Effect of Formulation Factors on the Skin Penetration of Drugs

ELEFTHERIOS G. SAMARAS

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

First Supervisor: Dr. Taravat Ghafourian

Second Supervisor: Prof. Iain Cumming

Medway School of Pharmacy
Universities of Kent and Greenwich
2/2012

Acknowledgements

I would like to thank my supervisor Dr. Taravat Ghafourian for all the guidance, support and kindness. Many thanks also to my second supervisor Prof. Iain Cumming, and Prof. Alistair Mathie. In addition I would like to thank Professor J.E. Riviere for providing the dataset for the third study (Chapter 5).

Eleftherios G. Samaras

Table of Contents

List of Tables

List of Figures

List of Schemes

List of Conferences Attended

14th UKICRS Annual symposium 2008 on 'Predicting Drug Delivery – Opportunities and Challenges' hosted by Merck Sharp and Dohme UK, Jan 2008

Chemical Computing Group, European User Group Meeting 2008, October 23 – 24 at Welcome Trust Genome Campus Conference Centre, Cambridge, UK

11th Annual Meeting of the Skin Forum which was held in Edinburgh on July 6th and 7th, 2010 at Surgeons' Hall, Edinburgh, EH8 9DW, United Kingdom

12th Annual Meeting of the Skin Forum which was held in Frankfurt on March 28th and 29th, 2011 at Campus Westend, Goethe Universitat Frankfurt, Germany

UK-QSAR Meeting held in Manchester on May 26th 2011 at Paterson Institute for Cancer Research The University of Manchester Wilmslow Road Withington Manchester M20 4BX England

List of Publications

Ghafourian, T., Samaras, E.G., Brooks, J.D., Riviere, J.E., 2010. Modelling the effect of mixture components on permeation through skin. *International Journal of Pharmaceutics* 398, 28-32.

Ghafourian, T., Samaras, E.G., Brooks, J.D., Riviere, J.E., 2010. Validated models for predicting skin penetration from different vehicles. *European Journal of Pharmaceutical Sciences* 41, 612-616.

Favretto, M., Ngassam, L., Jabarri, A., Patel, S., Samaras, E., Pors, K., Kallinteri, P., Casely-Hayford, M.A., 2011. Anthraquinone Bearing Artemisunate Analogues: Anticancer Efficacy and Effects of Encapsulation into Liposomes. *Journal of Controlled Release* (under publication).

Samaras, E.G., Ghafourian, T., 2011. Statistical Evaluation of the Effect of Mixture and Experimental Conditions on *In Vitro* Human Skin Penetration - Data from updated EDETOX database. Submitted to the *Journal of Pharmaceutical Research.*

Samaras, E.G., Verheijen, R., Ghafourian, T., 2011. Validated models for estimation of in vitro flux through human skin. External Validation of the Experimental Conditions. Under Preparation.

Abstract

Permeation of chemicals through skin can lead to systemic effects of drugs or toxic exogenous chemicals. Skin permeation of compounds and the effect of mixture ingredients can be related to their chemical structures as expressed by molecular properties. This approach, known as Quantitative Structure-Activity Relationship (QSAR) analysis, is used to estimate the skin absorption of compounds from different mixtures and may aid the understanding of the mechanisms involved.

In this investigation, using datasets of skin absorption of a model penetrant formulated in different solvents and a model penetrant formulated in mixtures containing various penetration enhancers, models were developed that could unravel the connection between structural features of such mixture ingredients and the skin permeation of the permeants. This work was further developed by investigating the combined effect of solvents and the permeant molecules in complex mixtures in order to identify the (permeant-mixture ingredient) interaction parameters as well. In order to obtain more applicable models, compounds were suggested for experimental measurements of skin permeation from the solvent mixtures through the comparison of the applicability domain of the available dataset with a commonly used skin absorption dataset (in water). With the addition of new data, the resulting QSAR models were able to estimate skin absorption of permeants from a complex mixture of solvents with a satisfactory level of accuracy.

QSAR models are needed to predict the effect of the varying experimental conditions along with the effects of formulation components and the permeant molecular properties. For this purpose, the skin permeation database, EDETOX was updated with more recent data from 2001 to 2010 and a large dataset was constructed. This made it possible to investigate the effect of experimental conditions such as Occlusion, Hydration, Infinite/Finite dosing and Donor concentration at a large scale for the first time. In conclusion, the project resulted in well validated universal QSAR models for the prediction of skin absorption of compounds. The most notable predictor variables were donor concentration, wetting level of skin (affected by occlusion, donor volume and pre-hydration), the gap between melting and boiling points of the vehicle, molecular size and lipophilicity and electronic descriptors of the permeants.

1. Introduction

Skin, the largest organ of the human body, is in continuous contact with exogenous molecules in everyday life. Some of these chemicals pass the skin barrier via specific pathways, ending in the blood and hence affecting metabolism and health. Skin's essential role is to protect the body from absorption of exogenous toxic material such as pesticides that target toxicological endpoints and can have local and systemic effects (Nielsen et al., 2004). Skin's protective nature is mainly dependent on the *stratum corneum* (SC), the outermost layer which limits the penetration of toxicants and drugs alike.

Skin absorption of toxic chemicals may be high enough to lead to poisoning. For example, skin exposure to paraquat can cause breathlessness, high fever with interstitial infiltrates in the upper and middle lung fields leading to rapid evolution to pulmonary fibrosis (Parisis et al., 1995) or persistent cholestasis (Bataller et al., 1999). Currently the effect of dermal re-entry of pesticides e.g. when workers re-enter an area previously treated with a pesticide is also a cause for concern (Belsey et al., 2011). The European Commission regulations known as REACH (Registration, Evaluation and Authorisation/restriction of Chemicals) requires extensive risk assessments of all existing chemicals, including exposure via dermal contact (Commission of the European Communities, 2003).

Skin is also the focus of research by drug formulators as a site of drug administration, due to the advantages it may offer over other routes of drug delivery (Barry, 2007) such as topical effect which bypasses the circulatory system having a much faster and targeted effect. Dermatological drug products are either dermal or topical and are intended for a localized pharmacological effect or transdermal drug products intended for the treatment or prevention of a systemic disease (Shah, 1993). Research effort has been invested in dermatological products, especially in transdermals, to enhance drug penetration through the SC in order to achieve the desired pharmacological effects (Cooper, 1985). Systemic drug delivery through skin is non-invasive and it bypasses the

first pass metabolism leading to potentially higher bioavailability in comparison with orally administered drugs which are carried to the liver via the portal vein where they may be metabolised significantly before entering the systemic circulation.

1.1.The Skin Structure

The human skin consists of three tissue layers, epidermis, dermis and subcutaneous fat, and contains appendages (hair follicles, sebaceous glands and sweat glands) that provide aqueous channels into the skin (Figure 1.1). Hairy skin, in contrast to soles and palms, contains hair follicles and sebaceous glands (MacGrath et al., 2010; Barry, 2007; Rosen, 2005).

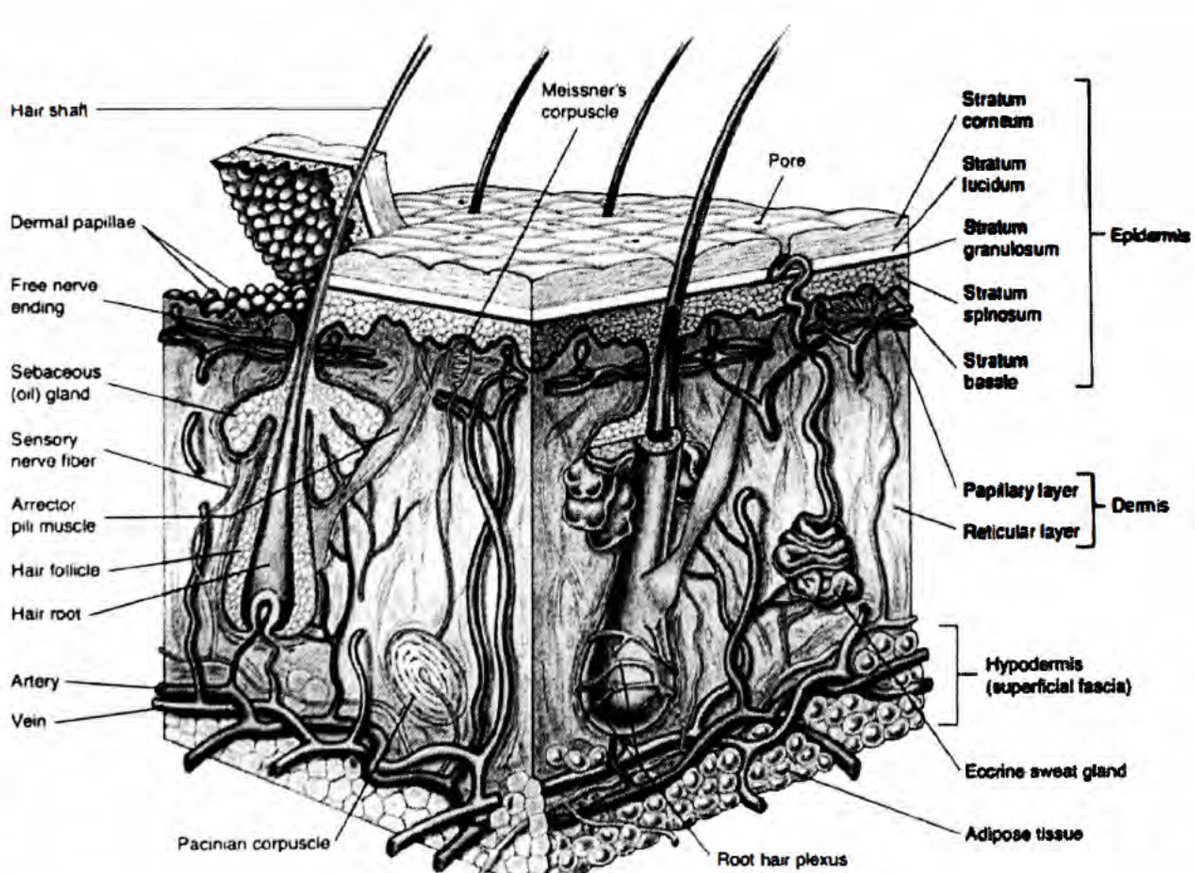


Figure 1.1. A general picture of a section of human skin (http://www.anti-aging-skin-care.com/forever-young-how-it-works-diagram.html).

1.1.1. Epidermis

The epidermis consists of four different layers, the SC, the *stratum granulosum*, the *stratum spinosum*, and the *stratum basale* (McGrath et al., 2010). The cells in the basal layer divide and migrate upwards to produce the SC. More specifically, in the *stratum basale* there is a single layer of more or less cuboidal karatinocytes with the majority of them having cellular projections extending into the dermis and resulting in a highly convoluted epidermal-dermal interface. Some of the basal keratinocytes in the deeper parts have a much flatter interface with the dermis. In the next layer the cells begin to become wider and flatter as they move outward toward the surface of the skin, their spiny appearance has given the name *stratum spinosum* to this layer. The spiny appearance of these cells reflects the presence of many desmosomal connections between cells as well as artifactual shrinkage of the cell bodies that occur during dehydration. In *stratum granulosum* the desmosomes become less prominent, but the cells are characterized by the presence of irregular dense proteinaceous granules known as keratohyalin granules. The end product of the keratinization process, extremely flattened cornified cells, result in the formation of the *stratum corneum* (Rosen, 2005; MacGrath et al., 2010).

The basal layer also includes melanocytes which are responsible for the production and distribution of melanin granules to keratinocytes. The langerhans cells, which are controlled by the immune system and are important in the body's defense mechanism, are also located in the epidermis (McGrath et al., 2010; Barry, 2007). In order for a substance to be absorbed into the body following dermal exposure, first it must be dissolved in the SC (the outermost sub-layer of the skin), and then diffuse through the remaining sub-layers of the epidermis and into the dermis, where it will finally diffuse into the blood capillaries (Barry, 2007).

1.1.1.1. The Stratum Corneum

The first (outermost) layer of the multilayered epidermis, the SC, is the most important layer of the skin in terms of its barrier function. It is almost an impermeable dead dense layer that protects humans from exogenous 'intruders' such as allergens and is responsible for the permeability of human skin to dermally and transdermally delivered drugs. SC also serves as the principle regulatory barrier to the transcutaneous traffic of water and xenobiotics (Hadgraft, 2004), aiding also in temperature regulation and UV protection (Rerek et al., 2005). SC varies in thickness throughout the body, from 0.8 mm on the palms to 0.006 mm on the eyelids. SC can be divided into two types according to thickness and purpose of use, type 1 is represented by the pads of the palms and soles that are adapted especially for friction and weight bearing, type 2 is represented by the remaining flexible and rather impermeable membranous layer (Barry, 2007).

The SC consists of 10-15 layers of vertically stacked corneocytes (cornified epithelial cells) surrounded by the lipid-rich extracellular matrix (Michaels et al., 1975; Wegener et al., 1997). The SC lipids are arranged in multiple bilayers providing alternate hydrophobic and hydrophilic barriers (Williams and Barry, 1991). Composition analysis shows that the SC lipid sheets are comprised mostly of ceramides (Figure 1.2), long chain free fatty acids, and cholesterol, present in approximately equimolar concentrations (Laugel et al., 2005).

More than one third of ceramides have chains longer than 22 carbon atoms and rod-like cylindrical shapes. The long chains of lipids, along with their spatial arrangement, have profound effects on the formation of the highly ordered and relatively impermeable lipid layer in the skin membrane (Forslind et al., 1997). Nine classes of ceramides have been identified and isolated from human SC. These are referred to as ceramides 1-9 (Baroli, 2010). Among these, ceramide 1 has an exceptional molecular structure with linoleic acid linked to a ω-hydroxy fatty acid and is the least polar, acting as a stabilizer of the intercellular lipid lamellae (Bouwstra et al., 2003).

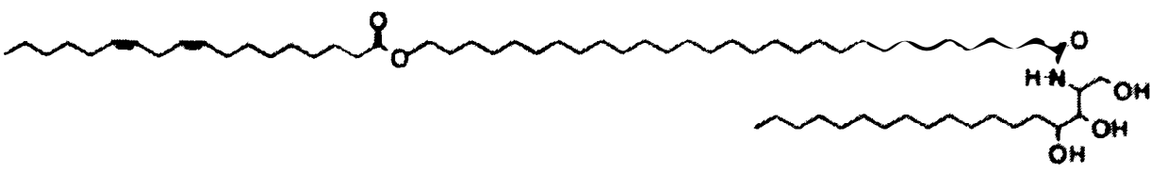


Figure 1.2 Structure of a ceramide

SC can absorb up to 10 times its dry weight when soaked in water (Anigbogu et al., 1995). When hydrated, keratinocytes swell as they absorb water into the intracellular keratin matrix, disrupting the organized layers of the SC. The permeability of SC has been shown to increase rapidly with water uptake, and reach a steady-state of diffusion (Roberts and Walker, 1993).

Skin hydration can be achieved via occlusion with a dressing that captures insensible water loss, thus promoting skin hydration (Riviere et al., 2001; Zhai and Maibach, 2001). Hydration contributes to the mechanism by which transdermal delivery systems or 'patches' achieve acceptable systemic drug concentrations. Hydration of the SC increases its permeability to many compounds. Complete hydration of the barrier, obtained by using an occlusive dressing, drastically increases penetration. Most transdermal delivery devices, by their very nature, function as occlusive systems (Chang and Riviere, 1993).

1.1.2. Dermis

The dermis is 3-5 mm thick and a layer responsible for the skin's structural integrity, elasticity and resilience. It consists of a matrix of connective tissue woven from fibrous proteins (collagen, elastin, and reticulin) that are embedded in an amorphous ground substance of mucopolysaccharide. Collagen is responsible for the structural support and elastin for the structural resilience of the skin (Barry, 2007). The most important type of cells in dermis are the fibroblasts which synthesize collagen, elastin and other structural proteins. Dermis also contains tiny blood vessels and lymph nodes which produce

immune cells, sebaceous glands, sweat glands, hair follicles, nerve and muscle cells. The blood vessel branches deliver blood to dermis, sweat glands, hair follicles, and subcutaneous fat. They are also important to carry toxins or systemic drugs away into the body. This blood supply reaches to within 0.2 mm of the skin surface, thus most compounds that pass the epidermis are quickly absorbed and systematically diluted. In this way the diffusing molecules reach quickly the capillaries, the concentration of penetrants in the dermis is kept very low, the epidermal concentration gradients is maximized, and percutaneous absorption is promoted (Rosen, 2005; McGrath et al., 2010; Hunter et al., 2002).

1.1.3. Subcutaneous tissue (hypodermis)

The subcutaneous tissue consists mainly of fat cells. It acts as shock absorber and heat insulator, protecting underlying tissues from cold and trauma. It provides a mechanical cushion and a thermal barrier; it synthesizes and stores readily available high-energy chemicals (Mills et al., 2006).

1.2. Routes of drug penetration across the skin

For the absorption of chemicals from any biological membranes, compounds need to cross the biological cell membrane. The basic cell membrane structure consists of a bimolecular lipid leaflet that contains phospholipids, cholesterol and fatty acid esters oriented with their hydrophobic portions inside and their hydrophilic portions facing the outside aqueous environment. Associated with the lipid molecules are globular protein molecules embedded into or passing through the membrane. Compounds can cross cell membranes in several ways: passive permeation (diffusion) through the lipid bilayer, passive transport through membrane channels or pores, active transport, facilitated transport (carrier-mediated transport) and phagocytosis (Escuder-Gilabert et al., 2003). However, most drug substances cross cells by passive permeation. In this process, a

substance dissolves in the membrane lipid bilayer, permeates through the membrane and enters into the cytoplasm of the cell. To establish an adequate concentration gradient for passive permeation the substance not only must be soluble in lipids but also must be sufficiently soluble in water due to the aqueous nature of the extracellular and intracellular spaces. Hence lipid-water partitioning may be the most important factor governing a substance's ability to diffuse through cell membranes (De Vito, 2000; Ho et al., 1977).

A molecule can penetrate healthy human skin via three different macro-routes; the sweat ducts, through the hair follicles with their associated sebaceous glands or across the continuous SC between these appendages (Figure 1.3). The fractional appendageal area available for transport is only about 0.1% (Barry, 2001).

Skin is a very specific type of biological membrane, owing to the unique properties of the SC. Diffusion through the highly lipophilic SC may occur by passive diffusion and the passage through the remaining sub-layers of the skin maybe more rapid due to the lower lipophilicity (De Vito, 2000).

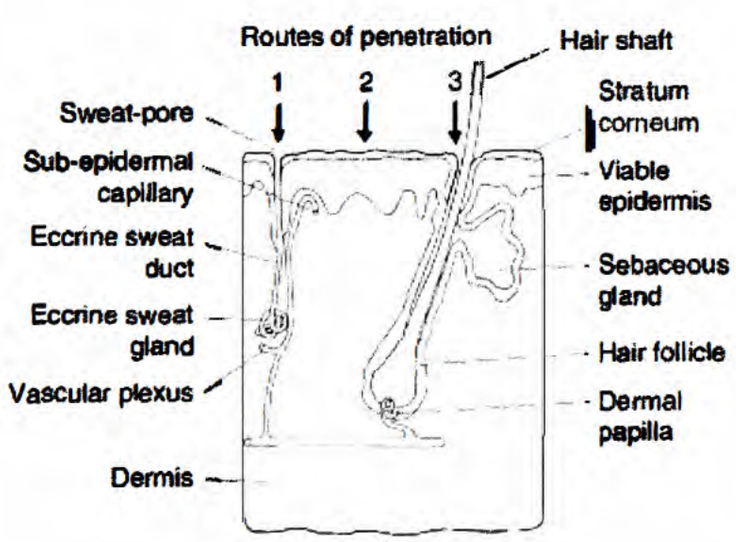


Figure 1.3: Macro-routes of penetration, taken from (Barry, 2001).

As it was mentioned before, SC is the most difficult barrier for a molecule to penetrate the skin, be absorbed and metabolised. There are two micro-routes for a molecule to penetrate through SC, the intercellular and the transcellular pathway (Barry and Williams, 1995). In the intercellular pathway a molecule penetrates via the intercellular space which consists of lipids and aqueous environment. In the transcellular pathway the penetrant penetrates the SC in a straight line, passing through cells and intercellular space (Figure 1.4). Most molecules penetrate through skin via the intercellular micro-route and therefore many enhancing techniques aim to disrupt or bypass SC's molecular architecture (Barry, 2001). Skin permeation of chemicals is often described using the 'brick and mortar' analogy. The SC's 'brick and mortar' structure is analogous to a wall (Figure 1.4). The corneocytes of hydrated keratin comprise the 'bricks', embedded in a 'mortar' that is composed of multiple lipid bilayers of ceramides, fatty acids, cholesterol and cholesterol esters (Michaels et al., 1975; Barry, 2001).

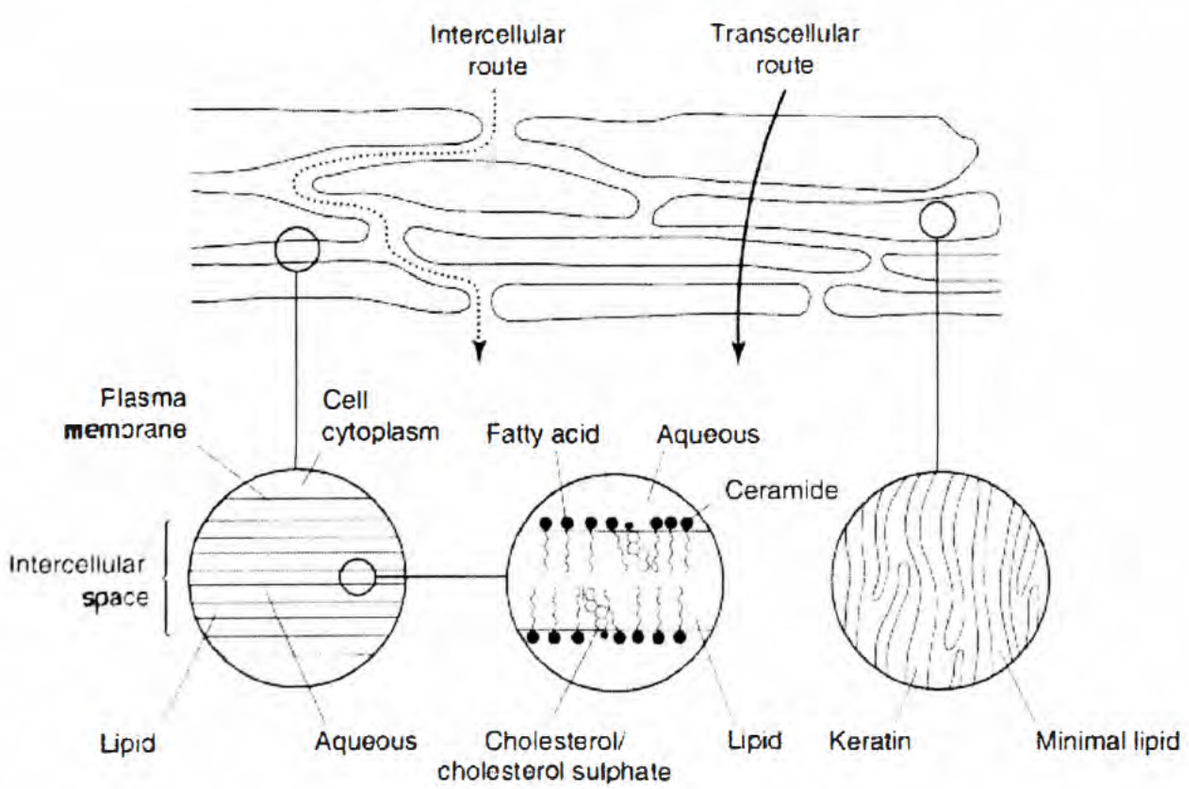


Figure 1.4: Micro-routes of penetration, taken from (Barry and Williams, 1995).

1.3. Factors affecting skin absorption

The rate of absorption of a drug is mainly influenced by age, sex, race, anatomical region, skin health status, temperature and blood flow rate, hydration, dose, duration of exposure, occlusion and the synergistic effect of a penetrant with a solvent in case of a mixture (Barry, 2007; Riviere et al., 2001). The race and sex of a human do not show noticeable differences in barrier properties of the skin (Lotte et al., 1993). The age of a human can influence skin absorption because of the differences in SC dryness, sebaceous glands activity and amount of skin surface lipids, due to flattening of the dermal-epidermal junction and attenuation of blood supply to viable epidermis caused by atrophy of the skin capillary network (Roskos et al., 1989).

The anatomical site on a human is also important because of the differences in thickness and composition of the SC, sebum composition and size and numbers of follicles (Scheuplein and Blank, 1971; Barry, 1983). The major route of permeation in SC is around the corneocytes, therefore, the larger the corneocytes the longer the route for the permeation. The size of the corneocytes is dependent on the site on the body it belongs to and this may affect skin's permeability. For example, the skin on the face is thinner and corneocyte size is smaller than a corneocyte located in the SC of the thicker abdominal area. This results in a shorter path for a drug to penetrate (Handgraft and Lane, 2009).

1.4. Assessment of skin absorption

1.4.1. *In vivo* methods

In vivo methods make use of living animals including humans. Generally animals' skins differ significantly from the skin of humans especially in characteristics that affect percutaneous absorption, such as the thickness and nature of the SC, the density of hair follicles and sweat glands, the nature of the pelt, the capillary blood supply and

biochemical aspects. Therefore, results taken from animal studies will need additional data from human studies (clinical assessment) in order for the regulatory bodies to grant a product licence (OECD 2004a, b).

Several methods can be used for the *in vivo* assessment of percutaneous absorption; the best is measuring the drug levels in the blood after dermal exposure. Some of the most common methods are explained below:

1. Plasma and/or excreta measurement: The dermal absorption of a chemical can be assessed by analyzing its concentration and/or its metabolite(s) in plasma, exhaled air, or urine (Akrill et al., 2002; Engstrom et al., 1977). The amount of the chemical that is absorbed can be measured from the concentration of a marker or from radioactivity in the case radiolabeled compounds have been used. By comparing either the totally excreted amount of a chemical or the area under the plasma concentration-time profile (AUC) of the two routes of administration (dermal and ref, the latter could be an oral dose), the dermal absorbed dose can be calculated from Eq. (1.1) (Kezic, 2008).

$$Absorption\ dose_{dermal} = \left(\frac{AUC\ or\ Excretion_{dermal}}{AUC\ or\ Excretion_{ref}}\right) \cdot Dose_{ref} \quad \text{Eq. (1.1)}$$

From the absorbed amount, exposed skin area, and exposure duration, the average absorption rate throughout the exposure can be deduced from eq. (1.2) (Kezic, 2008).

$$Absorption\ rate_{dermal} = \frac{Absorption\ dose_{dermal}}{(Area \cdot Time)} \quad \text{Eq. (1.2)}$$

From both equations only limited information on dermal kinetics can be extracted. It is preferred to use methods that estimate the dermal absorption rate time profile, such as (de)convolution methods, using the plasma concentration-time profile obtained from a dermal and a reference exposure of known rate and duration (Opdam, 1991). This approach can provide us with the average and maximum absorption rate as well as permeability coefficient (k_p) and lag time in case of steady-state absorption. In case of

radiolabeled chemicals being used, a mass balance approach may be preferred (Hueber-Becker et al., 2004).

2. Surface loss: Determination of the flux of material into the skin can be performed by measuring the loss rate from the vehicle. High performance liquid chromatography is a method that can be used for loss determination. The drawback of this method is that because of the skin's low permeability, the vehicle may be changed by evaporation or by dilution with sweat or transepidermal water more likely than partitioning of the drug into the skin or deposition to the skin surface and SC (Barry, 2007).

3. Histology: A common method of tissue analysis is the histological examination of the excised tissue. However, cutting, handling and development of skin sections encourage leaching and translocation of materials away from their original sites. Penetrants that fluoresce (e.g. tetracycline) can reveal the location of skin penetration routes via electronic microscopy. The use and combination of confocal microscopy techniques, such as laser scanning, fast digital signal capture/processing and image stabilization, can provide information at different depths in the skin (Rajadhyaksha et al., 1999).

4. Microdialysis: This is a semi-invasive sampling technique with probes inserted in the dermis and perfused with buffer for measuring concentrations of a substance in the deeper skin layers (Wagner et al., 2000). The principle of the technique is based on the passive diffusion of a chemical across the semi-permeable membrane of a microdialysis probe that is introduced into the dermis parallel to the exposed skin surface. The probe is slowly perfused with a tissue compatible sterile buffer (the perfusate), mimicking the blood flow. Molecules able to pass the probe membrane will diffuse across the membrane into the perfusate, which is collected at timed intervals for analysis (Kezic, 2008). The microdialysis method is less suitable for very lipophilic of highly protein-bound compounds because of the low recoveries of these compounds (Benfeldt et al., 1998). Microdialysis can be employed *ex vivo* as well as *in vivo* (Leveque et al., 2004).

5. Analysis of body tissues (e.g. biopsy of viable or non viable epidermis) or fluids e.g. circulating blood or urinary analysis is a major method for determining the skin absorption of chemicals *in vivo*. In addition, tape stripping of human SC has frequently been used for investigation of skin penetration barrier function and the factors involved in skin pathologies. Tape stripping is simple, inexpensive, with low invasive ability and can be used in humans and animals (Bashir et al., 2001).

The tape stripping method is based on the determination of the amount of chemical in the separate layers of the outermost layer of the skin, the SC. Generally a predetermined area of the skin is exposed to a chemical for a certain period of time. After the end of exposure, SC layers are removed sequentially by adhesive tape. The amount of recovered substance in each tape strip is determined with an appropriate analytical technique (Kezic, 2008). Due to variability in the amount of SC removed by tape caused from several factors, such as adhesive properties of the tape, pressure applied on the tape or even inter-individual anatomical differences, the measurement of the amount of protein on the tape is also essential, by monitoring trans-epidermal water loss or spectrophotometrically (Dreher et al., 1998). It should be mentioned that tape stripping is unsuitable for determination of volatile and rapidly penetrating chemicals, where the time needed to remove the entire SC is critical (Reddy, 2002).

6. Observation of a pharmacological or physiological response: Penetration kinetics can be determined when a drug stimulates a reaction in the viable tissue. Such a reaction can be a local allergic, toxic or physiological reaction (e.g. sweat gland secretion), pigmentation, sebaceous gland activity, vasodilatation, vasoconstriction, vascular permeability, epidermal proliferation and keratinisation changes in blood pressure, reduction in pain threshold and production of convulsions (Barry, 2007).

7. Physical properties of the skin: An example of this is the measurement of the amount of water loss from within the skin to the external atmosphere (transepidermal water loss) (Handgraft and Lane, 2009). Other methods for determining the physical properties of the skin include thermal determinations (e.g. Differential Scanning Calorimetry),

ultrasound (a technique of cyclic sound pressure revealing inner structures of a medium or animal), classification of function and dimension, spectral analysis (infrared and Raman) and the use of photoacoustic and electrical properties (Barry, 2007).

8. Bioassays: These involve measuring of the effects of a substance on a living organism, for example screening of topical formulations prior to clinical trial e.g. antibacterials (Hadgraft and Lane, 2009). Topical corticosteroid bioassays are the most sophisticated and refined of all such bioassays, and include, antigranuloma, thymus involution, inflammation, cytological techniques and psoriasis bioassays (Barry, 2007).

1.4.2. *In vitro* methods

As *in vivo* techniques require ethical approval, *in vitro* approaches have been used for dermal absorption studies and are also cheaper and faster (European Comission Guidelines and Guidance Documents, 2002; SCCNFP, 2003; ECB). Dermal absorption is an area in which *in vitro* approaches have a significant role to play as skin is a relatively easily accessible tissue (Williams, 2006). *In vitro* methods can use skin samples from humans (either viable or cadaver), excised animal skins (e.g. from rats or pigs) or artificial membranes.

Artificial membranes such as cellulose acetate, silicone rubber of isopropyl myristate, or lamellar systems, or skin equivalent in which artificial skin is grown from keratinocytes and fibroblasts are designed to mimic the intercellular lipid of the SC (Chilcott et al., 2005). Because such artificial membranes are not as complex as the human skin, experimental results need special consideration in the case of extrapolation to clinical trials. Similarly, the differences in the permeability of human and different animal skins have been observed in various studies (Moody, 1995).

In vitro approaches have a specific role in investigating the mechanism of interactions during absorption, such as the effects of multiple doses, mixtures, vehicles and dose

(Pendlington et al., 2004). Using *in vitro* approaches, dermal absorption may be measured using either flow through or static diffusion cell systems, both described by OECD guideline 428 (OECD, 2004b). The static diffusion cells have some advantages compared to the flow-through system because of their much simpler design, having only a magnetic stirrer as a technical feature. Thus their cost is lower than a flow-through diffusion system. On the other hand, the flow-through system simulates better the real physiological conditions due to the continuous replacement of receptor fluid (Bronaugh, 2004) resembling the systemic uptake of the drugs/chemicals in the blood vessels. However many comparative studies of static and flow-through diffusion cells have been carried out in the past with no difference in skin penetration measurements (Bronaugh and Maibach, 1985; Bronaugh and Stewart, 1985; Chilcott et al., 2005; van de Sandt et al., 2004).

The Franz diffusion cell is a widely used system for *in vitro* skin permeation studies (Friend, 1992) and may be run as static or flow through (ECETOC, 1993). In the donor chamber aliquots of penetrants can be applied only limited by the chamber's maximum volume. A skin membrane, which can be of various thickness, is applied between the donor and receptor chambers, the upper area of the skin is in contact with the donor chamber and the lower area of the skin is in contact with the receptor chamber. The heater circulator in the receptor chamber maintains a steady temperature of around 32°C, similar to the temperature of skin in a living human being. In the receptor chamber under the skin, the receptor fluid can manually be sampled by removing aliquots periodically for analysis (Bronaugh, 2004).

1.4.2.1. Static diffusion cells

For the static system (Figure 1.5), the sink conditions have to be arranged in a way that the concentration in the receptor fluid is not sufficient to inhibit absorption (Williams, 2006). It may be important to determine the levels and fate of chemicals remaining in the skin and a full mass balance in the *in vitro* studies.

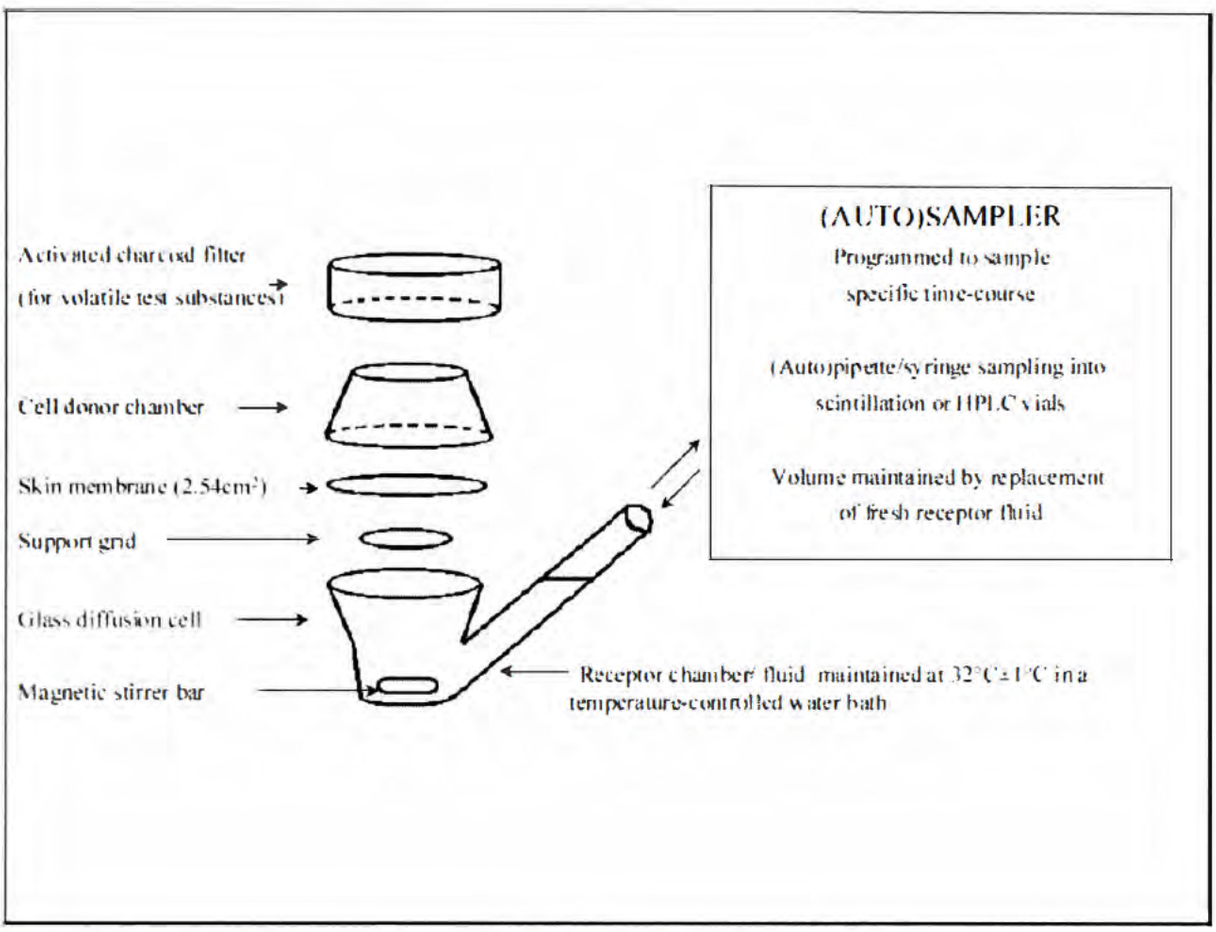


Figure 1.5 Typical design of a static diffusion cell (OECDb, 2004).

1.4.2.2. Flow-through diffusion cells

The flow-through system mimics very well the *in vivo* situation as receptor fluid flows below the skin. The skin can remain viable for up to 48 hours and the local metabolism can be studied, however, on the other hand it leads to dilution of the absorbed material limiting sensitivity. Flow-through cells are the same as static cells with the addition of sampling port in the receptor chamber for the continuous replacement of the receptor fluid in order to mimic continuous blood flow, as in *in vivo* conditions. A flow-through diffusion cell system has been developed by Bronaugh and Steward (1985) (Figure 1.6). The flow-through diffusion cell method is also useful when a permeant has a very low solubility in the receptor medium. Sink conditions are maximized, as the fluid is continually replaced using a suitable pump (at a rate of about 1.5 ml/h) (Brain et al., 1998).

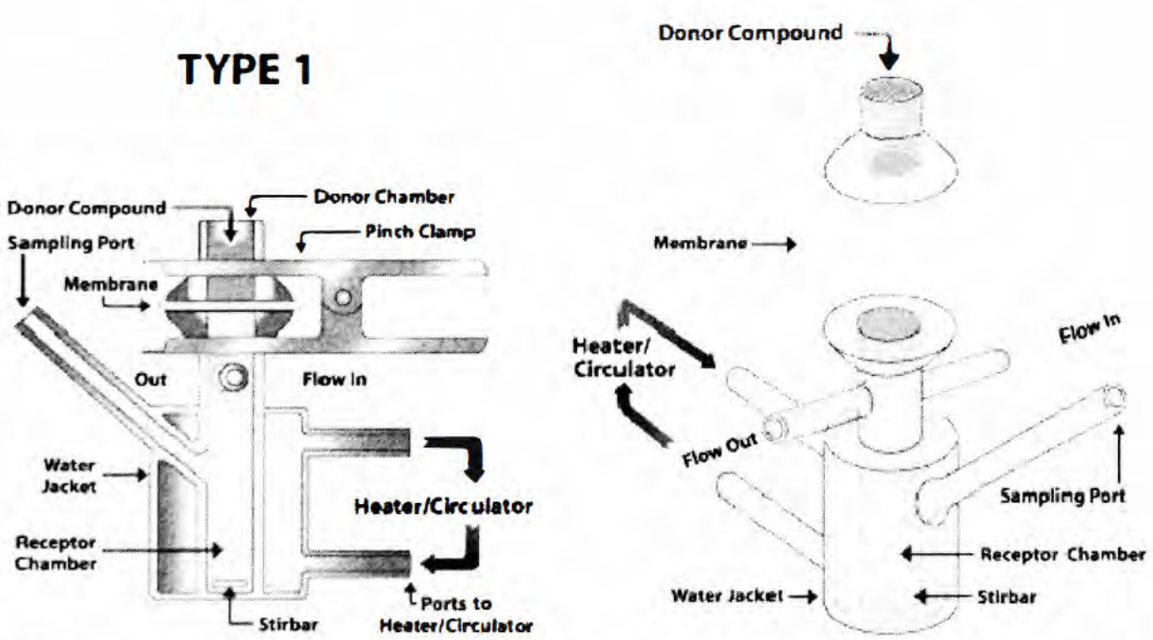


Figure 1.6. An example of a flow-through diffusion cell (http://www.permegear.com/franzflow.htm)

Usually the Franz cells are constructed from Teflon and contain a glass window in the receptor chamber for viewing the receptor contents. A number of flow-through diffusion cell systems that are similar to the Bronaugh cell system exist (Clowes et al., 1994; Tanojo et al., 1997).

The flow-through diffusion cell system is associated with a lag time, this is due to the low flow rate compared with the volume of the receptor chamber and outlet. Therefore, the necessity of minimizing the lag time effect with altered conditions is important (Anissimov and Roberts, 2001). Today, with the commercial availability of automated flow-through systems, the sampling can be left unattended because such systems employ tube-shaped skin permeation cells that fit directly into standard 2-ml glass auto-sampler vials (Moody, 1997, 2000).

1.4.3. Analysis of *in vitro* diffusion results

The percutaneous absorption of chemical substances and drugs maybe characterised by the total amount penetrated, the percentage of absorption of applied dose, the flux and the permeability coefficient k_p (Korinth et al., 2005).

In infinite dose assays the permeation rate reaches a steady state flux. The slope of the plot of the accumulated amount penetrated versus time (Figure 1.7) is the steady state flux, which can be used to calculate the permeability coefficient (k_p). There is usually an interval (lag time) between applying the dose and the steady state being reached. The 'lag time' is derived from the graph of the cumulative absorbed dose versus time (Figure 1.7) and it is the intercept (on the time axis) of the tangent to the linear part of the absorption profile (Jones et al., 2004).

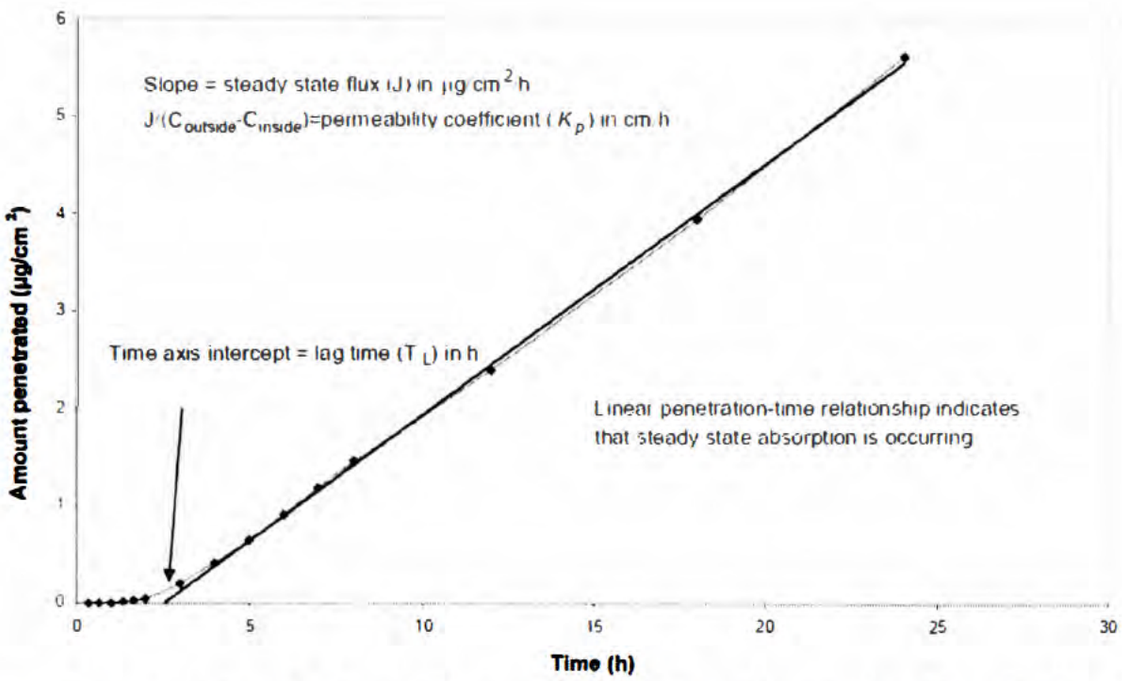


Figure 1.7. Illustrarion of flux and k_p and lag time. Steady state is achieved when the plot becomes linear, extrapolation of the linear portion to the time axis yields the lag time T_L (Jones et al., 2004).

In percutaneous absorption studies the penetration of chemical substances is often described by the flux. The word flux comes from Latin '*fluxus*' that means flow. Flux in drug delivery is defined as the diffusion of molecules across a unit area in a given amount of time (μg/cm^2/h) or permeation rate, and can be described by Fick's first law of diffusion (Buist et al., 2005).

The mechanism by which chemicals cross skin's barriers is diffusion. Fick's law of diffusion states that passive diffusion is due to the concentration gradient and the rate of transport (diffusion) is proportional to the concentration gradient. In many experimental designs a membrane is separating two compartments (donor and receptor) with a concentration gradient operating during a run and 'sink' condition (initially zero concentration for the receptor compartment). By measuring the cumulative mass of diffusant m which passes per unit area through the membrane as a function of time t (Figure 1.7), a straight line is obtained, the slope of which is the steady state flux d*m*/d*t* (Eq. 1.3).

$$\frac{dm}{dt} = \frac{DC_0 \cdot K}{h} \qquad \text{Eq. (1.3)}$$

Where C_0 is the constant concentration of drug in the donor phase, K is the partition coefficient of the membrane and the bathing solution (because in a steady state condition the concentration in the two media might be different, the ratio of these concentration is described by K (Jones et al., 2004)) and h is the membrane's thickness. Also T_L is the time axis intercept 'lag time' when the amount penetrated (diffused) is zero (Eq. 1.4).

$$T_L = \frac{h^2}{6D} \qquad \text{Eq. (1.4)}$$

Where D can only be estimated when thickness h is available. By analogy with chemical kinetic operations eq. (1.3) represents a zero-order process with a rate constant of DK/h (Barry, 2007).

Sometimes with biological membranes (such as skin), it is not easy to differentiate between the value of D and that of K. Therefore, a composite parameter, the permeability coefficient, P, where P = KD or P = KD/h is commonly used. The latter definition is used when h is uncertain, e.g. diffusion through skin.

The permeation coefficient, k_p of a molecule is characterized according to Fick's first law of diffusion by dividing the molecule's steady state flux by the initial concentration in the donor phase ($k_p = J_{ss}/C_{donor}$, kp is measured in cm/h and C_{donor} in mg/cm^3) (Environmental Protection Agency of US, 1992) and can only be calculated under infinite experimental conditions.

Fick's law is only valid for homogenous membranes, when the concentration gradient is constant in time, and diffusivity is independent of concentration (Buist et al., 2005). Korinth et al. (2005) stated that Fick's first law of diffusion is definitely able to describe

the distribution of chemicals in fluids and is used in various technical fields but its applicability in complex biological systems, with exception of the total gas exchange in the lung, is limited. This is because the skin is not a simple porous membrane but a complex entity designed to protect the organism from the environment. The experimental evidence for the general validity of physical diffusion laws for the assessment of percutaneous absorption is missing thus far, making k_p an artificial parameter, which is not applicable to specify or even to quantify the influence of the concentration of chemicals on percutaneous absorption. A few studies also indicate that k_p is unable to adjust the flux to the concentration of some chemicals. For example Liron and Cohen (1984) demonstrate that k_p of alkanoic acids enhances with increasing concentration of this compound in n-heptane as vehicle.

It is believed that the permeability coefficient is a more reliable parameter than flux taking the concentration into consideration and should remain constant for each compound independent from the grade of dilution, and can also be used for concentrations other than the one used in the experiment (Environmental Protection Agency of US, 1992). The chemical uptake rate, relating the crossing of the barrier of the chemical itself in terms of the bulk concentration, then becomes C times k_p times the surface area exposed (Environmental Protection Agency of US, 1992). In addition the permeation coefficient has been selected to quantify dermal absorption in QSAR studies (Fitzpatrick et al., 2004; Potts and Guy 1992), in comparison of various species studies (Lundh et al., 1997; Mehta et al., 1991) in vehicle effects studies (Liron and Cohen 1984; Sloan et al., 1986) and in barrier integrity of human skin studies (Bronaugh et al., 1986).

1.4.4. Sources of variability with *in vitro* measures of skin absorption

In vitro methods can vary greatly in terms of the skin samples, experimental procedures, and the resulting measurements. The inter- and intra-laboratory variation with *in vitro* percutaneous absorption methodology has been investigated to some extent in the past.

In a recent study by van de Sandt et al. (2004) the *in vitro* absorption of several compounds through human skin were determined in nine different laboratories and through rat skin in one laboratory. In all laboratories the studies were undertaken according to the detailed protocols of OECD (2004b) in terms of dose, exposure time, vehicle, receptor fluid, preparation of membranes and analysis. Results of this study showed noticeable differences that may be attributed to the inter-individual variability in absorption between samples of human skin and differences in skin site and source. Skin thickness only slightly influenced the absorption of benzoic acid and caffeine; however the maximum absorption rate of the most lipophilic compound, testosterone, was clearly higher in laboratories using thin, dermatomed skin membranes.

1.4.4.1. The Skin Samples

Not only have human and several animal skin types been used in *in vitro* studies, reconstituted human skin and several types of artificial membranes have also been suggested to mimic human skin (Chilcott et al., 2005). Human, monkey, rat, mouse, pig and guinea pig skin are some of the skin types have been used in dermal absorption studies (EDETOX, 2010). Interspecies comparisons of the skin absorption of a chemical substance have demonstrated skin permeability differences among species and sex. For example in the study of Moody (1995), rat and guinea pig skin overestimated the degree of dermal absorption in human skin after application of benzo[a]pyrene. Also the study of Bronaugh et al. (1983) concluded that skin from the back of male rats was approximately two-fold less permeable than that of female rats of the same strain after application of benzo[a]pyrene, because of the two fold thicker SC.

Rat skin is often used for toxicological studies. It is also required for special protocols, such as that for pesticide testing in the USA (USEPA, 1998; Zendian, 2000). According to the Scientific Committee on Cosmetic Products and Non-Food Products Intended for Consumers (SCCNFP), human skin (abdomen or breast) and pig skin are recommended for cosmetic testing (Steiling et al., 2001; SCCNFP, 2003).

In terms of human skin samples, the skin can be cadaver human skin or surgically removed skin which may be used fresh as viable skin or after a certain period of freezing. These can all lead to variability in the experimental results. For example, the absorption of benzoic acid and para-aminobenzoic acid were significantly greater in non-viable, compared with viable, metabolically active hairless guinea pig skin (Nathan et al., 1990). Moreover, skin samples could be used as full thickness skin (epidermal membranes) or dermatomed with varying thicknesses such as SC, epidermis, epidermis with dermis layer/s (Wilkinson et al., 2006). Epidermal membranes can be enzymically, heat or chemically separated (OECD, 2004b).

1.4.4.2. The receptor fluid

An important source of potential variability in *in vitro* tests is the composition of the receptor fluid. The receptor fluid is intended to mimic the capillary blood circulation (in which the chemical must be soluble) and not affect the skin barrier integrity (OECD, 2004a). The solubility between the test compound and the receptor fluid is important, especially in static diffusion systems, as it may affect the sink capacity and can have an influence on the receptor chamber dimensions or sampling frequency (Brain et al., 1998). Thus the selection of the receptor fluid and/ or the flow rate in flow-through cells is based on the solubility of the test compound, and the volume of the receptor chamber.

As stated by Skelly et al. (1987), the criteria that establish a suitable flow rate are 1) the concentration in the receptor fluid should not be greater than about 10% of the solubility limit (as for static diffusion cells) but large enough to be detected and 2) there must be adequate mixing in the receptor chamber. In most studies the receptor phase for hydrophilic compounds is usually saline or isotonic buffered saline solution. However for lipophilic compounds it can contain solvent mixtures such as water:ethanol (Greaves et al., 2002), or 5% bovine serum albumin (Sartorelli et al., 2000; Bronough, 2004) to aid solubility in receptor phase; it may also contain surfactants, proteins or even non-microbial growth factors (Collier et al., 1989). The effect of pH of the receptor fluid has

been studied on the absorption of Nicardipine and it was concluded that the receiver's pH effect, for weakly ionizable drugs, is more significant for compounds having high partition coefficients (Kou et al., 1993).

1.4.4.3. Duration of exposure and sampling time

The exposure time should reflect in-use conditions. The exposure time may therefore vary between a few minutes for a rinse-off product up to a maximum of 24h for a leave-on product and 6-8 h for industrial products (OECD, 2004b). For finite dose experiments, the wash-off from the skin should be performed with an aqueous soap solution and the time of the wash-off determines the exposure (OECD, 2004b). It is important to sample the receptor fluid for at least a 24h period. Increased exposure times are appropriate only in cases of long lag phases or for infinite applications in order to achieve a steady-state flux (OECD, 2004b). Periods of 24h or 48h should be adequate to study dermal absorption. Unless adequately preserved, the skin can deteriorate beyond this time (Jewell et al., 2000). Standard exposure times for test substances in finite dosing experiments are up to 24h and shorter in the case of rinse-off products (although the measurement of penetration of material continues for at least 24h).

1.4.4.4. Skin Hydration

Experimental approaches vary from studies employing prior hydration of the skin before the start of the experiment; to those using infinite doses (aqueous donor phase) which leads to skin hydration during the period of the experiment; studies using occlusion of the skin which may lead to some levels of hydration during the experiment; or those investigations employing finite dosing without occlusion which limits the skin hydration.

It has been demonstrated that the SC can significantly change its dimensions when exposed for long periods to water (Bouwstra et al., 2003). When in water or exposed to high humidity conditions, the SC swells and expands to a greater extent in the plane perpendicular to the skin surface rather than in a lateral dimension. This, naturally, has an impact on the tortuosity of the SC and, as such, can certainly influence SC barrier function (Crowther et al., 2008).

Skin occlusion has been found to enhance the percutaneous absorption of many, but not all topically applied compounds (Zhai and Maibach, 2001). On the other hand, un-occluded conditions can simulate the normal exposure situations in everyday life. However, volatile compounds may evaporate under un-occluded conditions and infinite dosing can only take place under occluded conditions (Kligman, 1983; Bronough and Stewart, 1985; Baker, 1986).

1.4.4.5. Types of exposure

With regard to dermal absorption there are two types of exposure, infinite and finite dose. In infinite dose experimental conditions the amount of test preparation that is applied to the skin achieves a maximal flux that is maintained (steady state) over time, thus it can be referred to as the steady state flux (J_{ss}). In finite dose experimental conditions the amount of test preparation is insufficient to maintain a maximal flux (Buist et al., 2005). The finite dose condition is used for simulation of everyday exposures of skin to penetrants. The dose solution is applied in a volume sufficient to cover the skin and normally remains unoccluded. In the infinite dose condition the solution is applied in excess and can be occluded for the duration of the study (OECDb, 2004; Sartorelli et al., 2000). According to the OECD (2004b), for finite dose experiments, a dose of 1-5 mg/cm^2 or 10 µl/cm^2 should be spread on the skin surface and for infinite dose experiments, a dose higher than 10 mg/cm^2 or 100 µl/cm^2 is needed in order to obtain steady state-conditions from which the flux and k_p can be calculated. In the literature, depending on the aims of the experiment, a full spectrum of application

doses can be found with varying duration of exposure and sampling time. However, in practice, dermal exposure will mostly occur under finite rather than infinite dose conditions, during which the concentration of the solution on the skin changes significantly over time (Buist et al, 2010).

1.4.5. Comparison of *in vitro* and *in vivo* skin absorption results

A vast number of historical studies have compared *in vitro* and *in vivo* methods for skin absorption. *In vitro* absorption of the pesticide isofenphos using flow through diffusion cells with human cadaver skin gave a percentage of dose absorbed similar to that of *in vivo* human volunteer studies (Wester et al., 1992). The dermal absorption of lindane in acetone and a white spirit-based formulation was investigated *in vitro* and *in vivo* by Dick et al., (1997a,b), where *in vitro* studies predicted the 40-fold difference in absorption between the two applications observed in the human volunteers. In other studies, *in vivo* and *in vitro* predictions with human skin using the same dose, vehicle, and application time for 50% 2-butoxyethanol/water, resulted in a close relationship between absorption rate determined *in vitro* (2.0 $\mu g/cm^2/h$) (Wilkinson and Williams, 2002;) and *in vivo* (2.7 $\mu g/cm^2/h$) (Jakasa et al., 2004). In the study of Trauer et al. (2009) on the permeation of topically applied caffeine *in vivo* and *in vitro* through human skin, *in vitro* methodology proved valuable for estimating *in vivo* dermal absorption. However, they concluded that *in vivo* experiments should not be abandoned as with *in vitro* tests, structural changes of skin occur and blood flow and metabolism are absent, probably accounting for reduced penetration rates *in vitro*.

In cases where there is a lack of correlation between *in vitro* and *in vivo* data, the discrepancy may be due to not correctly controlled experimental conditions or because they were conducted prior to the adoption of OECD Test Guidance 428 (OECD, 2004b). The OECD Test Guideline 428 has confirmed that *in vitro* studies can predict *in vivo* absorption when the correct methodology for both tests is used.

1.5. Effect of mixture components

According to OECD guidelines, application of a drug to the skin can be neat, diluted or with a vehicle (OECD, 2004a, b). A vehicle can be a single component, e.g. water, or a mixture of components e.g. water and ethanol. A vehicle can play a very important role in the penetration of a chemical through the skin. According to Roberts et al. (2002), vehicles can affect skin permeability by a range of mechanisms including delipidization, dehydration, fluidization, desmosome disruption in the SC and also by changing the polarity of the mixture followed by change in penetrant solubility and partitioning into SC. Pure solutes, can in some cases, enhance skin permeability by a direct corrosive effect (Roberts et al., 2002; Zinke et al., 2002). Generally, substitution of organic vehicle has the potential to enhance maximal flux. In cases when there are components in the vehicle that can interact with the intercellular lipids of the SC, then it is possible that permeation may be enhanced or suppressed (Davis et. al. 2002). Formulation ingredients can alter the skin penetration of a compound by affecting the barrier properties of the skin or by changing the partitioning of the compound into the SC. Two of the major formulation components in pharmaceutical preparations are surfactants and chemical enhancers (Roberts et al., 2002).

The solubility coefficient is one of the most important factors in transdermal drug delivery. The solubility of a chemical is different in different vehicles hence resulting in different flux and k_p values due to varying levels of saturation. Theoretically, when a chemical is saturated in a vehicle then the flux should be the same, no matter the vehicle, provided that the vehicle does not affect the skin membrane hypothetically and super-saturation does not occur (Roberts et al, 2002).

A vehicle can promote the penetration of a chemical by having low solubility, in this way a chemical will not be retained in the vehicle (Baker, 1986). In order to maximize the penetration of a drug through skin the formulation must not be over-solubilised but should be at a near saturation stage. This is because the partition coefficient of a drug between the skin and the solvent mixture generally falls as the solubility in the solvent

rises. Polar co-solvent mixtures such as propylene glycol with water may produce saturated drug solutions and so maximize the concentration gradient across the SC (Barry, 2007; Wiechers, 2005), although they may also affect SC directly.

The effect of mixture/formulation components on the skin penetration of a compound depends on the nature of the components, i.e. their chemical structure and physico-chemical properties. The relationship between chemical structures of the formulation ingredients and the skin penetration modification can be studied quantitatively using Quantitative Structure–Activity Relationship (QSAR) techniques (Ghafourian et al., 2004; 2010a, b).

1.5.1 Solvents

It is widely acknowledged by the risk assessment community that most occupational and environmental exposures to chemicals are to complex mixtures, and not individual compounds in defined aqueous vehicles (Riviere and Brooks, 2005). Experimental skin absorption studies that have been conducted in the past have shown that mixture/vehicle effects may significantly modify a chemical's dermal absorption and sometimes in an extent that may even overshadow the magnitude of permeability differences between individual compounds (Baynes et al., 2002b; Brooks and Riviere, 1996; Idson, 1983; Qiao et al., 1996; Riviere and Monteiro-Riviere, 2002; Riviere et al., 2001, 2003; Rosado et al., 2003). Many solvents open up the complex dense structure of the horny layer. Mixtures of non-polar and polar solvents such as chloroform and methanol remove the lipid fraction, forming artificial shunts through which molecules pass more easily (Barry, 2007).

From a topical drug formulation point of view, there are a large number of vehicles that, in spite of being very good enhancers of skin penetration, can very rarely be used because of their deleterious effects upon the skin, such as dimethyl sulfoxide (DMSO). DMSO is a powerful solvent that increases drug penetration very well but at the same

time it alters the biochemical and structural integrity of the skin and operates by direct insult to the SC (Barry, 1983). Ethanol is very often used as an enhancer/cosolvent (i.e. in estrogen patches) and increases drug penetration (Walters et al., 1989). For example in the study of Stinecipher and Shah (1997) dermal absorption of DEET (*N,N*-diethyl-*meta*-toluamide) was significantly higher from 30-45% ethanolic solutions than that from pure DEET and also from 60-90% ethanolic solutions.

1.5.2. Skin permeation enhancers

A drug that is prepared for systemic drug delivery by the transdermal route is considered successful when it penetrates the skin barrier in sufficient quantity to achieve its desired therapeutic effect. Two methods that are used for the modification of the properties of the SC, in order to enhance drug penetration and consequently absorption through the skin, are chemical and physical methods of enhancement (Rolf, 1988). The iontophoresis and ultrasound (also known as phonophoresis or sonophoresis) techniques are examples of physical means of enhancement that have been used for enhancing percutaneous penetration (and absorption) of various therapeutic agents (Shah, 1993). In iontophoresis a small electric charge is used to deliver a molecule (e.g. medicine) through the skin, whereas in the ultrasound technique cyclic sound pressure with a frequency greater than the upper limit of human hearing is used. With the iontophoresis technique, there is a possibility that the device may cause painful destruction of the skin with high current settings. Therefore, it is essential to use high quality electrodes with adequate skin adhesion, uniform current distribution, and well-controlled ionic properties (Shah, 1993).

Chemical penetration enhancement is the most common technique used in topical drug delivery (Williams and Barry, 2004). Chemicals that enhance the penetration of topically applied drugs are usually named as accelerants, absorption promoters, or simply as penetration enhancers. When designing a transdermal drug formulation it is highly important to firstly identify chemicals that significantly enhance drug penetration

through the epidermis and at the same time irritate or damage the skin the least along with having no serious side effects in general. Important properties of such chemical permeation enhancers along with their desirable attributes have been referred and described by Barry (Barry, 1983) and reviewed (Williams and Barry, 2004). It is optimal that such chemicals are safe and non-toxic, pharmacologically inert, non-irritating, and non-allergic (Kat and Poulsen, 1971). In addition, the skin tissue should come back to its normal state along with the barrier properties when the chemical has been removed.

Generally enhancers increase the drug permeability through the skin by causing reversible damage to the SC, increase (and optimize) the thermodynamic activity of the drug when functioning as co-solvent, increase the partition coefficient of the drug thus promoting its release from the vehicle into the skin, operate by conditioning the SC to promote drug diffusion and promote penetration and establish drug reservoir in the SC (Shah, 1993).

As discussed earlier, the routes of drug penetration through the skin are the polar, the non-polar, and the polar/non-polar pathways. The chemical enhancers penetrate the skin and enhance permeation of other chemicals by altering one of the above three pathways. The polar pathway is altered in reaction to protein conformational change or solvent swelling. The non-polar pathway is altered after differentiation to the rigidity of the lipid structure and fluidization of the crystalline pathway (a fact that substantially increases diffusion). The fatty acid enhancers increase the fluidity of the lipid portion of the SC (Knutson et al., 1985; Potts, 1989). Binary vehicle enhancers may act on both polar and non-polar pathways by altering the multilaminate pathway for penetrants. The enhancers may increase the drug diffusivity in the SC by simply dissolving the lipids of the skin or by denaturing the proteins of the skin. So when designing and developing a product, the type of enhancer that will be used is of significant importance.

According to Barry (2007), there are many points that should be considered and evaluated when penetration enhancers are used, these are:

- the mechanism of action
- if it is metabolised and generally the fate of enhancers in the body
- the influence of the degraded moieties from the enhancer, if any
- if the enhancer is powerful enough to increase penetration of other excipients
- the level of irritation that it causes
- the effect of the enhancer over the full duration of dosage form application
- if its action is reversible
- the interaction potentials with formulation ingredients
- the effect on multiple application under occlusion
- the effect of repeat application of enhancers at the same site and if it alters the integrity of the SC and drug penetration properties upon application
- if the enhancer selectively enhances permeation of specific ingredients
- if the enhancer is inert or it causes contact irritation or promote bacterial growth
- how safe the enhancer is and if it causes any toxicity

In addition, the use of strong enhancers may leach other components from a patch which may penetrate the SC. If this occurs, the fate of such chemicals in the body must be known in advance (Shah, 1993).

Chemical permeation enhancers may increase transdermal drug penetration but it is critical that this task can be accomplished without skin irritation or sensitization. The key is to find an enhancer that will combine both disruption of the impermeable SC barrier membrane without altering the fragile living tissue underneath (Cooper and Berner, 1987).

Chemical permeation enhancers may increase the permeation of other formulation excipients along with the drug as well. Their own intrinsic skin diffusivity can be increased too. Such effects must be carefully considered and evaluated in order to avoid

any toxicological implications and skin irritation. It has been observed that occlusion causes irritation that it is linked to SC hydration and decreases diffusional resistance to formulation components such as enhancers (Sloan et al., 1986). Some well known enhancers are Azone (1-dodecylazacycloheptan-2-one or laurocapram), pyrrolidones such as N-methyl-2-pyrrolidone (NMP) and ethanol (Williams and Barry, 2004).

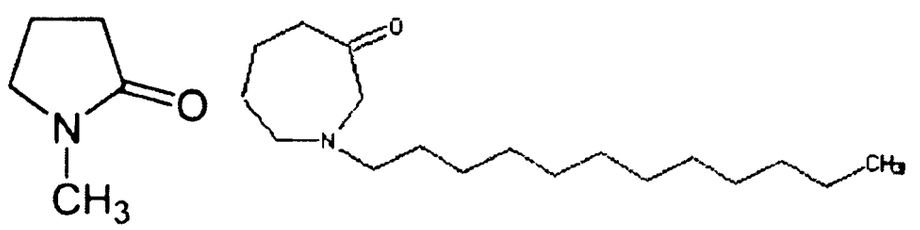


Figure 1.8. Molecular structures of NMP and Azone®.

Terpenes, a class of hydrocarbons, have been widely used in transdermal drug delivery as enhancers of drug penetration through human skin. Terpenes are naturally occurring volatile oils that appear to be promising candidates for use as clinically acceptable enhancers (Williams & Barry 1991). They have been reported to have good toxicological profiles, high percutaneous enhancement abilities and low cutaneous irritancy at low concentrations (Okabe et al. 1990). Kang et al. (2007) compared the human skin penetration effect of 49 terpenes and terpenoids on the *in vitro* permeability coefficient of haloperidol through excised human skin. They concluded that an ideal terpene enhancer should possess at least one, or combinations of, the following properties: hydrophobicity, be in liquid form at room temperature, contain an ester or aldehyde but not acid functional group, and is neither a triterpene nor tetraterpene.

1.5.3. Surfactants

The term surfactant comes from the phrase "**sur**face **act**ive age**nt**". Surfactants (or tensides) are wetting agents that are used to lower the surface tension of a liquid and the interfacial tension that leads to higher solubility of lipophilic compounds and easier spreading. Surfactants are used in many commercial formulations including skin

products. Surfactants are organic compounds that are amphiphilic, containing lipophilic ("tail") and hydrophilic ("head") groups, a property that makes them soluble in both organic solvents and water. The lipophilic part is usually a saturated or unsaturated hydrocarbon chain(s) and can be a heterocyclic or aromatic ring system (Attwood, 2007). Natural fats and oils, petroleum fractions, short synthetic polymers of high molecular weight synthetic alcohols are usually the source of such lipophilic parts.

The head (hydrophilic) part of a surfactant defines its classification e.g. anionic, cationic, non-ionic or amphoteric. Anionic (hydrophilic) surfactant include carboxylates, sulphates, sulphonates and phosphates. Cationic surfactants are usually forms of amine products. Non-ionic 'heads' associate with water at the ether oxygens of a polyethylene glycol chain. Amphoteric surfactants are not widely used as emulsifying agents although lecithin is used to stabilize intravenous fat emulsions.

The classification of surfactant decides whether it can bind cooperatively to proteins and thus denature them. With few exceptions (Otzen, et al., 2009), neutral surfactants do not denature proteins, whereas ionic surfactants do so at very low concentrations.

Surfactants have many uses in pharmaceutical formulation as wetting agents, solubilisers, emulsifiers, detergents and antibacterials (Attwood, 2007). Different concentrations of surfactants have been shown to increase the skin permeability of drugs such as diazepam (Shokri et al., 2001) and lorazepam (Nokhodchi et al., 2003). Laughlin (1978) has stated that surfactants with hydrophilic head groups enhance more effectively the percutaneous penetration of polar molecules, while those of lesser hydrophilicity should be less effective, a finding that was confirmed by Nokhodchi et al. (2003). In general cationic surfactants are more damaging and cause a greater increase in flux than anionic surfactants. Anionic surfactants cause greater enhancement and damage than nonionic surfactants (Stoughton, 1982; Cooper, 1984; Cumming and Winfield, 1994).

1.6 Quantitative Structure-Activity Relationships (QSARs)

The biological behavior of a compound is a function of its molecular structure. Therefore, one should be able to relate the biological activity to the molecular properties (molecular descriptors) of compounds quantitatively using Quantitative Structure-Activity Relationships (QSAR) techniques. These models may be able to estimate the biological activity of new compounds based solely on the properties of their molecular structures. As an integral part of the human health risk assessment of chemicals and also to be able to aid the drug delivery through skin, it is essential to be able to estimate absorption of chemicals via the dermal route. This is because despite the requirement by REACH for extensive risk assessment of chemicals, it is not practical to measure dermal absorption of the many thousands of industrial chemicals. Besides it is a well established fact that in order to enhance the permeability of a drug the choice of the vehicles and enhancers is crucial as the enhancing activities of enhancers towards different drugs are different (Ghafourian et al., 2004). The workload that is needed in order to identify just one good enhancer for only one specific drug is not a negligible factor. In order to accelerate this process a possible approach may be modeling the relationship of skin permeation measures of a series of compounds with their physico-chemical and/or structural properties.

Physico-chemical properties of substances are critical determinants of their ability to be absorbed or affect the absorption of other compounds. QSARs are mathematical models that statistically relate the biological activity of a compound to its physico-chemical properties (Magnusson et al., 2004). With QSARs we assume that the biological, physical and chemical activity of a molecule is closely related to its structure (i.e. its geometric, steric and electronic properties). So the biological activity or physicochemical properties (e.g. melting and boiling point) of a new untested molecule can be inferred from previously assessed compounds with similar molecular structure (Gramatica, 2011). The modelling of dermal contacts using QSAR leads to the development of expert systems that may be capable of reliably predicting the extent to

which a molecule will be percutaneously absorbed, without the need to make experimental measurements.

QSARs have found use in the regulatory risk assessment of new and existing compounds. The new regulatory framework, REACH, foresees the use of non-testing approaches, such as read-across chemical categories, structure-activity relationships (SARs) and QSARs (Bouwman et al., 2008; European Chemicals Bureau, 2008).

1.6.1. Molecular Descriptors

In order for the biological activity of a molecule to be related quantitatively to its physicochemical properties, calculations of descriptors that characterize its molecular structure are essential. A penetrant, after it is transported to the appropriate site of action, has to interact with that site to trigger the appropriate response. To trigger a biological response or, in this case, to be able to permeate skin, a compound must have the appropriate size and shape, or a specific range of lipophylicity, polarity or hydrogen bonding ability. Molecular descriptors such as molecular weight, melting and boiling point, number of double or triple bonds, number of carbon or oxygen molecules, are quantitative representatives of a molecule's chemical structure. Molecular descriptors can be discussed under three categories, hydrophobic, electronic and steric (size and shape) factors.

1.6.1.1. Hydrophobic descriptors

Hydrophobic descriptors indicate the comparative preference of a compound to be dissolved in non-polar organic solvents or in water. The partition coefficient (log P), where P is the octanol-water partition coefficient, is the most common descriptor. The aqueous solubility (S_{aq}) that is inversely related to P can be used in its place but cannot be measured or calculated as accurately. Chromatographic parameters, such as R_m (a

retention descriptor derived from reversed-phase thin-layer chromatography retention factor), are related to lipophilicity and are used for the measurement of the partition coefficient of molecules that are difficult to be determined experimentally (e.g. surfactants). Chromatographic parameters can only be obtained experimentally and are restricted to compounds in the range log P < 5.

Since the 1970s, various methods for the calculation of log P have been proposed (Leo. 1993; Carrut et al., 1997; Eros et al., 2002) and can be separated in two categories. Property-based methods that compute the log P as a function of molecular physico-chemical properties, such as molecular surface, volume, dipole moment, partial charges, and HOMO/LUMO energies. Alternatively additive methods use basic structural building blocks directly as descriptors. Additive methods calculate the log P value of a given molecule by summing up the contributions from all building blocks of its structure plus the required correction factors.

1.6.1.2. Electronic descriptors

Electronic descriptors represent properties such as atomic charge densities and molecular dipoles. Such descriptors are the Hammet substituent constant (σ) that is derived from pK_a values and shows the electron directing effect of an aromatic substituent. Molecular orbital descriptors such as dipole moment, atomic charges, the energy of the highest occupied molecular orbital (E_{HOMO}) represent a molecule's electron-donating ability and the lowest unoccupied molecular orbital (E_{LUMO}) represent a molecule's electron-accepting ability, can be calculated from the molecular orbital theory (Hansch and Leo, 1995). Hydrogen bonding descriptors are further electronic descriptors that can be split into H-bond donor (a hydrogen atom that is attached to relatively electronegative atom) and H-bond acceptor (when a hydrogen atom is missing) ability (Dearden and Ghafourian, 1999).

1.6.1.3. Steric descriptors

Steric descriptors characterize the size and shape of molecules. Molecular size is a much simpler property to measure than molecular shape. Examples of descriptors for molecular size are molecular volume and molecular surface area that can be calculated by summing the van der Waals volumes or surface areas of a molecule's constituent atoms (Dearden and Cronin, 2005).

Molar refractivity is the total polarizability of a mole of a substance, is derived from refractive (polarizability) index and has the units of molar volume.

Shape descriptors are also quite common steric descriptors. Sterimol descriptors represent the length of a substituent and its widths in different directions (Hansch and Leo, 1995; Dearden and Cronin, 2005). The topological shape descriptors known as the Kier and Hall shape indices represent the number of bond fragments in a non hydrogen molecular skeleton. The Kier and Hall kappa descriptors are the basis of a method of molecular structure quantification in which attributes of molecular shape are encoded into three indices ($^{1}\kappa$, $^{2}\kappa$, and $^{3}\kappa$). These Kappa values are derived from counts of one-bond, two-bond and three-bond fragments, each count being made relative to fragment counts in reference structures which possess a maximum and minimum value for that number of atoms (Hall and Kier, 1977; 1986; 1991).

1.6.1.4. 3-D descriptors

The 3-D descriptors represent three dimensionality of a molecule. Such descriptors can be as simple as inter-atomic distances or torsion angles or as complex as the distribution of electrostatic potential around a molecule. Also similarity descriptors allow comparison of the similarity of a molecule with a set of standard active molecules, on the bases of either electrostatic potential or steric parameters (Dearden and Cronin, 2005).

1.6.1.5. Topological descriptors

Topological descriptors represent the way that atoms are connected in a molecule, they include molecular connectivity descriptors of different orders such as the zero connectivity index (Chi0) and is dependent on the number of non-hydrogen atoms attached (Hall and Kier, 1986). Information content descriptors have been useful in correlating biological activity and predicting molecular similarity. The atom type electrotopological state indices are atomic level indices, combining the electronic character and the topological environment for each skeletal atom in a molecule (Kier and Hall, 1999).

1.6.2. Development and validation of QSARs

According to Gramatica (2011) 'an ideal QSAR should: 1) consider an adequate number of molecules for sufficient statistical representation, 2) have a wide range of quantified end-point potency (i.e. several orders of magnitude) for regression models or adequate distribution of molecules in each class (i.e. active and inactive) for classification models, 3) be applicable for reliable predictions of new chemicals (validation and applicability domain) and 4) allow to obtain mechanistic information on the modelled end-point.'

The interpretability of a QSAR model depends on the model's application, e.g. when predicted data are needed for screening of large libraries of chemicals, a validated mathematical model relating a target property to chemical features may be all that is necessary. In other more complex cases such as the dermal absorption of a molecule, it is obvious the need for also understanding and explaining the mechanism of action in chemical terms (Zefirov and Palyulin 2001; Livingstone, 2000).

The best fit models are not the best ones for prediction. In fact, a QSAR model must, first of all, be a real model, robust and predictive, to be considered a reliable model.

Only a stable and predictive model can be usefully interpreted for its mechanistic meaning, even so this is not always easy or feasible (Gramatica, 2011).

1.6.2.1. Multiple Linear Regression (MLR)

MLR is one of the earliest methods used for developing QSAR/QSPR (Quantitative Structure Property Relationship) models. MLR is simple in form and creates an easily interpretable mathematical expression. The weakness of MLR is that in case of correlated descriptors, it is incapable of deciding which correlated sets of descriptors are the most significant to the model. Because of that new improved methods based on MLR have been developed in order to minimize that weakness. Such methods are stepwise regression analysis the Best Multiple Linear Regression (BMLR), Heuristic Method (HM) and Genetic Algorithm based Multiple Linear Regression (GA-MLR) for the selection of the most significant descriptors (Liu and Long, 2009).

1.6.2.2. Partial Least Squares (PLS)

PLS is a statistical analysis method that is used in various fields and also in QSAR/QSPR model development. It combines features from Principle Component Analysis (PCA) and MLR. PLS can predict/analyse a set of dependent variables from a set of independent variables (predictors) (Abdi, 2003). PLS have evolved by combination with other statistical methods to Genetic Partial Least Squares (G/PLS), Factor Analysis Partial Least Squares (FA-PLS) and Orthogonal Signal Correction Partial Least Squares (OSC-PLS) (Liu and Long, 2009). In PCA, which was mentioned above, the descriptors are combined into a smaller number of terms (principle components) each of which is orthogonal (uncorrelated) to the others (Dearden and Cronin, 2005).

1.6.2.3. Artificial Neural Networks (ANN)

ANN have the ability to model non-linear correlations better than MLR (Dearden and Cronin, 2005). ANN tend to over-fit the data and have difficulty identifying the most significant descriptors in the resulting model. Radial Basis Function Neural Network (RBFNN) and General Regression Neural Network (GRNN) are the most frequently used ANN methods (Liu and Long, 2009).

1.6.2.4. Support Vector Machine (SVM)

SVM is a machine learning method, originally developed by Vapnik (1999) for pattern recognition problems in classification and regression analysis. Newly developed SVM methods are: Least Square Support Vector Machine (GS-SVM), Grid Search Support Vector Machine (GS-SVM), Potential Support Vector Machine (P-SVM) and Genetic Algorithms Support Vector Machine (GA-SVM) (Liu and Long, 2009).

1.6.2.5. Classification and Regression Trees

Classification trees are used to predict membership of cases or objects in the classes of a categorical dependent variable from their measurements on one or more predictor variables. Classification tree analysis is one of the main techniques used in so called data mining. The goal of classification trees is to predict or explain responses on a categorical dependent variable.

1.6.3. QSAR models of skin absorption

Most QSARs for skin absorption contain descriptors of hydrophobicity since the passage of a molecule through the SC involves a lipophilic pathway. As it is also a

diffusion process, descriptors of molecular size, such as molecular weight (MW) or volume (V), are also commonly included (Ghafourian and Fooladi, 2001). QSARs for the skin permeation are of two types. The 'general' type where the algorithm is intended to encompass a wide range of substances without reference to their chemical nature. The 'specific' type that have been developed to cover a smaller number of homologous chemical molecules or to investigate the influence of a particular physico-chemical property that is common to a group of molecules (Magnusson et al., 2004).

A pre-requisite to QSAR studies is the availability of a reliable dataset. In addition its statistical validity is related to the validity of the data and should be applied only within the model's applicability domain (Gramatica, 2011). Dermal absorption data using infinite doses, generated mostly by *in vitro* approaches with human skin have been combined in a number of datasets (Flynn, 1990; Cronin et al., 1999; Fitzpatrick et al., 2004; Riviere and Brooks (2005); EDETOX, 2010). Recently a database of *in vitro* and *in vivo* dermal penetration studies from peer-reviewed journals in which the quality of data were considered has been established (EDETOX, 2010). However, only about 50% of the entries satisfy the criteria for study design laid down by Soyei and Williams (2004).

Several QSAR models are currently available for the estimation of the skin penetration of drugs from the available infinite dose data. Potts and Guy (1992) developed a simple two-parameter model based on the hydrophobicity (expressed as octanol/water partition coefficient) and molecular weight of compounds. Several other investigators have shown the involvement of hydrophobicity and molecular size in the transdermal penetration of chemicals (Lien and Gao, 1995). Other prediction methods include those based on the Linear Free Energy Relationships (LFER) developed by Abraham and Martins (2004) and models based on theoretical structural descriptors (Ghafourian and Fooladi, 2000; Dearden et al., 2000). Recently, more complicated non-linear models have been proposed by several investigators. Examples of these include models based on Artificial Neural Networks (Degim et al., 2003; Katritzky et al., 2006) or an ensemble model using k-nearest-neighbour models and ridge regression (Neumann et

al., 2006). In addition a step forward has been achieved in understanding, modelling and predicting the effect of complex mixtures (Ghafourian et al., 2010a, b; Riviere and Brooks, 2011).

In reality, estimation of skin absorption is complicated due to the inconsistency of the methods and therefore the inconsistent and sometimes even controversial results of *in vitro/in vivo* tests. The inter-laboratory and inter-individual variability is often high (van de Sandt et al., 2004; Chilcott et al., 2005) which may be explained by the huge variety of methods and test systems used for skin permeation experiments. Moreover, methods of calculation and interpretation of results from the complex experimental set-up also varies (Henning et al., 2009).

2. Aims and Objectives

The aim of the research was to develop predictive models for the estimation of skin absorption of mixtures. This required characterization of the effect of formulation factors, experimental variables, and molecular properties on the permeation of compounds through the skin. Appropriate molecular modelling, Quantitative Structure – Activity Relationships and data mining techniques were used in order to rationalize the effect of molecular structure of the permeant and properties of the ingredients present in the formulations/mixtures on the penetration of chemicals through the skin under different experimental conditions represented by well defined parameters. Due to the availability of the *in vitro* data and also due to the fact that *in vivo* measures of skin absorption are more complex involving pharmacokinetics principles such as metabolism, distribution and excretion, this work focuses on estimation of *in vitro* endpoints such as flux and permeability coefficient.

Using QSAR techniques it is possible to estimate permeation of compounds through the skin and also the effect of formulation ingredients on the permeation of a chemical. Significant progress has been made in predicting the dermal absorption/penetration of topically applied compounds using QSAR models (Linusson et al., 2010). One aim of the current work was to expand the application of QSAR to the prediction of the effect of formulation factors such as solvents, permeation enhancers and other mixture ingredients. The effect of structural characteristics of the chemical enhancers on the permeation of drugs has been investigated (Ghafourian et al., 2004; Pugh et al., 2005) and different mechanisms have been proposed for the effect of enhancers on different penetrants. QSAR models have been developed for the effect of terpenes and other enhancers on 5-fluorouracil, estradiol, lorazepam, sodium diclofenac (Ghafourian and Goudarzi, 2009; Nokhodchi et al., 2007). In this project, QSAR studies were extended to allow for the development of models analyzing the effect of enhancers on other drugs. Kang et al. (2007) compared the Human Skin Penetration effect of 49 terpenes and terpenoids on the *in vitro* permeability coefficient of haloperidol through excised human skin. Also, in a recent publication the effect of thirty-eight different solvents including

terpenes, alcohols, and fatty acid solvents on the transdermal absorption of formoterol has been systematically studied with the aim of finding the best vehicle system for drug delivery (Kakubari et al., 2006). These have each provided sets of data which were explored using QSAR and data mining techniques in order to develop 'local' models.

From a toxicological point of view, there have been efforts to study the effect of mixture ingredients on the penetration of environmental pollutants and to use QSAR to predict the penetration of complex chemical mixtures (Riviere and Brooks, 2005; 2011; Xia et al., 2007). The study of Riviere and Brooks (2005) investigated the effects of solvents on skin permeability. Their skin permeability dataset consisted of permeability coefficients of 12 different penetrants each blended in 24 different solvent mixtures measured from finite-dose diffusion cell studies using porcine skin. The focus of this current work was to model the effect of mixture components on the skin penetration of chemicals. Considering the huge number of possible combinations of chemicals that might be intended for delivery through skin, or the skin might be exposed to unintentionally, a systematic method for characterizing the effects of formulation factors and or mixtures of chemicals was of great interest.

The final aim of this study was to develop models that can also incorporate the effects of experimental conditions such as membrane thickness, occlusion, hydration, vehicle ingredients and mode of dose application (finite or infinite dosing) on skin permeation flux. This was achieved through the use of a large dataset extracted from the EDETOX database and collated from more recent publications. The dataset was large enough to investigate statistically the combined effects of experimental conditions, mixture components and molecular structures of the penetrants using several data mining tools. The QSARs focused on skin flux as the response variable and linked this to the chemical structures of the penetrants and the physico-chemical properties of the vehicle mixtures.

Linear regression was used to develop QSARs where the dataset was relatively small (sections 3 and 4). For larger datasets, it was possible to employ some of the more complex non-linear methods including boosted trees, or simple regression trees.

Validation of a QSAR model is an essential part of *in silico* predictability. Since REACH highlighted the need to use QSAR models in order to reduce experimental and animal testing, the OECD (2004a, b) has defined the validation principles for QSAR models, for regulatory purposes (Commission of the European Communities, 2003). Depending on the size of the dataset validation of the models were performed using leave-a-group-out internal validation by employing a percentage of datapoints as external validation sets.

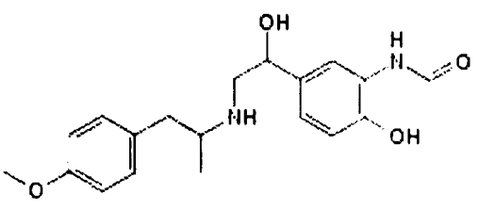
3. Effects of Solvents on Skin Permeation of Formoterol Fumarate

The aim of this investigation was to develop QSAR models for the effect of solvents on the skin penetration behaviour of a model penetrant. Such a model may identify the mechanisms involved in the penetration through skin and the effects of formulation factors. Kakubari et al. (2006) have measured the effect of several liquids, some of which are known penetration enhancers, on the penetration of Formoterol Fumarate (FF) through excised rat skin. In the experiment, FF dissolved in 4 ml of different solvents containing 2µl of ethanol was used as the donor phase in the diffusion cell. The receiver cells were filled with saline and solutions in both cells were stirred using magnetic stirrers. This dataset was selected for QSAR model development due to the consistency of the experiments and availability of results for a large number of solvents that is required for a QSAR study.

The model penetrant in this study was FF (Figure 3.1), a catecholamine analogue that possesses β_2 adreoceptor agonist potential and compliments it with high potency and prolonged efficacy (Bartow and Brogden, 1998; Faulds et al., 1991). Clinical studies have proved that FF causes bronchodilation for at least 12 hours after a single oral administration. Thus, in order to maintain a 24 hour inhibition of bronchoconstriction and effective plasma concentrations leading to suppression of asthmatic fits in asthma patients, transdermal drug delivery has been considered as a highly efficient method (Kakubari et al., 2006).

In this *in vitro* study the results are reported by Kakubari et al. (2006) as the amount permeated at different time intervals. The results showed a high enhancement from terpenes, fatty acid esters and higher alcohols while on the other hand there was no significant influence from lower alcohols, pyrrolidones and amines. More specifically the cineole/*N*-methyl -2-pyrrolidone (NMP) mixture enhanced slightly more the skin permeation of FF than with cineole alone. *l*-menthol/NMP mixed solvent system caused significant further increase. The maximum skin permeation of FF was detected when the ratio of l-menthol/NMP was 60/40 w/w. As an outcome of the *in vitro* study, l-menthol/NMP and isopropyl myristate (IPM)/NMP mixed

solvent systems were suggested as effective systems for augmenting skin permeation of FF, with potential applications in transdermal delivery of the drug. QSARs was used to relate the penetration changing behaviour of the solvents to their molecular structures as measured by molecular descriptors.



Log P = 2.03
MW = 344.4
pK_a = 8.95 (basic)

Figure 3.1. Molecular structure of formoterol fumarate

3.1. Methods

3.1.1. The dataset

The dataset consisted of the cumulative amount of FF that penetrated through abdominal rat skin at different time intervals of 2, 4, 6, 8 and 24 h as reported by Kakubari et al. (2006). The flux was calculated from the original data as the slope of the best fit line of the linear section of the plot between the cumulative amount absorbed and time. The donor phase was the same concentration of the drug (12.5 μg/ml) dissolved in the thirty-eight different solvents including terpenes, alcohols, and fatty acids. The volume applied was 4 ml. The rat skin used was full thickness, without pre-hydration, and non-occluded. Table 3.1 shows the dataset used in this study.

3.1.2. Structural descriptors

Structural descriptors included electronic parameters such as the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) energies, dipole moment, calculated by Vamp (using the AM1 Hamiltonian) using TSAR 3D, and atom and group counts, molecular weight and surface area and

volume calculated by TSAR 3D. Log P was calculated by the ACD/lab logD Suite 7.05 release (ACD/LogD, 2008).

3.1.3. Development and Validation of QSARs

Stepwise regression analysis was used to develop models for the cumulative amount penetrated at different time intervals, the flux and the enhancement ratio (ER) calculated as below:

$$ER_{solvent\ 1} = formoterol\ flux\ from\ solvent\ 1 / formoterol\ flux\ from\ water$$

All statistical analyses were performed in the MINITAB Statistical software version 13.20 (Minitab Statistical Software, 2008). Stepwise regression analyses were cut short to allow not more than 3-4 descriptors in the models. All the model parameters were significant with $P<0.05$.

The predictability of the models was examined by a leave-a-group – out procedure. As such, chemicals were sorted according to the ascending flux values; for each set of 4 solvents, the first compound was allocated to group a, the second to group b, the third to group c and the fourth to group d. This ensured that each group covered similar ranges of the skin penetration parameters. The regression was carried for the chemicals in groups a, b and c (as the training set), and the resulting equation was used to calculate the skin penetration parameter for the remaining group d (as the test set). The procedure was carried on to leave one group out at a time (all the possible combinations of groups making the training set). The Mean Absolute Error (MAE) of prediction was used as the error measure for assessing the validity of the models and was calculated as below:

$$MAE = \frac{\sum|observed - predicted|}{n}$$

Where observed and predicted in this case are log flux values obtained from the literature and that calculated by the models, and n is the number of data points.

Table 3.1. Dataset used for the QSAR study; Q2-Q8 are the cumulative amount permeated at time intervals of 2-8 h ($\mu g/cm^2$), and flux ($\mu g/cm^2/h$) of formoterol.

Vehicle	Q2	Q4	Q6	Q8	Q24	Equation	flux	group
Butyl myristate	0.8	2.6	3.9	4.6	10.2	y = 0.635x - 0.2	0.635	a
Cetyl isooctanate	0.1	0.3	0.8	1.2	3.6	y = 0.19x - 0.35	0.19	a
Cineole	8.4	21.8	30.1	34.5	48.2	y = 4.33x + 2.05	4.33	a
Diisopropyl sebacate	0	0.5	1.2	2.3	10.5	y = 0.38x - 0.9	0.38	a
Ethanol (50%)	0	0	0	0	0	y = 0	0	a
Ethyl Linoleate	0.5	2.3	6	8.8	17.1	y = 1.43x - 2.75	1.43	a
N-Methyl-2-Pyrrolidone	0.1	0.2	0.4	0.6	2.4	y = 0.085x - 0.1	0.085	a
Oleic acid	0	0.1	0.2	0.2	1	y = 0.035x - 0.05	0.035	a
Propylene glycol	0	0	0	0	0	y = 0	0	a
Terpinolene	2.1	6	7.9	8.7	13.7	y = 1.085x + 0.75	1.085	a
α-Terpineol	0.1	0.4	1.8	2.4	9.6	y = 0.415x - 0.9	0.415	b
Diethyl Phthalate	0	0.1	0.2	0.3	0.8	y = 0.05x - 0.1	0.05	b
d-Limonene	1.4	3.1	4.8	5.5	12.1	y = 0.7x + 0.2	0.7	b
Formamide	0	0	0	0	0	y = 0	0	b
l-Decanol	0.7	2.8	5.4	7.5	17.8	y = 1.15x - 1.65	1.15	b
Linalool	0.5	3.1	6.2	9.4	23.1	y = 1.49x - 2.65	1.49	b
Liquid paraffin	0	0	0	0.1	0.2	y = 0.015x - 0.05	0.015	b
Menthone	2.6	13.9	22.6	28.6	36.7	y = 4.335x - 4.75	4.335	b
Octyldodecanol	0.1	0.3	0.5	0.6	1.9	y = 0.085x - 0.05	0.085	b
Oleyl alcohol	0.1	0.3	0.9	1.3	4.9	y = 0.21x - 0.4	0.21	b
α-Terpinene	5.6	11.8	15.8	18.3	26.8	y = 2.105x + 2.35	2.105	c
Decyl oleate	0.1	0.2	0.5	0.7	2.4	y = 0.105x - 0.15	0.105	c
Diisopropyl adipate	0.1	0.3	0.9	1.7	12.3	y = 0.27x - 0.6	0.27	c
dl-Menthol	0.2	0.9	1.9	2.7	8.5	y = 0.425x - 0.7	0.425	c
Ethyl oleate	0.5	1.6	3	4.8	11.8	y = 0.715x - 1.1	0.715	c
Hexadecyl isostearate	0	0.1	0.2	0.3	0.8	y = 0.05x - 0.1	0.05	c
Isopropanol	0	0	0	0	0	y = 0	0	c
Isopropyl myristate	0.4	2.5	5.5	7.8	16.4	y = 1.26x - 2.25	1.26	c
Oleyl oleate	0	0.1	0.1	0.1	0.4	y = 0.015x + 9E-17	0.015	c
Diethyl sebacate	0.3	1.9	4.7	8.1	22.4	y = 1.31x - 2.8	1.31	d
Glyceryl triisooctanate	0.1	0.4	0.8	0.8	3.6	y = 0.125x - 0.1	0.125	d
Isopropyl palmitate	0.5	1.2	2.6	3.2	6.3	y = 0.475x - 0.5	0.475	d
l-Dodecanol	0.1	0.4	1.2	2.3	10.6	y = 0.37x - 0.85	0.37	d
Linalyl acetate	0.6	7	15.4	23.5	45.3	y = 3.855x - 7.65	3.855	d
l-Octanol	0.8	2.6	4.3	5.9	15.8	y = 0.85x - 0.85	0.85	d
N,N-Dimethyl-3-toluamide	0	0	0	0	0	y = 0	0	d
Saline	0	0.1	0.2	0.3	1.2	y = 0.05x - 0.1	0.05	d
n-Butyl alcohol	0	0.1	0.2	0.2	3.3	y = 0.035x - 0.05	0.035	d

3.2. Results

The histogram of Q and log Q at different time intervals showed that log Q has a less skewed normal distribution than Q. However, due to Q values being zero for many solvents, especially at smaller time intervals, which indicates a lag time, it was not possible to calculate the log Q for many data points, leading to the unnecessarily exclusion of data from the analysis. Therefore, a small value of 0.01 was added to all the Q values and the logarithm of those (log Q+0.01) were used for the model development. Models were developed for amount permeated at different time intervals, as the efficiency of a particular solvent system may differ at different time intervals and some solvents may work faster. Moreover, QSAR models were also developed for flux and log ER using stepwise regression analysis. Equations (3.1)-(3.4) below are the regression models obtained for the cumulative amount permeated from rat skin at various time intervals. Equations (3.5) and (3.6) are the QSAR models for log flux and log ER.

Log Q4+0.01 = - 2.01 + 3.60 6-aliphatic rings + 0.392 Balaban + 0.0831 Total Lipole – 35.3 Chiv6_r (3.1)
S = 0.6684 r^2 = 56.6% F = 10.76 P = 0.000 N =38

Log Q6+0.01 = - 0.703 + 1.20 Chiv4_pc + 0.140 Total Lipole – 1.20 6-aromatic rings – 0.260 Dipole (3.2)
S = 0.7091 r^2 = 57.8% F = 11.30 P = 0.000 N = 38

Log Q8+0.01 = - 0.409 + 1.15 Chiv4_pc + 0.138 Total Lipole – 0.329 Dipole – 1.10 6-aromatic rings (3.3)
S = 0.7207 r^2 = 56.8% F = 10.85 P = 0.000 N = 38

Log Q24+0.01 = 0.112 + 0.185 Chiv4_pc + 0.298 log P – 12.4 Chiv10_p – 1.76 6-aromatic rings (3.4)
S = 0.6808 r^2 = 66.4% F = 16.30 P = 0.000 N = 38

Log flux = 8.51 – 6.74 Chi9_p + 0.233 Chi0 – 7.06 Index of refraction (3.5)
S = 0.498 r^2 = 49.3 % F = 11.01 P = 0.000 N = 38

Log ER = 11.0 – 6.98 Chi9_p + 0.236 Chi0 – 7.87 Index of refraction + 0.0388 Total Lipole (3.6)
S = 0.484 r^2 = 53.6% F = 9.53 P = 0.000 N = 38

The parameters of equations 3.1-3.6 have been explained in Table 3.2. The statistical parameters of these regression equations were S, r^2, F, P, and N. Standard deviation (S) is a measure of spread of data, sensitive to shape of distribution, and shows how widely spread the values in the data set are from the regression line. r^2 is the square of the correlation coefficient, the closer the value is to unity, the higher the percentage of variation in y is explained by the equation. Fisher's statistic (F or F-statistic) is calculated for the regression. Probability error (P or P-value) is simply the probability that the hypothesis being tested is true. The hypothesis is that there is a less than one in twenty (0.05%) chance that there is no significant relationship. N is the number of data points.

Table 3.2. Descriptors selected by stepwise regression analysis in each equation and the step in the analysis in which they were selected

Descriptor	Eq.3.1	Eq.3.2	Eq.3.3	Eq.3.4	Eq.3.5	Eq.3.6	Description
6-aliphatic rings	1						Number of 6-membered aliphatic rings
Balaban	2						Balaban topological index
Total Lipole	3	2	2			4	Total lipole moment
Chiv6_r	4						6^{th} order valence corrected ring molecular connectivity index
Chiv4_pc		1	1	1			4^{th} order valence corrected path/ cluster molecular connectivity index
6-aromatic rings		3	4	4			Number of 6-membered aromatic rings
Dipole		4	3				Dipole moment
Log P				2			Logarithm of octanol/ water partition coefficient
Chiv10_p				3			10^{th} order valence corrected path molecular connectivity index
Chi9_p					1	1	9^{th} order simple path molecular connectivity index
Chi0					2	2	zero order simple molecular connectivity index
Index of refraction					3	3	Index of refraction calculated by ACD/labs Software

Looking at the statistical fit of equations 1-6 and judging by the r^2 values it can be seen that the QSAR for log Q24 has the best fit (equation 3.4). The predictive power of this equation was examined by the procedure of the leave-group-out (as explained

in section 3.1). Figure 3.2 shows the graph between flux and the corresponding absolute error values. It can be seen that for the majority of vehicles the absolute error is below 1. The solvents with large error of prediction can be seen in Table 3.3 where the absolute error values are listed.

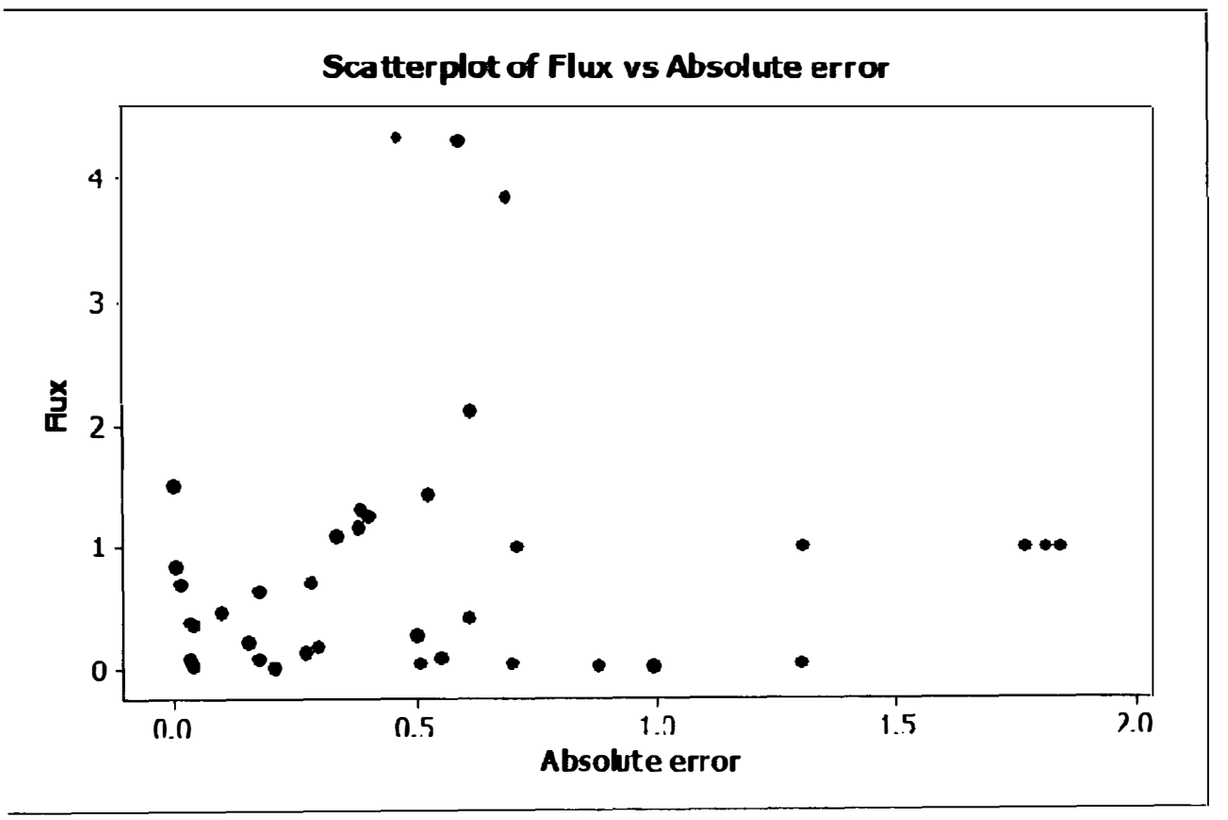


Figure 3.2: Graph between flux and the corresponding Absolute Error values

Table 3.3. Descriptor values used in the QSAR models (equations 3.1-3.6)

	6-aliphatic rings	Wiener	Balaban	VAMP Heat of Formation	Chiv6_r	Total Lipole	Chi9_p
Butyl myristate	0	1265	3.05	-197.03	0.00	8.03	0.34
Cetyl isooctanate	0	2783	3.10	-235.47	0.00	8.40	0.51
Cineole	2	128	1.87	-70.76	0.13	0.84	0.00
Diisopropyl sebacate	0	1184	3.24	-248.00	0.00	0.20	0.30
Ethanol (50%)	0	4	1.63	-62.65	0.00	1.31	0.00
Ethyl Linoleate	0	1720	3.00	-152.22	0.00	12.79	0.41
N-Methyl-2-Pyrrolidone	0	40	2.08	-40.24	0.00	1.46	0.00
Oleic acid	0	1313	2.92	-182.40	0.00	8.84	0.36
Propylene glycol	0	18	2.54	-116.83	0.00	1.43	0.00
Terpinolene	1	120	2.01	-9.16	0.05	0.28	0.00
α-Terpineol	1	152	2.15	-73.83	0.06	2.49	0.00
Diethyl Phthalate	0	446	2.40	-148.95	0.03	5.56	0.23
d-Limonene	1	120	2.01	-4.21	0.06	0.57	0.00
Formamide	0	4	1.63	-44.75	0.00	0.55	0.00
l-Decanol	0	220	2.69	-117.35	0.00	7.04	0.09
Linalool	0	180	3.38	-50.02	0.00	1.34	0.00
Liquid paraffin	0	560	2.81	-106.38	0.00	0.05	0.21
Menthone	1	150	2.17	-80.49	0.06	1.23	0.00
Octyldodecanol	0	1380	3.28	-180.71	0.00	2.62	0.42
Oleyl alcohol	0	1140	2.88	-142.83	0.00	11.15	0.34
α-Terpinene	1	120	2.01	-8.10	0.04	0.06	0.00
Decyl oleate	0	4308	3.12	-196.70	0.00	3.63	0.67
Diisopropyl adipate	0	574	3.38	-220.50	0.00	0.00	0.25
dl-Menthol	1	150	2.17	-97.99	0.07	1.82	0.00
Ethyl oleate	0	1720	3.00	-181.58	0.00	12.55	0.41
Hexadecyl isostearate	0	6836	3.14	-297.03	0.00	0.73	0.81
Isopropanol	0	9	2.32	-69.52	0.00	1.04	0.00
Isopropyl myristate	0	1072	3.06	-187.45	0.00	9.13	0.30
Oleyl oleate	0	8816	3.13	-261.70	0.00	2.77	0.92
Diethyl sebacate	0	1016	3.23	-242.76	0.00	2.82	0.28
Glyceryl triisooctanate	0	4180	4.40	-384.81	0.00	6.48	0.72
Isopropyl palmitate	0	1462	3.05	-201.14	0.00	10.68	0.37
l-Dodecanol	0	364	2.76	-131.04	0.00	8.42	0.15
Linalyl acetate	0	341	3.86	-82.78	0.00	3.30	0.00
l-Octanol	0	120	2.60	-103.66	0.00	5.66	0.00
N,N-Dimethyl-3-toluamide	0	197	2.16	-12.79	0.03	6.46	0.00
Saline	0	0	0.00	-59.24	0.00	0.22	0.00
n-Butyl alcohol	0	20	2.19	-76.30	0.00	2.87	0.00

Table 3.3. Continued

	Kappa1 index	Chiv4 _PC	VAMP Total Energy	Chi5_P	Chi0	Index of refraction	Absolute error
Butyl myristate	20.00	0.08	-3445.59	2.00	14.89	1.44	0.178
Cetyl isooctanate	26.00	0.37	-4380.48	2.72	19.30	1.45	0.297
Cineole	7.64	2.41	-1849.82	2.53	8.11	1.46	0.59
Diisopropyl sebacate	20.00	0.39	-3746.78	1.71	15.38	1.44	0.033
Ethanol (50%)	3.00	0.00	-659.79	0.00	2.71	1.35	1.843
Ethyl Linoleate	22.00	0.08	-3700.10	2.15	16.31	1.47	0.529
N-Methyl-2-Pyrrolidone	5.14	0.44	-1291.77	0.47	5.28	1.47	0.175
Oleic acid	20.00	0.05	-3417.64	1.94	14.89	1.47	0.877
Propylene glycol	5.00	0.08	-1136.36	0.00	4.28	1.43	1.305
Terpinolene	8.10	1.05	-1501.15	1.45	7.56	1.48	0.338
α-Terpineol	9.09	1.60	-1849.95	1.52	8.48	1.48	0.61
Diethyl Phthalate	14.06	0.51	-3011.08	2.70	11.97	1.51	0.702
d-Limonene	8.10	1.01	-1500.94	1.45	7.56	1.47	0.016
Formamide	3.00	0.00	-697.12	0.00	2.71	1.37	1.768
l-Decanol	11.00	0.00	-1906.46	0.85	8.36	1.44	0.381
Linalool	11.00	0.82	-1848.92	0.61	8.91	1.46	0.005
Liquid paraffin	15.00	0.00	-2364.99	1.35	11.19	1.43	0.993
Menthone	9.09	1.36	-1850.24	1.46	8.43	1.44	0.464
Octyldodecanol	21.00	0.38	-3464.59	2.21	15.60	1.45	0.033
Oleyl alcohol	19.00	0.00	-3124.56	1.85	14.02	1.46	0.155
α-Terpinene	8.10	1.02	-1501.11	1.45	7.56	1.48	0.615
Decyl oleate	30.00	0.14	-4810.50	3.25	21.96	1.46	0.554
Diisopropyl adipate	16.00	0.39	-3123.43	1.21	12.55	1.43	0.503
dl-Menthol	9.09	1.52	-1878.31	1.46	8.43	1.46	0.61
Ethyl oleate	22.00	0.08	-3728.68	2.15	16.31	1.46	0.286
Hexadecyl isostearate	35.00	0.37	-5783.00	3.84	25.66	1.46	0.51
Isopropanol	4.00	0.00	-815.62	0.00	3.58	1.38	1.813
Isopropyl myristate	19.00	0.19	-3289.64	1.72	14.35	1.44	0.4
Oleyl oleate	38.00	0.08	-6193.45	4.25	27.62	1.47	0.205
Diethyl sebacate	19.00	0.43	-3591.01	1.62	14.51	1.46	0.385
Glyceryl triisooctanate	33.00	1.26	-6073.68	3.99	25.06	1.46	0.271
Isopropyl palmitate	21.00	0.19	-3601.31	1.97	15.76	1.44	0.097
l-Dodecanol	13.00	0.00	-2218.13	1.10	9.78	1.44	0.041
Linalyl acetate	14.00	1.01	-2452.79	1.14	11.19	1.45	0.69
l-Octanol	9.00	0.00	-1594.79	0.60	6.95	1.43	0.005
N,N-Dimethyl-3-toluamide	10.08	0.69	-1986.33	1.77	9.14	1.53	0.712
Saline	1.00	0.00	-348.56	0.00	0.00	1.33	1.301
n-Butyl alcohol	5.00	0.00	-971.45	0.00	4.12	1.40	0.042

3.3. Discussion

In Q4 hours (equation 3.1) there is a positive correlation between the response and the number of 6-membered aliphatic rings, which can be translated as, the more 6-aliphatic rings in the molecular structure of the solvents the higher the penetration after 4h. The solvents that have 6-membered aliphatic rings in the dataset belong to the group of terpenes; these are cineole, terpinolene, α-terpineol, d-limonene, menthone, α-terpinene, dl-menthol. Cineole contains two 6-membered aliphatic rings and the other terpenes have one each. All terpenes in this study have 10 carbon atoms with similar molecular weight, 136 Da as minimum and 156 Da as maximum. As an exception linalool is a terpene with molecular weight of 154.24; which does not contain a ring. The molar refractivity of the above terpenes is also quite close with 45.35 m^3 mol^{-1} as minimum and 47.83 m^3 mol^{-1} as maximum. Their molar volume is similarly close as well with 159.6 m^3/mol as minimum and 175.5 m^3/mol as maximum. Polarizability of the drugs is very close with 17.98 cm^3 as minimum and 18.96 cm^3 as maximum. Terpenes as vehicles result in FF having an average flux of 1.91 which is higher than the average of all the other data points at 0.65. From all the above we can assume that terpenes have high enhancement activity which could be due to any of the structural or physico-chemical properties selected by stepwise regression explained in Table 3.2. Terpenes have been shown to increase skin permeation by interacting with stratum corneum lipids and/or keratin and increasing the solubility of drug into the stratum corneum lipids (Williams and Barry, 1991). Therefore, terpenes with different chemical structures will be expected to have different activities in terms of the effects on the skin penetration of a chemical. Although all but one of the terpenes in this list have equal number of 6-membered aliphatic rings, the remaining three descriptors of equation 3.1 will be able to differentiate different potencies of the terpenes. In order to affect the drug permeability, solvents will need to penetrate the SC. Kang et al. (2007) determined permeability coefficient of different terpenes experimentally using human skin. Their results suggested that liquid terpenes tend to produce better enhancing effects than solid terpenes and triterpenes (C_{30}), while tetraterpenes (C_{40}) generally had poor penetration effect than other terpenes. The terpenes in this investigation contained 10 carbon atoms (monoterpenes).

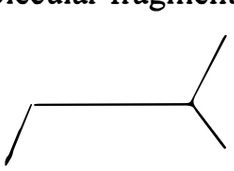
The balaban index, a highly discriminating topological index whose values does not substantially increase with the molecular size and represents extended connectivity and the shape of molecules (Thakur et al., 2004), is the second best descriptor in equations 3.1. Balaban topological index shows a positive correlation with the Q4 hours. The enhancer glyseryl triisooctanate has the highest balaban topological index with a value of 4.40 and saline has the lowest balaban topological index with a value of 0. The terpenes with 6-aliphatic rings have relatively average values, 1.87 minimum and 2.16 as maximum. The balaban topological descriptor is a molecular size dependant descriptor (Todeschini and Consonni, 2000) meaning that the higher the molecular size, the higher the balaban values become, resulting in a higher penetration of FF after 4h. For this dataset, the correlation between balaban and molecular weight has an r^2 value of 51.4%. In terms of terpenes, due to the equal number of carbon atoms, the difference between the molecular weights will be attributed to the presence of oxygen atoms, cyclicity and unsaturation. It has been shown that terpenes with aldehyde and ester functional groups are often better enhancers than the hydrocarbon terpenes (Vaddi, et al., 2002; Williams and Barry, 2004). Furthermore, smaller alcoholic terpenes with a higher degree of unsaturation appeared to be good candidates for enhancing the permeation of hydrophilic drugs (Ghafourian et al., 2004). Oxygen containing terpenes have been found to be more effective than hydrocarbon terpenes, with those containing a bicyclic structure displaying a lesser permeation enhancing effect (Fang et al., 2007). It has also been reported that terpenes with a minimal degree of unsaturation, like menthol and cineole, are good absorption promoters for polar and water-soluble drugs (Jain et al., 2002).

The descriptor total lipole is the second best descriptor in both equations 3.2 and 3.3, the third best descriptor of equation 3.1 and the fourth best descriptor of equation 3.6. Total lipole has positive correlation with the Q4, Q6, Q8 and Enhancement Ratio. From table 3.2 ethyl linoleate has the highest total lipole value 12.78 and diisopropyl adipate the minimum value 0. Total lipole shows lipophilicity of the molecule in a specific direction. For example, surfactants are expected to have high total lipole values and they are known enhancers of drug skin penetration (Ma et al., 2007). According to equations 3.1-3.3 and 3.6, such a lipolar property for the solvent is a contributing factor to the increase of skin absorption of FF. Oxygen containing

terpenes have lipolar (surfactant-like) property, which will increase their enhancer activities according to the QSAR equation. Evidence from DSC (differential scanning calorimetry) and ATR-FT-IR (attenuated total reflectance Fourier transform infrared spectroscopy) analysis revealed that the 1,8 cineole and menthol enhanced permeation of zidovudine by transforming stratum corneum lipids from a highly ordered orthorhombic perpendicular subcellular packing to a less ordered, hexagonal subcellular packing. Both terpenes show effects on both lipid alkyl tails and polar head groups, suggesting that the above terpenes mainly act at polar head groups and break interlamellar and intralamellar hydrogen bonding networks (Narishetty and Panchagnula, 2005). This mechanism is also shared with some of the suggested mechanisms of penetration enhancements by surfactants (Aungst et al., 1986; Barry, 1987). Accordingly, such amphipathic molecules can align themselves with the polar head groups close to the polar head groups in the lipid bilayers and have their alkyl chains descending into the hydrophobic region of the bilayer.

Chiv6_r is the fourth best descriptor in equation 3.1 with a high negative correlation with the response 'Log Q4+0.01'. The enhancers that have high Chiv6_r values are all the terpenes containing aliphatic rings plus enhancers diethyl phthalate and N,N-dimethyl-m-toluamide with diethyl phthalate and N,N-dimethyl-m-toluamide having the lowest value of 0.028 and cineole having the highest value. Here it can be assumed that every enhancer with 6-aliphatic rings has a non-zero Chiv6_r as well, but every enhancer that has Chiv6_r does not have 6-aliphatic rings (the six-membered ring may be aromatic). Therefore, it seems that solvents containing aromatic rings will lead to lower enhancement through the skin.

Chiv4_pc index is the best descriptor in equations 3.2, 3.3 and 3.4 having a positive correlation with all three responses. Chiv4_pc indicates the presence of following molecular fragments in the molecule:



Cineole has the highest Chiv4_pc index of 2.41. From Table 3.3 it is very interesting to see that the group of terpenes with 6-membered aliphatic rings in addition to

glyceryl triisooctanate and linalyl acetate, are the only solvents with Chiv4_pc values above 1, while the rest of the solvents in most cases have values well below 1.

The presence of a 6-membered aromatic ring is the third best descriptor in equation 3.2 and the fourth best descriptor in equations 3.3 and 3.4. It has negative correlation with responses, 'Log Q6+0.01', 'Log Q8+0.01' and 'Log Q24+0.01'. The solvents diethyl phthalate and N,N-dimethyl-m-toluamide are the only compounds with 6-membered aromatic rings. It is interesting that these two solvents have a lot in common such the high dipole moment values, 3.82 and 3.85 respectively, the highest Index of Refraction values, high surface tension and density values. In line with this, total dipole moment is the third best descriptor of equation 3.3 and the fourth best descriptor of equation 3.2 with a negative correlation with the response. Enhancer diisopropyl adipate has the lowest dipole moment value of 0.00002 and the enhancer N-methyl-2-pyrrolidone has the highest (3.91). This could be due to a better partitioning of solvents with low dipole moment into the SC, carrying along the drug and hence increasing its penetration into the skin.

Log P is the second best descriptor of equation 3.4. It has a positive correlation with the response (Log Q24+0.01). Solvents propylene glycol and oleyl oleate have the lowest and highest log P values, -6.44 and 16.67 respectively. This correlation shows that the more lipophilic the vehicles are, the higher the amount of the drug absorbed within 24h of application. Better penetration from lipophilic solvents could be attributed to the effect of such solvents on the SC. For example, previous studies have shown that terpenes with larger log P values are more effective enhancers than those with lower log P values, as it is easier for lipophilic terpenes to mix with the SC intercellular lipids (for extraction or for lipid transition) (Narishetty and Panchagnula, 2004; Ghafourian et al., 2004; Kang et al., 2007; Williams and Barry, 2004).

Chiv10_p is the third best descriptor of equation 3.4. It has a negative correlation with the response (Log Q24+0.01). Cineole, ethanol (50%), N-methyl-2-pyrrolidone, propylene glycol, terpinolene, α-terpineol, d-limonene, formamide, linalool, menthone, α-terpinene, dl-menthol, isopropanol, linalyl acetate, l-octanol, N,N-dimethyl-m-toluamide, saline, n-butyl alcohol have a value of 0 for Chiv10_p.

Enhancer Hexadecyl isostearate has the highest Chiv10_p value (0.41). Chiv10_p indicates the presence of long (10-atom) path fragments (not-branched). Equation 3.4 with a positive coefficient for log P and negative coefficient for Chiv10_p shows that the more lipophilic solvents that are not linear (non-branched) increase the absorption of FF.

Chi9_p is the best descriptor in equations 3.5 and 3.6. Chi9_p is the 9^{th} order path molecular connectivity index indicating the presence of 9 atom chains (Hall and Kier, 1991). Chi9_p has a negative correlation with the log flux and log ER. Terpenes with 6-aliphatic rings, along with the enhancers ethanol (50%), N-methyl-2-pyrrolidone, propylene glycol, formamide, linalool, isopropanol, linalyl acetate, l-octanol, N,N-dimethyl-m-toluamide, saline, n-butyl alcohol have a Chi9_p value of 0. Oleyl oleate has the highest Chi9_p value of 0.915712. According to this relationship, solvents with long linear (non-branched) chain are not good enhancers. This is in accordance with the finding from equation 3.4 as explained before.

Chi0 is the second best descriptor in equations 3.5 and 3.6 and has a positive correlation with responses log flux and log ER. Saline has a 0 value and the maximum value has been observed with oleyl oleate. Chi0 is strongly related to molecular size, the higher the molecular weight the more the Chi0 is (Hall and Kier, 1986). Chi0 can also be an indicator of lipophilicity in specific series of compounds, as growing atom numbers indicates growing number of carbon atoms which increase lipophilicity. Therefore, joint presence of Chi9_p and Chi0 in equations 3.5 and 3.6 is in accordance with presence of Chiv10_p and log P in equation 3.4.

The Index of refraction is the third best descriptor in equations 3.5 and 3.6; it also has a negative correlation with the response in both equations. Saline has the lowest value of 1.329 and N,N-dimethyl-m-toluamide has the highest observed value of the dataset (1.526). Index of refraction is an indicator of molecular polarisability; it shows polarisable molecules such as N,N-dimethyl-m-toluamide are not good promoters of skin penetration. This effect is similar to the effect of 6-aromatic rings in equations 3.2 3.3 and 3.4 and probably reflects the higher partitioning rates of FF from less polar solvents.

In the *in vitro* study of Kakubari et al. (2006), remarkable enhancement was noted with terpenes, fatty acids and higher alcohols whereas no significant influence was observed in the case of lower alcohols, pyrrolidones and amines. In the case of terpenes the greatest permeation of FF was seen with cineole, menthone and linalyl acetate after 8h. According to the QSAR estimation, terpenes (when used as the vehicles) are the most efficient ehancers of FF absorption, as indicated by regression equations 3.1, 3.2, 3.3 and 3.4.

The equations give useful information regarding the effect of structural characteristics on the penetration altering behaviour of the vehicles. However, the statistical parameters show an average fit and therefore it is expected to obtain a rather poor predictive power. Table 3.3 and Figure 3.2 show that the absolute errors are acceptable for majority of vehicles. Unfortunately, the outliers cannot be explained sufficiently by their molecular properties. Without exclusion of outliers, the mean absolute error was calculated to be 0.535.

3.4. Conclusion

This study showed that when specific terpenes, esters and higher alcohols are used as vehicles they can enhance substantially the penetration of FF through rat skin with the enhancement effect quantitatively related to specific properties of their molecular structures. It became evident that QSAR analysis can help identify those specific molecular characteristics for the vehicles that will aid and promote the absorption of a specific penetrant. Solvents with a lipolar (surfactant like) structure increase drug penetration at various time intervals, especially at the earlier times of 4, 6 and 8h. This effect is also evident in the enhancement ratio of the flux by solvents. According to the QSARs, lipophilic solvents are more effective in increasing penetration of the drug. This was shown by the presence of descriptors such as partition coefficient and zero order connectivity index in the QSARs. On the other hand, linear molecules (non-branched molecules with high Chiv10_p and Chi9_p) have a negative contribution to the absorption rate of the drug. This was evidenced by the concomitant presence of lipophilicity descriptors (positive coefficients) with descriptors such as high order connectivity indices (negative coefficients). An

interesting conclusion can be drawn from the QSARs that presence of aliphatic rings either represented by the number of 6-membered aliphatic rings or by the fourth order valence corrected path/cluster connectivity index (Chiv4_pc) is a positive factor for the solvents leading to increased drug penetration at various time intervals of 4, 6, 8 and 24h. The chemicals with 6-membered aliphatic rings and therefore high Chiv4_pc are mainly those belonging to the group of terpenes, which are known for their penetration enhancement activities. Solvents containing aromatic rings (6-aromatic ring) have a negative effect on the permeation of FF. This has been indicated indirectly by the descriptor 'index of refraction' in some of the reported QSAR.

It must be noted however that the molecular characteristics of 'good' enhancers of a specific drug (FF here) may not be the same as the good enhancers of a different penetrant. Here is the scope for more study to compare QSARs obtained for the penetration changing behaviour of vehicles towards different penetrants with different molecular properties. It can be seen that QSAR methods can support *in vitro* and *in vivo* analysis.

4. Effects of terpenes on the skin permeation of Haloperidol

Terpenes are a class of naturally occurring hydrocarbons known as essential oils. They are widely used in transdermal drug delivery as enhancers of drug permeation through human skin. They have been reported to have good toxicological profiles, high percutaneous enhancement abilities and low cutaneous irritancy at low concentrations (Okabe et al., 1990). The human skin permeation effect of 49 terpenes and terpenoids (chemically modified terpenes by oxidation or rearrangement of their carbon skeleton) on the *in vitro* permeability coefficient of haloperidol has been reported by Kang et al. (2007).

The permeability coefficients were determined with standardised methods using human skin samples from the abdomen areas of three healthy female donors. A full spectrum of terpenes was selected to include monoterpene, sesquiterpene, diterpene, triterpene and tetraterpene with various functional groups ranging from hydrocarbons, alcohols, aldehydes, ester, ketones, to oxides (Figure 4.1). The dataset offers an opportunity for Quantitative Structure- Activity Relationship studies to investigate the molecular properties in terpenes that lead to high levels of skin permeation of haloperidol. The advantage of this dataset is consistency of the methods used for the measurement of permeability coefficients of the model drug haloperidol.

A selected set of physicochemical properties of terpenes was used as the predictor variables and the permeability coefficient (k_p) of the model drug, haloperidol, exposed to skin in vehicles containing different terpenes, was chosen as the response variable. Haloperidol is an anti-psychotic agent suitable for transdermal delivery and there is a clinical need to develop such a dosage form (Vaddi et al., 2002; Whitehead, 1975). By nature, it is a hydrophobic molecule with low molecular weight (Figure 4.2). The only long-lasting formulation is its ester, the haloperidol decanoate, for intramuscular injection, which, however, has disadvantages such as injection pain, marked inter-individual variation and complex administration regimen. It is important to develop an alternative for its maintenance therapy to

prevent the relapse of psychosis. The solvent, propylene glycol (PG), is commonly used in skin care products.

In this research the focus was to model the effect of chemical enhancers on the skin penetration of chemicals. Considering the huge number of possible combinations of chemicals that might be intended for delivery through skin or the skin might be exposed to unintentionally, a systematic method for characterizing the effects of formulation ingredients is of great interest.

4.1. Methods

4.1.1. The Dataset

The dataset consisted of the permeation coefficient of haloperidol (HP) through excised human skin with the aid of 49 terpenes and terpenoids listed in Figure 4.1. The data were generated by Kang et al. (2007) using standardised experimental protocols. The donor mixture was HP dissolved in propylene glycol with a concentration of 3 mg/ml containing 5% (w/v) of the respective terpene. The receptor solution was 500 ml of 0.03% (v/v) lactic acid solution containing 1% (v/v) antibacterial antimycotic solution. Human skin samples were taken from abdomen area of three healthy female donors. The skin samples were used as dermatomed (epidermis), pre-hydrated, under occluded conditions, and the donor phase volume applied on the skin was 1ml. *In vitro* experiments were performed using the same set of automated flow-through diffusion cells.

4.1.2. Molecular Descriptors

A total of 205 descriptors for 49 terpenes (Citral was obtained from the average of Citral-α and Citral-β) were calculated using the following software packages: ACD labs/LogD Suite 7.0.5 (ACD/LogD, 2008), TSAR 3D version 3.3 (TSAR 3D, 2008) and Symyx QSAR version 2.3.0.0.12 (MDL QSAR, 2009). The descriptors included

electronic descriptors such as the highest occupied molecular orbital (E_{HOMO}) and the lowest unoccupied molecular orbital (E_{LUMO}) energies and dipole moment calculated using the AM1 Hamiltonian, atom and group counts, molecular weight, topological indices, surface area, volume and atom-type electrotopological indices calculated by TSAR 3D and Symyx QSAR. The logarithm of the octanol-water partition coefficient (log P) was calculated by the ACD/labs logD suite.

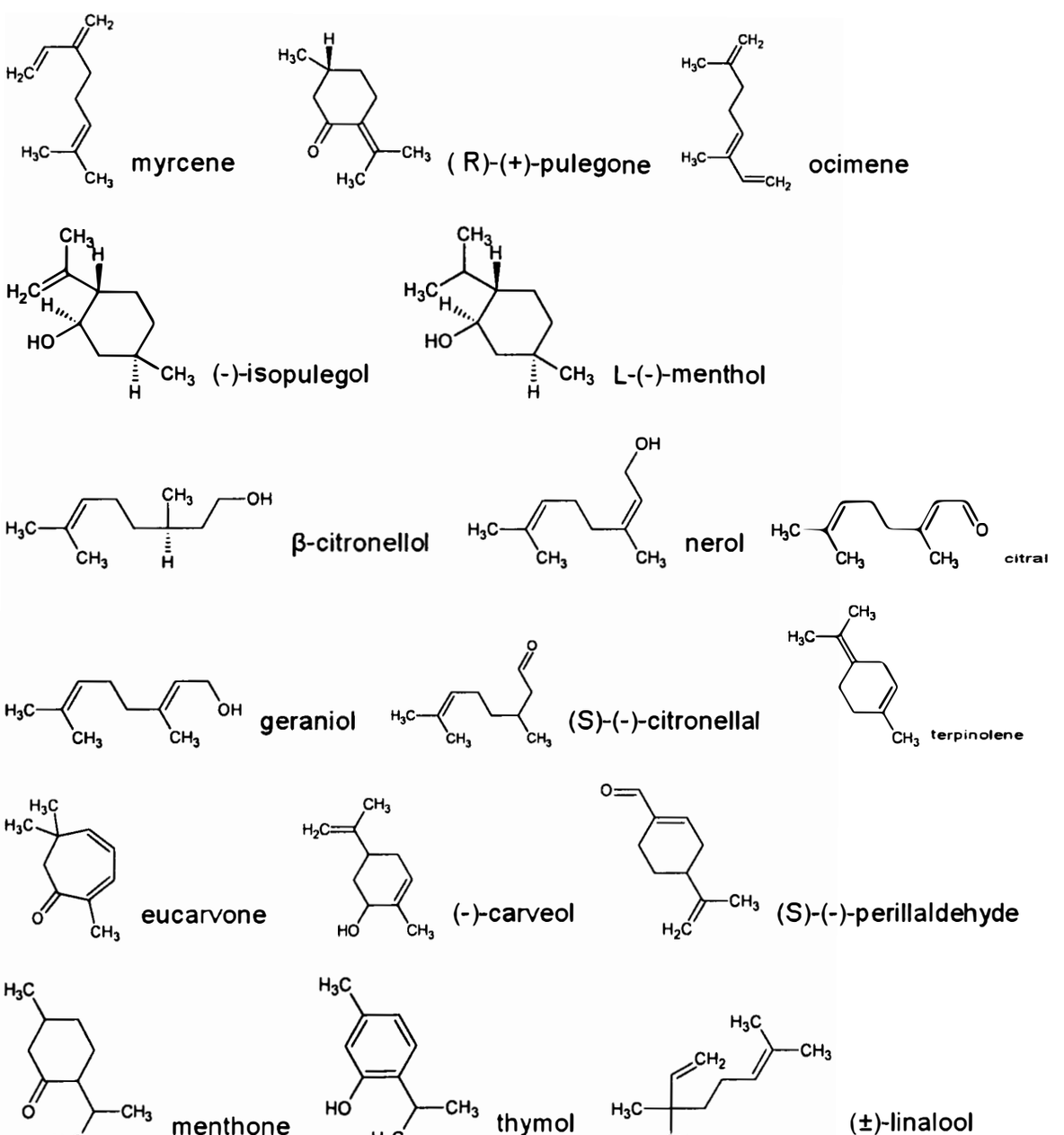
4.1.3. Development of QSAR models

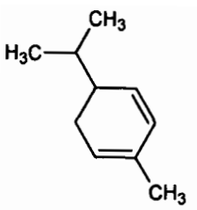
Stepwise regression analysis and Support Vector Machines (SVM) were used for the development of QSAR models. The enhancers were divided into training and test sets. QSAR models were developed using the enhancers in the training set and the models were used for the estimation of log k_p of haloperidol when chemicals in the test set were used as the enhancers. In order to divide the chemicals, data were sorted according to ascending log k_p values, and from each group of four, the first two chemicals were allocated to the training set, the third into the test set and the fourth into the training set. This was continued until all the chemicals were allocated into these two groups.

Stepwise regression analysis was performed with log k_p as the response against all molecular descriptors as the predictors. In order to obtain several other regression models, stepwise regression was repeated with the exclusion of some of the descriptors. In this way, three different regression models were developed and the prediction powers of the models were investigated.

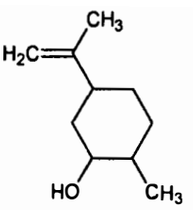
STATISTICA data mining version 8.0 (StatSoft, Inc., 2010) was used to carry out Support Vector Machine (SVM) analyses. In these analyses, log k_p was used as the continuous dependent and the descriptors selected by the three stepwise regression analyses or the feature selection tool in STATISTICA as the independent variables. Feature selection in STATISTICA is a data mining tool specifically designed to handle extremely large sets of continuous and/or categorical predictors for regression or classification type problems. It assumes no linear or monotone relationships between predictors and the dependent variables of interest, instead the software

applies a generalised 'notion of relationship' while screening the predictors one by one. In order to optimise the outcome, SVM types 1 and 2 (ε-SVM and ν-SVM), with different kernels of RBF (Radial Basis Function) and sigmoid were investigated. Several starting values for the capacity factor (*C*), ε and ν were investigated. Analyses were performed on the training set and the resulting models were used for the estimation of prediction error of the test set group. The best models were selected according to the lowest mean absolute prediction error for the test set group.

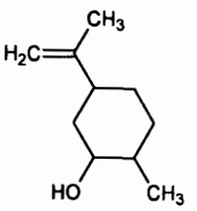




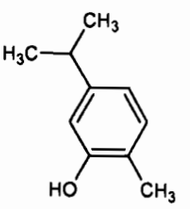
α-phellandrene



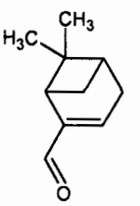
(-)-dihydrocarveol



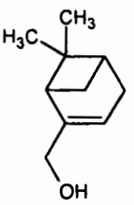
(+)-dihydrocarveol



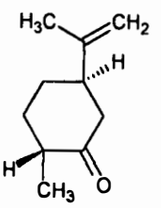
carvacrol



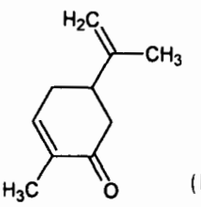
(1R)-(-)-myrtenal



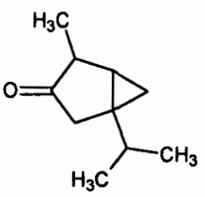
(1R)-(-)-myrtenol



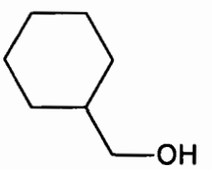
(+)-dihydrocarvone



(R)-(-)-carvone

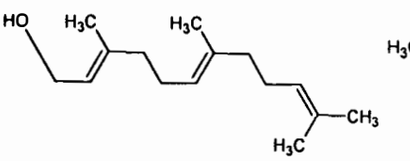


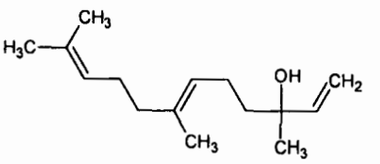
(-)-α-thujone



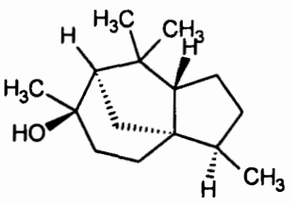
cyclohexanemethanol

farnesol

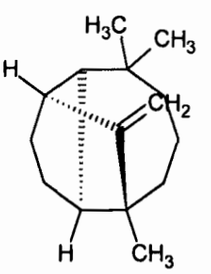




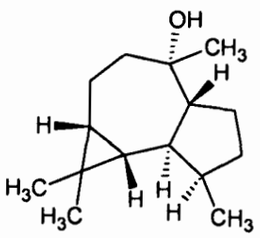
(±)-nerolidol



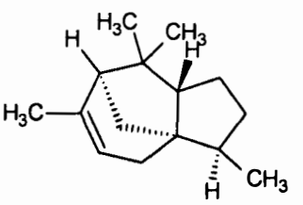
(+)-cedrol



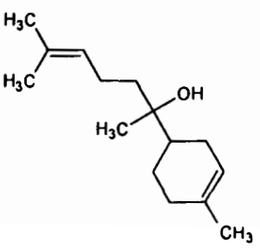
(+)-longifolene



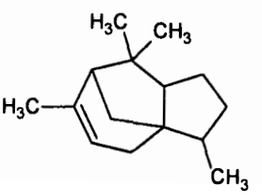
(-)-epiglobulol



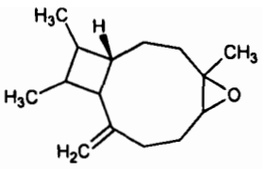
(-)-α-cedrene



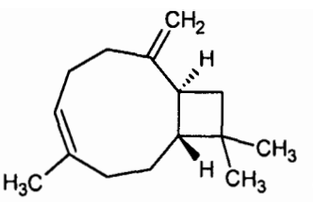
(±)-α-bisabolol



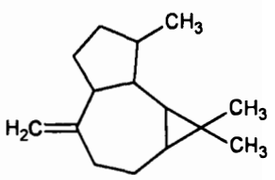
(+)-β-cedrene



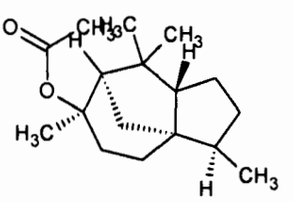
(-)-caryophyllene oxide



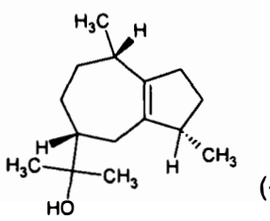
(-)-trans-caryophyllene



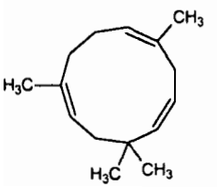
(+)-aromadendrene



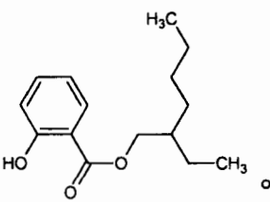
(+)-cedryl acetate



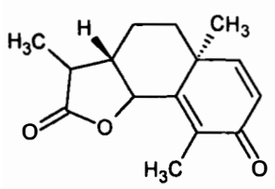
(-)-guaiol



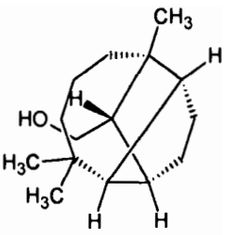
α-humulene



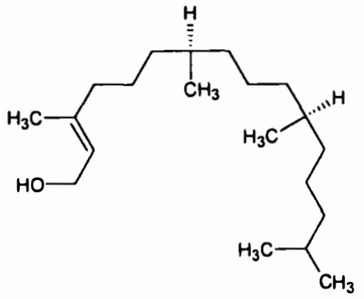
octisalate



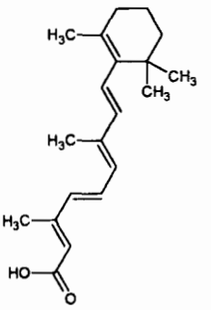
(-)-α-santonin



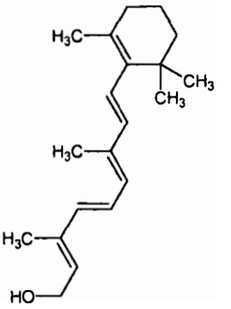
(-)-solongifolol



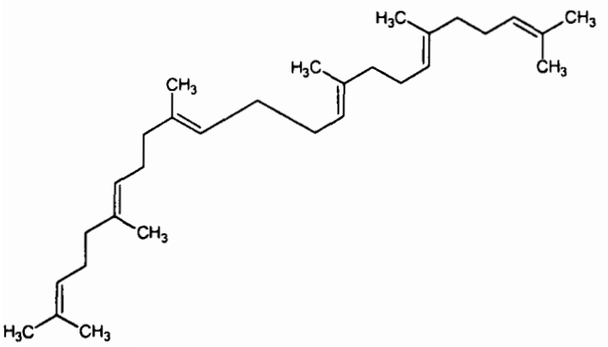
phytol



retinoic acid



retinol



squalene

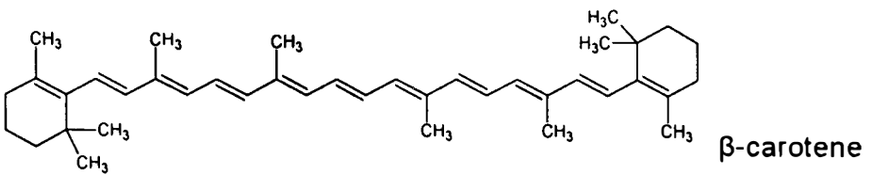
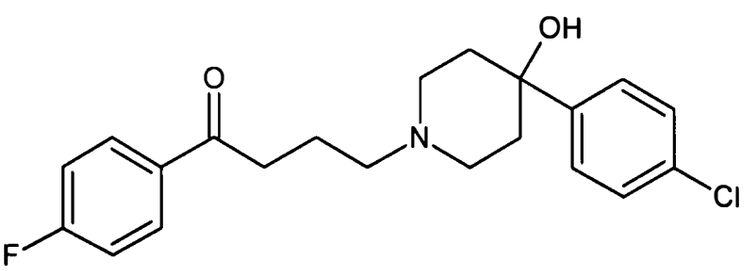


Figure 4.1. Molecular structures of the 49 terpenes



log P = 3.76
pKa = 8.04 (basic)
MW = 375.9

Figure 4.2. Molecular structure of Haloperidol

4.2. Results

4.2.1. Linear models

Skin permeation of haloperidol is affected by the addition of various terpenes (5% w/v) to the donor mixtures during *in vitro* tests. The resulting permeability coefficients are expected to be related to the molecular structure of the terpenes. Stepwise regression analyses in Minitab selected the most significantly related molecular descriptors to log k_p (as the response) and produced the QSAR models 4.1-4.3.

$$\text{Log } k_p = -11.4 - 0.00144 \text{ Wiener} + 0.0205 \text{ SA} - 2.73 \text{ SHHBd} + 31.1 \text{ MaxHp} - 1.18 \text{ dipole1} + 5.09 \text{ ABSQon} \quad (4.1)$$

$S = 0.999 \quad r^2 = 0.662 \quad F = 13.7 \quad N = 49 \quad P = 0.000$

$$\text{Log } k_p = -6.98 - 0.0147 \text{ Bp} + 0.0273 \text{ SA} + 2.63 \text{ SssO\_acnt} + 1.56 \text{ N}_{\text{aldehyde}} + 7.40 \text{ chi4\_C} - 0.559 \text{ Chi2\_p} \quad (4.2)$$

$S = 1.114 \quad r^2 = 0.583 \quad F = 9.6 \quad N = 49 \quad P = 0.000$

$$\text{Log } k_p = 0.79 + 3.07 \text{ N}_{\text{aldehyde}} - 0.240 \text{ Solubility} - 0.134 \text{ surface tension} + 2.08 \text{ SssO\_acnt} - 0.307 \text{ dipole2} - 46.3 \text{ SpcPolarizability} \quad (4.3)$$

$S = 1.168 \quad r^2 = 0.538 \quad F = 8.1 \quad N = 49 \quad P = 0.000$

The statistical parameters of the models are S, standard deviation, r^2, squared correlation coefficient, F, Fisher's statistic, and N, number of data points. P-values for equations (4.1)-(4.3) were <0.0005. The molecular descriptors have been described in Table 4.1. The descriptors in equations (4.1)-(4.3) are in the order that they were selected by stepwise regression analysis.

Table 4.1. Description of the molecular descriptors used in the QSAR models

Descriptors	Model No.	Description
Wiener	4.1	The sum of the number of edges in the shortest paths in a chemical graph between all pairs of non-hydrogen atoms in a molecule
SA	4.1 & 4.2	Surface area
SHHBd	4.1	e-State descriptor that emphasises hydrogen bond donating ability
MaxHp	4.1	The largest positive charge on a hydrogen atom
Dipole1	4.1	Dipole moment calculated by AM1
ABSQon	4.1	The sum of absolute charges on nitrogen and oxygen atoms
Bp	4.2 & 4.7	Boiling point of the compound
SssO_acnt	4.2 & 4.3	Count of all (-O-) groups of a molecule
$N_{aldehyde}$	4.2, 4.3, 4.7 & 4.8	Number of aldehyde groups
Chi4_c	4.2	4^{th} order cluster molecular connectivity index
Chi2_p	4.2	2^{nd} order simple path molecular connectivity index
Solubility	4.3	solubility of HP in Propylene Glycol at 37 °C with or without 5% (w/v) enhancer
Surface tension	4.3	
Dipole2	4.3	Dipole moment calculated by ACD
SpcPolarizability	4.3	Self-Consistent Point Polarisability
E_{LUMO}	4.7 & 4.8	Energy of the lowest unoccupied molecular orbital
SHother	4.7	Sum of the Hydrogen E-State values for hydrogens on carbon atoms
SdsCH	4.7 & 4.8	Sum of electrotopological state values for (=CH–) groups in a molecule
SEI	4.8	Total value of e-state indexes
Chi0	4.8	Zero order simple molecular connectivity index

4.2.2. Non-Linear models

The results of SVM analysis have been tabulated in Table 4.2. Models (4.4)-(4.6) in this table were generated using descriptors selected by stepwise regression analyses used in equations (4.1)-(4.3), respectively. In order to find the best SVM models, several combinations of the model parameters were investigated. *C* (Capacity) is a regularisation parameter that controls the trade-off between maximising the margin and minimizing the training error. If *C* is too small the insufficient stress will be placed on fitting the training data. If *C* is too large, then the algorithm will over-fit the training data (Luan et al. 2008). The minimum value for *C* was set at 1.0 with maximum at 10, 100 or 1000 with an increment of 1, 10 or 100. The kernel

parameters ν and ε were given maximum values of 1.0 and 10 and minimum values of 0.1 with increments of 0.1 respectively.

Table 4.2. Selected SVM models generated by STATISTICA for log k_p

			SVM					
No	Descriptors	Type	Kernel	γ	Coefficient	C	ε	ν
4.4	Wiener, SA, SHHBd, MaxHP, dipole1, ABSQon	2	Sigmoid	0.167	0.000	10	–	0.5
4.5	Bp, SA, SssO_acnt, $N_{aldehyde}$, Chi4_c, Chi2_p	1	RBF	0.167	–	68	0.2	–
4.6	$N_{aldehyde}$, Solubility, surface tension, SssO_acnt, dipole2, SpcPolarizability	1	RBF	0.167	–	10	0.4	–
4.7	E_{LUMO}, Bp, $N_{aldehyde}$, SHother, SdsCH	2	RBF	0.2	–	691	–	0.8
4.8	$N_{aldehyde}$, SdsCH, *SEI*, E_{LUMO}, Chi0	2	RBF	0.2	–	721	–	0.8

Models (4.7) and (4.8) were derived using descriptors selected by feature selection and variable screening tool of STATISTICA software. Model 4.7 resulted when the number of cuts for continuous predictors was set to 5 (the range of values for each continuous predictor was divided in 5 intervals) and model 8 resulted when the number of cuts for continuous predictors was settled to 10.

Table 4.3 shows a comparison of the statistical significance of the linear and non-linear regression models. In general non-linear models (4.4, 4.5, 4.6) show increased r^2 values in comparison with the analogous linear models (4.1, 4.2 and 4.3), which employ the same descriptors respectively. However these non-linear models also have higher MAE (mean absolute error) values. Also non-linear models (4.7) and (4.8) have higher MAE, but similar or better r^2 values than the linear models. The log k_p values for haloperidol and the predicted values by the selected models have been presented in appendix I.

Table 4.3. Differences in linear and non-linear regression analysis

Model No.	Mean Absolute Error %			Correlation Coefficient		
	Train	Test	Overall	Train	Test	Overall
4.1	0.662	0.760	0.686	0.679	0.620	0.662
4.2	0.867	0.744	0.837	0.596	0.574	0.583
4.3	0.894	0.711	0.849	0.527	0.64	0.538
4.4	1.074	0.868	1.024	0.682	0.740	0.691
4.5	1.026	0.835	0.980	0.784	0.711	0.764
4.6	1.053	0.662	0.959	0.675	0.791	0.693
4.7	1.119	1.049	1.101	0.676	0.648	0.670
4.8	1.073	1.140	1.090	0.685	0.588	0.665

4.3. Discussion

Terpenes are well-documented enhancers of drug absorption through skin (Barry, 2007). The effect of terpenes on the skin absorption depends on the chemical structure of the terpene and varies depending on the chemical structure of the penetrant (Ghafourian et al., 2004). Quantification of these effects can improve the process of drug formulation by facilitating the choice of the correct enhancer. Moreover, such computational models will help toxicologists with the human health risk assessment of skin exposure to chemical mixtures. Unfortunately, development of the computational models is greatly restricted due to the limitations in the availability of data. In this investigation, QSAR models were developed for the effect of terpene penetration enhancers on the penetration of haloperidol through excised human skin. Two statistical methods, stepwise regression analysis and STATISTICA Feature Selection were used for the selection of relevant descriptors with significant influence on the permeability coefficient from among a total of 205 molecular descriptors of the chemical enhancers calculated by computer software or determined experimentally (such as boiling point).

The Wiener topological index is the first descriptor of the QSAR model (4.1). In the chemical graph theory, the Wiener index is a topological index defined as the sum of the number of edges in the shortest paths in a chemical graph between all pairs of non-hydrogen atoms in a molecule (Diudea and Gutman, 1998). The Wiener index is the oldest topological index related to molecular branching (Todeschini and

Consonni, 2000). This index correlates with the van der Waals surface area of molecules (Gutman and Körtvélyesi, 1995). The second descriptor in model 4.1 is surface area which has a positive correlation with log k_p, meaning that the higher the surface area of a terpene the higher the log k_p of haloperidol. This combined with the negative coefficient of Wiener which has information on branching as well as molecular size, may indicate that branched terpenes will lead to high absorption of haloperidol through skin. It must be noted that in terms of molecular size, the molecular weight of all the terpenes in this study is below 500 Da and therefore they should not have penetration problems (Lipinski, 1997), in order to penetrate into the SC and change the barrier properties. Model (4.1) also shows the negative effect of hydrogen bond donor ability (SHHBd) of terpenes on the penetration rate of haloperidol. On the other hand, the descriptor MaxHP which represents the largest positive charge on a hydrogen atom in molecules has a positive correlation with log k_p. In other words, terpenes with a lower number of hydrogen bonding groups (e.g. non-hydroxylated terpenes), containing a positively charged hydrogen (such as those containing hydrogens close, but not directly attached to heteroatoms thus not counting as hydrogen bond donors) are better enhancers of haloperidol penetration. These are either molecules with an ether or ester group. In the study of Jeffrey et al. (1999), irritant esters of the skin showed lower density, lower water solubility, lower sum of partial positive charges, higher Hansen hydrogen bonding parameter and higher Hansen dispersion parameter when compared with non-irritant esters. Hydrogen bonding and the sum of partial positive charges is closely related to intermolecular interactions. The terpenes with the highest SHHBd and lowest MaxHP are the two esters octisalate and cedryl acetate respectively which according to Jeffrey et al. (1999) are also potential irritants. However, skin irritation is a property that depends not only on the skin penetration, but also on specific interactions of the chemical with skin (Roberts and Patlewicz, 2002).

Another descriptor of model (4.1), dipole1 is the dipole moment calculated by AM1 semi-empirical method, representative of the strength and orientation behaviour of a molecule in an electrostatic field. It has a negative correlation with log k_p indicating the negative effect of terpene dipolarity. Descriptor ABSQon with a positive coefficient in this same model represents the sum of absolute charges on nitrogen and oxygen atoms, which in case of terpenes only oxygen atoms may be present.

This indicates the positive effects of the oxygen atom on enhancement activity of terpenes. These findings are consistent with the work of Ghafourian et al. (2004) who demonstrated that low molecular weight terpenes containing heteroatoms (oxygen) are better penetration enhancers for 5-fluorouracil.

The descriptor Bp is the first descriptor of QSAR model (4.2) and represents the boiling point of terpenes. Boiling point can indicate the intermolecular interaction energy of a terpene which may result from its polarity or polarisability or most significantly from the hydrogen bonding. Therefore the negative relationship with the permeation coefficient indicates a higher skin permeation rate with the non-hydrogen-bond-donor terpenes. This is similar to the descriptor SHHBd in equation (4.1). Boiling point was also selected by STATISTICA feature selection (non-linear model (4.7)). SssO_acnt is the count of other oxygen groups in terpene molecules. It has a positive correlation with log k_p in QSAR models (4.2) and (4.3), Indicating higher skin permeation of haloperidol with ether or ester terpenes in the donor phase, analogous to ABSQon in equation (4.1). Moreover, the number of aldehyde groups ($N_{aldehyde}$) also has a positive correlation with log k_p in QSAR models (4.2) and (4.3). This descriptor is also present in both non-linear models (4.7) and (4.8), indicating a good enhancement effect on haloperidol penetration by aldehyde terpenes. These are again in accordance with Ghafourian et al. (2007) on the permeation enhancement of terpenes towards 5-fluorouracil.

The other descriptor of model (4.2), Chi4_c is the 4^{th} order cluster molecular connectivity index that is positively correlated with log k_p. It indicates presence of atoms that are attached to four other non-hydrogen atoms. Terpenes such as cedrol which are highly branched and cyclic have high values of Chi4_c. Moreover, the descriptor Chi2_p in this model is the 2^{nd} order simple path molecular connectivity index (Hall and Kier, 1991) that is negatively correlated with log k_p. This descriptor in conjuction with surface area which has a positive coefficient indicates that branched and more complex terpenes lead to higher penetration rates of haloperidol.

In model (4.3), apart from the number of aldehyde and etheric oxygen atoms which are similar to equation (4.2), solubility (mg/ml) of haloperidol in each terpene solution (measured by Kang et al., 2007) has been selected. This parameter has a

negative coefficient, this is in accordance with the scientific literature suggesting that a permeant should not be over-solubilised in the donor mixture, as this leads to lower degrees of saturation (ratio of concentration/solubility) when comparing the same drug concentrations. A near saturation level will maximise the concentration gradient across the SC because the partition coefficient of a drug between the skin and the solvent mixture generally falls as the solubility in the solvent rises (Roberts et al., 2002; Baker, 1986).

The descriptor surface tension in model (4.3) represents the surface energy of a liquid (a measure of how difficult it is to stretch or break the surface of a liquid). Surface tension is a result of higher intermolecular interaction forces between a liquid's molecule, than between the liquid molecule and air. In model (4.3), it shows a negative correlation with log k_p which is in fact analogous to the negative effect of Bp in equation (4.2). Terpenes such as (S)-(-)-perillaldehyde and (1R)-(-)-myrtenal with conjugated systems have high surface tensions and low skin penetration enhancement activities.

The other descriptor of model (4.3), Dipole2, is dipole moment calculated by MDL QSAR software, with a negative coefficient in equation (4.3) similar to that in equation (4.1). This indicates the negative effect of terpene dipolarity on the penetration enhancement property. Also in this equation SpcPolarizability is Self-Consistent Point Polarisability which is defined as specific polarisability of a molecule equal to Polarisability/Volume. This is the last descriptor of model (4.3) with a negative correlation with log k_p and terpenes thymol and (±)-linalool have the highest and the lowest values of this descriptor, respectively.

In models (4.7) and (4.8), E_{LUMO} is a descriptor that represents the energy of the lowest unoccupied molecular orbital and measures the electrophilicity of a molecule. E_{LUMO} has been shown to be related to hydrogen bond donor ability as well (Dearden and Ghafourian, 1999). The descriptor SHother represents the sum of the Hydrogen e-State values for hydrogens on groups other than –OH, -NH2, =NH, -SH, -NH-, ≡CH, or, if attached to chlorine or fluorine, -CH3, -CH2-, >CH-, =CH-. In this case, it can be seen from Figure 4.1 that SHother value will be high for molecules

containing hydrogen atoms attached to carbon atoms without a Cl or F atom attachments. In this dataset squalene, retinoic acid and retinol have a value of above 12 and β-carotene has a value of 26; on the contrary, (R)-(+)-pulegone and Cyclohexanemethanol have the lowest values. SdsCH is the atom type electrotopological state index for =CH– groups in a molecule. SEI is the sum of e-State indices; it represents the availability of electrons in atoms of a molecule. Although it is an electronic descriptor, it can also be related to the molecular size, as larger molecules have higher number of atoms with different levels of electrotopological state index. The descriptor Chi0 is the zero order simple molecular connectivity index, a molecular size descriptor.

Kang et al. (2007) suggested that an ideal terpene enhancer should possess at least one or combination of the following properties: hydrophobic, in liquid form at room temperature, with an ester or aldehyde but not acid functional group, and is neither a triterpene nor tetraterpene. Our findings agree well with this conclusion and in addition provide a quantitative model for estimation of the enhancement effect.

There are several previous QSAR works on penetration enhancers. The work published by Pugh et al. (2005), relates the enhancement effect to the simple descriptors of hydrogen bonding with a negative effect, chain lengths with a positive effect, and molecular weight with a negative effect. However, this model failed to predict the activity of some of the most effective enhancers. The models generated in our work employ more complex descriptors such as those related to molecular shape.

Karande et al. (2005) took a different approach by investigating the barrier disruption potential and skin safety of several classes of enhancers. They concluded that disruption of the skin barrier occurs by fluidisation of lipid bilayers or by extraction. It was suggested that irritation response correlated with the denaturation of SC proteins, while the fluidisation potential correlated with hydrophobicity of molecules, and the extraction potential correlated with the ratio of the hydrogen bonding component of solubility parameter (δ_h) to the square root of cohesive energy density. It is difficult to compare the results of this investigation with the findings of Karande et al. (2005), as this study is focused on a specific class of enhancers as opposed to the broad dataset employed by Karande et al. However as the relationship

with log P is not significant, it can be concluded that the terpenes in this investigation are either equal in their fluidisation property or their mechanism is mostly through extraction. The latter explanation can be confirmed by the negative effects of descriptors SHHBd, Bp and surface tension in the models. In terms of the fluidisation ability, on the other hand, the log P range of the terpenes varies from 1.6 to 15.51 with the average at 4.55. Therefore with this wide range of log P, similar fluidisation activity seems highly unlikely.

Linear methods, such as multiple linear regression (MLR) or Partial Least-Squares regression (PLS), are the most widely used statistical methods in the QSAR/QSPR area based on an assumption of linearity between the descriptors and the investigated experimental property (Gramatica et al., 2002; Hadjmohammadi et al., 2007; Roy et al., 2006; Tuppurainen et al., 2006). Support vector machine (SVM) is a non-linear algorithm developed for regression and classification problems (Vapnik, 1998). Due to its remarkable generalisation performance, SVM has attracted increased attention and gained extensive applications, including applications in QSAR and QSPR for drug design (Burbidge et al., 2001; Czerminski et al., 2001; Liu et al., 2003). Goudarzi and Goudarzi (2009) compared SVM with PLS, Principle Component Regression (PCR), and MLR, and found that SVM model was much better for modelling and predicting the acidic dissociation constants of some organic compounds. They concluded that SVM can be used as a powerful chemometrics tool for QSPR studies and that SVM drastically enhances the prediction ability of QSAR models. SVM has been used not only to construct activity and toxicity models (Yao et al., 2004) but also to build quantitative structure-retention index models (Du et al., 2009). In the findings of Chen (2008) it was shown that non-linear SVM derived statistical models have similar prediction ability to those of radial basis function neural network and MLR methods. It was also indicated that SVM can be used as an alternative modelling tool for quantitative structure-property/ activity relationships (QSPR/QSAR) studies. In our findings SVM gave better r^2 but higher MAE when compared with linear regression analysis (Table 4.3). According to the internal validation results in Table 4.3, the best estimation accuracy occurs through the use of the linear model (4.3) for the test and validation sets.

4.4. Conclusion

According to the findings of our *in silico* study, an ideal terpene enhancer should preferably contain an aldehyde, ester or ether group, while the presence of hydrogen bonding donor groups is not advantageous. The chemical range of the enhancers investigated in this work is limited to terpenes with similar structural properties. Therefore the conclusions drawn from this work might not be applicable to enhancers from other chemical classes such as ionic surfactants, Azone® like compounds (Figure 1.8) or fatty acids.

SVM models did not prove to be better than the linear regression models in terms of average prediction error, a general finding that contradicts those of Chen (2008), Goudarzi and Goudarzi (2009) and Luan et al. (2008) and indicates that the relation of log k_p with the selected descriptors is linear.

5. Modelling the Effect of Mixture Components

Chemicals that come into contact with skin are, more often than not, formulated in a mixture, or in the simplest form, dissolved in at least one solvent. Skin absorption of a chemical depends not only on the nature of the compound but also on the nature of the other ingredients present in the mixture (formulation). As a result, varying range of penetration rates are achieved by altering the formulations. However, the data used in most prediction models are based on experimental data for aqueous solutions of chemicals. The reason is because very limited experimental data are available for chemical mixtures. While it is frequently difficult to assess the absorption of individual chemicals, it is challenging to quantitatively assess the absorption from chemical mixtures. The aim of this investigation was to develop QSAR models to study the effect of mixture components on skin absorption of penetrants. The model will help identify the mechanisms involved in the penetration through skin and the effect of formulation factors. Data for this study were acquired from Riviere (2008) who collaborated with us in this project.

5.1. Methods

5.1.1. The dataset

Skin permeation data of 12 different penetrants (Table 5.1) each blended in 24 different solvent mixtures (Table 5.2), were used in this investigation. Experimental details are fully described in Riviere and Brooks (2005). The permeability data consisted of apparent skin permeation rate constants (k_p) in cm/h measured using finite-dose *in vitro* porcine-skin flow through diffusion cells. The skin was perfused using a Krebs-Ringer bicarbonate buffer spiked with dextrose and bovine serum albumin, and topically dosed unoccluded with 20 μL of one of 12 marker penetrant compounds (target dose of 10-20 $\mu g/cm^2$) formulated in one of 24 specified mixtures (Table 5.2). Trace amounts of methanol and toluene were used to solubilise radiolabelled penetrants before dilution with non-radiolabelled compounds.

Table 5.1. Penetrants

Atrazine	Pentachlorophenol
Chlorpyrifos	Phenol
Ethylparathion	ρ-Nitrophenol
Fenthion	Propazine
Methylparathion	Simazine
Nonylphenol	Triazine

This dataset was used for the development of QSARs. In order to evaluate the applicability of the model the chemical space of the penetrants of this dataset was compared with the combined datasets of Flynn (1990) and Wilschut et al. (1995). The combined dataset contains *in vitro* human skin permeability data (log k_p) for 112 compounds.

5.1.2. Structural descriptors

The predictors (descriptors) of penetrants included connectivity indices, quantum molecular descriptors and group counts calculated using the TSAR 3D software (Accelrys Ltd version 3.3). The physico-chemical properties of the mixture components including boiling point, melting point, solubility, vapour pressure and Henry's law constant were obtained through ChemBioFinder (CambridgeSoft, 2009) online software and the SRC PhysProp database (Syracuse Research Corporation, 2009). Log P for solvent components and for the penetrants was calculated by the ACD labs/logD Suite 7.0.5 (ACD/LogD, 2008). Averages of physico-chemical properties for solvent mixtures were calculated using the fractions of each component e.g. boiling point of the mixture.

5.1.3. Development and validation of QSARs

Stepwise regression analysis was used to develop the QSAR models in MINITAB version 15.1.0.0 (Minitab Statistical Software, 2009) The predictability of the models was examined by a leave – many - out procedure. As such, chemicals were sorted according to the ascending log k_p values; for each set of four solvents, the first compound was allocated to group a, the second to group b, the third to group c and the fourth to group d. This ensured that each group covered similar ranges of the skin permeation kinetics. The regression analysis was carried for the chemicals in groups a, b and c (as the training set), and the resulting equation was used to calculate the skin permeation parameter for the remaining group d (as the test set). The procedure was carried on to leave one group out at a time (all the possible combinations of groups making the training set). The Mean Absolute Error (MAE) of prediction was calculated as a measure of the model accuracy.

Table 5.2. Mean Composition of the 24 mixtures

Mixture	%EtOH	%Water	%PG	%MNA	%SLS
Et	99.67	0	0	0	0
Et+MNA	99.51	0	0	0.16	0
Et+SLS	62.59	26.53	0	0	10.61
Et+MNA+SLS	62.50	26.49	0	0.13	10.60
Et+Wa	42.66	55.86	0	0	0
Et+Wa+MNA	43.79	55.78	0	0.14	0
Et+Wa+SLS	39.44	50.25	0	0	10.05
Et+Wa+MNA+SLS	39.39	50.18	0	0.13	10.04
Wa	0	99.75	0	0	0
Wa+MNA	3.03	96.59	0	0.13	0
Wa+SLS	0	90.70	0	0	9.07
Wa+MNA+SLS	2.75	87.77	0	0.12	9.13
Et+PG	42.99	0	56.73	0	0
Et+PG+MNA	42.92	0	56.65	0.14	0
Et+PG+SLS	28.39	24.15	37.54	0	9.66
Et+PG+MNA+SLS	28.36	24.13	37.50	0.12	9.65
PG	0	0	99.75	0	0
PG+MNA	2.93	0	96.70	0.12	0
PG+SLS	0	22.13	68.79	0	8.85
PG+MNA+SLS	2.69	22.29	65.76	0.11	8.92
Wa+PG	0	48.99	50.76	0	0
Wa+PG+MNA	2.98	47.46	49.18	0.13	0
Wa+PG+SLS	0	44.62	46.23	0	8.92
Wa+PG+MNA+SLS	2.71	43.20	44.76	0.11	8.98

Et-Ethanol; PG-Propylene glycol; MNA-Methyl nicotinate; SLS-Sodium lauryl sulphate; Wa-Water.

The chemical space of the present dataset was compared with that of the skin permeability dataset drawn from Flynn (1990) and Wilschut et al. (1995). The comparison was made using descriptor spaces of Potts and Guy (1992) model (i.e molecular weight and octanol/water partition coefficient), Principal Component Analysis (PCA) scores plot with all the descriptors being included in the analysis and PCA scores plot using the descriptors selected by stepwise regression analysis for the Flynn (1990) and Wilschut et al. (1995) dataset. PCA was carried out using MINITAB statistical software.

5.2. Results and Discussion

The combined effect of chemical structures of the penetrants and the properties of the mixture components on the permeation rate through porcine skin was studied using QSAR. Stepwise regression analysis performed on the data set of 288 penetrant/ mixture-component combinations resulted in equation 5.1, in which the descriptors were limited to two penetrant descriptors and one solvent mixture descriptor.

$$\text{Log } k_p = - 0.909 - 0.610 \log P + 2.62 \text{ Chi9\_p} - 0.00917 (\text{SolBP - SolMP}) \quad (5.1)$$

$$S = 0.438,\ r^2 = 0.729,\ F = 255.2,\ P = 0.000,\ N = 288$$

In equation (5.1), Log k_p represents permeation rate constant of compounds dissolved in various solvent mixtures from porcine skin, log P is the octanol/ water partition coefficient of the solute (the penetrant), Chi9_p is the 9^{th} order path molecular connectivity index of the penetrant, and SolBP – SolMP is the difference between the boiling point and the melting point of the solvent system.

Log P was the most significant descriptor of the equation (the first to be selected by the stepwise regression analysis). It can be seen in equation (5.1) that log P of penetrants has a negative effect on the skin permeation rate. This is opposite to the common knowledge that lipophilic compounds have higher skin permeation profiles, as evidenced also in Potts and Guy's model (1992). The negative relationship between log k_p and log P could be due to the fact that most of the drugs in this particular dataset are more lipophilic than the compounds in the datasets normally used in QSAR studies of skin permeability. Figure 5.1 shows a graph between log k_p and log P for the penetrants of this study and the penetrants of Wilschut et al. (1995) and Flynn (1990). The opposite trends of the relationships between log k_p and log P for the two datasets are evident despite the poor correlations. The figure also shows that compounds of the present dataset have relatively higher log P values than compounds in the combined datasets of Wilschut et al. (1995) and Flynn (1990). This follows the well established nonlinear relationship of biological activity with lipophilicity described by parabolic (Hansch and Clayton, 1973) or bilinear

(Kubinyi, 1977) models. Compounds with extreme lipophilicity can be expected to partition into the skin and remain there, with little permeation to the aqueous receptor phase. This has been shown for example for tetrahydrocannabinol (Challapalli and Stinchcomb, 2002), with extremely high log P value of 6.84 as calculated by ACD/logD Suite. López et al. (1998) showed a bilinear relationship between lipophilicity of phenyl alcohols and the permeability coefficient through rat skin, where the k_p was reduced for compounds with log P values higher than around 5.

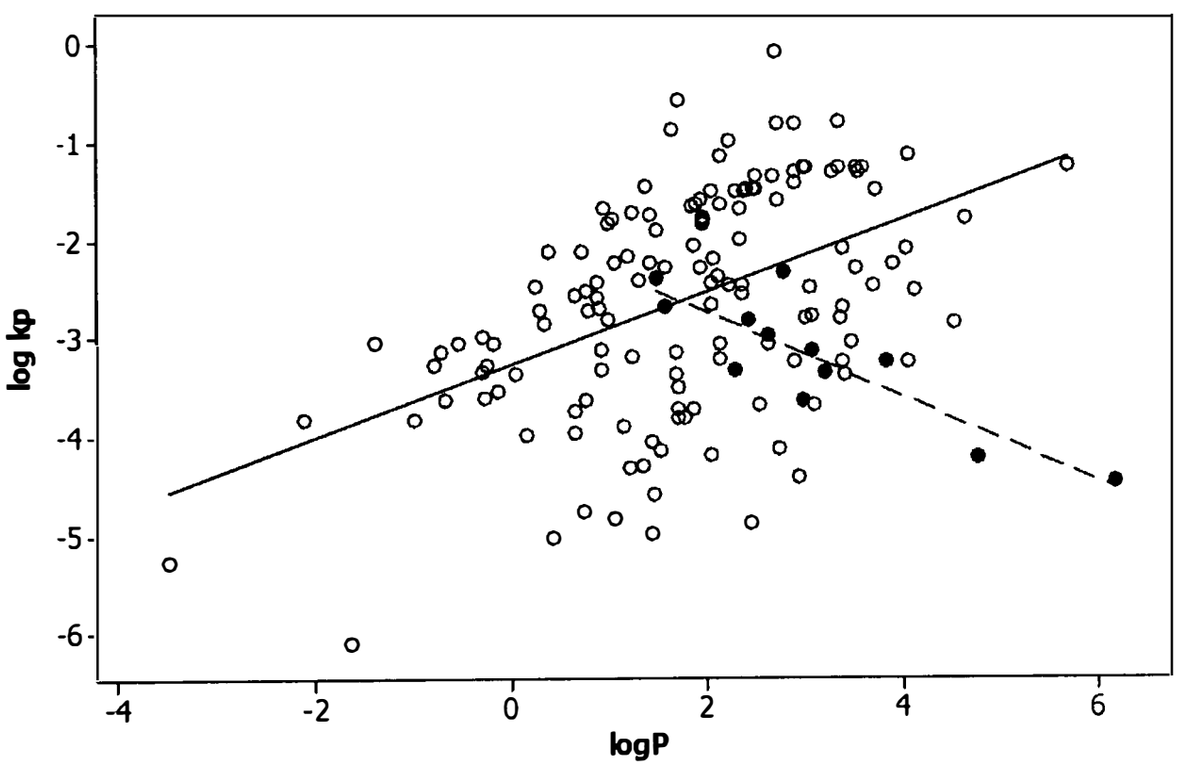


Figure 5.1. Comparison of the lipophilicity of the drugs in the two datasets of Riviere's (solid circles) and Flynn (1990) and Wilschut et al. (1995) dataset (empty circles)

There are a number of other factors that may have contributed to the observed negative relationship between log k_p and lipophilicity. One is the finite-dose nature of the experiments with skin dosed with a limited amount of drug. The limited availability of the compound could result in a large fraction of the lipophilic compounds being concentrated in the skin according to their skin/water partition

coefficients. A second factor is the differing nature of the receptor phase which contained albumin.

Chi9_p is the second most significant descriptor of equation (5.1). This molecular connectivity descriptor indicates the presence of nine-atom chains in the molecules. The positive coefficient of this descriptor indicates a better permeation of compounds containing long chain fragments. The penetrants with the highest Chi9_p values were chlorpyrifos, fenthion and nonylphenol. These penetrants have the maximum molecular weight of 350 Da which is still smaller than the size expected to limit the absorption. According to Barry (2007) a molecule's ideal molecular mass, in order to penetrate the SC is less than 600 Da. In addition, according to Lipinski's rule of five, chemicals with molecular weight of above 500 Da may have biological membrane penetration problems (Lipinski et al., 1997).

The third descriptor of the equation, SolBP – SolMP, represents the difference between melting and boiling points of the solvent mixtures, where the higher the difference, the lower the skin absorption of compounds from the vehicle. It is therefore expected that compounds formulated in vehicles with small boiling and melting point gaps will have better permeation through skin. The difference between these two properties has been attributed to the molecular symmetry, with highly symmetrical molecules having much larger melting points and decreased boiling points (Slovokhotov, 2007). In the solvents used in this study, the biggest difference in the melting and boiling points is for propylene glycol. Therefore the vehicles containing higher concentrations of this solvent will lead to lower permeation of the penetrants studied in this investigation.

In order to validate the reported QSAR, a leave-many-out procedure as explained in the methods section was used and mean absolute error calculated. Figure 5.2 is the graph between observed and predicted log k_p. The r^2 between observed and predicted log k_p and the MAE were 0.654 and 0.396, respectively.

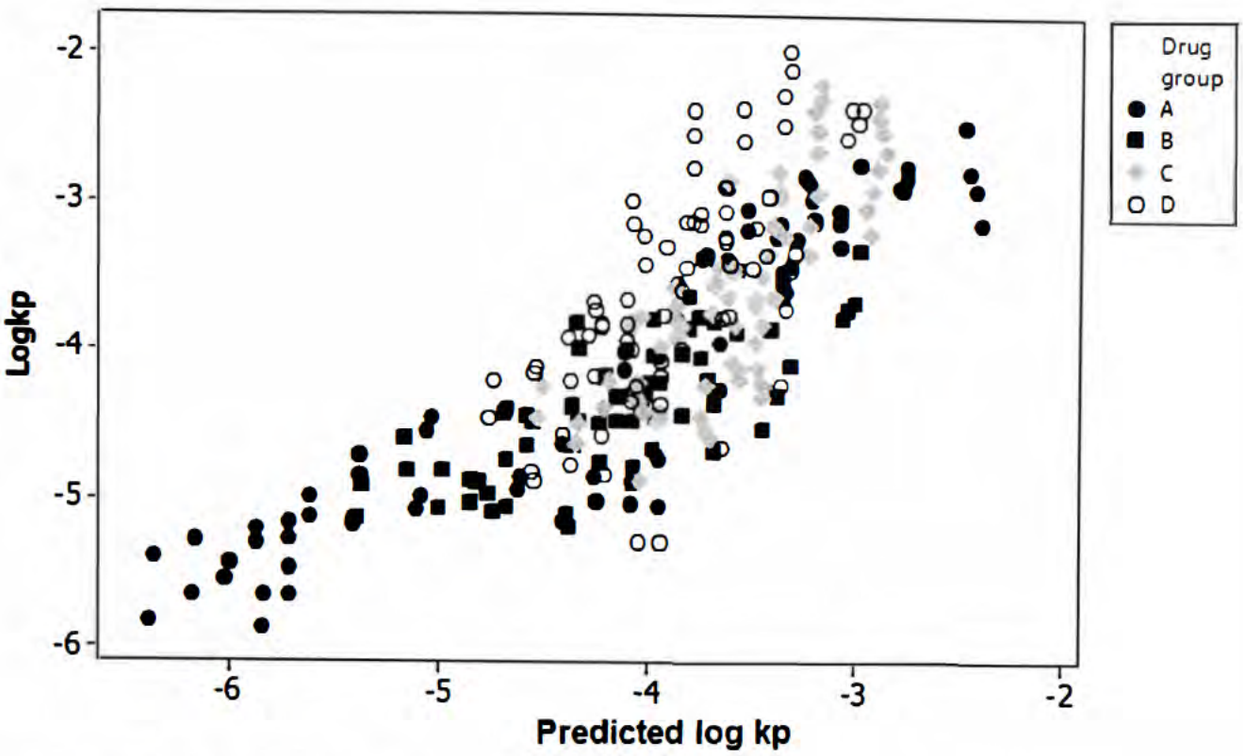


Figure 5.2. Plot of observed Log k_p against predicted Log k_p.

The level of uncertainty in predictions made by any QSAR is characterised by the validation tests, but it also depends on the diversity of the training set which defines the domain of applicability. Any QSAR model is expected to perform best for the chemicals that are similar to those in the training set (Weaver and Gleeson, 2008). The applicability of equation (5.1) is be limited to the prediction of log k_p for new molecules that are similar to those of our dataset. Therefore, the chemical space of the penetrants included in this dataset was compared to that of the combined datasets of Flynn (1990) and Wilschut et al. (1995). Comparison of the physicochemical properties of the penetrants in the two datasets above were made first by looking at the molecular descriptors of the widely accepted Potts and Guy model (1992) consisting of log P and molecular weight (MW) as in Figure 5.3. The figure shows that our dataset does not include any hydrophilic chemicals and log P values are all above 1.5. The other limitation of the dataset is the relatively low molecular weights of the chemicals in comparison with the datasets of Flynn (1990) and Wilschut et al. (1995). Therefore, a few high molecular weight and low lipophilicity chemicals could be identified for future measurements. Examples are hydrocortison octanoate, caffeine and methanol, as can be seen in the figure.

As a second strategy, the two datasets were compared using all the calculated molecular descriptors, a total of 128. This was made possible through the use of Principal Component Analysis (PCA). PCA is a data reduction method which takes the information from original molecular descriptors and generates the same number of new descriptors (PCs), with the first PC containing the maximum information of the original dataset, and the second PC being the second most informative. Therefore, the plot between PC1 and PC2 (the scores plot) provides a good overview of the information content of the dataset. The first two principal component score vectors, PC1 and PC2, are plotted in Figure 5.4. The figure shows that the chemicals of the current dataset are located in the bottom left quarter of the plot, with relatively low PC1 and PC2 values. By visual inspection of the graph, several groups of chemicals belonging to datasets of Flynn (1990) and Wilschut et al. (1995) were identified in the plot to cover various ranges of PC1/PC2. These are chemicals with high PC1 and PC2 values such as codeine and morphine, compounds with high PC1 and varying values of PC2 including steroids such as testosterone and hydrocortison octanoate, and compounds with very low PC1 and PC2 values such as octanol.

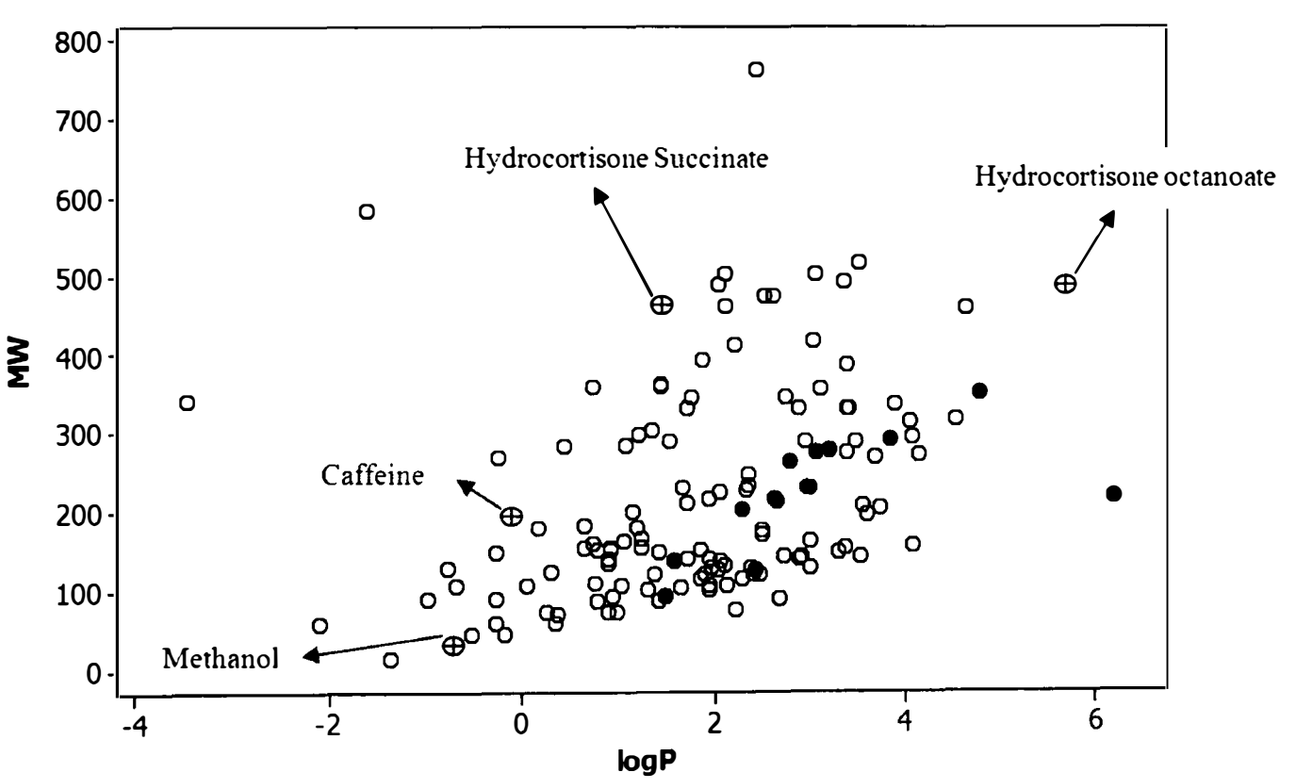


Figure 5.3. Comparison of the chemical diversity of the penetrants of the present dataset (solid circles) with that of the combined dataset of Flynn (1990) and Wilschut et al. (1995) (empty circles), using log P and molecular weight.

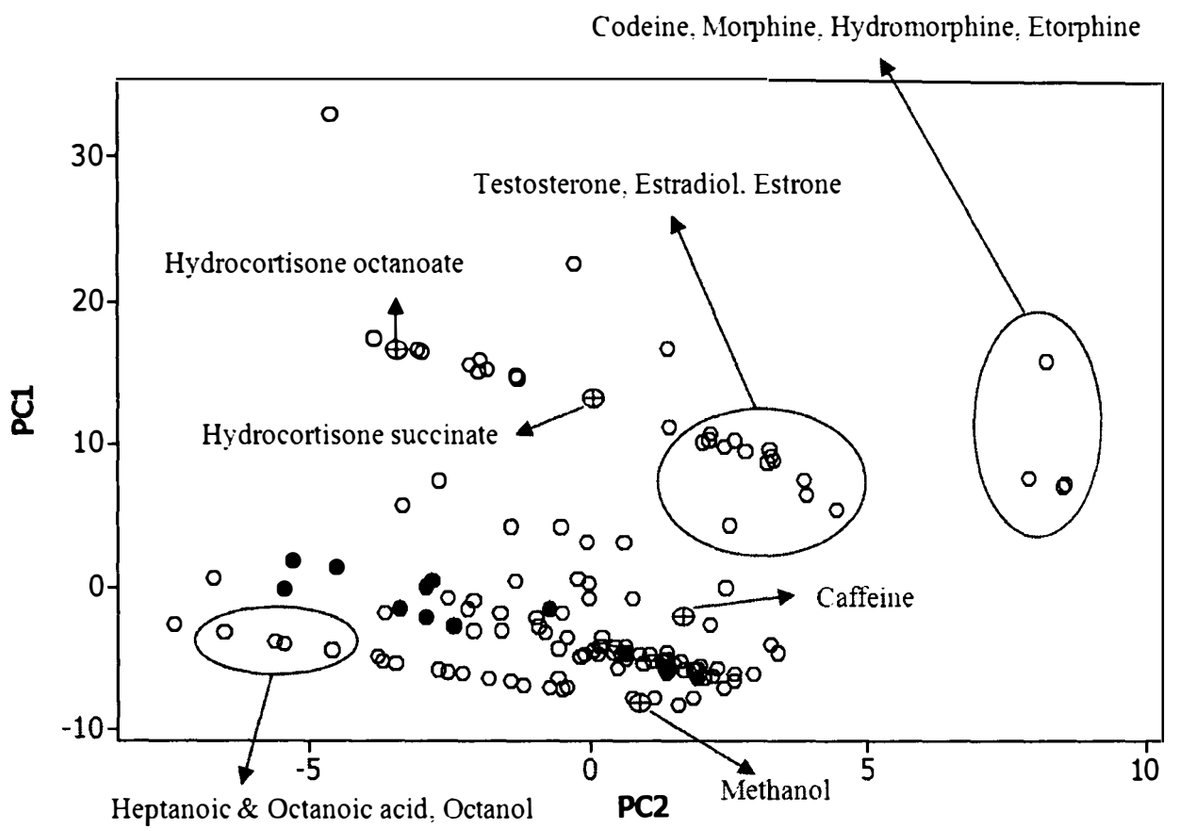


Figure 5.4. Comparison of the chemical diversity of the penetrants of the present dataset (solid circles) with that of the combined dataset of Flynn (1990) and Wilschut et al. (1995) (empty circles), using PCA scores plot incorporating all the descriptors.

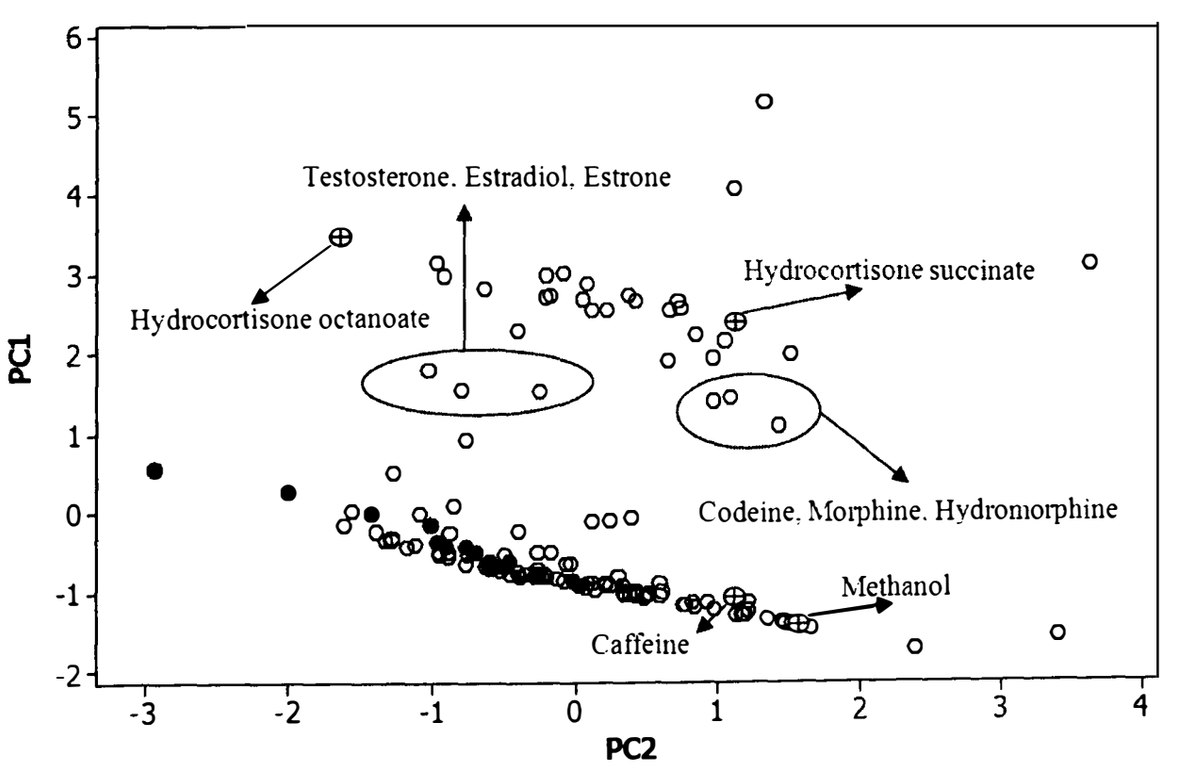


Figure 5.5. Comparison of the chemical diversity of the penetrants of the present dataset (solid circles) with that of the combined dataset of Flynn (1990) and Wilschut et al. (1995) (empty circles), using PCA scores plot incorporating the descriptors of equation 5.2.

The third method for comparison of the datasets involved the use of a selection of molecular properties that are specifically involved in the skin permeation of compounds. To this end, stepwise regression analysis was used for the selection of molecular descriptors affecting compounds' absorption through skin. In this analysis, the dataset of Flynn (1990) and Wilschut et al. (1995) containing the skin permeation rate constant through human skin was used. In stepwise regression analysis, the skin permeation rate constant (log k_p) was the dependant variable and all the molecular descriptors were the independent variables. Stepwise regression analysis selected three descriptors and resulted in equation (5.2) below.

$$\text{Log } k_p = -2.91 + 0.62 \log P + 5.21 \text{ Chiv10\_p} - 1.64 \text{ Chiv6\_p} \qquad (5.2)$$

S = 0.548, r^2 = 0.757, F = 140, P = 0.000, N = 139

In equation 5.2, log P is the octanol/water partition coefficient, Chiv10_p and Chiv6_p are 10^{th} and 6^{th} order valence corrected path molecular connectivity indices of the penetrants. Molecular connectivity indexes are topological descriptors of molecular structures indicating the frequencies of occurrence of certain fragments in the molecules. Path molecular connectivity indexes indicate the frequency of non-branched chains of certain lengths, in this case six-atom and ten-atom chains as shown in scheme 5.1 below (Todeschini and Consonni, 2000).

Scheme 5.1. Six-atom and ten-atom fragments for the calculation of path molecular connectivity indexes, Chiv6_p and Chiv10_p.

The three descriptors selected by stepwise regression analysis were used in PCA and the scores plot between the first and the second PCs (Figure 5.5) was used to compare the datasets. Figure 5.5 is similar to Figure 5.4 in identifying certain compounds from the dataset of Flynn (1990) and Wilschut et al. (1995) such as steroids, narcotic analgesics and small polar molecules such as caffeine and methanol which are not present in the current dataset.

Therefore, an overview of Figures 5.3-5.5 can identify several areas of the chemical space that are missing from the present dataset. From these groups of chemicals, caffeine, 1-octanol, testosterone and codeine were selected for further studies.

5.3. Conclusion

In conclusion, skin permeation of drugs from different vehicle systems can be modelled using QSAR given the availability of an appropriate dataset containing diverse permeants and vehicles. Vehicle effects were well predicted in this work. However, rigorous validation of such models for estimation purposes will require a large volume of data. In this study, the negative relationship was obtained between log k_p and log P. This was attributed to the fact that most of the drugs in this particular dataset are more lipophilic than the compounds in the common permeability datasets used in QSAR studies of skin permeability. Therefore, it can be envisaged that these highly lipophilic agents concentrate in the SC with little ability to partition into the aqueous receptor phase. This scenario is relevant for many

pesticides and lipophilic contaminants encountered in environmental exposure scenarios. For further validation of this model, skin permeation of the compounds identified through the comparison of the datasets is necessary to be determined in similar solvent mixtures.

6. Validating the effects of Mixture Components

For a rigorous study analysing the effect of mixture components as well as the effect of molecular structures of the penetrants a chemically diverse dataset is required. As suggested from the study comparing the datasets of Riviere (2008) and combined dataset of Flynn (1990) and Wilschut et al. (1995), caffeine, codeine, octanol and testosterone were the compounds with relatively high distance from Riviere (2008) compounds in the chemical space. These were selected as representatives of the chemical space of the dataset when compared with the skin permeability dataset of Flynn (1990) and Wilschut et al. (1995). The skin permeability of the above four different penetrants, each blended in the 24 different solvent mixtures presented in Table 5.2, were determined using diffusion cell studies employing porcine skin. The resulting 96 k_p values were combined with the previous dataset of 288 k_p data and used for QSAR analysis.

6.1. Material and Methods

6.1.1. Materials

Caffeine [8-14C] Specific Activity: 50-60 mCi/mmol 1.85-2.22 GBq/mmol, n-Octanol [1-14C] Specific Activity : 2-10 mCi/mmol 74-370 MBq/mmol, Testosterone [4-14C] Specific Activity : 50-60 mCi/mmol 1.85-2.22 GBq/mmol, Codeine [N-methyl-14C], obtained from American Radiolabeled Chemicals, Inc, St. Louis USA. Absolute ethyl alcohol was obtained from Aaper Alcohol and Chemical Co. Shelbyville, K Y, USA. Propylene glycol (purity = 99%), Sodium lauryl sulphate (purity = 99%), and Methyl nicotinic acid (purity = 99%) were obtained from Sigma Chemical Co. St. Louis, MO, USA. Water was distilled in our in-house still.

6.1.2. Skin penetration studies

Apparent permeability coefficient (k_p) of caffeine, codeine, octanol and testosterone, each blended in the 24 different mixtures, as presented in Table 5.2, were obtained through flow-through diffusion cell using porcine skin. The flow-through diffusion cell was used to perfuse skin obtained from the dorsal area of weanling female Yorkshire pigs according to protocols approved by the North Carolina State University Institutional Animal Care and Use Committee. Skin was dermatomed to a thickness of 500 μm with a Padgett dermatome. Each circular skin disk was punched to provide a dosing surface area of 0.64 cm^2 and then placed into a two-compartment Teflon Bronaugh flow-through diffusion cell. Skin was perfused using a Krebs-Ringer bicarbonate buffer spiked with dextrose and bovine serum albumin, and topically dosed non-occluded with 20 μl of one of the four marker penetrant compounds (10μg/cm^2) formulated in one of 24 specified mixtures listed in Table 5.2. This resulted in a total of 96 treatments with n = 4-5 replicates/treatment designed as a randomised complete factorial experiment.

6.1.3. QSAR studies

The k_p values measured in this study for caffeine, codeine, octanol and testosterone were merged with the previous dataset of log k_p values for 12 other compounds (Table 5.1) also blended with the same mixture components as Table 5.2 (Riviere and Brooks, 2005). These log kp values are measured using the same experimental procedures as in this study. Therefore, the dataset used for the QSAR studies consisted of a total of 384 unique measurements of k_p for the penetrant/ components combinations. Table 6.1 is the list of the 16 penetrants used in QSAR study.

Table 6.1. Complete set of Penetrants

Atrazine	Pentachlorophenol	Caffeine
Chlorpyrifos	Phenol	Codeine
Ethylparathion	ρ-Nitrophenol	Octanol
Fenthion	Propazine	Testosterone
Methylparathion	Simazine	
Nonylphenol	Triazine	

For the development of QSAR models, properties of the penetrants and the solvent mixtures were assembled. The molecular descriptors (properties) of the penetrants were calculated using two software packages of ACD labs/LogD Suite 7.0.5 release (ACD/LogD, 2008) and TSAR 3D (Accelrys Ltd version 3.3). The molecular descriptors included octanol/ water partition coefficient, molecular connectivity indices, quantum molecular descriptors, and various atom and group counts. The physico-chemical properties of mixture components including boiling point, melting point, solubility, vapour pressure and Henry's law constant were obtained through ChemBioFinder (CambridgeSoft, 2009) online software and SRC PhysProp database (Syracuse Research Corporation, 2009). Hildebrand solubility parameters (δ) were obtained from Hansen (1967) for the solvents and calculated according to Fedors (1974) group contribution method for the penetrants. As there was a mixture of a

number of solvents in the vehicles, averages of physicochemical properties for solvent mixtures were calculated using the fractions of each component.
Stepwise regression analysis was performed between log k_p as the dependant variable and the molecular descriptors of the penetrants and the mixture components as the predictors. This enabled the identification of the significant molecular descriptors affecting skin penetration of chemicals. Several stepwise regression analyses using various sets of penetrant molecular descriptors and solvent properties were performed and several regression models were generated. In order to minimise the risk of chance correlations, the number of descriptors in the regression models were limited to four.

The models were validated for penetrants using a leave-many-out cross validation procedure. To do this, the penetrants were divided into four groups with similar ranges of lipophilicity (log P values) in each group. Regression analyses were performed four times, each time leaving one group out. The log k_p values of the test sets were estimated using the equations obtained for the training sets and the mean absolute error was calculated from the difference between the observed and the predicted log k_p values of the test sets.

6.2. Results and Discussion

Skin penetration of drugs is controlled by the molecular structures and physico-chemical properties of the intended penetrants and the mixture ingredients in the vehicle. In order to rationalise the combined effect of structural characteristics of the penetrants and the physico-chemical properties of the mixture components, this investigation focused on QSAR model development for a dataset of skin permeation of chemicals dissolved into a combination of several solvents, surfactants and methyl nicotinic acid. Permeation coefficients were measured for four compounds that were rationally selected in order to add a high level of diversity to the existing dataset (Ghafourian et al., 2010). Tables 5.2 and 6.1 provide the list of the vehicles and the permeants, respectively. The k_p data measured in this investigation (n = 96) were merged with the previously obtained dataset of k_p (n = 288) and the resulting dataset

was used for the QSAR development. This data set is available online through (Riviere and Brooks, 2010).

Stepwise regression analysis of different combinations of solvent properties and molecular descriptors of the penetrants resulted in a number of QSAR models from which four were selected based on the goodness of fit (r^2 values). In order to reduce the risk of chance correlations, only four descriptors were allowed in the equations. The selected equations are listed in Table 6.2. In equations (6.1)-(6.4), the letter in the brackets indicates if the variable is a descriptor for the penetrant (P) or for the vehicle (V). It can be seen that each equation consists of 2-3 penetrant descriptors and 1-2 vehicle descriptors, with equations 6.1-6.3 containing 1 combined vehicle-penetrant descriptor. In equations 6.1-6.4, Δmp is the difference between the melting point of the penetrant and that of the solvent, Wiener is the Wiener topological index (the sum of distances between all pairs of vertices in the molecular graph of an alkane (Diudea and Gutman, 1998)), δ is the Hildebrand solubility parameter, E_{HOMO} is the energy of the highest occupied molecular orbital, BP is the boiling point, N_{atoms} is the total number of atoms in the molecules, BP-MP is the difference between the boiling and melting points of a compound, and Lipole is the total lipole moment of the penetrants.

Table 6.2. QSAR models obtained from stepwise regression; N is the number of datapoints (penetrant/ vehicle combinations); S, the standard deviation; r^2, the squared correlation coefficient

EQ		N	S	r^2
6.1	Log k_p = - 0.956 - 0.00322 Δmp - 0.000320 Wiener(P) - 0.0121 BP(V) - 0.114 Lipole(P)	384	0.478	0.701
6.2	Log k_p = - 310 - 0.000315 Wiener(P) – 0.00771 δ(V).E_{HOMO}(P) – 0.0102 BP(V) – 0.0750 Lipole(P)	384	0.494	0.681
6.3	Log k_p = - 2.48 – 0.0474 N_{atoms}(P) – 0.00798 δ(V).E_{HOMO}(P) – 0.0102 BP(V) – 0.0723 Lipole(P)	384	0.516	0.653
6.4	Log k_p = -4.29 – 0.0474 N_{atoms}(P) – 0.00904 BP-MP(V) – 0.345 E_{HOMO}(P) – 0.0790 Lipole(P)	384	0.522	0.644

Considering that N_{atoms} and Wiener (Diudea and Gutman, 1998) can be regarded as size descriptors, it can be seen from Table 6.2 that all QSAR models indicate the negative effect of the penetrant's molecular size on the log k_p. Moreover, there is a negative contribution from the total lipole of the penetrants in equations (6.1)-(6.4).

Total lipole is a measure of lipophilicity distribution calculated as sum of local values of log P, similar to dipole moment (Pedretti et al., 2002). It shows lipophilicity of the molecule in a specific direction. Surfactants are expected to have high total lipole values and they are known enhancers of drug skin penetration (Ma et al., 2007). Thus, according to these equations, the less lipolar penetrants will have higher permeation rates. Chlorpyrifos has the highest total lipole value of 10.0 and caffeine has the lowest value of 0.19.

The other penetrant descriptor, which can be seen in most of the equations, is E_{HOMO}. This molecular descriptor represents the energy of the highest occupied molecular orbital. E_{HOMO} measures the nucleophilicity of a molecule. The negative relationship of this descriptor with the logarithm of the permeation rate indicates that the electron rich nucleophilic compounds such as those containing aromatic rings are the least permeable. In equations (6.2) and (6.3), the product of the penetrants' E_{HOMO} and the vehicles' solubility parameter is used. The parameter $\delta(V).E_{HOMO}(P)$ is a solvent/penetrant interaction term. This descriptor indicates that a highly nucleophilic penetrant will have a lower penetration from highly associated vehicles, i.e. those vehicles with high intermolecular interaction forces such as hydrogen bonding.

The most persistent vehicle descriptor in the QSARs is the boiling point, with a negative effect on the permeation rate of chemicals. The solubility parameter is also present in some equations. Both solubility parameter and boiling point can represent the intermolecular interaction energy of the vehicle which can result from the polarity of the solvents. Therefore the negative relationship indicates a higher skin permeation rate with the less polar vehicles. Similar results have been shown previously where the permeation coefficients of highly lipophilic compounds, nicardipine and nimodipine was increased in the less polar solvent mixtures of ethanol-water (Krishnaiah et al., 2002; Krishnaiah et al., 2004). In the case of nimopidine, 60:40 (v/v) ratio of ethanol: water was an optimum solvent mixture leading to the highest permeation rate of nimopidine, with the k_p dropping slightly at higher ethanol concentrations (Krishnaiah et al., 2004).

Solubility parameters of the solvents and the permeants have been implicated as important factors controlling skin penetration of compounds (Dias et al., 2007; Roy and Flynn, 1989). From a parabolic relationship between skin permeation rate and the solubility parameter it has been concluded that the skin has a solubility parameter of around 10 $(cal/cm^3)^{1/2}$ (Liron and Cohen, 1984). In equations (6.2)-(6.4), the solubility parameter of the vehicle has been selected by stepwise regression analysis as a significant contributor to skin permeation rate. The solvent mixtures in this study have solubility parameters of >12. Therefore, it is expected that the lower the δ value of the vehicle, the closer the value to the skin δ and therefore the higher the permeation constant should be. In equations (6.2) and (6.3), δ(V) E_{HOMO}(P) has a negative effect on the skin absorption, implying a lower skin absorption of nucleophilic drugs from polar solvents, as explained before.

In equation (6.1), the difference between melting points of the vehicle and penetrant has been selected as the most significant of all descriptors. The negative coefficient of the descriptor Δmp indicates that the melting point of the penetrants should be close to the melting point of the vehicle for better skin absorption. Therefore, since the vehicles are all liquids, this implies that penetrants with low melting points are likely to have higher absorption rates. This finding is in agreement with a previous observation where it has been shown for two optical isomers of ibuprofen that the low melting point, S enantiomer has a higher skin permeation rate than the high melting point R enantiomer (Cilurzo et al., 2010a). A similar conclusion has been made in a different investigation, when it was observed that among the alkyl analogues of cyclizine the analogue with the lowest melting point, and not the most lipophilic one, showed the highest skin penetration rate (Monene et al., 2005). Melting point has a similar effect on the intestinal absorption of drugs with low melting point drugs generally showing a higher fraction of dose absorbed from the GI tract (Chu and Yalkowsky, 2009).

In equation (6.4), the descriptor BP-MP(V), with a negative coefficient, indicates that penetration rate is slower from solvents with large boiling and meting point gaps. The difference between these two properties has been attributed to the molecular symmetry, with highly symmetrical molecules having much larger melting

points and decreased boiling points (Slovokhotov, 2007). In the solvents used in this study, the biggest difference in melting and boiling points is for propylene glycol. Therefore the vehicles containing higher concentrations of this solvent will have higher difference between melting and boiling points, leading to lower penetration of the penetrants.

Table 6.2 shows that equation 6.1 with the highest r^2 and the lowest S value has the best fit to the data. The goodness of fit is reduced from equation (6.1) to (6.4). The predictive powers of the equations were tested by an internal validation procedure explained in the methods section. Table 6.3 shows the statistical parameters obtained from this exercise.

Table 6.3. Statistical parameters obtained from internal validation of QSAR equations 6.1-6.4; N is the number of datapoints, S is the standard deviation and r^2 is the squared correlation coefficient between observed and predicted log k_p for the test sets, and MAE is mean absolute error of prediction

Equation	N	S	r^2	MAE
1	384	0.557	0.592	0.454
2	384	0.594	0.535	0.496
3	384	0.605	0.517	0.497
4	384	0.618	0.497	0.493

It must be noted that during validation tests, each set of four penetrants (dissolved in any solvent mixture) were removed once as the test set and the equation obtained for the remaining 12 penetrants were used for the estimation of the log k_p of the test set. Thus, Table 6.3 represents the results of such log k_p estimations for all the four test sets (each set containing 4 penetrants). According to the table, the predictivity of equation (6.1) is the highest among all the equations with a mean absolute error of 0.454. Equations (6.2)-(6.4) show slightly higher prediction errors for the test sets, with MAE increasing in the order of eq. (6.4) > eq. (6.2) > eq. (6.3). Therefore, it appears that the QSAR model (6.1) is robust in terms of prediction of log k_p for those new penetrants that fall within the applicability domain of the model. Figure 6.1 is the plot of observed versus predicted log k_p values using Equation 6.1.

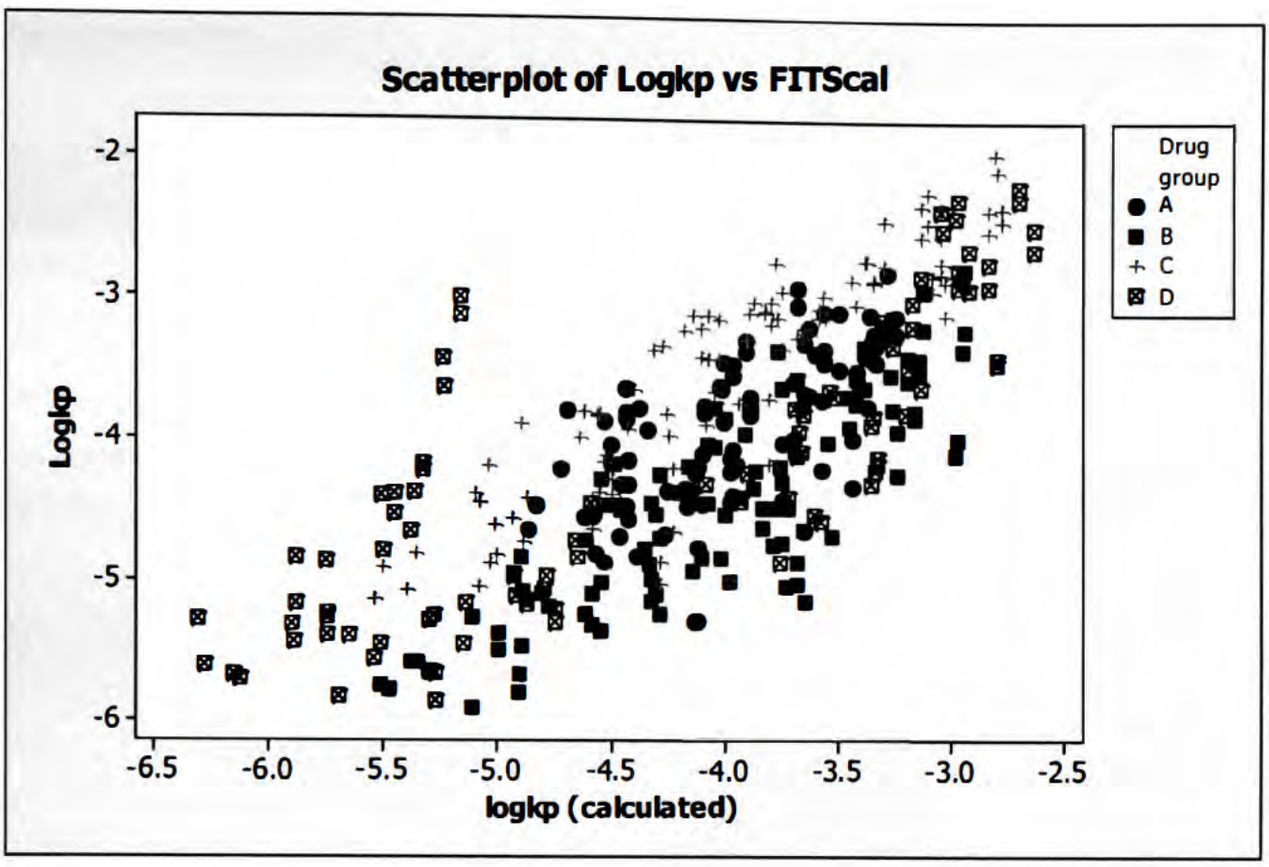


Figure 6.1. Scatterplot of observed log k_p vs. log k_p calculated by equation (6.1)

The outliers in the graph with underestimated log k_p values are testosterone dissolved in different vehicles, most notably in ethanol-water, ethanol-water-methyl nicotinate, water and water-methyl nicotinate. Testosterone is the largest molecule in the dataset in terms of the total number of atoms and the Wiener topological index, where the negative coefficient of Wiener in equation (6.1) describes the negative effect of the molecular size on absorption. Testosterone is also the most affected penetrant by the variation in solvent mixture, as it presents the largest gap between the highest and the lowest k_p from different mixture components. Considering that only k_p values from certain mixture components is higher than expected, it is concluded that the enhancing effects of these vehicles are not fully accounted for by the model and further investigations are required to explain the observed effect. Comparing equation (6.1) with equation (5.1), it can be seen that although equation 5.1 has a higher r^2 value of 0.729, the prediction accuracy is higher for equation 6.1 (MAE of 0.454 compared with 0.654 for equation 5.1). This emphasises the importance of using a diverse data set in developing a valid QSAR model.

6.3. Conclusion

In order to further understand and predict the combined effect of solvents and permeant structures on the skin absorption of chemicals, the diversity of the dataset (both penetrants and solvents) is of utmost importance. In this work, addition of the four selected chemicals expanded the diversity of our dataset. This had a positive impact on the estimation accuracy of the QSAR models, as it is evident from the results of internal validation. It can be assumed that the current QSAR could perform better in the estimation of the external data which fall within the applicability domain.

7. Statistical Evaluation of the Effect of Mixtures and Experimental Conditions on *In Vitro* Human Skin Penetration – Data from the updated EDETOX database

In vitro methods are commonly used in order to estimate the extent of systemic absorption of chemicals through skin. Due to the wide variability of experimental procedures, types of skin and data analytical methods, the resulting permeation measures e.g. flux, permeability coefficient or percentage absorbed during a certain period of exposure varies significantly between laboratories and individuals. The OECD test guidelines allow the use of a relatively flexible range of *in vivo* and *in vitro* methods for the assessment of human health risk of chemicals. A comprehensive assessment of the reliability and consistency of the available data to date for extrapolations to *in vivo* situations has been restricted due to the limitations in the availability of a wide-ranging database. The aim of this study was to investigate the effects of experimental conditions such as membrane thickness, occlusion, hydration, vehicle ingredients and mode of dose application (finite or infinite dosing) on the skin permeation flux, by using the collated data from the literature and the largest database of *in vitro* studies that is publically available, in the EDETOX project (EDETOX, 2010). In this work, an exhaustive literature review covering the period 2001-2010 was performed and the *in vitro* skin permeation data were collated and combined with the dataset extracted from EDETOX database. Only flux data obtained using human skin were analysed at this stage. The investigation focused on the effects of experimental conditions including the membrane thickness, type of exposure (finite/ infinite), pre-hydration and occlusion of the skin on the *in vitro* flux. These parameters were investigated in combination with the effects of the chemical structures of the penetrants and the formulation/mixture ingredients. The statistical techniques consisted of linear and non-linear methods of stepwise regression analysis and Regression Trees (RT).

7.1. Methods

7.1.1 The dataset

The *in vitro* fluxes of chemicals from human skin measured by flow-through or static cells were obtained from the recent literature (2001-2010) and EDETOX (Evaluations and Predictions of Dermal Absorption of Toxic Chemicals) database (EDETOX, 2010). The EDETOX database contains data from *in vitro* and *in vivo* percutaneous penetration studies involving use of different species, cell types and chemicals with a total of 2501 records for *in vitro* and *in vivo* data taken from 341 penetration studies. The EDETOX database gave information about chemical name, vehicle used, origin of the skin sample, membrane thickness, area of the membrane to which the dose was applied, cell type (flow through or static), exposure time, length of study, percentage of dose absorbed, percentage recovery, flux, permeation rate (k_p), lag time, where available, and the references to the original publications. Further information with regards to the hydration state, occlusion condition, the volume applied ($\mu l/cm^2$), dose applied ($\mu g/cm^2$) and donor concentration ($\mu g/ml$) was added to the dataset by careful inspection of the original publications. In a small number of cases, flux values were calculated by multiplying k_p by drug concentration that remained relatively constant in the vehicle, assuming that under sink conditions the drug concentration in the receiver compartment was negligible compared to that in the donor compartment.

An exhaustive literature survey was carried and the data from recent publications (2001-2010) were added to the dataset extracted from EDETOX database. The literature survey was carried out in the Web of Knowledge with the key words; skin absorption, skin penetration, skin permeation, skin permeability, dermal absorption, dermal penetration, dermal permeation and dermal permeability. From the resulting 1800 scientific publications all human *in vitro* data were extracted.

Data concerning the pre-treated skin samples with a solvent or a penetration enhancer were discarded but pre-treatment with water (hydration) was allowed in the dataset. Absorption measurements from commercial mixtures with unknown constituents or complicated formulations such as liposomes and emulsions were not

used. The final working dataset consisted of 884 studies (536 with flux values) containing 272 unique chemicals. The chemicals were either applied as neat (around 10% of the data) or formulated in simple mixtures with the majority of vehicles containing water as a constituent. In a few cases that the formulations were gels, the percentage of constituents were known. In majority of cases the exposure time was 24 h, but it varied from 0.167 to 336 h, and the sampling time between 0.167 and 336 h. The composition of the receptor fluid could vary to allow different additives, pH or solvent types. In many cases, the finite or infinite dosing conditions were explicitly specified in the literature. In other cases, if the application volume was above 100µl it was taken as 'infinite', if donor volume was between 50-100µl then provided that the percentage absorbed was less than 20% it was considered as 'infinite' or otherwise a 'finite' application. An indicator variable was generated to indicate finite or infinite dosing in the statistical analyses with the values of 2 for finite and 1 for infinite dosing. Experimental conditions under which flux was measured were explored further and whether the skin was hydrated prior to the experiment (minimum of 1 hr hydration) and whether the donor compartment was occluded was recorded in the dataset. In order to incorporate these in statistical analysis, states of pre-hydration or occlusion were given a value of '1' where skin was hydrated or occluded, and '0' when the skin was not pre-hydrated or occluded.

In the dataset, the preparation of the skin may vary from full thickness or dermatomed skin, to epidermal membranes and from frozen/ thawed to fresh skin. Skin thickness measurements were specified in many publications in mm. If only a description was provided in the literature, the full thickness skin was taken as 2 mm, epidermis as 0.8 mm and SC as 0.2 mm thick.

7.1.2. Molecular descriptors of permeants

Simplified Molecular Input Line Specifications (SMILES) of penetrants were obtained online from ChemSpider, PubChem (accessed Sep 2010), and Sigma-Aldrich (accessed Oct 2010). If the compound structure was not available in these databases, reference books or ChemBioFinder (CambridgeSoft, 2011) were used to find the molecular structure, then the structure was generated by drawing in

ACD/ChemSketch freeware software (ACD/ChemSketch, 2010) and SMILES codes were obtained. The molecular descriptors (375) were calculated using ACD labs/LogD Suite version 12.01 (ACD/LogD, 2009) and Molecular Operating Environment (MOE) version 2011.10 (Chemical Computing Group, 2011). The molecular descriptors included physical properties (e.g. partition coefficient and molecular weight), subdivided surface areas, atom and bond counts, molecular connectivity and kappa shape indices, adjacency and distance matrix descriptors, partial charge descriptors, potential energy descriptors, MOPAC descriptors, and conformation dependent charge descriptors.

7.1.3. Properties of the mixture (vehicle)

The physico-chemical properties of mixture components such as boiling point, melting point, density, log P, and solubility were obtained through SRC PhysProp database (Syracuse Research Corporation, 2010), Sigma Aldrich website, and ChemSpider. For pharmaceutical excipients such as polyethyleneglycols (PEGs), petrolatum and mineral oil the properties were obtained from Rowe et al. (2009). Average of the physicochemical properties for every solvent mixture was calculated for the liquid ingredients, e.g. boiling point of the vehicle. The effect of solid solutes (including the permeants) on boiling and melting points were calculated using the principles of the colligative properties (Sinko et al., 2011). Therefore, boiling point elevation (ΔT_b) and freezing point depression (ΔT_f) due to the dissolved material can be calculated by equations (7.1) and (7.2) respectively.

$$\Delta T_b = \text{molality} * K_b * i \qquad (7.1)$$

$$\Delta T_f = \text{molality} * K_f * i \qquad (7.2)$$

In equation (7.1) and (7.2), K_b is ebullioscopic constant specific for the solvent and i is Van 't Hoff factor, K_f is cryoscopic constant, specific for the solvent. The ebullioscopic (K_b) and cryoscopic constant (K_f) were obtained from the literature (Moore, 1972) and was averaged for the solvent mixtures.

7.1.4. Development and Validation of models

Logarithm of steady state flux showed normal distribution and therefore this was used for statistical analysis and development of the mathematical models. Before model development, the data were assessed using a simple semi-mechanistic model involving a linear relationship between log flux and simple parameters such as donor concentration (as in Fick's law of diffusion), partition coefficient and molecular size (as in Potts and Guy, 1992) and an index of molecular polarity. After establishing a preliminary linear relationship, the outliers were identified and, where appropriate, the identified outliers were removed from the dataset.

The dataset was sorted according to log flux values and partitioned into training and test sets with the ratio of three to one for training and test sets. Two main methods were used to investigate the effect of experimental variables (indicator variables for skin pre-hydration, occlusion, and finite/ infinite dosing together with the values of skin thickness), mixture properties and the structural descriptors of the permeants on the flux. These were stepwise regression analysis using MINITAB statistical software version 15.1.0.0 (Minitab Statistical Software, 2010) and non-linear method of RT (Regression Trees) in STATISTICA Data Miner software 9.1 (StatSoft, Inc., 2010). These methods can be considered as variable selection tools for the development of linear (stepwise regression) and non-linear (RT) models with best fit to the training set data. Each of these methods can also allow the user to manipulate the statistically selected variables. Therefore, interactive RT data-mining tool was utilised to evaluate the variables of experimental conditions for each split.

In the RT method, several stopping criteria were examined, including the STATISTICA default settings. These included either the minimum number of 11, 22 and 40 compounds, or the minimum fraction of 0.05, 0.02 and 0.01 to the total number of compounds for partitioning. The default values were used for the maximum number of levels set at 10 and the maximum number of nodes at 1000. The V-fold cross-validation with default settings were used in which the v value was 10.

In order to compare the validity of the RT and regression models, models were generated using training set compounds and the prediction accuracy was assessed by comparing the average error levels of the estimation of log flux for the test set compounds. The error criterion was Mean Absolute Error (MAE) calculated as the average of the absolute difference between observed and predicted log flux for the test set.

7.2. Results

7.2.1. The dataset

The collated dataset comprised work reported in a wide range of literature where the skin permeation measurements pursued a large variety of goals ranging from *in vivo / in vitro* correlation studies (Dick et al., 1995) to the pharmaceutical formulation optimisation for systemic absorption of drugs (Dias et al., 1999) which could include chemical enhancers (Copovi et al., 2006; Patil et al., 1996). Furthermore, a large volume of the literature concerns the study of the effect of experimental conditions such as the skin type and area of the skin (Wilkinson and Williams, 2002), clothing (Wester et al., 2000), pH (Sznitowska et al., 2001), mixture components (Santos et al., 2010), and receptor phase composition (Surber et al., 1991) on the *in vitro* absorption of drugs, pesticides, solvents and other compounds. Therefore large inter-laboratory and inter-individual variations are very common. In the current exercise, a data set with the greatest internal consistency is required in order to investigate the effects of some of the variable experimental conditions as well as the vehicle and the permeant chemical structures. Therefore, the dataset was initially assessed through the use of a simple semi-mechanistic model and the extreme outliers were identified. In accordance with the Fick's Law of diffusion and the well-accepted model of Potts and Guy (1992), this initial semi-mechanistic model for flux was formulated to comprise donor concentration (according to Fick's Law of diffusion), a size descriptor and lipophilicity index (log P). In addition a polarity descriptor was also incorporated in the model as informed by previous studies (Tayar et al., 1991; Liou et al., 2009). Multiple regression analysis was used to fit the data and only the

statistically significant parameters with P values below 0.05 were allowed in the model (equation (7.3)).

$$\log \text{flux} = 1.63 + 0.000002\,[\text{donor}] - 2.83\,\text{PSA/SA} - 0.00417\,\text{MV} \qquad (7.3)$$

$S = 1.25,\ r^2 = 0.376,\ N = 499,\ F = 99.5,\ P = 0.000$

In equation (7.3), [donor] is the donor concentration, PSA/SA is the polarity index represented by the ratio of polar surface area to the total surface area, and MV is the size descriptor, molar volume. It must be noted that despite the fact that lipophilicity of permeants is believed to be a major factor in skin permeation of compounds, in this case octanol/water partition coefficient (log P) was not statistically significant and therefore not included in equation (7.3). Figure 7.1 shows the plot between observed and calculated log flux by equation (7.3). The regression line between observed and calculated log flux has an intercept of 0.000 and a coefficient of 1 (Equation (7.4)).

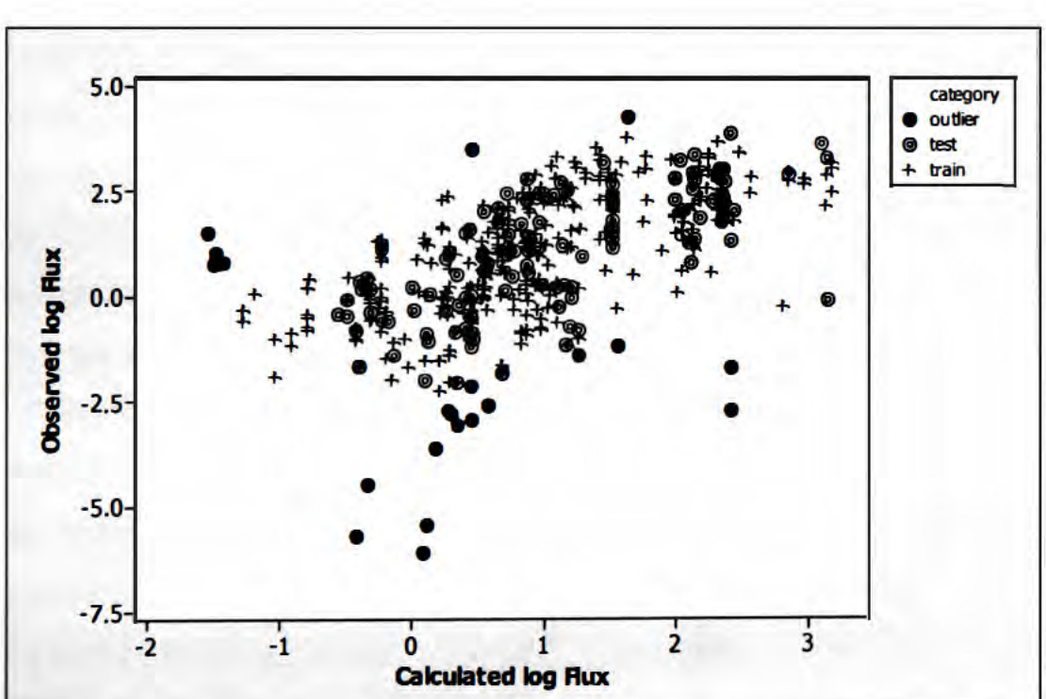


Figure 7.1. Observed log flux against calculated log flux by equation (7.3).

$$\log \text{flux (obs.)} = -0.0000 + 1.00\ \log \text{flux (cal.)} \qquad (7.4)$$

$S = 1.25,\ r^2 = 0.376,\ N = 499,\ F = 299.7,\ P = 0.000$

Chemicals with standardised residuals greater than 1.5 and less than -1.5 were marked as outliers; these are highlighted in Figure 7.1. Table 7.1 provides the list of these chemicals and the corresponding references. The outliers reported in Table 7.1 have been explained below and were removed from subsequent statistical analyses.

In Table (7.1) chemicals 1-9 are steroids with the flux values taken from Scheuplein et al. (1969) being much lower than what was expected according to equation (7.3). In fact the skin permeability measurements of Scheuplein et al have been reported to be consistently lower than those measured by other groups (Johnson et al., 1995). For example eight independent measurements of the permeability of estradiol are in good agreement with one another, but are greater than the value reported by Scheuplein et al. Therefore we decided not to use any flux values reported in Scheuplein et al. (1969) in the analysis.

Other compounds with lower than expected flux values are benzene, dibutylphthalate, cyclohexanol, mannitol and testosterone. In the dataset, there are three other sources reporting flux values for benzene of 2.4-4.4 times higher than this reference. The low flux of dibutylphthalate reported by Scott et al. (1987) may be due to the enzymatic degradation in the skin (Beydon et al., 2010), although the skin flux reported by Beydon et al. (2010) is eight times higher despite the use of full thickness as opposed to dermatomed skin. The small flux values of cyclohexanol in the two records (Table 7.1) may be due to the unusually small duration of exposure of 10 and 60 minutes (Fasano and McDougal, 2008). The flux value for mannitol by Patil et al. (1996) (reported in the table of outliers) is around four times lower than the value reported by Akhter et al. (1984) despite using full thickness as opposed to the dermatomed skin used by Patil et al. (1996). Average log flux value (-2.19) from Lee et al. (2001) was much lower when compared with other references (Patil et al., 1996; Sznitowska et al., 2001; Buist et al., 2010) that had a log flux range of -0.74 to 0.26.

Flux values for triclosan, a nicotine derivative, propranolol oleate isomers and tetrahydrofuran, were higher than those estimated by equation (7.3). The log flux reported by Moss et al. (2000) for triclosan is quite high at 3.56 which is the highest flux in the dataset after tetrahydrofuran, dimethylformamide and water. The

deviation of the nicotinic acid derivative (C5H4NCOOR, R = (CH2CHCH3O)7.29H) could be due to the poly- disperse nature of this derivative which may lead to higher absorption of lower molecular weight molecules, not captured by the molecular model generated for the average molecule. There are four data points for propranolol oleate, namely (RS)-propranolol oleate and (S)-propranolol oleate administered using saline or mineral oil (Cilurzo et al. 2010b), where the flux values are higher than those expected by equation (7.3). Although flux values for these chemicals were not available from other sources for a direct comparison, the log flux values for propranolol reported in this reference (1.26 and 1.38 from saline or mineral oil respectively) is much higher than the value reported by Ritshel et al. (1989) of -1.517. It must be noted that the latter log flux value is measured using a lower donor concentration (Ritshel et al., 1989), but even after correction for the concentration using the coefficient of [donor] in equation (7.3), the corrected log flux (-1.517) is still much lower than that in Cilurzo et al. (2010b). The log flux value of tetrahydrofuran reported by Fasano and McDougal (2008) was also quite high at 4.3 which is the highest log flux in the dataset. A possible change to the barrier property during the course of the experiment has been suggested in this publication.

Table 7.1. Outliers from equation (7.3).

No.	Chemical	Vehicle	St Residual	Reference
1	progesterone	water	-2.69	Scheuplein et al., (1969)
2	pregnenolone	water	-2.48	Scheuplein et al., (1969)
3	testosterone	water	-2.69	Scheuplein et al., (1969)
4	aldosterone	water	-4.97	Scheuplein et al., (1969)
5	hydrocortisone	water	-4.33	Scheuplein et al., (1969)
6	corticosterone	water	-3.38	Scheuplein et al., (1969)
7	cortison	water	-4.48	Scheuplein et al., (1969)
8	cortexolone	water	-3.02	Scheuplein et al., (1969)
9	17α-hydroxyprogesterone	water	-2.38	Scheuplein et al., (1969)
10	Nicotinic acid derivative (C_5H_4NCOOR, $R=(CH_2CHCH_3O)_{7.29}H$)	water	2.24	Dal Pozzo et al., (1991)
11	benzene	hexadecane	-2.01	Blank and McAuliffe, (1985)
12	dibutylphthalate	neat	-2.46	Scott et al., (1987)
13	(RS)-propranolol oleate	saline	1.60	Cilurzo et al., (2010b)
14	(S)-propranolol oleate	saline	1.59	Cilurzo et al., (2010b)
15	(RS)-propranolol oleate	mineral oil	1.61	Cilurzo et al., (2010b)
16	(S)-propranolol oleate	mineral oil	1.79	Cilurzo et al., (2010b)
17	cyclohexanol	neat	-4.23	Fasano and McDougal, (2008)
18	tetrahydrofuran	water	2.20	Fasano and McDougal, (2008)
19	cyclohexanol	neat	-3.42	Fasano and McDougal, (2008)
20	triclosan	90% aqueous ethanol	2.50	Moss et al., (2000)
21	mannitol	water	-2.48	Patil et al., (1996)
22	testosterone (average)	ethanol	-2.03	Lee et al., (2001)

7.2.2. Parameters selected by linear and non-linear methods

Indicator variables for experimental conditions of log flux measurement, namely skin pre-hydration, occlusion, finite/infinite dosing and also the skin thickness in mm were used in the stepwise regression analysis along with the molecular descriptors of the penetrants and properties of the vehicles. None of these indicator variables or the skin thickness was selected by stepwise regression analysis. However, donor concentration was the first variable to be selected (equation (7.5)).

$$\log \text{flux} = -1.92 + 0.000001\,[\text{donor}] - 0.00570\,\text{MW} + 0.00235\,\text{BP-MP(mix)} + 3.96\,\text{vsurf\_G} + 0.0137\,\text{SlogP\_VSA4} - 1.93\,\text{fiAB} - 0.343\,\text{VAdjMa} \qquad (7.5)$$

$$S = 0.948,\ r^2 = 0.558,\ N = 454,\ F = 80.41,\ P = 0.000$$

The parameters of equation (7.5) consist of donor concentration, a vehicle property representing the difference between boiling and melting points of the vehicle mixture (BP-MP(mix)), and five other parameters representing molecular descriptors of the permeants. Table 7.2 gives a brief description of these parameters.

The relevance of the donor concentration of the permeant to the flux is clear as the higher the concentration the higher the flux according to Fick's law of diffusion. The difference between the boiling and the melting points of the donor mixture has a positive effect on the flux. This had been observed previously for Riviere's dataset (Riviere and Brooks, 2010) that involved various combinations of five different vehicle ingredients and 16 permeants (Ghafourian et al., 2010a). The reason for this effect can be attributed to the better penetration of low melting point vehicles carrying the drug along into and out of the skin (Monene et al., 2005). For example, the formation of eutectic mixtures between drug and some vehicles has been proposed as the reason for skin penetration enhancement by some terpenes (Kaplun-Frischoff and Touitou, 1997; Stott et al., 1997). On the other hand, the magnitude of the gap between melting and boiling points indicates certain characteristics in the molecular structure as it is believed that more symmetrical molecules have larger melting points and decreased boiling points (Slovokhotov et al., 2007). Molecular weight of the permeants has been selected in agreement with the model proposed by

Potts and Guy (1992). The molecular descriptor vsurf_G represents the molecular globularity (Cruciani et al., 2000), and with a positive coefficient, indicates higher flux values of non-spherical molecules that may be elongated or planar shaped. SlogP_VSA4 is a lipophilicity descriptor for the permeants which is known to play a major role in skin permeation (Bouwman et al., 2008). FiAB is a molecular descriptor that describes the fraction of ionisation of molecules at pH 7.4. In the lipophilic environment of the stratum corneum ionised molecules are expected to permeate more slowly than unionised molecules (Watkinson et al., 2009). The last molecular descriptor (VadjMa) of the permeant represents the number of strong bonds (ionic, covalent, polar covalent). The number of strong bonds is related to the size of the molecule therefore the larger molecules are expected to have low permeation rates.

Table 7.2. Brief description of the parameters of regression and RT models

Descriptors	**Description**
[donor]	donor concentration (μg/ml)
BP-MP(mix)	difference between the boiling and melting points of the mixture (donor phase)
chi0	Atomic connectivity index (order 0). This is calculated as the sum of 1/sqrt(d_i) over all heavy atoms *i* with $d_i > 0$
chi0v	Atomic valence connectivity index (order 0). This is calculated as the sum of 1/sqrt(v_i) over all heavy atoms i with $v_i > 0$
chi1v_C	Carbon valence connectivity index (order 1). This is calculated as the sum of 1/sqrt($v_i v_j$) over all bonds between carbon atoms *i* and *j* where $i < j$
fiAB	fraction of molecules ionised as anion and cation at pH 7.4
GCUT_PEOE_1	The GCUT descriptors are calculated from the eigenvalues of a modified graph distance adjacency matrix. Each *ij* entry of the adjacency matrix takes the value 1/sqr(d_{ij}) where d_{ij} is the (modified) graph distance between atoms *i* and *j*. The diagonal takes the value of the PEOE partial charges. The resulting eigenvalues are sorted and the smallest, 1/3-ile, 2/3-ile and the largest eigenvalues are reported
GCUT_SLOGP_1	The GCUT descriptors using atomic contribution to log P instead of partial charge
GCUT_SMR_0	The GCUT descriptors using atomic contribution to molar refractivity using the instead of partial charge
GCUT_SMR_3	The GCUT descriptors using atomic contribution to molar refractivity instead of partial charge
Infinite/Finite	Indicator variable indicating infinite or finite exposures taking a value of 2 for finite and 1 for infinite dosing

Descriptors	Description
KierA1	First order alpha modified shape index, also correlated with molecular size (Hall and Kier, 1991)
KierA3	Third order alpha modified shape index, informing centrality of branching with large values representing location of branching at the extremities of the molecule (Hall and Kier, 1991)
logS	Log of the aqueous solubility (mol/L) calculated by MOE from an atom contribution linear atom type model with r^2 = 0.90, ~1,200 molecules
MW	molecular weight
MV	Molar volume
Occlusion	Indicator variable for occlusion of the skin during *in vitro* test
PEOE_RPC+	Relative positive partial charge: the largest positive atomic partial charge divided by the sum of the positive partial charges
PEOE_VSA_POS	Total positive van der Waals surface area. This is the sum of the van der Waals surface area of atoms with non-negative partial charges
Pre-hydration	Indicator variable for pre-hydration of the skin prior to the *in vitro* test
SlogP_VSA4	sum of van der Waals surface area of atoms with log P contributions in the range of (0.1-0.15)
SMR_VSA6	Sum of the van der Waals surface area of atoms with atomic contribution to molar refractivity in the range (0.485, 0.56)
Thickness	Skin thickness
VAdjMa	vertex adjacency information which depends on the number of heavy-heavy bonds
vsa_acc	Approximation to the sum of VDW surface areas of pure hydrogen bond acceptors (not counting acidic atoms and atoms that are both hydrogen bond donors and acceptors such as -OH)
vsa_hyd	Approximation to the sum of VDW surface areas of hydrophobic atoms
vsurf_CW3	Capacity factor representing the ratio of the hydrophilic surface over the total molecular surface. These are calculated at eight different energy levels (from -0.2 to -6.0 kcal/mol) (Cruciani et al., 2000)
vsurf_D6	Volume that can generate hydrophobic interactions. VolSurf computes hydrophobic descriptors at eight different energy levels (from -0.2 to -1.6 kcal/mol) (Cruciani et al., 2000)
Vsurf_EWmin1	The lowest hydrophilic interaction energy
vsurf_G	The molecular globularity – how spherical a molecule is, where values above 1 is non-perfect spheres (Cruciani et al., 2000)
vsurf_HB5	H-bond donor capacity, representing the molecular envelope which can generate attractive H-donor interactions with carbonyl oxygen probe. The descriptors are computed at six

Descriptors	Description
	different energy levels (from -1 to -6 kcal/mol) (Cruciani et al., 2000)
vsurf_W1	Hydrophilic volume describing the molecular envelope which is accessible to and attractively interacts with water molecules at eight different energy levels (from -0.2 to -6.0 kcal/mol) (Cruciani et al., 2000)
vsurf_Wp2	Polar volume (Cruciani et al., 2000)
wienerPath	Wiener path number: half the sum of all the distance matrix entries (Cruciani et al., 2000)

It can be noted that skin thickness or the indicator variables for the experimental conditions, namely pre-hydration, finite/ infinite dosing or occlusion, are not selected by stepwise regression analysis (equation (7.5)). This indicates that, in comparison with some of permeant or vehicle properties, these experimental conditions are more minor contributors to the overall observed flux values from many sources. On the other hand, we may also attribute this to the incorporation of extremely high number of permeant molecular descriptors in stepwise regression analysis (a total of 375 descriptors) in comparison with the few variables of experimental conditions, leading to inadequate variable selection by this statistical method.

Similarly the model generated by RT did not include any experimental conditions, only donor concentration and three molecular descriptors for the penetrants (Figure 7.2). The selected descriptors have been defined in Table 7.2. Table 7.3 gives the statistical parameters of the RT model. According to RT model (7.1), the requirements of a high skin flux are high donor concentration, small positively charged molecular surface area (PEOE_VSA_POS), large surface area of non-acidic hydrogen bond acceptors such as ether and ketone groups (vsa_acc), with a complex effect of hydrophobic volume (vsurf-D6) probably indicating the negative effect of molecular size at nodes ID 4, 5, 14 and 15, and the positive effect of hydrophobicity at nodes ID 6, 7, 10 and 11.

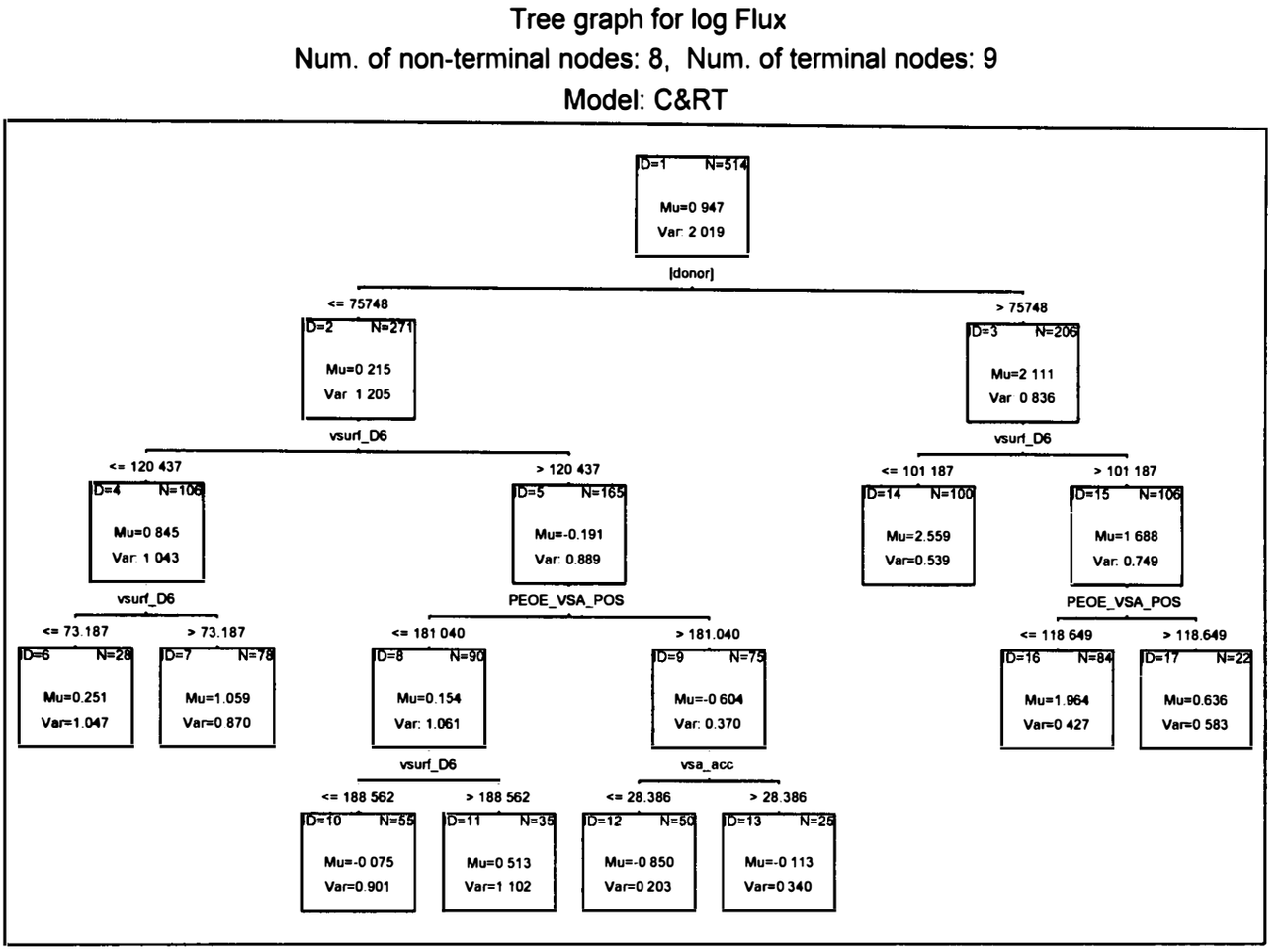


Figure 7.2. The RT model (7.1), N is the number of data points, Mu is the average and Var is the variance of log flux.

Table 7.3. Statistical parameters of RT models

RT model	MAE	Risk Estimate	Standard Error	
(7.1)	0.625	0.643	0.047	Train
		0.828	0.057	V-fold
(7.2)	0.638	0.676	0.050	Train
		0.859	0.061	V-fold
(7.3)	0.628	0.655	0.051	Train
		0.859	0.061	V-fold
(7.4)	0.597	0.569	0.039	Train
		0.859	0.061	V-fold
(7.5)	0.587	0.585	0.046	Train
		0.859	0.061	V-fold
(7.6)	0.573	0.552	0.041	Train
		0.859	0.061	V-fold

7.2.3. Effect of membrane thickness

Skin thickness is thought to play a significant role in dermal absorption of chemicals. Permeation through viable full thickness skin membranes has been shown to be less than permeation through only the epidermis (Cnubben et al., 2002). The membrane thickness in the dataset varied between 0.2 for SC and 2 mm for the full thickness human skin. The statistical significance of membrane thickness was investigated by linear and non-linear procedures. As membrane thickness was not automatically selected by stepwise regression analysis (equation (7.5)) it was manually incorporated in the regression analysis and the statistics were inspected. Regression analysis with inclusion of thickness (in mm) resulted in equation (7.6).

Log flux = - 1.67 + 0.000001 [donor] - 0.00561 MW + 0.0140 SlogP_VSA4 - 1.95 fiAB + 0.00192 BP-MP(mix) + 3.82 vsurf_G - 0.312 VAdjMa - 0.201 Thickness (7.6)

S = 0.943, r^2 = 0.564, N = 454, F = 71.9, P = 0.000

Although r^2 of equation (7.6) shows only a moderate improvement to equation (7.5), the thickness parameter is statistically significant in this equation (P = 0.014). Wilkinson et al. (2004, 2006) studied the influence of skin thickness on percutaneous penetration using caffeine, testosterone, butoxyethanol and propoxur. They concluded that a complex relationship exists between skin thickness, lipophilicity of the penetrant, and percutaneous penetration and distribution. Therefore, due to the uneven effect of skin thickness on the penetration of different chemicals of varied lipophilicity (or other physicochemical properties), a linear relationship such as equation (7.6) cannot adequately represent the effect of thickness.

Accordingly, skin thickness was incorporated in the RT analysis in the first split and the tree was allowed to select other parameters of highest statistical significance. RT model (7.2) presented in Figure 7.3 involves skin thickness, donor concentration and four molecular descriptors of the penetrants (see Table 7.2 for the description of parameters and Table 7.3 for the statistical parameters). According to this model, *in vitro* skin flux is higher with thin skin samples, large donor concentrations, low polarity index (high GCUT_PEOE_1), small positively charged molecular surface

area (PEOE_VSA_POL), small hydrophilic volume (vsurf_W1) and molecular size (chi1v_C).

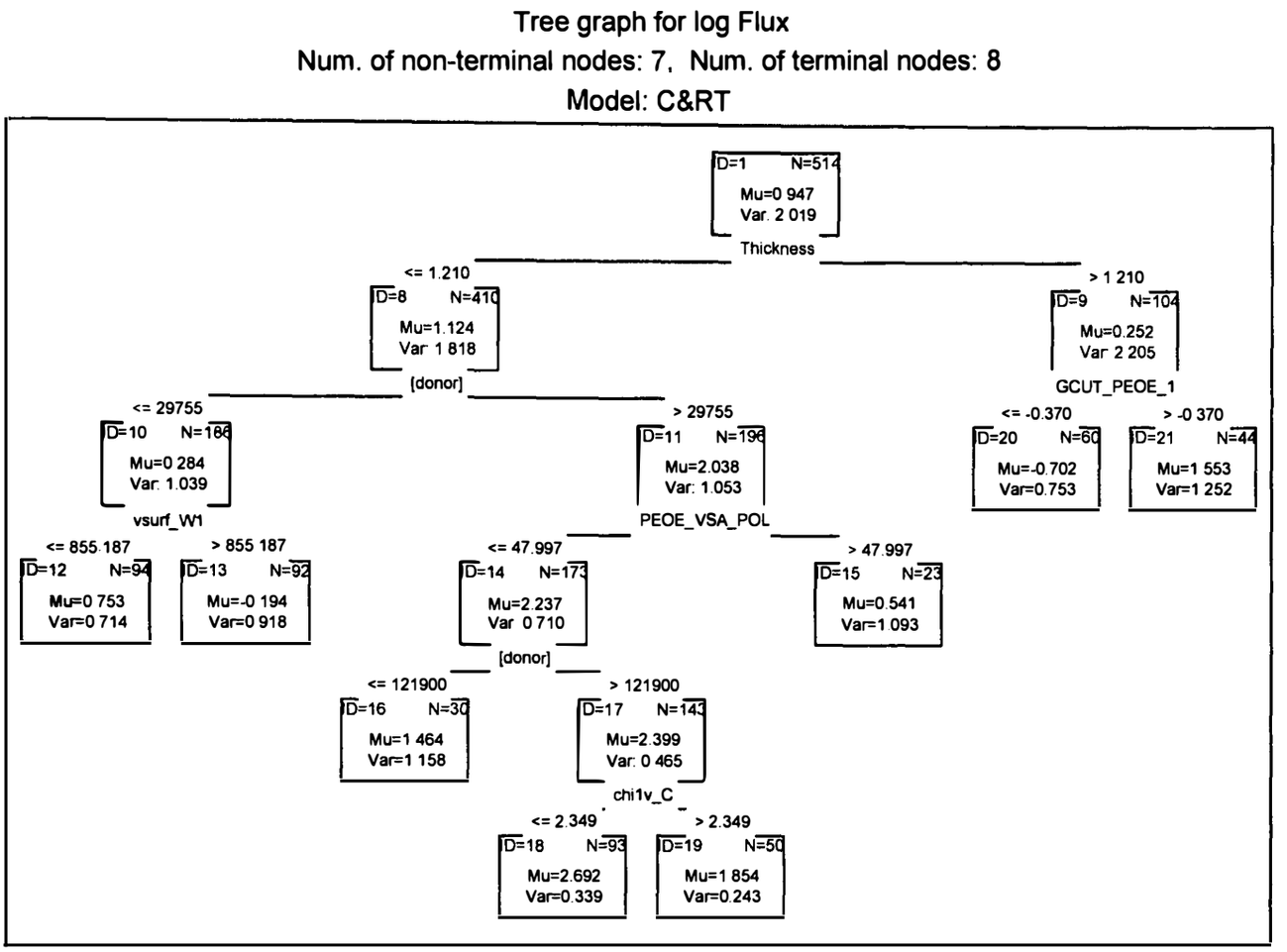


Figure 7.3. RT model (7.2) incorporating membrane thickness for the first partitioning, N is the number of data points, Mu is the average and Var is the variance of log flux.

7.2.4. Effect of finite or infinite dosing

The dataset included in this work contained steady state and maximum flux obtained under finite and infinite dose exposures. In infinite dose, the concentration of the solution applied to the skin does not significantly change over time. Therefore a maximum flux can be achieved and maintained during the course of the experiment (steady state flux). However in finite dose exposures the amount of test preparation applied to the skin will reduce over time and therefore the maximum flux cannot be maintained. In order to incorporate the dose exposure condition, an indicator variable taking a value of 2 for finite and 1 for infinite condition was used. Out of 513 flux

values with known exposure conditions, 143 and 370 were obtained under finite and infinite exposure conditions, respectively. This indicator variable was not selected by stepwise regression (equation (7.5)) or RT method (Figure 7.2). However, the graph between observed log flux values and the log flux calculated by equation (7.5) indicates differing lines of best fit to the data obtained under infinite or finite dose exposures (Figure 7.4). These regression lines can be compared using general linear model (GLM). The results, reported in Table 7.4, indicate statistically different slopes and intercepts for finite and infinite exposures. Moreover, when included in the regression analysis, the indicator variable for finite/ infinite dose was statistically significant with a P value of 0.000 (equation (7.7)).

Log flux = - 1.08 + 0.000001 [donor] - 0.00592 Weight + 0.00992 SlogP_VSA4 - 1.85 fiAB + 0.00230 BP-MP(mix) + 3.55 vsurf_G - 0.293 VadjMa - 0.391 Infinite/Finite (7.7)

S = 0.936, r^2 = 0.572, N = 453, F = 74.1, P = 0.000

The negative coefficient of Infinite/Finite (indicator variable) indicates higher flux values when measured under infinite dose exposures.

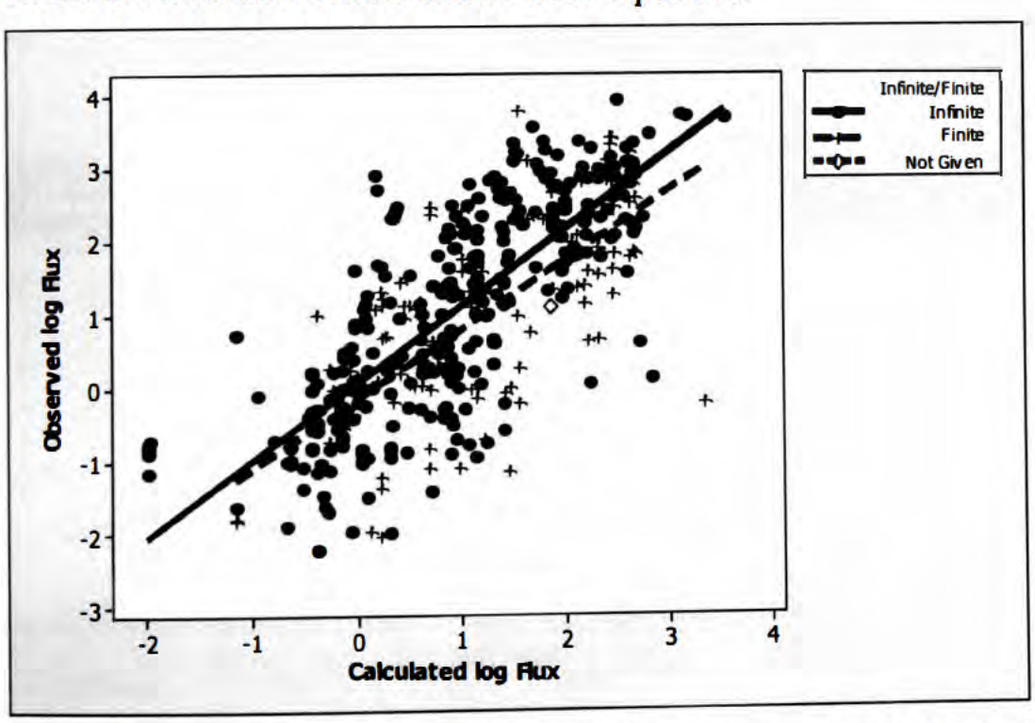


Figure 7.4. Observed against calculated log flux under infinite or finite dosing conditions.

The RT model with the inclusion of this indicator variable (RT (7.3)) is presented in Figure 7.5. The RT model indicates somewhat higher average flux for infinite exposure experiments. This tree also incorporates donor concentration and five molecular descriptors for the penetrants (descriptors explained in Table 7.2 and the statistical parameters of the tree in Table 7.3). According to this model, compounds will have higher *in vitro* flux values when applied in infinite doses, with high donor concentrations, and if they are more lipophilic (GCUT_SLOGP_1), have smaller molecular size (chi0 and GCUT_SMR_3) with high polarisability (SMR_VSA6) or hydrophilic surface (vsurf_CW3).

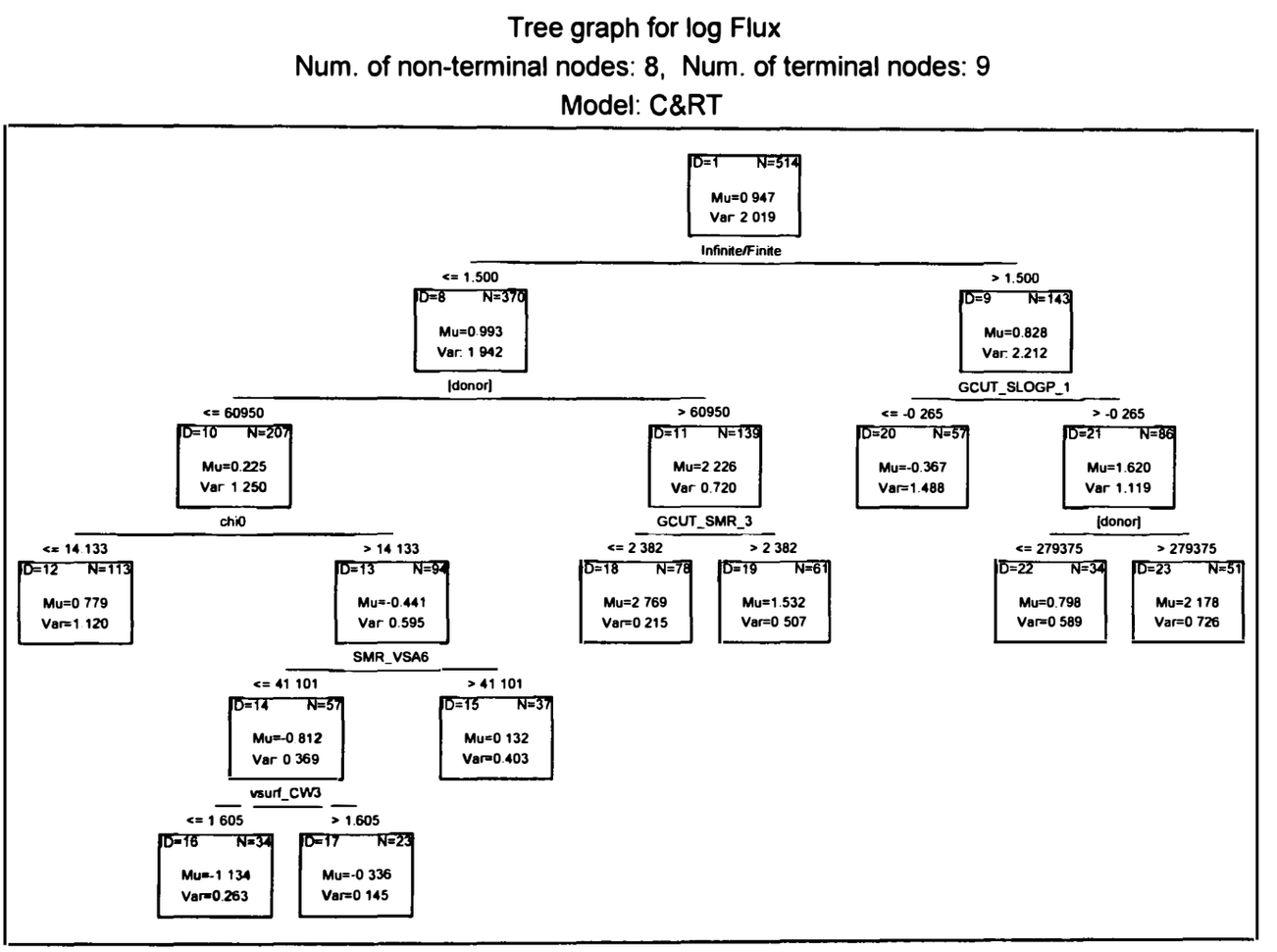


Figure 7.5. RT model (7.3) incorporating indicator variable for infinite or finite dose application, N is the number of data points, Mu is the average and Var is the variance of log flux.

7.2.5. The effect of skin pre-hydration

The stratum corneum normally contains 5-20% water but, when hydrated, it can contain up to 50% water. Hydration can affect the permeability of the skin to chemicals (Scheuplein and Blank, 1971; Roberts and Walker, 1993; Rawlings and Matts, 2005). In many skin permeation studies, the experimental procedures involve pre-hydration of the skin before the start of the experiment. This is most common in infinite dose procedures in order to maintain the consistency of the membrane during the course of the experiment. In this dataset, 187 data points used pre-hydrated skin and 317 data points employed dry skin. An indicator variable was used for skin pre-hydration taking the value of 1 when the skin was pre-hydrated for at least one hour prior to the experiment and the value of 0 when this procedure was not used. It can be seen in equation (7.5) that this indicator variable has not been selected by stepwise regression analysis. Figure 7.6 identifies the lines of best fit to the data obtained with pre-hydrated or dry skin. Comparing these lines by GLM (Table 7.4) shows that pre-hydration of the skin does not affect the slope of the line, although the intercepts are statistically different. When used in combination with descriptors of equation (7.5), the indicator variable for skin pre-hydration is not statistically significant at $P<0.05$. However, with a P value of 0.077 pre-hydration of skin has a positive effect on skin flux.

One reason for the insignificant effect of pre-hydration could be attributed to the fact that at least with infinite dose experiments, SC can quickly hydrate during the course of the experiment. Moreover, even in finite dose *in vitro* experiments many studies are conducted under occluded conditions, which may lead to some levels of hydration. On the other hand, the extent of the hydration-related permeability change for different chemicals is not well elucidated. For example, it has been shown that the increase of hydration (because of occluded conditions) does not always guarantee an increase in penetration rates (Bucks et al., 1991). Comparing the skin permeability measure obtained from a Fickian diffusion model and that obtained from the transient skin permeation profiles, Tang et al. (2002) reported significantly increased skin permeability due to hydration for highly hydrophilic compounds while skin permeation of lipophilic compounds were comparable between the hydrated and non-hydrated states of skin.

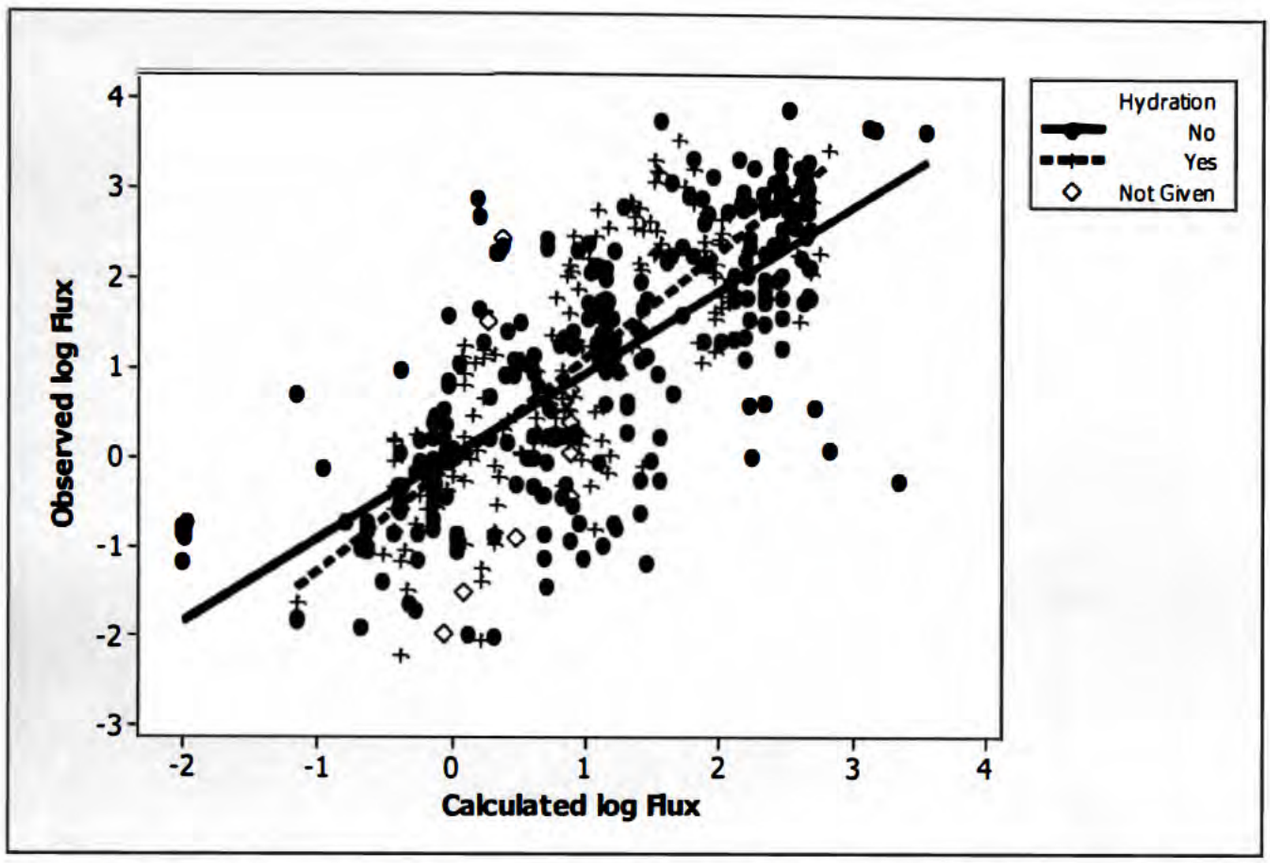


Figure 7.6. The effect of skin pre-hydration before the dose application in *in vitro* tests

In conclusion, as with membrane thickness, hydration of skin may have different levels of effects on the flux of compounds with varying physico-chemical properties. Therefore, the effect is expected to be captured better in the non-linear models such as RT. The RT model incorporating the indicator variable for skin pre-hydration has been presented in Figure 7.7 (RT (7.4)). The model indicates higher average flux from pre-hydrated skin (average log flux difference for the two groups is 0.144). RT (7.4) also includes donor concentration, five molecular descriptors for the permeants and boiling point of the donor mixture. The parameters of the model have been described in Table 7.2 and the statistical parameters are presented in Table 7.3. According to this RT model, compounds with small molecular size (KierA1), applied in higher concentrations, especially those with small hydrophobic surface area (vsa_hyd), will be absorbed very well from pre-hydrated skin samples. Similarly, from dry skin samples, compounds with small molecular size (chi0v) applied in high concentrations are absorbed with the highest flux, while those applied in lower concentrations, with a small molecular size (wienerPath), having a

high polarisability (GCUT_SMR_0), if administered in a low boiling point vehicle are absorbed the least.

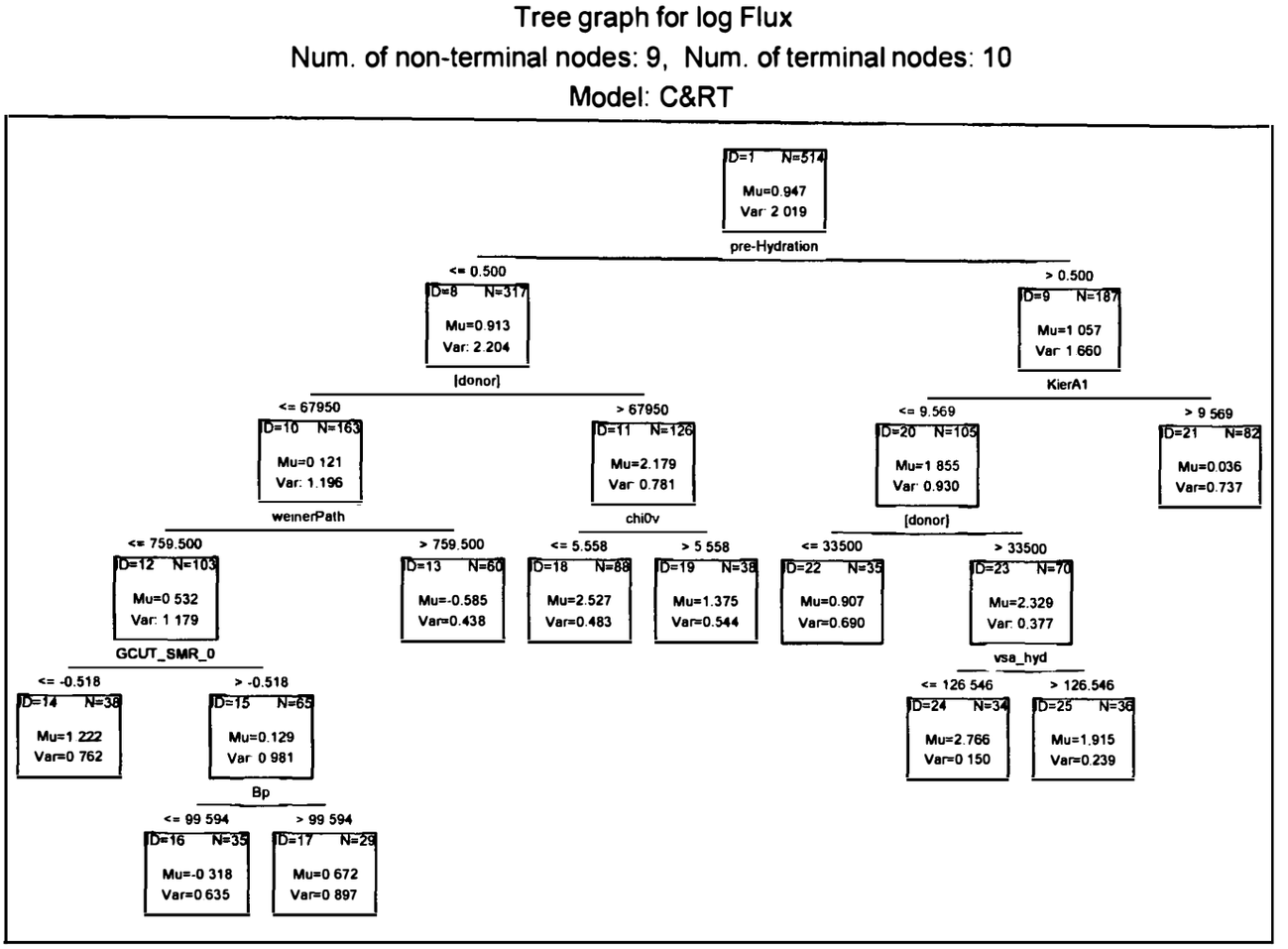


Figure 7.7. RT (7.4) incorporating indicator variable for skin pre-hydration for the first partitioning, N is the number of data points, Mu is the average and Var is the variance of log flux.

7.2.6. Effect of occlusion of the skin

Occlusion of the skin can lead to the gradual hydration of the skin during the course of the *in vitro* tests even during finite exposures where only a small volume of dose is applied. Therefore, as with hydration, flux values of various compounds are expected to be affected by occlusion of the skin. Out of 481 flux values with reported occlusion condition, 287 and 194 were performed under occluded and non-occluded conditions, respectively. Occlusion was represented by an indicator variable taking a value of 1 for occluded and 0 for non-occluded conditions. This indicator variable

was not selected by stepwise regression or RT models. Moreover, when it was used in the regression analysis along with the descriptors of equation (7.5), this parameter was not statistically significant (P = 0.506). Also, the graph between observed and calculated log flux using equation (7.5) gives similar slopes but differing intercepts for the data obtained under occluded or non-occluded conditions (see Figure 7.8 and GLM results reported in Table 7.4). To explore the effect of the occlusion parameter further, it was incorporated in RT analysis and RT model (7.5) was obtained (Figure 7.9). RT (7.5) indicates that the average flux obtained under occluded conditions is higher than the average flux measured under non-occluded conditions by 0.66 log units. This may be explained by the effect of occlusion on skin hydration, or evaporation of the volatile penetrants and/or the vehicles. In the study of Sartorelli et al. (2000) a 5 to 10 fold increase in permeability of the SC was noted when the skin was occluded. In the dermal absorption study of Jung et al. (2003) where catechol was applied in ethanol, occluded conditions resulted in 78% of the applied dose permeating into the receptor fluid, compared with 55% in dermal samples that were not occluded.

On the other hand, most finite-exposure skin absorption experiments are performed under non-occluded conditions and hence the lower flux values are expected to be achieved due to the lower available doses in the donor compartment. Therefore the statistical significance of occlusion may only represent the importance of donor concentration as in RT model (7.1).

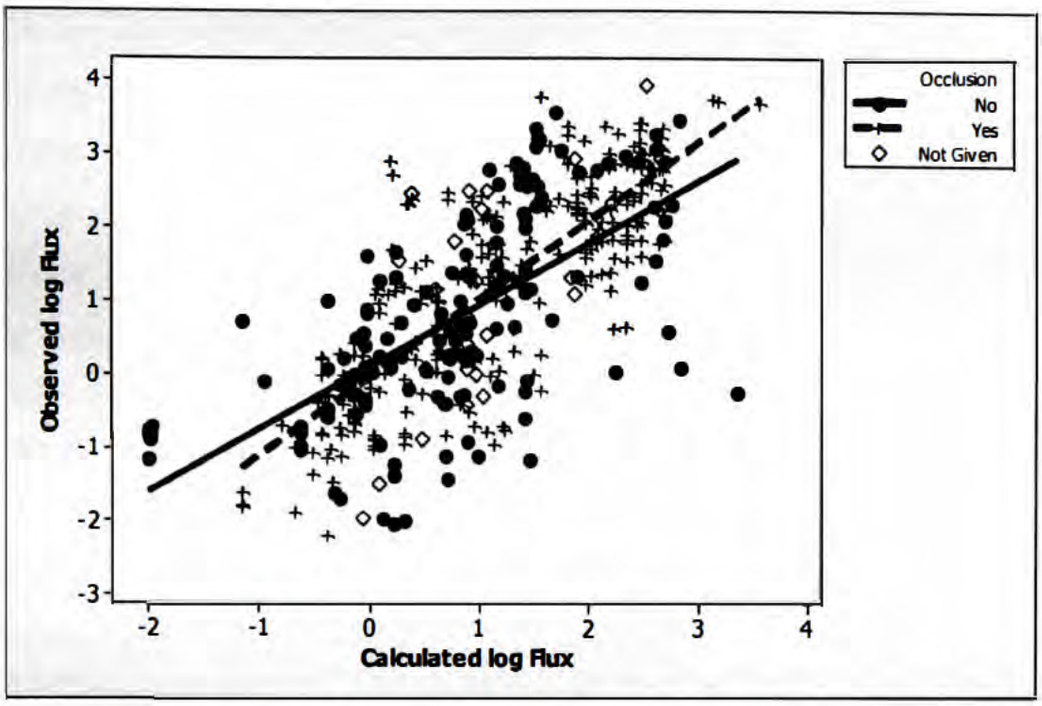


Figure 7.8. The effect of skin occlusion during the *in vitro* measurements

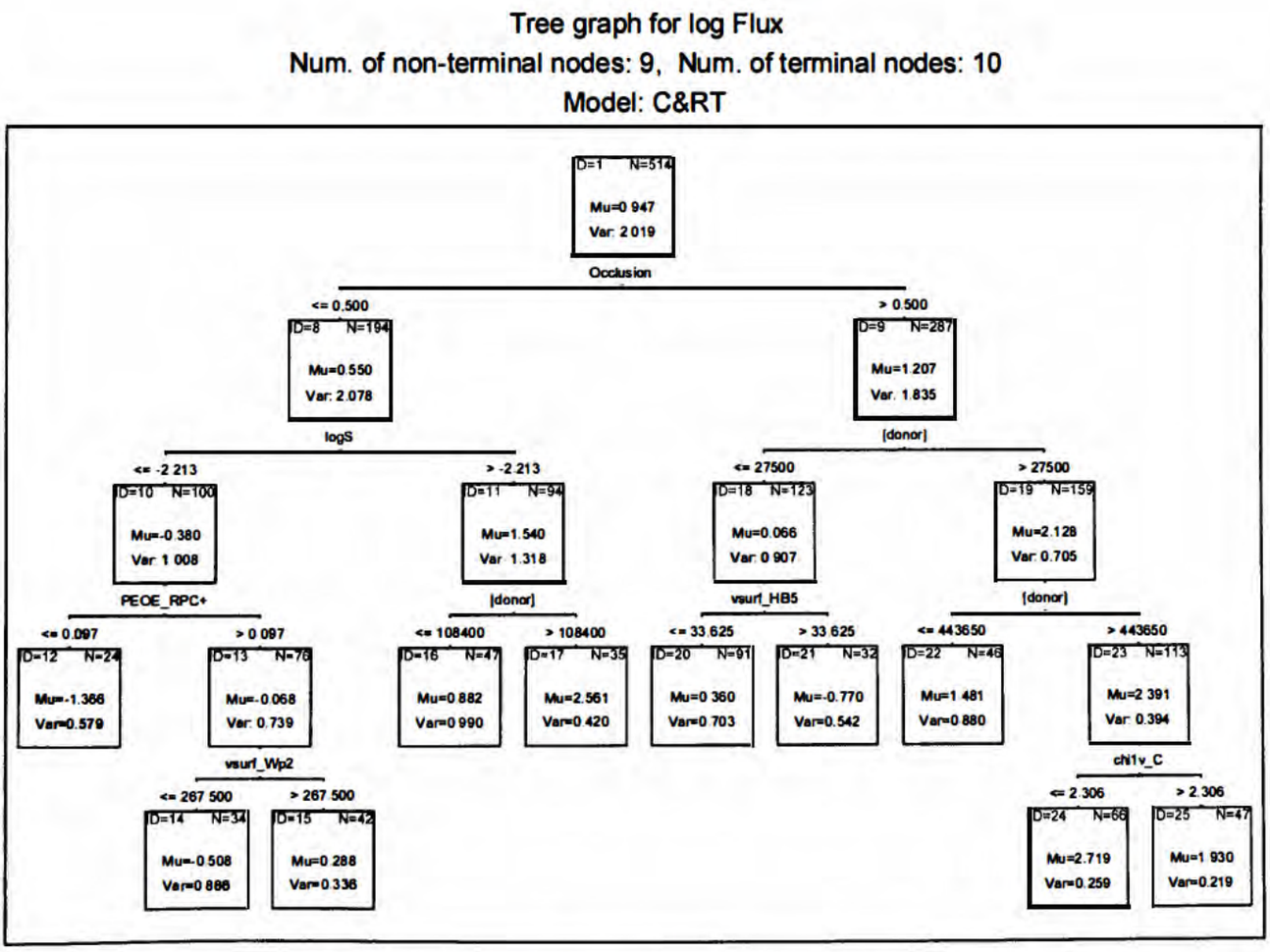


Figure 7.9. RT model (7.5) incorporating occlusion in the first partitioning, N is the number of data points, Mu is the average and Var is the variance of log flux.

Other parameters of RT (7.5) indicate a higher flux from occluded skin samples for chemicals applied in higher concentrations having a small molecular size (chiv_C), while those applied in lower doses will have low flux values especially if the molecules have a high hydrogen bonding donor capacity (vsurf_HB5). From the non-occluded skin samples, compounds with relatively high solubility (logS) applied in high doses also show high flux values, while those with low solubility, having a large sum of positive atomic charges (leading to low relative positive charge, PEOE_RPC+) or low polar volume (vsurf_Wp2) show the least flux values.

Table 7.4. Results of GLM comparing the regression lines for data obtained under various experimental conditions

Parameter	Intercept		Slope	
	F	P	F	P
Hydration	0.62	0.43	7.71	0.006
Occlusion	1.25	0.264	6.42	0.012
Finite/ infinite dosing	2.07	0.151	1.15	0.285

7.2.7. Effect of Vehicle

Stepwise regression analysis selected the difference between boiling and melting points of the vehicle mixtures to represent the effect of vehicle on the flux of the permeants (equation (7.5)). However, a vehicle descriptor is missing from most RT models with the exception of RT (7.4) which includes a vehicle descriptor (boiling point of vehicles). As with the other experimental parameters, effect of vehicle properties on the flux was further analysed by incorporating boiling point, melting point and the gap between these two in the interactive trees. The tree obtained using BP-MP(mix) was selected and reported in Figure 7.10 (RT model (7.6)), as it had the lowest standard error. The descriptors of the model have been described in Table 7.2. According to this model, flux is higher if the donor mixture has a higher gap between melting and boiling points and the molecular size of the permeant is small (large chiv_C), especially if applied in higher concentrations. If BP-MP(mix) is small, then molecules with small molecular size (KierA1) still have higher flux values than the large molecular size compounds specially if applied in higher concentrations (ID 13 in figure 10), or lower concentration when the KierA3 is large (ID 15). According to

Table 7.3, the RT (7.6) has the lowest mean absolute error amongst all RT models, which may indicate the high significance of the vehicle properties in the skin absorption.

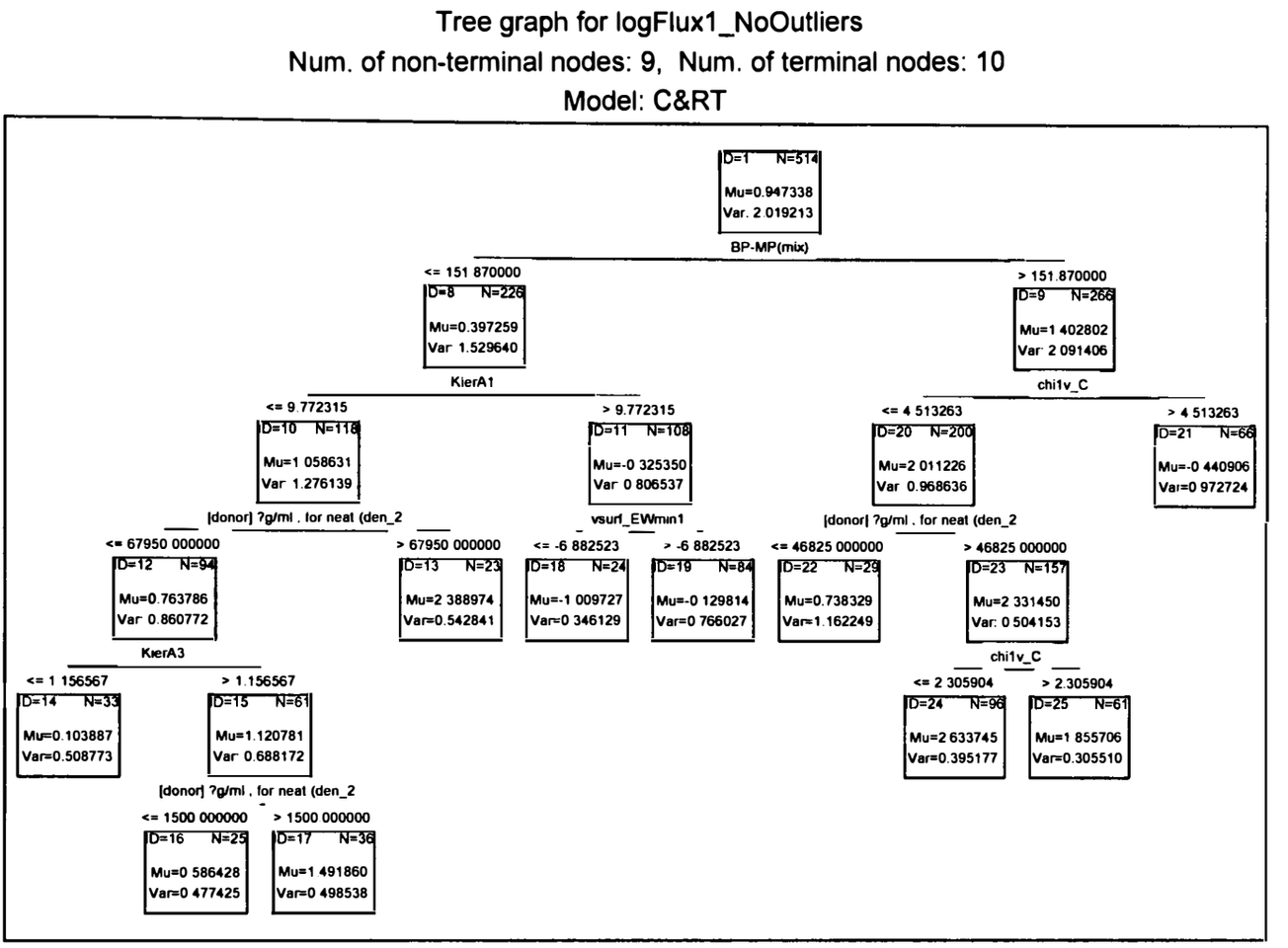


Figure 7.10. RT model (7.6) incorporating BP-MP(mix) as the first parameter for partitioning, N is the number of data points, Mu is the average and Var is the variance of log flux.

7.2.8. Validation of the models

Regression models (equations (7.5) – (7.7)) and RT models (7.1) – (7.6) were validated by developing the regression equations and RT trees for the training set, and then estimating the flux values for the test set using these models. The models obtained for the training set have been reported in the supporting material (Appendix II). Table 7.5 gives the mean absolute error of log flux estimation for the test set compounds and the number of test set compounds for which the models are able to provide the estimation. The lowest average error is achieved with RT model (7.5)

followed by the RT (7.4) and then regression equation (7.7) and RT (7.3). The most valid model, RT (7.5) involves the occlusion indicator variable, donor concentration, and five permeant descriptors implying the negative effects of molecular size and hydrogen bonding donor ability on flux from occluded skin, and positive effect of aqueous solubility, and polar volume and negative effect of relative positive charge when the skin is not occluded. It must be noted that the number of test compounds for which estimation has been made possible is different, depending on the availability of the model parameters.

Table 7.5. Mean Absolute Error (MAE) for the test set using regression and RT models, N is the number of test set data estimated by the model

			Test Set		Training Set	
Method	Experimental Parameter	Model	MAE	N	MAE	N
Regression	-	Equation (7.5)	0.689	112	0.761	342
	Skin thickness	Equation (7.6)	0.697	112	0.754	342
	Exposure type	Equation (7.7)	0.682	112	0.752	341
RT	-	RT (7.1)	0.747	119	0.708	358
	Skin thickness	RT (7.2)	0.807	119	0.731	358
	Exposure type	RT (7.3)	0.682	119	0.653	357
	Pre-hydration	RT (7.4)	0.607	110	0.610	335
	Occlusion	RT (7.5)	0.578	108	0.589	336
	BP-MP(mix)	RT (7.6)	0.722	112	0.650	343

7.2.9. Conclusion

Inter-laboratory and inter-individual variations are very common in the *in vitro* measures of skin permeation. This has been attributed to a number of experimental variables including skin samples' thickness differences, skin hydration, occlusion of the skin, infinite or finite dosing and vehicle ingredients. The dataset gathered here provided an excellent resource for investigating the effects of various parameters on the skin penetration flux from *in vitro* studies. The statistical analysis involved linear regression with stepwise variable selection and non-linear regression tree (RT in STATISTICA). The indicator variables for skin pre-hydration, occlusion, finite/

infinite dose and skin thickness were not selected by the variable selection methods of stepwise regression or RT. The selected parameters were donor concentration, several molecular descriptors of the permeants and, in the case of stepwise regression, the difference between boiling and melting points of the donor mixtures. This could mean a higher impact of the permeant concentration and molecular structure, and the vehicle properties on the skin flux.

The significance of each of the experimental conditions and BP-MP(mix) were studied further by inclusion of these parameters in regression and interactive RT models. Within the linear regression equations the order of significance of experimental variables, as deduced from the P-values, were BP-MP(mix), indicator variable for mode of exposure (finite or infinite) ($P = 0.000$), membrane thickness ($P = 0.014$), and skin pre-hydration ($P = 0.077$), with occlusion of the skin not being significant ($P > 0.10$). Moreover, validation of the models also revealed a lower average error (for test set) for the model incorporating mode of exposure as opposed to the skin thickness.

According to the mean absolute error of the calculated values (reported in Table 7.3), the order of significance of the non-linear RT models is models incorporating: BP-MP(mix), skin occlusion, skin pre-hydration, and mode of exposure. However, validation of these models indicated the highest validity of the model incorporating occlusion state of the skin (RT (7.5)) followed by the model incorporating both skin pre-hydration and boiling point of the donor phase as experimental parameter (RT (7.4)). It is interesting to note that linear and non-linear models incorporating mode of exposure (finite or infinite dosing) result in a similar level of estimation error. This conclusion, in association with the fact that experimental parameters of occlusion and pre-hydration states perform better in non-linear models, indicates that the effects of occlusion and skin pre-hydration on flux are complex and non-linear.

In both regression and RT models donor concentration of the permeant is the first and the most important parameter related to the skin flux. This is expected since concentration is the driving force for passive diffusion of molecules across the skin. A variety of permeant parameters have been employed in the models, with majority implying the negative impacts of the large molecular size and hydrophilicity, while a

certain level of lipophilicity and polarisability has been indicated as a positive effect on the flux.

The statistical analyses and models reported in this work provide a suitable method for homogenising the *in vitro* flux values measured under varying experimental and exposure conditions and will provide reasonable estimates of the flux values under other experimental conditions. It is expected that the results will benefit *in vivo* estimations using the *in vitro* flux estimates.

8. Validated models for estimation of *in vitro* flux through human skin

Validation is an imperative part of QSAR studies. Although many QSAR specialists still apply internal validation approaches as an adequate and reliable measure of model predictivity (Hawkins et al, 2003; Kraker et al, 2006; Helma 2004), the predictive ability of the model for the external chemicals that have never been included for the training of the model cannot be verified. In fact, while a good performance in internal validation is a necessary condition for robustness of a model, it alone may not be sufficient indicator of model predictivity for external chemicals (Gramatica, 2007).

The external validation set can be set aside at the beginning of the model development, so that they are not involved in variable selection. This procedure will provide a more robust measure of the model validity. The collated dataset (from EDETOX database and other sources) is large enough to allow external validation. Accordingly, in this current work, QSAR models were developed for the training set and validated externally. Several statistical techniques for variable selection and model development were examined. The methods included stepwise regression analysis, regression trees, boosted trees and support vector machines. The experimental conditions, vehicle properties and molecular descriptors, as explained in Chapter 7, were the model predictors. In addition two new scales were defined to account for the skin hydration level or, in case the donor phase was not aqueous, the level of skin wetting.

8.1. Methods

8.1.1. The dataset

The dataset has been explained in depth in section 7.1. The experimental parameters, vehicle properties and molecular descriptors have been explained in section 7.2, with the addition of two parameters indicating the wetting and hydration states of the skin. The dataset contains information on whether the skin was hydrated before the

experiment, if it was occluded during the experiment, the type of the solvent (vehicle), and the volume of donor phase applied to the skin. So in order to estimate the extent of hydration and, in case the solvent is not water, the extents of wetting, two scales were developed as described below in Table 8.1.

In order to assign the hydration levels, parameters such as pre-hydration, occlusion and volume applied were taken into account (Table 8.1). When the volume applied was less than 100 μl and the skin was not pre-hydrated or occluded, a 0 value was assigned to both hydration and wetting scales. If the volume applied was between 100 and 500 μl and the skin was not pre-hydrated nor occluded, the scale was 1. When the volume applied was between 100 and 500 μl and the skin was not pre-hydrated but it was occluded, the wetting scale was given a value of 2. If the volume applied was 500 μl or larger and the skin was not pre-hydrated, the wetting scale was given a value of 3 regardless of occlusion state. Finally if the skin was pre-hydrated before the experiment, its data points were given a 4 on the scale if it was also occluded, regardless of solvent volume. Table 8.1 provides the details of the rules.

Table 8.1. Criteria for hydration estimation scale

Pre-hydration	occluded	Applied volume	Wetting scale	Hydration scale
Yes	Yes	Any	4	Same as the wetting
No	Both	500 or more	3	scale, unless the
No	Yes	100-500	2	vehicle contained less
No	No	100-500	1	than 25% water, in
No	Yes	Below 100	1	which case the
Yes	No	Below 100	1	hydration scale was 0
No	Both	Below 100	0	

8.1.2. Development and validation of models

The models were developed for log flux using the linear stepwise regression analysis in the MINITAB statistical software version 15.1.0.0 (Minitab Statistical Software, 2010). Furthermore, non-linear models were developed in STATISTICA Data Miner software version 9.1 (StatSoft, Inc., 2010); these were general and interactive

regression trees (RT), boosted trees and support vector machines (SVM). For the development of predictive models, external validation was conducted. Therefore, the validation set was excluded from analysis from the beginning. Mean Absolute Error (MAE) values were calculated for the training and validation sets.

Regression trees (RT)

Stepwise regression and RT methods have been described in section 7.4. General RTs were made with log flux as the dependent variable and the predictors were selected by this statistical analysis. Several stopping criteria were examined, including the STATISTICA default settings. These included either, the minimum number of 11, 22 and 40 compounds, or the minimum fraction of 0.05, 0.02 and 0.01 to the total number of compounds for partitioning. The default values were used for the maximum number of levels set at 10 and the maximum number of nodes at 1000. The V-fold cross-validation with default settings were used in which seed for random number generator was set to 1 and the V value to 10.

After the development of the general RT models, interactive RT models were examined where the experimental parameters, membrane thickness, wetting and hydration scales, indicator variables for finite or infinite dosing, pre-hydration and occlusion were manually selected as the first splitting parameter. With wetting and hydration scales as the partitioning descriptors, the splits were made at 3.5, 2.5, 1.5 and 0.5 values. The stopping criteria were: Prune on variance, Min n 22, Min n in child node 22, Max n levels 3 and Max n nodes 1000.

Boosted trees

Boosted trees were developed using STATISTICA software. Each tree made only one split (stopping criterion of maximum number of nodes = 3). Various combinations of parameters, learning rate, number of additive terms, and subsample proportion, were examined and the best results were selected. The number of additive terms was adjusted to the values shown in the table, so that the errors for the training and the test sets were similar.

Support Vector Machines

Machine learning using support vector machines was performed using STATISTICA software. The analysis was regression type ε-SVM (type 1). The parameters were capacity value of 10, ε value of 0.1, and kernel type of RBF. Log flux was used as the dependent variable. Independent variables were selected based on the results of stepwise regression analysis, regression trees, and boosted trees. Combinations of some of the experimental parameters with the descriptors of the above models were also examined and their effect on the model and the errors associated with them were considered for the selection of the best SVM models.

8.2. Results and Discussion

8.2.1. Regression models

Stepwise regression analyses using the training set data resulted in model (8.1), and model (8.2) was obtained when molecular weight was manually incorporated in the analysis. The parameters of these models have been described in Table 8.2.

Log flux = 3.65 + 0.000001 [donor] - 0.522 InfiniteFinite - 1.41 dens - 0.0359 b_count - 0.00741 SlogP_VSA5 - 1.59 fiAB + 0.00188 BP-MP(mix) (8.1)

S = 0.939, r^2 = 0.567, N = 341, F = 62.4, P = 0.000

Log flux = 2.10 - 0.00413 MW + 0.000001 [donor] - 0.421 InfiniteFinite - 0.00911 SlogP_VSA5 - 1.49 glob + 0.00196 BP-MP(mix) - 1.81 fiAB - 0.00535 SlogP_VSA8 (8.2)

S = 0.932, r^2 = 0.57.5, N = 341, F = 56.1, P=0.000

According to model (8.1), a donor mixture with high permeant concentration, high difference between boiling and melting points (BP-MP(mix)), and applied under infinite conditions will lead to increased flux values. In addition properties of the permeants for a high flux values are small molecular size as evidenced by the negative coefficient of b_count, low molecular density (which occurs in molecules

with low number of heteroatoms such as N and O), low surface area of lipophilic groups (SlogP_VSA5), and low fraction of molecules ionised as both anion and cation at pH 7.4 (fiAB).

Model (8.2) employs the same parameters for the donor mixture as model (8.1); these are [donor], InfiniteFinite, and BP-MP(mix). The molecular descriptors are also similar, with molecular weight possibly replacing b_count in equation (8.1), and addition of molecular globularity and highly lipophilic surface area (SlogP_VSA8) instead of molecular density. The descriptor glob indicates that linear or planar molecules should have a higher flux values than spherical molecules.

Table 8.2. Brief description of the parameters of regression and RT models

Descriptors	**Description**
AM1_dipole	The dipole moment calculated using the AM1 Hamiltonian
AM1_E	The total energy (kcal/mol) calculated using the AM1 Hamiltonian
AM1_HF	The heat of formation (kcal/mol) calculated using the AM1 Hamiltonian
[donor]	donor concentration (μg/ml)
b_count	Number of bonds (including implicit hydrogens)
b_rotN	Number of rotatable bonds. A bond is rotatable if it has order 1, is not in a ring, and has at least two heavy neighbours
b_rotR	Fractions of rotatable bonds, it is calculated by dividing the number of rotatable bonds with the number of bonds between heavy atoms.
BP	Boiling point of the mixture
Bp-Mp(mix)	difference between the boiling and melting points of the mixture (donor phase)
chi0_C	Carbon connectivity index (order 0). This is calculated as the sum of 1/sqrt(d_i) over all carbon atoms i with $d_i > 0$
chi1v	Bond valence connectivity index (order 1) from (Hall and Kier, 1991) and (Hall and Kier, 1977). This is calculated as the sum of 1/sqrt($v_i v_j$) over all bonds between heavy atoms i and j where i<j
dens	Mass density (molecular weight divided by molecular volume)
fiAB	fraction of molecules ionised as anion and cation at pH 7.4
GCUT_SLOGP_2 GCUT_SLOGP_3	The GCUT descriptors using atomic contribution to log P instead of partial charge

Descriptors	Description
GCUT_SMR_0	The GCUT descriptor using atomic contribution to molar refractivity instead of partial charge
glob	Globularity, value of 1 indicates a perfect sphere while a value of 0 indicates a two or one dimensional object.
pre-Hydration	Indicator variable for pre-hydration of the skin prior to the *in vitro* test
InfiniteFinite	Indicator variable indicating infinite or finite exposures taking a value of 2 for finite and 1 for infinite dosing
Kier2	Second kappa shape index: $(n\text{-}1)^2/m^2$
KierA1	First alpha modified shape index, also correlated with molecular size (Hall and Kier, 1991)
LogD(5.5)	Apparent partition coefficient (distribution coefficient) at pH 5.5
LogS	Log of the aqueous solubility (mol/L). This property is calculated from an atom contribution linear atom type model (Hou 2004) with $r^2 = 0.90$, to 1,200 molecules
Thickness	Skin thickness in millimetres (mm)
Occlusion	Indicator variable for occlusion of the skin during *in vitro* test
PC+	Total positive partial charge
PEOE_VSA-4	Sum of van der waals surface area where atomic charge is in the range (-0.25, -0.2)
PEOE_VSA_HYD	Total hydrophobic van der Waals surface area. This is the sum of the van der waals surface area such that absolute value of atomic charge is less than or equal to 0.2
petitjean	Value of (diameter – radius) / diameter
PSA	Polar surface area
Q_VSA_POL	Total polar van der Waals surface area. This is the sum of the van der waals surface area such that absolute value of atomic charge is greater than 0.2
SlogP_VSA5	sum of van der Waals surface area of atoms with log P contributions in the range of (0.15-0.2)
SlogP_VSA8	sum of van der Waals surface area of atoms with log P contributions in the range of (0.3-0.4)
SMR_VSA6	Sum of the van der Waals surface area of atoms with atomic contribution to molar refractivity in the range (0.485, 0.56)
Wetting scale	Level of skin sample hydration
VDistMa	If m is the sum of the distance matrix entries then VDistMa is defined to be the sum of $\log_2 m - D_{ij} \log_2 D_{ij} / m$ over all i and j
vsa_hyd	Approximation to the sum of van der waals surface areas of hydrophobic atoms ($Å^2$)
vsurf_CW1	Capacity factor representing the ratio of the hydrophilic surface over the total molecular surface. These are calculated at eight

Descriptors	Description
	different energy levels (from -0.2 to -6.0 kcal/mol) (Cruciani et al., 2000)
vsurf_D1 vsurf_D3 vsurf_D4 vsurf_D5	Volume that can generate hydrophobic interactions. VolSurf computes hydrophobic descriptors at eight different energy levels (from -0.2 to -1.6 kcal/mol) (Cruciani et al., 2000)
vsurf_DW23	Contact distances of lowest hydrophilic energy descriptors
vsurf_HB5	H-bond donor capacity, representing the molecular envelope which can generate attractive H-donor interactions with carbonyl oxygen probe. The descriptors are computed at six different energy levels (from -1 to -6 kcal/mol) (Cruciani et al., 2000)
MW	Molecular weight (including implicit hydrogens) in atomic mass units

Table 8.3 gives the predictive abilities of the regression models above and compares them with the internal validation of equation (7.5) where the descriptors were selected using all the data points (as opposed to only the training set used for equations (8.1) and (8.2)). According to this Table, equation (8.1) is the most accurate predictor of the log flux values of external validation set.

Table 8.3. The statistical parameters of regression equations

Model	Validation	Training set		Validation set	
		MAE	n	MAE	n
Eq (7.5)	Internal	0.761	454	0.689	112
Eq (8.1)	External	0.736	341	0.667	112
Eq (8.2)	External	0.718	341	0.688	112

8.2.2. Regression Trees (RT) analyses

Several regression trees were generated using a combination of descriptors. The selected trees have been described in Table 8.4. RT (8.1) was obtained with the stopping criteria of: prune on variance, minimum number for splitting 40, and minimum number in child node 11. RT (8.2) was obtained when the stopping criteria were relaxed to allow expansion of the tree: prune on variance, minimum number for splitting 10, and minimum number in child node 11. RT (8.3)-(8.5) were obtained using the descriptors of equations (7.5), (8.1) and (8.2). Using RT analysis, the

experimental parameters such as the indicator variables for finite or infinite dosing or the wetting or hydration scales were not selected as a splitting variable. Hereafter, interactive trees were generated with the manual selection of these parameters. In this way, RT models (8.6)-(8.13) are the regression trees obtained when an experimental parameter was manually selected as the first partitioning descriptor. It must be noted that the hydration scale was not significant. Table 8.4 gives the statistical parameters of the models.

Table 8.4. Description of regression trees, risk estimates and standard error for both training and test set; RT is regression trees and SWR is stepwise regression analysis

Model	Parameters	Manually selected experimental parameter	Variable selection	Risk estimate		Standard error	
				Train	Test	train	test
RT (8.1)	All	None	RT	0.471	0.726	0.042	0.118
RT (8.2)	All	None	RT	0.429	0.719	0.040	0.117
RT (8.3)	Eq (7.5)	None	SWR	0.714	0.877	0.061	0.139
RT (8.4)	Eq (8.1)	None	SWR	0.733	0.941	0.061	0.138
RT (8.5)	Eq (8.2)	None	SWR	0.721	0.891	0.058	0.130
RT (8.6)	All	FiniteInfinite	RT	0.725	0.955	0.060	0.157
RT (8.7)	All	Thickness	RT	0.508	0.831	0.042	0.123
RT (8.8)	All	Pre-hydration	RT	0.720	0.887	0.056	0.127
RT (8.9)	All	Occlusion	RT	0.540	0.720	0.051	0.116
RT (8.10)	All	Wetting scale (split at 3.5)	RT	0.679	0.851	0.058	0.120
RT (8.11)	All	Wetting scale (split at 2.5)	RT	0.622	0.868	0.051	0.141
RT (8.12)	All	Wetting scale (split at 1.5)	RT	0.667	0.960	0.057	0.156
RT (8.13)	All	Wetting scale (split at 0.5)	RT	0.582	0.832	0.053	0.154

For regression-type problems with a continuous dependent variable, risk is calculated as the residual variance (Breinman et al., 1984). Standard error is the standard deviation of a mean and is calculated by the square root of the (sample variance)^2 when divided by the sample mean.

Table 8.5 gives a summary of the models and the Mean Absolute Error (MAE) for the training and validation sets using these models. According to Table 8.5, all RT models have lower MAE values for the training set than for the test set. RT models with the lowest MAE for the validation set are RT (8.9), (8.1), (8.2), followed by RT (8.13). RT models (8.2), (8.9) and (8.13) have been presented in Figures 8.1-8.3, respectively. RT (8.1) is a shorter version of RT (8.2).

Table 8.5. Statistical parameters of the models; RT is regression trees and SWR is stepwise regression analysis

Model	Parameters	Manually selected experimental parameter	Variable selection	Training set		validation set	
				MAE	n	MAE	n
RT (8.1)	All	None	RT	0.533	343	0.659	112
RT (8.2)	All	None	RT	0.500	343	0.660	112
RT (8.3)	Eq (7.5)	None	SWR	0.664	342	0.730	112
RT (8.4)	Eq (8.1)	None	SWR	0.668	341	0.765	112
RT (8.5)	Eq (8.2)	None	SWR	0.664	341	0.735	112
RT (8.6)	All	Finite/Infinite	RT	0.654	357	0.726	119
RT (8.7)	All	Membrane thickness	RT	0.567	343	0.715	111
RT (8.8)	All	Pre-hydration	RT	0.666	350	0.742	117
RT (8.9)	All	Occlusion	RT	0.551	336	0.624	108
RT (8.10)	All	Wetting scale (split at 3.5)	RT	0.637	341	0.732	111
RT (8.11)	All	Wetting scale (split at 2.5)	RT	0.618	341	0.697	111
RT (8.12)	All	Wetting scale (split at 1.5)	RT	0.631	341	0.759	111
RT (8.13)	All	Wetting scale (split at 0.5)	RT	0.581	341	0.686	111

RT (8.2) is the second most accurate RT for the estimation of the flux of the external validation set. According to RT (8.2) presented in Figure (8.1), flux is much higher with higher donor concentration ([donor]) and low value of Kier2 (indicating small molecular size) especially if GCUT_SLOGP_3 is small. Otherwise if the concentration is low then a combination of low GCUT_SLOGP_3, high vehicle boiling point (BP), low polarity (PSA, which can also indicate molecular size), high Kier2 and small lipophilicity (vsurf_DW23) are needed for a high Flux.

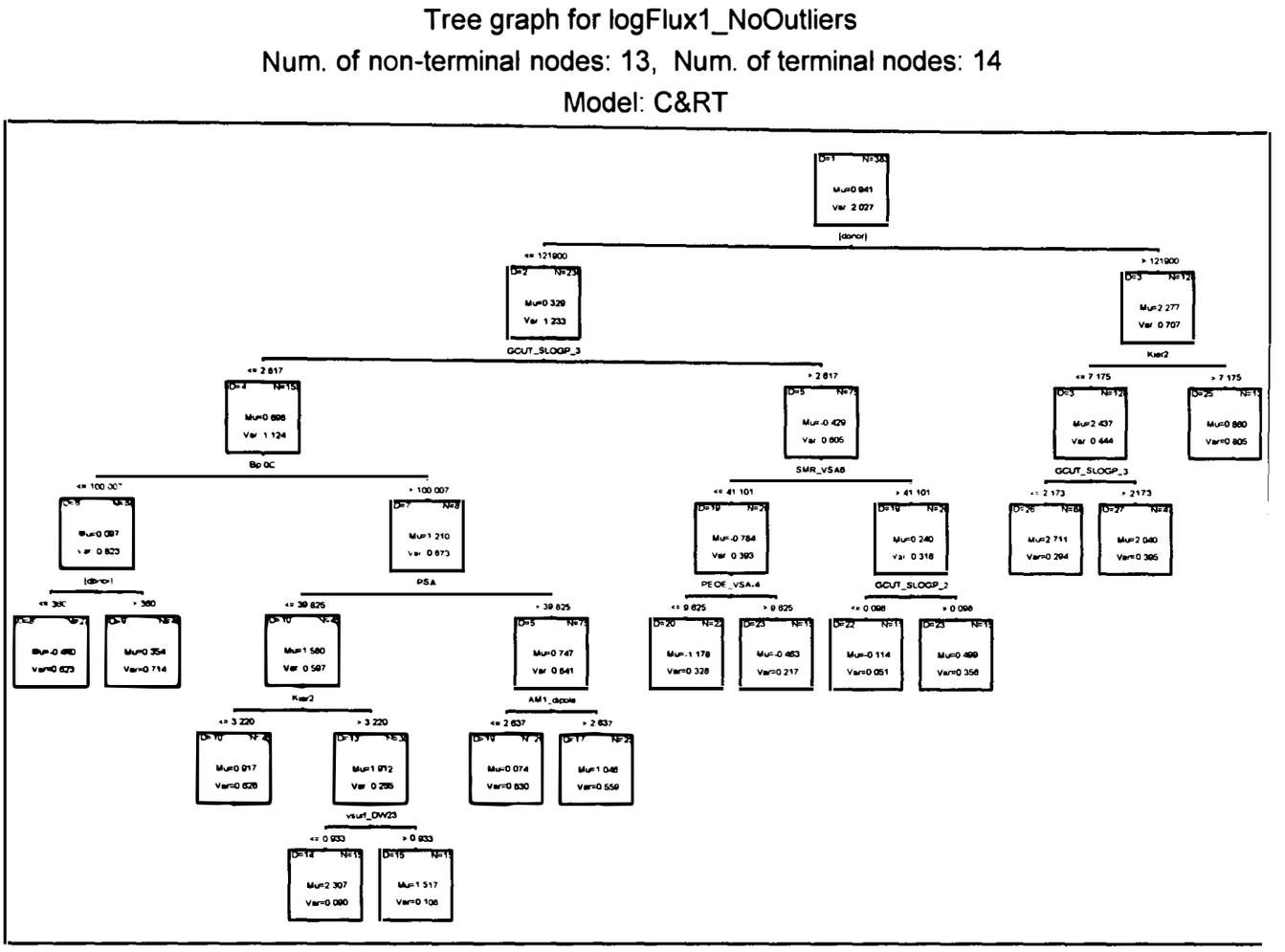


Figure 8.1. RT model (8.2) using all parameters with the stopping criteria set at minimum number of splitting 10 and minimum number in child node: 11.

RT(8.9) is the lowest error regression tree for the estimation of flux. According to this model (presented in Figure 8.2), flux is much higher when the skin is occluded and combined with a high donor concentration and high ratio of hydrophilic surface to the total surface area (vsurf_CW1) it leads to highest average flux. Otherwise in case of low donor concentration a combination of low hydrogen bonding donor ability (vsurf_HB5), high flexibility (b_rotN) and low polar surface area (Q_VSA_POL) is required for high flux through occluded skin. If un-occluded then for permeants with low hydrophobic volume (vsurf_D5) high flux values are still possible if the skin is pre-hydrated. Pre-hydration of the skin is most common with infinite exposures where the donor phase is appliedin high volumes, hence gradual hydration of the skin during the course of the occlusion state.

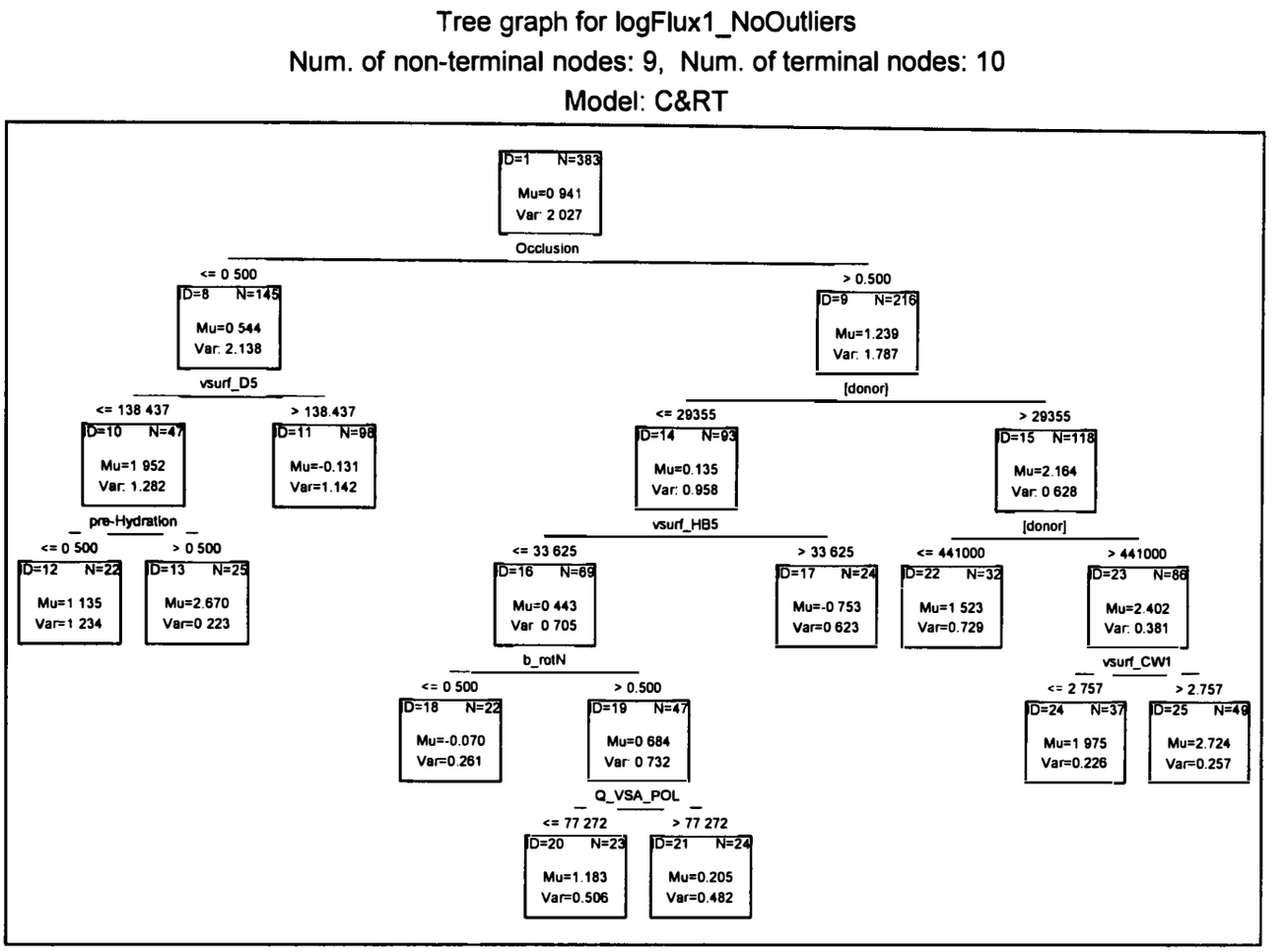


Figure 8.2. RT model (8.9) using indicator variable for occlusion as the first parameter for splitting the data

Figure 8.3 shows the third best RT for the estimation of flux (RT 8.13). RT (8.13) shows that the wetting scale is also a good descriptor which aids the estimation of flux values. From amongst several cut-off points for this scale, the best MAE was achieved by the interactive tree which used the cut-off value of 0.5 for splitting the data. Since only applied volumes of less than 100 score 0 on the scale, splitting the data at wetting scale of 0.5 could be an indirect method for splitting finite from infinite dosing. The scale, however, also takes into account pre-hydration and occlusion, since pre-hydration and occlusion would give a higher score on the scale regardless of the applied volume. Other parameters of RT (8.13) indicate that high donor concentration and small molecular size (Chi0_C) would result in high flux values when the wetting scale is 1, 2, 3, or 4. When the donor concentration is low and the wetting level of the skin is 1, 2, 3, or 4 in the scale, a relatively high flux can still be achieved for small size (KierA1) flexible molecules (b_rotR) (see ID 71 in figure 8.13). In case of low wetting scale of 0, there are 41 cases of low total positive

charge (which can indicate low hydrogen bonding donor ability according to Dearden and Ghafourian (1999)) which can have high flux values (according to descriptor PC+).

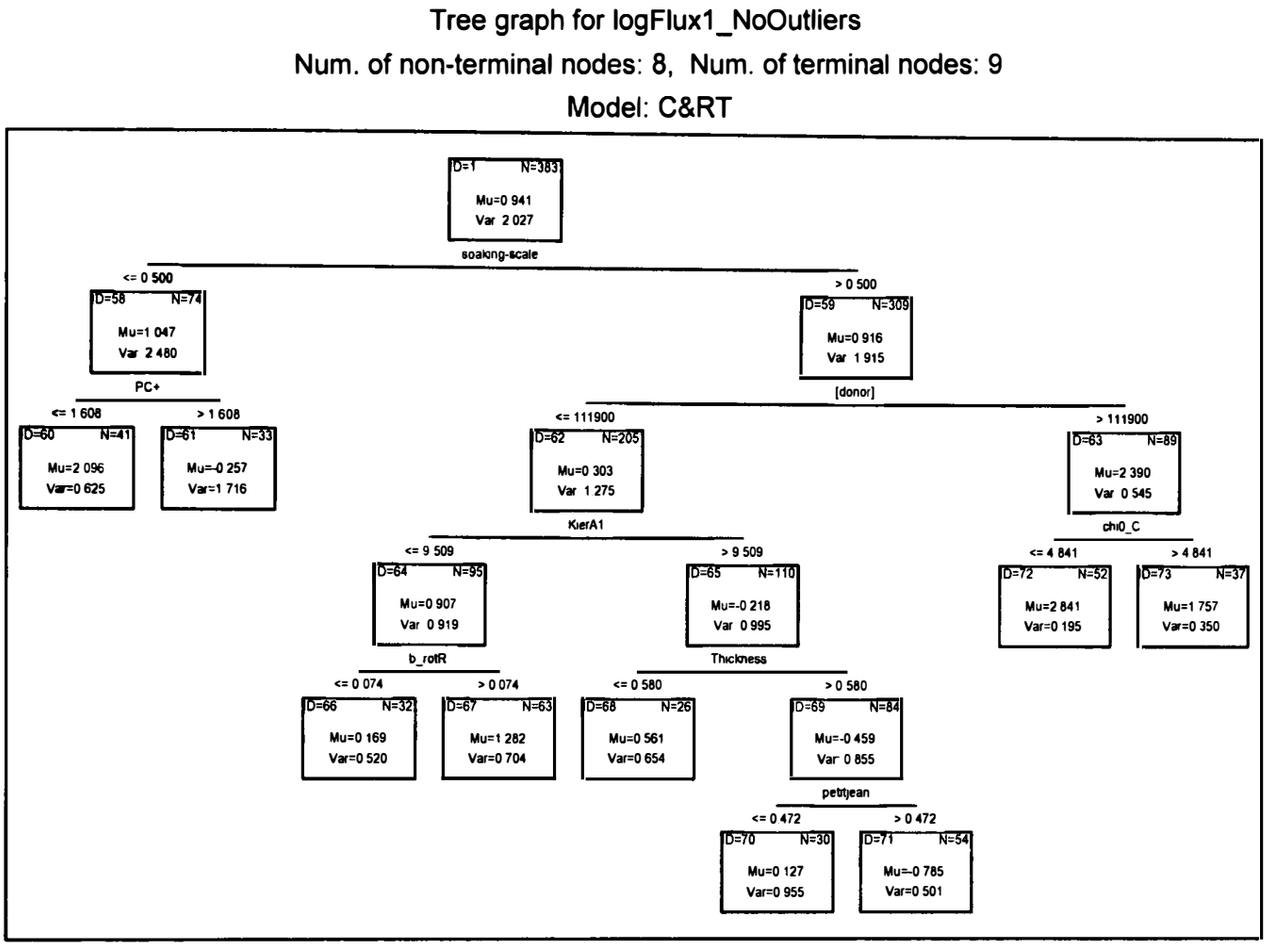


Figure 8.3. RT model (8.13) using wetting scale as the first parameter for splitting the data at the value of 0.5.

8.2.3. Boosted trees

The boosted tree module allows multiple tree models to be generated for prediction. It computes a sequence of simple trees, where each successive tree is built for the prediction of the residuals of the preceding tree. For each predictor variable in the database the sum of squares regression is calculated and the best predictor variable is used for the split (Friedman, 1999a, b).

In boosted trees analysis using STATISTICA, many trees were generated each containing only 1 split (stopping criterion was maximum nodes of 3). To avoid over

training in this analysis, the number of additive terms was selected at the point where the error for the training and test set was similar.

The statistical properties and parameters of the selected boosted trees have been presented in Table 8.6. Also presented in this table are accuracy estimates of the models. The lowest MAE for the test set was calculated for model (8.4) using 90 trees, a learning rate of 0.05 and a subsample proportion of 0.5. Further reduction of the learning rate and increasing the number of additive terms did not significantly reduce error, but did increase computing time. The second lowest estimation error was for model (8.6). In general there was little difference in the error of the selected boosted trees.

Table 8.6. Parameters used for boosted tree analyses and the resulting risk estimates, standard error, and MAE of log flux estimation for training and test sets

Model	Learning rate	Number of additive terms	Subsample proportion	Risk Estimate		Standard Error		MAE	
				train	test	train	test	Train	Test
8.1	0.1	40	0.5	0.671	0.676	0.050	0.104	0.643	0.638
8.2	0.005	800	0.5	0.662	0.660	0.049	0.102	0.637	0.627
8.3	0.005	800	0.4	0.656	0.658	0.048	0.102	0.638	0.627
8.4	0.05	90	0.5	0.655	0.635	0.049	0.098	0.632	0.612
8.5	0.05	90	0.45	0.650	0.656	0.047	0.100	0.635	0.628
8.6	0.005	800	0.55	0.656	0.656	0.048	0.101	0.635	0.624

Boosted trees analysis may employ many predictor variables in the trees. It can also provide a variable importance table where the variables are sorted according to the importance in the boosted tree model. Variable importance (predictor importance) is calculated as the relative (scaled) average value of the predictor statistic over all trees and nodes; hence these values reflect on the strength of the relationship between the predictors and the dependent variable of interest, over the successive boosting steps (STATISTICA Help file, StatSoft Inc.).

The top 10 highly important predictors in the best and the second best boosted tree models are shown in Table 8.7. Most of these predictor variables were different from those selected by other methods (stepwise regression and regression trees analyses). For example the donor concentration, which was consistently selected by RT and stepwise regression methods, is not amongst the top 10 most important descriptors of the boosted trees models. The variables that were selected as the most important were all (indirectly) related to molecular size and/or hydrophobicity. On the other hand, the important descriptors of the two boosted trees models were similar, but not the same, as can be seen in the table.

Tabel 8.7. The most important variables of the best and second best boosted trees and their order of importance.

predictor	Order of importance	
	Model (8.4)	Model (8.6)
AM1_HF	1	1
GCUT_SMR_0	2	2
KierA1	3	7
vsa_hyd	4	12
LogS	5	3
AM1_E	6	17
chi1v	7	14
PEOE_VSA_HYD	8	6
GCUT_SLOGP_3	9	20
VDistMa	10	27
vsurf_D5	26	4
vsurf_D4	16	5
vsurf_D3	29	10
vsurf_D1	30	8
LogD(5.5)	50	9

Table 8.8 shows the order of importance for some selected parameters. Out of all the descriptors of experimental conditions, membrane thickness has the highest importance according to this table.

Table 8.8. Importance of experimental conditions and some selected mixture properties in the selected boosted trees

Predictor	Order of importance	
	Model (8.4)	Model (8.6)
Hydration	289	292
Occlusion	293	295
Wetting scale	177	165
Membrane thickness	67	47
[Donor]	141	157
Bp(mix)	148	148
Mp(mix)	171	160
Finite/infinite dosing	251	212

Support Vector Machines

Various methods were used to select variables to feed into ε-SVM as the predictor variables. Table 8.9 describes the different combinations of variables that were used for the analysis. The table also shows the MAE values of log flux estimations for the test and training sets.

According to Table 8.9, the MAE for the SVM model developed using variables from equation (7.5) (SVM (8.5)) was 0.627 for the test set and 0.688 for the training set. Different descriptors of experimental conditions and log P were added to this set of variables to study their effect on the predictive capabilities of the models. Adding Finite/Infinite and occlusion reduced the MAE for the test set to 0.622 and 0.624 respectively. Addition of the pre-hydration indicator, the wetting scale and membrane thickness increased the error to 0.648, 0.635 and 0.640 respectively.

The parameters of equation (7.3) gave poorer prediction of the log flux for the test set, with MAE of 0.738. These variables included the donor concentration, the polar surface area divided by the total surface area (PSA/SA) and molecular volume. Addition of the variables of experimental conditions lowered the error. Addition of the difference between the boiling and melting points of the donor mixture and log P gave the lowest error with MAE of 0.686 for the test set. This value however is still much higher than the one achieved by the analysis using the variables of equation (7.5).

Although boosted trees had very low MAE, the incorporation of the top 10 variables of the selected boosted tree models listed in Table 8.7 in SVM analysis gave a model with a very poor prediction of log flux (SVM (8.11)). Boosted trees employ many variables and not just the 10 highly ranked in the importance scale. It seems that for the variable set to show an acceptable performance in SVM, all these variables may need to be incorporated. On the other hand, the number of descriptors that can be incorporated in SVM is restricted.

The descriptors of selected RTs did not perform well in SVMs. MAE values of log flux estimation for test set using RT (7.1) and RT (7.4) are 0.747 and 0.607 respectively (Table 7.3). The latter is an interactive tree with pre-hydration manually selected as the first partition. The estimation error of the SVM developed using the variables of RT (7.1) were very similar to the original tree (MAE of 0.746). The SVM using descriptors of RT (7.4) was much higher than the original at 0.693.

In conclusion, Table 8.9 shows that the SVM developed using the parameters of equation (7.5) and Finite/Infinite dosing variable leads to the lowest MAE of estimation for the test set (MAE of 0.622). Parameters of equation (7.5) are donor concentration, molecular weight, the difference between the boiling and melting points of the donor mixture, and four molecular descriptors of the permeants, vsurf_G, SlogPVSA4, fiAB and VadjMa. The latter four variables are chemical parameters related to molecular size and hydrophobicity.

Table 8.9. MAE of log flux estimation for the test and training sets using ε-SVM regression models

SVM Model	Parameters	MAE	
		Train	**Test**
8.1	Top 10 important variables of boosted trees model (8.4)	0.804	0.885
8.2	RT (8.13)	0.679	0.723
8.3	RT (7.1)	0.741	0.746
8.4	RT (7.4)	0.683	0.693
8.5	eq (7.5)	0.688	0.627
8.6	**eq (7.5) + Infinite/Finite**	**0.657**	**0.622**
8.7	eq **(7.5)** + wetting scale	0.664	0.635
8.8	eq **(7.5)** + Occlusion	0.687	0.624
8.9	eq **(7.5)** + Hydration	0.676	0.648
8.10	eq **(7.5)** + Membrane	0.678	0.640
8.11	eq **(7.3)**	0.744	0.738
8.12	eq **(7.3)** + BP-MP(mix)	0.722	0.697
8.13	eq **(7.3)** + log P	0.740	0.734
8.14	eq **(7.3)** + Bp-Mp(mix) + Infinite/Finite	0.708	0.706
8.15	eq **(7.3)** +Bp-Mp(mix) + log P	0.721	0.686

The selected models

A diverse selection of models with high predictive power has been summarized in Table 8.10.

Table 8.10. The selected models, their errors, N, number of outliers and parameters

Model	Descriptors	Train		Test		No. Outliers		Description
		MAE	N	MAE	N	MAE > 1.5	MAE > 2.0	
Eq 7.5	All	0.761	342	0.689	112	50	19	Stepwise regression on all data
Eq 8.1	All	0.736	341	0.667	112	48	15	Stepwise regression on training set
Eq 8.2	All	0.718	341	0.688	112	48	15	Stepwise regression on training set
RT 8.1	All	0.533	343	0.659	112	20	9	General interactive tree
RT 8.9	All	0.551	336	0.624	108	31	11	First partition based on occlusion
RT 8.13	All	0.581	341	0.686	109	28	11	Scale split at 0.5 as first partition
Boosted trees 8.4	All	0.632	383	0.612	111	34	10	Learning rate 0.05, 90 trees, subsampling rate 0.5
Boosted trees 8.6	All	0.635	383	0.624	111	34	11	Learning rate 0.005, 800 trees, subsampling rate 0.55
SVM 8.5	From RT 7.4	0.683	343	0.693	112	50	12	number of support vectors 231 (207 bounded) Gamma = 0.125
SVM 8.6	From Eq 7.5 + finiteInfinte	0.657	341	0.622	112	39	16	number of support vectors 230 (201 bounded) Gamma = 0.125
SVM 8.12	From Eq 7.3 + Bp-Mp (mix)	0.722	343	0.697	112	40	13	number of support vectors 253 (235 bounded) Gamma = 0.250

The selection of the models above was based mostly on the MAE of log flux estimation for the external validation set. The table also shows the number of data points with high estimation errors using these models (data with Absolute Error (AE) above 1.5 and 2.0). The lowest mean absolute error in the table is achieved by boosted trees (8.4) with MAE of 0.612. In addition, the boosted trees models were able to provide estimations for a higher number of (test and training set) data. Inclusion of 383 training data in these models compared to 336-343 training data used in the remaining models may have resulted in a better model training and hence lower prediction error. Figure 8.4 shows the graph of log flux against the absolute error of log flux estimation using boosted trees (8.4). The figure shows that most compounds have been predicted with AE<1. It can be observed that test and training sets show similar levels of error (no over-fitting).

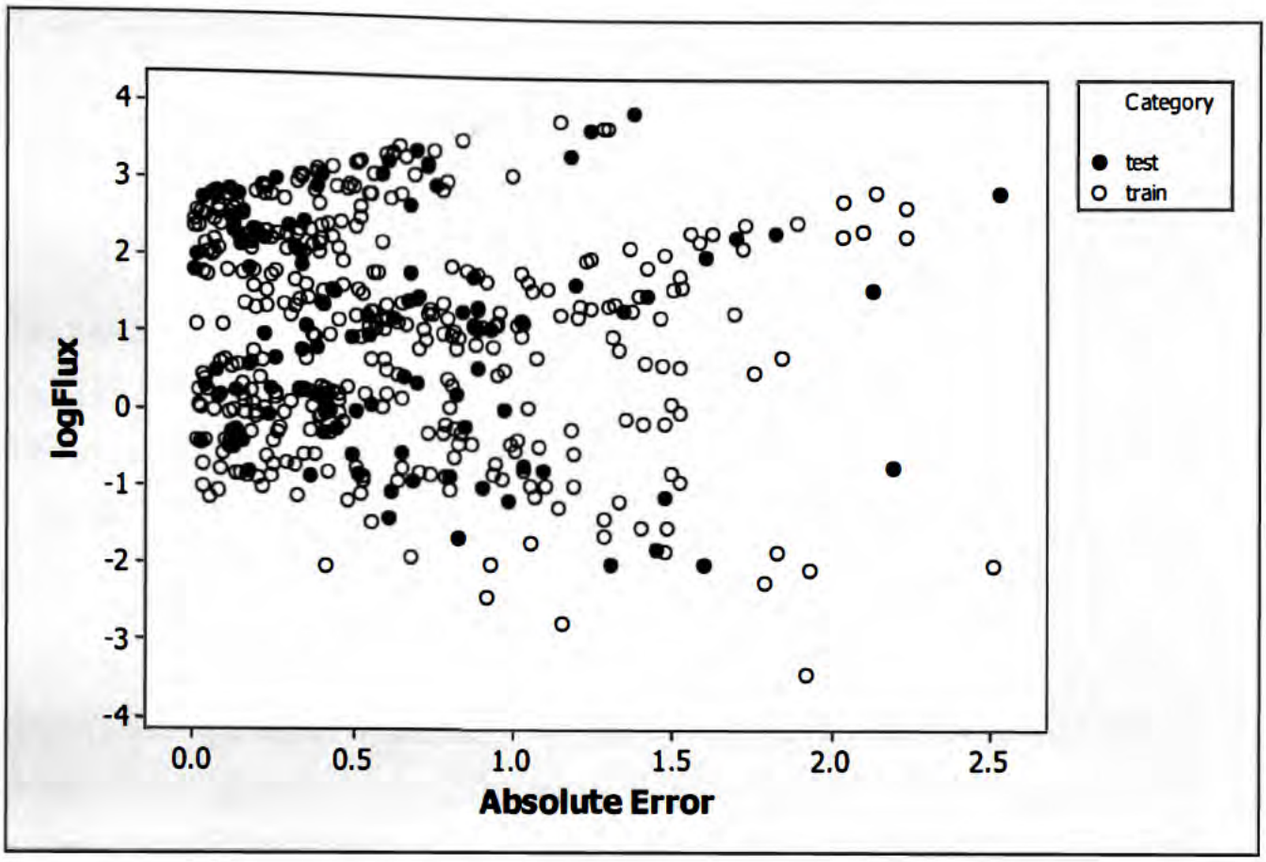


Figure 8.4. The plot of log flux against absolute error of estimation by boosted tree (8.4)

The interactive trees analysis has a lower number of high error data (outliers) compared to the other methods, but the MAE for the validation set is relatively high. It can also be observed that the training set MAE values are very low. This may be explained by the fact that more data have been excluded from the interactive trees due to the model parameters not being available (e.g. the number of data points in training and test sets). RT (8.9) shows one of the lowest validation set errors. In this tree the first splitting parameter is the indicator variable for occlusion (manually selected in the interactive RT).

The support vector machine with the variables from equation (7.5) plus the indicator of finite or infinite dosing (SVM (8.6)) has the second lowest MAE for the test set (0.622). It seems that SVM analysis gives a more accurate estimation of log flux than regression analysis using the same descriptors (comparison of SVM (8.1) with equation (7.5)). Support vector machines (SVM) are a non-linear algorithm developed for regression and classification problems (Vapnik, 1998). This method performs regression and classification tasks by constructing nonlinear decision boundaries. Because of the nature of the feature space in which these boundaries are

found, Support Vector Machines can exhibit a large degree of flexibility in handling classification and regression tasks of varied complexities. However, SVMs using the variables from RTs are not as accurate.

The outliers

It can be seen in Table 8.10 that log flux estimations by the models for a number of data points can deviate by more than 1.5 (absolute error) from the observed values. Table 8.11 shows the outliers (AE>1.5) in 8 or more of the 11 selected models, from test or training tests.

Interestingly most outliers (except for zidovudine and terbinafine) are those flux values that have been measured after very low applied dose (volume applied is below 30 μl). It must be noted that many entries with low donor volumes are not outliers; for examples refer to the graph between absolute error of estimation by boosted tree (8.4) and the donor volume at Figure 8.5. The error in the prediction of log flux at these extremely low exposure quantities could be a result of limitations of the models or limitations in the dermal absorption measurements. When the quantities of applied chemicals are small the experimental error may also be more significant, resulting in a larger variation in the observed flux values. In cases where the database contains other entries of the same compounds which are not outliers, it is unlikely that the large errors are a result of the chemical structure of these compounds or the way these compounds interact with human skin.

Tabel 8.11. Data with log flux estimation absolute error of >1.5 in 8 or more of the 11 selected models

Chemical	Log flux	Number of models	Reference
Propoxur	-2.036	10	van de Sandt et al. (1993)
Benzoic acid	-0.789	9	Patil et al. (1996)
Salicylic acid (average)	-0.921	8	Ademola et al. (1993)
Zidovudine (AZT)	2.701	10	Narishetty et al. (2005)
Zidovudine (AZT)	2.890	10	Narishetty et al. (2005)
Terbinafine	-2.000	8	Schmook et al. (2001)
Methyl-Parathion	-1.979	9	Sartorelli et al. (1997)

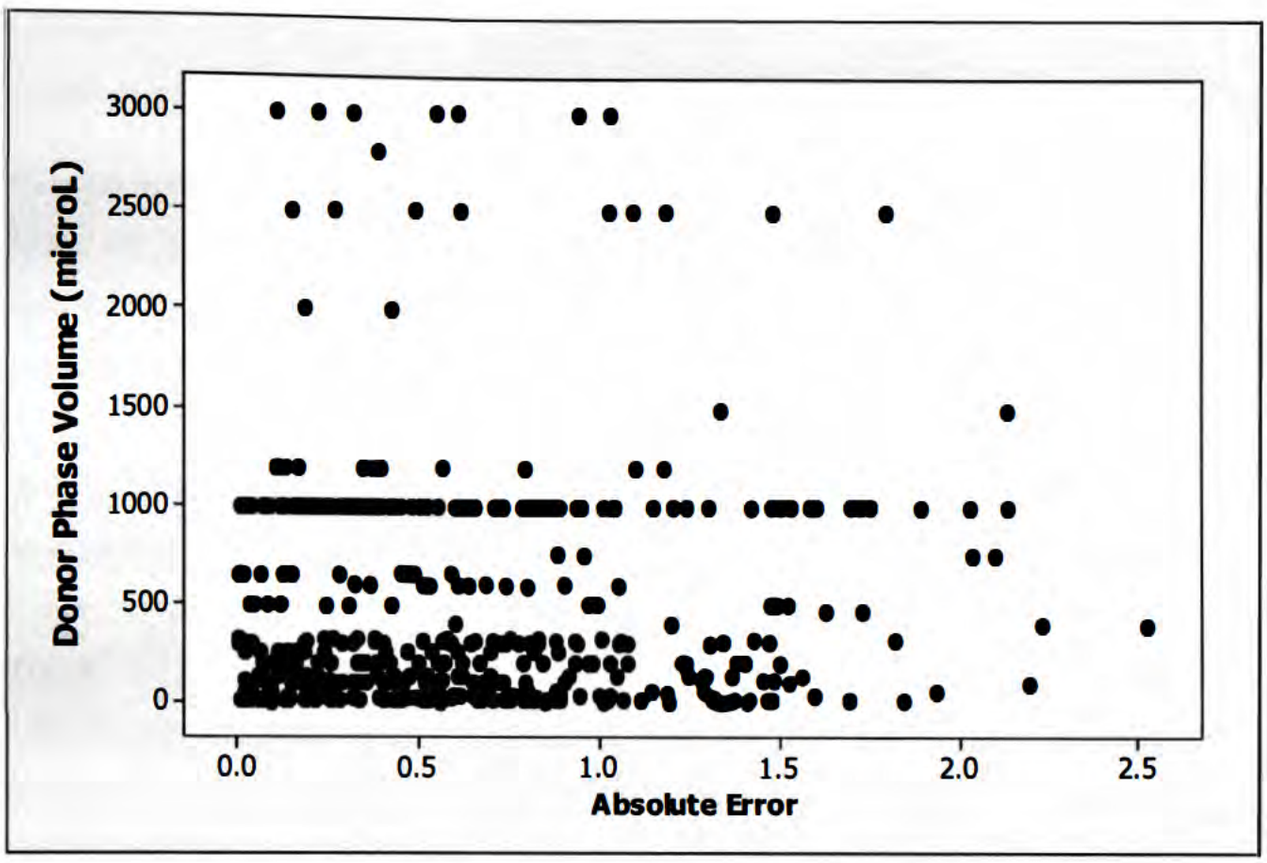


Figure 8.5. The plot of donor phase volume against absolute error of estimation by boosted tree (8.4)

The database contained 9 entries for propoxur, with only one entry having high prediction errors in most models. The average log flux of the remaining 8 entries was 0.291, varying between -1.390 and 1.199, compared with -2.035 for the outlier. The large absolute error of the outlying value of propoxur may be explained by the very low donor concentration compared to the other entries (320 compared to 1280-2750 μg/ml), resulting in the much lower observed flux which is consistently overestimated by the models.

Benzoic acid flux values were found 16 times in the database. One value of benzoic acid was an outlier in 9 out of 11 models. This flux value also was obtained following a much lower donor concentration compared to the other entries (122 compared to 2600-34010 μg/ml). The average log flux for the remaining benzoic acid entries (after removing this outlier) was 2.329.

20 entries of salicylic acid were found which contained log flux values. One entry had an AE greater than 1.5 in 8 out of 11 models. The donor concentration used was 90 compared to 500-368200 μg/ml for the other entries. The average log flux of the

other salicylic acid entries was much higher at 1.864 in comparison with the outlier value listed in Table 8.11.

Four log flux entries were found in the database for zidovudine. The average log flux of the two values which were not outliers was 1.368, ranging from 1.068 to1.668. Two entries were outliers with much higher observed log flux values than the predicted values. This may be due to the presence of permeation enhancers, namely l-menthol and 1.8% cineole. Both these compounds are terpenes which are known penetration enhancers (Williams and Barry, 2004) and were present at 5% (w/v) concentration in the vehicle (66% ethanol in water). In the other entries for zidovudine same vehicle was used without the penetration enhancers, explaining their lower log flux values. Importantly this observation indicates that the effect of these penetration enhancers was not accurately predicted by the models, even though the models used various parameters for the vehicles (such as the difference between boiling and melting points of the vehicle). This may be due to the specific mode of action of these penetration enhancers such as specific interactions with stratum corneum lipids (Cornwell et al., 1996) or proteins (Barry, 1991), which is not fully understood. It must be noted that are other entries with penetration enhancers in the donor phase that are not outliers. Other mixtures containing various terpene penetration enhancers show AE values of 0.655-1.15 using boosted tree (8.4).

Terbinafine is an outlier in 8 out of 11 models. The original paper by Schmoock et al. (2001) describes that the flux value of <0.01 $\mu g/cm^2/h$ was below the detection limit of the used analytical method. However the predicted value of terbinafine in the models is higher than the observed value. The log P value of terbinafine is 5.58. It seems that the log flux of chemicals with high log P values (above 5) is less well predicted by the models. The average percentage of outliers for chemicals with log P above 5 is 10.7% while the average in the rest of the database is 5.7%. The discrepancy between observed and predicted values could also be due to the epidermal metabolism of terbinafine. Although the skin biotransformation of terbinafine has not been investigated (Web of Science records, September 2011), this drug is extensively distributed in human and the systemic clearance of orally administered terbinafine is primarily dependent on biotransformation (Humbert et al., 1995) with at least seven CYP enzymes being involved in terbinafine, most importantly CYP2C9, CYP1A2, and CYP3A4 (Vickers et al., 1999).

Immunohistochemical studies and catalytic activity measurements have indicated cutaneous cytochrome P450 (CYPs), as the most important phase I enzymes that are mainly localised in the epidermis hair follicles, and sebaceous glands (Baron et al., 2008; Zhang et al., 2009).

Estimated absorption for methyl-parathion is much higher than the actual reported value by Sartorelli et al. (1997). The log k_p value corresponding to this flux is in accordance with the average log k_p value reported by Riviere and Brooks (2010). In both investigations the volume applied to the skin is quite low at 30 or 10 µl and the skin is not occluded. The low absorption could result from solvent evaporation (in this case acetone) and precipitation of the permeant at the top of the skin or at the upper layers of SC cells.

8.3. Conclusion

In this investigation, out of 514 flux entries, 131 entries were set aside as the external validation set from beginning of the analysis and the remaining 383 entries were used as the training set for the development of models. Many of the resulting models showed mean absolute error of prediction at an acceptable range 0.612-0.697. The best model in terms of the MAE for test set was the boosted tree model employing 90 trees at a learning rate of 0.05. This model also had one of the lowest number of outliers (absolute error of >2) in test and training sets. RT model (8.9) and SVM (8.6) were the other selected models based on MAE for the validation set. The RT model employed occlusion and pre-hydration indicator variable as well as donor concentration and several permeant properties. The SVM model employed the donor concentration, boiling and melting point gap of the vehicle, the indicator of finite or infinite dosing and several molecular descriptors of the permeants. The donor concentration was the most important descriptor of flux values in most models, selected as first choice in simple regression and RTs.

Non-linear models generally perform better than the linear regression models. The non-linear relationship is especially evident with the descriptors of experimental conditions, occlusion, pre-hydration and wetting scale. These descriptors play

important roles in the RT models, while being less significant in linear regression models (not selected by stepwise regression analysis). The importance of occlusion is emphasised through the RT model (8.9) which also agrees with the conclusion from chapter 7.

The outlier analysis identified seven entries with absolute error of more than 2 using 8 or more out of 11 selected models. The majority of the outliers were flux values measured with small donor volumes and non-occluded conditions implying that models may fail to take the solvent evaporation into account. It was also concluded that models perform less well for highly lipophilic compounds (log P > 5) and the applications involving specific penetration enhancers in the donor mixtures.

As a general conclusion of this study, the need for a widely accepted wetting scale to indicate the effect of skin hydration/wetting on the skin absorption is outlined. The scale should be derived from the experimental conditions, such as occlusion, pre-hydration, application volume, and the type of the vehicle.

9. Conclusion

This investigation aimed at modelling the effect of formulation factors and mixture ingredients on the skin penetration of compounds. A major obstacle with these studies is the availability of reliable and consistent data. *In vivo* skin permeation data are scarce and, in addition, they are affected by several other factors, apart from the skin absorption, including pharmacokinetics behaviour of the permeant, e.g. metabolism and stability, distribution and excretion. This has led to *in vitro* measures to be employed as acceptable skin absorption assessment criteria. Most modelling efforts have also been directed at the estimation of the *in vitro* measures of the skin permeation such as flux and permeability coefficient. In this project Quantitative Structure-Activity Relationship (QSAR) techniques were employed to model the *in vitro* measures of skin absorption. Chapters 3 and 4 use two very consistent datasets where skin absorption is measured under the same experimental conditions in the same laboratory. The dataset in Chapter 3 is the flux enhancement of a model drug, formoterol, by the use of solvents alternative to water from excised rat skin. The dataset in Chapter 4 consisted of permeation coefficient of the model drug haloperidol, under the effect of terpenes and terpenoids through excised human skin.

QSAR analyses in these two chapters were limited to the use of linear regression and Support Vector Machines, as they are more suitable for smaller datasets. The results of analysis showed that solvents with lipolar (surfactant like) structure, containing aliphatic rings and non-linear structures that are lipophilic in nature result in the highest flux values for formoterol (Chapter 3). On the other hand, terpene enhancers required different characteristics to perform as good skin penetration enhancers for haloperidol. These characteristics involve hydrogen bonding acceptor abilities, branched or cyclic structures, low dipolarity, and hydrogen bonding donor ability; while lipophilicity expressed by log P has no effect (Chapter 4). The differing structural requirements observed in these two studies can be attributed to the differing experimental set up; studies in Chapter 3 used different liquids including terpenes as the vehicle while studies in Chapter 4 used only 5% solutions of the liquids (mainly terpenes) in propylene glycol. The other important difference is the

chemical space of the two datasets. Compounds in Chapter 3 contain alcohols, esters, fatty acids and terpenes with a wide range of molecular weights from ethanol to dodecanol and isopropyl palmitate to toluamide. On the contrary, compounds in Chapter 4 are comprised of terpenes mostly. Also, the model drugs in these two studies are different, although they have similar chemical structure, molecular weight, and pk_a values. According to Karande et al. (2005), chemical enhancers that work through fluidisation of SC lipids are more potent if they are more lipophilic. The findings of Chapter 3 agree well with this definition of fluidisers. The results obtained for terpene enhancers in Chapter 4 are not consistent with the expectations from fluidisers, but with extractors according to Karande et al. definition.

The remaining chapters are dedicated to complex mixtures of permeants and several varying mixture components. The use of such datasets meant that concomitant effects of mixture components and chemical structures of the penetrants could be analysed. Chapters 5 and 6 employed the dataset of Riviere and Brooks (2010), where the effects of certain solvent mixtures and a surfactant on permeation of several compounds from excised porcine skin has been investigated. The limited number of penetrant/vehicle systems led to the conclusion that more experimental measurements of skin permeation rates were necessary to obtain an appropriately validated QSAR model. After analysing the chemical space, four permeant molecules were selected for *in vitro* permeation studies from the same formulation mixtures. The QSAR models using the expanded dataset showed good prediction abilities with improved mean absolute error of 0.45 for the internal validation sets. In these models, the effect of vehicles were modelled by the difference between their melting and boiling points, a property related to molecular symmetry, and the difference between melting points of the permeant and the vehicle. The selected permeant descriptors in these linear models showed the negative impact of permeant molecular size, dipolarity and presence of electron rich nucleophilic groups on the skin permeability coefficients.

Inter-laboratory and inter-individual variations are very common in the *in vitro* measures of skin permeation. This has been attributed to a number of experimental conditions including sin thickness, pre-hydration, and occlusion. The last dataset used for QSAR studies in Chapter 7 and 8 allowed for the effect of experimental

conditions to be investigated, as well as the effects of mixture components and permeant descriptors. From the EDETOX database, the *in vitro* skin flux data from excised human skin were extracted. This dataset was updated with more recent publications (2001-November 2010). Experimental conditions of the flux measurements were extracted from original publications. These included membrane thickness, type of exposure (Finite/Infinite), pre-hydration and occlusion of the skin, donor concentration, volume and dose applied. The dataset was large enough to allow more sophisticated data mining tools for the analysis. As the first step, the effect of experimental conditions were analysed on the *in vitro* skin flux by incorporating appropriate indicators of these conditions (Chapter 7). Linear regression and non-linear Regression Trees (RT) showed the substantial dependability of flux on donor concentration. The effects of the remaining experimental factors, although significant, were not as considerable. The results indicated a higher impact of the permeant concentration and molecular structure, and the vehicle properties on the skin flux in comparison to the experimental variables of skin thickness, finite or infinite dosing, occlusion or pre-hydration. The effect of occlusion was more pronounced than the other experimental conditions, but this was only evident in the non-linear analysis.

The importance of external validation of QSAR models is well recognised for the applicability of the models to real-life prediction scenarios. This latter dataset was large enough to allow for a set of entries to be excluded from the QSAR development and serve as the external validation set (Chapter 8). The training set of 383 data-points was used in several statistical methods. Some of the resulting models showed good prediction abilities for the external validation set. The best model, in terms of the MAE for the validation set and the number of outliers, was the boosted tree model employing 90 trees at a learning rate of 0.05. Regression Tree model (8.9) and SVM (8.6) were the other selected models. The RT model employed occlusion and pre-hydration indicator variable as well as donor concentration and several permeant properties. The SVM model employed the donor concentration, boiling and melting point gap of the vehicle, the indicator of finite or infinite dosing and several molecular descriptors of the permeants.

Non-linear models generally perform better than the linear regression models. The non-linear relationship is especially evident with the descriptors of experimental conditions, occlusion, pre-hydration and wetting scale. These descriptors play important roles in the RT models, while being less significant in linear regression models (not selected by stepwise regression analysis).

Finally, based on the current results it can be concluded that modelling can aid the understanding of the mechanisms involved in the skin penetration of compounds and with the appropriate data it can lead to validated models with acceptable prediction accuracy. The *in silico* prediction of skin penetration can aid drug formulators and regulatory risk assessment bodies. It is expected that the results of this work may benefit *in vivo* estimations using the *in vitro* flux estimates.

10. Future Work

Following the research carried out so far, QSAR modelling can be continued. Many investigators employ different animal skin types in *in vitro* assessments. Despite random comparisons available in the literature, a systematic comparison of the animal data with human skin and that with *in vivo* results is required. The EDETOX database has a wealth of *in vitro* and *in vivo* data that can be updated and analysed. Such investigations can aid validation of various *in vitro* and *in vivo* techniques. Undoubtedly, availability of large datasets, either *in vivo* or *in vitro*, along with cutting edge QSAR software and data-mining tools can provide better models for the prediction of relationships between chemical properties and skin permeation potential.

The dataset can be populated with the constituents of the receptor phase. It has been suggested that permeation may be limited from skin to the receptor phase for compounds with low aqueous solubility. Researchers often use varying pH levels, co-solvent mixtures and albumin in the receptor phase, in order to aid the sink conditions. Moreover, the constituents of the receptor phase may affect the skin itself. Therefore it is essential, especially with static diffusion cells, to study the effect of receptor phase ingredients.

11. References

Abdi, H., 2003. PLS-Regression; Multivariate analysis. In: M. Lewis-Beck, A. Bryman, & T. Futing (Eds): Encyclopedia for research methods for the social sciences. Thousand Oaks: Sage.

Abraham, M.H., Chadha, H.S., Martins, F., Mitchell, R.C., Bradbury, M.W., Gratton, J.A., 1999. Hydrogen bonding part 46. A Review of the correlation and prediction of transport properties by an LFER method: Physicochemical properties, brain penetration and skin permeability. Pestic. Sci., 55, 78–88.

Abraham, M.H., Chadha, H.S., Mitchell, R.C., 1995. The factors that influence skin penetration of solutes. J. Pharm. Pharmacol., 47, 8–16.

Abraham, M.H., Martins, F., 2004. Human skin permeation and partition: General linear free-energy relationship analyses, J. Pharm. Sci. 93, 1508-1523.

ACD/ChemSketch Freeware, version 12.01, Advanced Chemistry Development, Inc., Toronto, ON, Canada, www.acdlabs.com, 2010.

ACD/logD version 11, Advanced Chemistry Development, Inc., Toronto, ON, Canada, www.acdlabs.com, 2008.

ACD/logD version 12.01, Advanced Chemistry Development, Inc., Toronto, ON, Canada, www.acdlabs.com, 2009.

Ademola, J.I., Bloom, E., Maczulak, A.E., Maibach, H.I., 1993. Skin Penetration and Metabolism – Comparative – Evaluation of Skin Equivalent, Cell Culture, and Human Skin. J. Toxicol., Cutaneous Ocul. Toxicol. 12, 129-138.

Akhter, S.A., Bennett, S.L., Waller, I.L., Barry, B.W., 1984. An Automated Diffusion Apparatus For Studying Skin Penetration. Int. J. Pharm. 21, 17-26.

Akrill, P., Cocker, J., Dixon, S., 2002. Dermal exposure to aqueous solutions of N-methyl pyrrolidone, Toxicol. Lett. 134, 265-269.

Anderson, B.D., Higuchi, W.I., Raykar, P.V., 1988. Heterogeneity effects on permeability-partition coefficient relationships in human stratum corneum. Pharm. Res., 5, 566-573.

Anigbogu, A.N.C., Williams, A.C., Barry, B.W., Edwards, H.G.M., 1995. Fourier transform raman spectroscopy of interactions between the penetration enhancer dimethyl sulfoxide and human stratum corneum. Int. J. Pharm. 125, 265-282.

Anissimov, Y.G., Roberts, M.S., 2001. Diffusion modelling of percutaneous absorption kinetics. 2. Finite vehicle volume and solvent deposited solids. J. Pharm. Sci. 90, 504-520.

Attwood, D., 2007. Disperse systems. In: M. E. Aulton (Ed.). Aulton's Pharmaceutics, The Design and Manufacture of Medicines, Churchill Livingstone, Elsevier, 85-90.

Aungst, B.J., Rogers, N.J., Shefter, E., 1986. Enhancement of naloxon penetration through human skin in vitro using fatty acids, fatty alcohols, surfactants, sulfoxides and amides. Int. J. Pharm. 33, 225-234.

Baker, H., 1986. The skin as a barrier. In: Rock, A., (Ed.). Textbook of dermatology. Oxford, Blackwell Scientific, 355-365.

Baroli, B., 2010. Penetration of nanoparticles and nanomaterials in the skin: Fiction or reality? J. Pharm. Sci. 99, 21-50.

Baron, J.M., Wiederholt, T., Heise, R., Merk, H.F., Bickers, D.R., 2008. Expression and function of cytochrome p450-dependent enzymes in human skin cells. Curr. Med. Chem. 15, 2258-2264.

Barry, B.W., 1983. Properties that influence percutaneous absorption, in Dermatological Formulation, Percutaneous Absorption. Marcel Dekker, New York, 127-233.

Barry, B.W., 1987. Mode of action of penetration enhancers in human skin. J. Cont. Rel. 6, 85-97.

Barry, B.W., 1991. Lipid–protein-partititioning theory of skin penetration enhancement, J. Controlled. Release 15, 237– 248.

Barry B.W., 2001. Review: Novel mechanisms and devices to enable successful transdermal drug delivery. Eur. J. Pharm. Sci. 14, 101–114.

Barry, B.W., 2007. Transdermal drug delivery. In: M. E. Aulton (Ed.). Aulton's Pharmaceutics, The Design and Manufacture of Medicines, Churchill Livingstone, Elsevier Chapter 38, 580-585.

Barry, B.W., Williams, A.C., 1995. Permeation enhancement through skin. In: Swarbrick, J., Boylan, J.C. (Eds.). Encyclopedia of Pharmaceutical Technology, Vol. 11. Marcel Dekker, New York, 449–493.

Bartow, R.A., Brogden, R.N., 1998. Formoterol. An update of its pharmacological properties and therapeutic efficacy in the management of asthma. Drugs 55, 303-322.

Bashir, S.J., Chew, A.L., Anigbogu, A., Dreher, F., Maibach, H.I., 2001. Physical and physiological effects of stratum corneum tape stripping. Skin. Res. Technol. 7, 40-48.

Bataller, R., Bragulat, E., Nogué, S., Görbig M., N., Bruguera, M., Rodés, J. 1999. Prolonged cholestasis after acute paraquat poisoning through skin absorption. Am. J. Gastroenterol. 95, 1340-1343.

Baynes, R.E., Brooks, J.D., Mumtaz, M., Riviere, J.E., 2002. Effect of chemical interactions in pentachlorophenol mixtures on skin and membrane transport. Toxicol. Sci. 69, 295– 305.

Belsey, N.A., Cordery, S.F., Bunge, A.L., Guy, R.H. 2011. Assessment of dermal exposure to pesticide residues during re-entry. Environ. Sci. Technol. 45, 4609-4615

Benfeldt, E., Groth, L., 1998. Feasibility of measuring lipopbilic or protein-bound drugs in the dermis by in vivo microdialysis after topical or systemic drug administration. Acta. Derm. –Venereol. 78, 274-278.

Beydon, D., Payan, J.P., Grandclaude, M.C. 2010. Comparison of percutaneous absorption and metabolism of di-n-butylphthalate in various species. Toxicol. in Vitro 24, 71–78.

Bezema, F.R., Marttin, E., Roemele, P.E.H., Brussee, J., Bodde, H.E., de Groot, H.J.M., 1996. H-2 NMR evidence for dynamic disorder in human skin induced by the penetration enhancer Azone. Spectrochim. Acta. Mol. Biomol. Spectros., 52, 785–791.

Blank, I.H., McAuliffe, D.J., 1985. Penetration of benzene through human skin. J. Invest. Dermatol. 85, 522-526.

Bouwman, T., Cronin, M.T.D., Bessems, J.GM., van de Sandt, J.J.M., 2008. Improving the applicability of (Q)SARs for percutaneous penetration in regulatory risk assessment. Hum. Exp. Toxicol. 27:269-276.

Bouwstra, J.A., de Graaff, A., Gooris, G.S., Nijsse, J., Wiechers, J.W., van Aelst, A.C., 2003. Water distribution and related morphology in human stratum corneum at different hydration levels. J. Invest. Dermatol. 120, 750-758.

Bouwstra, J.A., Honeywell-Nguyen, P.L., Gooris, G.S., Ponec, M., 2003. Structure of the skin barrier and its modulation by vesicular formulations. Prog. Lipid Res. 42, 1-36.

Brain, K.R., Walters, K.A., Watkinson, A.C., 1998. Investigation of skin permeation in vitro. In: Roberts, M.S., Walter, K.A., (Eds.). Dermal absorption and toxicity assessment. Drugs and the Pharmaceutical Sciences Vol. 91. New York, Marcel Dekker, 161-187.

Bronaugh, R.L., 2004. Methods for in vitro percutaneous absorption. In: Zhai, H., Maibach, H.I., (Eds.). Dermatotoxicology, 6th ed. New York, CEC Press, 520-526.

Bronaugh, R.L., Maibach, H.I., 1985. Percutaneous absorption of nitroaromatic compounds: in vivo and in vitro studies in the human and monkey. J. Invest. Dermatol. 84, 180-183.

Bronaugh, R.L., Stewart, R.F., Congdon, E.R., 1983. Differences in permeability of rat skin related to sex and body site. J. Sot. Cosmet. Chem. 34, 127-135.

Bronaugh, R.L., Stewart, R.F., 1985. Methods for *in vitro* percutaneous absorption studies IV: the flow-through diffusion cell. J. Pharm. Sci. 74, 64-67.

Bronaugh, R.L., Stewart, R.F., Simon, M., 1986. Methods for in vitro percutaneous absorption studies. VII: Use of excised human skin. J. Pharm. Sci. 75, 1094–1097.

Brooks, J.D., Riviere, J.E., 1996. Quantitative percutaneous absorption and cutaneous distribution of binary mixtures of phenol and para-nitrophenol in isolated perfused porcine skin. Fundam. Appl. Toxicol. 32, 233– 243.

Bucks, D., Guy, R., Maibach, H., 1991. Effects of occlusion. In: Bronaugh, R.L., Maibach, H.I., (Eds.). In vitro percutaneous absorption: principles, fundamentals, and applications. Boca Raton, FL, CRC Press, 85-114.

Buist, H.E., de Heer, C., Bessems, J.G.M., Bouwman, T., Chou, S., Pohl, H.R., van de Sandt, J.J.M., 2005. Dermal absorption in risk assessment: the use of relative absorption versus permeation coefficient (Kp). Occupational and environmental exposures of skin to chemicals - 2005. Abstract for Poster 25, URL: http://www.cdc.gov/niosh/topics/skin/OEESC2/AbPost025Buist.html.

Buist, H.E., van Burgsteden, J.A., Freidig, A.P., Maas, W.J.M., van de Sandt, J.J.M., 2010. New in vitro dermal absorption database and the prediction of dermal absorption under finite conditions for risk assessment purposes. Regul. Toxicol. Pharmacol. 57, 200-209.

Burbidge, R., Trotter, M., Buxton, B., Holden, S., 2001. Drug design by machine learning: support vector machines for pharmaceutical data analysis. Comput. Chem. 26(1), 5-14.

CambridgeSoft, 2009. ChemBioFinder, http://www.cambridgesoft.com/databases, Retrieved September 2009.

CambridgeSoft, 2011. ChemBioFinder, http://www.cambridgesoft.com/databases, retrieved November 2010 - February 2011.

Carrupt, P.A., Testa, B., Gaillard, P., 1997. Computational Approaches to Lipophilicity: Methods and Applications. In Lipkowitz, K.B., Boyd, D.B., Eds. ReViews in Computational Chemistry, Wiley, New York, Vol. 11, 241-315.

Challapalli, P.V.N., Stinchcomb, A.L., 2002. In vitro experiment optimization for measuring tetrahydrocannabicol skin permeation. Int. J. Pharm., 241, 329-339.

Chang, S., Riviere, J.E., 1993. Effect of humidity and occlusion on the percutaneous absorption of parathion in vitro. Pharm. Res. 10, 152-155.

ChemSpider Home Page. http://www.chemspider.com/ (accessed Sept 2010).

Chen, H.F., 2008. Quantitative predictions of gas chromatography retention indexes with support vector machines, radial basis neural networks and multiple linear regression. Anal. Chim. Acta 609, 24-36.

Cheng, T.J., Zhao, Y., Li, X., Lin, F., Xu, Y., Zhang, X.L., Li, Y., Wang, R.X., Lai, L.H., 2007. Computation of octanol-water partition coefficient by guiding an additive model. J. Chem. Inf. Model. 47, 2140-2148.

Chilcott, R.P., Barai, N., Beezer, A.E., Brain, S.I., Brown, M.B., Bunge, A.L., Burgess, S.E., Cross, S., Dalton, C.H., Dias, M., Farinha, A., Finnin, B.C., Gallagher, S.J., Green, D.M., Gunt, H., Gwyther, R.L., Heard, C.M., Jarvis, C.A., Kamiyama, F., Kasting, G.B., Ley, E.E., Lim, S.T., McNaughton, G.S., Morris, A., Nazemi, M.H., Pellett, M.A., du Plessis, J., Quan, Y.S., Raghavan, S.L., Roberts, M., Romonchuk, W., Roper, C.S., Schenk, D., Simonsen, L., Simpson, A., Traversa, B.D., Trottet, L., Watkinson, A., Wilkinson, S.C., Williams, F.M., Yamamoto, A. Hadgraft, J., 2005. Inter- and intra-laboratory variation of *in vitro* diffusion cell measurements: An international multi-centre study using quasi-standardised methods and materials. J. Pharm. Sci. 94, 632-638.

Chu, K.A., Yalkowsky, S.H., 2009. An interesting relationship between drug absorption and melting point. Int. J. Pharm. 373, 24-40

Cilurzo, F., Alberti, E., Minghetti, P., Gennari, C.G.M., Casiraghi, A., Montanari, L., 2010a. Effect of drug chirality on the skin permeability of ibuprofen, Int. J. Pharm. 386, 71–76.

Cilurzo, F., Minghetti, P., Alberti, E., Gennari, C.G.M., Pallavicini, M., Valoti, E., Montanari, L., 2010b. An investigation into the influence of Counterion on the RS-Propranolol and S-Propranolol skin permeability. J. Pharm. Sci. 99, 1217-1224.

Clowes, H.M., Scott, R.C., Heylings, J.R., 1994. Skin absorption: Flow-through or static diffusion cells. Toxicol. in Vitro, 827-830.

Cnubben, N.H., Elliot G.R., Hakkert, B.C., Meuling, W.J., van de Sandt, J.J., 2002. Comparative *in vitro – in vivo* percutaneous penetration of the fungicide *ortho*-phenylphenol. Regul. Toxicol. Pharmacol. 35, 198-208.

Collier, S.W., Sheikh, N.M., Sakr, A., Lichtin, J.L., Stewart, R.F., Bronaugh, R.L., 1989. Maintenance of skin viability during in vitro percutaneous absorption/metabolism studies. Toxicol. Appl. Pharmacol. 99, 522-533.

Commission of the European Communities. 2003. Proposal for a Regulation of the European Parliament and of the Council Concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency and amending Directive 1999/45/EC and Regulation (EC) {on Persistent Organic Pollutants}. http://europa.eu.int/comm/enterprise/chemicals/chempol/whitepaper/reach.htm.

Cooper, E.R., 1984. Increased skin permeability of lipophilic molecules. J. Pharm. Sci. 73, 1153-1155.

Cooper, E.R., 1985. Vehicle effects on skin penetration, In:. Bronaugh, R.L., Maibach, H.I., (Eds.). Percutaneous Absorption. Marcel Dekker, New York, 525-530.

Cooper, E.R., Berner, B., 1987. In: Kydonieus A.F., Berner, B., (Eds.). Penetration enhancers, in Transdermal Delivey of Drugs, Vol. 2. CRC Press, Boca Ratom, FL, 57-62.

Copovi, A., Diez-Sales, O., Herraez-Dominguez, J. V., Herraez-Dominguez, M., 2006. Enhancing effect of alpha-hydroxyacids on "in vitro" permeation across the human skin of compounds with different lipophilicity. Int. J. Pharm. 314, 31-36.

Cornwell, P.A., Barry, B.W., Bouwstra, J.A., Gooris, G.S., 1996. Modes of action of terpene penetration enhancers in human skin; differential scanning calorimetry, small-angle X-ray diffraction and enhancer uptake studies, Int. J. Pharm. 127, 9 – 26.

Cronin, M.T.D., Dearden, J.C., Moss, G.P., Murray-Dichinson, G., 1999. Investigation of the mechanism of flux across human skin in vitro by quantitative structure permeability relationships. Eur. J. Pharm. Sci. 7, 325-330.

Crowther, J.M., Sieg, A., Blenkiron, P., Marcott, C., Matts, P.J., Kaczvinsky, J.R., Rawlings, A.V., 2008. Measuring the effects of topical moisturizers on changes in stratum corneum thickness, water gradients and hydration in vivo. Br. J. Dermatol. 159, 567-577.

Cruciani, G., Crivori, P., Carrupt, P.A, Testa, B., 2000. Molecular fields in quantitative structure permeation relationships: the VolSurf approach. J. Mol Struct. - THEOCHEM 503, 17-30.

Cumming, K.I., Winfield, A.J., 1994. In-vitro evaluation of sodium carboxylates as dermal penetration enhancers. Int. J. Pharm. 108, 141-148.

Czerminski, R., Yasri, A., Hartsourgh, D., 2001. Use of Support Vector Machine in pattern classification: Application to QSAR studies. Quant. Struct.-Act. Relat. 20, 227-240.

Dal Pozzo, A., Donzelli, G., Liggeri, E., Rodriguez, L., 1991. Percutaneous absorption of nicotinic acid derivatives *In vitro*. J. Pharm. Sci. 80, 54-56.

Davis, A.F., Gyurik, R.J., Hadgraft, J., Pellett, M.A., Walters, K.A., 2002. Formulation strategies for modulating skin permeation. In Walters, K. A., ed. Dermatological and transdermal formulations. New York, Marcel Dekker, pp 271-317 (Drugs and the Pharmaceutical Sciences Vol. 119).

Dearden, J.C., Cronin, M.T.D., 2005. Quantitative structure-activity relationships (QSAR) in drug design. In: Smith H.J., (ed.) Smith and Williams' Introduction to the Principles of Drug Design and Action. 4 th Edition. Taylor and Francis, Boca Raton FL, USA, 185-209.

Dearden, J. C., Cronin, M.T.D., Patel, H., Raevsky, O.A., 2000. QSAR prediction of human skin permeability coefficient. J. Pharm. Pharmacol. 52 (S), 221.

Dearden, J.C., Ghafourian, T., 1999. Hydrogen bonding parameters for QSAR: Comparison of indicator variables, hydrogen bond counts, molecular orbital and other parameters. J. Chem. Inf. Comp. Sci. 30, 231-235.

Degim, T., Hadgraft, J., Ilbasmis, S., Ozkan, Y., 2003. Prediction of skin penetration using artificial neural network (ANN) modeling. J. Pharm. Sci. 92, 656-664.

De Vito, S. C., 2000. Absorption through cellular membrane. In Boethling, R.S., Mackay, D., (Eds.). Handbook of Property Estimation Methods for Chemicals. CRC Press, Boca Raton, 261.

Dias, M., Farinha. A., Faustino. E., Hadgraft. J., Pais, J., Toscano, C., 1999. Topical delivery of caffeine from some commercial formulations. Int. J. Pharm. 182 41–47.

Dick, I.P., Blain, P.G., Williams, F.M., 1997. The percutaneous absorption and skin distribution of lindane in man. I. In vivo studies. Hum. Exp. Toxicol. 16, 645-651.

Dick, I.P., Blain, P.G., Williams, F.M., 1997. The percutaneous absorption and skin distribution of lindane in man. II. In vivo studies. Hum. Exp Toxicol 16, 652-657.

Dick, D., Ng, K.M.E., Sauder, G.N., Chu, I., 1995. In Vitro and In Vivo Percutaneous Absorption of 14C-Chloroform in Humans. Hum. Exp. Toxicol. 14, 260-265.

Diudea, M.V., Gutman, I., 1998. Wiener-type topological indices. Croat. Chem. Acta 71, 21-51.

Dreher, F., Arens, A., Hostynek, J.J., Mudumba, S., Ademóla, J., Maibach, H.I., 1998. Colorimetric method for quantifying buman Stratum corneum removed by adhesive-tape stripping. Acta. Derm. Venereol. 78, 186-189.

Du, H.Y., Wang, J., Yao, X.J., Hu, Z.D. 2009. Quantitative Structure-Retention Relationship models for the prediction of the reversed-phase HPLC gradient retention based on the heuristic method and Support Vector Machine. J. Chromatogr. Sci. 47, 396-404.

ECB Technical Guidance Document on Risk assessment in support of commission Directive 93/67/EEC on Risk Assessment for new substances and Commission Regulation (EC) No. 1488/94 on Risk Assessment for existing substance. EUR 20418 EN/1.

ECETOC, 1993. Percutaneous absorption. Brussels, European Centre for Ecotoxicology and Toxicology of Chemicals, 1-80 (Monograph No. 20).

EDETOX, 2010. Database created by the University of Newcastle (http://edetox.ncl.ac.uk/), accessed June 2010.

Elias, P.M., 1983. Epidermal lipids, barrier function and desquamation. J. Invest. Dermatol. 80, 44-49.

Engstrom, K., Husman, K., Riihimaki, V., 1977. Percutaneous absorption of m-xylene in man. Int. Arch. Occup. Environ. Health 39, 181-189.

Environmental Protection Agency of U.S., 1992. Risk assessment forum. Guidelines for exposure assessment Published on May 29, 1992, Federal Register 57(104):22888-22938. Washington, DC, URL: http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=15263#Download

Eros, D., Kovesdi, I., Orfi, L., Takacs-Novak, K., Acsady, G., Keri, G., 2002. Reliability of log P predictions based on calculated molecular descriptors: A critical review. Curr. Med. Chem. 9, 1819- 1829.

Escuder-Gilbert, L., Martinez-Pla, J.J., Sagrado, S., Villanueva-Camanas, R.M., Medina-Hernandez, M.J., 2003. Biopartitioning micellar separation methods: modelling drug absorption. J. Chromatogr. 797, 21-35.

European Chemicals Bureau, Ispra, Italy, website accessed March 2008, http://ecb.jrc.it/reach/reach-legislation/.

European Commission, 2002. Guidance Document on Dermal Absorption Directorate E1 – Plant Health. Sanco/333/2000, Rev. 6, November 2002.

Fang J.Y., Tsai T.H., Lin Y.Y., Wong W.W., Wang M.N., Huang J.F., 2007. Transdermal delivery of tea catechins and theophylline enhanced by terpenes: A mechanistic study. Biol. Pharm. Bull. 30, 343-349.

Fasano, W.J., McDougal, J.N., 2008. In vitro dermal absorption rate testing of certain chemicals of interest to the Occupational Safety and Health Administration: Summary and evaluation of USEPA's mandated testing. Regul. Toxicol. Pharmacol. 51, 181-194.

Faulds, D., Hollingshead, L.M., Goa, K.L., 1991. Formoterol. A review of its pharmacological properties and therapeutic potential in reversible obstructure airways disease. Drugs, 42, 115 – 137.

Fedors, R.F., 1974. Methods for estimating both solubility parameters and molar volumes of liquids. Polym. Eng. Sci. 14, 47-154.

Fitzpatrick, D., Corish, J., Hayes, B., 2004. Modelling skin permeability in risk assessment – the future. Chemosphere 55, 1309 – 1314.

Flynn, G. L., 1990. Physicochemical determinants of skin absorption. In: Gerrity, T.R. Henry, C.J. (Eds.). Principles of route to route extrapolitation for risk assessment extraction for skin assessment, Elsevier, NewYork, NY, 93-127.

Forslind, B., Engstrom, S., Engblom, J., Norlen, L., 1997. A novel approach to the understanding of human skin barrier function. J. Dermatol. Sci. 14, 115-125.

Friend, D.R., 1992. In vitro skin permeation techniques. J. Controlled Release 18, 235-248.

Ghafourian, T., Fooladi, S., 2001. The effect of structural QSAR parameters on skin penetration. Int. J. Pharm., 217, 1-11.

Ghafourian, T., Samaras, E.G., Brooks, J.D., Riviere, J.E., 2010a. Modelling the effect of micture components on permeation through skin. Int. J. Pharm. 398, 28-32.

Ghafourian, T., Samaras, E.G., Brooks, J.D., Riviere, J.E., 2010b. Validated models for predicting skin penetration from different vehicles. Eur. J. Pharm. Sci. 41, 612-616.

Ghafourian, T., Zandasrar, P., Hamishekar, H., Nokhodchi, A., 2004. The effect of penetration enhancers on drug delivery through skin: a QSAR study. J. Controlled Release 99, 113-125.

Goudarzi, N., Goudarzi. M., 2009. Prediction of the acidic dissociation constant (pka) of some organic compounds using linear and nonlinear QSPR methods. Mol. PhyS. 107(14), 1495-1503.

Gramatica, P., 2007. Principles of QSAR models validation: internal and external. QSAR Comb. Sci. 26, 694-701.

Gramatica, P. A short history of QSAR evolution. QSAR Research Unit in Environmental Chemistry and Ecotoxicology, DBSF, Insubria University, Varese, Italy. Obtained online from: http://www.qsarworld.com/Temp_Fileupload/Shorthistoryofqsar.pdf Accessed July 2011.

Gramatica, P., Pilutti, P., Papa, E., 2002. Ranking of volatile organic compounds for tropospheric degradability by oxidants: a QSPR Approach. SAR QSAR Environ. Res. 13, 743-753.

Greaves, L.C., Wilkinson, S.C., Williams, F.M., 2002. Factors Affecting Percutaneous Absorption of Caffeine in Vitro. Toxicology, 178, 65-66.

Gutman, I., Körtvélyesi, T., 1995. Wiener indices and molecular surfaces. Zeitschrift für Naturforschung. A, A journal of physical sciences 50a, 669-671.

Hadgraft, J., Lane, M.E., 2009. Transepidermal water loss and skin site: A hypothesis. Int. J. Pharm. 373, 1-3.

Hadgraft, J., (2004). Review article: Skin deep. Eur. J. Pharm. Biopharm. 58, 291–299.

Hadjmohammadi, M.R., Fatemi, M.H., Kamel, K., 2007. Quantitative structure-property relationship study of retention time for some pesticides in gas chromatography. J. Chromatogr. Sci. 45, 400-404.

Hall L.H., Kier L.B., 1977. The nature of Structure-Activity Relationships and their relation to molecular connectivity. Eur. J. Med. Chem 12, 307.

Hall L.H., Kier L.B., 1986. Molecular Connectivity in Structure-Activity Analysis, J. Wiley and Sons, New York.

Hall, L.H., Kier, L.B., 1991. The molecular connectivity chi indices and kappa shape indices in structure-property modeling. In: Boyd, D., Lipkowitz, K., (Eds.). Reviews in Computational Chemistry, VCH, New York, 384–385.

Hansch, C., Clayton, J.M., 1973. Lipophilic character and biological-activity of drugs .2. Parabolic case. J. Pharm. Sci., 62, 1-21.

Hansch, C., Leo, A., 1995. Exploring QSAR: Fundamentals and Applications in Chemistry and Biology; American Chemical Society: Washington DC.

Hansen, C.M., 1967. The three dimensional solubility parameter – key to paint component affinities I. Solvents, plasticizers, polymers and resins. J. Paint Technol. 39, 104-117.

Hawkins, D.M., Basak, S.C., Mills, D., 2003. Assessing model fit by cross-validation. J. Chem. Inf. Comp. Sci. 43, 579 – 586.

Helma, C., 2004. SAR QSAR Environ. Res. 15, 367 – 383.

Henning, A., Schaefer, U.F., Neumann, D., 2009. Potential pitfalls in skin permeation experiments: Influence of experimental factors and subsequent data evaluation. Eur. J. Pharm. Biopharm. 72, 324-331.

Ho, N.F.H., Park, J.Y., Morozowich, W., Higuichi. W.I., 1997. Physical model approach to the design of drugs with improved intestinal absorption. In: Roche E.B. (Ed.) Design of Biopharmaceutical Properties through Prodrugs and Analogues, AphA/APS, Washington, 136.

Hou, T.J., Xia, K., Zhang, W., Xu, X.J., 2004. ADME Evaluation in Drug Discovery. 4. Prediction of Aqueous Solubility Based on Atom Contribution Approach. J. Chem. Inf. Comput. Sci. 44. 266-275.

Humbert, H, Cabiac M.D., Denouel, J., Kirkesseli, S., 1995. Pharmacokinetics of terbinafine and of its five main metabolites in plasma and urine, following a single oral dose in healthy subjects. Biopharm. Drug Dispos. 16, 685–694.

Hueber-Becker, F., Nohynek, G.J., Meuling, W.J., Benech-Kieffer, F., Toutain, H., 2004. Human systemic exposure to a [14G]-para-phenylenediamine-containing oxidative bair dye and correlation with in vitro percutaneous absorption in human or pig skin. Food Chem. Toxicol. 42, 1227-1236.

Hunter, J., Savin, J., Dahl, M., 2002. Clinical Dermatology. In: Elliot, J. (Ed.) 3rd ed. Oxford: Wiley-Blackwell.

Idson, B., 1983. Vehicle effects in percutaneous absorption. Drug Metab. Rev. 14, 207– 222.

Jain, A.K., Thomas, S.N., Panchagnula, R., 2002. Transdermal drug delivery of imipramine hydrochloride. I. Effect of terpenes. J. Controlled Release 79, 93-101.

Jakasa, I., Mohammadi, N., Kruse, J., Kezic, S., 2004. Percutaneous absorption of neat and aqueous solutions of 2-butoxyethanol in volunteers. Int. Arch. Occup. Environ. Health 77, 79-84.

Jewell, C., Heylings, J.R., Clowes, H.M., Williams, F.M., 2000. Percutaneous absorption and metabolism of dinitrochlorobenzene in vitro. Arch. Toxicol. 74, 356-365.

Johnson, M.E., Blankschtein, D., Langer, R., 1995. Permeation of steroids through human skin. J. Pharm. Sci. 84, 1144-1146.

Jones, A.D., Dick, I.P., Cherrie, J.W., Cronin, M.T.D., Van De Sandt, J.J.M., Esdaile, D.J., Lyengar, S., ten Berge, W., Wilkinson, S.C., Roper, C.S., Semple, S., de Heer, C., Williams, F.M., 2004. CEFIC Workshop on methods to determine dermal permeation for human risk assessment. Held in Utrecht 13-15th June. Research Report TM/04/07.

Jung, C.T., Wickett, R.R., Desai, P.B., Bronaugh, R.L., 2003. In vitro and in vivo percutaneous absorption of catechol. Food Chem. Toxicol. 41, 885-895.

Kakubari, I., Nakamura, N., Takayasu, T., Yamauchi, H., Takayama, S., Takayama, K., 2006. Effects of solvents on Skin Permeation of Formoterol Fumarate. Biol. Pharm. Bull. 29, 146 – 149.

Kai, T., Mak, V.H.W., Potts, R.O., Guy, R.H., 1990. Mechanisms of percutaneous penetration enhancement: effect of n-alkanols on the permeability barrier of hairless mouse skin. J. Control. Release 12, 103–112.

Kang, L., Yap, C.W., Lim, P.F.C., Chen Y.Z., Ho, P.C., Chan, Y.W., Wong, G.P., Chan, S.Y., 2007. Formulation development of transdermal dosage forms: Quantitative structure-activity relationship model for predicting activities of terpenes that enhance drug penetration through houman skin. J. Controlled Release 120, 211-219.

Kanikkannan, N., Kandimalla, K., Lamba, S.S., Singh, M., 2000. Structure-activity relationship of chemical penetration enhancers in transdermal drug delivery. Curr. Med. Chem.7, 593-608.

Kaplun-Frischoff, Y., Touitou, E., 1997. Testosterone skin permeation enhancement by menthol through the formation of a eutectic with drug and interaction with skin lipids. J. Pharm. Sci. 86, 1349–99.

Karande, P., Jain, A., Ergun, K., Kispersky, V., Mitragotri, S., 2005. Design principles of chemical penetration enhancers for transdermal drug delivery. Proc. Natl. Acad. Sci. U. S. A. 102, 4688-4693.

Kat, M., Poulsen, B.J., 1971. In: Brodie B.B., Gillette J., (Eds.). Handbook of Experimental Pharmacology, Vol. 28. Springer-Verlag, New York, Part 1, 103.

Katritzky, A.R., Dobchev, D.A., Fara, D.C., Hur, E., Tamm, K., Kurunczi, L., Karelson, M., Varnek, A., Solov'ev, V.P., 2006. Skin Permeation Rate as a Function of Chemical Structure. J. Med. Chem. 49, 3305-3314

Kezic, S., 2008. Eurotox Article: Methods for measuring in-vitro percutaneous absorption in humans. Hum. Exp. Toxicol. 27, 289-295.

Kier, L.B., Hall, L.H., 1999. Molecular Structure Description: the Electrotopological State; Academic Press: San Diego.

Kligman, A.M., 1983. A biological brief on percutaneous absorption. Drug Dev. Ind. Pharm. 9, 521-560.

Knutson, K., Potts, R.O., Gusek, D.B., Golden, G.M., McKie, J.E., Lambert, W.J., Higuchi, W.I., 1985. Macro- and molecular physical-chemical considerations in understanding drug transport in the stratum corneum, J. Controlled Release 2, 67-87.

Kou, J.H., Roy, S.D., Du, J., Fujiki, J., 1993. Effect of receiver fluid pH on in Vitro skin flux of weakly ionizable drugs. Pharm. Res. 10, 986-990.

Korinth, G., Schaller, K.H., Drexler, H., 2005. Is the permeability coefficient k_p a reliable tool in percutaneous absorption studies? Arch. Toxicol. 79, 155-159.

Kraker, J.J., Hawkins, D.M., Basak, S.C., Nataralan, R., Mills, D., 2006. Chemom. Intell. Lab. Syst., online: doi:10.1016/ j.chemolab.2006.03.001.

Krishnaiah, Y.S.R., Bhaskar, P., Satyanarayana, V., 2004. Penetration-enhancing effect of ethanol-water solvent system and ethanolic solution of carvone on transdermal permeability of nimodipine from HPMC gel across rat abdominal skin. Pharm. Dev. Technol. 9, 63-74.

Krishnaiah, Y.S.R., Satyanarayana, V., Karthikeyan, R.S., 2002. Effect of the solvent system on the in vitro permeability of nicardipine hydrochloride through excised rat epidermis. J. Pharm. Pharmaceut. Sci. 5, 123-130.

Kubinyi, H., 1977. Quantitative structure-activity-relationships 7. Bilinear model, a new model for nonlinear dependence of biological-activity on hydrophobic character. J. Med. Chem., 20, 625-629.

Laugel, C., Yagoubi, N., Baillet, A., 2005. 'ATR-FTIR spectroscopy: a chemometric approach for studying the lipid organisation of the stratum corneum.' Chem. Phys. Lipids, 135, 55-68.

Laughlin, R.G., 1978. Relative hydrophilicities among surfactants hydrophilic groups. In: Brown, G.H. (Ed.). Advances in Liquid Crystals. Academic Press, New York.

Lee, F.W., Earl, L., Williams, F.M., 2001. Interindividual variability in the percutaneous penetration of testosterone through human skin *in vitro*. Toxicology 168, 63.

Leo, A.J. 1993. Calculating $\log P_{oct}$ from Structures. Chem. ReV. 93, 1281-1306.

Leveque, N., Makki, S., Hadgraft, J., Humbert, P., (2004). Comparison of franz cells and microdialysis for assessing salicylic acid penetration through human skin. Int. J. Pharm. 269, 323-328.

Lien, E.J., Gao, H., 1995. QSAR analysis of the skin permeability of various drugs in man as compared to in vivo and in vitro studies in rodent. Pharm. Res. 4, 583-587.

Linusson, A., Elofsson, M., Andersson, I.E., Dahlgren, M.K., 2010. Statistical molecular design of balanced compounds libraries for QSAR modelling. Curr. Med. Chem. 17, 2001-2016.

Liou Y.B., Ho, H.O., Yang, C.J., Lin, Y.K., Sheu, M.T., 2009. Construction of a quantitative structure-permeability relationship (QSPR) for the transdermal delivery of NSAIDs. J. Controlled Release 138, 260-267.

Lipinski, C.A., Lombardo, F., Dominy, B.W., Feeney, P.J., 1997. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Del. Rev., 23, 3–25.

Liron, Z., Cohen, S., 1984. Percutaneous absorption of alkanoic acids. I: A study of operational conditions. J. Pharm. Sci. 73, 534–537

Liron, Z., Cohen, S., 1984. Percutaneous absorption of alkanoic acids II. application of regular solution theory. J. Pharm. Sci. 73, 538-542.

Liu, P., Long, W., 2009. Current mathematical methods used in QSAR/QSPR studies. Int. J. Mol. Sci. 10, 1978-1998.

Liu, H.X., Zhang, R.S., Yao, X.J., Liu, M.C., Hu, Z.D., Fan B.T., 2003. QSAR study of ethyl 2-[(3-Methyl-2,5-dioxo(3-pyrrolinyl))amino]-4-(trifluoromethyl)pyrimidine-5- carboxylate : an inhibitor of AP-1 and NF-kB mediated gene expression based on support vector machines. J. Chem. Inf. Comput. Sci. 43, 1288-1296.

Livingstone, D.J., 2000. The characterization of chemical structures using molecular properties. A survey. J. Chem. Inf. Comput. Sci. 40, 195-209.

López, A., Faus, V., Díez-Sales, O., Herráez, M., 1998. Skin permeation model of phenyl alcohols: comparison of experimental conditions. Int. J. Pharm., 173, pp.183-191.

Lotte, C., Wester, R.C., Rougier, A., Maibach, H.I., 1993. Racial differences in the in vivo percutaneous absorption of some organic compounds: a comparison between black, Caucasian and Asian subjects. Arch. Dermatol. Res. 284, 456-459.

Luan, F., Liu, H.T., Wen, Y.Y., Zhang, X.Y., 2008. Classification of the fragrance properties of chemical compounds based on support vector machine and linear discriminant analysis. Flavour Fragrance J. 23, 232-238.

Lundh, T., Boman, A., Akesson, B., 1997. Skin absorption of the industrial catalyst dimethylethylamine in vitro in guinea pig and human skin, and of gaseous dimethylethylamine in human volunteers. Int. Arch. Occup. Environ. Health 70, 309–313.

Ma, Q.H., Hao, X.Z., Zhou, H.F., Gu, N., 2007. Effect of surfactants on preparation and skin penetration of tea polyphenols liposomes. IEEE/ICME International conference on complex medical engineering, Vol. 1-4, 209-212.

MacGrath, J.A., Uitto, J., 2010. Anatomy and organization of human skin. In: Burns, T., Breathnack, S., Cox, N., Griffitshs, C., (Eds.). Rook's Textbook of Dermatology, 8th edn., ch. 3. Blackwell, Oxford.

Magnusson, B.M., Pugh, W.J., Roberts, M.S., 2004. Simple Rules Defining the Potential of Compounds for Transdermal Delivery or Toxicity. Pharm. Res. 21, 1047-1054.

MDL QSAR, 2.3.0.0.12, Symyx Technologies Inc. merged with Accelrys, Inc. in 2009, Accelrys Software Inc., San Diego, http://accelrys.com.

Mehta, M., Kemppainen, B.W., Stafford, R.G. 1991. In vitro penetration of tritium-labelled water (THO) and [3H]PbTx-3 (a red tide toxin) through monkey buccal mucosa and skin. Toxicol. Lett. 55, 185–194.

Michaels, A.S., Chandrasekaran, S.K., Shaw J.E., 1975. Drug permeation through human skin: theory and in vitro experimental measurement. Am. Inst. Chem. Eng. J. 21, 985-996.

Mills, P.C. Cross, S.E., (2006) Transdermal drug delivery: Basic principles for the veterinarian. Vet. J. 172, 218-233.

Minitab 13.1.0.0 Statistical Software (2008). [Computer software]. State College, PA: Minitab, Inc. (www.minitab.com)

Minitab 15.1.0.0 Statistical Software (2010). [Computer software]. State College, PA: Minitab, Inc. (www.minitab.com)

MOE version 2011.10. Chemical Computing Group Inc. Montreal, Quebec, Canada, http://www.chemcomp.com/, 2011.

Monene, L.M., Goosen, C., Breytenbach, J.C., Hadgraft, J., 2005. Percutaneous absorption of cyclizine and its alkyl analogues. Eur. J. Pharm. Sci. 24, 239-244.

Monteiro-Riviere, N.A., 1986. Ultrastructural evaluation of the porcine integument. In: Tumbleson, M.E. (Ed.), Swine in Biomedical Research, vol. 1., Plenum, New York, pp. 641–655.

Monteiro-Riviere, N.A., Inman, A.O., Mak, V., Wertz, P., Riviere, J.E., 2001. Effect of selective lipid extraction from different body regions on epidermal barrier function. Pharm. Res., 18, 992–998.

Moody, R.P., 2000. Automated in vitro dermal absorption (AIVDA): predicting skin permeation of atrazine with finite and infinite (swimming/bathing) exposure models. Toxicol. in Vitro 14, 467-474.

Moody, R.P., 1997. Automated in vitro dermal absorption (AIVDA): A new in vitro method for investigating transdermal flux. Altern. Lab. Anim. 25, 347-357.

Moody R.P., Nadeau, B., Chu, Ih. 1995. In vivo and In vitro dermal absorption of benzo[a]pyrene In rat, quinea pig, human and tissue-cultured skin. J. Dermatol. Sci. 9, 48-58.

Moore, W.J., 1972. Physical Chemistry. 5th edition, Longman publishing, Prentice-Hall.

Moss, G.P., Cronin, M.T.D., 2002. Quantitative structure-permeability relationships for percutaneous absorption: re-analysis of steroid data. Int. J. Pharm. 238, 105–109.

Moss, P.G., Dearden, J.C., Patel, H., Cronin, M.T.D., 2002. Quantitative structure-permeability relationships (QSPRs) for percutaneous absorption. Toxicol. In-Vitro 16, 299-317.

Moss, T., Howess, D., Williams, F.M., 2000. Percutaneous penetration and dermal metabolism of triclosan. J. Pharmacol. Exp. Ther. 38, 361-370.

Narishetty, S.T.K., Panchagnula, R., 2004. Transdermal delivery of zidovudine: Effect of terpenes and their mechanism of action. J. Controlled Release 95, 367-379.

Narishetty, S.T.K., Panchagnula, R., 2005. Effect of L-menthol and 1,8-cineole on phase behaviour and molecular organization of SC lipids and skin permeation of zidovudine. J. Controlled Release 102, 59-70.

Nathan, D., Sakr, A., Lichtin, J.L., Bronaugh, R.L., 1990. In vitro skin absorption and metabolism of benzoic-acid, p-aminobenzoic acid, and benzocaine in the hairless guinea-pig. Pharm. Res. 7, 1147-1151.

Neumann, D., Kohlbacher, O., Merkwirth, C., Lengauer, T., 2006. A Fully Computational Model for Predicting Percutaneous Drug Absorption, J. Chem. Inf. Model. 46, 424-429.

Nielsen, J.B., Nielsen, F., Sorensen, J.A., 2004. In vitro percutaneous penetration of five pesticides – Effects of molecular weight and solubility characteristics. Ann. Occup. Hyg. 48, 697-705.

Nokhodchi, A., Sharabiani, K., Rashidi, M.R., Ghafourian, T., 2007. The effect of terpene encentration on the skin penetration of diclofenac sodium. Int. J. Pharm. 335, 97-105.

Nokhodchi, A., Shokri, J., Barzegar-Jalali, M., Ghafourian, T., 2003. The enhancement effect of surfactants on the penetration of lorazepam through rat skin, Int. J. Pharm. 250, 359-369.

OECD, 2004a. OECD Guidance Document No.28: Guidance Document for the Conduct of Skin absorption studies, OECD Paris, 2004.

OECD, 2004b. OECD guideline for the testing of chemicals. Skin absorption: *in vitro* method. 428. Adopted: 13 April 2004. Paris, Organisation for Economic Co-operation and Development. 1-8.

Okabe, H., Obata, Y., Takayama, K., Nagai, T., 1990. Percutaneous absorption enhancing effect and skin irritation of monocyclic monoterpenes. Drug Des. Delivery 6, 229-238.

Opdam, J.J., 1991. Linear systems dynamics in toxicokinetic studies. Ann. Occup. Hyg. 35, 633-649.

Otzen, D.E., Sehgal, P., Westh, P., 2009. α-Lactalbumin is unfolded by all classes of detergents but with different mechanisms, J. Coll. Int. Sci. 329, 273–283.

Parisis, S.A., Maniati, M.A., Kyriakidis, V., Constantopoulos, S.H., 1995. Pulmonary damage due to paraquat poisoning through skin absorption. Respiration 62, 101-103.

Patil, S., Singh, P., Szolar-Platzer, C., Maibach, H.I., 1996. Epidermal enzymes as penetration enhancers in transdermal drug delivery. J. Pharm. Sci. 85, 249-252.

Pedretti, A., Villa, L., Vistoli, G., 2002. Modeling of binding modes and inhibition mechanism of some natural ligands of farnesyl transferase using molecular docking. J. Med. Chem. 45, 1460-1465.

Pendlington R.U., Sanders, D.J., Bourner, C.B., Saunders, D.R., Peace, C.K., 2004. Development of a repeat dose in vitro skin penetration model. In: Brain, K.R., Walters, K.A. (Eds.). Perspectives in Percutaneous Penetration vol. 9a. Abstracts of presentations at the ninth international perspectives in percutaneous penetration conference held in La Grande Motte, April 2004. Cardiff, STS Publishing, 79-92.

Potts, R.O., 1989. Physical characterization of the stratum corneum: the relationship of echanical and barrier properties to lipid and protein structure. In: Hadgraft, J., Guy, R.H., (Eds.). Transdermal Drug Delivery. Marcel Dekker, New York, 23-58.

Potts, R.O., Guy, R.H., 1992. Predicting skin permeability. Pharm. Res. 9, 663-669.

Potts, R.O., Guy, R.H., 1995. A predictive algorithm for skin permeability – The effects of molecular size and hydrogen bond activity. Pharm. Res. 12, 1628-1633.

PubChem Home Page. http://pubchem.ncbi.nlm.nih.gov/ (Accessed Sep 2010).

Pugh, W.J., Roberts, M.S., Hadgraft, J., 1996. Epidermal permeability – penetrant structure relationships: 3. The effect of hydrogen bonding interactions and molecular size on diffusion across the stratum corneum. Int. J. Pharm. 138, 149-165.

Pugh, W.J., Wong, F.F., Michniak, B.B., Moss, G.P., 2005. Discriminant analysis as a tool to identify compounds with potential as transdermal enhancers. J. Pharm. Pharmacol. 57, 1389-1396.

Qiao, G.L., Brooks, J.D., Baynes, R.E., Monteiro-Riviere, N.A., Williams, P.L., Riviere, J.E., 1996. The use of mechanistically defined chemical mixtures (MDCM) to assess component effects on the percutaneous absorption and cutaneous disposition of topically-exposed chemicals: I. Studies with parathion mixtures in isolated perfused porcine skin. Toxicol. Appl. Pharmacol. 141, 473– 486.

Rajadhyaksha, M., Gonzalez, S., Zavislan, J.M., Anderson, R.R., Webb, R.H., 1999. In vivo confocal scanning laser microscopy of human skin II: Advances in instrumentation and comparison with histology. J. Invest. Dermatol. 113, 293-303.

Rawlings, A.V., Matts, P.J., 2005. Stratum corneum moisturization at the molecular level: an update in relation to the dry skin cycle. J. Invest. Dermatol. 124, 1099-1110.

Raykar, P.V., Fung, M.C., Anderson, B.D., 1988. The role of protein and lipid domains in the uptake of solutes by human stratum corneum. Pharm. Res., 5, 140–150.

Reddy, M.B., Stinchcomb, A.L., Guy, R.H., Bunge, A.L., 2002. Determining dermal absorption parameters in vivo from tape stripping. Pharm. Res. 19, 292-298.

Rerek, M.E., Wyck, D.V., Mendelsohn, R., Moore, D.J., 2005. 'FTIR spectroscopic studies of lipid dynamics in phytosphingosine ceramide models of the stratum corneum lipid matrix.' Chem. Phys. Lipids, 134, 51-58.

Ritshel, W.A., Sabouni, A., Hussain, A.S., 1989. Percutaneous absorption of coumarin, griseofulvin and propranolo across human scalp and abdominal skin. Methods Find. Exp. Clin. Pharmacol. 11, 643-646.

Riviere, J.E., Baynes, R.E., Brooks, J.D., Yeatts, J.L., Monteiro-Riviere, N.A., 2003. Percutaneous absorption of topical diethyl-m-toluamide(-DEET): effects of exposure variables and coadministered toxicants. J. Toxicol. Environ. Health A 66, 131–151.

Riviere, J.E., Brooks, J.D., 2005. Predicting skin permeability from complex chemical mixtures. Toxicol. Appl. Pharmacol. 208, 99-110.

Riviere, J.E., Brooks, J.D., 2007. Prediction of dermal absorption from complex chemical mixtures: incorporation of vehicle effects and interactions into a QSPR framework. SAR QSAR Environ. Res., 18, 31-44.

Riviere, J.E., Brooks, J.D. 2010. Chemical mixture absorption dataset, http://www.lib.ncsu.edu/resolver/1840.2/2297

Riviere, J.E., Brooks, J.D., 2011. Predicting skin permeability from complex chemical mixtures: Dependency of quantuitative structure permeability relationshiops (QPSR) on biology of skin model used. Toxicol. Sci. 119, 224-232.

Riviere, J.E., Monteiro-Riviere, N.A., 2002. Gulf War related exposure factors influencing topical absorption of ^{14}C permethrin. Toxicol. Lett. 135, 61– 71.

Riviere, J.E., Papich, M.G., 2001. 'Potential and problems of developing transdermal patches for veterinary applications.' Adv. Drug Delivery Rev. 50, 175-203.

Riviere, J.E., Qiao, G.L., Baynes, R.E., Brooks, J.D., Mumtaz, M., 2001. Mixture component effects on the in vitro dermal absorption of pentachlorophenol. Arch. Toxicol. 75, 329– 334.

Roberts, D.W., Patlewicz, G., 2002. Mechanism based structure-activity relationships for skin sensitisation – The carbonyl group domain. SAR QSAR Environ. Res. 13, 145-152.

Roberts, M.S., Cross, S.E., Pellett, M.A., 2002. Skin transport. In: Walters, K.A., (Ed.). Dermatological and transdermal formulations. New York, Marcel Dekker, 89-195.

Roberts, M.S., Walker, M., 1993. Water – the most natural penetration enhancer. In: Walters K. A., Hadgraft J., (Eds.). Pharmaceutical skin penetration enhancement. New York, Marcel Dekker. 1-30.

Rolf, D., 1988. Chemical and physical methods of enhancing transdermal drug delivery, Pharm. Technol. 12, 131-140.

Rosado, C., Cross, S.E., Pugh, W.J., Roberts, M.S., Hadgraft, J., 2003. Effect of vehicle pretreatment on the flux, retention, and diffusion of topically applied penetrants in vitro. Pharm. Res. 20, 1502– 1507.

Rosen, M. 2005. Technology, Applications and Formulations (Personal Care and Cosmetic Technology). In Rosen, M., (Ed.). Delivery System Handbook for Personal Care and Cosmetic Products. New York, USA: William Andrew, Inc.

Roskos, K.V., Maibach, H.I., Guy, R.H., 1989. The effect of aging on percutaneous absorption in man. J. Pharmacokinet. Biopharm. 17, 617-630.

Rowe, R.C., Sheskey, P.J., Quinn, M.E., 2009. Handbook of pharmaceutical excipients. 6th edition. Pharmaceutical press.

Roy, K., Sanyal, I., Roy, P.P., 2006. QSPR of the bioconcentration factors of non-ionic organic compounds in fish using extended topochemical atom (ETA) indices. SAR QSAR Environ. Res. 17, 563-582.

Santos, P., Watkinson, A.C., Hadgraft, J., Lane, M.E., 2010. Oxybutynin permeation in skin: The influence of drug and solvent activity. Int. J. Pharm. 384, 67-72.

Sartorelli, P., Andersen, H.R., Angerer, J., Corish, J., Drexler, H., Goen, T., Griffin, P., Hotchkiss, S.A.M., Larese, F., Montomoli, L., Perkins, J., Schmeiz, M., van de Sandt, J., Williams, F., 2000. Percutaneous penetration studies for risk assessment. Environ. Toxicol. Pharmacol. 8, 133-152.

Sartorelli, P., Aprea, C., Bussani, R., Novelli, M.T., Orsi, D., Sciarra, G., 1997. In vitro percutaneous penetration of methyl-parathion from a commercial formulation throiugh the human skin. Occup. Environ. Med. 54, 524-525.

SCCNFP, 2003. Basic criteria for the in vitro assessment of dermal absorption of cosmetic ingredients, updated October 2003, adopted by the SCCNFP during the 25th plenary meeting of 20 October 2003. Scientific Committee on cosmetic products and non-food products intended for consumers, pp 1-9 (SCCNFP/0750/03).

Scheuplein, R.J., Blank, I.H., 1971. Permeability of the skin. Physiol. Rev. 51, 702-747.

Scheuplein, R.J., Blank, I.H., Brauner, G.J., MacFarlane, D.J., 1969. Percutaneous Absorption of Steroids. J. Invest. Dermatol. 52, 63-70.

Schmook, F.P., Meingassner, J.G., Billich, A., 2001. Comparison of human skin or epidermis models with human and animal skin in in-vitro percutaneous absorption. Int. J. Pharm. 215, 51-56.

Scott, R.C., Dugard, P.H., Ramsey, J.D., Rhodes, C., 1987. *In vitro* absorption of some o=phthalate diesters through human and rat skin. Environ. Health Perspect. 74, 223-227.

Shah, P. Vinod., 1994. Skin Penetration Enhancers: Scientific Perspectives. In Hsieh D.s., (Ed.). Drug Permeation Enhancement Theory and Applications, Center for Drug Evaluation and Research, Food and Drug Administration, Rockville, Maryland, Decker, Vol 62, 19-24.

Shokri, J., Nokhodchi, A., Dashbolaghi, A., Hassan-Zadeh, D., Ghafourian, T., Barzegar-Jalali, M., 2001. The effect of surfactants on the skin penetration of diazepam. Int. J. Pharm. 228, 99-107.

Sigma-Aldrich Home Page. http://www.sigmaaldrich.com/sigma-aldrich/home.html (Accessed Sep 2010).

Sinko, P.J., 2011. Martin's Physical Pharmacy and Pharmaceutical Sciences, 6th edition, Lippincott Williams & Wilkins publications 226.

Skelly, J.P., Shan, V.P., Guy, R.H., Wester, R.C., Flynn, G., Yacobi, A., 1987. FDA and AAPS report of the workshop on principles and practises of in vitro percutaneous penetration studies: relevance to bioavailability and bioequivalence. Pharm. Res. 4, 265-267.

Sloan, K.B., Koch, S.A., Siver, K.G., Flowers, F.P., 1986. Use of solubility parameters of drug and vehicle to predict flux through skin. J. Invest. Dermatol. 87, 244–252.

Slovokhotov, Y.L., Batsanov, A.S., Howard, J.A.K., 2007. Molecular van der Waals symmetry affecting bulk properties of condensed phases: melting and boiling points. Struct. Chem. 18, 477-491.

Soyei, S., Williams, F., 2004. A database of percutaneous absorption, distribution and physicochemical parameters. In: Brain, K.R., Walters, K.A. (Eds.). Perspectives in percutaneous penetration. Vol. 9a. Abstracts of presentations at the ninth international perspectives in percutaneous penetration conference held in La Grante Motte, April 2004. Cardiff, STS Publishing, 84.

StatSoft, Inc. (2010). STATISTICA (data analysis software system), version 9.1. www.statsoft.com.

Steiling, W., Kreutz, J., Hofer, H., 2001. Percutaneous penetration/dermal absorption of hair dyes in vitro. Toxicol. in Vitro 15, 565-570.

Stinecipher, J., Shah, J., 1997. Percutaneous permeation of N,N-diethyl-m-toluamide (DEET) from commercial mosquito repellents and the effect of solvent. J. Toxicol. Environ. Health, 52, 119-135.

Stott, W., Williams, A.C., Barry, B.W., 1997. Transdermal delivery from eutectic systems: Enhanced permeation of a model drug, ibuprofen. J. Controlled Release 50, 297–308.

Stoughton, R.B., 1982. In: Farber, E.H., (Ed.). Psoriasis, Grune and Stration, New York, 346-398.

Surber, C., Wilhelm, K.P., Maibach, H.I., 1991. In-Vitro Skin Pharmacokinetic Of Acitretin: Percutaneous Absorption Studies In Intact And Modified Skin From Three Different Species Using Different Receptor Solutions. J. Pharm. Pharmacol. 43, 836-840.

Syracuse Research Corporation, 2009. SRC PhysProp database, http://www.syrres.com/what-we-do/databaseforms.aspx?id=386, Accessed September 2009.

Syracuse Research Corporation, 2009. SRC PhysProp database, http://www.syrres.com/what-we-do/databaseforms.aspx?id=386, retrieved November 2010 – February 2011.

Sznitowska, M., Janicki, S., Baczek, A., 2001. Studies on the effect of pH on the lipoid al route of penetration across stratum corneum. J. Controlled Release 76, 327-335.

Tang, H., Blankschtein, D., Langer, R., 2002. Prediction of steady-state skin permeabilities of polar and nonpolar permeants across excised pig skin based on measurements of transient diffusion: Characterization of hydration effects on the skin porous pathway. J. Pharm. Sci. 91, 1891-1907.

Tanojo, H., Bouwstra, J.A., Junginger, H.E., Bodde, H.E., 1997. In vitro human skin barrier modulations by fatty acids: skin permeation and thermal analysis studies. Pharm. Res. 14, 42-49.

Tayar, N.El., Tsai, R.S., Testa, B., Carrupt, P.A., Hansch, C., Leo, A., 1991. Percutaneous penetration of drugs – A quantitative structure permeability relationship study. J. Pharm. Sci. 80, 744-749.

Thakur, A., Thakur, M., Khadikar, P.V., Supuran, C.T., Sudele, P., 2004. QSAR study on benzenesulphonamide carbonic anhydrase inhibitors: topological approach using Balaban index. Bioorg. Med. Chem. 12, 789–793.

Todeschini, R., Consonni, V., 2000. Handbook of molecular descriptors. Wiley-VCH

Trauer, S., Patzelt, A., Otberg, N., Knorr, F., Rozycki, C., Balizs, G., Buttemeyer, R., Linscheid, M., Liebsch, M., Lademann, J., 2009. Permeation of topically applied caffeine through human skin – a comparison of in vivo and in vitro data. Br. J. Clin. Pharmacol. 68, 181-186.

TSAR 3D, Release 3.3, 2008. Accelrys Software Inc., San Diego, http://accelrys.com.

Tuppurainen, K., Korhonen, S.P., Ruuskanen, J., 2006. Performance of multi component self-organizing regression (MCSOR) in QSAR, QSPR, and multivariate calibration: comparison with partial least-squares (PLS) and validation with large external data sets. SAR QSAR Environ. Res. 17, 549-561.

Tur, E., Maibach, H., Guy, R.H., 1991. Percutaneous penetration of methyl nicotinate at 3 anatomic sites – evidence for an appendageal contribution to transport. Skin Pharmacol., 4, 230-234.

USEPA, 1998. Health effects test quidelines. OPPTS 870.7600. Dermal penetration. Washington, DC, United States Environmental Protection Agency, Office of Prevention, Pesticides and Toxic Substances, 1-12 (http://www.epa.gov/opptsfrs/publications/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Series/870-7600.pdf).

Vaddi, H.K., Ho, P.H., Chan, Y.W., Chan, S.Y., 2002. Terpenes in ethanol: haloperidol permeation and partition through human skin and stratum corneum changes. J. Controlled Release 81, 121-133.

van der Merwe, D., Riviere, J.E., 2005. Comparative studies on the effects of water, ethanol and water/ethanol mixtures on chemical partitioning into porcine stratum corneum and silastic membrane. Toxicol. in Vitro, 19, 69-77.

van de Sandt, J.J.M., van Burgsteden, J.A., Carmichael, P.L., Dick, I., Kenyon, S., Korinth, G., Larese, F., Limasset, J.C., Maas, W.J.M., Montomoli, L., Nielsen, J.B., Payan, J.P., Robinson, E., Sartorelli, P., Schaller, K.H., Wilkinson, S.C., Williams, F.M., 2004. In vitro predictions of skin absorption of caffeine, testosterone, and benzoic acid: a multi-centre comparison study. Regul. Toxicol. Pharmacol. 39, 271-281.

van de Sandt, J.J.M., Rutten, A.A.J.J.L., Van Ommen, B., 1993. Species-Specific Cutaneous Biotransformation of the Pesticide Propoxure during Percutaneous Absorption in Vitro. Toxicol. Appl. Pharmacol. 123, 144-150.

Vapnik, V., 1998. Statistical learning theory. Wiley, NewYork, 1998.

Vapnik, V., 1999. The support Vector method of function estimations. U.S. Patent 5,950,146.

Vickers, A.E.M., Sinclair, J.R., Zollinger, M., Heitz, F., Glänzel, U., Johanson, L., Fischer, V., 1999. Multiple cytochrome p-450s involved in the metabolism of terbinafine suggest a limited potential for drug-drug interactions, Drug Metab. Dispos. 27, 1029-1038.

Wagner, H., Kostka, K.H., Lehr, C.M., Schaefer, U.F., 2000. Drug distribution in human skin using two different in vitro test systems: comparison with in vivo data. Pharm. Res. 17, 1475–1481.

Walters, K. A., 1989. Penetration enhancers and their use in transdermal therapeutic systems, in Transdermal Drug Delivery. In Hadgraft, J., Guy, R.H., (Eds.). Marcel Dekker, New York, 197-246.

Watkinson, R.M., Herkenne, C., Guy, R.H., Hadgraft, J., Oliveira, G., Lane, M.E., 2009. Influence of ethanol on the solubility, ionization and permeation characteristics of ibuprofen in silicon and human skin. Skin Pharmacol. Physiol. 22, 15-21.

Whitehead, T., 1975. Letter: long-acting phenothiazines, Br. Med. J. 2, 502.

Weaver, S., Gleeson, M.P., 2008. The importance of the domain of applicability in QSAR modeling. J. Mol. Graph. Model., 26, 1315–1326.

Wegener, M., Neubert, R., Rettig, W., Wartewig, S., 1997. Structure of stratum corneum lipids characterized by FT-Raman spectroscopy and DSC. III. Mixtures of ceramides and cholesterol. Chem. Phys. Lipids 88, 73-82.

Wester, R.C., Maibach, H.I., Melendres, J., Sedik, L., Knaak, J., Wang, R., 1992. In vivo and In vitro percutaneous absorption and skin evaporation of isofenphos in man. Fundam. Appl. Toxicol. 19, 521-526.

Wester, R.M., Tanojo, H., Maibach, H.I., Wester, R.C., 2000. Predicted chemical warfare agent VX toxicity to uniformed soldier using parathion in vitro human skin exposure and absorption, Toxicol. Appl. Pharmacol. 168, 149–152.

Wiechers, J. W., 2005. Optimizing Skin Delivery of Active Ingredients From Emulsions. In: Rosen M. R., (Ed.). Delivery system handbook for personal care and cosmetic products. William Andrew, 410-433.

Wilkinson, S.C., Mass, W.J.M., Nielsen, J.B., Greaves, L.C., van de Sandtt J.J.M., Williams F.M., 2004. Influence of skin thickness on percutaneous penetration in vitro. In: Brain K.R., Walters K.A., (Eds.). Perspectives in percutaneous penetration. Vol. 9a. Abstracts of presentations at the ninth international perspectives in percutaneous penetration conference held in La Grande Motte, Appril 2004. Cardiff, STS Publishing, 83.

Wilkinson, S.C., Mass, W.J., Nielsen, J.B., Greaves, L.C., van de Sandt J.J., Williams F.M., 2006. Interactions of skin of skin thickness and physicochemical properties of test compounds in percutaneous penetration studies. Int. Arch. Occup. Environ. Health 79, 405-413.

Wilkinson S.C., Williams, F.M., 2002. Effects of Experimental Conditions on Absorption of Glycol Ethers Through Human Skin In Vitro. Int. Arch. Occup. Environ. Health 75, 519-527.

Williams, A .C, Barry, B.W., 1991. Terpenes and the lipid-protein-partitioning theory of skin penetration enhancement. Pharm. Res. 8, 17-24.

Williams, A.C., Barry, B.W., 2004. Penetration enhancers. Adv. Drug Delivery Rev. 56, 603-618.

Williams, F.M., 2006. In vitro studies – how good are they at replacing in vivo studies for measurement of skin absorption? Environ. Toxicol. Pharmacol. 21, 199-203.

Wilschut, A., Berge, W.F., Robinson, P.J., McKone, T.E., 1995. Estimating skin permeation. The validation of five mathematical skin permeation models. Chemosphere, 30, 1275-1296.

Xia, X.R., Baynes, N.A., Monteiro-Riviere, N.A., Riviere, J.E., 2007. A system coefficient approach for quantitative assessment of the solvent effects on membrane absorption from chemical mixtures. SAR and QSAR Environ. Res. 18, 579-593.

Yao, X.J., Panaye, A., Doucet, J.P., Zhang, R.S., Chen, H.F., Liu, M.C., Hu, Z.D., Fan, B.T., 2004. J. Comparative study of QSAR/QSPR correlations using support vector machines, radial basis function neural networks and multiple linear regression. Chem. Inf. Comput. Sci. 44, 1257-1266.

Zhai, H., Maibach, H.I., (2001). Effects of skin occlusion on percutaneous absorption: an overview. Skin Pharmacol. Appl. Skin Physiol. 14, 1-10.

Zhang, Q., Grice, J.E., Wang, G., Roberts, M.S., 2009. Cutaneous metabolism in transdermal drug delivery, Curr. Drug Metab. 10, 227-235.

Zefirov, N.S., Palyulin, V.A., 2001. QSAR for boiling points of 'small' sulphides. Are the 'high-quality structure-property-activity regressions' the real high quality QSAR models? J. Chem. Inf. Comput. Sci. 41, 1022-1027.

Zendian, R.P., 2000. Dermal absorption of pesticides in the rat. Am. Ind. Hyg. Assoc. J. 61, 473-483.

Zinke, S., Gerner I., and Schlede, E., 2002. Evaluation of a rule base for identifying contact allergens by using a regulatory database: Comparison of data on chemicals notified in the European Union with 'structural alerts' used in the DEREK expert system. ATLA, Altern. Lab. Anim. 30, 285-298.

Appendix I. (Chapter 4)

Table of Observed and Predicted values of log k_p using selected models (4.2), (4.3) and (4.6)

Terpene	Observed log k_p	Calculated log k_p eq. 4.2	Calculated log k_p eq. 4.3	Calculated log k_p eq. 4.6
(±)-linalool	-8.97	-8.2507	-7.41963	-9.06905
(+)-dihydrocarveol	-8.71	-8.4224	-8.31341	-7.26405
Carvacrol	-8.44	-7.4005	-7.59137	-7.17269
Cyclohexanemethanol	-8.08	-8.324	-8.10577	-6.88104
Eucarvone	-7.6	-7.8927	-7.54669	-7.96921
longifolene	-7.42	-6.1803	-6.45695	-6.83037
(+)-dihydrocarvone	-7.17	-8.3596	-8.31665	-7.20122
(-)-caryophyllene oxide	-6.78	-5.7301	-6.15407	-6.56110
(S)-(-)-perillaldehyde	-6.59	-5.7767	-5.90326	-7.28409
Mycrene	-5.43	-5.9501	-6.07925	-5.58414
Octisalate	-5.19	-5.5916	-5.10692	-6.40621
α-phellandrene	-4.96	-6.0948	-6.46148	-5.61650
Retinoic acid	-12.13	-10.2614	-9.99137	-8.86625
β-carotene	-11.15	-9.6582	-9.97142	-9.60389
(-)-guaiol	-8.88	-6.5885	-6.53369	-7.10427
(-)-dihydrocarveol	-8.87	-8.423	-8.31393	-7.21064
Menthone	-8.72	-7.8591	-7.80519	-7.17515
squalene	-8.56	-9.5242	-9.48878	-7.01425
isolongifolol	-8.55	*	*	-8.41447
(-)-α-thujone	-8.52	-7.6361	-7.67455	-8.40008
(-)-isopulegol	-8.35	-7.4614	-7.41958	-7.13739
Thymol	-8.29	-7.0478	-7.29079	-7.25570
(1R)-(-)-myrtenol	-8.28	-7.2154	-7.07304	-7.27875
(+)-cedrol	-7.96	-6.9266	-6.67125	-8.26703
Nerol	-7.8	-7.0573	-6.85331	-6.52506
β-citronellol	-7.66	-7.1314	-6.84855	-6.45056
(-)-α-santonin	-7.58	-8.0124	-8.14328	-7.21100
(R)-(-)-carvone	-7.56	-8.2472	-8.3026	-7.12701
Geraniol	-7.43	-7.1316	-6.91752	-6.71124
(+)-aromadendrene	-7.4	-7.6803	-8.08491	-7.14859
L-(-)-menthol	-7.34	-7.709	-7.64647	-7.01487
(-)-trans-caryophyllene	-7.28	-6.7815	-7.09781	-6.56528
(+)-β-cedrene	-7.01	-7.5759	-7.84226	-7.52366
Terpinolene	-7.01	-6.2493	-6.63375	-5.46699
(-)-α-cedrene	-6.89	-7.3558	-7.65221	-7.20580
Farnesol	-6.72	-6.5146	-6.40013	-6.57075
Retinol	-6.71	-9.1062	-8.88538	-8.25553
(R)-(+)-pulegone	-6.63	-6.9729	-7.16196	-6.58421
(-)-carveol	-6.45	-8.4006	-8.38836	-7.14288
α-humulene	-6.23	-7.3645	-7.31997	-6.53951
(+)-cedryl acetate	-5.52	-5.0876	-4.95672	-7.06468
Ocimene	-5.41	-5.4926	-5.71284	-5.64836
(1R)-(-)-myrtenal	-5.29	-5.5196	-5.53068	-6.83553
(±)-α-bisabolol	-5.25	-6.477	-6.47828	-7.22693
Phytol	-5.13	-5.6963	-5.55615	-6.84787
Citral	-5.08	-4.8277	-4.87863	-6.016113
(S)-(-)-citronellal	-4.83	-5.1049	-4.99207	-5.90627
(±)-nerolidol	-4.59	-7.0292	-6.46598	-6.71934
(-)-epiglobulol	-4.41	-6.0449	-6.13801	-6.75915

Appendix II. Training Set Models (Chapter 7)

Models obtained for the training set have been presented here. These models were used for the estimation of log flux for the test set and the Mean Absolute Error (MAE) of estimation for test set was calculated (Table 5 in the paper).

Model based on equation (7.5)

Log flux = - 1.35 + 0.000001 [donor] -0.00489 Weight + 0.0140 SlogP_VSA4 - 1.73 fiAB + 0.00254 BP-MP(mix) + 3.69 vsurf_G - 0.435 VAdjMa

N=342, S = 0.962649, r^2 = 0.544, F = 56.8, P = 0.000

Model based on equation (7.6)

Log flux = - 1.09 + 0.000001 [donor] - 0.00478 Weight + 0.0142 SlogP_VSA4 - 1.74 fiAB + 0.00197 BP-MP(mix) + 3.54 vsurf_G - 0.391 VAdjMa - 0.228 Thickness

N = 342, S = 0.955921, r^2 = 0.551, F = 51.2, P = 0.000

Model based on equation (7.7)

Log flux = - 0.52 + 0.000001 [donor] - 0.00510 Weight + 0.0107 SlogP_VSA4 - 1.72 fiAB + 0.00239 BP-MP(mix) + 3.26 vsurf_G - 0.382 VAdjMa - 0.358 Infinite/Finite

N = 341, S = 0.953, r^2 = 0.555, F=51.8, P = 0.000

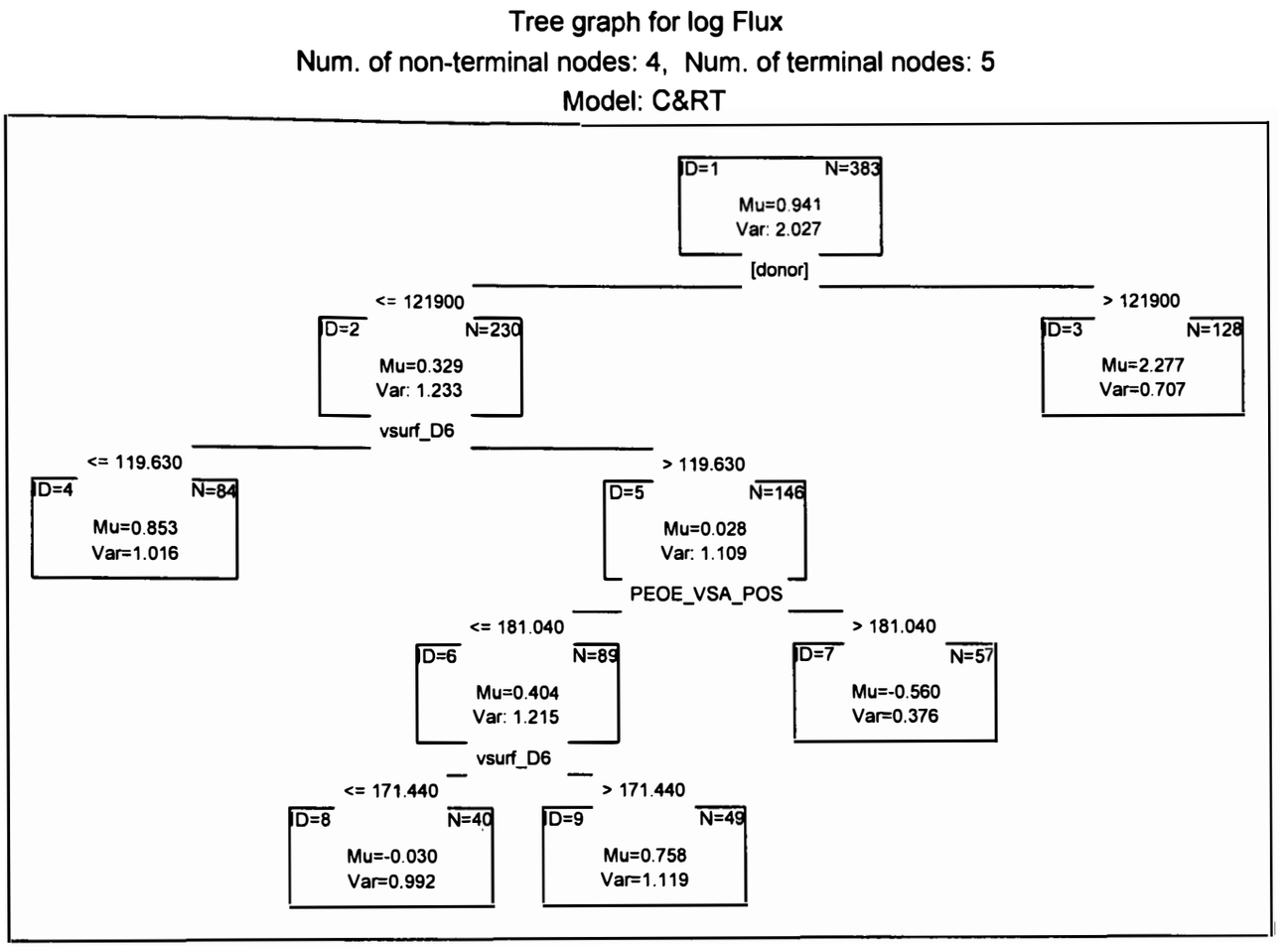


Figure 1. Model based on C&RT (7.1)

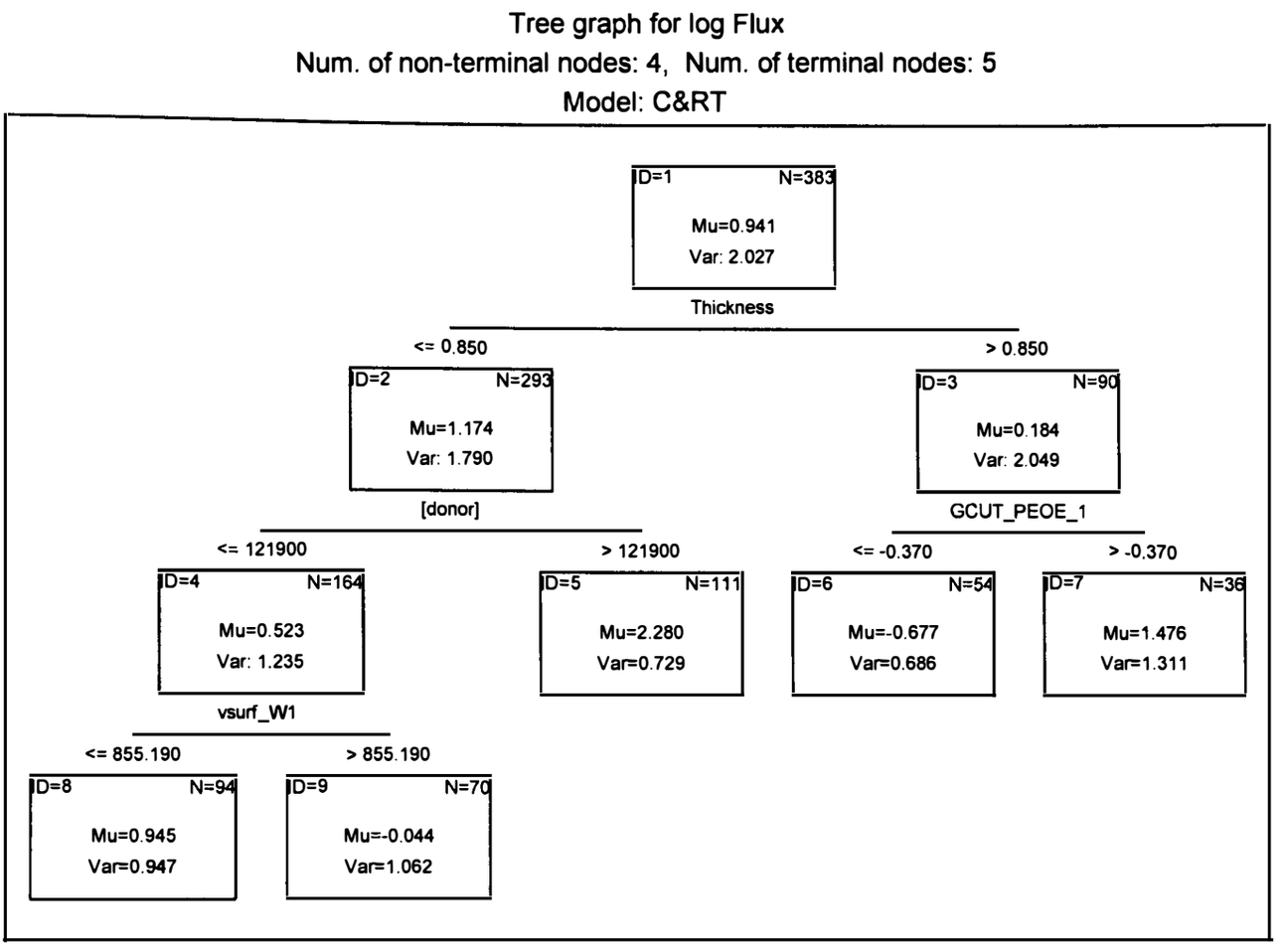


Figure 2. Model based on C&RT (7.2)

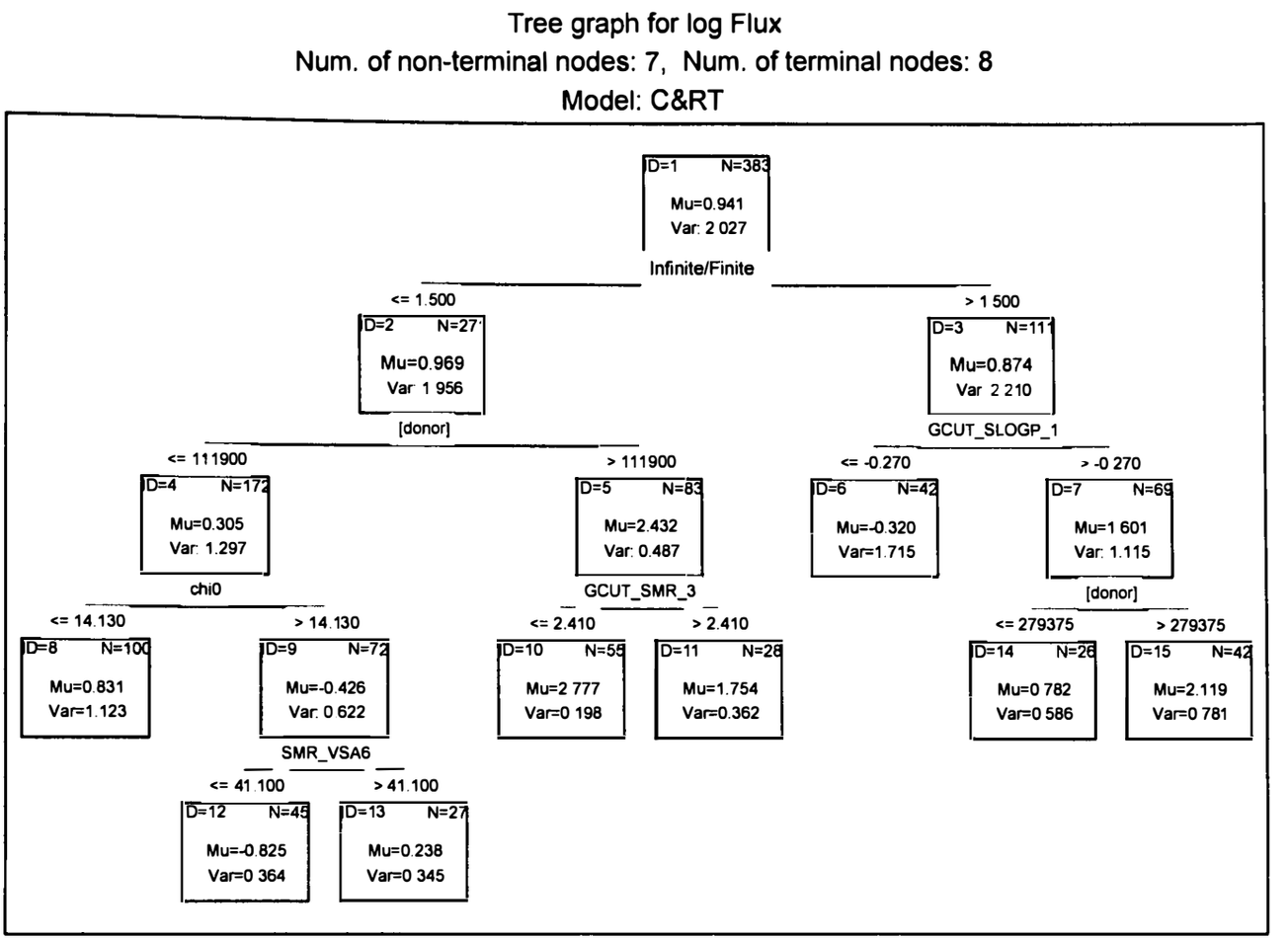


Figure 3. Model based on C&RT (7.3)

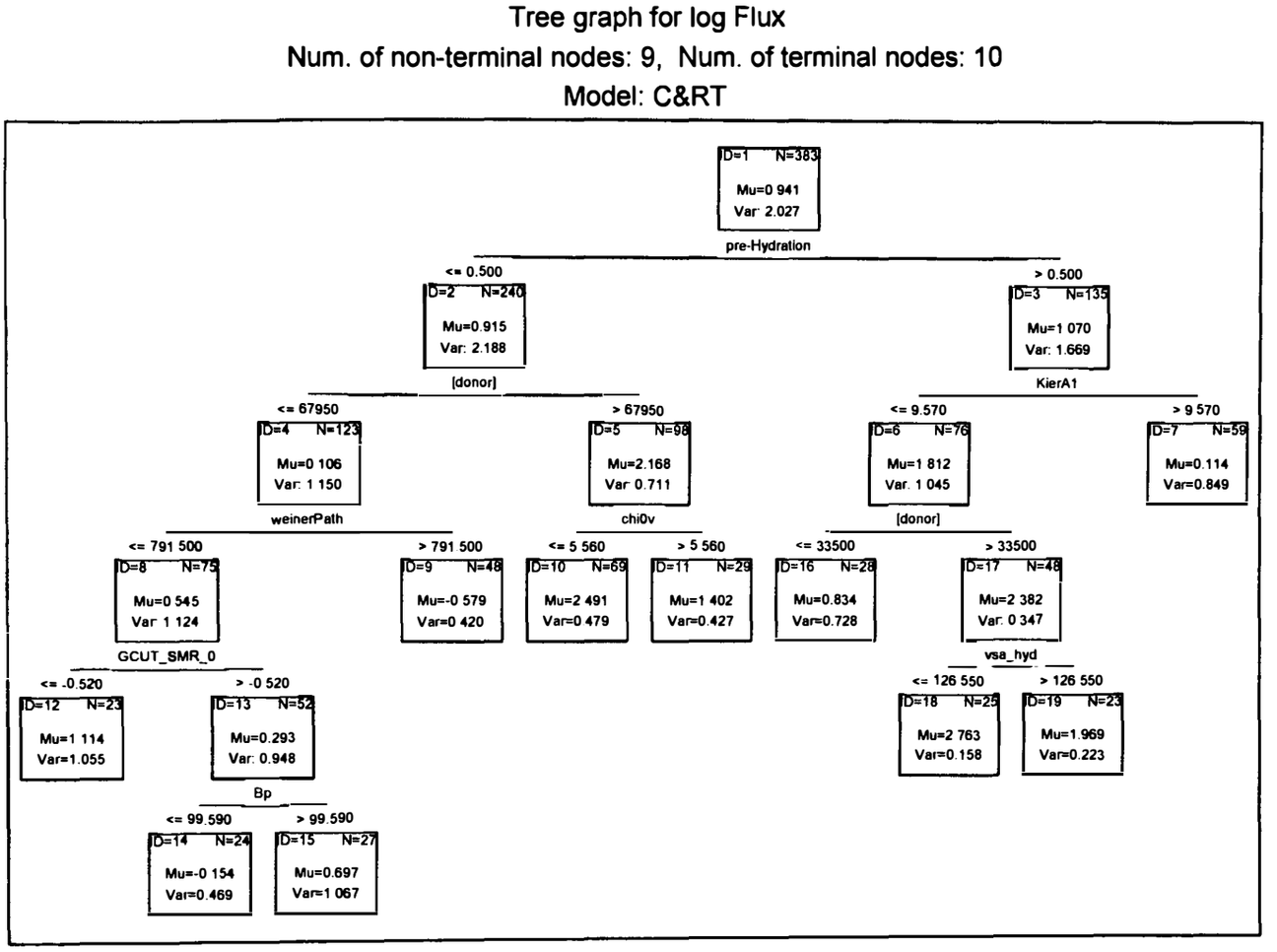


Figure 4. Model based on C&RT (7.4)

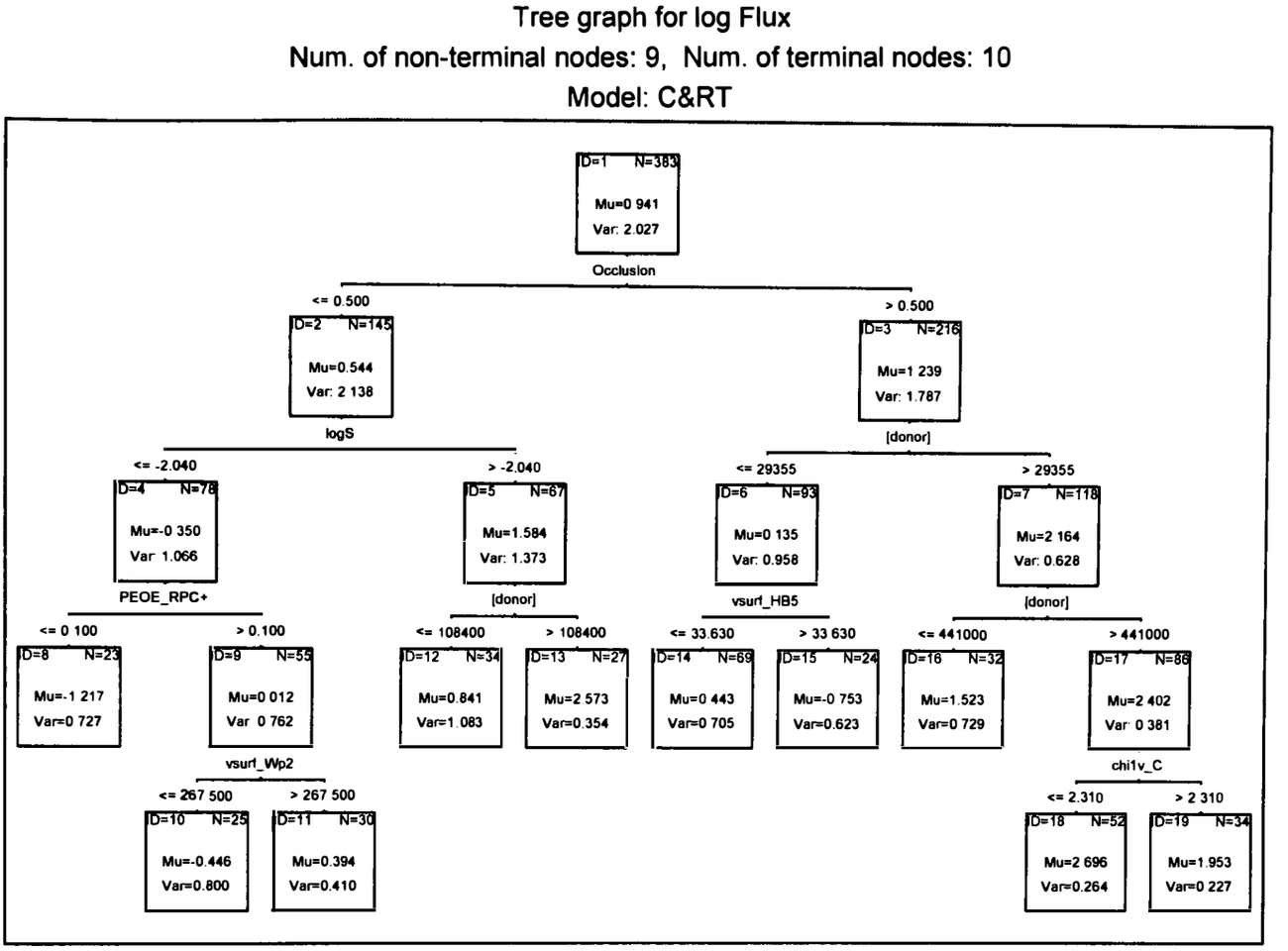


Figure 5. Model based on C&RT (7.5)

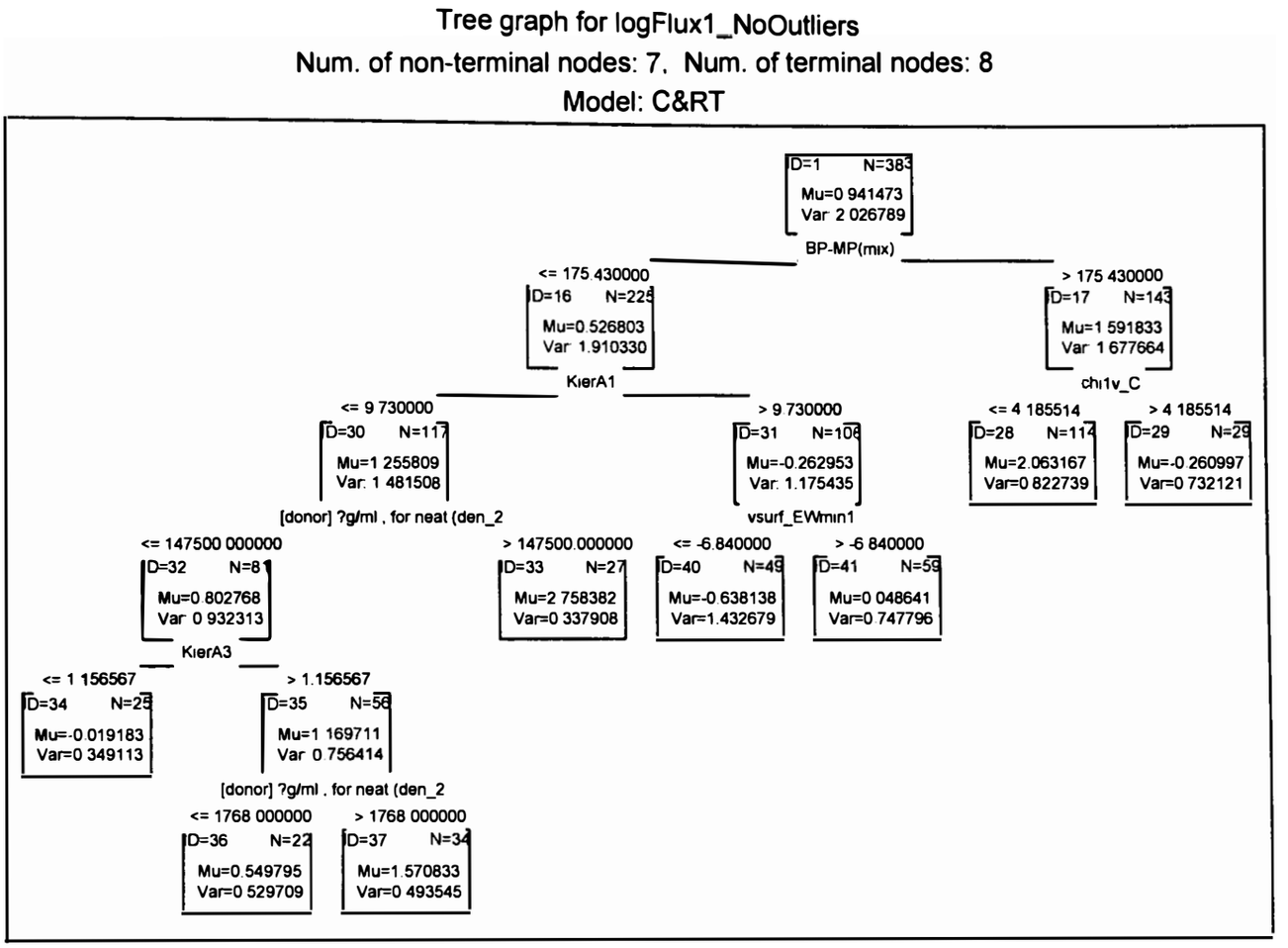


Figure 6. Model based on C&RT (7.6)

Appendix III. Posters

Modelling the Effect of Solvents on the Skin Penetration of Formoterol Fumarate

Eleftherios Samaras and Taravat Ghafourian
Medway School of Pharmacy, Universities of Kent and Greenwich
Central Ave., Chatham, Kent ME4 4TB, England

Abstract

A vehicle influences the concentration of penetrant within the membrane, affecting its diffusivity in the skin and rate of transport. Despite the huge amount of effort made for the understanding and modelling the skin absorption of chemicals, a reliable estimation of the skin penetration potential from formulations remains a challenging objective.

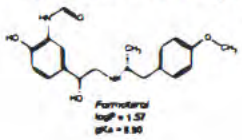
In this investigation, Quantitative Structure – Permeability Relationships (QSPeR) were employed to relate the effects of solvents on the skin permeation of formoterol fumarate to the molecular descriptors of the solvents. The permeability data consisted of the cumulative amount of permeated drug across excised rat skin at various time intervals and the corresponding flux. Donor phase was the drug dissolved in the different solvents including terpenes, alcohols, and fatty acids. The structural descriptors were calculated using specialised molecular modelling and QSAR software. Stepwise regression analysis was used to select the statistically significant descriptors and generate the QSPeR. The results indicated a negative effect of molecular size and aromatic rings of the solvents on the penetration rate of formoterol fumarate.

Introduction

The success of a transdermal drug formulation depends on the ability of drug to penetrate the skin in sufficient quantities in order to achieve the desired therapeutic effect. From a drug delivery perspective, penetration of the drug depends not only on the nature of the drug but also on the nature of the other ingredients present in the mixture (formulation). As a result, desirable penetration rates could be achieved by altering the formulation of a drug. However, the data used in most prediction models are based on experimental data for individual chemicals because very limited experimental data are available for chemical mixtures. While it is frequently difficult to assess the absorption of individual chemicals, it is challenging to quantitatively assess the absorption from chemical mixtures.

Aims

The aim of this investigation was to develop QSAR models for the effect of solvents on the skin penetration behaviour of a model penetrant (formoterol fumarate). The model will help identify the mechanisms involved in the penetration through skin and the effect of formulation factors.



Methods

The dataset:

The dataset consisted of the cumulative amount of formoterol penetrated through abdominal rat skin at different time intervals of 2, 4, 6, 8 and 24 h as reported by Kakubari et al (2006). The flux was calculated from the original data. Donor phase was the drug dissolved in the thirty-eight different solvents including terpenes, alcohols, and fatty acids.

Structural descriptors:

Structural descriptors were calculated as follows:

- Electronic parameters such as HOMO and LUMO energies and dipole moment calculated by Vamp (using the AM1 Hamiltonian) in TSAR 3D software.
- Atom and group counts, molecular weight and surface area and volume calculated by TSAR 3D.
- Log P calculated by the ACD/labs logD Suite.

Development and Validation of QSARs:

Stepwise regression analysis was used to develop models for cumulative amount penetrated and the flux. The predictivity of the models were examined by a leave – many – out procedure.

References

Kakubari et al. (2006) Biol.Pharm. Bull., 29, 146 – 149.
Ghafourian and Fooladi (2000) Int. J. Pharm., 217, 1-11.

Results

In the QSARs below N is the number of data points, s is standard deviation and R^2 is the correlation coefficient.

QSAR 1:

$$\log Q_{24} = -0.937\,atoms - 2.51\,{}^{9}\chi^{v}_{p} - 1.72\,aromatics - 0.177\,E_{LUMO} + 1.58$$

$$N = 32 \quad S = 0.401 \quad R^2 = 0.611$$

Q_{24}: Cumulative amount of permeated formoterol within 24 h
Atoms: number of atoms
${}^{9}\chi^{v}_{p}$: 9th order valence – corrected path molecular connectivity index
Aromatics: number of aromatic rings
E_{LUMO}: energy of the lowest unoccupied molecular orbital

QSAR 2:

$$\log J = -2.87 - 9.96\,{}^{9}\chi_{p} + 0.0191\,MW + 0.0515\,lipole$$

$$n = 32 \quad s = 0.430 \quad R^2 = 0.654$$

J: Flux
${}^{9}\chi_{p}$:9th order path molecular connectivity index
MW: molecular weight
Lipole: lipole moment calculated by TSAR software

QSAR 3:

$$\log Q_{8} = -1.95 - 8.84\,{}^{9}\chi_{p} + 0.0171\,MW + 0.0604\,lipole + 0.327\,aliphatic\ rings$$

$$n = 32 \quad S = 0.440 \quad R^2 = 0.667$$

Q_8: Cumulative amount of permeated formoterol within 8 h
Aliphatic rings: number of 6-membered aliphatic rings

Discussion

The QSARs show that solvents with longer alkyl chains (those with higher ${}^{9}\chi_{p}$ or ${}^{9}\chi^{v}_{p}$ Values) lower the penetration rate of formoterol. Negative effect of molecular size can also be seen from the negative coefficient of 'atoms' in QSAR1. This might be due to a lower partition coefficient of formoterol between the stratum corneum and these solvents. Molecular weight (MW) of solvents has a positive effect on the skin penetration flux and cumulative amount of formoterol permeated in 8 h. However, this is true only if MW is accompanied by ${}^{9}\chi_{p}$ in the equation, otherwise the equation between log J and MW alone will have a negative slope. This shows that when MW is accompanied by ${}^{9}\chi_{p}$, it does not represent molecular size, but the presence of heavier heteroatoms in the solvent molecules, increasing the penetration of formoterol.

Table below shows the fold error of prediction of J using a leave-many-out procedure where 25% of chemicals were left out at a time and the J values were predicted using the descriptors in QSAR2. Average error of prediction was 2.8 fold.

The fold error of prediction for the skin penetration of formoterol dissolved in different solvents

Solvent	Fold error	Solvent	Fold error
Butyl myristate	1.51	Oleyl alcohol	1.43
Cetyl isooctanate	1.98	α-Terpinene	4.12
Cineole	3.89	Decyl oleate	3.58
Diisopropyl sebacate	1.08	Diisopropyl adipate	3.18
Ethyl Linoleate	3.38	dl-Menthol	4.07
N-Methyl-2-Pyrrolidone	1.50	Ethyl oleate	1.93
Oleic acid	7.53	Hexadecyl isostearate	3.24
Terpinolene	2.18	Isopropyl myristate	2.51
?-Terpineol	4.08	Oleyl oleate	1.60
Diethyl Phthalate	5.03	Diethyl sebacate	2.43
d-Limonene	1.04	Glyceryl triisooctanate	1.87
l-Decanol	2.40	Isopropyl palmitate	1.25
Linalool	1.01	l-Dodecanol	1.10
Liquid paraffin	9.84	Linalyl acetate	4.89
Menthone	2.91	l-Octanol	1.01
Octyldodecanol	1.08	n-Butyl alcohol	1.10

Conclusion

In conclusion skin penetration of drugs from different simple vehicle systems can be predicted using QSAR. However, complex mixture of solvents and formulation types will require a large volume of data for any estimation purposes.

Modelling the Effect of Solvents on the Skin Penetration of Formoterol Fumarate

Eleftherios Samaras and Taravat Ghafourian
Medway School of Pharmacy, Universities of Kent and Greenwich
Central Ave., Chatham, Kent ME4 4TB, England

Abstract

A vehicle influences the concentration of penetrant within the membrane, affecting its diffusivity in the skin and rate of transport. Despite the huge amount of effort made for the understanding and modelling the skin absorption of chemicals, a reliable estimation of the skin penetration potential from formulations remains a challenging objective. In this investigation, Quantitative Structure Activity Relationships (QSAR) were employed to relate the effects of different concentrations of solvents blended with various penetrants, upon skin absorption. The permeability data consisted from studies that took place using In-vitro porcine-skin flow through diffusion cells. Average of physicochemical properties for every solvent mixture were calculated. QSAR software was used fir the calculation of the structural descriptors. Stepwise regression analysis and regression analysis were used for modelling. The penetrant descriptors log P, the ninth order path connectivity index and the solvent predictor ,namely the difference between boiling and melting points, gave the best correlation and the most reliable QSAR

Introduction

The success of a transdermal drug formulation depends on the ability of drug to penetrate the skin in sufficient quantities in order to achieve the desired therapeutic effect. From a drug delivery perspective, penetration of the drug depends not only on the nature of the drug but also on the nature of the other ingredients present in the mixture (formulation). As a result, desirable penetration rates could be achieved by altering the formulation of a drug. However, the data used in most prediction models are based on experimental data for individual chemicals because very limited experimental data are available for chemical mixtures. While it is frequently difficult to assess the absorption of individual chemicals, it is challenging to quantitatively assess the absorption from chemical mixtures.

Aims

The aim of this investigation was to develop QSAR models for the effect of mixture components on skin absorption of penetrants. The model will help identify the mechanisms involved in the penetration through skin and the effect of formulation factors.

Methods

1. The dataset:

The permeability data consisted from studies that took place using In-vitro porcine-skin flow through diffusion cells. Absorption penetration data of 12 different penetrants (Table 1) each blended in 24 different solvent mixtures (Table 2), as reported by Riviere et al (2005).

Table1; Penetrants	
Atrazine	p-Chlorophenol
Chlorpyrifos	Phenol
Ethylparathion	p-Nitrophenol
Fenthion	Propazine
Methylparathion	Simazine
Nonylphenol	Triazine

Table2 Solvent Mixtures
Ethanol
Methylnicotinic acid
Propylene glycerol
Sodium lauryl sulfate
Water

2. Structural descriptors:

- The predictors (descriptors) of penetrants included connectivity indexes, quantum molecular descriptors, group counts (e.g. Vamp) were calculated in the TSAR 3D (Accelrys) software.
- The physico-chemical properties of mixture components (boiling point, melting point, solubility, vapor pressure and Henry's law constant) were obtained through Chemfinder online software and the Syracuse Research Corporation (SRC) website.
- Log P calculated by the ACD/labs logD Suite.

3. Development and Validation of QSARs:

Average of physicochemical properties for every solvent mixture were calculated, e.g. boiling point of the mixture. Stepwise regression analysis was used to develop models. The predictivity of the models were examined by a leave – many – out procedure.

Results

In the QSARs below N is the number of data points, s is standard deviation and R^2 is the correlation coefficient.

QSAR :

$$\text{Logkp} = -0.909 - 0.610\ \text{logP} + 2.62\ {}^{9}\chi_p - 0.00917\ (\text{SolBP-SolMP})$$

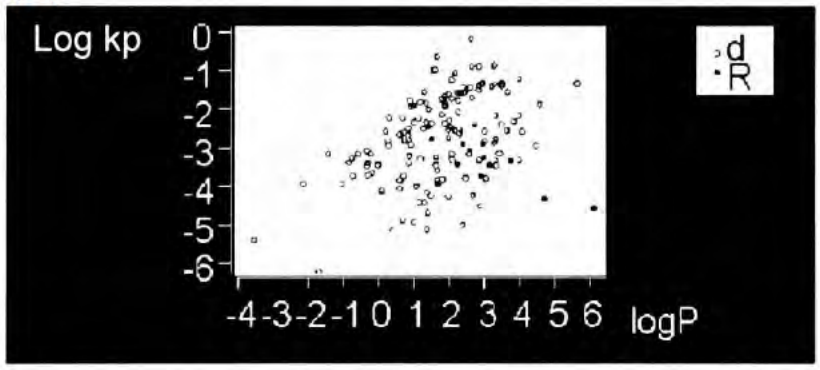
$$N = 288 \quad S = 0.4378 \quad r^2 = 0.729 \quad F = 255.20$$

Logp: Partition coefficient

${}^{9}\chi_p$:9th order path molecular connectivity index

SolBP-SolMP: the difference between the boiling point of the solvent and the melting point of the solvent

Diagram 1 below shows the relationship between log kp and liopophilicity of the skin absorption dataset of Wilchut et al (group d) and the chemicals used in this study (group R)



Discussion

↑logP → ↓logkp and ↑SolBP-SolMP → ↓logkp

When ↑logP → ↑lipophilicity → ↓trandermal absorption and vice versa

The negative relationship between log kp and log P could be due to the fact that most of the drugs in this particular dataset are more lipophilic than the compounds in the datasets normally used in QSAR studies of skin permeability.

Conclusion

In conclusion skin penetration of drugs from different simple vehicle systems can be predicted using QSAR. However, complex mixture of solvents and formulation types will require a large volume of data for any estimation purposes.

References

Riviere, J. E., and Brooks, J. D., 2005. Predicting the skin permeability from complex chemical mixtures. Toxicology and Applied Pharmacology 208, 99-100

Modelling the Effect of Mixture Components on the Skin Penetration

T. Ghafourian[1], E. Samaras[1], J. Riviere[2]
[1]: Medway School of Pharmacy, Universities of Kent and Greenwich, Chatham, Kent ME4 4TB, UK; [2]: Center for Chemical Toxicology Research and Pharmacokinetics, 4700 Hillsborough Street, North Carolina State University, Raleigh, USA

Abstract

A vehicle influences the concentration of penetrant within the membrane, affecting its diffusivity in the skin and rate of transport. Despite the huge amount of effort made for the understanding and modelling the skin absorption of chemicals, a reliable estimation of the skin penetration potential from formulations remains a challenging objective. In this investigation, Quantitative Structure Activity Relationships (QSAR) were employed to relate the effects of different concentrations of solvents blended with various penetrants, upon skin absorption. The permeability dataconsisted from studies that took place using In-vitro porcine-skin flow through diffusion cells. Average of physicochemical properties for every solvent mixture were calculated. QSAR software was used fir the calculation of the structural descriptors. Stepwise regression analysis and regression analysis were used for modelling. The penetrant descriptors log P, the ninth order path connectivity index and the solvent predictor,namely the difference between boiling and melting points, gave the best correlation and the most reliable QSAR

Introduction

The success of a transdermal drug formulation depends on the ability of drug to penetrate the skin in sufficient quantities in order to achieve the desired therapeutic effect. From a drug delivery perspective, penetration of the drug depends not only on the nature of the drug but also on the nature of the other ingredients present in the mixture (formulation). As a result, desirable penetration rates could be achieved by altering the formulation of a drug. However, the data used in most prediction models are based on experimental data for individual chemicals because very limited experimental data are available for chemical mixtures. While it is frequently difficult to assess the absorption of individual chemicals, it is challenging to quantitatively assess the absorption from chemical mixtures.

Aims

The aim of this investigation was to develop QSAR models for the effect of mixture components on skin absorption of penetrants. The model will help identify the mechanisms involved in the penetration through skin and the effect of formulation factors.

Methods

1. The dataset:

The permeability data was obtained from diffusion cell studies using porcine skin. Penetration data was available for 12 different penetrants (Table 1) each blended in 24 different solvent mixtures (Table 2).

Table1: Penetrants

Atrazine	p-Chlorophenol
Chlorpyrifos	Phenol
Ethylparathion	p-Nitrophenol
Fenthion	Propazine
Methylparathion	Simazine
Nonylphenol	Triazine

Table2: Solvent Mixtures

Ethanol
Methylnicotinic acid
Propylene glycerol
Sodium lauryl sulfate
Water

2. Molecular descriptors:

- The predictors (descriptors) of penetrants included connectivity indexes, quantum molecular descriptors, group counts (e.g. Vamp) were calculated in the TSAR 3D (Accelrys) software.
- The physico-chemical properties of mixture components (boiling point, melting point, solubility, vapor pressure and Henry's law constant) were obtained through Chemfinder online software and the Syracuse Research Corporation (SRC) website.
- Log P for solvent components and for the penetrants was calculated by the ACD/labs logD Suite.
- Average of physicochemical properties for every solvent mixture were calculated, e.g. boiling point of the mixture.

3. Development and Validation of QSARs:

Stepwise regression analysis was used to develop models.

The present dataset was compared with the skin permeability dataset drawn from Flynn (1990) and Wilschut et al (1995) in order to identify compounds that are needed to be tested for their skin permeability from similar solvent mixtures, in order for the model to be validated. The comparison was made visually using descriptor spaces of Potts and Guy (1992) model, Principal component analysis (PCA) scores plot with all the descriptors being included in the analysis and PCA scores plot using the descriptors selected by stepwise regression analysis for the Flynn (1990) and Wilschut et al (1995) dataset.

Results

$$\text{Logkp} = -0.909 - 0.610\ \text{logP} + 2.62\ {}^{9}\chi_{p} - 0.00917\ (\text{SolBP-SolMP})$$

$$N = 288 \quad S = 0.438 \quad r^2 = 0.729 \quad F = 255.2$$

log P: Partition coefficient
$^{9}\chi_{p}$: 9th order path molecular connectivity index
SolBP-SolMP: the difference between the boiling and the melting points of the solvent
N: the number of data points
S: standard deviation
R^2: the squared correlation coefficient

- ***The negative correlation between skin penetration and lipophilicity: Penetrants too lipophilic***

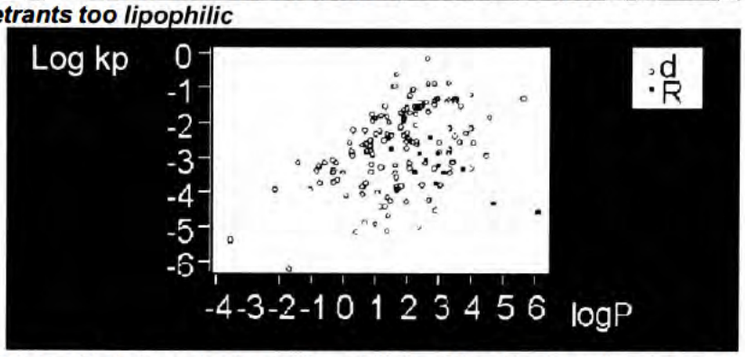
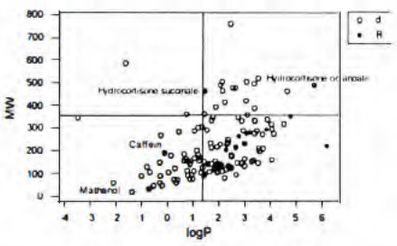


Fig 1. Relationship between log kp and liopophilicity of the skin absorption dataset of Wilschut et al (1995) (group d) and the chemicals used in this study (group R)

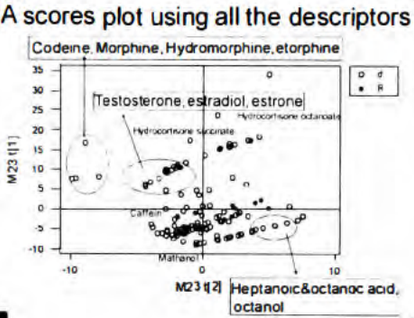
- ***The need for rigorous validation:***

Measurements needed for external validation set:

Potts and Guy (1992) model

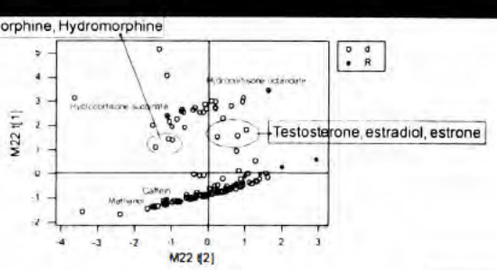


PCA scores plot using all the descriptors



[redacted]

PCA scores plot using the selected descriptors



Conclusion

In conclusion skin penetration of drugs from different vehicle systems can be modelled using QSAR. However, rigorous validation of such models for estimation purposes will require a large volume of data. The negative relationship between log kp and log P could be due to the fact that most of the drugs in this particular dataset are more lipophilic than the compounds in the common permeability datasets used in QSAR studies of skin permeability. For validation of this model, skin penetartion of the drugs identified in the graphs (or one compound out of each group identified) is necessary to be determined in similar solvent mixtures.

References

Flynn GL (1990) In: Gerrity TR and Henry CJ. Principles of route to route extraction for risk assessment. Elsevier, New York.
Potts RO and Guy RH (1992) Pharm. Res. 9: 663-669.
Riviere JE and Brooks JD (2005) Toxicology and Applied Pharmacology 208: 99-100.
Syracuse Research Corporation (SRC), http://www.syrres.com
Chemfinder, Cambridge software, http://chemfinder.cambridgesoft.com
Wilschut et al (1995) Chemosphere, 30: 1275-1296.

Modelling the Effect of Mixtures on the In-vitro Human Skin Penetration

Eleftherios Samaras and Taravat Ghafourian
Medway School of Pharmacy, Universities of Kent and Greenwich
Central Ave., Chatham, Kent ME4 4TB, UK

Abstract

Absorption of chemicals from skin is a function of the chemical structure of the penetrant and the formulation/ mixture ingredients. The experimental conditions such as dosing procedure (finite/infinite) and occlusion and hydration of skin can also affect the skin flux. The aim of this study was to investigate the effects of experimental conditions, vehicle ingredients, and chemical structure of the penetrants on the skin permeation flux using Quantitative Structure-Activity Relationships (QSAR) technique. The European research project EDETOX has put together a database of skin penetration of chemicals. For this research, skin penetration data measured *in vitro* through human skin by flow-through or static cells were extracted from the database and data from more recent publications were also incorporated. The publications were explored to identify the experimental conditions such as hydration state of the skin. Linking the skin permeation parameter to the chemical structures of the penetrants and the physico-chemical properties of the vehicle mixtures resulted in quantitative models with good accuracy for the estimation of flux. The models reported here are linear regression model and a regression tree model. Both models involve the donor concentration as a variable which follows Fick's law of diffusion. Moreover, both models involve a mixture (donor phase) parameter which is the boiling point in the Tree model and the difference between melting and boiling points in the egression model. The remaining variables of both models are molecular descriptors of the permeants which indicate the negative effects of molecular size and the polar surface area and the positive impact of certain levels of hydrophobicity.

Introduction and aims

As an integral part of the human health risk assessment of chemicals and also to be able to aid the drug delivery through skin, it is essential to be able to estimate absorption of chemicals via the dermal route. From a drug delivery perspective, penetration of the drug depends not only on the nature of the drug but also on the nature of the other ingredients present in the mixture (formulation). As a result, desirable penetration rates could be achieved by altering the formulation of a drug. Such modelling efforts are often restricted due to the limitations in the availability of data (1). As such, most skin permeation models are based on the chemical structures of the permeants often ignoring the effects of mixture components or the experimental conditions. The aim of this study was to investigate the effects of experimental conditions such as membrane thickness, diffusion cell type, occlusion, hydration, vehicle ingredients and mode of dose application (finite or infinite dosing) on the skin permeation flux. This was combined with Quantitative Structure Activity Relationships (QSAR) linking the skin flux to the chemical structures of the penetrants and the physico-chemical properties of the vehicle mixtures.

Methods

1. *The dataset:*

The *In-vitro* flux of chemicals from human skin measured by flow-through or static cells were taken from the EDETOX dataset (2). An exhaustive literature survey was carried and the data from more recent publications was added to this dataset (up to November 2010). Experimental conditions such as the occlusion and skin hydration was added for the new data and the data taken from EDETOX. In many cases, the finite or infinite dosing conditions were explicitly specified in the publication. In other cases, if the application volume>100 µl it was taken as infinite, if donor volume was 50-100 then provided that the %absorbed was <20% it was infinite or otherwise finite application.

2. *Molecular descriptors of permeants:*

The molecular descriptors were calculated using ACD labs LogD Suite (ver 12) and MOE (Chemical Computing Group).

3. *Properties of the mixture (vehicle):*

The physico-chemical properties of mixture components such as boiling point, melting point, solubility, vapor pressure and Henry's law constant were obtained through Syracuse Research Corporation (SRC) website, Sigma Aldrich website, or Chemfinder online. Average of the physicochemical properties for every solvent mixture were calculated for the liquid ingredients, e.g. boiling point of the vehicle. The effect of solid solutes (including the permeants) on boiling and melting points were calculates using the principles of the colligative properties.

4. *Development and Validation of models:*

The data was assessed using a simple mechanistic model involving a linear relationship between log Flux and donor concentration, ratio of polar surface area to the total surface area, and molar volume. The outliers were identified and the reasons were investigated. Stepwise regression analysis was used to investigate the effect of experimental variables, mixture properties and the structural descriptors of the permeants, and to develop the linear models using MINITAB Statistical Software. Non linear method of C&RT (Regression and Classification Trees) was also used for model development. The predictivity of the models were examined by a leave – many – out procedure.

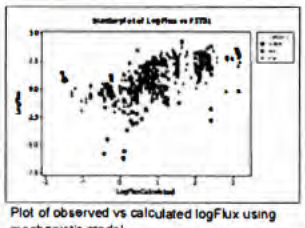
Results

1. The initial mechanistic model:

$$\log Flux = 1.53 - 0.175\ psa/sa - 0.00428\ vol + 0.000002\ [donor]$$

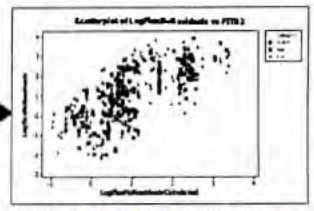
$N = 465$ $S = 0.997$ $r^2 = 0.484$ $F = 144$

psa/sa is polar surface area/surface area and *vol* is molar volume of the permeant; *[donor]* is donor concentration (µg/ml). Note that *log P* is not included as it was not significant.



Plot of observed vs calculated logFlux using mechanistic model

The outliers of the mechanistic model were identified and the values of flux where compared with data from other sources using human or animal models. The controversial data thus identified and removed before the model building exercise.



Plot of observed vs calculated logFlux using mechanistic model after exclusion of outliers

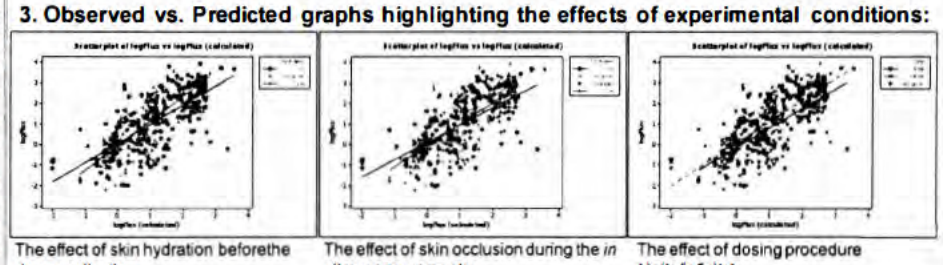
2. The selected linear regression model (stepwise regression):

$$\log Flux = -1.92 + 0.000001\ [donor] - 0.00570\ MW + 0.00235\ BP\text{-}MP(mix) + 3.96\ vsurf\_G + 0.0137\ SlogP\_VSA4 - 1.93\ fiAB - 0.343\ VAdjMa$$

$N = 454$ $S = 0.948$ $r^2 = 0.558$ $F = 80.41$

[donor]: donor concentration (µg/ml), *MW*: molecular weight, *BP-MP(mix)*: the difference between the boiling and melting point of the mixture in donor phase, *vsurf_G*: the molecular globularity – how spherical a molecule is (values above 1 is non-perfect spheres), *SlogP_VSA4*: sum of van der Waals surface area of atoms with log P contributions in the range of (0.1-0.15), *fiAB*: fraction of molecules ionized both as anion and cation at pH 7.4, *VAdjMa*: vertex adjacency information depending on number of heavy-heavy bonds (increases with increase of MW).

3. Observed vs. Predicted graphs highlighting the effects of experimental conditions:

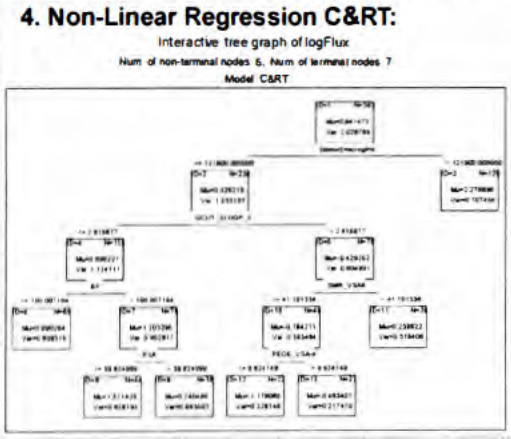


The effect of skin hydration beforethe dose application

The effect of skin occlusion during the *in vitro* measurements

The effect of dosing procedure (inite/infinite)

4. Non-Linear Regression C&RT:



Plot of observed vs predicted log Flux from the C&RT model

In the tree graph *[donor]* is the concentration of permeant in the donor phase, *GCUT_SLOGP_3* is the GCUT descriptors (calculated from the eigenvalues of a graph distance adjacency matrix) using atomic contribution to logP, *BP* is the boiling point of the donor phase, *SMR_VSA6* is sum of van der Waals surface area where atomic contributions to molar refractivity is in the range [0.485, 0.56], *PSA* is the polar surface area of the permeant and *PEOE_VSA_4* is sum of van der Waals surface area where atomic charges are in the range [-0.25, -0.20].

5. Validation of the models

Mean Absolute Error (MAE) for test and training sets

	Linear regression	C&RT
Training Set	0.812	0.607
Test Set	0.774	0.885

Conclusion

The dataset gathered here provide an excellent resource for investigating the effects of various parameters on the skin penetration parameters from in vitro study. The examples reported here involve a linear regression and a non-linear regression tree. The indicator variables for skin hydration and occlusion were not selected by the variable selection methods of stepwise regression or C&RT. Nevertheless, the graphs between observed and predicted log Flux indicate different slopes for the lines of fit. The future work will look at other modelling techniques, and also the accuracy of separate models for the data obtained under different experimental conditions. The MAE of the linear model is higher in both training and test set, again due to the fact that a larger set of data-points has been used.

References

1. Ghafourian, T., Samaras, E., Brooks, J. D., Riviere, J. E., 2010. Modelling the effect of micture components on permeation through skin. *Int. J. Pharm.*, 398, 28-32.
2. The EDETOX Database, EU Framework 5, 2005. Available at: http://research/edetox/theedetoxdatabase.html [accessed: 20 Oct 2010]

Appendix IV. Publications

International Journal of Pharmaceutics 398 (2010) 28–32

Contents lists available at ScienceDirect

International Journal of Pharmaceutics

journal homepage: www.elsevier.com/locate/ijpharm

Modelling the effect of mixture components on permeation through skin

T. Ghafourian[a,*], E.G. Samaras[a], J.D. Brooks[b], J.E. Riviere[b]

[a] Medway School of Pharmacy, Universities of Kent and Greenwich, Central Avenue, Chatham, Kent ME4 4TB, UK
[b] Center for Chemical Toxicology Research and Pharmacokinetics, 4700 Hillsborough Street, North Carolina State University, Raleigh, USA

ARTICLE INFO

Article history:
Received 4 May 2010
Received in revised form 8 July 2010
Accepted 9 July 2010
Available online 17 July 2010

Keywords:
Skin
QSAR
Formulation
Mixture
Permeation
Penetration

ABSTRACT

A vehicle influences the concentration of penetrant within the membrane, affecting its diffusivity in the skin and rate of transport. Despite the huge amount of effort made for the understanding and modelling of the skin absorption of chemicals, a reliable estimation of the skin penetration potential from formulations remains a challenging objective. In this investigation, quantitative structure–activity relationship (QSAR) was employed to relate the skin permeation of compounds to the chemical properties of the mixture ingredients and the molecular structures of the penetrants. The skin permeability dataset consisted of permeability coefficients of 12 different penetrants each blended in 24 different solvent mixtures measured from finite-dose diffusion cell studies using porcine skin. Stepwise regression analysis resulted in a QSAR employing two penetrant descriptors and one solvent property. The penetrant descriptors were octanol/water partition coefficient, log *P* and the ninth order path molecular connectivity index, and the solvent property was the difference between boiling and melting points. The negative relationship between skin permeability coefficient and log *P* was attributed to the fact that most of the drugs in this particular dataset are extremely lipophilic in comparison with the compounds in the common skin permeability datasets used in QSAR. The findings show that compounds formulated in vehicles with small boiling and melting point gaps will be expected to have higher permeation through skin. The QSAR was validated internally, using a leave-many-out procedure, giving a mean absolute error of 0.396. The chemical space of the dataset was compared with that of the known skin permeability datasets and gaps were identified for future skin permeability measurements.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Skin, the largest organ of the human body, is constantly exposed to various compounds and chemical substances in everyday life. Some of these pass the skin barrier ending in blood and hence affecting metabolism and health. Understanding the mechanism of skin penetration, and therefore the level of toxicity and irritancy of the chemicals may provide the knowledge to enhance transdermal drug delivery, boost the cosmetic industry and increase the broadness and reliability of the risk assessment on dermal exposure to toxic substances (Barry, 2007).

In order for a substance to be absorbed into the body following dermal exposure, first it must be dissolved-dissipated in the stratum corneum (SC)—the outermost sub-layer of the skin, and then diffuse through the remaining sub-layers of the epidermis and into the dermis, where it will finally diffuse into the blood capillaries. The SC is the most important barrier of the skin. SC consists of layers of tightly packed, flattened, keratin-enriched, anucleate corneocytes that are embedded in an intercellular lipid matrix (Bouwstra et al., 2002). Long chain ceramides, fatty acids, cholesterol and triglycerides are lipid matrix's main constituents (Monteiro-Riviere et al., 2001). These lipids form long lamellae parallel to the corneocyte surfaces. The lipids are arranged in bilayers consisting of ordered, crystalline phases on both sides of a narrow, central band of fluid lipids (Bouwstra et al., 2002; Monteiro-Riviere, 1986). Lipophilicity of a compound dictates the partitioning behavior into corneocytes. It can be postulated that hydrophilic compounds tend to partition into the corneocyte proteins while more lipophilic compounds into the SC lipids (Raykar et al., 1988; Van der Merwe and Riviere, 2005). There is evidence in the literature (Anderson et al., 1988) which is also supported by quantitative structure–activity relationship (QSAR) studies (Ghafourian and Fooladi, 2001) which indicates the higher partitioning of more lipophilic compounds containing fewer heteroatoms into the lipid domain of SC, while partitioning to protein domain is less sensitive to the size and number of heteroatoms of penetrants.

Permeation of chemicals through skin is not only dependent on the chemical structure and properties of the penetrant itself, but also it is affected by the other chemicals present in the mixture. Solvents and other mixture ingredients can alter the permeation profile of a chemical by changing the properties of the lipid and

* Corresponding author. Tel.: +44 1634202952; fax: +44 1634883927.
E-mail address: t.ghafourian@kent.ac.uk (T. Ghafourian).

0378-5173/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijpharm.2010.07.014

protein domains of SC, solubility and therefore the thermodynamic activity of the penetrant in the mixture, and partitioning of the penetrant from the vehicle into the SC. Chemical enhancers, for example, can cause a dynamic structural disorder in the SC lipid domain that will lead to enhanced transdermal permeation (Bezema et al., 1996).

A problem that emerges at this stage is the difficulty of accurately predicting the diffusivity and partitioning as they are both ultimately dependent on the skin structure, changes to the skin caused by various solvents and permeants, changes of the formulation containing the permeant, and the effect of metabolizing enzymes on permeants. From the above, given also the almost unlimited possible combinations of solvent mixtures and permeants, it can be assumed that the accurate prediction of diffusion and partition from permeant and solvent chemical data is uncertain (Van der Merwe and Riviere, 2005). On the other hand large sets of empirical data provide us with valuable certainty in the process of identifying characteristics of permeant and vehicle systems that have consistent effects across a wide range of experimental conditions.

Empirical data of permeant and solvent has been used extensively with success in predicting skin permeability. This is often carried out through the use of QSAR where skin permeation profile is related to the molecular properties of compounds, given that the skin permeation is measured at consistent experimental conditions. QSAR has been efficiently used to model skin permeation of chemicals from simple systems such as saturated aqueous solutions (El Tayar et al., 1991; Abraham et al., 1995, 1999; Ghafourian and Fooladi, 2001; Moss and Cronin, 2002). Potts and Guy (1992) developed the first widely accepted QSAR model for predicting skin permeability coefficient (k_p), a linear regression model that consisted of lipophilicity measured by octanol/water partition coefficient and molecular weight. The Potts and Guy model is based on the data collated by Flynn (1990) consisting *in vitro* skin permeability coefficients of 94 compounds from aqueous solutions. On the other hand, a systematic approach to investigate the effect of mixture components is essential. This is not only due to the fact that most chemicals that the skin is exposed to are in mixtures, but also because of the impact of such mixture constituents on the skin absorption. Despite the availability of QSAR models representing the effect of chemical enhancers on the permeation of drugs (Ghafourian et al., 2004; Pugh et al., 2005), due to the lack of sufficient high quality data, such models for the effect of other mixture ingredients such as solvents are not available in the literature.

In a recent work, Riviere and Brooks (2005, 2007) investigated *in vitro* permeation of several chemicals from chemical mixtures containing various concentrations of different solvents, a surfactant (sodium lauryl sulfate, SLS) and a vasodilator (methyl nicotinic acid) (Tur et al., 1991). This comprehensive dataset provides an opportunity for understanding the effect of mixture components on the skin permeation through QSAR modelling. The aim of this investigation was to develop QSAR models to study the effect of mixture components on skin absorption of penetrants. The model will help identify the mechanisms involved in the permeation through skin and the effect of formulation factors.

2. Methods

2.1. The dataset

Skin permeation data of 12 different penetrants (Table 1) each blended in 24 different solvent mixtures (Table 2), were used in this investigation. Experimental details are fully described in Riviere and Brooks (2005). The permeability data consisted of apparent skin permeation rate constants (k_p) in cm/h measured using finite-dose *in vitro* porcine skin flow through diffusion cells. The skin was perfused using a Krebs–Ringer bicarbonate buffer spiked with dextrose and bovine serum albumin, and topically dosed nonoccluded with 20 μL of one of the 12 marker penetrant compounds (target dose of 10–20 μg/cm^2) formulated in one of the 24 specified mixtures (Table 2). Trace amounts of methanol and toluene were used to solubilize radiolabelled penetrants before dilution with nonradiolabelled compounds.

Table 1
Penetrants.

Atrazine	Pentachlorophenol
Chlorpyrifos	Phenol
Ethylparathion	p-Nitrophenol
Fenthion	Propazine
Methylparathion	Simazine
Nonylphenol	Triazine

This dataset was compared in terms of the chemical space of the penetrants with the combined datasets of Flynn (1990) and Wilschut et al. (1995). The combined dataset contains *in vitro* human skin permeability data ($\log k_p$) for 112 compounds.

2.2. Structural descriptors

The predictors (descriptors) of penetrants included connectivity indexes, quantum molecular descriptors, and group counts calculated using TSAR 3D software (Accelrys Ltd. version 3.3). The physicochemical properties of mixture components including boiling point, melting point, solubility, vapour pressure and Henry's law constant were obtained through ChemBioFinder (CambridgeSoft, 2009) online software and the SRC PhysProp database (Syracuse Research Corporation, 2009). Log *P* for solvent components and for the penetrants was calculated by the ACD/labs log D Suite (7.0.5 release). Averages of physicochemical properties for solvent mixtures were calculated using the fractions of each component e.g. boiling point of the mixture.

2.3. Development and validation of QSARs

Stepwise regression analysis was used to develop the models in MINITAB (version 15.1.0.0). The predictability of the models was

Table 2
Mean Composition of the 24 mixtures.

Mixture	%EtOH	%Water	%PG	%MNA	%SLS
Et	99.67	0	0	0	0
Et + MNA	99.51	0	0	0.16	0
Et + SLS	62.59	26.53	0	0	10.61
Et + MNA + SLS	62.50	26.49	0	0.13	10.60
Et + Wa	42.66	55.86	0	0	0
Et + Wa + MNA	43.79	55.78	0	0.14	0
Et + Wa + SLS	39.44	50.25	0	0	10.05
Et + Wa + MNA + SLS	39.39	50.18	0	0.13	10.04
Wa	0	99.75	0	0	0
Wa + MNA	3.03	96.59	0	0.13	0
Wa + SLS	0	90.70	0	0	9.07
Wa + MNA + SLS	2.75	87.77	0	0.12	9.13
Et + PG	42.99	0	56.73	0	0
Et + PG + MNA	42.92	0	56.65	0.14	0
Et + PG + SLS	28.39	24.15	37.54	0	9.66
Et + PG + MNA + SLS	28.36	24.13	37.50	0.12	9.65
PG	0	0	99.75	0	0
PG + MNA	2.93	0	96.70	0.12	0
PG + SLS	0	22.13	68.79	0	8.85
PG + MNA + SLS	2.69	22.29	65.76	0.11	8.92
Wa + PG	0	48.99	50.76	0	0
Wa + PG + MNA	2.98	47.46	49.18	0.13	0
Wa + PG + SLS	0	44.62	46.23	0	8.92
Wa + PG + MNA + SLS	2.71	43.20	44.76	0.11	8.98

EtOH, ethanol; PG, propylene glycol; MNA, methyl nicotinate; SLS, sodium lauryl sulfate.

examined by a leave-many-out procedure. As such, chemicals were sorted according to the ascending $\log k_p$ values; for each set of 4 solvents, the first compound was allocated to group a, the second to group b, the third to group c and the fourth to group d. This ensured that each group covered similar ranges of the skin permeation kinetics. The regression was carried for the chemicals in groups a, b and c (as the training set), and the resulting equation was used to calculate the skin permeation parameter for the remaining group d (as the test set). The procedure was carried on to leave one group out at a time (all the possible combinations of groups making the training set). The mean absolute error (MAE) of prediction was calculated as a measure of the model accuracy.

The chemical space of the present dataset was compared with that of the skin permeability dataset drawn from Flynn (1990) and Wilschut et al. (1995). The comparison was made using descriptor spaces of Potts and Guy (1992) model (i.e. molecular weight and octanol/water partition coefficient), principal component analysis (PCA) scores plot with all the descriptors being included in the analysis and PCA scores plot using the descriptors selected by stepwise regression analysis for the Flynn (1990) and Wilschut et al. (1995) dataset. PCA was carried out using MINITAB statistical software.

3. Results and discussion

The combined effect of chemical structures of the penetrants and the properties of the mixture components on the permeation rate through porcine skin was studied using QSAR. Stepwise regression analysis performed on the dataset of 288 penetrant/mixture component combinations resulted in Eq. (1), in which the descriptors were limited to two penetrant descriptors and one solvent mixture descriptor.

$$\text{Log}\,k_p = -0.909 - 0.610\,\log P + 2.629\chi_p - 0.00917(\text{SolBP} - \text{SolMP}) \quad (1)$$

$$S = 0.438, \quad r^2 = 0.729, \quad F = 255.2, \quad P = 0.000, \quad N = 288$$

In Eq. (1), $\log k_p$ represents permeation rate constant of compounds dissolved in various solvent mixtures from porcine skin, $\log P$ is the octanol/water partition coefficient of the solute (the penetrant), ${}^9\chi_p$ is the 9th order path molecular connectivity index of the penetrant, and SolBP – SolMP is the difference between the boiling point and the melting point of the solvent system.

Log P was the most significant descriptor of the equation (the first to be selected by the stepwise regression analysis). It can be seen in Eq. (1) that $\log P$ of penetrants has a negative effect on the skin permeation rate. This is opposite to the common knowledge that lipophilic compounds have higher skin permeation profiles, as evidenced also in Potts and Guy's model (1992). The negative relationship between $\log k_p$ and $\log P$ could be due to the fact that most of the drugs in this particular dataset are more lipophilic than the compounds in the datasets normally used in QSAR studies of skin permeability. Fig. 1 shows a graph between $\log k_p$ and $\log P$ for the penetrants of this study and the penetrants of Wilschut et al. (1995) and Flynn (1990). The opposite trends of the relationships between $\log k_p$ and $\log P$ for the two datasets are evident despite the poor correlations. The figure also shows that compounds of the present dataset have relatively higher $\log P$ values than compounds in the combined datasets of Wilschut et al. (1995) and Flynn (1990). This follows the well established nonlinear relationship of biological activity with lipophilicity described by parabolic (Hansch and Clayton, 1973) or bilinear (Kubinyi, 1977) models. Compounds with extreme lipophilicity can be expected to partition into the skin and remain there, with little permeation to the aqueous receptor phase. This has been shown for example for tetrahydrocannabinol (Challapalli and Stinchcomb, 2002), with extremely high $\log P$ value of 6.84 as calculated by ACD log D/Suite. López et al. (1998)

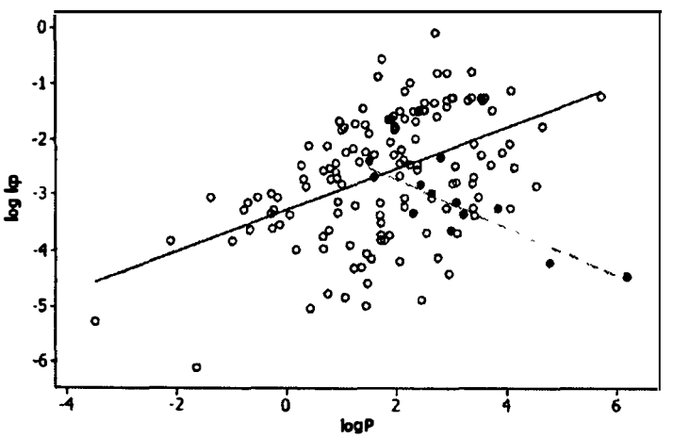


Fig. 1. Comparison of the lipophilicity of the drugs in the two datasets of Riviere's (solid circles) and Flynn (1990) and Wilschut et al. (1995) dataset (empty circles).

showed a bilinear relationship between lipophilicity of phenyl alcohols and the permeability coefficient through rat skin, where the k_p was reduced for compounds with log P values higher than around 5.

There are a number of other factors that may have contributed to the observed negative relationship between $\log k_p$ and lipophilicity. One is the finite-dose nature of the experiments with skin dosed with a limited amount of drug. The limited availability of the compound could result in a large fraction of the lipophilic compounds being concentrated in the skin according to their skin/water partition coefficients. A second factor is the differing nature of the receptor phase which contained albumin.

${}^9\chi_p$ is the second most significant descriptor of Eq. (2). This molecular connectivity descriptor indicates the presence of nine-atom chains in the molecules. The positive coefficient of this descriptor indicates a better permeation of compounds containing long chain fragments. The penetrants with the highest ${}^9\chi_p$ values were chlorpyrifos, fenthion and nonylphenol. These penetrants have the maximum molecular weight of 350 Da which is still smaller than the size expected to limit the absorption. According to Barry (2007) a molecule's ideal molecular mass, in order to penetrate the SC is less than 600 Da. In addition, according to Lipinski's rule of five, chemicals with molecular weight of above 500 Da may have biological membrane penetration problems (Lipinski et al., 1997).

The third descriptor of the equation, SolBP–SolMP, represents the difference between melting and boiling points of the solvent mixtures, where the higher the difference, the lower the skin absorption of compounds from the vehicle. It is therefore expected that compounds formulated in vehicles with small boiling and melting point gaps will have better permeation through skin. The difference between these two properties has been attributed to the molecular symmetry, with highly symmetrical molecules having much larger melting points and decreased boiling points (Slovokhotov et al., 2007). In the solvents used in this study, the biggest difference in the melting and boiling points is for propylene glycol. Therefore the vehicles containing higher concentrations of this solvent will lead to lower permeation of the penetrants studied in this investigation.

In order to validate the reported QSAR, a leave-many-out procedure as explained in Section 2 was used and mean absolute error calculated. Fig. 2 is the graph between observed and predicted $\log k_p$. The r^2 between observed and predicted $\log k_p$ and the MAE were 0.654 and 0.396, respectively.

The level of uncertainty in predictions made by any QSAR is characterized by the validity tests, but it also depends on the diversity of the training set which defines the domain of applicability. Any QSAR model is expected to perform best for the chemicals that are similar

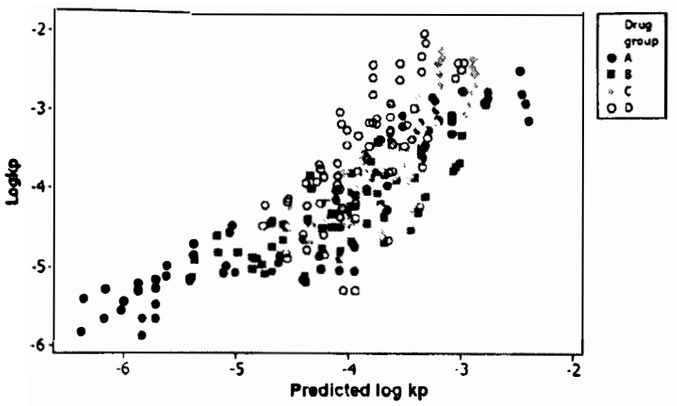


Fig. 2. Plot of observed $\log k_p$ against predicted $\log k_p$.

to those in the training set (Weaver and Gleeson, 2008). Applicability of Eq. (1) will be limited to the prediction of $\log k_p$ for new molecules that are similar to those of our dataset. Therefore, the chemical space of the penetrants included in this dataset was compared to that of the combined datasets of Flynn (1990) and Wilschut et al. (1995). Comparison of the physicochemical properties of the penetrants in the two datasets above were made first by looking at the molecular descriptors of the widely accepted Potts and Guy model (1992) consisting of $\log P$ and molecular weight (MW) as in Fig. 3a. The figure shows that our dataset does not include any hydrophilic chemicals and $\log P$ values are all above 1.5. The other limitation of the dataset is the relatively low molecular weights of the chemicals in comparison with datasets of Flynn and Wilschut et al. Therefore, a few high molecular weight and low lipophilicity chemicals can be identified for future measurements. Examples are hydrocortisone octanoate, caffeine and methanol, as it can be seen in the figure.

As a second strategy, the two datasets were compared using all the calculated molecular descriptors, a total of 128. This was made possible through the use of principal component analysis (PCA). PCA is a data reduction method which takes the information from original molecular descriptors and generates the same number of new descriptors (PCs), with the first PC containing the maximum information of the original dataset, and the second PC being the second most informative. Therefore, the plot between PC1 and PC2 (the scores plot) provides a good overview of the information content of the dataset. The first two principal component score vectors, PC1 and PC2, are plotted in Fig. 2b. The figure shows that the chemicals of the current dataset are located in the bottom left quarter of the plot, with relatively low PC1 and PC2 values. By visual inspection of the graph, several groups of chemicals belonging to datasets of Flynn and Wilschut et al. were identified in the plot to cover various ranges of PC1/PC2. These are chemicals with high PC1 and PC2 values such as codeine and morphine, compounds with high PC1 and varying values of PC2 including steroids such as testosterone and hydrocortisone octanoate, and compounds with very low PC1 and PC2 values such as octanol.

The third method for comparison of the datasets involved the use of a selection of molecular properties that are specifically involved in the skin permeation of compounds. To this end, stepwise regression analysis was used for the selection of molecular descriptors affecting compounds' absorption through skin. In this analysis, the dataset of Flynn and Wilschut et al. containing the skin permeation rate constant through human skin using the saturated aqueous solutions as the donor phase was used. In stepwise regression analysis, the skin permeation rate constant ($\log k_p$) was the dependent variable and all the molecular descriptors were the independent variables. Stepwise regression analysis selected three

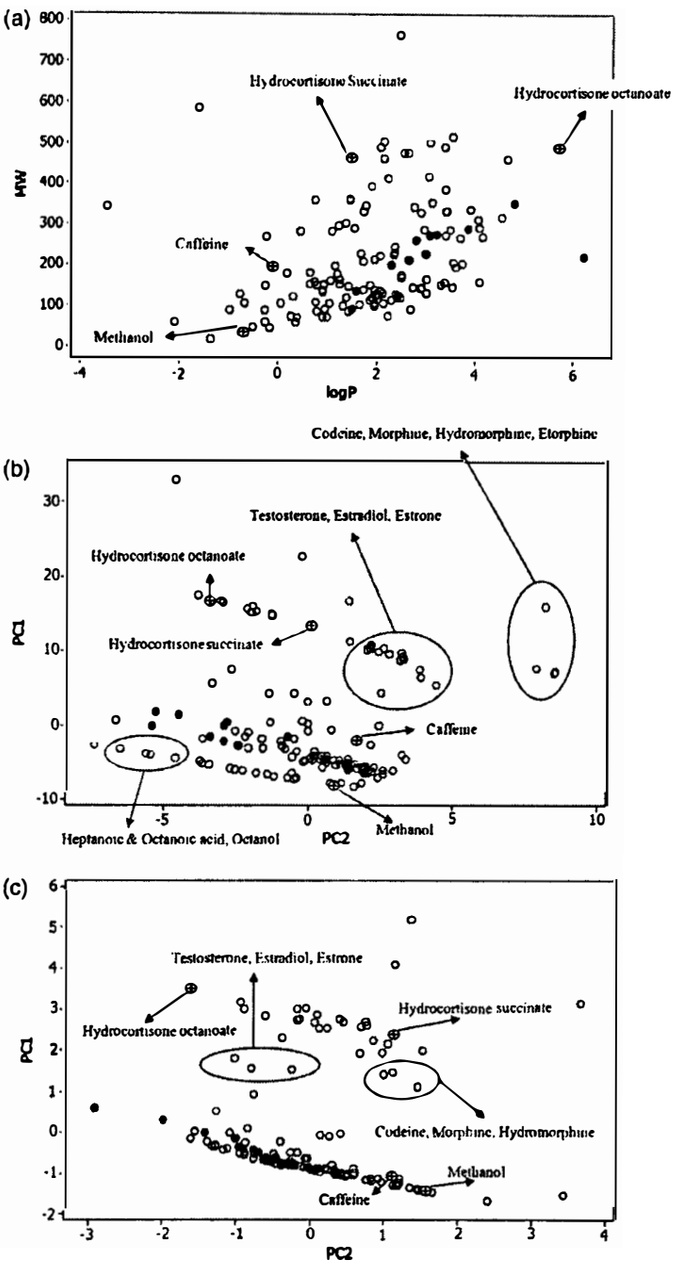
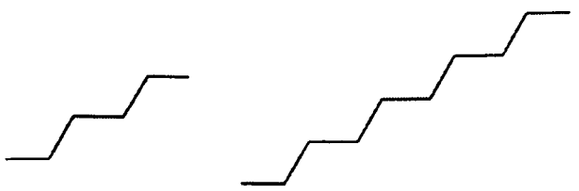


Fig. 3. Plots comparing chemical diversity of the penetrants of the present dataset (solid circles) with that of the combined dataset of Flynn (1990) and Wilschut et al. (1995) (empty circles). (a) Plot between $\log P$ and molecular weight; (b) scores plot between the first and the second principal components of PCA using all the descriptors; (c) scores plot between the first and the second principal components of PCA using descriptors of Eq. (2).

descriptors and resulted in Eq. (2) below.

$$\text{Log}\,k_p = -2.91 + 0.62\ \log P + 5.21\,{}^{10}\chi_p^v - 1.64\,{}^{6}\chi_p^v$$
$$S = 0.548,\quad r^2 = 0.757,\quad F = 140,\quad P = 0.000,\quad N = 139 \qquad (2)$$

In Eq. (2), $\log P$ is the octanol/water partition coefficient, ${}^{10}\chi_p^v$ and ${}^{6}\chi_p^v$ are 10th and 6th order valence corrected path molecular connectivity indexes of the penetrants, respectively. Molecular connectivity indexes are topological descriptors of molecular structures indicating the frequencies of occurrence of certain fragments in the molecules. Path molecular connectivity indexes indicate the frequency of non-branched chains of certain lengths, in this case six-atom and ten-atom chains as shown in Scheme 1 below (Todeschini and Consonni, 2000).



Scheme 1. Six-atom and ten-atom fragments for the calculation of path molecular connectivity indexes, $^{6}\chi_{p}^{v}$ and $^{10}\chi_{p}^{v}$, respectively.

The three descriptors selected by stepwise regression analysis were used in PCA and the scores plot between the first and the second PCs (Fig. 3c) was used to compare the datasets. Fig. 3c is similar to Fig. 3b in identifying certain compounds from the dataset of Flynn and Wilschut et al. such as steroids, narcotic analgesics and small polar molecules such as caffeine and methanol which are not present in the current dataset.

Therefore, an overview of Fig. 3a and b can identify several areas of the chemical space that are missing from the present dataset. From these groups of chemicals, caffeine, 1-octanol, testosterone and codeine were selected for further studies and the *in vitro* measurements are currently being undertaken.

4. Conclusion

In conclusion skin permeation of drugs from different vehicle systems can be modelled using QSAR given the availability of an appropriate dataset containing diverse permeants and vehicles. Vehicle effects were well predicted in this work. However, rigorous validation of such models for estimation purposes will require a large volume of data. In this study, the negative relationship was obtained between $\log k_p$ and $\log P$. This was attributed to the fact that most of the drugs in this particular dataset are more lipophilic than the compounds in the common permeability datasets used in QSAR studies of skin permeability. Therefore, it can be envisaged that these highly lipophilic agents concentrate in the SC with little ability to partition into the aqueous receptor phase. This scenario is relevant for many pesticides and lipophilic contaminants encountered in environmental exposure scenarios. For further validation of this model, skin permeation of the compounds identified through the comparison of the datasets is necessary to be determined in similar solvent mixtures.

References

Abraham, M.H., Chadha, H.S., Martins, F., Mitchell, R.C., Bradbury, M.W., Gratton, J.A., 1999. Hydrogen bonding part 46. A review of the correlation and prediction of transport properties by an LFER method: physicochemical properties, brain penetration and skin permeability. Pestic. Sci. 55, 78–88.

Abraham, M.H., Chadha, H.S., Mitchell, R.C., 1995. The factors that influence skin penetration of solutes. J. Pharm. Pharmacol. 47, 8–16.

Anderson, B.D., Higuchi, W.I., Raykar, P.V., 1988. Heterogeneity effects on permeability–partition coefficient relationships in human stratum corneum. Pharm. Res. 5, 566–573.

Barry, B.W., 2007. Transdermal drug delivery. In: Aulton, M.E. (Ed.), Aulton's Pharmaceutics, the Design and Manufacture of Medicines, 3rd ed. Churchill Livingstone, Elsevier. pp. 580–585.

Bezema, F.R., Marttin, E., Roemele, P.E.H., Brussee, J., Bodde, H.E., deGroot, H.J.M., 1996. H-2 NMR evidence for dynamic disorder in human skin induced by the penetration enhancer Azone. Spectrochim. Acta. Mol. Biomol. Spectrosc. 52, 785–791.

Bouwstra, J.A., Honeywell-Nguyen, P.L., Gooris, G.S., Ponec, M., 2002. Structure of the skin barrier and its modulation by vesicular formulations. Prog. Lipid Res. 42, 1–36.

CambridgeSoft, 2009. ChemBioFinder, http://www.cambridgesoft.com/databases, retrieved September 19–22, 2009.

Challapalli, P.V.N., Stinchcomb, A.L., 2002. In vitro experiment optimization for measuring tetrahydrocannabicol skin permeation. Int. J. Pharm. 241, 329–339.

El Tayar, N., Tsai, R.S., Testa, B., Carrupt, P.A., Hansch, C., Leo, A., 1991. Percutaneous penetration of drugs: a quantitative structure–permeability relationship study. J. Pharm. Sci. 80, 744–749.

Flynn, G.L., 1990. Physicochemical determinants of skin absorption. In: Gerrity, T.R., Henry, C.J. (Eds.), Principles of Route-to-route Extrapolitation for Risk Assessment. Elsevier, New York, NY, pp. 93–127.

Ghafourian, T., Fooladi, S., 2001. The effect of structural QSAR parameters on skin penetration. Int. J. Pharm. 217, 1–11.

Ghafourian, T., Zandasrar, P., Hamishekar, H., Nokhodchi, A., 2004. The effect of penetration enhancers on drug delivery through skin: a QSAR study. J. Control. Release 99, 113–125.

Hansch, C., Clayton, J.M., 1973. Lipophilic character and biological-activity of drugs. 2. Parabolic case. J. Pharm. Sci. 62, 1–21.

Kubinyi, H., 1977. Quantitative structure–activity-relationships. 7. Bilinear model, a new model for nonlinear dependence of biological-activity on hydrophobic character. J. Med. Chem. 20, 625–629.

Lipinski, C.A., Lombardo, F., Dominy, B.W., Feeney, P.J., 1997. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Del. Rev. 23, 3–25.

López, A., Faus, V., Díez-Sales, O., Herráez, M., 1998. Skin permeation model of phenyl alcohols: comparison of experimental conditions. Int. J. Pharm. 173, 183–191.

Monteiro-Riviere, N.A., Inman, A.O., Mak, V., Wertz, P., Riviere, J.E., 2001. Effect of selective lipid extraction from different body regions on epidermal barrier function. Pharm. Res. 18, 992–998.

Monteiro-Riviere, N.A., 1986. Ultrastructural evaluation of the porcine integument. In: Tumbleson, M.E. (Ed.), Swine in Biomedical Research, vol. 1. Plenum, New York, pp. 641–655.

Moss, G.P., Cronin, M.T., 2002. Quantitative structure-permeability relationships for percutaneous absorption: re-analysis of steroid data. Int. J. Pharm. 238, 105–109.

Potts, R.O., Guy, R.H., 1992. Predicting skin permeability. Pharm. Res. 9, 663–669.

Pugh, W.J., Wong, F.F., Michniak, B.B., Moss, G.P., 2005. Discriminant analysis as a tool to identify compounds with potential as transdermal enhancers. J. Pharm. Pharmacol. 57, 1389–1396.

Raykar, P.V., Fung, M.C., Anderson, B.D., 1988. The role of protein and lipid domains in the uptake of solutes by human stratum corneum. Pharm. Res. 5, 140–150.

Riviere, J.E., Brooks, J.D., 2007. Prediction of dermal absorption from complex chemical mixtures: incorporation of vehicle effects and interactions into a QSPR framework. SAR QSAR Environ. Res. 18, 31–44.

Riviere, J.E., Brooks, J.D., 2005. Predicting skin permeability from complex chemical mixtures. Toxicol. Appl. Pharmacol. 208, 99–100.

Slovokhotov, Y.L., Batsanov, A.S., Howard, J.A.K., 2007. Molecular van der Waals symmetry affecting bulk properties of condensed phases: melting and boiling points. Struct. Chem. 18, 477–491.

Syracuse Research Corporation, 2009. SRC PhysProp database, http://www.syrres.com/what-we-do/databaseforms.aspx?id=386, retrieved September 20, 2009.

Todeschini, R., Consonni, V., 2000. Handbook of Molecular Descriptors. Wiley-VCH.

Tur, E., Maibach, H., Guy, R.H., 1991. Percutaneous penetration of methyl nicotinate at 3 anatomic sites—evidence for an appendageal contribution to transport. Skin Pharmacol. 4, 230–234.

Van der Merwe, D., Riviere, J.E., 2005. Comparative studies on the effects of water, ethanol and water/ethanol mixtures on chemical partitioning into porcine stratum corneum and silastic membrane. Toxicol. In Vitro 19, 69–77.

Weaver, S., Gleeson, M.P., 2008. The importance of the domain of applicability in QSAR modeling. J. Mol. Graph. Model. 26, 1315–1326.

Wilschut, A., Berge, W.F., Robinson, P.J., McKone, T.E., 1995. Estimating skin permeation. The validation of five mathematical skin permeation models. Chemosphere 30, 1275–1296.

European Journal of Pharmaceutical Sciences 41 (2010) 612–616

Contents lists available at ScienceDirect

European Journal of Pharmaceutical Sciences

journal homepage: www.elsevier.com/locate/ejps

Validated models for predicting skin penetration from different vehicles☆

Taravat Ghafourian[a,*], Eleftherios G. Samaras[a], James D. Brooks[b], Jim E. Riviere[b]

[a] Medway School of Pharmacy, Universities of Kent and Greenwich, Central Avenue, Chatham Maritime, Kent ME4 4TB, UK
[b] Center for Chemical Toxicology Research and Pharmacokinetics, 4700 Hillsborough Street, North Carolina State University, Raleigh, USA

ARTICLE INFO

Article history:
Received 14 May 2010
Received in revised form 10 August 2010
Accepted 28 August 2010
Available online 15 September 2010

Keywords:
Skin
Penetration
QSAR
Permeation
Formulation
Mixture

ABSTRACT

The permeability of a penetrant though skin is controlled by the properties of the penetrants and the mixture components, which in turn relates to the molecular structures. Despite the well-investigated models for compound permeation through skin, the effect of vehicles and mixture components has not received much attention. The aim of this Quantitative Structure Activity Relationship (QSAR) study was to develop a statistically validated model for the prediction of skin permeability coefficients of compounds dissolved in different vehicles. Furthermore, the model can help with the elucidation of the mechanisms involved in the permeation process. With this goal in mind, the skin permeability of four different penetrants each blended in 24 different solvent mixtures were determined from diffusion cell studies using porcine skin. The resulting 96 *kp* values were combined with a previous dataset of 288 *kp* data for QSAR analysis. Stepwise regression analysis was used for the selection of the most significant molecular descriptors and development of several regression models. The selected QSAR employed two penetrant descriptors of Wiener topological index and total lipole moment, boiling point of the solvent and the difference between the melting point of the penetrant and the melting point of the solvent. The QSAR was validated internally, using a leave-many-out procedure, giving a mean absolute error of 0.454 for the log *kp* value of the test set.
© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Skin is the largest organ of all mammals protecting the underlying muscles, bones, ligaments and internal organs as well as guarding the whole organism from exogenous molecules. Within skin the outermost layer, stratum corneum, is the formidable barrier to the exogenous compounds, limiting penetration of toxicants and drugs alike. Skin penetration of chemicals is an integral part of human health risk assessment of chemicals exposed via the dermal route (Shah, 1994). Skin has also been the focus of research by drug formulators as a site of drug administration, due to the advantages it may offer over other routes of drug delivery (Barry, 2007). Topical delivery affects the tissues under the site of application, while systemic delivery has an effect after distribution to the circulatory system. The rate of drug delivery through skin is influenced by factors including skin health status, age, race, anatomical region, thermodynamic activity of the drug in the formulation and interactions of the drug and formulation with the skin. Drug in the formulation needs to pass the stratum corneum's intercellular lipids that surround dead keratin-filled corneocytes and also the subcutaneous fat to reach the blood capillaries (Elias, 1983).

Penetration of a compound into skin is controlled by its physicochemical properties and the chemical structure. For example, it has been shown that lipophilicity and hydrogen bonding ability of a compound plays a major role on the skin absorption profile (Pugh et al., 1996; El Tayar et al., 1991). On the other hand, formulation ingredients can alter the skin penetration of a compound by affecting the barrier properties of the skin or by changing the partitioning of the compound into the SC. Therefore, penetration of the drug depends not only on the nature of the drug but also on the nature of the other ingredients present in the formulation. The vehicle in which a penetrant is dissolved or dispersed is of outmost significance. Vehicles can affect skin permeability by a range of mechanisms including delipidization, hydration, fluidization and desmosome disruption in the stratum corneum, and also by changing the polarity of the formulation mixture which is followed by a change in the penetrant solubility and partitioning to stratum corneum (Roberts et al., 2002). Solvents are also likely to affect the conformation of stratum corneum in a way that the diffusion and partitioning of the penetrants are modified (Kai et al., 1990; Raykar et al., 1988; Rosado et al., 2003). Pure solutes can in some cases enhance the skin permeability by a direct corrosive effect (Roberts et al., 2002; Zinke et al., 2002). Other common mixture components are surfactants and, in case of drug formulations,

☆ Supported by NIOSH grant R01-OH07555.
* Corresponding author at: Medway School of Pharmacy, Universities of Kent and Greenwich, Central Avenue, Chatham Maritime, Kent ME4 4TB, UK.
Tel.: +44 1634 202952; fax: +44 1634 883927.
E-mail address: t.ghafourian@kent.ac.uk (T. Ghafourian).

0928-0987/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejps.2010.08.014

penetration enhancers. Surfactants are used in the pharmaceutical/cosmetic preparations, agrochemical products (e.g. herbicides) and industrial solutions. In industry surfactants are added to formulations in order to solubilise lipophilic active ingredients, and in transdermal drug delivery to solubilise lipids within the stratum corneum. Penetration enhancers may increase the diffusion coefficient of drugs in the stratum corneum (i.e. disrupt the barrier nature of the stratum corneum), may act to increase the effective concentration of the drug in the vehicle (for example, acting as an anti-solvent), could improve partitioning between the formulation and the stratum corneum (perhaps by altering the solvent nature of the skin membrane to improve partitioning into the tissue) or, less likely, by decreasing the skin thickness (perhaps by providing a permeation 'shortcut' as opposed to a tortuous pathway for a permeant) (Williams and Barry, 2004).

The effect of mixture/formulation components on the skin penetration of a compound depends on the nature of the component, i.e. its chemical structure and physicochemical properties. In other words, chemical structure of a formulation component can determine the effect that it will have on the stratum corneum or on the partitioning of the penetrant, leading to the observed changes in the skin penetration profile of the penetrant. The relationship between chemical structures of the formulation ingredients and the skin penetration modification can be studied quantitatively using Quantitative Structure–Activity Relationship (QSAR) techniques. QSAR has been previously applied to study the effect of structural variation of chemical enhancers on the skin penetration of various drugs (Ghafourian et al., 2004; Moss et al., 2002).

Most mechanistic studies on skin penetration are based on the penetration of individual chemicals (Flynn, 1990), with only few attempts towards a comprehensive investigation on the effect of chemical mixtures. Such a systematic study requires a large volume of tedious experimental measurements involving various penetrant/mixture-component combinations. Riviere and Brooks (2005, 2007) have determined skin permeation coefficient of 12 compounds from a mixture of several solvents, a surfactant and methyl nicotinic acid (288 combinations). A QSAR analysis of the data revealed several penetrant/solvent properties that are significant contributors to the skin permeation coefficients (Ghafourian et al., 2010). The study also revealed several gaps in the chemical space of Riviere's penetrants in comparison with the well-established datasets of Flynn (1990) and Wilschut et al. (1995) which contain skin penetration data of aqueous solutions of over 100 compounds. In this investigation, four chemicals were selected from Flynn and Wilschut et al. datasets for further skin penetration studies using Riviere's experimental protocol which involves blending of each chemical with 24 mixture combinations. The selections were made from the identified gaps in the chemical space and the compounds are expected to add a high level of diversity to the dataset. These new measurements facilitated the development of statistically validated QSAR models. Statistically validated QSAR models can be used for the estimation of skin penetration of new compounds or the effect of new mixture components on the penetration of a penetrant. The models can aid the understanding of the mechanisms involved in skin penetration of compounds and the effect of mixture components.

2. Materials and methods

2.1. Materials

Caffeine [8-14C] specific activity: 50–60 mCi/mmol, 1.85–2.22 GBq/mmol, n-octanol [1-14C] specific activity: 2–10 mCi/mmol, 74–370 MBq/mmol, testosterone [4-14C] specific activity: 50–60 mCi/mmol 1.85–2.22 GBq/mmol, codeine [N-methyl-14C], obtained from American Radiolabeled Chemicals, Inc., St. Louis, USA. Absolute ethyl alcohol was obtained from Aaper Alcohol and Chemical Co., Shelbyville, KY, USA. Propylene glycol (purity = 99%), sodium lauryl sulphate (purity = 99%), and methyl nicotinic acid (purity = 99%) were obtained from Sigma Chemical Co., St. Louis, MO, USA. Water was distilled in our in-house still.

Table 1
Composition of the 24 mixtures.

EtOH	PG
EtOH + MNA	PG + MNA
EtOH + SLS	PG + SLS
EtOH + MNA + SLS	PG + MNA + SLS
EtOH + Water	PG + Water
EtOH + Water + MNA	PG + Water + MNA
EtOH + Water + SLS	PG + Water + SLS
EtOH + Water + MNA + SLS	PG + Water + MNA + SLS
EtOH + PG + Water	Water
EtOH + PG + Water + MNA	Water + MNA
EtOH + PG + Water + SLS	Water + SLS
EtOH + PG + Water + MNA + SLS	Water + MNA + SLS

EtOH—ethanol; PG—propylene glycol; MNA—methyl nicotinate; SLS—sodium lauryl sulphate.

2.2. Skin penetration studies

Apparent permeability coefficient (*kp*) of caffeine, codeine, octanol and testosterone each blended in 24 different mixtures, as presented in Table 1, were obtained through flow-through diffusion cell using porcine skin. The flow-through diffusion cell was used to perfuse skin obtained from the dorsal area of weanling female Yorkshire pigs according to protocols approved by the North Carolina State University Institutional Animal Care and Use Committee. Skin was dermatomed to a thickness of 500 μm with a Padgett dermatome. Each circular skin disk was punched to provide a dosing surface area of 0.64 cm^2 and then placed into a two-compartment Teflon Bronaugh flow-through diffusion cell. Skin was perfused using a Krebs–Ringer bicarbonate buffer spiked with dextrose and bovine serum albumin, and topically dosed nonoccluded with 20 μl of one of the four marker penetrant compounds (10 $\mu g/cm^2$) formulated in one of 24 specified mixtures listed in Table 1. This resulted in a total of 96 treatments with $n = 4–5$ replicates/treatment designed as a randomized complete factorial experiment.

2.3. QSAR studies

The *kp* values measured in this study for caffeine, codeine, octanol and testosterone were merged with the previous dataset of *kp* values for 12 other compounds blended with the same mixture components as Table 1 (Riviere and Brooks, 2005). These *kp* values are measured using the same experimental procedures as in this study. Therefore, the dataset used for the QSAR studies consisted of a total of 384 unique measurements of *kp* for the penetrant/components combinations. Table 2 is the list of the 16 penetrants used in QSAR study.

For the development of QSAR models, properties of the penetrants and the solvent mixtures were assembled. The molecular descriptors (properties) of the penetrants were calculated using two software packages of ACD labs/LogD Suite (7.0.5 release) and TSAR 3D (Accelrys Ltd. version 3.3). The molecular descriptors included octanol/water partition coefficient, molecular connectivity indexes, quantum molecular descriptors, and various atom and group counts. The physicochemical properties of mixture components including boiling point, melting point, solubility, vapour

Table 2
Penetrants.

Atrazine	Pentachlorophenol
Chlorpyrifos	Phenol
Ethylparathion	p-Nitrophenol
Fenthion	Propazine
Methylparathion	Simazine
Nonylphenol	Triazine
Caffeine	Octanol
Codeine	Testosterone

pressure and Henry's law constant were obtained through ChemBioFinder (CambridgeSoft, 2009) online software and SRC PhysProp database (Syracuse Research Corporation, 2009). Hildebrand solubility parameters (δ) were obtained from Hansen (1967) for the solvents and calculated according to Fedors group contribution method (1974) for the penetrants. As there was a mixture of a number of solvents in the vehicles, averages of physicochemical properties for solvent mixtures were calculated using the fractions of each component.

Stepwise regression analysis was performed between log *kp* as the dependant variable and the molecular descriptors of the penetrants and the mixture components as the predictors. This enabled the identification of the significant molecular descriptors affecting skin penetration of chemicals. Several stepwise regression analyses using various sets of penetrant molecular descriptors and solvent properties were performed and several regression models were generated. In order to minimise the risk of chance correlations, the number of descriptors in the regression models was limited to four.

The models were validated for penetrants using a leave-many-out cross validation procedure. To do this, the penetrants were divided into four groups with similar ranges of lipophilicity (log *P* values) in each group. Regression analyses were performed four times, each time leaving one group out. The log *kp* values of the test sets were estimated using the equations obtained for the training sets and the mean absolute error was calculated from the difference between the observed and the predicted log *kp* values of the test sets.

3. Results and discussion

Skin penetration of drugs is controlled by the molecular structures and physicochemical properties of the intended penetrants and the mixture ingredients in the vehicle. In order to rationalize the combined effect of structural characteristics of the penetrants and the physicochemical properties of the mixture components, this investigation focused on the QSAR model development for a dataset of skin permeation of chemicals dissolved into a combination of several solvents, surfactant and methyl nicotinic acid. Permeation coefficients were measured for four compounds that were rationally selected in order to add a high level of diversity to the existing dataset (Ghafourian et al., 2010). Tables 1 and 2 provide the list of the vehicles and the permeants, respectively. The *kp* data measured in this investigation ($n = 96$) was merged with the previously obtained dataset of *kp* ($n = 288$) and the resulting dataset was used for the QSAR development (Riviere and Brooks, 2010).

Stepwise regression analysis of different combinations of solvent properties and molecular descriptors of the penetrants resulted in a number of QSAR models from which four were selected based on the goodness of fit (R^2 values). In order to reduce the risk of chance correlations, only four descriptors were allowed in the equations. The selected equations have been listed in Table 3. In Eqs. (1)–(4), the letter in the brackets indicates if the variable is a descriptor for the penetrant (*P*) or for the vehicle (*V*). It can be seen that each equation consists of 2–3 penetrant descriptors and 1–2 vehicle descriptors, with Eqs. (1)–(3) containing 1 combined vehicle-penetrant descriptor. In Eqs. (1)–(4), Δmp is the difference between the melting point of the penetrant and that of the solvent, *W* is the Wiener topological index (the sum of distances between all pairs of vertices in the molecular graph of an alkane (Diudea and Gutman, 1998)), δ is the Hildebrand solubility parameter, E_{HOMO} is the energy of the highest occupied molecular orbital, BP is the boiling point, N_{atoms} is the total number of atoms in the molecules, BP – MP is the difference between the boiling and melting points of a compound, and Lipole is the total lipole moment of the penetrants.

Considering that N_{atoms} and *W* (Diudea and Gutman, 1998) can be regarded as size descriptors, it can be seen from Table 3 that all QSAR models indicate the negative effect of the penetrant's molecular size on the log *kp*. Moreover, there is a negative contribution by total lipole of the penetrants in Eqs. (1)–(4). Total lipole is a measure of lipophilicity distribution calculated as sum of local values of log *P*, like dipole moment (Pedretti et al., 2002). It shows lipophilicity of the molecule in a specific direction. Surfactants are expected to have high total lipole values and they are known enhancers of drug skin penetration (Ma et al., 2007). Thus, according to these equations, the less lipolar penetrants will have higher permeation rates. Chlorpyrifos has the highest total lipole value of 10.0 and caffeine has the lowest value of 0.19.

The other penetrant descriptor, which can be seen in majority of the equations, is E_{HOMO}. This molecular descriptor represents the energy of the highest occupied molecular orbital. E_{HOMO} measures the nucleophilicity of a molecule. The negative relationship of this descriptor with the logarithm of the permeation rate indicates that the electron rich neucleophilic compounds such as those containing aromatic rings are the least permeable. In Eqs. (2) and (3), the product of the penetrants' E_{HOMO} and the vehicles' solubility parameter is used. $\delta(V) \cdot E_{HOMO}(P)$ is a solvent/penetrant interaction term. This descriptor indicates that a highly nucleophilic penetrant will have a lower penetration from highly associated vehicles, i.e. those vehicles with high intermolecular interaction forces such as hydrogen bonding.

The most persistent vehicle descriptor in the QSARs is the boiling point, with a negative effect on permeation rate of chemicals. Solubility parameter is also present in some equations. Both solubility parameter and boiling point can represent the intermolecular interaction energy of the vehicle which can result from the polarity of the solvents. Therefore the negative relationship indicates a higher skin permeation rate with the less polar vehicles. Similar results have been shown previously where the permeation coefficients of highly lipophilic compounds, nicardipine and nimodipine was increased in the less polar solvent mixtures of ethanol–water

Table 3
QSAR models obtained from stepwise regression; *N* is the number of datapoints (penetrant/vehicle combinations); *S*, the standard deviation; R^2, the squared correlation coefficient.

		N	*S*	R^2
$\log kp = -0.956 - 0.00322\,\Delta mp - 0.000320\,W(P) - 0.0121\,BP(V) - 0.114\,Lipole(P)$	(1)	384	0.478	0.701
$\log kp = -310 - 0.000315\,W(P) - 0.00771\,\delta(V) \cdot E_{HOMO}(P) - 0.0102\,BP(V) - 0.0750\,Lipole(P)$	(2)	384	0.494	0.681
$\log kp = -2.48 - 0.0474\,N_{atoms}(P) - 0.00798\,\delta(V) \cdot E_{HOMO}(P) - 0.0102\,BP(V) - 0.0723\,Lipole(P)$	(3)	384	0.516	0.653
$\log kp = -4.29 - 0.0474\,N_{atoms}(P) - 0.00904\,BP - MP(V) - 0.345\,E_{HOMO}(P) - 0.0790\,Lipole(P)$	(4)	384	0.522	0.644

Table 4
Statistical parameters obtained from internal validation of QSAR equations (1)–(4); N is the number of datapoints, S is the standard deviation and R^2 is the squared correlation coefficient between observed and predicted log kp for the test sets, and MAE is mean absolute error of prediction.

Equation	N	S	R^2	MAE
(1)	384	0.557	0.592	0.454
(2)	384	0.594	0.535	0.496
(3)	384	0.605	0.517	0.497
(4)	384	0.618	0.497	0.493

(Krishnaiah et al., 2002; Krishnaiah et al., 2004). In the case on nimopidine, 60:40 (v/v) ratio of ethanol:water was an optimum solvent mixture leading to the highest permeation rate of nimopidine, with the kp dropping slightly at higher ethanol concentrations (Krishnaiah et al., 2004). Solubility parameters of the solvents and the permeants have been implicated as important factors controlling skin penetration of compounds (Dias et al., 2007; Roy and Flynn, 1989). From the parabolic relationship between skin permeation rate and the solubility parameter it has been concluded that the skin has a solubility parameter of around 10 $(cal/cm^3)^{1/2}$ (Liron and Cohen, 1984). In Eqs. (2)–(4), solubility parameter of the vehicle has been selected by stepwise regression analysis as a significant contributor to skin permeation rate. The solvent mixtures in this study have solubility parameters of >12. Therefore, it is expected that the lower the δ value of the vehicle, the closer the value to the skin δ and therefore the higher the permeation constant should be. In Eqs. (2) and (3), $\delta(V)\cdot E_{HOMO}(P)$ has a negative effect on the skin absorption, implying a lower skin absorption of nucleophilic drugs from polar solvents, as explained before.

In Eq. (1), the difference between melting points of the vehicle and penetrant has been selected as the most significant of all descriptors. The negative coefficient of the descriptor Δmp indicates that the melting point of the penetrants should be close to the melting point of the vehicle for a better skin absorption. Therefore, since the vehicles are all liquids, this implies that penetrants with low melting points are likely to have a higher absorption rate. This finding is in agreement with a previous observation where it has been shown for two optical isomers of ibuprofen that the low melting point S enantiomer has a higher skin permeation rate than the high melting point R enantiomer (Cilurzo et al., 2010). A similar conclusion has been made in a different investigation, when it was observed that among the alkyl analogues of cyclizine the analogue with the lowest melting point, and not the most lipophilic one, showed the highest skin penetration rate (Monene et al., 2005). Melting point has a similar effect on the intestinal absorption of drugs with low melting point drugs generally showing a higher fraction of dose absorbed from GI tract (Chu and Yalkowsky, 2009).

In Eq. (4), the descriptor BP – MP(V) with negative coefficient indicates that penetration rate is slower from solvents with large boiling and meting point gaps. The difference between these two properties has been attributed to the molecular symmetry, with highly symmetrical molecules having much larger melting points and decreased boiling points (Slovokhotov et al., 2007). In the solvents used in this study, the biggest difference in melting and boiling points is for propylene glycol. Therefore the vehicles containing higher concentrations of this solvent will have higher difference between melting and boiling points, leading to lower penetration of the penetrants.

Table 3 shows that Eq. (1) with the highest R^2 and the lowest S value has the best fit to the data. The goodness of fit is reduced from Eqs. (1) to (4). The predictive powers of the equations were tested by an internal validation procedure explained in the methods section. Table 4 shows the statistical parameters obtained from this exercise. It must be noted that during validation tests, each set of

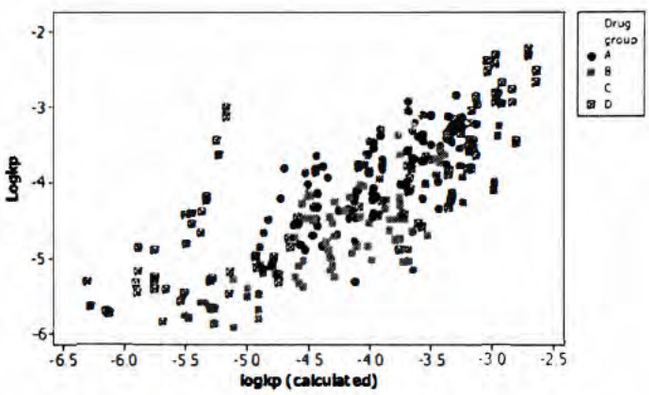


Fig. 1. Scatterplot of observed log kp versus log kp calculated by Eq. (1).

four penetrants (dissolved in any solvent mixture) were removed once as the test set and the equation obtained for the remaining 12 penetrants were used for the estimation of the log kp of the test set. Thus, Table 4 represents the results of such log kp estimations for all the four test sets (each set containing four penetrants). According to the table, the predictivity of Eq. (1) is the highest among all the equations with a mean absolute error of 0.454. Eqs. (2)–(4) show slightly higher prediction errors for the test sets, with MAE increasing in the order of Eq. (4) > Eq. (2) > Eq. (3). Therefore, it appears that the QSAR model 1 is robust in terms of prediction of log kp for those new penetrants that fall within the applicability domain of the model. Fig. 1 is the plot of observed versus predicted log kp values using Eq. (1). The outliers in the graph with underestimated log kp values are testosterone dissolved in different vehicles, most notably in ethanol–water, ethanol–water–methyl nicotinate, water and water–methyl nicotinate. Testosterone is the largest molecule in the dataset in terms of the total number of atoms and the topological index W, where the negative coefficient of W in Eq. (1) describes the negative effect of the molecular size on absorption. Testosterone is also the most affected penetrant by the variation in solvent mixture, as it presents the largest gap between the highest and the lowest kp from different mixture components. Considering that only kp values from certain mixtures components is higher than expected, it is concluded that the enhancing effects of these vehicles are not fully accounted for by the model and further investigations are required to explain the observed effect.

References

Barry, B.W., 2007. Transdermal drug delivery. In: Aulton, M.E. (Ed.), Aulton's Pharmaceutics, The Design and Manufacture of Medicines, third ed. Elsevier, Churchill Livingstone, pp. 580–585.
CambridgeSoft, (accessed 09.09), http://www.cambridgesoft.com/databases, 2009.
Chu, K.A., Yalkowsky, S.H., 2009. An interesting relationship between drug absorption and melting point. Int. J. Pharm. 373, 24–40.
Cilurzo, F., Alberti, E., Minghetti, P., Gennari, C.G.M., Casiraghi, A., Montanari, L., 2010. Effect of drug chirality on the skin permeability of ibuprofen. Int. J. Pharm. 386, 71–76.
Dias, M., Hadgraft, J., Lane, M.E., 2007. Influence of membrane–solvent–solute interactions on solute permeation in skin. Int. J. Pharm. 340, 65–70.
Diudea, M.V., Gutman, I., 1998. Wiener-type topological indices. Croat. Chem. Acta 71, 21–51.
Elias, P.M., 1983. Epidermal lipids, barrier function and desquamation. J. Invest. Dermatol. 80, 44–49.
El Tayar, N., Tsai, R.-S., Testa, B., Carrupt, P.-A., Hansch, C., Leo, A., 1991. Percutaneous penetration of drugs: a quantitative structure–permeability relationship study. J. Pharm. Sci. 80, 744–749.
Fedors, R.F., 1974. Methods for estimating both solubility parameters and molar volumes of liquids. Polym. Eng. Sci. 14, 47–154.
Flynn, G.L., 1990. Physicochemical determinants of skin absorption. In: Gerrity, T.R., Henry, C.J. (Eds.), Principles of Route to Route Extrapolitation for Risk Assessment Extraction for Skin Assessment. Elsevier, New York, pp. 93–127.

Ghafourian, T., Zandasrar, P., Hamishekar, H., Nokhodchi, A., 2004. The effect of penetration enhancers on drug delivery through skin: a QSAR study. J. Control. Release 99, 113–125.

Ghafourian, T., Samaras, E.G., Brooks, J.D., Riviere, J.E., 2010. Modelling the effect of mixture components on permeation through skin. Int. J. Pharm. 398, 28–32.

Hansen, C.M., 1967. The three dimensional solubility parameter—key to paint component affinities. I. Solvents, plasticizers, polymers and resins. J. Paint Technol. 39, 104–117.

Kai, T., Mak, V.H.W., Potts, R.O., Guy, R.H., 1990. Mechanisms of percutaneous penetration enhancement: effect of n-alkanols on the permeability barrier of hairless mouse skin. J. Control. Release 12, 103–112.

Krishnaiah, Y.S.R., Bhaskar, P., Satyanarayana, V., 2004. Penetration-enhancing effect of ethanol–water solvent system and ethanolic solution of carvone on transdermal permeability of nimodipine from HPMC gel across rat abdominal skin. Pharm. Dev. Technol. 9, 63–74.

Krishnaiah, Y.S.R., Satyanarayana, V., Karthikeyan, R.S., 2002. Effect of the solvent system on the in vitro permeability of nicardipine hydrochloride through excised rat epidermis. J. Pharm. Pharmaceut. Sci. 5, 123–130.

Liron, Z., Cohen, S., 1984. Percutaneous absorption of alkanoic acids. II. Application of regular solution theory. J. Pharm. Sci. 73, 538–542.

Ma, Q.H., Hao, X.Z., Zhou, H.F., Gu, N., 2007. Effect of surfactants on preparation and skin penetration of tea polyphenols liposomes. In: IEEE/ICME International Conference on Complex Medical Engineering, vols. 1–4, pp. 209–212.

Monene, L.M., Goosen, C., Breytenbach, J.C., Hadgraft, J., du Plessis, J., 2005. Percutaneous absorption of cyclizine and its alkyl analogues. Eur. J. Pharm. Sci. 24, 239–244.

Moss, P.G., Dearden, J.C., Patel, H., Cronin, M.T.D., 2002. Quantitative structure–permeability relationships (QSPRs) for percutaneous absorption. Toxicol. In Vitro 16, 299–317.

Pedretti, A., Villa, L., Vistoli, G., 2002. Modeling of binding modes and inhibition mechanism of some natural ligands of farnesyl transferase using molecular docking. J. Med. Chem. 45, 1460–1465.

Pugh, W.J., Roberts, M.S., Hadgraft, J., 1996. Epidermal permeability—penetrant structure relationships. 3. The effect of hydrogen bonding interactions and molecular size on diffusion across the stratum corneum. Int. J. Pharm. 138, 149–165.

Raykar, P.V., Fung, M.C., Anderson, B.D., 1988. The role of protein and lipid domains in the uptake of solutes by human stratum corneum. Pharm. Res. 5, 140–150.

Riviere, J.E., Brooks, J.D., 2010. Chemical Mixture Absorption Dataset., http://www.lib.ncsu.edu/resolver/1840.2/2297.

Riviere, J.E., Brooks, J.D., 2007. Prediction of dermal absorption from complex chemical mixtures: incorporation of vehicle effects and interactions into a QSPR framework. SAR QSAR Environ. Res. 18, 31–44.

Riviere, J.E., Brooks, J.D., 2005. Predicting skin permeability from complex chemical mixtures. Toxicol. Appl. Pharmacol. 208, 99–100.

Roberts, M., Cross, S., Pellet, M., 2002. Skin transport. In: Walters, K.A. (Ed.), Dermatological and Transdermal Formulations. Marcel Dekker, New York, pp. 89–195.

Rosado, C., Cross, S.E., Pugh, W.J., Roberts, M.S., Hadgraft, J., 2003. Effect of vehicle pretreatment on the flux, retention, and diffusion of topically applied penetrants in vitro. Pharm. Res. 20, 1502–1507.

Roy, S.D., Flynn, G.L., 1989. Transdermal delivery of narcotic analgesics: comparative permeabilities of narcotic analgesics through human cadaver skin. Pharm. Res. 6, 825–832.

Shah, V.P., 1994. Skin penetration enhancers: scientific perspectives. In: Hsieh, D.S. (Ed.), Drug Permeation Enhancement Theory and Applications. Marcel Decker, New York, pp. 19–24.

Slovokhotov, Y.L., Batsanov, A.S., Howard, J.A.K., 2007. Molecular van der Waals symmetry affecting bulk properties of condensed phases: melting and boiling points. Struct. Chem. 18, 477–491.

Syracuse Research Corporation, 2009. SRC PhysProp Database (retrieved 20.09.09) http://www.syrres.com/what-we-do/databaseforms.aspx?id=386.

Williams, A.C., Barry, B.W., 2004. Penetration enhancers. Adv. Drug Deliv. Rev. 56, 603–618.

Wilschut, A., Berge, W.F., Robinson, P.J., McKone, T.E., 1995. Estimating skin permeation. The validation of five mathematical skin permeation models. Chemosphere 30, 1275–1296.

Zinke, S., Gerner, I., Schlede, E., 2002. Evaluation of a rule base for identifying contact allergens by using a regulatory database: comparison of data on chemicals notified in the European Union with with 'structural alerts' used in the DEREK expert system. Altern. Lab. Anim. 30, 285–298.