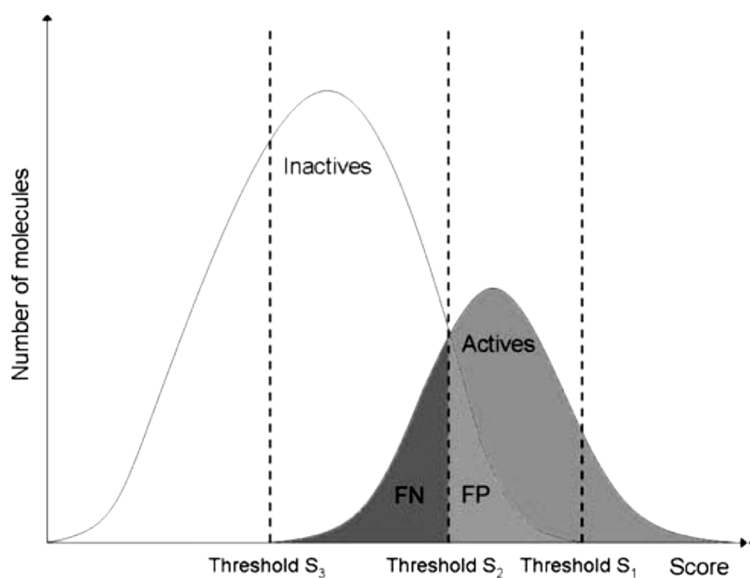
Figure 3.1.3 Theoretical distributions for active molecules and decoys according to their score. Due to an overlap in distribution, different ratios of FP and FN are retrieved depending on the selection threshold S (Kirchmair *et al.*, 2008).
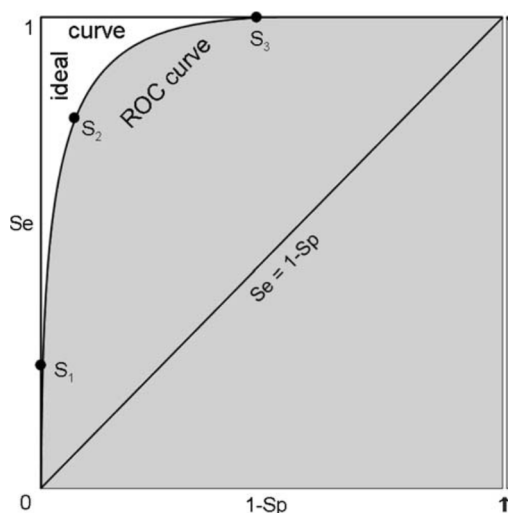
Figure 3.1.4 The ROC curves for ideal and overlapping distributions of actives and decoys. The three ROC curve points S1, S2 and S3 are representative of the corresponding thresholds illustrated in figure 3.1.3. A random distribution causes a ROC curve that tends toward the Se = 1-Sp line asymptotically with increasing numbers of actives and decoys (Kirchmair *et al.*, 2008).

In this study pharmacophore models will be created to screen virtual libraries of the FDA's approved drugs (DrugBank) and the US Environmental Protection Agency's (EPA) maximum daily dose drug databases for compounds that may inhibit MAO-A and MAO-B. The virtual library of FDA approved drugs will be obtained from the DrugBank and the screening of this library as well as the maximum daily dose drug database will be carried out with the molecular modelling software, Discovery Studio® 3.1. The steps that will be followed may be summarized as follows:

- Using X-ray crystal structures of MAO-A and MAO-B with ligands co-crystallized in the active site, *structure-based* pharmacophore models will be generated.
- The interactions of the co-crystallized ligands with MAO will be analysed to gain insight into interactions that are important for inhibitor binding.
- The abilities of the models to identify known MAO inhibitors will be determined using a test set of inhibitors known to inhibit MAO-A and/or MAO-B.
- Different conformations of the drugs within the virtual libraries of FDA approved drugs and EPA's maximum daily dose databases will be generated.
- The conformers will be mapped to the pharmacophore models and the hits identified.
- Selected hits will be evaluated *in vitro* as inhibitors of MAO*.*
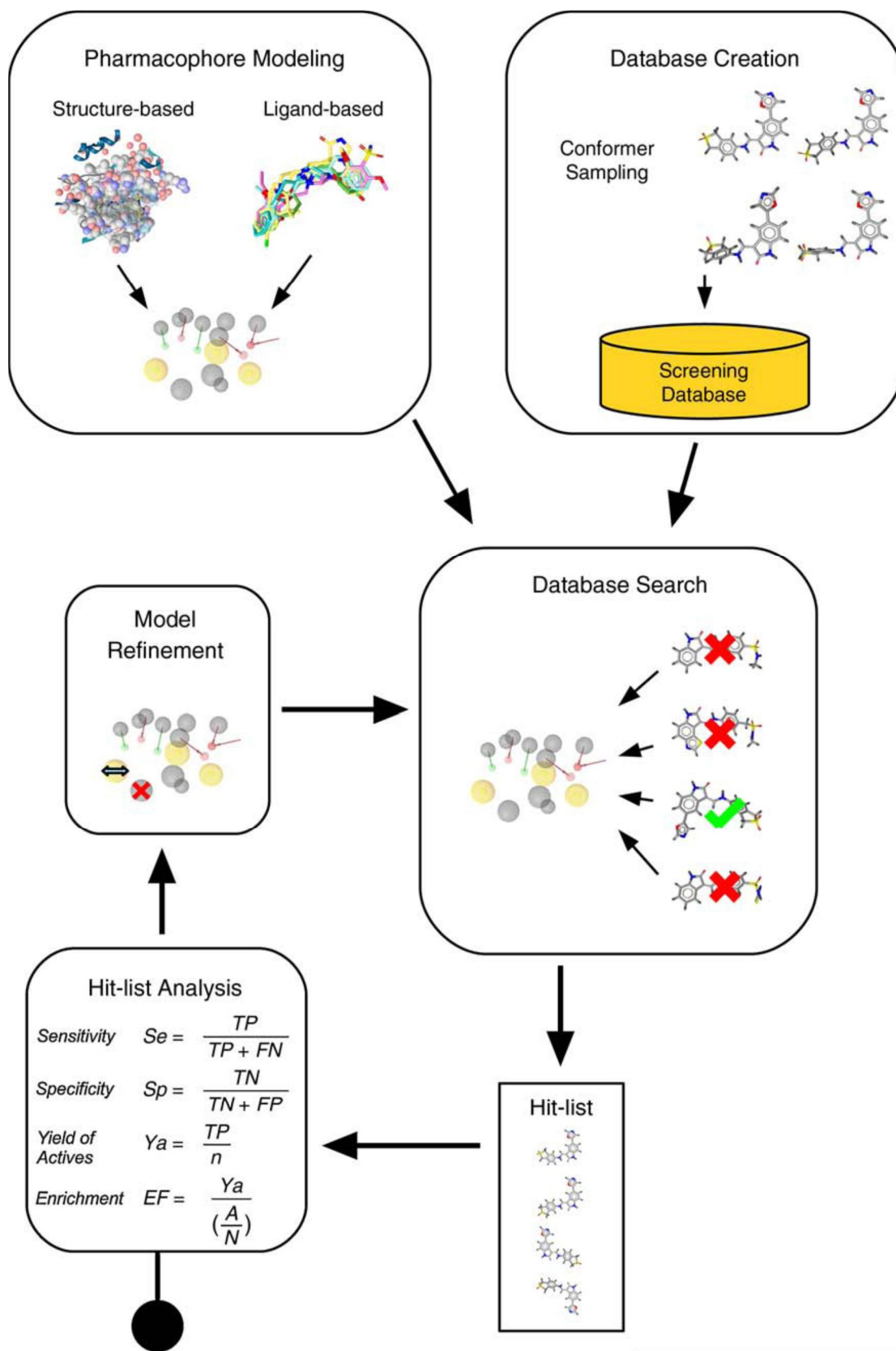
Figure 3.1.5 The 3D pharmacophore-based screening workflow (Seidel *et al.*, 2010)

3.2   Experimental methods:

3.2.1   Construction and screening of the pharmacophore models:

Structure-based pharmacophore generation is dependent on the availability of the target protein structures. In this study *structure-based* models will be generated using:

- The X-ray crystal structure of human MAO-A with harmine co-crystallized in the active site (PDB code: 2ZX5)
- The X-ray crystal structure of human MAO-B with safinamide co-crystallized in the active site (PDB code: 2V5Z).

The following steps were followed:

(a) Protein Preparation:

- The crystallographic structures of the MAO enzymes (as given above), were retrieved from the Brookhaven Protein Data Bank (www.rcsb.org/pdb).
- The correctness of the valences of the FAD cofactor and the co-crystallized ligands were verified and the protein models were automatically typed with the Momany and Rone CHARMm force field.
- Within the Discovery Studio® environment, the pH was set to 7.4 and hydrogen atoms were added to the FAD cofactor, and the co-crystallized ligands.
- The pKa values and protonation states of the ionisable amino acids were calculated and hydrogen atoms were added at pH 7.4 to the protein models, including the waters.
- A fixed atom constraint was applied to the backbone of the enzymes and the models were energy minimized using the Smart Minimizer algorithm with the maximum steps set to 50 000. For this procedure the implicit generalized Born solvation model with molecular volume was used.

(b) Pharmacophore construction:

- All crystal waters were deleted from the protein, except those that undergo hydrogen bonding with the co-crystallized ligands. The binding sites of the MAO enzymes were defined based on the location of the co-crystallized ligands using a sphere with a radius of 9 Å around the active site.

- To determine important interactions between the co-crystallized ligand and amino acid residues, an interaction map between the ligands and amino acid residues was calculated.

- Based on the interaction analysis, pharmacophore features were added to the model. The hydrogen bond acceptor, hydrogen bond donor and hydrophobic features were clustered in turn. Location constraints were added to the features.
- An exclusion constraint was applied within 2 Å of the active site.

(c) Pharmacophore validation and library screening:

- After the pharmacophore models had been constructed a test set was compiled. The test set consists of ligands which are known to inhibit MAO-A and ligands known to inhibit MAO-B. Conformations of the test set were generated (250 of each ligand) by using the BEST conformation method.
- The generated conformations were then queried by the pharmacophore model to determine which conformations fit the features best.
- The structures of the FDA approved drugs (DrugBank) and the EPA's maximum daily dose drug databases were subsequently screened with the pharmacophore models.

Retrieve the co-crystallized structure of MAO-A and MAO-B.

Set pH to 7.4 and add hydrogen atoms to the FAD and the ligand

Calculate the protein ionization and residue pKa at pH 7.4, and add hydrogen atoms to the amino acid residues

Generate an interaction map, add the pharmacophore features and location constraints.

Generate conformations of the test set and the drug libraries.

Query the generated conformations with the pharmacophore model.

Select hits for *in vitro* evaluation.

Figure 3.2.1 Workflow for the construction of a *structure-based* pharmacophore model and screening of a virtual library.

3.2.2 **Molecular docking:**

Molecular docking was carried out with Discovery studio® 3.1. As already mentioned the crystallographic structures of the MAO enzymes were retrieved from the Brookhaven Protein Data Bank. The following structures were used for these studies:

- o Human MAO-A co-crystallized with harmine (pdb. file 2ZX5)
- o Human MAO-B co-crystallized with safinamide (pdb file 2V5Z).

(a) Protein Preparation:

- The protonation states of the ionisable amino acid residues were calculated at pH 7.4 and hydrogen atoms were added to the receptor model.
- The valences of the FAD cofactor and co-crystallized ligands were corrected and hydrogen atoms were added according to the appropriate protonation states at pH 7.4. The structures were typed automatically with the Momany and Rone CHARMm force field.
- A fixed atom constraint was applied to the backbone of the enzymes and the models were energy minimized using the Smart Minimizer algorithm with the maximum steps set to 50 000. For this procedure the implicit generalized Born solvation model with molecular volume was used.
- From the X-ray crystallographic structure of MAO-B, it can be seen that 3 active site water molecules are conserved. Therefore all the crystal waters except these 3 waters in MAO-A and MAO-B were removed.

(b) Docking:

- The co-crystallized ligands and the backbone constraints were subsequently removed from the models and the binding sites were identified by a floodfilling algorithm.
- Structures of the ligands to be docked were constructed within Discovery Studio 3.1®, and their hydrogen atoms were added according to the appropriate protonation states at pH 7.4. The geometries of the ligands were briefly optimized using a fast Dreiding-like force field (1000 interactions) and the atom potential types and partial charges were assigned with the Momany and Rone CHARMm force field.
- Docking of the ligands was carried out with the CDOCKER algorithm with the generation of 10 random ligand conformations and a heating target temperature of 700 K in full potential mode.

- The docking solutions were refined using the Smart Minimizer algorithm. Ten possible binding solutions were computed for each docked ligand and the best-ranked binding conformation of each ligand was determined according to the DockScore values.

Retrieve the crystal structures of MAO-A and MAO-B.

Calculate the protonation states of the ionisable amino acid residues at pH 7.4.

Add hydrogen atoms to the receptor model.

Apply a fixed atom constraint to the enzyme backbone.

Do an energy minimization.

Remove all the crystal waters except the 3 conserved waters.

Remove the co-crystallized ligand and the backbone constraint. Identify the binding site.

Construct the ligands to be docked and dock them with the CDOCKER algorithm.

Refine orientations with the Smart Minimizer algorithm.

Figure 3.2.2 Workflow for docking ligands into the active site of the MAO enzymes.

3.3   Results:

3.3.1   *Structure-based* pharmacophore of MAO-A:
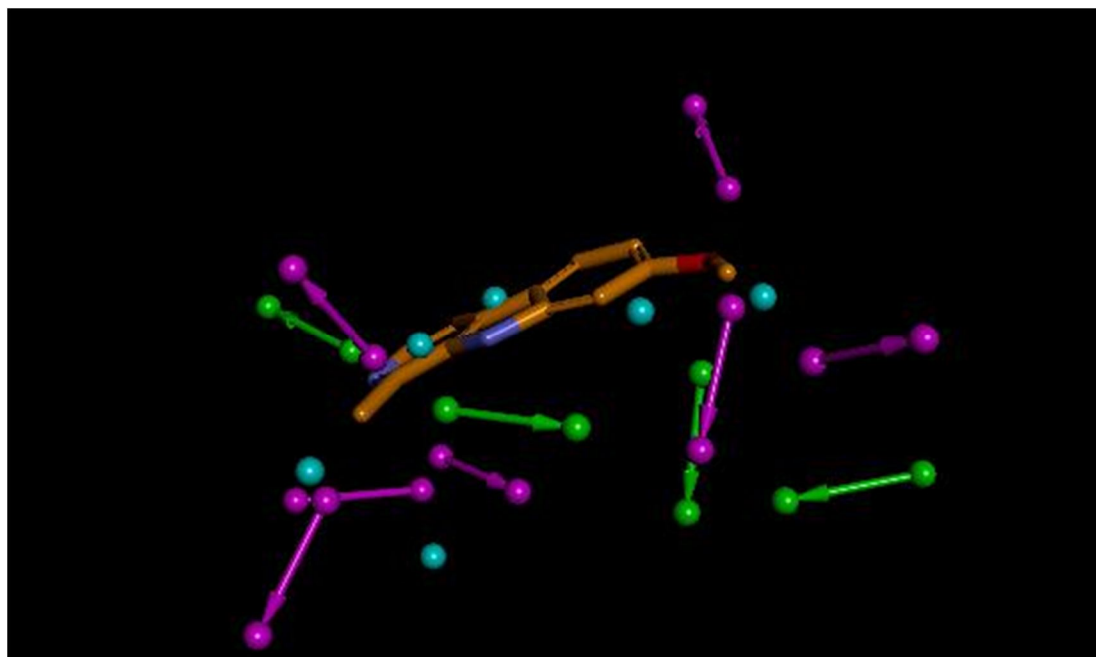
Figure 3.3.1 Graphical representation of the *structure-based* pharmacophore model of MAO-A, which was constructed using the structure of the co-crystallized ligand, harmine. The model may be used to screen virtual libraries for ligands that bind to MAO-A. The green arrows represent hydrogen bond acceptor features, the purple arrows represent hydrogen bond donor features and the cyan spheres represent hydrophobic features.

For the construction of the pharmacophore model in figure 3.3.1, the X-ray crystal structure of human MAO-A co-crystallized with harmine in the active site was used (PDB code 2ZX5). All the calculations were carried out with Discovery Studio® 3.1. The software calculates the interactions between the harmine and the amino acid residues of the active site. The software also determines any additional interactions that may exist between a ligand and the active site of MAO-A. Based on these possible interactions, pharmacophore features are placed in the active site. These features are hydrogen bond acceptor features, hydrogen bond donor features and hydrophobic features. The user then clusters the features into groups. With this step a group of features that represent the same interaction is combined into a single feature. Location constraints are subsequently added to each feature. These are spheres which are placed around the center of the features and they define the ideal location for the ligand atoms. The sphere represents the tolerance of the allowable deviation of the ligand atoms from the ideal position. In the last step, exclusion constraint features is placed around the model. While the software screens a virtual database for ligands that may map to the pharmacophore model, the algorithm attempts to find ligands that do not bind in the excluded zones. The binding of ligands is thus confined to the active site only. Since

65

harmine is relatively large and fills the active site cavity of MAO, the exclusion feature is representative of the amino acid residues surrounding the active site.
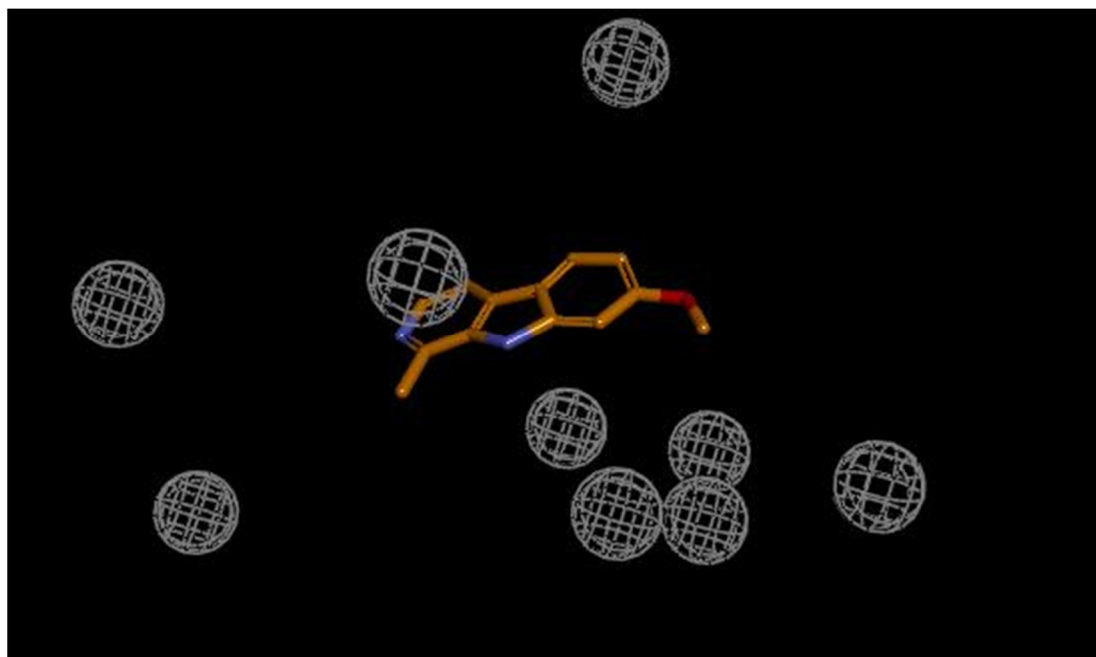
Figure 3.3.2 Graphical representation of the MAO-A pharmacophore model derived from the structure of harmine using the *structure-based* approach. In this representation, only the exclusion constraint features are illustrated.
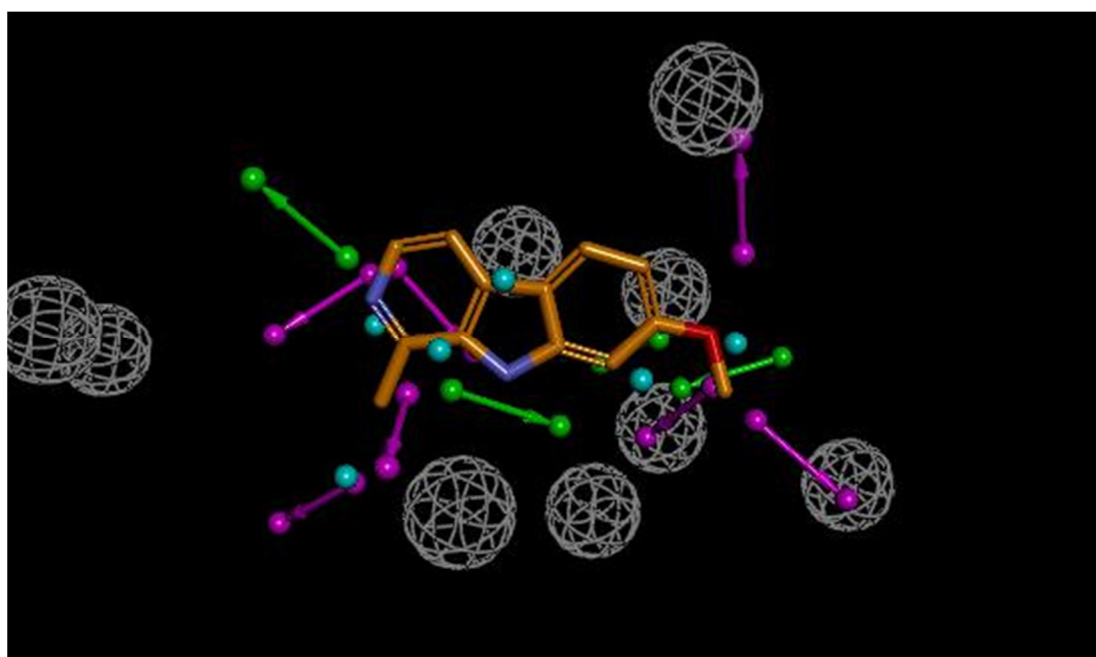
Figure 3.3.3 Graphical representation of the MAO-A pharmacophore model derived from the structure of harmine using the *structure-based* approach. In this representation, the exclusion constraint features as well as the hydrogen bond acceptor, hydrogen bond donor and hydrophobic features are shown.

To gain insight into the pharmacophore features of the model shown in Fig 3.3.1, the interactions between the co-crystallized ligand (harmine) and the active site of MAO-A were analyzed. This analysis was done by displaying the interactions in two-dimensions as well as calculating the interaction energies between the ligand and individual residues of the active site.

As shown in the two-dimensional representation of these interactions, there are no pi-pi interactions between harmine and the residues in the active site of MAO. The two-dimensional representation also shows that hydrophobic interactions exist between the ligand and Ile-180, Phe-208, Gln-215 and Ile-335 (as indicated by the cyan shading of the spheres). The interaction energies show that these amino acid residues contribute significantly to the total binding energy of the ligand (-2.876, -3.389, -4.724 and -2.527 kcal/mol, respectively). Based on the more negative energy, the interaction with Gln-215 is especially important. There is a strong interaction between harmine and the FAD cofactor (-2.521 kcal/mol) which illustrates the importance of its role in substrate and inhibitor binding. The two-dimensional representation also indicates that a hydrogen bond exists between the amine hydrogen of harmine and the oxygen of an active site water (HOH-746), where harmine acts as a hydrogen bond donor.
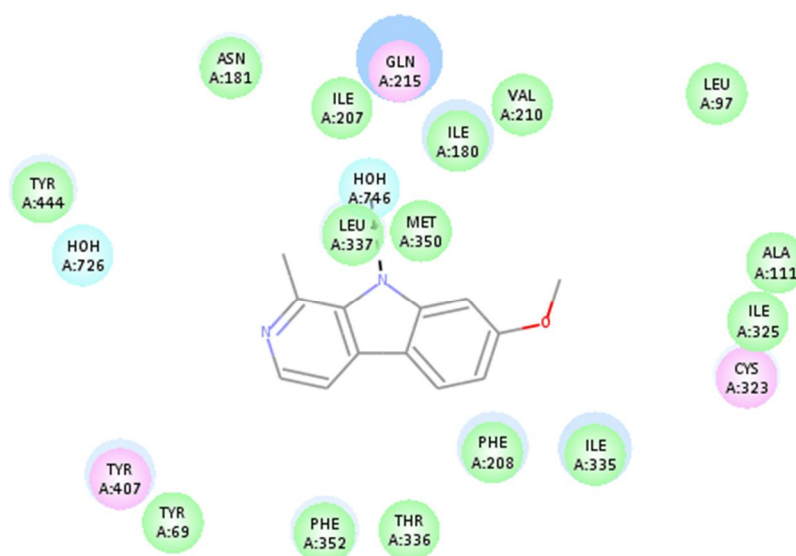
Figure 3.3.4 A two-dimensional representation of the binding of harmine in the MAO-A active site.

Table 3.3.1 The interaction energies between harmine and the active site residues and waters of MAO-A. Selected interactions, those that are the most productive, are shaded:

Name	Forcefield	Total Interaction Energy (kcal/mol)	Total VDW Interaction Energy (kcal/mol)	Total Electrostatic Interaction Energy (kcal/mol)
2Z5X	CHARMm	-27.63292	-25.67569	-1.95723

Interaction Energies			
Residue	Interaction Energy (kcal/mol)	VDW Interaction Energy (kcal/mol)	Electrostatic Interaction Energy (kcal/mol)
A_ALA68	-0.167062	-0.124832	-0.042230
A_TYR69	-1.319910	-1.465940	0.146028
A_LEU97	-0.223802	-0.239213	0.015411
A_PHE108	-0.076243	-0.070406	-0.005837
A_ALA111	-0.429997	-0.216156	-0.213841
A_ILE180	-2.876600	-2.836630	-0.039970
A_ASN181	-1.170640	-1.174680	0.004037
A_TYR197	-0.068383	-0.094819	0.026436
A_ILE207	-0.477119	-0.634126	0.157007
A_PHE208	-3.389060	-2.990900	-0.398163
A_SER209	-0.365134	-0.364331	-0.000803
A_VAL210	-0.556786	-0.585740	0.028954
A_GLN215	-4.724010	-4.767820	0.043810
A_CYS323	-0.476047	-0.490002	0.013955
A_ILE325	-1.162810	-1.127090	-0.035718
A_ILE335	-2.527040	-2.516640	-0.010401
A_THR336	-0.789534	-0.796514	0.006980
A_LEU337	-2.243600	-2.277100	0.033497
A_MET350	-1.145140	-1.117930	-0.027214
A_PHE352	-1.013900	-1.101110	0.087211
A_TYR407	-2.125410	-2.323700	0.198287
A_TYR444	-0.451864	-0.487063	0.035199
A_FAD600	-2.521630	-1.890300	-0.631330
A_HOH706	-0.199003	-0.097508	-0.101495
A_HOH710	-0.039960	-0.174450	0.134490
A_HOH718	-0.500084	-0.499793	-0.000291
A_HOH725	-0.130260	-0.343501	0.213241
A_HOH726	-0.177478	-0.688920	0.511442
A_HOH729	-0.068252	-0.066552	-0.001700
A_HOH739	-0.415425	-0.466452	0.051027
A_HOH746	-0.380772	0.411755	-0.792527
A_HOH766	-0.987837	-0.916089	-0.071748
A_HOH805	-2.040050	-1.087720	-0.952329
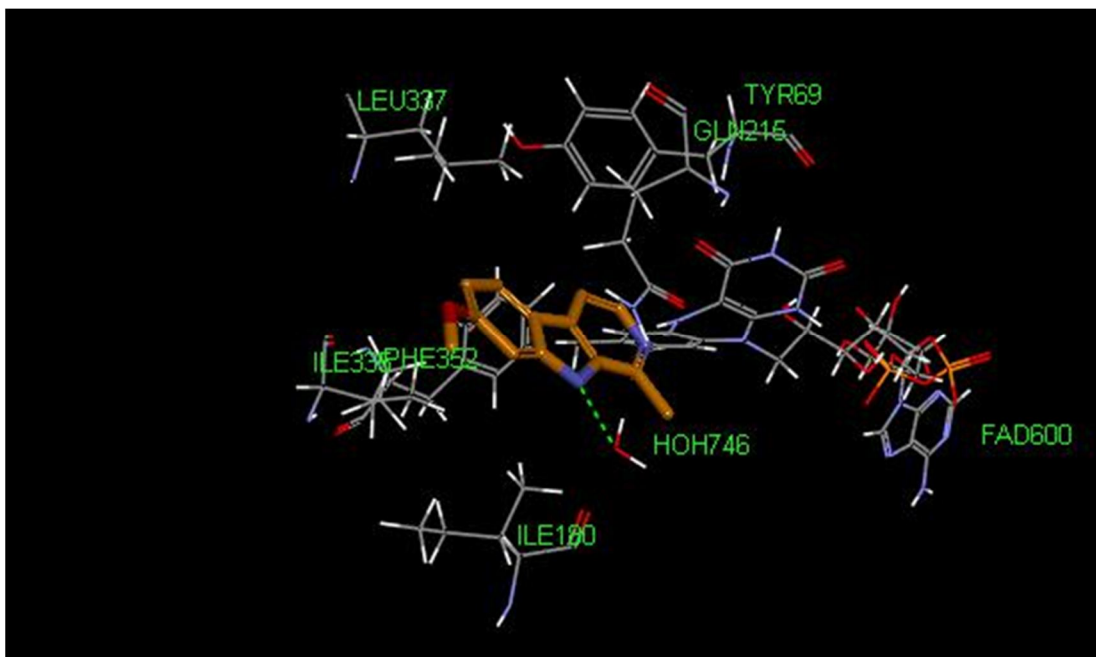
Figure 3.3.5 A three-dimensional representation of the interaction between harmine and the selected residues in the active site of MAO-A.

As shown in figure 3.3.6, the four acceptor features of the pharmacophore model correspond to interactions with the following residues:
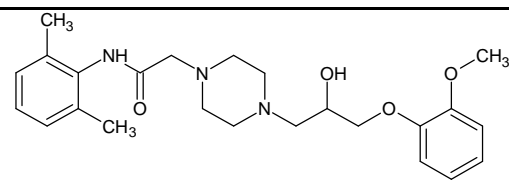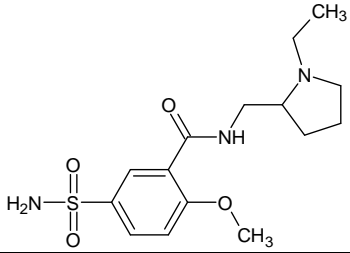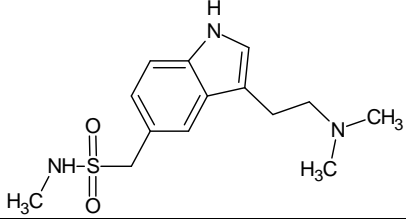
- Phe-208
- Ser-209
- Gln-215
- FAD
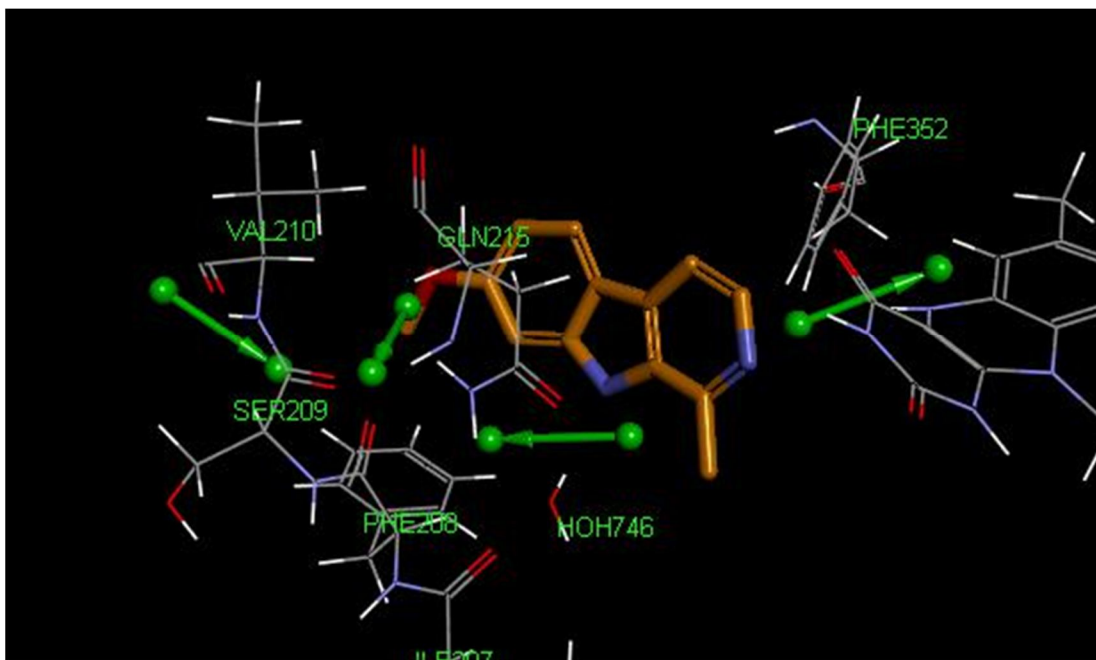
Figure 3.3.6 A three-dimensional representation of acceptor features and their corresponding residues.

As shown in figure 3.3.7, the seven donor features correspond to interactions with the following residues:

- Ala-111 (peptide carbonyl)
- Ile-180 (peptide carbonyl)
- Phe-208 (peptide carbonyl)
- Gln-215 (side chain carbonyl)
- Thr-336 (peptide carbonyl)
- Tyr-444 (phenolic oxygen)
- The active site water HOH-726
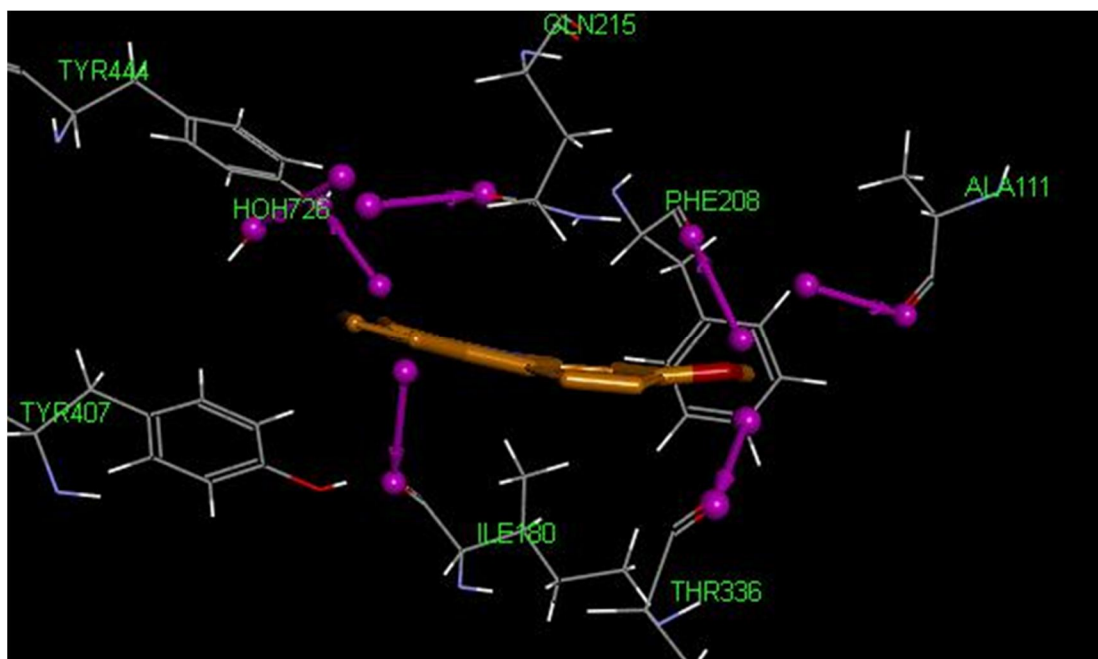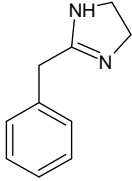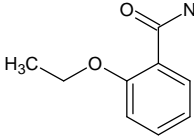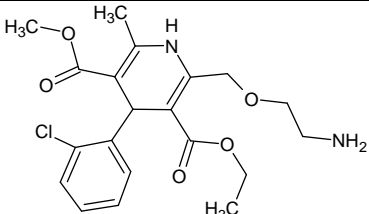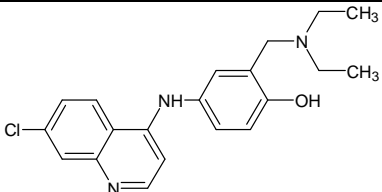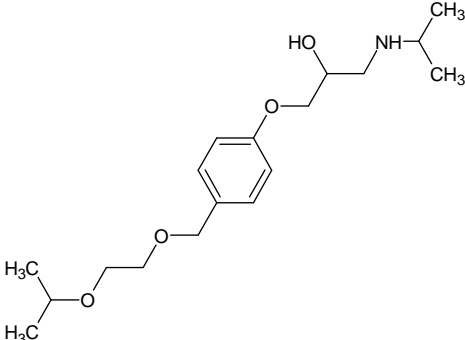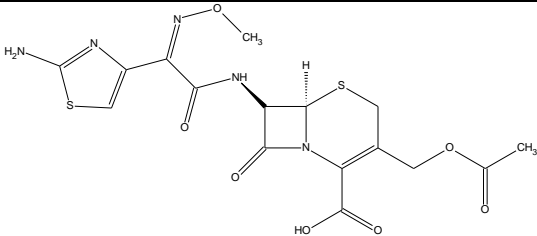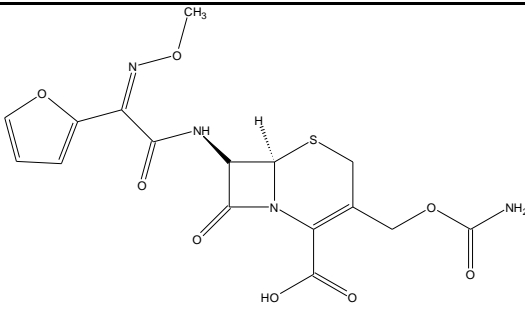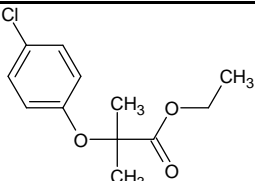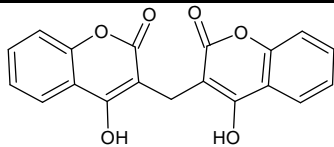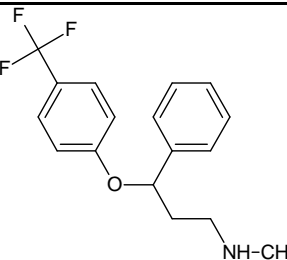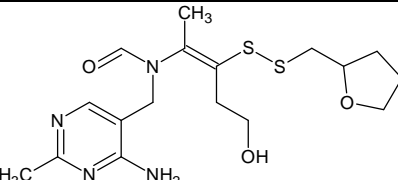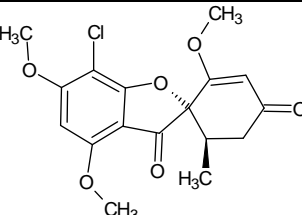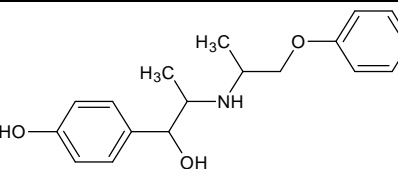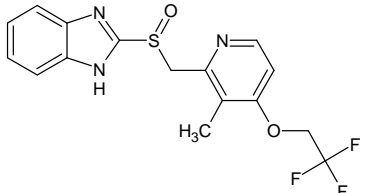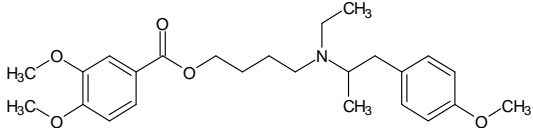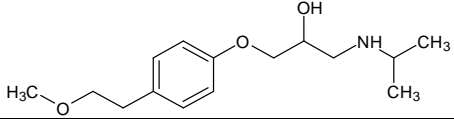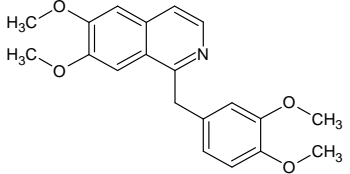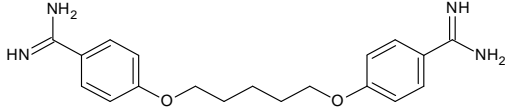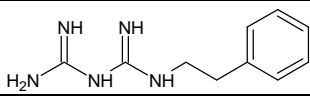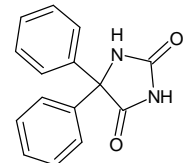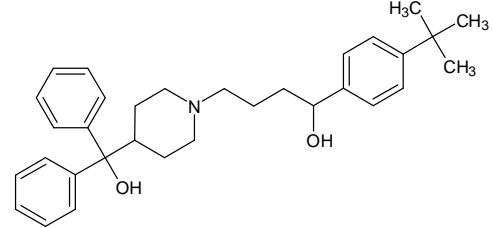
Figure 3.3.7 A three-dimensional representation of donor features and their corresponding residues.
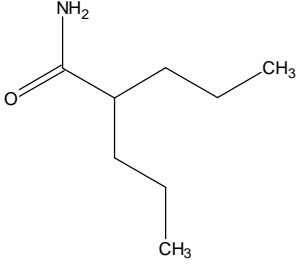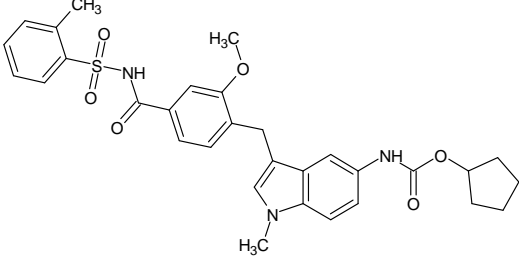
The pharmacophore model was subsequently used to screen a virtual drug library for molecules with the potential to bind to the active site of MAO-A. For this purpose, both the DrugBank library, which contains all of the United States FDA approved drug molecules, and the United States EPA's maximum daily dose database were used. First a set of conformations was calculated for each molecule in the libraries. For each compound, a maximum of 255 conformations was generated using the BEST algorithm of Discovery Studio®. Those conformations were then mapped to the pharmacophore model using the Screen Library protocol of Discovery Studio 3.1®. None of the features were set as required features and the fitting method was set to rigid. Both 4- and 5-feature hits were considered for *in vitro* evaluation as inhibitors of human MAO-A as shown in table 3.3.2.

Table 3.3.2 A list of the compounds in the DrugBank which mapped to the pharmacophore model derived from the structure of MAO-A (with harmine co-crystallized) and from the structure of MAO-B (with safinamide co-crystallized) using the *structure-based* approach. Also given are the fit-values of the respective compounds. These compounds represent drugs which are used systemically by humans. It is important to note that these are only the hits that were selected for further *in vitro* evaluation as MAO inhibitors.

Name	Structure	Fit-Value for MAO-A	Fit-Value for MAO-B
2-Benzyl-2-imidazoline (tolazoline)		3.74	3.81
2-Ethoxybenzamide		3.56	2.70
Amlodipine		3.71	3.85
Amodiaquine		3.90	3.80
Bisoprolol		3.89	3.68
Cefotaxime		3.63	3.67
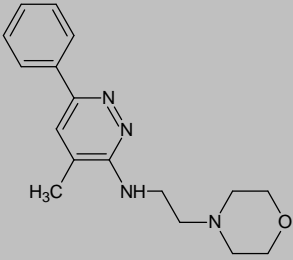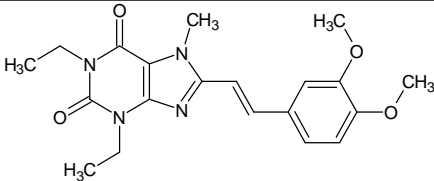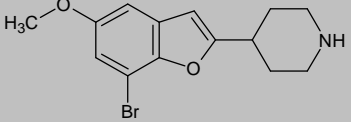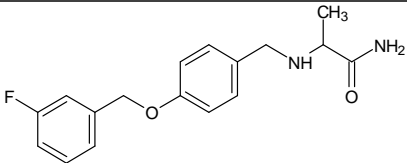
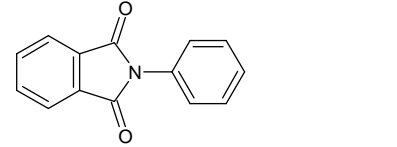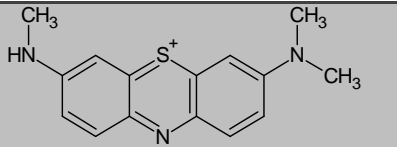| Cefuroxime		3.51	3.41
Clofibrate		3.92	3.70
Dicumarol		3.45	3.25
Fluoxetine		3.58	3.22
Fursultiamine		3.59	3.63
Griseofulvin		3.89	3.94
Isoxsuprine		3.72	3.74
Lansoprazole		3.83	3.80
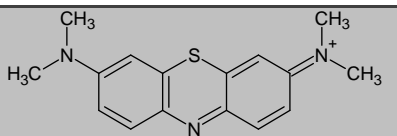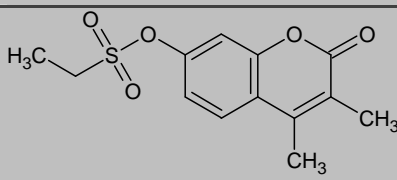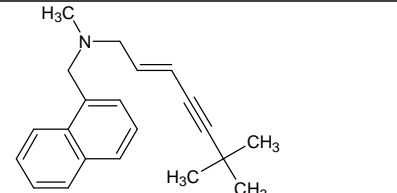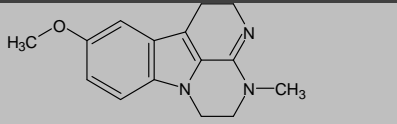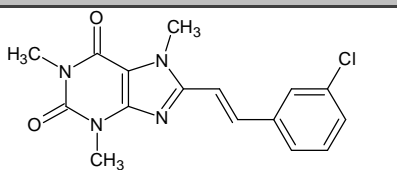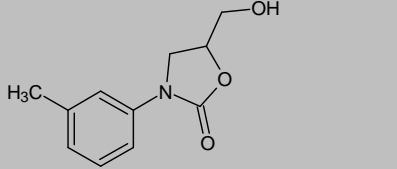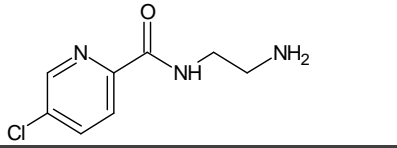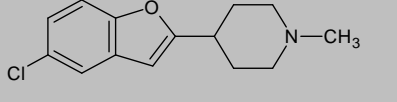
Mebeverine		3.83	3.83
Metoprolol		3.54	3.52
Papaverine		3.96	3.78
Pentamidine		3.92	3.82
Phenformin		3.56	3.13
Phenytoin		3.54	2.59
Ranolazine		-	3.78
Sulpiride		3.82	3.80
Sumatriptan		3.59	2.89
Terfenadine		3.77	3.71

Valpromide		3.81	3.86

Zafirlukast		3.78	3.90

To determine if the pharmacophore model also has the ability to identify known MAO-A inhibitors, a series of 18 test compounds were queried with the model.

Table 3.3.3 A virtual library of 18 test compounds (9 inhibitors and 9 non-inhibitors of MAO-A) was screened with the MAO-A pharmacophore model. Below is given a list of compounds that were found to be four feature hits. The pharmacophore model used was derived from the structure of harmine. The compounds not shaded are known not to inhibit MAO-A.

Name:	Structure:	Fit-Value
Minaprine		3.82
Istradefylline		3.77
Brofaromine		3.70
Safinamide		3.58
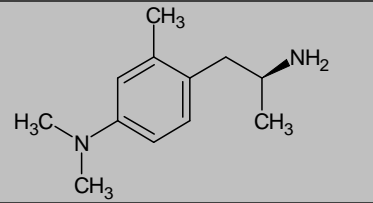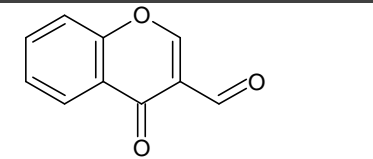
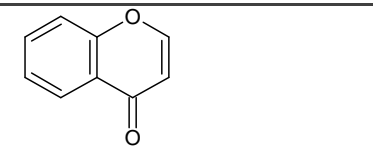|		
Methylene blue		3.54
Esuprone		3.53
Terbinafine		3.46
Metralindole		3.46
8-(3-chlorostyryl)caffeine		3.42
Toloxatone		3.41
Lazabemide		3.27
Serchloremine		3.17
N-phenylphthalimide		2.86
Azure B		2.86
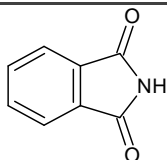
Phthalimide		2.66

Table 3.3.4 A virtual library of 18 test compounds (9 inhibitors and 9 non-inhibitors of MAO-A) was screened with the MAO-A pharmacophore model. Below is given a list of compounds that were found not to be hits. The pharmacophore model used was derived from the structure of harmine. The compounds not shaded are known not to inhibit MAO-A.

Name:	Structure:
Amiflamine	
Chromone-3-carboxaldehyde	
4-Chromone	

Analysis of results:

Sensitivity:

$$Se = \frac{TP}{TP + FN} = \frac{8}{8 + 1} = 0.8889$$

The Se value of 0.8889 indicates that 88.89% of the active compounds were identified by the model. Therefore the model has the ability to identify and retrieve truly active compounds.

Specificity:

$$Sp = \frac{TN}{TN + FP} = \frac{2}{2 + 7} = 0.2222$$

The Sp value of 0.2222 indicates that only 22.22% of decoys were discarded therefore the model also retrieves a high number of decoys in addition to the active molecules.

Yield of actives:

$$Ya = \frac{TP}{n} = \frac{8}{15} = 0.5333$$

The Ya value is calculated as 0.5333. This value shows that 53.33% of the hits were active.

Accuracy:

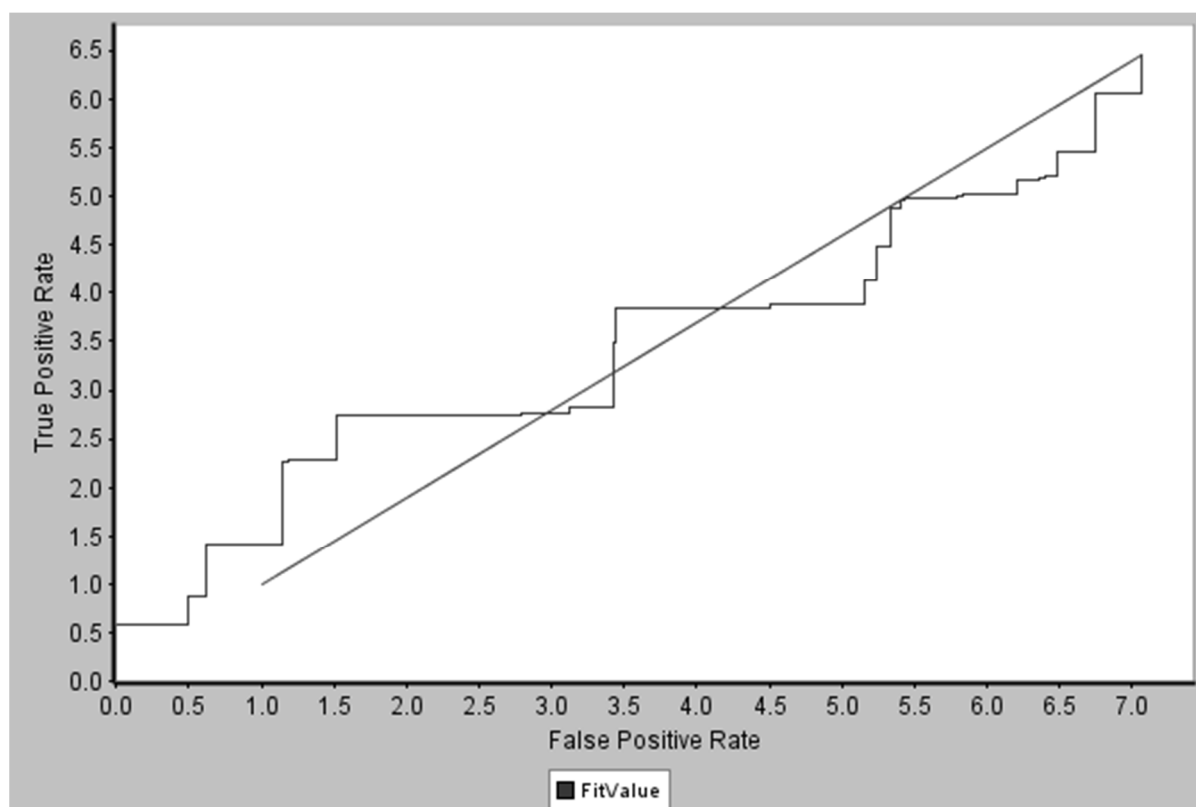$$Acc = \frac{TP + TN}{N} = \frac{8 + 2}{18} = 0.5556$$

This result shows that the model is 55.56% accurate.

Balanced labelling performance:

$$l_{bal} = \frac{1}{2} \cdot Se + \frac{1}{2} \cdot Sp = \frac{0.8888}{2} + \frac{0.2222}{2} = 0.5555$$

This result shows that the model is 55.55% accurate.

ROC curve:

Score Property	Area under ROC Curve	ROC Evaluation
FitValue	1.246	Excellent

Figure 3.3.8 ROC curve for the MAO-A pharmacophore model.

A ROC curve was constructed for the Fit-Values. As shown in figure 3.3.8, the ROC curve lies between the ideal curve and the Se = 1-Sp asymptote. It may thus be concluded that Fit-Values of the active compounds are significantly higher than those of the decoys. The analysis of the curve by Discovery Studio 3.1® indicates that the ROC evaluation is good.

Although the model is not very specific for the exclusion of compounds known not to inhibit MAO-A, it is highly sensitive and its results are more than 50% accurate and can therefore be useful if the compounds retrieved by the model are further screened with *in vitro* testing as MAO-A inhibitors.

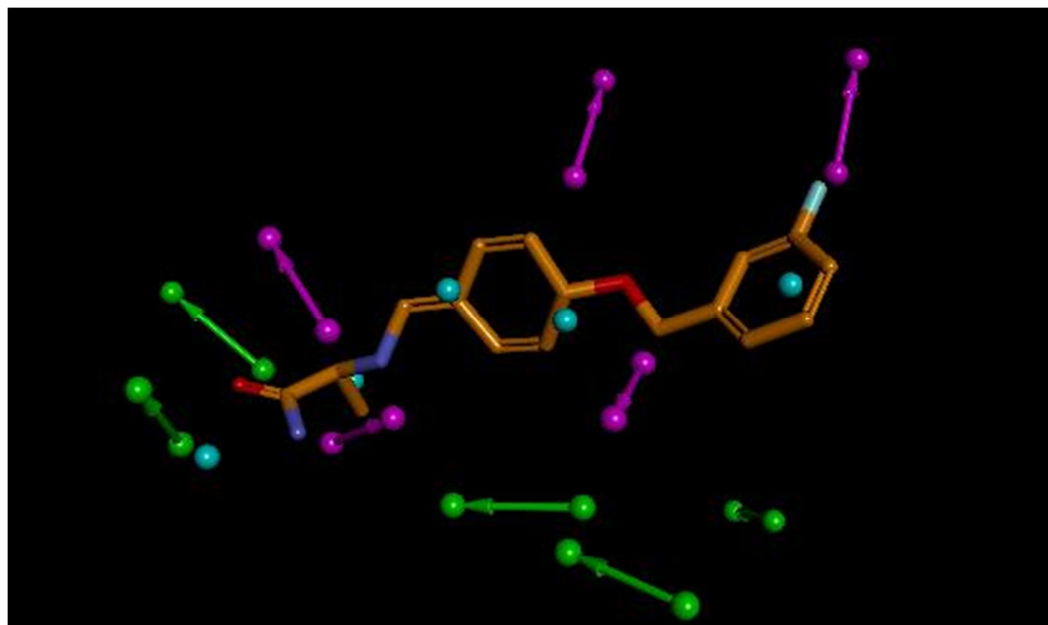### 3.3.2 *Structure-based* pharmacophore model of MAO-B:

Figure 3.3.9 Graphical representation of the *structure-based* pharmacophore model of MAO-B, which was constructed using the structure of the co-crystallized ligand, safinamide.

For the construction of the pharmacophore model in figure 3.3.9, the X-ray crystal structure of human MAO-B co-crystallized with a safinamide in the active site was used (PDB code 2V5Z). The calculations carried out were similar to those described in the previous section for MAO-A.
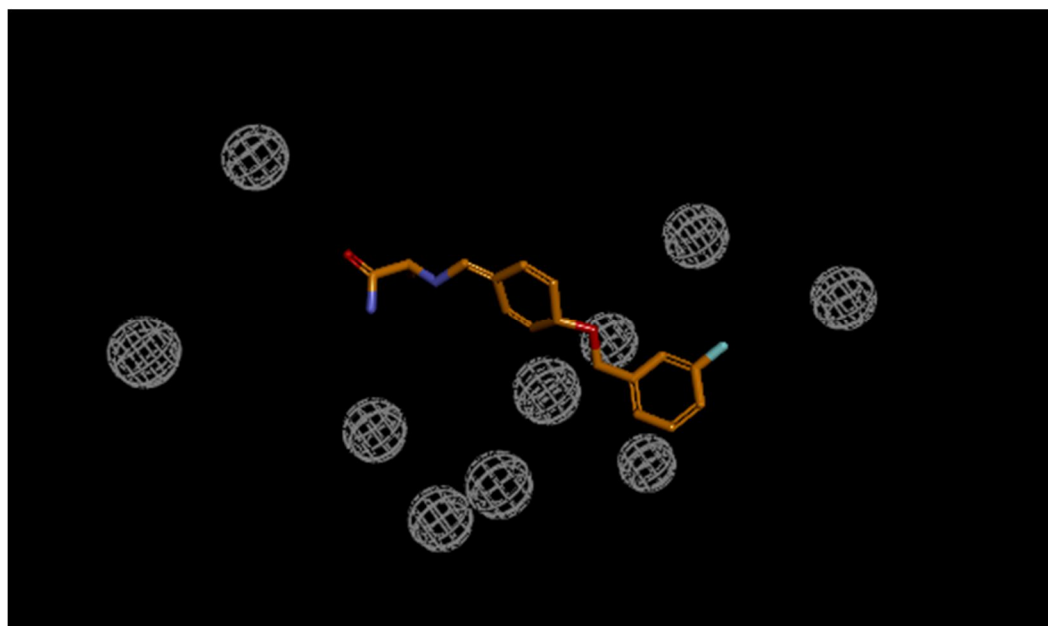
Figure 3.3.10 Graphical representation of the MAO-B pharmacophore model derived from the structure of safinamide using the *structure-based* approach. In this representation, only the exclusion constraint features are illustrated.
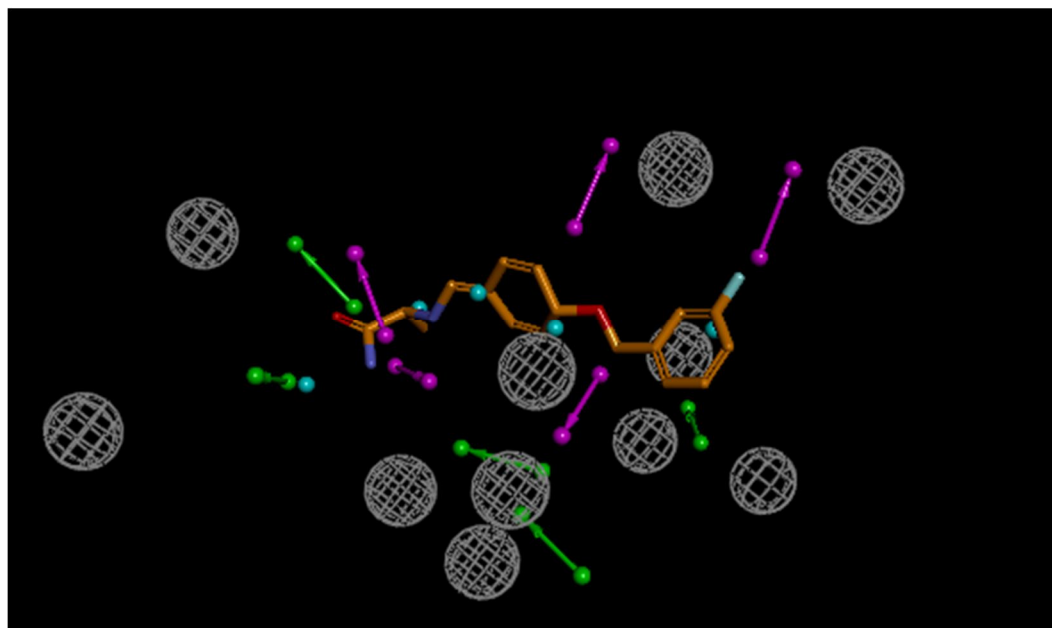
Figure 3.3.11 Graphical representation of the MAO-B pharmacophore model derived from the structure of safinamide using the *structure-based* approach. In this representation, the exclusion constraint features as well as the hydrogen bond acceptor, hydrogen bond donor and hydrophobic features are shown.

The interactions between safinamide and the active site of MAO-B were analyzed using a similar protocol to that used for the analysis of the interactions between harmine and MAO-A.

As shown in the two-dimensional representation of these interactions, there are no pi-pi interactions between safinamide and the residues in the active site of MAO. The two-dimensional representation also shows that hydrophobic interactions exist between the ligand and Leu-171, Ile-199, Gln-206, Ile-316, Tyr-326, Tyr-398 and Tyr-435 (as indicated by the cyan shading of the spheres). The interaction energies show that these amino acid residues contribute significantly to the total binding energy of the ligand (-4.365, -3.367, -5.184, -1.87, -3.041 and -2.971 kcal/mol, respectively). There is also an interaction with the active site water HOH-1351 (-1.796 kcal/mol). Based on the more negative energy, the interaction with Leu-171 and Gln-205 is especially important. The two-dimensional representation also indicates that hydrogen bonds exist between the amide functional group of safinamide and Gln-206. The carbonyl oxygen of the amide also acts as a hydrogen bond acceptor in an interaction with the active site water HOH-1351.
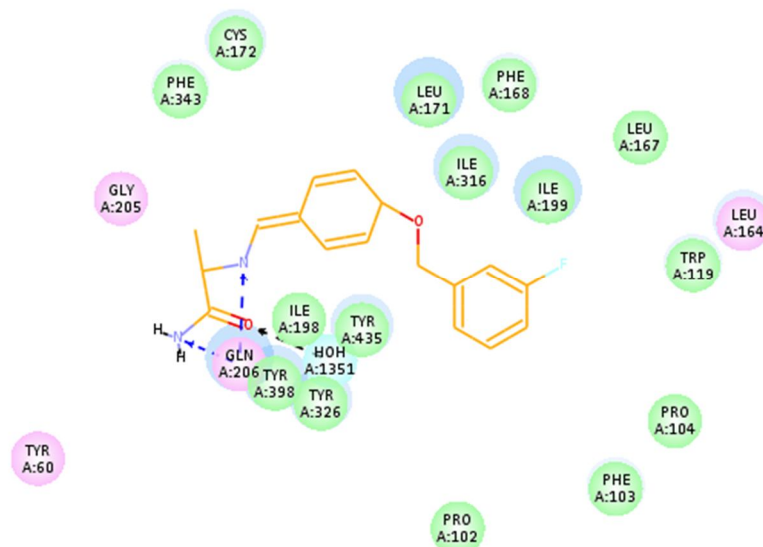
Figure 3.3.12 A two-dimensional representation of the binding of safinamide in the MAO-B active site.

Table 3.3.5 The interaction energies between safinamide and the active site residues and waters of MAO-B. Selected interactions, those that are the most productive, are shaded:

Name	Forcefield	Total Interaction Energy (kcal/mol)	Total VDW Interaction Energy (kcal/mol)	Total Electrostatic Interaction Energy (kcal/mol)
2V5Z	CHARMm	-39.18424	-34.92873	-4.25551

| **Interaction Energies**			
Residue	**Interaction Energy (kcal/mol)**	**VDW Interaction Energy (kcal/mol)**	**Electrostatic Interaction Energy (kcal/mol)**
A_GLY58	-0.136607	-0.086885	-0.049722
A_SER59	-0.222672	-0.169864	-0.052808
A_TYR60	-1.695590	-2.003590	0.308002
A_VAL61	-0.154478	-0.126010	-0.028468
A_GLN65	-0.041408	-0.098069	0.056662
A_PHE99	-0.138023	-0.159521	0.021498
A_PRO102	-0.940973	-0.497440	-0.443533
A_PHE103	-1.378000	-1.337310	-0.040691
A_PRO104	-0.852757	-0.903993	0.051236
A_TRP119	-1.556110	-1.413950	-0.142161
A_LEU164	-1.341830	-1.610150	0.268316
A_LEU167	-0.969116	-0.909259	-0.059857
A_PHE168	-2.461500	-2.232630	-0.228871
A_LEU171	-4.365330	-4.316890	-0.048437
A_CYS172	0.151385	0.259395	-0.108010
A_TYR188	-0.011600	-0.137433	0.125833
A_ILE198	-1.998230	-1.945430	-0.052798
A_ILE199	-3.367640	-2.985300	-0.382343
A_SER200	-0.463414	-0.363566	-0.099848
A_GLY205	-0.837361	-0.673937	-0.163424
A_GLN206	-5.184040	-2.597670	-2.586370
A_ILE316	-1.870310	-1.857520	-0.012795
A_TYR326	-3.041600	-2.566770	-0.474831
A_LEU328	-0.233845	-0.319285	0.085440
A_MET341	-0.082714	-0.093881	0.011168
A_PHE343	-1.222840	-1.279630	0.056791
A_TYR398	-2.971280	-2.919920	-0.051365
A_TYR435	-1.796350	-1.582220	-0.214126
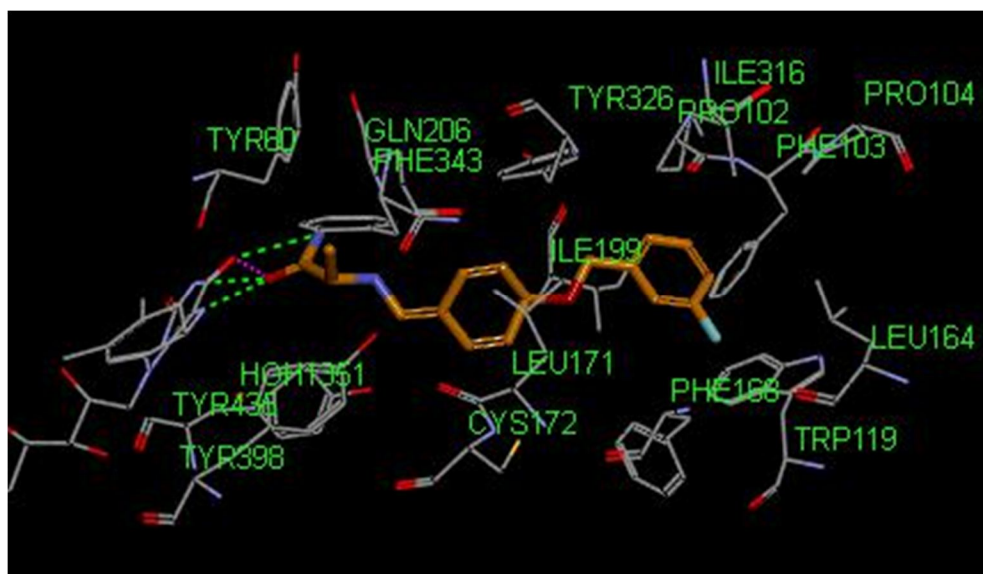
Figure 3.3.13 A three-dimensional representation of the interactions between safinamide and selected residues in the active site of MAO-B.

As shown in figure 3.3.14, the five acceptor features of the pharmacophore model correspond to interactions with the following residues:

- Gln-65
- Ser-200
- Gln-206
- Thr-327
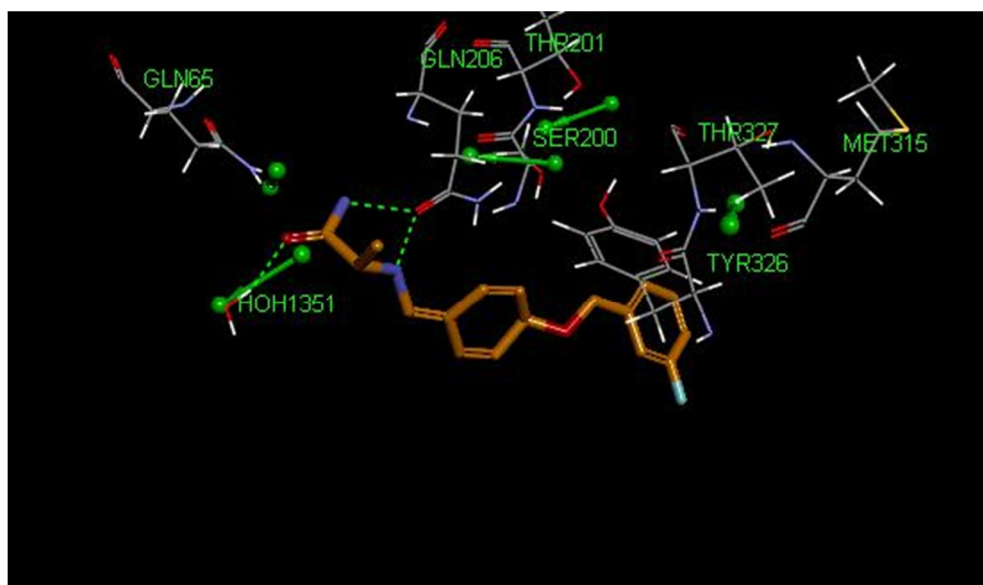- The active site water, HOH-1351

Figure 3.3.14 A three-dimensional representation of acceptor features and their corresponding interacting residues.

As shown in figure 3.3.15, the five donor features corresponds to interactions with the following residues:

- Leu-164 (peptide carbonyl)
- Phe-168 (peptide carbonyl)
- Ile-199 (peptide carbonyl)
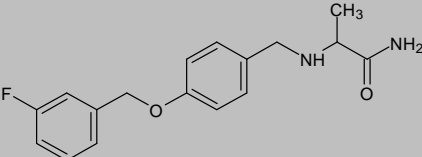- Gln-206 (side chain carbonyl)
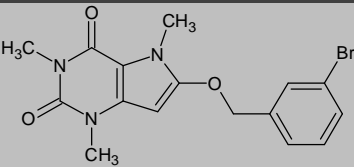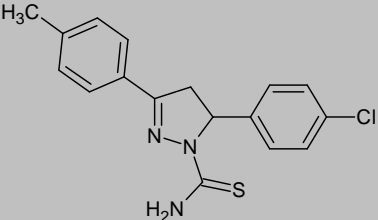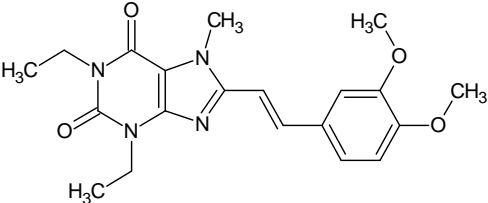- Tyr-435 (phenolic oxygen)
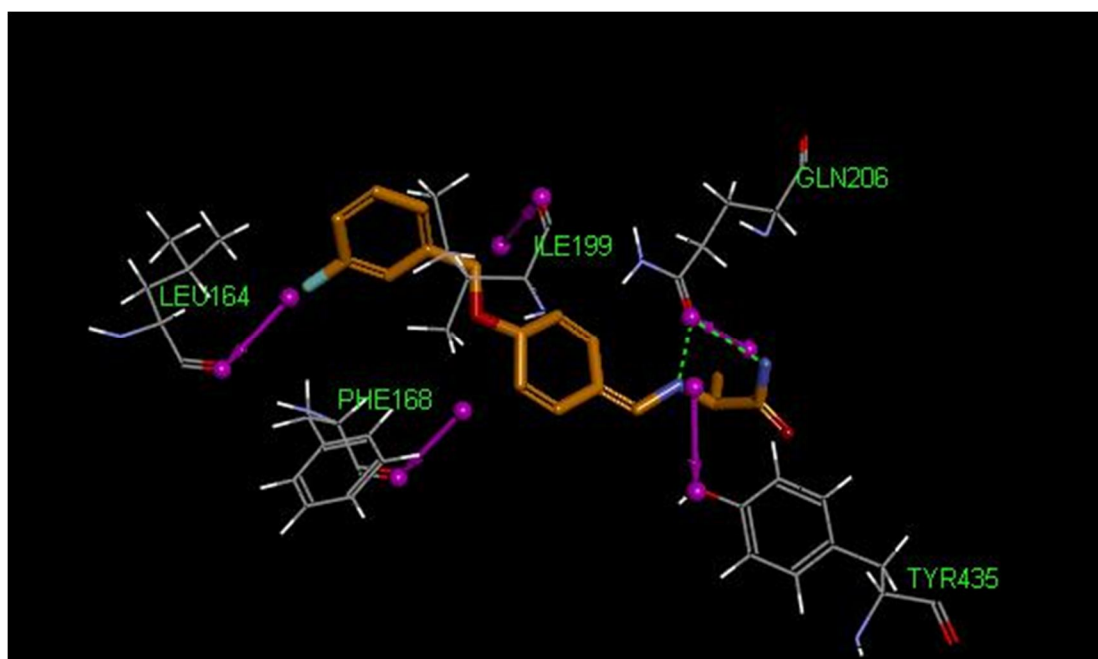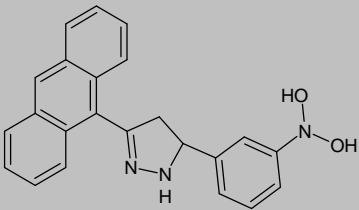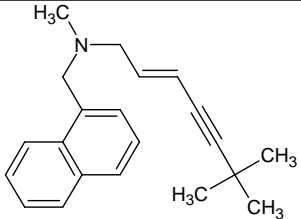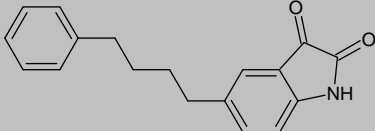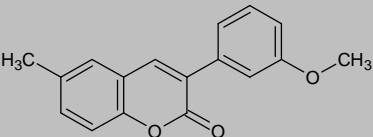
Figure 3.3.15 A three-dimensional representation of donor features and their corresponding interacting residues.

The pharmacophore model was subsequently used to screen a virtual drug library for molecules with the potential to bind to the active site of MAO-B. This was done in a similar manner to the screening of the MAO-A model described earlier. Both 4- and 5-feature hits were considered for *in vitro* evaluation as inhibitors of human MAO-B, as shown in table 3.3.6.

To determine if the pharmacophore model also has the ability to identify known MAO-B inhibitors, a series of 17 test compounds were queried by the model.

Table 3.3.6 A virtual library of 17 test compounds (10 known inhibitors and 7 non-inhibitors of MAO-B) was screened with the MAO-B pharmacophore model. Below is given a list of compounds that were found to be hits. The pharmacophore model used was derived from the structure of safinamide. The compounds not shaded are known not to inhibit MAO-B.

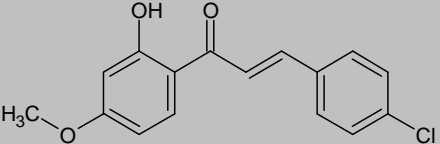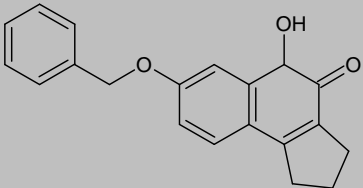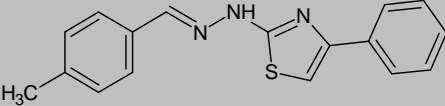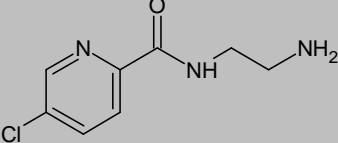| Name	Structure	Fit-Value
Pyrazoline derivative 2		3.71
Terbinafine		3.68
Isatin derivative		3.67
Coumarin derivative 2		3.59
Safinamide		3.33
Caffeine derivative		3.32
Pyrazoline derivative 1		3.30
Istradefylline		3.27
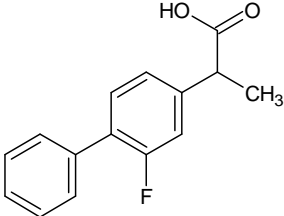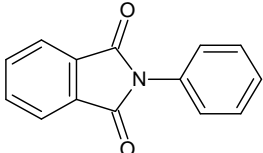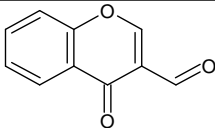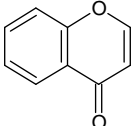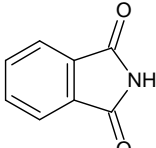
Chalcone derivative		3.13
Coumarine derivative 1		3.04
Hydrazinothiazole derivative		2.80
Lazabemide		2.77
Flurbiprofen		2.29

Table 3.3.7 A virtual library of 17 test compounds (10 known inhibitors and 17 non-inhibitors of MAO-B) was screened with the MAO-B pharmacophore model. Below is given a list of compounds that were found not to be hits. The pharmacophore model used was derived from the structure of safinamide. The compounds not shaded are known not to inhibit MAO-B.

Name:	Structure:
N-phenylphthalimide	
Chromone-3-carboxaldehyde	
4-Chromone	
Phthalimide	

Analysis of results:

Sensitivity:

$$Se = \frac{TP}{TP + FN} = \frac{10}{10 + 0} = 1$$

The Se value of 1 shows that the model has an excellent sensitivity for compounds that have the ability to inhibit MAO-B. The Se value is equal to 1, which demonstrates that the model retrieved all the active MAO-B inhibitors.

Specificity:

$$Sp = \frac{TN}{TN + FP} = \frac{4}{4 + 3} = 0.5714$$

The Sp value of 0.5714 indicates that more than half of the decoys were discarded. Therefore the model may also retrieve decoys in addition to truly active MAO-B inhibitors. The model therefore has a moderate ability to distinguish between inhibitors and non-inhibitors.

Yield of actives:

$$Ya = \frac{TP}{n} = \frac{10}{13} = 0.7692$$

The Ya value is calculated as 0.7692. Therefore 76.92% of the hits are truly active.

Accuracy:

$$Acc = \frac{TP + TN}{N} = \frac{10 + 4}{17} = 0.8235$$

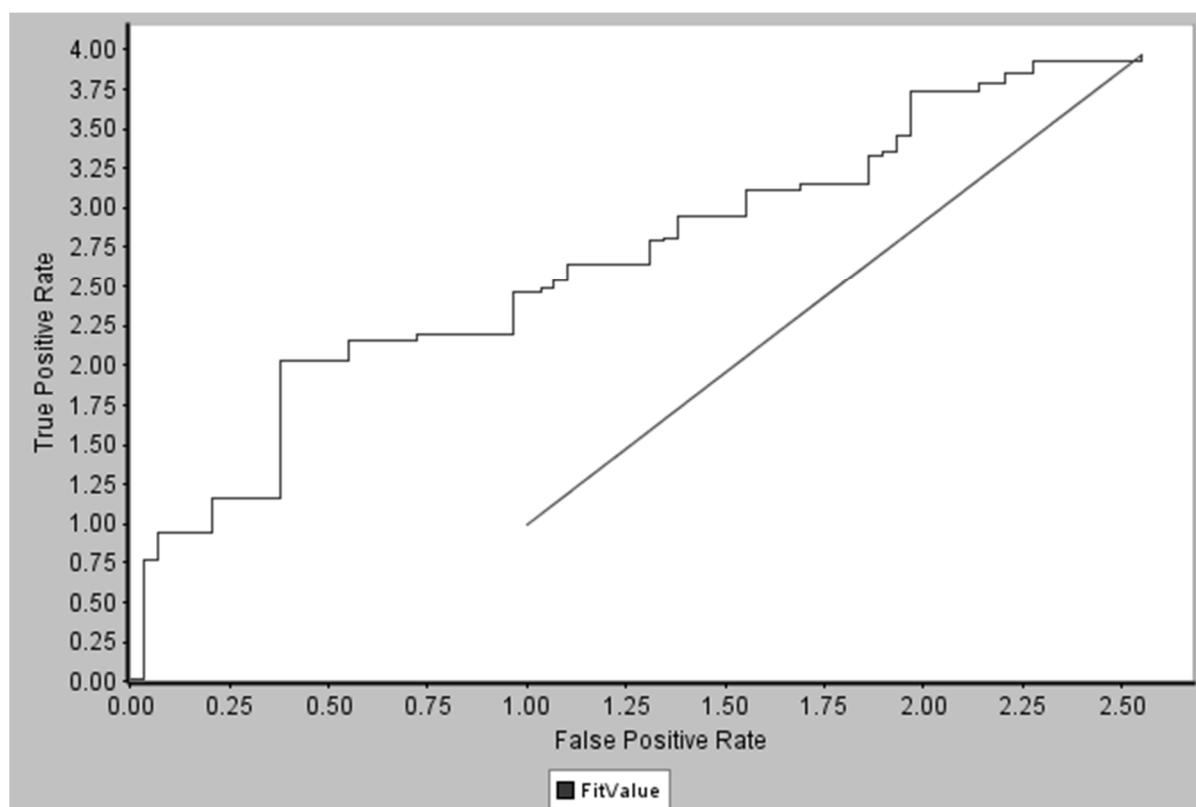This result shows that the model is 82.35% accurate.

Balanced labelling performance:

$$l_{bal} = \frac{1}{2}.Se + \frac{1}{2}.Sp = \frac{1}{2} + \frac{0.5714}{2} = 0.7857$$

The value of 0.7857 indicates a good balanced labelling performance.

ROC curve:

Score Property	Area under ROC Curve	ROC Evaluation
FitValue	2.926	Excellent

Figure 3.3.16 A ROC curve for the MAO-B pharmacophore model.

As seen in figure 3.3.16, the ROC curve of the pharmacophore model lies closer to the ideal curve than to the Se = 1-Sp asymptote and may therefore be considered to be an excellent curve.

The pharmacophore model constructed for MAO-B was sensitive to all of the known inhibitors of MAO-B in the test set and moderately specific with regards to identifying compounds that do not bind to MAO-B. This model has a very good accuracy and balanced labelling performance and therefore it can be a useful tool to identify compounds that can act as MAO-B inhibitors from a database.

3.4 <u>Summary:</u>

In this chapter pharmacophore models of MAO-A and MAO-B were constructed according to the given protocol. These models may be used for the virtual screening of the FDA approved drugs as potential MAO inhibitors. Descriptions of the pharmacophore models were given and key interacting residues in the active sites of MAO-A and MAO-B were identified. Evaluation of the pharmacophore models showed that they are accurate enough to be used in a virtual screening process. Table 3.3.2 in this chapter provides a list of selected drugs, which mapped to the MAO-A and MAO-B pharmacophore models. These drugs were selected for *in vitro* evaluation. It was interesting to note that with the exception of one drug, ranolazine, all of the 26 selected hits mapped to both the MAO-A and MAO-B pharmacophore models. This result is prabably due to the fact that the MAO-A and MAO-B active sites are highly similar with only six of sixteen residues differing between the two isoforms (Son *et al.*, 2008).