Chemometric Analysis to Probe Blood Brain Barrier Permeability of New Chemical Entities (NCEs)

Inam Tahir

(Fall 2015-MS CS&E 00000119555)

Supervisor: **Dr. Ishrat Jabeen**

Research Center for Modeling and Simulation,

National University of Sciences and Technology,

Islamabad,

Pakistan.

2018

Chemometric Analysis to probe Blood Brain Barrier Permeability of New Chemical Entities (NCEs)

This thesis is submitted in partial fulfillment of the requirements for the degree of

Master of Science

In

Computational Sciences and Engineering

Inam Tahir
(Fall 2015-MS CS&E 00000119555)

Supervisor: **Dr. Ishrat Jabeen**

Research Center for Modeling and Simulation,

National University of Sciences and Technology Islamabad

Pakistan.

2018

I dedicate this thesis to ....

HUMANITY

Acknowledgements

In the name of Allah, the Most Gracious and the Most Merciful, Alhamdulillah, all praises to Allah for the strengths and His blessings upon me in completing my degree.

Firstly, I would like to thank my supervisor and mentor Dr. Ishrat Jabeen. I am immensely grateful and indebted to her for her expert, sincere and valuable guidance and encouragement throughout my degree.

I cordially thank my family especially my grand-father Muhammad Sadiq Khan, my grandmother Naik Akhter, my father Tahir Kamal, my mother Riffat Kamal, my brother Ikram Tahir, my sister Habiba Tahir and my fiancé Samia Shafqat for their prayers, love and support throughout my life. this journey would not have been possible without the support of my beloved family.

I would like to acknowledge the efforts Miss Samia Shafqat for helping me in troubleshooting research problems.

My sincere gratitude to miss Ayesha Arif, miss Maryam Nisar and miss Sanam Banaras for their guidance and kind support in hour of need. I am immensely grateful to them for their advices and all the help in completing my degree.

I would like to thank my lab fellows for their support. A special thanks to Miss Sadia, Miss Saba and Miss Usrah for their precious time and help in completing my project.

I would like to thank National University of Sciences and Technology, for equipping me with required knowledge to do this project at Research Center for Modeling and Simulation.

Table of Contents

List of Abbreviations

2D	2 Dimensional
3D	3 Dimensional
Å	Angstrom (10^{-10} Meter)
ACC	Hydrogen Bond Acceptor
AMANDA	Algorithm
BBB	Blood Brain Barrier
BMEC	Brain Micro Vessel Endothelial Cell Culture
Cbr	Brain Tissue Concentration
CLACC	Consistently Large Auto And Cross Correlation
CNS	Central Nervous System
Cp	Plasma Concentration
CSF	Cerebrospinal Fluid
DON	Hydrogen Bond Donor
Ees	Electrostatic Energy
Ehb	Hydrogen Bond Energy
Elj	Lennard-Jones Energy
FFD	Fractional Factorial Design
GRIND	GRID Independent Molecular Descriptors
HPLC	High-Performance Liquid Chromatography
IAMs	Immobilized Artificial Membrane
JAMs	Junctional Adhesion Molecules
Log BB	Log Of Blood Brain Barrier Permeability
LOO	Leave One Out
MIF	Molecular Interaction Field
MOE	Molecular Operating Environment
PAMPA	Parallel Artificial Membrane Permeability Assay
PC	Principal Component
PCA	Principal Component Analysis
Pgp	P-Glycoprotein
PLS	Partial Least Square Analysis
QSAR	Quantitative Structure Activity Relationship
RMSD	Root Mean Square Deviations
TJs	Tight Junctions

List of Figures

List of Tables

Abstract

In present world one of greatest challenges faced by pharmaceutical industry is to maximize efficacy of various neuroactive agents and minimize the risk of neurotoxicity of drugs designed for peripheral body systems. Bioavailability of a neuroactive active agent depends upon its transport across blood brain barrier and thus, is a factor of vital importance in determining drug efficacy. Over the past few decades, intensive research efforts have been made to elucidate blood brain barrier permeability of compounds. However, these experiments are very costly and time demanding endeavors. Therefore, in this study various chemometeric models, including Principal Component Analysis (PCA) and Partial Least Square Analysis (PLS) using Grid Independent Descriptors have been developed to predict blood brain barrier permeability of a diverse dataset of 218 compounds. Our model elucidates two hydrogen bond donor groups at a mutual distance of 6.00 Å to 6.40 Å and a hydrophobic group at a mutual distance of 10 Å to 10.4 Å from one of the hydrogen bond donor groups may have a positive impact on blood brain barrier permeability of already marketed neuroactive agents. Overall, this study can prove to be useful for prediction of blood brain barrier permeability of new chemical entities in early stages of drug development.

1. Introduction

Vesicular system is responsible for the transport of materials such as food and oxygen, throughout the human body. This system consists of a network of capillaries which supply blood to every organ of the body including the brain. Brain is a delicate organ of the human body and it needs to be protected in every possible way. Considering the delicacy of the brain the capillaries supplying blood to brain are also specialized in transporting materials selectively. These specialized capillaries form blood brain barrier (Abbott, 2013). Blood brain barrier is a selective barrier which consists of capillary forming endothelial cells surrounded by astrocytic feet processes, pericytes and perivascular neurons. These endothelial cells are joined to each other through transcellular proteins such as claudins, occludin and JAMs (Junctional adhesion molecules) forming tight junctions (TJs). Unlike a normal capillary these tight junctions restrict the passive movement of the molecules across blood brain barrier, providing protection to the brain from pathogens and toxins. To modulate paracellular transport these transmembrane proteins are joined to cytoskeleton through accessory proteins like singulin and zonula occludin (Ballabh, 2004). Numerous transporters, present on luminal and abluminal membranes of the cerebral endothelial cells are responsible for controlling transcellular traffic between brain and blood. The efflux of the waste products and harmful substances is also performed by these transporters especially P-glycoprotein (P-gp). Antibiotics, peptides and lipoproteins are transported through receptor mediated transcytosis and nonspecific adsorptive-mediated transcytosis. The transcytosis activity of the BBB is suspected to be up-regulated during disease conditions (Gloor *et al*., 2001). This highly selective nature of BBB restricts free movement of the molecules across it thus providing protection brain. However, a lot of potentially active drugs are also not allowed to enter brain through normal circulatory system due to tight junctions and high levels of efflux transporters (Francisco *et al*, 2011). This protective mechanism of blood brain barrier has been a major obstacle in designing drugs for various psychological and

neurological diseases for a couple of decades (Pangalos *et al.*, 2007). Therefore, it is very important to know the BBB permeability of drugs-like entities for CNS drug designing and also to avoid psychotropic effects of CNS inactive drugs. A gold standard parameter used to measure the BBB permeability experimentally is log BB. Log BB is described as the ratio of the concentration of drug inside the brain to the concentration of the drug in plasma. The Log of this ratio of concentration between brain and blood is signified as log BB and is given by Equation mentioned below.

$$\textbf{Log BB} = \textbf{Log} \left(C_{BRAIN}/C_{BLOOD}\right)$$

Higher the value of log BB for a compound, higher is its BBB permeability (Carpenter *et al*., 2014).Various traditional methods have been used previously to calculate Log BB experimentally. These include *in vivo* techniques which involve the dissection of lab rats and *in vitro* techniques such as parallel artificial membrane permeability assay PAMPA and immobilized artificial membrane (IAM). PAMPA has been able to show good prediction ability (Masungi *et al*., 2008; Mensch et al., 2010; Tsinman *et al*., 2011;Campbell *et al*., 2014; Könczöl *et al*., 2013). PAMPA was developed by Kansy *et al,* in 1998 to predict passive permeability through gastrointestinal (GI) tract but Di, *et al.* adapted it for BBB studies (Di *et al*., 2003). PAMPA consist of two compartments namely a donor and an acceptor compartment. The compound to be tested permeates through donor compartment to acceptor compartment to monitor its permeability. The IAMs mimic phospholipid environment of the BBB by anchoring synthetic lipid analogs to silica particles in monolayer density. These particles are used making HPLC column as packing material (Yang *et al*., 1997). Chromatic retention factors are then used as for making predictions. These methods produce reasonable predictions. However, these methods are time consuming and very expensive. With the advancement is the field of computational sciences this job has been made time and cost effective. In quest of finding the effect of structure on the BBB permeability of the molecules, 3D QSAR models have been developed in past few decades to predict BBB permeability (Liu *et al*., 2004; Bujak*et al*., 2015; Vilar*et al*., 2010; Katritzky*et al*., 2006; Wu*et al*., 2012). The first QSAR study is believed to be conducted by Young and co-workers on 20 histamine antagonists in 1988 (Young *et al*., 1988).

Later, the parameters such as lipophilicity, topological indices, solvatochromic parameters and combination of these parameters were used by Abraham *et al*.(a.1994, b.1995), Norinder *et al*.(1998), Brewster *et al.* (1996), Kelder *et al.* (1999),Clark *et al.* (1999), Subramanian *et al.* (2003), Luco *et al*. (1999)and Fehr *et al.* (2000)for the development of log BB models. These studies identify three main properties such as lipophilicity, hydrogen bonding potential and molecular volume to be responsible for permeability of molecules across BBB.

However, these classical 3D QSAR approaches were alignment dependent, time consuming and produced user biased results depending upon the alignment used (Carpenter *et al*., 2014; Kubinyi*et al*.1993; Kubinyi*et al*., 1997; Verma, J*et al*., 2010; Jagiello*et al*., 2016; F Morales*et al*., 2017). More over superimposition of a structurally diverse dataset is nearly impossible and may consume intensive computational power for large datasets. To counter this problem, alignment free approaches based on autocorrelation functions have been suggested by Broto *et al., 1984*, Gasteiger *et al*., 1996 and Clementi *et al*., 1993. Broto used classic autocorrelation transform to obtain autocorrelation vectors from 2D and 3D structures. Gasteiger used special autocorrelation function on molecular surface properties. Clementi also used autocorrelation vectors but only for planer compounds (Pastor *et al*., 2000). All these approaches were quite effective in providing solution to alignment problems faced by classical QSAR models yet necessary transformations produce difficulty in interpretation of resultant models in original descriptors space.

In present study, a 3D-QSAR model has been developed using molecular force field based chemically interpretable descriptors known as GRID independent descriptors (GRIND). These descriptors being highly relevant to biological properties of compounds and easy to compute are derived from Molecular interaction force field and require no superimposition of compounds. An important feature of GRIND is that they are independent of the position and orientation of molecular structure in space (Pastor *et al*., 2000). Instead of absolute 3D coordinates as in classical 3D QSAR, GRIND measures the distances between relevant groups (Pastor *et al*., 2000). Force fields are calculated using GRID program and these fields are simplified afterwards.

Using a particular transform function known as CLACC, results are encoded in the form of variable which are alignment independent. Descriptors are then used to obtain graphical diagrams called correlograms in which the products of node-node energies are plotted against the mutual distances between nodes. In this study Principal Component analysis and a powerful regression analysis method known as partial least square analysis have been developed using GRIND descriptors. Principal Component Analysis explained the structural variance of the dataset compounds. Whereas PLS model with q^2=0.50 and r^2= 0.63 after leave one out (LOO) cross validation explains the effect of structural features and their mutual distances from each other on Log BB. The model is externally validated using a test dataset of 44 diverse compounds along with their experimental log BB values. Almost all the compounds of the test set have been predicted well with a difference of less than 1.5 log unit between actual and predicted log BB values. Hydrogen bond donors/acceptors, hydrophobic features and shape based features of the molecules have been identified as features effecting blood brain barrier permeability of the compounds. Moreover, the significant outcome of this study is that the presence of hydrogen bond donor groups at a mutual distance of 6.00 to 6.40 Å and a hydrophobic group at a distance of 10Å-10.4Å from one of the hydrogen bond donor groups may have a positive impact on blood brain barrier permeability of already marketed neuroactive agents,

2. Literature Review

2.1 Blood Brain Barrier:

The central nervous system is surrounded by a unique microvasculature system termed as 'the blood brain barrier' possessing peculiar properties compared to rest of the body organs. BBB is composed of highly specialized cellular system possessing modified basal membrane embedded with large number of pericytes and astrocytic end feet regulating the movement of cells, ions and molecules (Daneman, 2012). Moreover, the endothelial cells surrounding the brain tissues differ in two ways from endothelial cells of other body organs i.e., presence of continuous tight junctions preventing paracellular movement of compounds and secondly there is no transendothelial pathways such as intracellular vesicles discovered to date as shown in figure 2.1 thus, BBB forms a strict barrier between blood and the brain tissues (Pramod D).

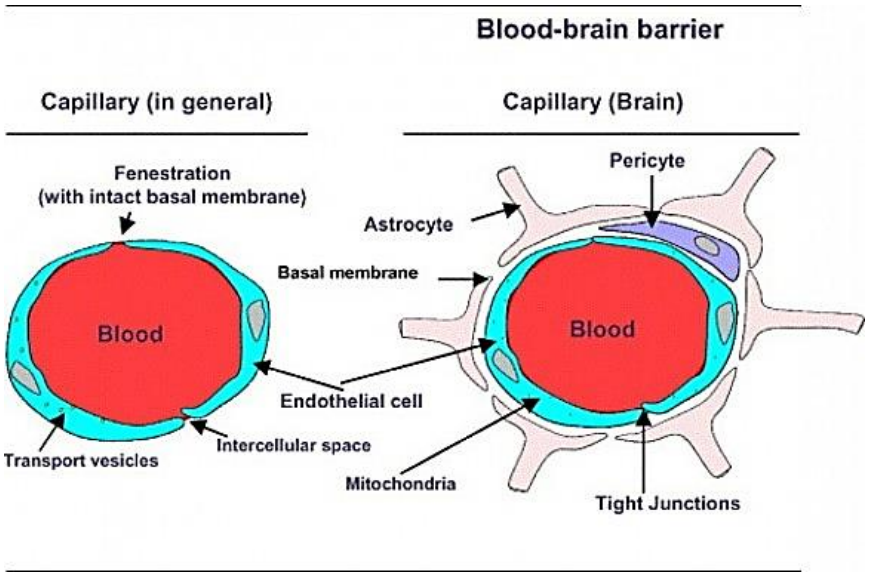

***Figure 2.1**: Difference between capillaries transporting blood to other tissues of the body and capillaries transporting blood to CNS; BBB is composed of specialized endothelial cells possessing continuous tight junctions and basal membrane embedded with large number of pericytes and astrocytic end feet forming a barrier (Adapted from Esparza C, 2015)*

The major role of BBB is to tightly regulate the homeostasis of the neuronal tissues and restricting entry of pathogens, toxins and other harmful substances that hinder normal neuronal functions. This protective nature of BBB is in turn a major obstacle in delivery of drugs to the central nervous system and many efforts have been made to understand the nature of BBB for efficient delivery of therapeutic drugs (Larsen *et al*., 2014).

2.2 Discovery of blood brain barrier:

The blood brain was first identified by Paul Ehrlich in 1880s, when he administered certain dyes such as trypan blue intravenously and observed the effect of dye on all organs of the body except the brain and spinal cord. The results were very astonishing thus he concluded that the dyes have lower affinity for brain tissues compared to other tissues of the body.

In 1913, another scientist, Edwin Goldman demonstrated the effect of the same dyes by injecting them directly into the cerebrospinal fluid, he found that the brain tissues were immediately stained but the rest of the body tissues were not stained depicting a physical barrier between the nervous tissues and rest of the body tissues.

The term ‘blood-brain barrier’ was first coined in 1898 by Lewandowsky. He conducted a group of experiments along with his colleagues to study the effect of neurotoxic drugs on brain function. They found that these neurotoxic agents affected the brain structures only when they were directly injected into the brain but not when they were injected into the vascular system. Later Reese and colleagues studied the structural features of blood-brain barrier using electron microscope.

2.3 Structural Components of BBB:

The vascular system is composed of blood vessels transporting blood to all parts of the body. These blood vessels are made of two types of cells. Endothelial cells that forms the walls of the blood vessels and mural cells that are present on the abluminal surface of the endothelial cell layer. The BBB is formed by these cells along with critical interaction with neural cells, immune cells and glial cells as shown in Figure 2.2 (A) and (B).

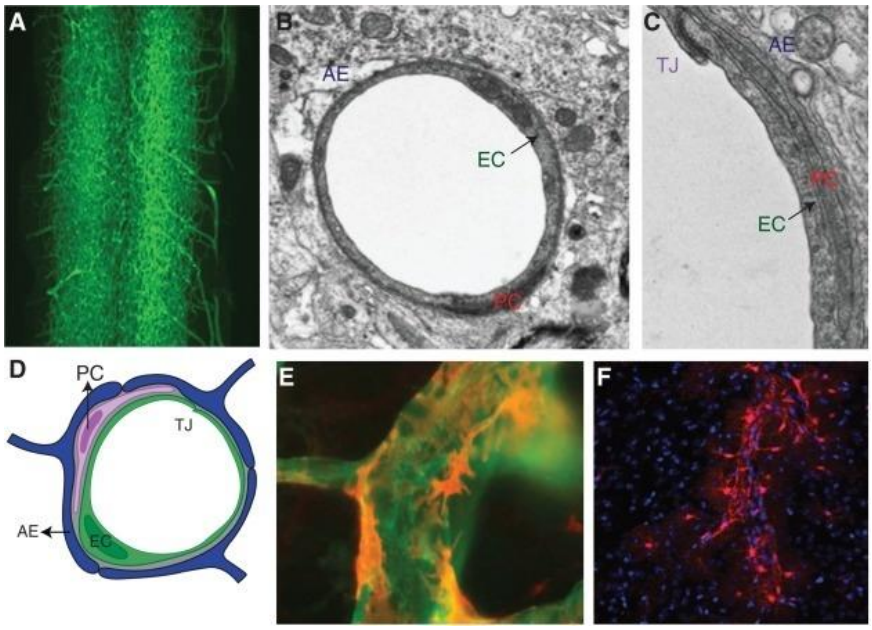

Figure 2.2: *Study of structural components of BBB. (A) vascular tissue cast of the spinal cord depicting dense network of vascular system in the central nervous system. (B) The electron micrograph of cross section of CNS blood vessel showing the BBB system formed by endothelial cells (ECs), pericytes (PCs), and astrocytes. (C) electron micrograph showing the tight junctions of endothelial cells, pericytes, basal membrane and astrocytes. (D) Types of cells in the BBB. (E) Immunofluorescence micrograph representing relationship of pericytes in red with endothelial cells in green. (F) Micrograph illustrating the association between astrocytic end feet red-labeled with GFAP-cre; Rosa-tdTomato with blood vessels. Astrocytes ensheath the blood vessels (Adapted from Daneman, R.& Prat, A. 2015)*

2.3.1 Endothelial cells:

Endothelial cells are simple form of squamous epithelial cells that forms the walls of the blood vessels. The large arteries and veins possesses up to dozens of endothelial cells while small capillaries are composed of a single cell that folds to form the lumen of the vessel (Aird, 2007) the endothelial cells present in the central nervous system bears unique properties compared to endothelial cells present in other tissues. The unique property is the presence of tight junctions (TJs) between adjacent endothelial cells. This property limits the movement of molecules paracellularly (Reese and Karnovsky, 1967; Westergaard and Brightman, 1973). Moreover, these endothelial cells restrict vesicle mediated transcellular transport of molecules (Coomber and Stewart, 1985). This transcellular and paracellular barrier restricts the movement of materials between the blood and the brain and transport is regulated by highly specialized transport systems (Betz *et al*., 1980).

2.3.2 Mural cells and pericytes:

These are basically vascular smooth muscle cells and pericytes that are involved in forming normal vasculature. They partially cover the walls of the blood vessels. Pericytes sits on the abluminal surface of the endothelial cells of micro vessels of central nervous system embedded in basal membrane (Sim 1986), due to lack of markers specific to pericytes there has been a difficulty studying them and are often confused to other cells that are present in perivascular space (Armulik *et al.*2011). Armulik *et al*. (2011) identified NG2 and PDGFR-b to have positive reactivity to pericytes in the central nervous system. Pericytes can extend along the abluminal surface of the endothelial cells and can span several endothelial cells; they contain a specific contractile protein that has the ability to modulate the diameter of the capillaries (Peppiatt *et al*., 2006; Hall *et al*., 2014). It is seen that although pericytes line the endothelial tube but they have no physical contact with the endothelium and are separated by basal membrane in which they are embedded. However, the specific points at with the pericytes form cellular adhesions with endothelium mediated by N-cadherin are termed as peg and socket junctions (Gerhardt *et al*., 2000). Pericytes of the BBB are derived from neural crest compared to pericytes of other tissues

which originate from mesoderm (Majesky, 2007). The BBB pericytes are responsible for angiogenesis, regulation of immune cells infiltration, wound healing, deposition of extracellular matrix (Armulik *et al*., 2011). Shepro and Morel (1993) determined that the endothelial: pericyte ratio is higher in central nervous system microvasculature which is approximately between 1:1 and 3:1 while ratio of 100:1in muscle blood vessels.

2.3.3 Basal membrane:

The blood vessels possess two types of basal membranes that is outer parenchymal basal membrane and inner vascular basal membrane. Vascular basal membrane is basically an extracellular matrix produced by pericytes and endothelial cells while astrocytes are responsible for the synthesis of parenchymal basal membrane (Sorokin, 2010). The basal membranes are responsible for several important roles such as secretion of glycoproteins, type IV collagen, proteoglycans; heparin sulfate, laminin and nidogen thus, help in signaling processes at the vasculature (Daneman, 2015). Damage to basal membrane can cause disruption of the BBB and leukocyte infiltration seen in various neurological diseases (Daneman, 2015).

2.3.4 Astrocytes:

Astrocytes are type of glial cells that ensheath the blood vessels and neuronal connections (Abbott *et al*., 2006). The end feed basal membrane contains three proteins that are aquaporin 4, dystroglycan and dystrophin and thus, completely ensheath the microvasculature. A complex is formed between the dystrophin protein and dystroglycan protein that links the end feet to basal membrane (Noell *et al*., 2011). The aquaporin protein is important for maintaining the water

homeostasis of central nervous system. The major role of astrocytes is to form a physical link between the neuronal tissues and blood vessels (Daneman, 2015). Astrocytes regulate the blood flow in the blood vessels in response to neuronal activity by dilation and contraction of vascular smooth muscle cells and pericytes surrounding the capillaries (Gordon *et al*., 2011). Janzerand Raff (1987) studied the important role of astrocytes in BBB since they produced the same properties in peripheral blood vessels in transplantation studies.

2.3.5 Immune cells:

The blood vessels of the central nervous system interact with the immune cells that are present within the brain tissues as well as the immune cells in blood. Macrophages and glial cells are the major immune cell population of central nervous system. Macrophages are present at the abluminal side of the microvasculature which is a fluid filled canal called Virchow–Robin space (Polfliet *et al*., 2001). Studies suggest that these cells have the ability to pass the BBB and provide the first line of defense against pathogens via phagocytosis (Williams *et al*., 2001). Microglial cells enter the central nervous system during embryonic development and are parenchymal immune cells that are involved in innate immune response and wound healing (Ginhoux *et al*., 2010). Other immune cells include neutrophils, macrophages and T lymphocytes that protect the BBB in response to pathogens, tissue injury and disease condition by altering vascular permeability by producing a ROS species (Persidsky *et al*., 1999; Hudson *et al*., 2005).

2.4 Transport mechanisms across the BBB:

The BBB possess a number of highly selective mechanisms for the transport of essential nutrients and molecules into the CNS.

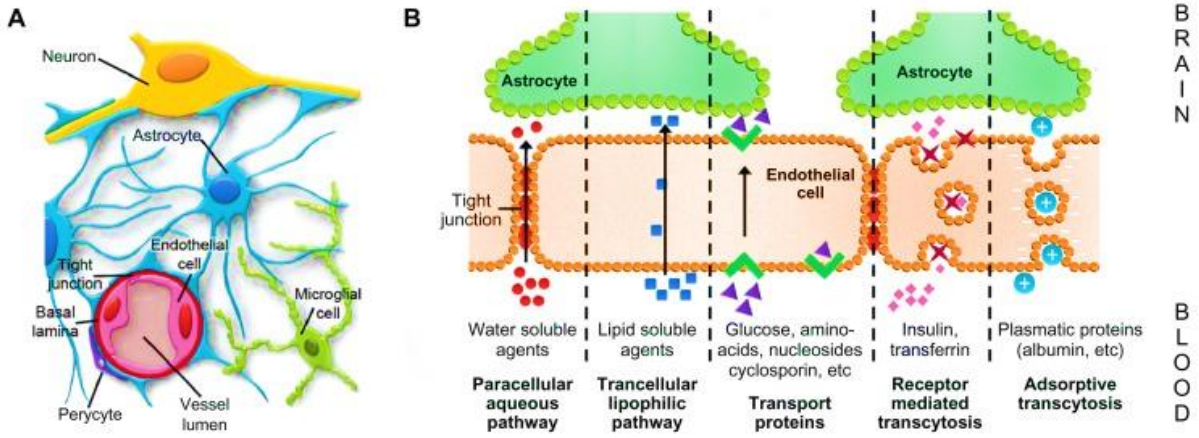

Figure 2.3: *(A) Anatomy of the brain microvasculature, (B) Transport mechanisms across the blood brain barrier. 1. Paracellular aqueous pathway that involves paracellular movement of small water-soluble compounds via diffusion. 2. Transcellular lipophilic pathway that involves movement of lipid soluble compounds across the BBB via diffusion. 3. Facilitated diffusion of selective compounds via transport proteins. 4. Receptor mediated transcytosis is movement of selective macromolecules such as insulin, transferin in clathrin coated vesicles. 5. Adsorptive transcytosis involves movement of positively charges proteins such as albumin across the plasma membrane. (Adapted from Ramos-Cabrer et al. 2013).*

2.4.1 Diffusion across the BBB:

The BBB allows the transport of certain molecules via diffusion. The diffusion across the BBB is divided into two types. One is Paracellular diffusion and second is Transcellular diffusion. The paracellular diffusion does not widely occur across the BBB due the presence of tight junctions in endothelial cells. Transcellular diffusion involves the movement of lipophilic compounds across the BBB. The greater the extent of lipophilicity of molecule the greater is its diffusion across the BBB (Pardridge, 1998). The other factor that effects the diffusion of molecules is the

molecular weight of compounds, the smaller the molecular weight of molecules the greater probability of their penetration across BBB such as O_2, NO, CO_2, H_2O are highly permeable compounds. The third factor is the presence of hydrogen bond donor groups will significantly enhance the membrane permeability of compounds (Burton *et al.*, 1993). Diffusion is a spontaneous process that involves the movement of molecules across the concentration gradient. The diffusion across the membrane is directly dependent on the concentration gradient described by the equation as follows

$$DG = (R) \times (T) \times (In \times [Ci] / [Co])$$

Where, DG is described as Gibbs free energy change, R is gas constant and is equal to 1.987 cal/mol·°K, T is the absolute temperature in degrees Kelvin and [Ci]/[Co] is the ratio of concentration of molecules on inside of membrane (Ci) and outside of membrane (Co) (Karp, 1999).

In case of electrolytes the two gradients must me considered the chemical gradient and electric potential gradient. The equation would be as follows:

$$D\ G = [(R)\ x\ (T)\ x\ (ln\ x\ [Ci] / [Co])] + [(z)\ x\ (F)\ x\ (D\ Em)]$$

The ‘z’ is the charge of the electrolyte, D Em is potential difference between membranes and ‘F’ is Faraday’s constant that equals to 23.06 kcal/V (Stein, 1967).

2.4.2 Facilitated diffusion via transporters:

Facilitated diffusion across the BBB is basically energy independent carrier mediated endocytosis that involves a carrier protein or transporter that undergoes a conformational change there by transporting the molecules across the membrane. The most common molecules that are

transported via transporters are amino acids, nucleosides, cyclosporine, glutathione, monocarboxylates, hexoses amines, small peptide molecules, etc. (Tsuji and Tamai, 1999).

2.4.3 Receptor mediated transcytosis:

Endocytosis or transcytosis are of two types; the receptor mediated transcytosis and adsorptive transcytosis. Receptor mediated transcytosis involves the selective uptake of macromolecules by receptors present on the surface of endothelial cells. The substances include hormones, growth factors and enzymes such as insulin (Duffy and Pardridge, 1987), transferrin (Fishman *et al*., 1987) and leptin (Banks *et al*., 1996). It is a highly specific ATP dependent transport mechanism in which the molecules to be transported becomes attached to the receptor in a region called coated pits. These coated pits contain clathrin protein, and other electron dense proteins (Moore *et al*., 1987). Once the molecule attaches to the receptor the coated pits invaginate into the cytoplasm then pinch free to form coated endosomes. Then clathrin coat is removed inside the cell to form smooth vesicles (Stahl and Schwartz, 1986). The ATPases causes the acidification of the vesicles resulting in detachment of molecules from the receptors.

2.4.4 Adsorptive transcytosis:

Adsorptive transcytosis possesses high capacity but is low affinity transport system as compared to receptor mediated transcytosis. It is mediated by electrostatic interaction between a positively charged molecule such as charged peptide and the negatively charged surface of plasma membrane that is glycocalyx (Gonatas *et al*., 1984).

2.5 BBB a challenge in drug designing:

Highly selective nature of the blood brain barrier poses a challenge to the drug designers. There are two major concerns regarding blood brain barrier in drug designing. First the drugs that needs to be delivered to central nervous system should be able to pass through blood brain barrier. Secondly, drugs designed for the peripheral body should not cross blood brain barrier to avoid possible psychotropic side effects. Therefore, it is very crucial to know the blood brain barrier permeability of the drug molecules in early stages of drug designing. The knowledge of the blood brain barrier permeability of a drug ensures the efficacy of CNS active drugs. It also helps in designing safe drugs for peripheral systems of the body. The permeability of a compound across blood brain barrier could be found through in vivo, in vitro and *in silico* methods. The most reliable reference information for validating other models is provided by the in vivo methods. However mechanistic information has been effectively provided by the in vitro techniques. Non-cell based assays prove to be important for passively permeable compounds. Transport through mediated systems may also be predicted by the combination of *in vitro* methods and physicochemical information. *In silico* methods rely upon the quantity as well as the quality of the datasets used to build models for the prediction of blood brain barrier permeability. Two experimental approaches have commonly been used to assess blood brain barrier penetration. First one is equilibrium distribution between brain and blood and the other one is blood brain barrier permeability. The equilibrium distribution approach being a gold standard parameter has been established to calculate the blood brain barrier permeability of a compound. This gold standard parameter has been denoted as Log BB. The Log BB has been defined as the ratio of the concentration of a drug inside the brain to the concentration in blood. The Log of this ratio of concentration between brain and blood is signified as log BB and is given by Equation mentioned below.

$$\text{Log BB} = \text{Log} (C_{BRAIN}/C_{BLOOD})$$

Higher the value of log BB for a compound, higher is its blood brain barrier permeability. Log BB has been determined at steady state and through blood and brain concentration curve (Kalvass and Maurer, 2002). Low permeability of a compound can be attributed to different reason like active efflux at blood brain barrier, frequent binding to plasma protein, the sink effect of cerebrospinal fluid or low partitioning to brain tissues (Kalvass and Maurer, 2002).

2.6 Methods to calculate log BB values of drugs:

***Table 2.1**: Methods to study brain uptake of drugs including in vivo, in vitro and in silico techniques with their advantages and disadvantages. (Adapted from Bickel U, 2005)*

Technique	**/ Estimated parameter**	**Advantages**	**Disadvantages**
***In vivo* methods**			
1. Intravenous injection/ brain sampling	Influx; Influx/efflux	Most physiological approach; highest sensitivity; low technical difficulty	May require good analytical tools to exclude metabolite uptake and careful pharmacokinetic analysis to discriminate unidirectional uptake versus bidirectional transfer

2.	Brain uptake index	Influx	Fast procedure; moderate technical difficulty; permits wide range of modifications of injectate composition; artifacts by metabolism largely excluded	Relatively insensitive (compared with intravenous injection and brain perfusion)
3.	Brain perfusion	Influx	Higher sensitivity compared with BUI; permits modification of both perfusate composition and flow rates; artifacts by peripheral metabolism excluded	Technically more difficult than intravenous experiments and BUI
4.	Quantitative autoradiograph	Influx	Excellent spatial resolution	Time-consuming evaluation; no proof of integrity of tracer
5.	External registration: MRI, SPECT, PET	Influx/efflux	Non-invasive and applicable in humans; allows time course measurements in individual subjects	Expensive equipment (MRI, PET) and tracers (PET); limited sensitivity (MRI) and availability of labeled tracers (MRI, PET); poor spatial resolution for small animals (SPECT)

6.	Microdialysis	Influx/efflux	Allows time course measurements in individual subjects; samples well suited for subsequent analytical procedures	Technically involved; *in vivo* probe calibration required for valid quantitative evaluation; local damage to BBB integrity
7.	CSF sampling	Influx/efflux	Readily accessible for sampling; applicable to humans	Reflects permeability of B-CSF-B and CSF fluid dynamics rather than BBB

***In vitro* methods**				
1.	Fresh isolated brain micro-vessels	Binding, uptake, efflux	Representing the in vivo expression of transporters and efflux systems at the BBB	Transcellular passage cannot be measured
2.	EC membrane vesicles	Carrier-mediated transport	Allows distinction of luminal versus abluminal transport activity	Large amounts of source material required, laborious preparation
3.	Endothelial cell culture Primary cultures, cell lines	Receptor binding; uptake; luminal to abluminal transfer (and opposite direction)	Permeability screening experiments (feasible with primary EC from bovine/porcine sources); effect of culture conditions on BBB transport properties may be studied (e.g., astroglial factors, serum effects, inflammatory stimuli, hypoxia/aglycemia	No system yet able to represent in vivo condition with respect to barrier tightness and BBB specific transporter expression; multitude of models makes comparison of results between studies difficult

***In silico* models**			
Rules of thumb Classification models **QSAR**	CNS active (+/-); Log BB; Log PS	Screening of large compound libraries (depending on model selection and computational resources); screening of virtual libraries	Many current models based on data, which may not represent BBB permeability as such (log BB; CNS activity); still very limited data bases for BBB transport (log PS models)

2.6.1 *In vivo* methods:

Two major methods namely Log BB and Log PS are used most widely for the assessment of blood brain barrier permeability through in vivo approach. The most effective technique is the **Intravenous brain uptake studies** that possess maximum sensitivity and specificity. Moreover, the intravenous experiments allow the calculation of tissue uptake for a long duration of time (Bickel U, 2005). In this case two pharmacokinetic approaches needs to be considered for brain uptake studies. One is unidirectional uptake of drug from blood plasma to brain tissue and second is reversible uptake of drug between blood and brain tissue that is efflux and influx occurring simultaneously (Bickel U, 2005). The intravenous blood uptake analysis is performed by sampling of arterial blood at suitable time intervals to determine the time curve of plasma concentration (Cp) and brain tissue concentration (Cbr) by obtaining a brain tissue sample. The Cbr of unidirectional brain uptake at time is calculated as,

$$\mathbf{Cbr = K_{in} \times AUC^{T}_{0}}$$

Where, K_{in} is unidirectional influx constant from plasma into brain, AUC^{T}_{0} is integral of plasma conc. from 0 to time T (Kennedy, 1997).

Brain uptake index and isolated brain perfusion are both related to intravenous brain uptake method with certain advantages. These methods allow the manipulation of intravascular fluid composition in brain microcirculation within broad range which is not possible in cases of intravenous studies (Bickel U, 2005)

Brain micro dialysis technique has been used to study the distribution of drug to the brain. This method involves the implantation of stereotaxic probe causing acute invasive injury under anesthesia. Once the probe is implanted the sampling can be done repeatedly (Deguchi and Morimoto, 2001).

Sokoloff's *et al*., (1977) developed the **Quantitative autoradiography (QAR)** method to measure the local glucose metabolism using [14C] 2-deoxyglucose and for blood flow with iodo [14C] antipyrine. It has advantages in brain tumor studies and ischemia. The 14C is labeled with radioactive tracer and is administered via intravenous route. Then the blood samples are taken to measure the concentration at specific time course.

2.6.2*In vitro* methods:

Isolated brain micro-vessels, is a technique that involves isolation of viable micro vessels from brain tissue of species including human autopsy brain. The capillaries that are isolated remain metabolically active despite reduction in ATP content. The expression of specific genes of transport proteins and endothelial cell surface receptors specific to BBB provides in vivo conditions for experimentation (Pardridge WM., 1998). This method provides an efficient system for studying transport mechanisms of BBB for transport of nutrients and proteins into the brain tissues.

Endothelial cell culture models have been used to study BBB permeability and drug development studies. BMEC culture (brain micro vessel endothelial cell culture) has been used for quantitative permeability studies (Gumbleton M and Audus KL, 2001).

Primary endothelial cell cultures have also been used and cultures have been developed from mouse, rat, bovine, porcine, primate, human brain but only porcine and bovine primary cultures have provided sufficient data for screening of drugs and studying of BBB permeability. Several variants of bovine primary endothelial cultures have been used for studying drug permeability (Audus KL *et al*., 1998).

2.6.3*In silico* methods:

In silico approach involves the development of a computational model to predict BBB permeability based on their structure. Lipophilicity and molecular weight are the two principle molecular properties upon which the in silico modeling of BBB permeability have been studied (Oldendorf WH, 1974; Levin VA, 1980). Permeability modeling is primarily based on passive transport across the BBB since the structural information needed for predicting active and facilitated diffusion is insufficient. Drugs are classified based on their pharmacological activity as CNS active and CNS inactive drugs (Ajay et al., 1999).

The data sets developed are usually large for up to hundreds and thousands of compounds and predictive models of test compound are built on one- and two-dimensional molecular descriptors (Engkvist O et al., 2003). The limitation of these models for predicting BBB permeability is that the CNS inactive drugs may actually be able to pass the BBB but does not possess any structural feature for permeability requirement. The accuracy of test sets for CNS active drugs is 80% while the accuracy of CNS inactive drugs is 10 to 20% lower (Clark DE, 2003).

Quantitative Structure Activity Relationship (QSAR) approach is based on pharmacokinetics that is the separation of blood and brain at steady state termed as **log BB.**

$$\text{Log BB} = \text{Log (c BRAIN/ c BLOOD)}$$

The limitations for accurate prediction of log BB values include the need of log BB values to be determined at steady state first. Taking blood and brain tissue samples before achieving steady state tissue distribution effects the brain uptake estimation. Secondly, whole brain tissue conc. is calculated that effect the estimation of free conc. and determination of specific binding sites of drug in the brain. Thirdly, the log BB value not only describes the BBB permeability but a number of other important processes in CNS such as local metabolism, binding of plasma protein, tissue binding, CSF re-absorption and efflux or clearance of unnecessary and toxic compounds via interstitial fluid flow therefore, the descriptors used to build QSAR models are still limited (Bickel U, 2005). Lipinski (2000) also stated that few molecular descriptors have been used to build QSAR models that are molecule size, polarity, lipophilicity, hydrogen bonding and charge. Computer algorithms calculating the Log P values, is one reason of its popularity which is closely accurate to experimentally calculated values (Hansch C, 2004). Many QSAR models have been developed in an attempt to increase the calculation speed with reduced molecular descriptors which are based on polar surface area/physico-chemical properties (Clark DE, 1999; Lobell M, 2003). PSA is calculated by analyzing the structure and hydrogen bonding capacity of compound. Van de Waterbeemd *et al*., (1998) described that a PSA less than 90 Å, and molecular weight less than 450 make drugs CNS active in a data set of total 125 drugs.

3. Methodology

3.1 Dataset Collection:

A structurally diverse dataset of 552 compounds along with their experimentally known log BB values has been collected from different literature sources including, Geldenhuys *et al*., (2015),Gulyaeva *et al*., (2003) and Liu *et al*., (2004). 3D structures of molecules were obtained using software MOE version v1.05, followed by computation of the partial charges and energy minimization using MMFF94 force field (Pastor *et al*., 2000). A complete data curation protocol is provided in figure 1. Briefly, at first step duplicates and fragments were removed from the data followed by the application of drug like filters as defined by Lipinski *et al, (*Lipinski et al. 1997, 2001 & 2004). This results in a final dataset of 218 compounds as training set for building the GRIND model as shown in annex 1. Interestingly, more than 53% of the training data include drugs already available in the market such as Bupropion, Sertraline, Maprotiline and Imipramine for the treatment of various neurological and psychological disorders (van de Waterbeemd*et al*., 1998). Additionally, a separate dataset of 44 compounds as shown in annex 2 has been taken from some different publication sources as test set

Computation of Physicochemical Parameters

In order to estimate the impact of physicochemical properties on blood brain barrier permeability, 2D physicochemical descriptors including log P (o/w), molecular weight (M.W) and topological polar surface area (TPSA) have been computed using software MOE version 2018-01 as provided in annex 1.

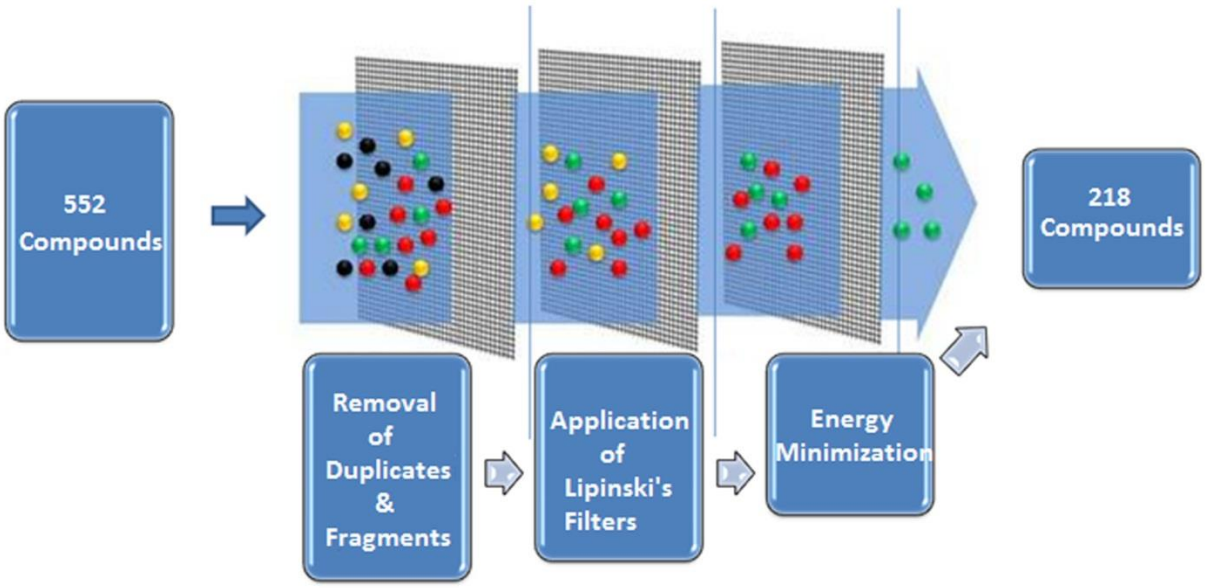

Figure 3.1: Steps involved in refinement of blood brain barrier dataset.

3.23D QSAR Modeling:

Standard 3D conformations of both training as well as test set were generated by using CORINA CLASSIC (Suenderhauf *et al*., 2012). These extended 3D conformations along with log BB values of dataset molecules were imported to software package pentacle v1.05 to compute GRID independent descriptors. GRIND approach aims to extract the information enclosed in the molecular interaction fields (MIFs) (Pastor *et al*., 2000) and compress it into new types of variables whose values are independent of the spatial position of the molecule studied. Default program values were used for the computation of MIF using four probes (DRY representing hydrophobic interaction, O (Carbonyl Oxygen) representing hydrogen bond acceptor group, N1 (Amide Nitrogen) representing hydrogen bond donor groups and TIP representing the shape descriptor). Most relevant regions were extracted from the MIF using an optimization algorithm (AMANDA) that uses the intensity of the field at a node and the mutual node-node distances between the chosen nodes as a scoring function (Pastor *et al*., 2000).At each point, the interaction energy (*Exyz)* is calculated as a sum of Lennard-Jones energy (Elj), hydrogen bond (Ehb) and Electrostatic (Eal) interactions.

$$\mathbf{E_{TOTAL}} = \sum \mathbf{E_{Lj}} + \sum \mathbf{E_{es}} + \sum \mathbf{E_{hb}}$$

Default values of probe cutoff (DRY= -0.5, O= -2.6, N1= -4.2, TIP= -0.74) were used for discretization of MIF. Nodes with an energy value under this cutoff were discarded. Finally, consistently large auto and cross-correlation (CLACC) method was used for encoding the pre-filtered nodes into GRIND. The values obtained from the analysis were represented directly in correlogram plots, where the products of node- node energies have been reported versus distance separating the nodes. In order to understand the structural variance of the data and its correlation with log BB values, Principle Component Analysis (PCA) and Partial Least Square (PLS) Analysis using GRIND variables have been performed. Briefly, Leave One Out (LOO) cross validation procedure (Gao *et al*., 2017) was utilized to correlate the observed verses predicted log BB values.

Table 3.1: Flowchart describes the steps of methodology.

Dataset Collection	•A dataset of 552 compounds along with their experimentally known log BB values was collected from literature.
Dataset Purification	•Duplicates and fragments in dataset were removed by application of molecular weight and log P filters which reduced the dataset to 218 compounds.
Building 3D structure	•3d Structures were generated in MOE using SMILE codes
GRIND MIF Computation	•In first step, the molecules were place in GRID with 0.5 as GRID step value. Each probe was placed iteratively at nodes to calculate total energy. $E_{TOTAL} = \sum E_{Lj} + \sum E_{es} + \sum E_{hb}$
GRIND Discretization	•AMANDA algorithm filtered MIF and extracted nodes of high importance.
GRIND Encoding	•Finally, pre-filtered nodes were encoded using consistently large auto and cross correlation (CLACC) algorithm.
Principal Component Analysis	•Structural variance of the data was analyzed by Principle Component Analysis (PCA)
Partial Least Square Analysis	•A partial least square analysis (PLS) model was developed to correlate the BBB permeability of molecules with 3D structural features**.**
External Validation	•A test set of 44 diverse compounds along with their Log BB values was used for external validation of model.

4. Results & Discussion

4.1 Principal Component Analysis:

Principle Component Analysis (PCA) of the GRIND variables revealed 40 % structural variance of the training data with the help of first two principle components. Briefly, 1st principle component define the data on the basis of explicit distance ranges of a hydrogen bond donor reference feature from other 3D structural features including, a hydrophobic group, a molecular steric hotspot, another hydrogen bond donor and from a hydrogen bond acceptor group within the respective extended conformation of a molecule as depicted by DRY-O, O-TIP O-O and O-N1correlograms respectively in figure 4.2 A. However, 2nd principle component differentiate the training data by delineating the distance of a hydrogen bond acceptor reference feature from same 3D structural features as characterized by 1st PC. These are illustrated by DRY-N1, N1-TIP N1-N1 and O-N1correlograms respectively as shown in figure 4.2 B. A summary, of the respective distance variables of different 3D structural features as defined by PC1 and PC2 are provided in table 4.1.

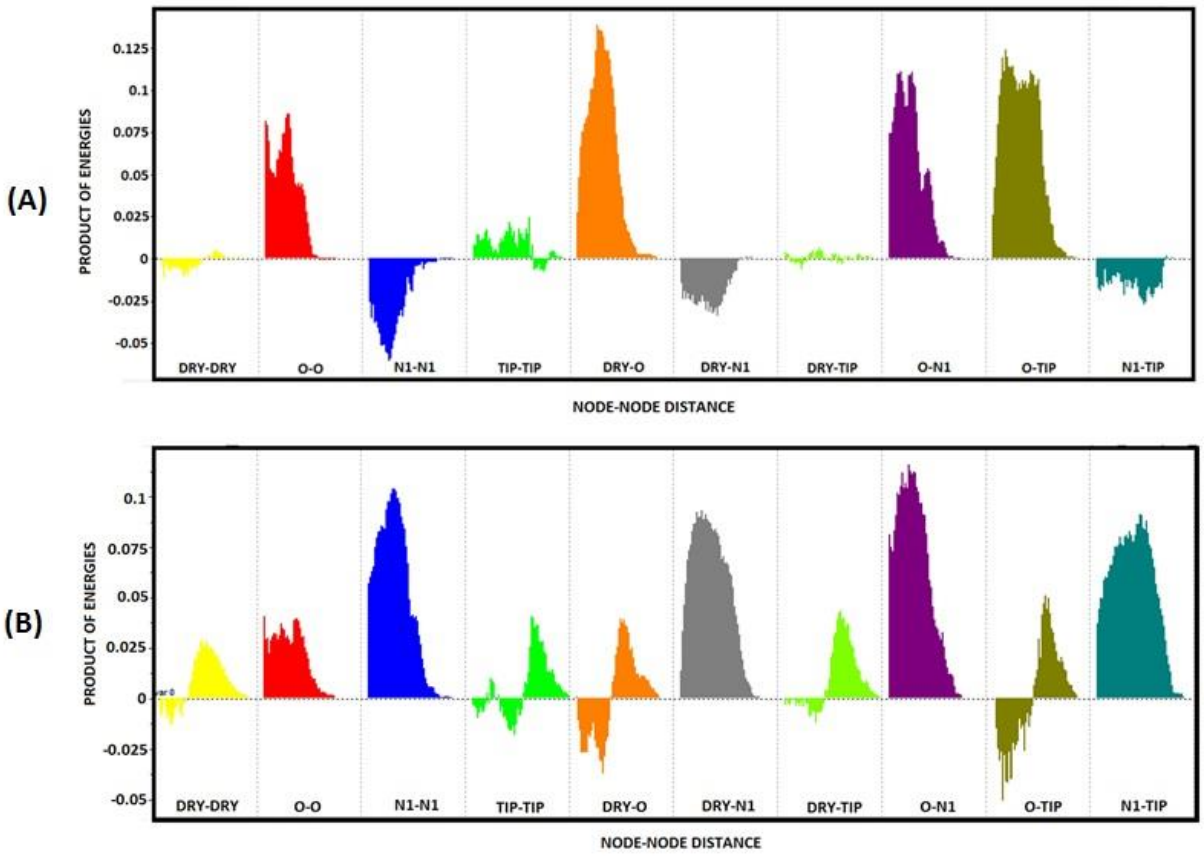

Figure 4.1: (A) Represents GRIND variables of the 1st PC where, O-O, DRY-O, O-N1 and O-TIP correlograms are characterized by the more significant variables defining the data. (B) Represents GRIND variables of the 2nd PC where, DRY-N1, N1-TIP N1-N1 and O-N1correlograms are depicted by the important variables defining the data.

Table 4.1: Representing distances in Angstrom (A) between important 3D structural features of the data depicted by first two principal components

PCs	Probes	Distances b/w GRIND Variables	Comments
PC1	O-O	6.80-7.20	Absent in cluster A & some compounds of cluster B
	DRY-O	6.40-6.80	Absent in only Cluster A
	O-N1	4.40-4.80	Absent in Cluster A & Cluster B
	O-TIP	4.80-5.20	Absent in only Cluster A
PC2	N1-N1	8.00-8.40	Absent in Cluster B & some compounds of Cluster A
	DRY-N1	6.80-7.20	Absent in only Cluster B
	O-N1	6.40-6.80	Absent in Cluster A & Cluster B
	N1-TIP	12.00-12.40	Absent in only Cluster B

A plot between 1st and 2nd principle component in figure 4.2 shows that a hydrogen bond donor group is absent within respective chemical scaffolds of compound in cluster A and therefore, the feature distances ranges defined by 1st PC on the basis of DRY-O, O-TIP O-O and O-N1 correlogram in figure 4.2 A are absent in this cluster of the data set. Similarly, a hydrogen bond acceptor feature is absent in all compounds encircled as cluster B in figure 4.2 and therefore, a reference point for mapping the distances of characterized 3D structural features defined by DRY-N1, N1-TIP N1-N1 and O-N1 correlogram is absent in cluster B. Nevertheless, rest of the data exhibit one to two hydrogen bond acceptor as well as donor group within respective chemical scaffold thus, are defined by the presence of all O-O, DRY-O, O-TIP, DRY-N1, N1-TIP, N1-N1 and O-N1 variables depicted by first two PCs as shown in figure 4.2.

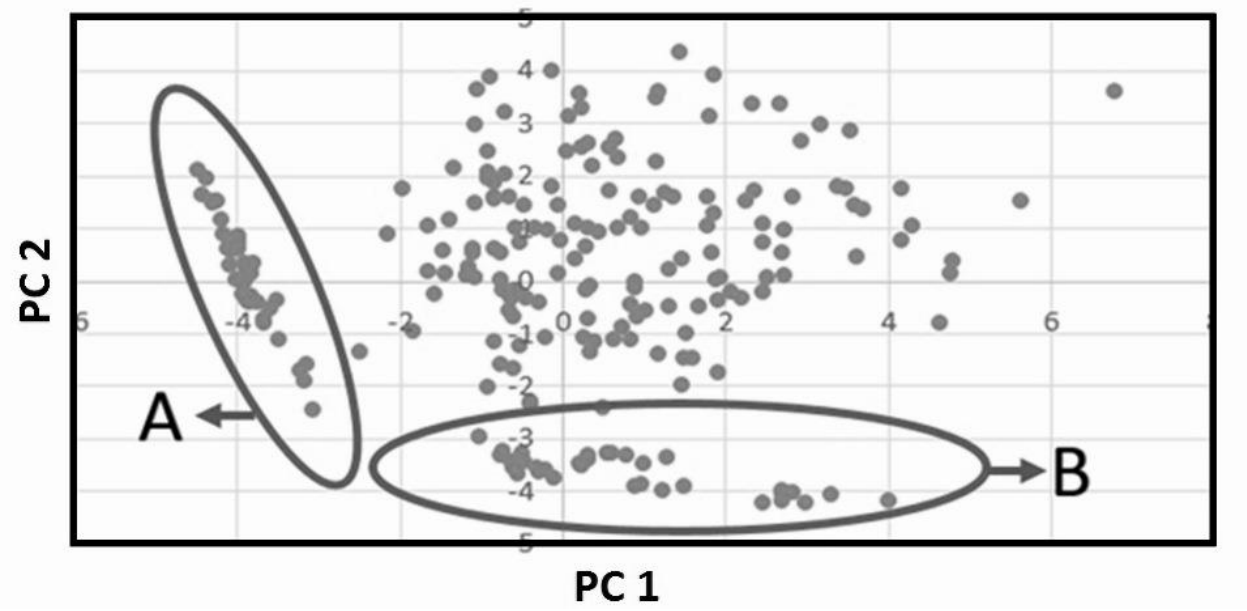

Figure 4.2: Plot produced by PC 1 and PC 2 showing Cluster A and B distinctively.

Furthermore, the impact of presence or absence of hydrogen bond acceptor and donor groups at specific distance from rest of the pharmacophoric features on blood brain barrier has been elucidated by Partial Least Square (PLS) Analysis.

4.2 Partial Least Square (PLS) Analysis

Partial Least Square (PLS) Analysis using Leave One Out (LOO) cross validation of the complete set of GRIND variables revealed a model with suboptimal statistical parameters of q2= 0.46, R2= 0.59 and SDEP=0.52. This is might be due to the presence of some inconsistent set of variables as defined by pastor *et al* (Pastor, Cruciani, McLay, Pickett, & Clementi, 2000). Therefore, a variable selection algorithm known as FFD (Fractional factorial design) was applied to remove inconsistent variable (Gunst & Mason, 2009). An improvement in the statistical parameters of different GRIND models after subsequent 1st and 2nd FFD run has been observed as shown in table 4.2. A final GRIND model was obtained after the 2nd FFD cycle with $q^2 = 0.50$, $r^2 = 0.63$ and standard error of prediction (SDEP) of 0.49 as shown in table 4.2.

Table 4.2 Blood brain barrier permeability model statistics after subsequent 1st, 2nd and 3rd FFD cycles.

FFD Cycle	Variables #	Q^2	R^2	SDEP
0	Complete	0.46	0.59	0.52
1st	440	0.48	0.61	0.51
2nd	408	0.50	0.63	0.50

The plot of observed versus predicted log BB values obtained from multiple linear regression analysis using leave one out (LOO) cross validation is shown in figure 4.3. It has been observed that almost all compounds in the training as well as test set (44 compounds) are predicted well with a difference of less than 1.5 log unit between actual and predicted log BB as shown in SM table 1.

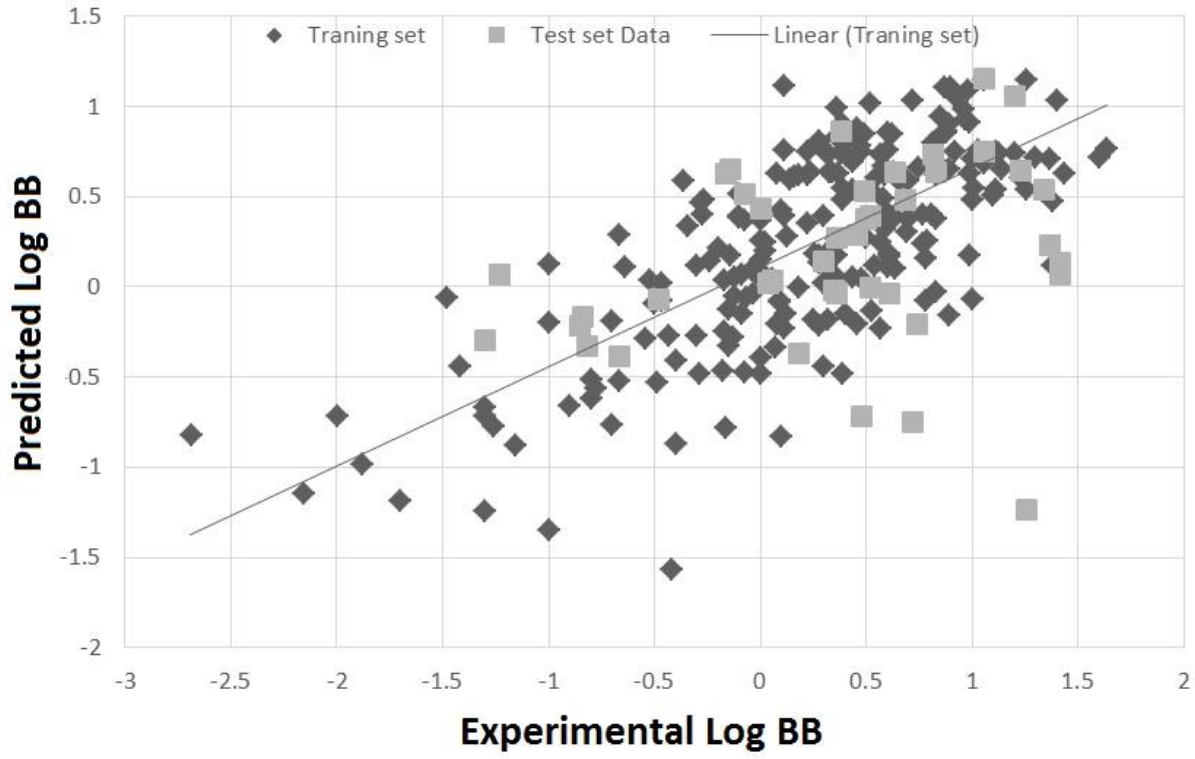

Figure 4.3: Plot of experimental versus predicted log BB values obtained from multiple Linear regression model.

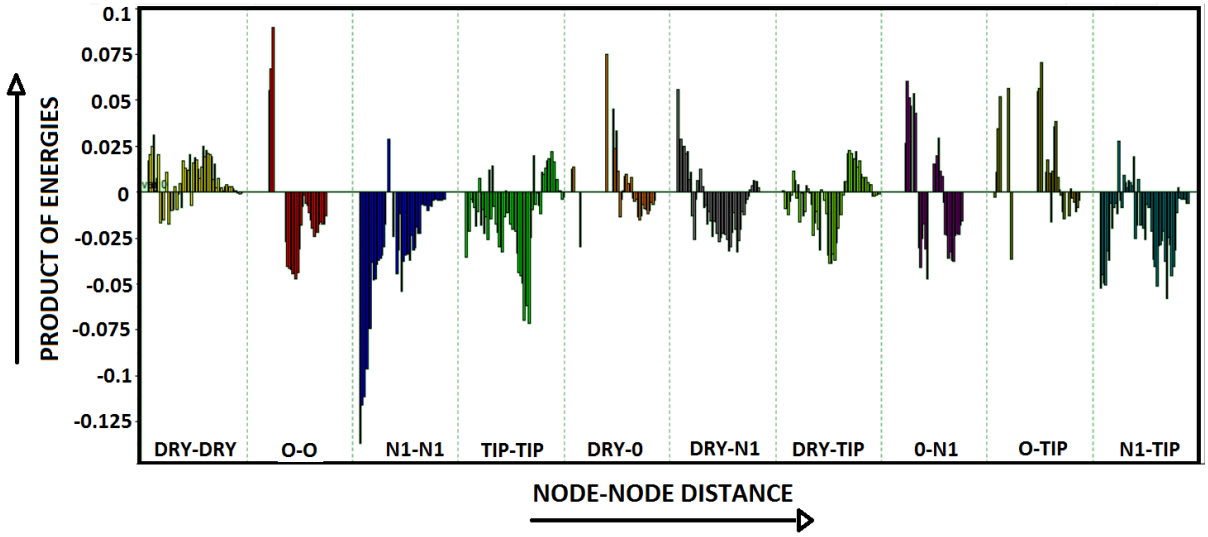

Figure 4.4: PLS coefficient correlograms plot representing the influence of 3D structural features on Log BB.

Figure 4.4 represent a PLS coefficient correlograms plot where positive and negative auto and cross correlogram peaks represent the 3D structural features having direct and inverse correlation with blood brain barrier permeability. Highly negative variable values of N1-N1, TIP-TIP, and N1-TIP correlograms in figure 4.4 represent the 3D features that are present in non-permeable compounds. However, highly positive O-O, DRY-O, O-N1 and O-TIP correlograms variables depict the 3D structural features at certain mutual distance having positive impact on blood brain barrier permeability.

Briefly, the most positive O-O correlogram depicts the presence of two hydrogen bond donor groups (Don1 and Don2) at a mutual distance of 6.00 - 6.40 Å in highly permeable compounds having log BB values range 0.32 to 1.40 as shown in figure 4.5. Within our training data, two hydrogen bond donor groups at a mutual distance of 6.00 - 6.40 Å have been observed in extended 3D conformations of various antipsychotic agents including nebivolol, bupropion, Bromperidol, Zanapezil, Tramadol, Oxazepam, Lubeluzole, biperiden and Donepezil exhibiting log BB of 0.48 to 1.44. However, a decrease in blood brain barrier permeability (log BB) has been observed as the distance between two hydrogen bond donor group increases.

For instance, the two features, are present at a longer distance range of 11.6-12.0 Å in least permeable compounds (log BB: 0.30 to -2.69) including 9-hydroxy risperidone, SKF89124 (7-hydroxy ropinirole) and SKF 93319 as shown in figure 4.5 and thus, are absent CNS active agents in our training data. Interestingly, O-N1 correlogram variables complement the O-O correlogram and represent a hydrogen bond donor group (Don1) at a distance of 6.0 to 6.4 Ǻ from a hydrogen bond acceptor group in highly permeable compounds having a log BB range of 0.39 to 1.64 Ǻ.

Similarly, the DRY-O correlogram shown in the figure 4.4 illustrates the presence of hydrophobic group at a mutual distance of 10.0–10.4 Ǻ from a hydrogen bond donor group (Don1) in only highly permeable compounds having log BB range between 0.30 to 1.6 including Sertraline, Desipramine, Tamozifin and Methadone. Thus, it highlights the influence of hydrophobic feature of the molecules on blood brain barrier permeability which is in accordance with previous study conducted by Gulyaeva *et al*, in which relative hydrophobicity and lipophilicity of drugs on log BB was measured by aqueous two-phase partitioning, octanol-buffer partitioning and HPLC (Gulyaeva et al., 2003). Nevertheless, the most positive variable in O-TIP correlogram depicts the presence of a hydrogen bond donor group (Don1) at a mutual distance of 4.80 to 5.20 Ǻ from molecular steric hotspots in highly permeable compounds exhibiting a log BB range of 0.28 to 1.60 Ǻ.

Overall our PLS model elucidated the importance of one hydrogen bond donor group represented by Don1 as shown in figure 4.5 that may act as an anchor for mapping the distances of other pharmacophoric features. Moreover, PC1 also separated our training data on the basis of presence/absence of a hydrogen bond donor reference feature. Therefore, for the compounds of cluster A, distance between 3D structural features defined by O-O, DRY-O, O-N1 and O-TIP correlogram could not be mapped mainly due to absence of a hydrogen bond donor reference point in their respective chemical scaffolds.

Moreover, the influence of hydrogen bond donor group on BBB permeability of the drug like compounds have been demonstrated in previous studies by considering descriptors like number of hydrogen bond donors/acceptors present within molecules and polarity (M. Abraham, Chadha, & Mitchell, 1995; Geldenhuys, Mohammad, Adkins, & Lockman, 2015; Kansy, Senner, & Gubernator, 1998). These studies associate higher blood brain barrier permeability with lower counts of hydrogen bond donors and acceptors within a molecule. Moreover, a study conducted by Prashant *et al*, correlates blood brain barrier permeability with the strength of hydrogen bond acceptor or donor groups within a molecule (M. H. Abraham, Chadha, & Mitchell, 1994; Geldenhuys et al., 2015; Katritzky et al., 2006; Kelder, Grootenhuis, Bayada, Delbressine, & Ploemen, 1999). Additionally, van de Waterbeemd *et al* demonstrated the importance of shape of the molecule towards blood brain barrier permeability using molecular size, shape and hydrogen bonding descriptors (van de Waterbeemd, Camenisch, Folkers, Chretien, & Raevsky, 1998). These previous investigations are in line with our study and thus, strengthen our models. Additionally, our models utilizes > 50% of already marketed drugs including CNS active agents which may provide a better scaffold to correlate and understand the 3D structural features important for BBB permeability. Moreover, our models were able to define and map the distances between these 3D structural features.

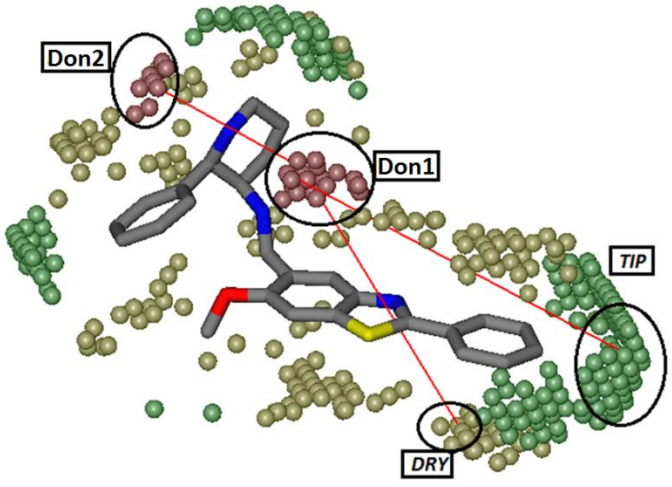

Figure 4.5: Represents a blood brain barrier permeable compound. In this figure blood red contour shows hydrogen bond donor hot-spots, Moss green contour represents hydrophobic hotspots and sea green contour represents molecular edges.

Table 4.3: Summary of GRIND variables, their mutual distance and impact on blood brain barrier permeability.

Correlogram	Distance	Influence on log BB
O-O	6.00-6.40 Å	+
DRY-O	10.00-10.40 Å	+
O-N1	6.00-6.40 Å	+
O-TIP	4.80-5.20 Å	+
N1-N1	1.60-2.00 Å	-

Likewise, N1-N1 correlogram variables in figure 4.4 depict the presence of two hydrogen bond acceptor groups (Acc1, Acc2) at a mutual distance of 1.60 to 2.0 Å in least permeable compounds (-2.6 to -0.18) of the training data as shown in figure 4.6. Similarly, the most negative N1-TIP correlogram variable illustrates a distance of 17.6 to 18 Å between a hydrogen bond acceptor group (Acc1) and a molecular steric hotspot edge in non-permeable compounds of our dataset. Highly negative TIP-TIP correlogram variable depicts the presence of two molecular steric boundaries (TIP1 and TIP2) at a mutual distance of 15.6 to16.0 Å in non-permeable (log BB -0.54 to -2.6) compounds of the dataset as shown in figure 4.6. Thus, it may suggests that as the distance between the molecular edges increase and molecule gets more flexible then the permeability of the molecule decreases which has also been mentioned in a previous study by Pardridge *et al* (Pardridge, 2005)

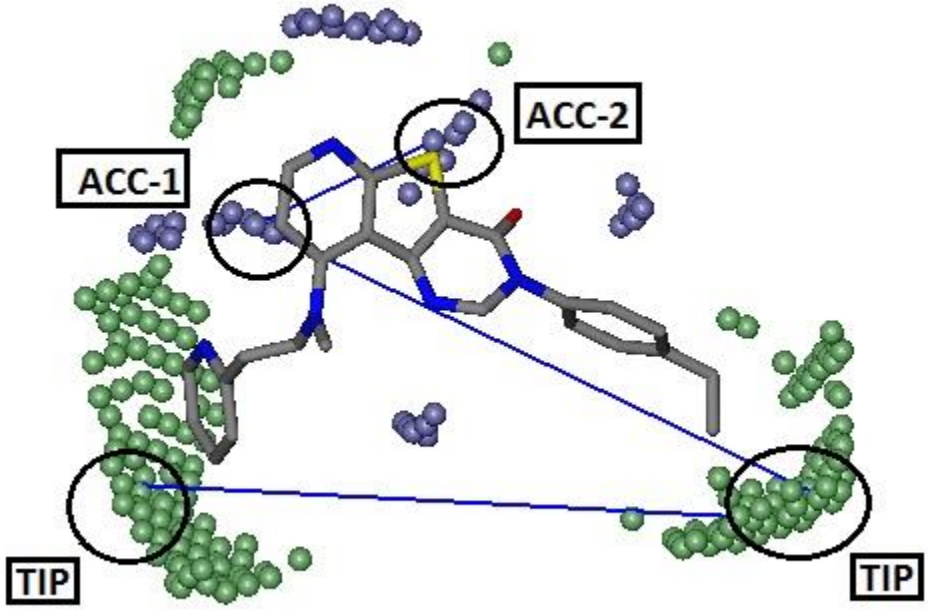

Figure 4.6: Represents a non-permeable compound in which blue contour represents hydrogen bond acceptor hotspots where as sea green contour refers to molecular edges.

Overall, GRIND model suggests that the permeability of the compounds has been influenced by the distance between four 3D structural features such as hydrogen bond donors, hydrogen bond acceptors, shape and hydrophobic features which is in agreement with previously developed QSAR models in which Lipophilicity and molecular volume have been identified as important properties along with hydrogen bonding potential to cross BBB (M. Abraham et al., 1995; Brewster, Pop, Huang, & Bodor, 1996; Clark, 1999; Feher, Sourial, & Schmidt, 2000; Kelder et al., 1999; Subramanian & Kitchen, 2003; Suenderhauf, Hammann, & Huwyler, 2012). Several QSAR models correlates physiochemical descriptors such as lipophilicity, water-accessible volume, topological polar surface area, hydrogen bonds, rotatable bonds, molecular weight, ionization potential, charge, among many others (Gao, Chen, Cai, & Xu, 2017). These descriptors reflects the presence of certain functional groups within the molecules (Geldenhuys et al., 2015). However, in our study, the effect of 3D structural features has been found to be dependent on the distances present between these features within the extended conformations of the molecules.

The presence of two hydrogen bond donor will show positive behavior towards BBB permeability only if they are present within a molecule at 6.00 - 6.40 Å distance from each other. However, hydrogen bond donor groups present at a mutual distance of 11.6-12.0 Å will show a negative effect on blood brain barrier permeability of compounds. Similarly, a molecule will exhibit permeability, if the hydrogen bond acceptor groups such as carbonyl oxygen is present at distance of 6.00 - 6.40 Å from hydrogen bond donor group with a molecule whereas, a negative effect will be observed for a mutual distance of 10.8-11.2 Å. The aromatic moieties present at distance of 10.0-10.4 Å from the donor groups within molecules will contribute positively towards BBB permeability of the molecule while a negative effect on BBB permeability of the molecules will be observed for a mutual distance of 20.0-20.4 Å. These facts signify the important role of relative distance between 3D structural features for BBB permeability of compounds. Moreover, a mutual distance of 4.80-5.20 Å between hydrogen bond donor and molecular edges will act as a positive contributor towards BBB permeability in our model. It is very interesting to note that protonated nitrogen in straight chain has been found as a common hydrogen bond donor in highly permeable compounds as shown in the figure 4.5.

Conclusion:

In the light of the predictions made by 3D QSAR model using GRID independent descriptors (GRIND) it has been demonstrated that the relative distance between important pharmacophoric features such as hydrogen bond donors/acceptors, hydrophobicity and shape based features may affect BBB permeability of a compound. Optimum distance of a hydrogen bond donor pharmacophore from the other features which include hydrophobic group, another donor group and molecular edges, determines the fate of molecule to cross the BBB. According to our model, two hydrogen bond donor groups are present at a mutual distance of 6.00 to 6.40 Å in CNS active agents having high log BB values. One of the hydrogen bond donor group, at a distance of 10Å-10.4Å from a hydrophobic group and at a distance of 6.00 - 6.40 Å from a hydrogen bond acceptor functional group also represent important attributes for blood brain barrier permeability. Similarly the optimum distance between the reference hydrogen bond donor group and molecular edges should be in a range of 4.8Å- 5.2Å for BBB permeability as suggested by our GRIND model. Overall, our GRIND model not only highlights the important 3D structural features of a molecule for BBB penetration but also maps the relative distance between the functional groups optimum for BBB permeability. Since the descriptors used to build this model, are highly relevant to biological activity, alignment free and easy to obtain therefore, this model can prove to be useful in early stages of drug development for prediction of blood brain barrier permeability of new chemical entities.

Annex 1.

Training Dataset

NAME OF COMPOUND	*log BBB*	*Predicted Log BB*	*M. Weight*	*Log P (o/w)*	*TPSA*	*Reference*
Org12962	1.64	0.64	265.67	2.18	28.16	14
sertraline	1.60	1.02	306.24	5.46	12.03	13
Trifluopromazine	1.44	0.59	352.42	4.97	6.48	11
bupropion	1.40	0.56	239.75	3.54	29.10	11
1-isobutyl-4-(4-iodophenyl)piperazine	1.38	0.87	344.24	3.85	6.48	3
bromperidol	1.38	-0.16	420.32	4.80	40.54	5
nor-1-chlorpromazine	1.37	1.08	304.85	4.36	15.27	5
Maprotiline	1.30	1.09	277.41	5.26	12.03	10
diphenhydramine	1.26	0.47	255.36	3.78	12.47	11
Pfizer Compound 23	1.26	0.37	429.59	6.14	46.18	3
promazine	1.23	0.96	284.43	3.99	6.48	4
desipramine	1.20	1.07	266.39	3.84	15.27	4
zanapezil	1.14	0.21	376.54	5.12	32.34	13

I-4 (44cyanophenoxymethylpiperidineniodoallyl)	1.13	0.47	382.25	3.59	36.26	3
dizocilpine (MK801)	1.11	0.44	221.30	3.85	12.03	9
CP-122721	1.11	0.97	380.41	5.25	42.52	13
Astemizole	1.10	0.00	458.58	5.99	42.32	10
desmethyldesipramine (didesipramine)	1.06	1.01	252.36	3.31	29.26	4
chlorpromazine	1.06	1.18	318.87	4.62	6.48	3
Org5222 (asenapine)	1.03	0.63	285.77	3.55	12.47	14
1-hexyl-4-(4-iodophenyl)piperazine	1.01	0.79	372.29	4.82	6.48	3
phenserine	1.00	0.59	337.42	3.37	44.81	10
Doxepin	1.00	0.34	279.38	4.17	12.47	4
tertbutylchlorambucil	1.00	-0.04	360.32	4.70	29.54	4
mianserin	0.99	0.04	264.37	3.33	6.48	9
Citalopram	0.99	0.95	324.40	4.40	36.26	9
Trimipramine	0.99	1.04	294.44	4.46	6.48	14
Pfizer Compound 6	0.98	0.70	352.52	5.35	33.29	11
nor-2-chlorpromazine (didemethylchlorpromazine)	0.97	1.17	290.82	3.83	29.26	5
Pfizer Compound 9	0.96	0.77	406.49	6.79	33.29	11

Pfizer Compound 8	0.92	0.68	338.50	4.99	33.29	11
tamoxifen	0.92	0.61	371.52	6.70	12.47	6
Methadone	0.90	0.54	309.45	5.24	20.31	10
10e	0.89	0.50	327.43	3.76	29.85	7
donezepil (donepezil)	0.89	0.96	379.50	4.15	38.77	4
amitriptyline	0.89	-0.17	277.41	4.56	3.24	13
rivastigmine	0.88	0.69	250.34	2.57	32.78	11
Pfizer Compound 10	0.88	0.53	416.39	5.86	42.52	13
Pfizer Compound 5	0.87	0.92	296.41	3.81	33.29	11
Meperidine	0.85	0.59	247.34	3.04	29.54	9
biperiden	0.85	0.52	311.47	3.87	23.47	11
Pfizer Compound 21	0.85	0.44	409.60	5.68	46.18	3
imipramine	0.83	-0.04	280.42	4.10	6.48	9
butaperazine (tyrylen)	0.83	0.95	409.60	4.33	26.79	13
MCHR antagonist 19	0.83	-0.06	384.87	4.64	62.41	14
beloxepin (Org4428)	0.82	0.52	295.38	3.20	32.70	14
MCHR antagonist 23c	0.81	0.10	488.68	4.70	47.10	9

fluvoxamine	0.79	0.15	318.34	3.23	56.84	13
sabeluzole	0.78	-0.31	415.53	4.13	48.83	9
MCHR antagonist 7c	0.78	-0.28	482.00	4.85	54.37	2
9b	0.77	0.14	311.36	3.12	39.08	7
loperamide	0.77	0.12	477.05	6.35	43.78	13
northioridazine	0.75	0.97	356.56	5.18	15.27	3
Tramadol	0.72	0.68	263.38	3.14	32.70	12
Clozapine	0.70	0.17	326.83	3.08	30.87	10
Perphenazine	0.70	0.18	403.98	3.45	29.95	10
BBcpd23 (ranitidine analog) SKB36 (2-pyridine-amine-n-33-n-piperidinyl-methylphenoxypropyl)	0.69	0.11	325.46	3.80	37.39	4
Budipine	0.66	0.93	293.45	5.92	3.24	9
doxylamine	0.64	0.36	270.38	2.92	25.36	3
terfenadine	0.64	-0.11	471.69	7.65	43.70	13
mefloquine	0.63	0.41	378.32	4.27	45.15	3
Pfizer Compound 24	0.62	0.40	393.55	4.83	46.18	11
rolipram	0.61	0.03	275.35	2.12	47.56	4
oxazepam	0.61	0.22	286.72	2.42	61.69	3

cocaine	0.60	0.69	303.36	2.42	55.84	9
5-Hydroxytryptamine (5-HT) receptor ligand 1b	0.60	0.49	259.69	2.84	36.42	10
amoxapine	0.60	0.17	313.79	3.04	36.86	10
Bremazocine	0.60	0.24	315.46	3.77	43.70	3
desmonomethylpromazine (monodesmethylpromazine)	0.59	1.02	270.40	3.73	15.27	3
lubeluzole	0.59	-0.23	433.52	4.32	48.83	2
S-1-Fluorocarazolol	0.58	0.85	316.38	3.27	57.28	9
TZ-19	0.58	0.36	397.90	3.75	36.42	11
fentanyl	0.58	1.02	336.48	4.28	23.55	3
nebivolol	0.57	-0.09	405.44	3.21	70.95	9
MCHR antagonist 28	0.57	0.16	380.45	4.00	71.64	2
triazolam	0.57	-0.08	343.22	4.59	43.07	3
AFE	0.55	0.94	290.41	3.80	29.26	3
pentazocine	0.54	0.42	285.43	3.72	23.47	3
mirtazapine	0.53	0.30	265.36	2.50	19.37	13
2-hydroxydesipramine	0.53	0.35	282.39	3.57	35.50	14
Org32104 (CID 22154175)	0.52	0.60	281.35	2.94	41.49	14

desmethyldiazepam (Nordazepam)	0.50	0.14	270.72	2.91	41.46	3
mepyramine (pyrilamine)	0.49	0.52	285.39	2.62	28.60	4
lamotrigine	0.48	0.01	256.10	2.49	90.71	11
Pfizer Compound 1	0.48	0.28	454.66	6.93	24.50	3
TZ-18	0.47	1.08	264.54	3.13	36.42	11
MSP (N-methylspiroperidol)	0.46	-0.27	409.51	3.53	43.86	3
Pfizer Compound 2	0.46	0.27	468.69	7.28	24.50	11
lorazepam	0.44	0.33	321.16	3.01	61.69	4
BBcpd24 (ranitidine analog) SKB37	0.44	-0.02	331.48	3.61	37.39	3
phencyclidine	0.44	0.95	243.39	4.97	3.24	3
TZ-21	0.41	0.57	308.99	3.34	36.42	11
Pfizer Compound 11	0.41	1.03	408.46	6.05	42.52	11
tibolone	0.40	-0.07	312.45	2.97	37.30	14
betaxolol	0.39	0.58	307.43	2.70	50.72	7
hydroxyzine	0.39	0.01	374.91	3.20	35.94	14
Org30526	0.39	-0.27	271.75	3.28	21.26	4
4g	0.39	0.91	309.39	3.91	29.85	4

TZ-11	0.38	1.03	230.10	2.51	36.42	11
Pfizer Compound 4	0.37	0.34	412.58	5.75	24.50	11
desmethylclobazam (norclobazam)	0.36	0.16	286.72	2.45	49.41	7
9a	0.36	0.07	307.40	3.27	39.08	11
midazolam	0.36	0.76	325.77	4.14	30.18	3
Pfizer Compound 15	0.36	0.15	379.53	4.46	46.18	3
fluoxetine	0.36	0.49	309.33	4.76	21.26	4
clobazam	0.35	0.11	300.75	2.65	40.62	3
lidocaine	0.34	0.41	234.34	2.22	32.34	3
TZ-2	0.33	0.95	319.00	2.88	36.42	11
galanthamine	0.32	-0.09	287.36	2.49	41.93	7
1	0.32	0.56	389.48	4.44	38.88	11
Pfizer Compound 22	0.32	0.35	395.57	5.28	46.18	6
AR 1D9859	0.31	0.78	275.39	2.87	21.70	6
Trazodone	0.30	-0.15	371.87	2.67	42.39	10
Diltiazem	0.30	0.90	414.53	3.26	59.08	4
pergolide	0.30	0.09	314.50	4.18	19.03	10

SB222200	0.30	-0.52	380.49	6.48	41.99	5
toliprolol	0.28	0.15	223.32	2.04	41.49	9
loreclezole	0.28	0.10	274.54	2.67	30.71	2
diazepam	0.28	-0.03	284.75	3.11	32.67	3
Ondansetron	0.28	0.63	292.38	3.14	34.36	2
R82913	0.28	0.54	321.88	3.18	50.60	3
spiperone	0.26	-0.02	395.48	3.33	52.65	3
ropinirole (SKF101468)	0.25	0.23	260.38	3.02	32.34	3
thioridazine	0.24	0.93	370.58	5.72	6.48	4
Dextromethorphan	0.22	0.19	271.40	3.56	12.47	4
BBcpd26 (ranitidine analog) SKB42	0.22	0.59	365.48	4.44	50.53	9
talsupram	0.22	1.09	311.49	6.13	12.03	13
sulforidazine	0.18	0.06	402.58	4.32	40.62	6
CID 10451635 (2-(3'-Iodo-4'-aminophenyl)-6-hydroxybenzothiazole)	0.18	0.63	368.20	4.44	59.14	3
ST 363	0.16	1.00	230.10	2.51	36.42	6
zolantidine (SKB41)	0.14	0.16	381.54	5.25	37.39	4
MCHR antagonist 2'	0.13	0.23	394.90	5.49	41.57	9

pentobarbital (5-Ethyl-5-(1-methylbutyl)barbital)	0.12	-0.18	226.28	2.13	75.27	4
amobarbital	0.11	-0.19	226.28	2.13	75.27	9
Dehydroevodiamine	0.11	0.99	304.37	2.35	35.35	11
clonidine	0.11	0.28	230.10	2.47	36.42	3
Pfizer Compound 20	0.11	0.52	390.53	4.03	51.11	4
Naproxen	0.10	-0.32	230.26	3.29	46.53	10
quinidine (quinine)	0.10	0.22	323.44	4.23	42.35	3
5-pentyl-5-ethyl barbituric acid	0.09	-0.12	226.28	2.21	75.27	3
MPPF	0.09	-0.12	434.51	3.39	48.91	9
Nimodipine	0.08	-0.14	418.45	3.08	119.68	9
terbinafine	0.08	0.68	291.44	5.89	3.24	13
BSP (p-bromospiperone)	0.07	-0.27	474.37	4.13	52.65	3
flunitrazepam	0.06	0.17	313.29	2.60	78.49	3
1-Butyl-3-phenylthiourea	0.03	0.82	208.33	2.84	56.15	9
Oxprenolol	0.02	0.23	265.35	2.07	50.72	12
DM5	0.00	0.28	293.75	2.79	55.40	8
Testosterone	0.00	-0.25	288.43	3.10	37.30	9

levorphanol (racemorphan)	0.00	0.52	257.38	3.29	23.47	14
Warfarin	0.00	0.13	308.33	3.43	63.60	10
Org34167 (CID 9861160)	0.00	0.21	264.33	3.71	52.05	10
1 del tercero	0.00	-0.15	314.47	5.51	29.46	5
BBcpd22 (ranitidine analog) SKB34 (propanol33npiperidinylmethylphenoxy)	-0.02	0.27	249.35	2.67	32.70	4
4-hydroxymidazolam	-0.03	-0.04	341.77	3.65	50.41	3
SB-656104-A	-0.04	0.01	488.05	4.88	65.64	16
atropine	-0.06	0.16	289.38	2.40	49.77	5
(+)-Tetrahydropalmatine	-0.07	-0.28	355.43	3.07	40.16	9
1-hydroxymidazolam	-0.07	0.15	343.79	3.66	47.86	3
KLP-440	-0.09	0.04	294.35	2.69	52.57	9
bretazenil	-0.09	0.10	418.29	3.52	64.43	3
phenytoin	-0.10	0.41	252.27	2.45	58.20	3
4-iodoantipyrine (iodophenazone)	-0.10	0.06	314.13	2.62	23.55	9
Rofecoxib	-0.10	0.34	314.36	3.09	60.44	3
BBcpd14 (cimetidine derivative)	-0.12	-0.06	368.46	3.94	86.53	4
bisphenol A	-0.12	0.31	228.29	4.71	40.46	3

5	-0.13	-0.21	391.56	4.93	28.48	8
MCHR antagonist 8	-0.14	0.16	410.90	4.75	50.80	9
Dexamethasone	-0.15	-0.18	392.47	2.15	94.83	9
daidzein	-0.15	-0.24	254.24	2.19	66.76	3
vorozole	-0.16	-0.57	324.77	2.07	61.42	2
1-(8-fluorooctyl)-2-nitroimidazole (FON) (MOLI001285)	-0.17	0.36	243.28	3.10	63.64	3
DM49	-0.17	-0.06	343.30	3.64	64.63	11
ibuprofen	-0.18	0.00	206.28	3.61	37.30	4
Indole-based MCHr1 antagonist(7a)	-0.20	-0.23	454.57	4.55	59.39	9
BBcpd21 (ranitidine analog) SKB31 (benzamiden33piperidinylmethylphenoxy)	-0.24	0.01	352.48	4.28	41.57	4
alaptide (MP 005-942-485)	-0.26	0.09	362.45	4.46	41.05	3
BBcpd18 (ranitidine analog)	-0.27	0.48	312.33	3.40	90.02	4
BBcpd19 (ranitidine analog) SKB29	-0.28	0.32	337.38	3.37	89.77	4
Campthothecin	-0.29	-0.37	348.36	2.30	79.73	9
L663581	-0.30	0.29	357.80	2.45	77.05	3
16alpha-fluoroestradiol (estradiol16afluoro)	-0.30	-0.22	290.38	3.65	40.46	4
carbamazepine epoxide	-0.34	0.26	252.27	2.13	58.86	3

mesoridazine	-0.36	0.15	386.58	3.73	42.76	5
Probenecid	-0.40	-0.23	285.36	2.24	74.68	10
risperidone	-0.40	-0.30	410.49	2.80	61.94	2
Topotecan	-0.42	-0.85	421.45	2.04	103.20	9
SKF89124 (7-hydroxy ropinirole)	-0.43	0.06	276.38	2.71	52.57	3
BBcpd20 (ranitidine analog) SKB30 (npropylacetamide33npiperidinylmethylphenoxy)	-0.46	0.11	290.41	2.62	41.57	4
Compound 4a	-0.47	-0.19	350.45	3.37	48.80	9
tretinoin	-0.49	-0.12	300.44	4.63	37.30	6
Gefitinib	-0.50	-0.22	446.91	3.15	68.74	9
phenylbutazone	-0.52	0.05	308.38	4.08	40.62	3
zolpidem	-0.54	-0.11	309.41	2.70	35.58	13
norverapamil	-0.64	-0.16	440.58	5.08	72.74	5
9-hydroxy risperidone	-0.67	-0.39	426.49	2.14	82.17	3
BBcpd12 (cimetidine derivative) (SKB15)	-0.67	0.24	357.23	2.69	86.53	4
Tolbutamide	-0.70	-0.44	270.35	2.09	75.27	10
levocabastine hydrochloride	-0.70	0.02	420.53	6.19	64.33	2
N-methylthiazolium-DCKA	-0.78	-0.32	398.29	4.20	59.28	9

Colchicine	-0.80	-0.36	399.44	2.21	83.09	10
Compound 13	-0.80	-0.59	441.56	3.98	61.69	9
Glibenclamide	-0.90	-0.52	494.01	4.14	113.60	10
DM44	-1.00	-0.13	304.30	2.17	101.22	11
liarozole	-1.00	0.12	308.77	3.78	46.50	2
ridogrel	-1.00	-0.68	366.34	4.33	71.78	2
29	-1.15	-0.53	473.64	3.24	71.85	8
indometacin (indomethacin)	-1.26	-0.58	357.79	4.37	68.53	4
SKF93319	-1.30	-0.49	448.59	4.00	69.87	9
diclazuril	-1.30	-0.84	407.64	4.48	85.56	4
Cerivastatin	-1.30	-0.53	459.56	4.78	99.88	2
r-etodolac (etodolac)	-1.42	-0.15	287.36	3.74	62.32	3
4-hydroxyalprazolam	-1.48	-0.14	324.77	3.50	63.30	5
chlorambucil	-1.70	-0.41	304.22	3.22	40.54	4
temelastine	-1.88	-0.90	442.36	3.40	79.27	4
icotidine	-2.00	-0.82	379.46	2.22	88.50	4
cetirizine	-2.15	-0.84	388.89	3.40	53.01	3
Sulphasalazine	-2.69	-0.64	398.40	3.24	141.31	2

Annex 2

Test Dataset

Name	*Exp* ***Log BB***	***Predicted Log BB***	***Molecular weight***	***Log P (o/w)***	***TPSA***	***References***
Acyclovir	-0.836	-0.1621	225.21	-0.062	0	PubChem CID: 2022
Alovudine	0.605	-0.03242	244.22	-0.062	0	PubChem CID: 33039
Amobarbital	0.04	0.018039	226.28	-0.062	0	PubChem CID: 2164
Atenolol	1.42	0.136751	266.34	-0.062	0	PubChem 2249
BCNU	0.52	0.392351	214.05	-0.062	0	PubChem 2578
Bromperidol	1.38	0.079255	420.32	-0.062	0	PubChem 2448
Bunitrolol	0.38	0.802454	248.33	-0.062	0	PubChem CID: 2473
CP-141938	-3.6	0.987173	403.55	-0.062	0	PubChem CID: 10363809
Chlorpromazine	1.06	0.782075	318.87	-0.062	0	PubChem CID: 443037
Cimetidine	1.42	0.073777	252.35	-0.062	0	PubChem 2756
Clobazam	0.35	-0.01937	300.75	-0.062	0	PubChem 2789
Desipramine	1.2	1.01714	266.39	-0.062	0	PubChem CID: 2995
Desmethyldesipramine	1.06	1.0927	252.36	-0.062	0	PubChem CID: 159642
Desmethyldiazepam	0.5	0.376081	270.72	-0.062	0	PubChem 2997
Flunitrazepam	0.06	0.033916	313.29	-0.062	0	PubChem 3380
Granisetron	0.687	0.588365	312.42	-0.062	0	PubChem CID: 3510

Haloperidol	1.34	0.47565	375.87	-0.062	0	PubChem CID: 3559
Ibuprofen	0.18	-0.3667	206.28	-0.062	0	PubChem 3672
Imipramine	0.83	0.657662	280.42	-0.062	0	PubChem CID: 3696
Indomethacin	1.26	-1.2362	357.79	-0.062	0	PubChem CID: 3715
Mesoridazine	0.36	0.305869	386.58	-0.062	0	PubChem 4078
Nor-1-chlorpromazine	1.37	0.228926	334.87	-0.062	0	PubChem CID: 443037
Org4428	0.82	0.699986	295.38	-0.062	0	PubChem CID: 166560
Phenylbutazone	0.52	-0.00377	308.38	-0.062	0	PubChem CID: 4781
Promazine	1.23	0.564733	284.43	-0.062	0	PubChem CID: 2726
Propanolol	0.64	0.779884	259.35	-0.062	0	PubChem 4946
Quinidine	0.46	0.323179	324.42	-0.062	0	PubChem CID: 441074
Ribavirin	-0.668	-0.38574	244.21	-0.062	0	PubChem CID: 37542
SB-222200	0.3	0.141727	380.49	-0.062	0	PubChem CID: 6604009
Stavudine	-0.48	-0.07508	224.22	-0.062	0	PubChem CID: 18283
Sulforidazine	0.18	0.556144	402.58	-0.062	0	PubChem 31765
Triazolam	0.74	-0.20987	343.22	-0.062	0	PubChem 5556
Zalcitabine	-0.85	-0.21703	211.22	-0.062	0	PubChem CID: 24066
Zidovudine	0.72	-0.74865	267.25	-0.062	0	PubChem CID: 35370
buspirone	0.48	-0.14324	385.51	-0.062	0	PubChem CID: 2477
desmethylclobazam	0.36	-0.03257	286.72	-0.062	0	PubChem CID: 89657

didanosine	-1.3	-0.29434	236.23	-0.062	0	PubChem CID: 50599
levorphanol	0	0.420926	257.38	-0.062	0	PubChem CID: 5359272
mepyramine	0.49	0.445086	285.39	-0.062	0	PubChem 4992
primidone	-0.07	0.521221	218.256	-0.062	0	PubChem CID: 4909
ranitidine	-1.23	0.093939	314.404	-0.062	0	PubChem CID: 3001055
thiopental	-0.14	0.650979	242.337	-0.062	0	PubChem CID: 3000715
thioperamide	-0.16	0.673994	292.445	-0.062	0	PubChem CID:3035905
tiotidine	-0.82	-0.32979	312.414	-0.062	0	PubChem CID:50287

References

- Abbott, N.J., *Blood–brain barrier structure and function and the challenges for CNS drug delivery.* Journal of inherited metabolic disease, 2013. **36**(3): p. 437-449.

- Ballabh, P., A. Braun, and M. Nedergaard, *The blood–brain barrier: an overview: structure, regulation, and clinical implications.* Neurobiology of disease, 2004. **16**(1): p. 1-13.

- Gloor, S.M., et al., *Molecular and cellular permeability control at the blood–brain barrier.* Brain research reviews, 2001. **36**(2): p. 258-264.

- Francisco, P.-P., E.-C. Manuel, and G.-M. Xerardo, *Review of Bioinformatics and QSAR Studies of β-Secretase Inhibitors.* Current Bioinformatics, 2011. **6**(1): p. 3-15.

- Pangalos, M.N., L.E. Schechter, and O. Hurko, *Drug development for CNS disorders: strategies for balancing risk and reducing attrition.* Nature Reviews Drug Discovery, 2007. **6**: p. 521.

- Carpenter, Timothy S., et al., *A Method to Predict Blood-Brain Barrier Permeability of Drug-Like Compounds Using Molecular Dynamics Simulations.* Biophysical Journal, 2014. **107**(3): p. 630-641.

- Masungi, C., et al., *Parallel artificial membrane permeability assay (PAMPA) combined with a 10-day multiscreen Caco-2 cell culture as a tool for assessing new drug candidates.* Die Pharmazie-An International Journal of Pharmaceutical Sciences, 2008. **63**(3): p. 194-199.

- Mensch, J., et al., *Evaluation of various PAMPA models to identify the most discriminating method for the prediction of BBB permeability.* European Journal of Pharmaceutics and Biopharmaceutics, 2010. **74**(3): p. 495-502.

- Tsinman, O., et al., *Physicochemical selectivity of the BBB microenvironment governing passive diffusion—matching with a porcine brain lipid extract artificial membrane permeability model.* Pharmaceutical research, 2011. **28**(2): p. 337-363.

- Campbell, S.D., K.J. Regina, and E.D. Kharasch, *Significance of lipid composition in a blood-brain barrier–mimetic pampa assay.* Journal of biomolecular screening, 2014. **19**(3): p. 437-444.

- Könczöl, A.r.d., et al., *Applicability of a blood–brain barrier specific artificial membrane permeability assay at the early stage of natural product-based CNS drug discovery.* Journal of natural products, 2013. **76**(4): p. 655-663.

- Kansy, M., F. Senner, and K. Gubernator, *Physicochemical high throughput screening: parallel artificial membrane permeation assay in the description of passive absorption processes.* J Med Chem, 1998. **41**(7): p. 1007-10.

- Di, L., et al., *High throughput artificial membrane permeability assay for blood–brain barrier.* European journal of medicinal chemistry, 2003. **38**(3): p. 223-232.

- Yang, C.Y., et al., *Immobilized artificial membranes—screens for drug membrane interactions.* Advanced Drug Delivery Reviews, 1997. **23**(1): p. 229-256.

- Liu, X., et al., *Development of a computational approach to predict blood-brain barrier permeability.* Drug metabolism and disposition, 2004. **32**(1): p. 132-139.

- Bujak, R., et al., *Blood–brain barrier permeability mechanisms in view of quantitative structure–activity relationships (QSAR).* Journal of pharmaceutical and biomedical analysis, 2015. **108**: p. 29-37.

- Vilar, S., M. Chakrabarti, and S. Costanzi, *Prediction of passive blood–brain partitioning: straightforward and effective classification models based on in silico derived physicochemical descriptors.* Journal of Molecular Graphics and Modelling, 2010. **28**(8): p. 899-903.

- Katritzky, A.R., et al., *Correlation of blood–brain penetration using structural descriptors.* Bioorganic & medicinal chemistry, 2006. **14**(14): p. 4888-4917.

- Wu, Z.-Y., et al., *Comparison of prediction models for blood brain barrier permeability and analysis of the molecular descriptors.* Die Pharmazie-An International Journal of Pharmaceutical Sciences, 2012. **67**(7): p. 628-634.

- Young, R.C., et al., *Development of a new physicochemical model for brain penetration and its application to the design of centrally acting H2 receptor histamine antagonists.* Journal of medicinal chemistry, 1988. **31**(3): p. 656-671.

- Abraham, M.H., H.S. Chadha, and R.C. Mitchell, *Hydrogen bonding. 33. Factors that influence the distribution of solutes between blood and brain.* Journal of pharmaceutical sciences, 1994. **83**(9): p. 1257-1268.

- Abraham, M., H. Chadha, and R. Mitchell, *Hydrogen-bonding. Part 36. Determination of blood brain distribution using octanol-water partition coefficients.* Drug design and discovery, 1995. **13**(2): p. 123-131.

- Norinder, U., P. Sjöberg, and T. Österberg, *Theoretical calculation and prediction of brain–blood partitioning of organic solutes using MolSurf parametrization and PLS statistics.* Journal of pharmaceutical sciences, 1998. **87**(8): p. 952-959.

- Brewster, M.E., et al., *AM1-based model system for estimation of brain/blood concentration ratios.* International journal of quantum chemistry, 1996. **60**(8): p. 1775-1787.

- Kelder, J., et al., *Polar molecular surface as a dominating determinant for oral absorption and brain penetration of drugs.* Pharmaceutical research, 1999. **16**(10): p. 1514-1519.

- Clark, D.E., *Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 2. Prediction of blood–brain barrier penetration.* Journal of pharmaceutical sciences, 1999. **88**(8): p. 815-821.

- Subramanian, G. and D.B. Kitchen, *Computational models to predict blood–brain barrier permeation and CNS activity.* Journal of computer-aided molecular design, 2003. **17**(10): p. 643-664.

- Luco, J.M., *Prediction of the brain− blood distribution of a large set of drugs from structurally derived descriptors using partial least-squares (PLS) modeling.* Journal of chemical information and computer sciences, 1999. **39**(2): p. 396-404.

- Feher, M., E. Sourial, and J.M. Schmidt, *A simple model for the prediction of blood–brain partitioning.* International journal of pharmaceutics, 2000. **201**(2): p. 239-247.

- Kubinyi, H., *3D QSAR in drug design: volume 1: theory methods and applications*. Vol. 1. 1993: Springer Science & Business Media.

- Kubinyi, H., *QSAR and 3D QSAR in drug design Part 1: methodology.* Drug discovery today, 1997. **2**(11): p. 457-467.

- Verma, J., V.M. Khedkar, and E.C. Coutinho, *3D-QSAR in drug design-a review.* Current topics in medicinal chemistry, 2010. **10**(1): p. 95-115.

- Jagiello, K., et al., *Advantages and limitations of classic and 3D QSAR approaches in nano-QSAR studies based on biological activity of fullerene derivatives.* Journal of Nanoparticle Research, 2016. **18**(9): p. 256.

- F Morales, J., et al., *Current State and Future Perspectives in QSAR Models to Predict Blood-Brain Barrier Penetration in Central Nervous System Drug R&D.* Mini reviews in medicinal chemistry, 2017. **17**(3): p. 247-257.

- Pastor, M., et al., *GRid-INdependent descriptors (GRIND): a novel class of alignment-independent three-dimensional molecular descriptors.* Journal of medicinal chemistry, 2000. **43**(17): p. 3233-3243.

- Daneman, R., & Prat, A. (2015). The Blood–Brain Barrier. Cold Spring Harbor Perspectives in Biology, 7(1), a020412. http://doi.org/10.1101/cshperspect.a020412.

- Larsen, J.M., Martin, D.R., Byrne, M.E. (2014). Recent advances in delivery through the blood–brain barrier. Curr Top Med Chem 14: 1148–1160.

- Ehrlich, P. (1885). Sauerstoff-Bedürfniss des Organismus Eine farbenanalytische Studie. Berlin.

- Goldmann, E.E. (1909). Die äussere und Sekretion des gesunden und kranken Organismus im Lichte der “vitalen Färbung”.Beitr Klin Chirurz. 64:192-265.

- Lewandowsky, M. (1900). Zur Lehre der Zerebrospinal flussigkeit. Z Klin Med 40: 480 – 484.

- Aird, WC. (2007). Phenotypic heterogeneity of the endothelium: I. Structure, function, and mechanisms. Circ Res 100: 158–173.

- Reese TS, Karnovsky MJ. 1967. Fine structural localization of a blood–brain barrier to exogenous peroxidase. J Cell Biol 34: 207–217.

- WestergaardE, Brightman MW. 1973. Transport of proteins across normal cerebral arterioles. J Comp Neurol 152:17– 44.

- Coomber BL, Stewart PA. 1985. Morphometric analysis of CNS microvascular endothelium. Microvasc Res 30: 99– 115.

- Betz AL, Firth JA, Goldstein GW. 1980. Polarity of the blood–brain barrier: Distribution of enzymes between the luminal and abluminal membranes of brain capillary endothelial cells. Brain Res 192: 17–28

- SimsDE. 1986. Thepericyte—Areview.TissueCell18:153– 174.

- Armulik A, Genove ´ G, Betsholtz C. 2011. Pericytes: Developmental, physiological, and pathological perspectives, problems, and promises. Dev Cell 21: 193–215.

- Hall CN, Reynell C, Gesslein B, Hamilton NB, Mishra A, Sutherland BA, O'Farrell FM, Buchan AM, Lauritzen M, Attwell D. 2014. Capillary pericytes regulate cerebral blood flow in health and disease. Nature 508: 55–60.

- Peppiatt CM, HowarthC, Mobbs P, Attwell D. 2006. Bidirectional control of CNS capillary diameter by pericytes. Nature 443: 700–704.

- Gerhardt H, Wolburg H, Redies C. 2000. N-cadherin mediates pericytic–endothelial interaction during brain angiogenesis in the chicken. Dev Dyn 218: 472–479.

- Majesky MW. 2007. Developmental basis of vascular smooth muscle diversity. Arterioscler Thromb Vasc Biol 27: 1248–1258

- Shepro D, Morel NM. 1993. Pericyte physiology. FASEB J 7: 1031–1038.

- Sorokin L. 2010. The impact of the extracellular matrix on inflammation. Nat Rev Immunol 10: 712–723.

- Abbott NJ, Ronnback L,HanssonE. 2006.Astrocyte–endothelial interactions at the blood–brain barrier. Nat Rev Neurosci 7: 41–53.

- Noell S,Wolburg-BuchholzK,MackAF,BeedleAM,SatzJS, Campbell KP, Wolburg H, Fallier-Becker P. 2011. Evidence fora roleof dystroglycanregulating themembrane architecture of astroglial endfeet. Eur J Neurosci 33: 2179–2186.
- Gordon GR, Howarth C, MacVicar BA. 2011. Bidirectional control of arteriole diameter by astrocytes. Exp Physiol 96: 393–399.

- Janzer RC, Raff MC. 1987. Astrocytes induce blood–brain barrier properties in endothelial cells. Nature 325: 253– 257.
- PolflietMM, Zwijnenburg PJ, van Furth AM,van der Poll T, Do ¨pp EA, Renardel de Lavalette C, van Kesteren-Hendrikx EM, van Rooijen N, Dijkstra CD, van den Berg TK. 2001. Meningeal and perivascular macrophages of the central nervous system play a protective role during bacterial meningitis. J Immunol 167: 4644–4650.
- Williams K, Alvarez X, Lackner AA. 2001. Central nervous system perivascularcells are immunoregulatory cellsthat connect the CNS with the peripheral immune system. Glia 36: 156–164
- Ginhoux F, Greter M, LeboeufM, NandiS, See P, Gokhan S, Mehler MF, Conway SJ, Ng LG, Stanley ER, et al. 2010. Fate mapping analysis reveals that adult microglia derive from primitive macrophages. Science 330: 841–845.
- Persidsky Y, Ghorpade A, Rasmussen J, Limoges J, Liu XJ, Stins M, Fiala M, Way D, Kim KS, Witte MH, et al. 1999. Microglial and astrocyte chemokines regulate monocyte migration through the blood–brain barrier in human immunodeficiency virus-1 encephalitis. Am J Pathol 155: 1599–1611.
- Hudson LC, Bragg DC, Tompkins MB, Meeker RB. 2005. Astrocytes and microglia differentially regulate trafficking of lymphocyte subsets across brain endothelial cells. Brain Res 1058: 148–160.
- Ramos-Cabrer, Pedro & Campos, Francisco. (2013). Liposomes and nanotechnology in drug development: Focus on neurological targets. International journal of nanomedicine. 8. 951-60. 10.2147/IJN.S30721.
- Pardridge,William M. (1998).CNS Drug Design Based on Principles of Blood-Brain Barrier Transport.Journal of Neurochemistry.70(5):1781-92.
- L. STEIN. 1967.Introducing Not-Self. The jounal of analytical psychology. Volume 12, Issue 2. Pages 97–114.
- Akira Tsuji, Ikumi Tamai,1999. Carrier-mediated or specialized transport of drugs across the blood–brain barrier,Advanced Drug Delivery Reviews,Volume 36, Issues 2–3, Pages 277-290,https://doi.org/10.1016/S0169-409X(98)00084-2.
- Duffy KR, Pardridge WM.1987. Blood-brain barrier transcytosis of insulin in developing rabbits.Brain research.420(1):32-8.

- Fishman J. B., Rubin J. B., Handrahan J. V., Connor J. R. and Fine R. E.(1987) Receptor-mediated transcytosis of transferrin across the blood±brain barrier. J. Neurosci. Res. 18, 299±304.
- Banks WA, Kastin AJ, Huang W, Jaspan JB, Maness LM. 1996 Leptin enters the brain by a saturable system independent of insulin. Peptides;17(2):305-11.
- Stahl, P. and Schwartz, A.L. (1986) Receptor-mediated Endocytosis. Journal odf Clinical Investigation. 77:657.
- Moore, M. S., Mahaffey, D. T., Brodsky, F. M., and Anderson, R. G. W. (1987) Science. 236, 558-563.
- N K Gonatas, A Stieber, W F Hickey, S H Herbert, J O Gonatas. (1984). Endosomes and Golgi vesicles in adsorptive and fluid phase endocytosis.Journal of Cell Biology. 99 (4): 1379.
- Burton, P. S., Conradi, R.A., Hilger, A.R. and Ho, N.F.H. (1993). Evidence for polarized efflux system for peptides in the apical membrane of Coca-2 cells. Biochem. Biophys. Res. Comm. 190:760-766.
- Cory Kalvass, J., & Maurer, T. S. (2002). Influence of nonspecific brain and plasma binding on CNS exposure: implications for rational drug discovery. Biopharmaceutics & drug disposition, 23(8), 327-338.
- Bickel, U. (2005). How to Measure Drug Transport across the Blood-Brain Barrier. *NeuroRx*, *2*(1), 15–26.
- Kennedy T. Managing the drug discovery/development interface. Drug Discovery Today 2:436–444, 1997.

- Deguchi Y, Morimoto K. Application of an in vivo brain microdialysis technique to studies of drug transport across the bloodbrain barrier. Curr Drug Metab 2:411–423, 2001.
- Sokoloff L, Reivich M, Kennedy C, Des Rosiers MH, Patlak CS, Pettigrew KD, et al. The [14C]deoxyglucose method for the measurement of local cerebral glucose utilization: theory, procedure, and normal values in the conscious and anesthetized albino rat. J Neurochem 28:897–916, 1977.

- Pardridge WM. Blood-brain barrier carrier-mediated transport and brain metabolism of amino acids. Neurochem Res 23:635– 644, 1998.
- Gumbleton M, Audus KL. Progress and limitations in the use of in vitro cell cultures to serve as a permeability screen for the blood-brain barrier. J Pharm Sci 90:1681–1698, 2001.
- Audus KL, Rose JM, Wang W, Borchardt RT. Brain microvessel endothelia cell culture systems. In: Introduction to the bloodbrain barrier: methodology, biology and pathology (Pardridge WM, ed), pp 86–93. Cambridge, UK: Cambridge University Press, 1998.

- Oldendorf WH. Lipid solubility and drug penetration of the blood brain barrier. Proc Soc Exp Biol Med 147:813–815, 1974.

- Levin VA. Relationship of octanol/water partition coefficient and molecular weight to rat brain capillary permeability. J Med Chem 23:682–684, 1980.

- Ajay, Bemis GW, Murcko MA. Designing libraries with CNS activity. J Med Chem 42:4942–4951, 1999.

- Engkvist O, Wrede P, Rester U. Prediction of CNS activity of compound libraries using substructure analysis. J Chem Inf Comput Sci 43:155–160, 2003
- Clark DE. Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 2. Prediction of blood-brain barrier penetration. J Pharm Sci 88:815– 821, 1999.

- Clark DE. In silico prediction of blood-brain barrier permeation. Drug Discov Today 8:927–933, 2003.

- Lipinski CA. Drug-like properties and the causes of poor solubility and poor permeability. J Pharmacol Toxicol Methods 44: 235–249, 2000.

- Hansch C, Leo A, Mekapati SB, Kurup A. QSAR and ADME. Bioorg Med Chem 12:3391–3400, 2004.

- Lobell M, Molnar L, Keseru GM. Recent advances in the prediction of blood-brain partitioning from molecular structure. J Pharm Sci 92:360–370, 2003.

- van de Waterbeemd H, Camenisch G, Folkers G, Chretien JR, Raevsky OA. Estimation of blood-brain barrier crossing of drugs using molecular size and shape, and H-bonding descriptors. J Drug Target 6:151–165, 1998.

- Ooms, F., et al., *A simple model to predict blood–brain barrier permeation from 3D molecular fields.* Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, 2002. **1587**(2): p. 118-125.
- Garg, P. and J. Verma, *In Silico Prediction of Blood Brain Barrier Permeability: An Artificial Neural Network Model.* Journal of Chemical Information and Modeling, 2006. **46**(1): p. 289-297.

- Huang, R., et al., *The NCGC pharmaceutical collection: a comprehensive resource of clinically approved drugs enabling repurposing and chemical genomics.* Science translational medicine, 2011. **3**(80): p. 80ps16-80ps16.
- Sadowski, J., *3D structure generation.* Handbook of Chemoinformatics: From Data to Knowledge in 4 Volumes, 2003: p. 231-261.

- Cawley, G.C. and N.L. Talbot, *Efficient leave-one-out cross-validation of kernel fisher discriminant classifiers.* Pattern Recognition, 2003. **36**(11): p. 2585-2592.

- Gunst, R.F. and R.L. Mason, *Fractional factorial design.* Wiley Interdisciplinary Reviews: Computational Statistics, 2009. **1**(2): p. 234-244.

- Geldenhuys, W.J., et al., *Molecular determinants of blood-brain barrier permeation.* Ther Deliv, 2015. **6**(8): p. 961-71.

- Gulyaeva, N., et al., *Relative hydrophobicity and lipophilicity of drugs measured by aqueous two-phase partitioning, octanol-buffer partitioning and HPLC. A simple model for predicting blood–brain distribution.* European journal of medicinal chemistry, 2003. **38**(4): p. 391-396.

- van de Waterbeemd, H., et al., *Estimation of blood-brain barrier crossing of drugs using molecular size and shape, and H-bonding descriptors.* Journal of drug targeting, 1998. **6**(2): p. 151-165.

- Suenderhauf, C., F. Hammann, and J. Huwyler, *Computational prediction of blood-brain barrier permeability using decision tree induction.* Molecules, 2012. **17**(9): p. 10429-10445.

- Pardridge, W.M., *The Blood-Brain Barrier: Bottleneck in Brain Drug Development.* NeuroRX, 2005. **2**(1): p. 3-14.

- Gao, Z., et al., *Predict drug permeability to blood-brain-barrier from clinical phenotypes: drug side effects and drug indications.* Bioinformatics, 2017. **33**(6): p. 901-908