VIRTUAL DESIGN OF CHEMICAL PENETRATION ENHANCERS

By

SHARATH GOLLA

Bachelor of Engineering (Honors)

Birla Institute of Technology and Science

Pilani, Rajasthan (India)

May 2005

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
July, 2008

VIRTUAL DESIGN OF CHEMICAL

PENETRATION ENHANCERS

Thesis Approved:

Dr. Khaled A.M. Gasem
Thesis Adviser

Dr. Robert L. Robinson, Jr.

Dr. Sundar Madihally

Dr. A. Gordon Emslie
Dean of the Graduate College

ACKNOWLEDGEMENTS

I thank all my teachers, colleagues, and friends who have been associated with me throughout the course of my graduate study. My sincere thanks to Dr. Khaled A. M. Gasem for providing me with this learning opportunity and guiding me through it. I would also like to thank my committee members, Dr. Robert L. Robinson, Jr. and Dr. Sundar Madihally for their guidance and valuable suggestions. Special thanks go to Dr. Brian Neely and Eric Whitebay for their help. Finally, I would like to thank my family members for all their support.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

Table Page

CHAPTER 4 QUANTITATIVE STRUCTURE-PROPERTY RELATIONSHIPS MODELING OF SKIN IRRITATION

LIST OF FIGURES

Figure Page

Figure Page

CHAPTER 1

INTRODUCTION

Traditional drug delivery techniques, such as oral or intravenous administration, are often associated with problems relating to over- and under-dosing, interactions with the harsh gastro-intestinal environment, and/or the production of toxic by-products through metabolism in the liver. Often, a large dose of a drug is required to attain therapeutic levels of the drug in the blood plasma, which may result in toxicity to other organs [1]. An alternate delivery technique which offers improved therapeutic control (both temporal and spatial) is required.

Recently, the technique of transdermal drug delivery (TDD) has gained popularity due to its ability to overcome most of the above problems. Transdermal technologies are employed for delivering a variety of therapeutic drugs. Currently, over 35 FDA-approved transdermal products are available for applications, including hormone replacement therapy, management of pain, angina, smoking cessation and neurological disorders such as Parkinson's diseases. Thus, a significant health benefit is derived from TDD and, consequently, there is a sizeable market potential for transdermally delivered therapeutic agents. For example, in 2005, the expected annual sales worldwide are $12.7 billion, which are projected to increase to $21.5 billion in 2010 and to $31.5 billion by 2015 [2].

The goal of TDD is to maximize the rate of transport of the therapeutic agent through the skin and into systemic circulation, while minimizing the retention and

metabolism of the drug in the skin [3]. Human skin is considered to be one of the most efficient natural polymers. It serves as a barrier to transport of chemicals both in and out of the human body [4, 5]. Figure 1.1 illustrates the cross section of human skin. Human skin is composed of three layers (a) the avascular dermis composed of metabolically active basal cells and the inactive stratum corneum (SC), (b) the vascular dermis consisting of the blood vessels, and (c) the subcutaneous tissue consisting of adipocytes, hair follicles, sweat and sebaceous glands. Each of these layers offers varying resistances to drug permeation [6, 7]. Several physical and chemical alternatives are currently being investigated for possible improvement of TDD. However, the economic viability and technical feasibility of using chemicals as penetration enhancers (CPEs) makes them the most attractive option [8].

The basic requirement of TDD is that the drug penetrates the SC and is absorbed into systemic circulation. In general, the drug has two potential routes of entry: (a) across the SC, and (b) through hair follicles and sweat ducts, as shown in Figure 1.2. To permeate across the SC, the drug must first partition into the SC and then diffuse through the protein-lipid matrix. The drug then diffuses through the epidermis into systemic circulation. Permeation through hair follicles and sweat ducts involves diffusion through the pores and epidermis into the systemic circulation. Hair follicles and sweat ducts occupy only a small fraction of the total skin area and are believed to be insignificant factors in TDD [9]. Some recent studies [10], however, suggest that they may be a significant pathway for large polar molecules.

Both physical and chemical methods have been proposed and implemented to increase permeation of drugs across the skin [11]. Physical approaches such as

iontophoresis [12] and sonophoresis [13, 14] are still at their inception and require further research before commercialization. One delivery mechanism that has been studied extensively and implemented commercially is the use of CPEs. Based on the mechanism of action, CPEs can be grouped into two classes: (a) chemicals that alter the structure of the skin lipids, and (b) chemicals that enhance the solubility of the drug in the skin lipids. Numerous CPEs have been identified and evaluated; however, none has proved to be truly effective or universally applicable [15]. Development of CPEs requires detailed analysis of several interrelated factors, including (a) structure and properties of human skin, (b) thermophysical properties of the penetration enhancer, and (c) the properties of the penetrant. Williams [16] provides a detailed review on CPEs and their desirable thermophysical attributes.

The current experimental techniques used for CPE development are time-consuming and expensive. An attractive alternative (widely used for drug design) is "virtual synthesis," in which structure-based QSPR models are coupled with powerful screening algorithms to identify viable drug molecules. However, the existing virtual screening methodology suffers from several limitations, including: (a) reliance on linear models in QSPR model development, (b) absence of a theoretical framework in the models used to describe thermophysical properties, (c) use of only off-the-shelf structural descriptors, (d) use of general-purpose heuristic algorithms for molecular screening, and (e) inadequate data for model development and testing. These combined limitations hinder progress toward effective virtual design algorithms for CPEs.

Objectives

The primary goal of our research is to integrate non-linear quantitative-structure-property-relationship (QSPR) modeling and robust genetic algorithms (GAs) to facilitate the design of improved CPEs. The specific objectives for accomplishing this goal were to:

1. Identify the thermophysical properties pertinent in CPE design, and assemble reliable QSPR models for these properties.
2. Develop improved QSPR models for skin sensitization and skin irritation using advanced non-linear modeling.
3. Develop GA algorithms for generating new potential CPEs.
4. Incorporate the property models and the GA algorithms into an effective platform for the virtual design of CPEs.

Our basic premise is that novel, effective mathematical models can be developed to describe accurately the relationship between a molecular structure of a chemical and its CPE behavior, and that these models can form the basis for the "virtual design" of promising molecular structures for use as CPEs. The innovative integration of non-linear, theory-based QSPR modeling and robust GAs removes existing barriers to the use of computational chemistry in CPE design, and yields structure-based models to delineate the specific structural features of CPEs that are responsible for improved permeation of a drug through skin. The scientific knowledge gained in developing the models will be significant in drug development and therapeutic agent delivery design.

Thesis Organization

This thesis is written in the "manuscript style" and is divided into three separate self-contained manuscripts. Since the same modeling strategy was adopted for skin sensitization and skin irritation, some sections of the individual chapters are repetitive. Also, the modeling methodology used in this study has been developed in collaboration with other members of the OSU Thermodynamics Research Group [17-19]. Consequently, similar documentation has been used. Chapters 2 deals with computer-aided molecular design of CPEs using GAs and QSPR models, and Chapters 3 and 4 deal with the methodology adopted and the results obtained for modeling skin sensitization and skin irritation, respectively. Conclusions based on the efforts undertaken are given in each of these chapters.

REFERENCES

[1] S. Chong and H. L. Fung, "Transdermal drug delivery systems: pharmacokinetics, clinical efficacy, and tolerance development," *Transdermal Drug Delivery: Developmental Issues and Research Initiatives,* pp. 135-153, 1989.

[2] M. Karablylkoglu, "New frontiers in transdermal drug delivery systems," in *PharmaVentures Ltd.,* Oxford Science Park, Oxford, UK, 2007.

[3] B. C. Finnin and T. M. Morgan, "Transdermal penetration enhancers: Applications, limitations, and potential," *Journal of Pharmaceutical Sciences,* vol. 88, pp. 955-958, 1999.

[4] N. A. Monteiro-Riviere, "Anatomical factors affecting barrier function," in *Dermatotoxicology*, 7 ed: CRC Press, 2008.

[5] N. A. Monteiro-Riviere, "Comparative anatomy, physiology, and biochemistry of mammalian skin," *Dermal and Ocular Toxicology: Fundamentals and Methods,* pp. 3–71, 1991.

[6] P. M. Elias, "Epidermal lipids, barrier function, and desquamation," *Journal of Investigative Dermatology,* vol. 80, pp. 44-49, 1983.

[7] S. K. Chang and J. E. Riviere, "Percutaneous absorption of parathion *in vitro* in porcine skin: Effects of dose, temperature, humidity, and perfusate composition on absorptive flux," *Toxicological Sciences,* vol. 17, pp. 494-504, 1991.

[8] R. H. Guy, "Current status and future prospects of transdermal drug delivery," *Pharmaceutical Research,* vol. 13, pp. 1765-1769, 1996.

[9] J. E. Wahlberg, "Transepidermal or transfollicular absorption? *In vivo* and *in vitro* studies in hairy and non-hairy guinea pig skin with sodium (22Na) and mercuric (203Hg) chlorides," *Acta Dermato-Venereologica,* vol. 48, pp. 336-344, 1968.

[10] B. W. Barry, "Mode of action of penetration enhancers in human skin," *Journal of Controlled Release,* vol. 6, pp. 85-97, 1987.

[11] B. W. Barry, "Novel mechanisms and devices to enable successful transdermal drug delivery," *European Journal of Pharmaceutical Sciences,* vol. 14, pp. 101-114, 2001.

[12] R. H. Guy, "Iontophoresis: Recent developments," *Journal of Pharmacy and Pharmacology,* vol. 50, pp. 371-374, 1998.

[13] H. Tang, S. Mitragotri, D. Blankschtein, and R. Langer, "Theoretical description of transdermal transport of hydrophilic permeants: Application to low-frequency sonophoresis," *Journal of Pharmaceutical Sciences,* vol. 90, pp. 545-568, 2001.

[14] S. S. Mitragotri, D. Blankschtein, and R. S. Langer, "Transdermal protein delivery using low-frequency sonophoresis," US Patent 6,002,961, 1999.

[15] P. Karande, A. Jain, and S. Mitragotri, "Discovery of transdermal penetration enhancers by high-throughput screening," *Nature Biotechnology,* vol. 22, pp. 192-197, 2004.

[16] A. C. Williams and B. W. Barry, "Penetration enhancers," *Advanced Drug Delivery Reviews,* vol. 56, pp. 603-618, 2004.

[17] S. S. Godavarthy, "Design of improved solvents for extractive distillation," Ph. D. Dissertation, Oklahoma State University, Stillwater, OK, 2004.

[18] S. S. Godavarthy, R. L. Robinson, and K. A. M. Gasem, "Improved structure–property relationship models for prediction of critical properties," *Fluid Phase Equilibria,* vol. 264, pp. 122-136, 2008.

[19] D. Ravindranath, B. J. Neely, R. L. Robinson, and K. A. M. Gasem, "QSPR generalization of activity coefficient models for predicting vapor–liquid equilibrium behavior," *Fluid Phase Equilibria,* vol. 257, pp. 53-62, 2007.

[20] "http://www.ahdesign-online.net," October 24, 2007.

[21] B. W. Barry and A. C. Williams, "Permeation enhancement through skin," *Encyclopedia of Pharmaceutical Technology,* vol. 11, pp. 449–493, 1995.

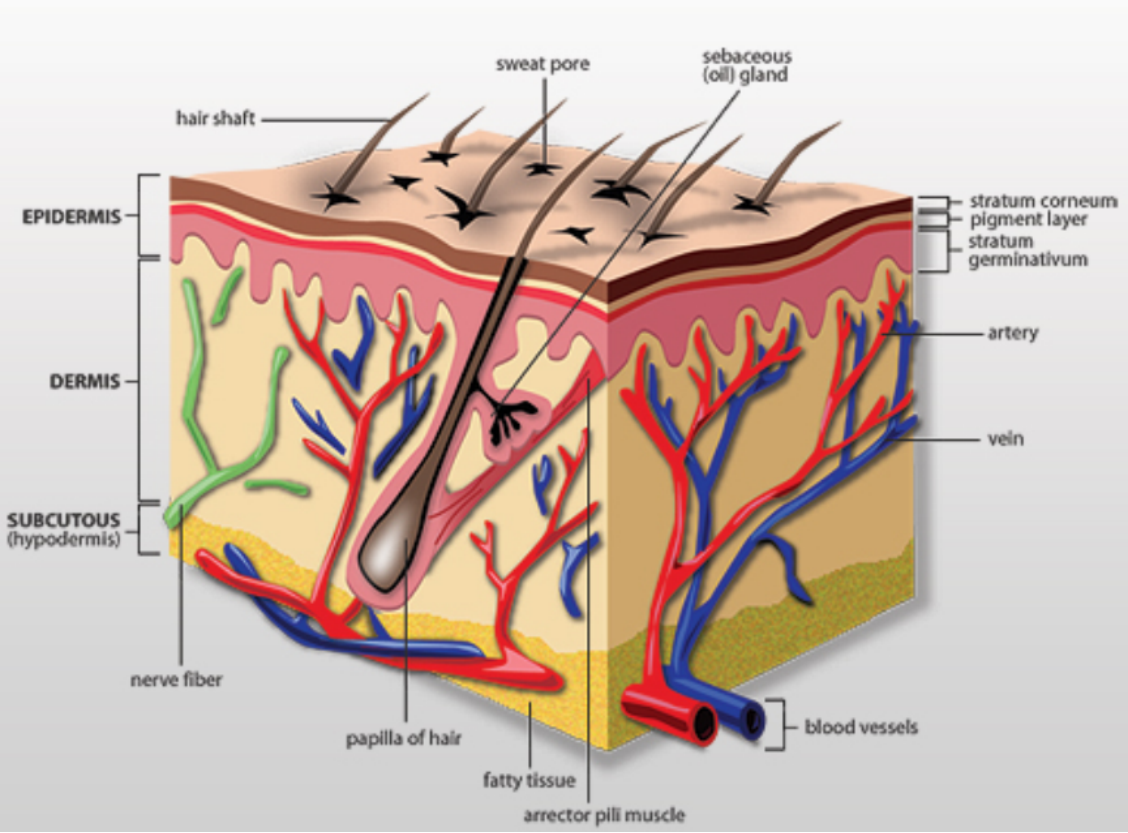


Figure 1.1. Human skin cross section [20]

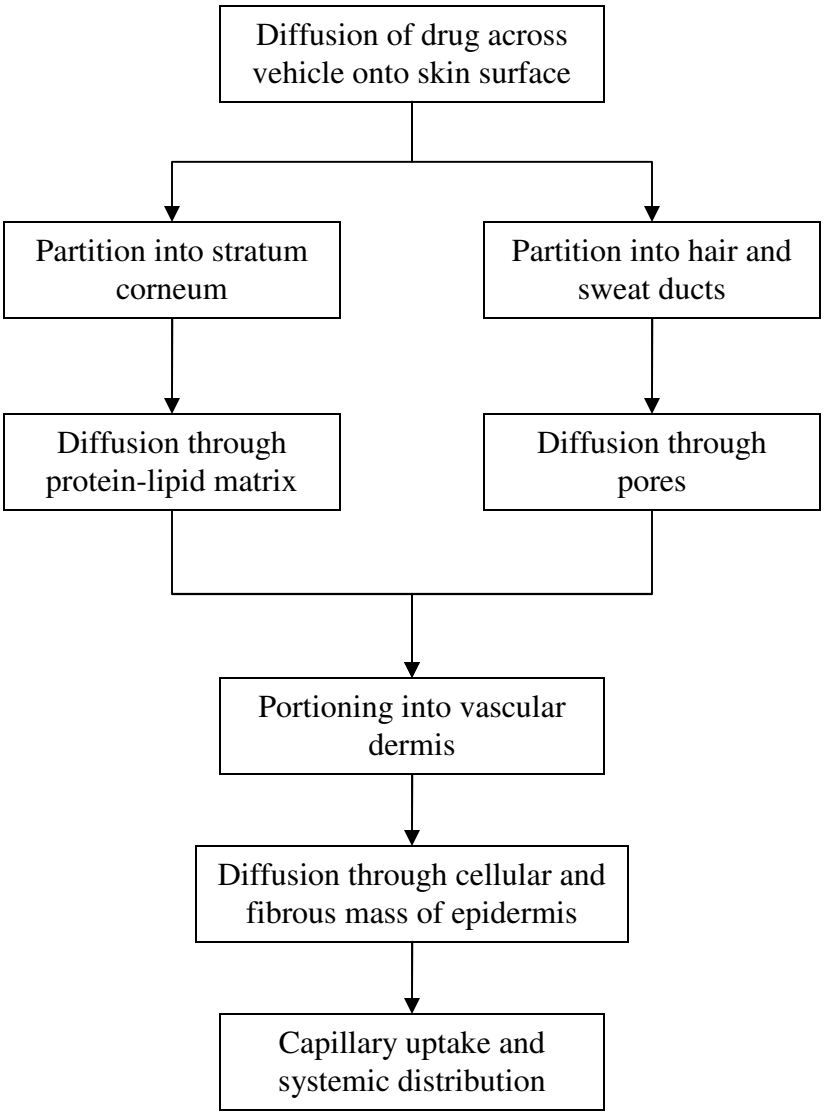


Figure 1.2. Potential routes for percutaneous absorption of drugs [21]

CHAPTER 2

VIRTUAL DESIGN OF CHEMICAL PENETRATION ENHANCERS

2.1. INTRODUCTION

The rational design of molecules with desired properties challenges engineers attempting to meet the needs of various industries, including pharmaceuticals, polymers, petrochemicals, and construction [1-3]; especially, since the demand for newly-designed molecules that enhance current processes and satisfy more stringent operating requirements in technology has been increasing [4]. The traditional approach of discovering molecules with desired properties involves testing thousands of molecules for their chemical and physical properties, which is an expensive and laborious undertaking. Hence, rational design techniques, such as computer-aided molecular design (CAMD), have found wide application in recent years [4, 5].

In contrast to traditional methodologies, CAMD methods expedite the design process by predicting the behavior of potential molecules using reliable property models. CAMD involves the design of new molecules based on a set of desired properties and can be classified as (a) forward CAMD, which involves computation of chemical, physical and biological properties from the molecular structure, and (b) inverse CAMD, which involves generation of a molecular structure with the desired properties [6, 7]. In pharmaceutical processes, CAMD is used to identify new drugs useful for targeted applications, while meeting design constraints such as minimal side effects and toxicity.

CAMD methods have been employed successfully for a large range of applications, including solvent design/selection [8], chloro-fluro-carbon (CFC) substitutes, alternative process fluids, polymer design [9], drug design [10], and design for novel molecules with superior properties [11]. A typical CAMD design algorithm features two key components, a method for generating candidate molecules and the models used to predict the pertinent physiochemical properties of the newly generated molecules. Although genetic algorithms (GA) have been used extensively for generating new molecular structures from seed molecules, the use of a large molecular database to identify potential molecules has also been reported [12, 13]. Property predictions of the generated molecules are usually done using group contribution methods, equation-of-state approaches, and quantitative structure-property relationship (QSPR) models. Figure 2.1 presents the various stages involved in CAMD, in general.

In this work, a combination of genetic algorithms and QSPR techniques has been used to develop the CAMD algorithm for virtual design of chemical penetration enhancers (CPEs) for transdermal drug delivery. Extensive efforts have been expended in search of chemicals that enhance the penetration of therapeutic drugs through human skin. Although such CPEs can be valuable in increasing the amount and/or rate of drug delivery, they can also have undesirable effects, including skin irritation and toxicity. Thus, a distinct need exists for effective methods to identify new CPEs that provide optimum penetration enhancement with minimal side effects.

The primary goal of this work is the integration of non-linear, QSPR modeling and robust GAs to facilitate the rational design of improved CPEs. Our basic premise is that novel, effective mathematical models can be developed to describe accurately the

relationship between a chemical's molecular structure and its CPE behavior, and that these models can form the basis for the "virtual design" of promising molecular structures for use as CPEs. Ultimate benefits of such a design capability include: (a) identifying novel CPEs; (b) reducing the need for expensive and time-consuming experiments; and (c) setting the stage for the synthesis and commercialization of improved CPEs for use by the medical community.

The work reported here proceeded in distinct stages, as described below. To begin, the target properties for design of CPEs were identified through a thorough literature survey and analysis of their molecular properties. Using a database of over 250 CPEs cited in the literature as seed molecules, new molecules were generated using genetic operators such as crossover, mutation and functional group addition. QSPR models developed using artificial neural networks (ANNs) were used to predict the physiochemical properties such as skin penetration coefficient, octanol/water partition coefficient, melting point, skin sensitization and skin irritation of the newly generated molecules. The molecules were scored and screened before being passed to the next generation. To further validate the design methodology results, all identified potential CPEs were tested for toxicity and skin permeation through carefully designed experimental techniques, as detailed elsewhere [14, 15].

2.2. COMPUTER-AIDED MOLECULAR DESIGN (CAMD) METHODS

The traditional sequential method of molecular discovery for developing new chemicals requires expensive chemical synthesis followed by time-consuming experimental thermophysical property measurements. Often several hundred (and in the

case of drug design, several thousand) new molecules may be tested before a viable chemical is identified. An attractive solution to these development problems is the use of virtual screening wherein (a) the physical synthesis of molecules is replaced by virtual synthesis, (b) the experimental property measurements are minimized through the use of accurate property prediction models, and (c) robust scoring modules guide the virtual screening algorithms toward the most feasible subset of molecules. The complexity of CAMD problems differ based on the targeted application and the computational techniques used.

Several approaches for CAMD design have been proposed for diverse applications that vary in their solution strategy, complexity and the range of properties considered. A multi-step and multi-level approach for CAMD was presented by Harper and Gani [16]. A three step approach involving problem formulation (pre-design), compound identification (design) and result analysis (post-design) was proposed to be effective in CAMD. The molecules are screened using macroscopic representation of molecules and the selected molecules are further screened using microscopic representation. Application examples, such as design of a replacement solvent for liquid-liquid extraction of phenol from wastewater and design of a benzene replacement, were presented to illustrate the application and efficacy of this approach. To overcome the limitations of group contribution methods for property predictions in CAMD, a new method using a combination of multi-level approach for molecular generation and property predictions using connectivity indices, fragments, and mixed methods has been proposed [5]. Cabezas [17] developed a CAMD technique that searches through a database of compounds to identify solvent molecules with desired properties, which

results in identification of molecules that already exist. Pretel *et al.* [18] employed a group contribution molecular design approach to synthesize molecular structures with desired solvent properties. The size of the combinatorial problem posed by the molecular synthesis procedure is reduced by group selection and physical and molecular constraints at different stages. Constantinou *et al.* [19] employed a group contribution approach to generate acyclic, cyclic and aromatic compounds of various degrees of complexity and size. Feasible solutions were found in each of the five case studies used to evaluate the efficacy of the CAMD technique.

CAMD techniques have been implemented successfully by our research group for more than a decade to design solvents for extractive distillation [8, 20]. Our third-generation chemical design algorithms developed for design of proprietary solvents [21] should be effective in CPE design, once calibrated properly for this application. A CAMD problem typically involves the following steps as proposed by Gani and coworkers [11, 16, 19, 22] and are described in greater detail in the following sections:

- Problem formulation – The target physiochemical properties and their desired values are determined. To design chemical compounds that enhance skin penetration, properties that affect the enhancement capability of a chemical are identified.
- Initial search – The list of molecules identified as potential CPEs in the literature are identified and introduced into the CAMD algorithm as parent molecules in the first generation. Thus, the genetic material that constitutes a good CPE can be provided to the algorithm at its inception.

- Molecule generation and testing – Using the list of candidate molecules, new molecules are generated and tested. The selected CAMD technique should be able to generate candidate molecules and evaluate the properties of the generated molecules.
- Verification – The efficacy of the selected molecules is confirmed through experimental validation.

2.2.1 Problem formulation

Identifying the desired target properties of the chemical compounds to be generated is the defining step in CAMD processes. A knowledge-based system is required to identify target properties, as well as their corresponding property values. Knowledge-based systems that specify the target properties and the acceptable values of the properties for solvent design have been extensively discussed [23]. Only a few systems that discuss problem formulation for novel drug design exist. Hence, the need for developing knowledge-based systems for novel drug discovery arises.

Lipinski's 'rule of 5' is one such expert system that predicts the solubility and permeability of the drug molecules based on four target properties [24], namely, the molecular weight, count of hydrogen bond donors, count of hydrogen bond acceptors and octanol/water partition coefficient (log K_{ow}). The World Drug Index, a large computerized database consisting of approximately 50,000 drug molecules, was used for identifying the target properties of the drug molecules. Since our target is the identification of novel potential CPEs, extensive knowledge of the properties of the CPEs and their corresponding functionalities is needed.

The target molecules should be able to enhance the permeation of a selected drug through the skin without causing any harmful effects. After thorough analysis of the currently available CPEs and their properties, the following properties were identified as being significant for transdermal drug delivery:

1. Molecular weight: Molecules with low molecular weights easily penetrate the skin due to their small size. Hence an upper limit of 500 was imposed on the molecular weight of potential CPEs [24-27].
2. Octanol/water partition coefficient (K_{ow}): Drugs with very low or high partition coefficient fail to reach systemic circulation [24, 26, 27]. Several ranges of log K_{ow} values have been proposed in the literature for effective permeation enhancement. In this work, molecules with log K_{ow} values in the range of 1-3 were accepted and considered to indicate good permeation enhancement [25].
3. Melting point: Molecules with high melting points, due to their low solubility both in water and fat, are ineffective in transdermal drug delivery (TDD) [26], and only molecules with melting points less than 200° C were accepted [25].
4. Skin sensitization: The CPE should not cause any skin irritation or sensitization upon application [25]. All the newly generated molecules are scored using three independent skin sensitization QSPR models, and only those molecules that are classified as being non-sensitizers in all three models are passed to the next generation.
5. Number of hydrogen donor groups: The sum of the hydrogen atoms linked to oxygen and nitrogen atoms in the molecule determines the total number of hydrogen-bond donor groups in a molecule. The permeability across the lipid bi-

layer has been identified to be significantly lower for drugs with an excessive number of theses groups [24, 25]. Hence, a hydrogen-bond donor number upper limit of five was specified for acceptance of a molecule as a CPE.

6. Number of hydrogen-acceptor groups: The total number of nitrogen, oxygen and fluorine atoms in the molecule (excluding nitrogen atoms with a formal positive charge, higher oxidation states and pyrrolyl forms) determines the total number of hydrogen-bond acceptor groups in a molecule. Presence of too many acceptor groups has been identified as a hindrance to the permeability across the lipid bi-layer [24], and therefore an upper limit of 10 was used for the hydrogen-bond acceptor number.

2.2.2 Initial search

The genetic material (chemical structures) identified as effective CPEs are utilized by the GA to generate new potential chemical structures. Accordingly, a thorough literature search is required to assemble available CPE data. An exhaustive literature search focused on database compilation of CPE molecules was completed by Osborne and Henke [28]. Over 400 technical and patent literature sources were reviewed, and a dataset of 275 CPE molecules was compiled. Molecules that enhance skin permeability by reversibly altering the skin or by changing the physiochemical nature of the skin were included in the database. Additives that enhance the skin penetration by altering the solubility or changing the ionization state of the drug were not classified as CPEs.

The chemical classes present in the database include fatty alcohols, fatty acids, fatty acid esters, fatty alcohol ethers, biologics, enzymes, amines, amides, complexing agents, macrocyclics, classical surfactants, pyrrolidones, ionic compounds, solvents and

azone-related compounds. In a recent article involving over 90 technical and patent literature sources, Thong *et al.* [29] studied CPE classification and mechanisms and provided a database of approximately 180 CPE molecules along with their chemical class, mechanism of action and examples of targeted drugs. The chemical classes of the CPE molecules in this database include sulfoxides, alkanones, alcohols, polyols, amides, pyrrolidones, fatty acids, fatty acid esters, surfactants, terpenes, organic acids, cyclodextrins and FDA-approved CPEs.

These two databases were studied carefully and a new database (Oklahoma State University (OSU) CPE database) consisting of over 400 CPE molecules was compiled. The molecular structures of 272 CPEs, identified using multiple software applications, were used as seed molecules in our GA approach for CPE design.

2.2.3 Molecular generation and testing

Genetic algorithms: While the desired properties and their target values, as well as the list of candidate molecules, depend on the specific CAMD application, the efficiency of the CAMD technique depends on the methods used for molecule generation and property evaluation [30]. Evolutionary techniques have been found useful in generating new molecules with desired properties. In molecular modeling, evolutionary computation involves searching for candidate molecules utilizing concepts developed in evolution and genetic science. One advantage of using evolutionary techniques is the ability to work simultaneously with a number of potential molecular candidates. Hence, the likelihood of identifying an optimum structure representing the global minimum in the search domain is high [31]. GAs introduced by Holland [32] are used widely and have proven effective as an evolutionary technique. For example, Venkatasubramanian *et al.* [1, 30] proposed

the use of GA for polymer design using CAMD. In spite of the large search space and the complex nonlinear group interactions, the genetic design methodology has been successful in the identification of target molecules.

A combination of GA and QSPR techniques has proven effective in the novel design of molecules [33]. Nachbar [6] used the GA-QSPR technique to generate molecules with desired properties, where a molecular encoding mechanism using valence states and relative probabilities for each atomic specie was used. A user interactive tool, "Molecule Evoluator," was develop by Lameijer *et al.* [34] to design drug molecules using a "TreeSMILES" representation scheme. By specifying an upper and lower bound for the descriptors, such as the number of hydrogen donors/acceptors, the molecular weight, the log P (lipophilicity), the polar surface area, the number of rotatable bonds, and the number of aromatic systems and substituents, novel potential structures were found. Similarly, Douguet *et al.* [35] developed an expert system that generates new drug molecules with the desired shape, lipophilic and electronic properties using GA.

GAs operate by generating new molecules in each generation through crossover and mutation of randomly-selected candidate molecules. All newly generated molecules undergo a scoring process where molecules are assigned a numerical score based on their property values. These molecules are screened, and those scoring well are passed to the next generation. Figure 2.2 summarizes the GA methodology for CPE design. Simplified Molecular Input Line Entry System (SMILES) notation was used for molecule representation in GA. Crossover and mutation operators were used for the generation of new molecules. Functional groups that were prominent in currently available CPEs were used during random mutation. Scoring and screening of the molecules was performed

using QSPR models for predicting properties such as octanol/water partition coefficient, permeation coefficient, melting point and toxicity. The non-linear QSPR models were developed using effective neural networks employing randomization of training data, random initialization of weights and random search for the best neural network for prediction of the desired property.

Molecular representation: Developing a GA for CPE molecular design requires an effective molecular representation scheme. Various methods for molecular representation are used in practice. Genetic graphs, MDL- file format, and SMILES [35, 36] are among the most popular molecular representation schemes. SMILES notation is a line/string notation that is human readable and can be transformed easily into a 2-D structure. Although the SMILES technique has a simple construction and few vocabulary rules, it encodes all the structural information found in an extended connection table. In this work, all seed molecules were converted to SMILES notation using OpenBabel software [37].

Genetic operators: GAs involve random selection of parent molecules to generate new offspring. To accomplish this, a variety of genetic operators and processes are used, as discussed below.

1. Selection: The genetic algorithm has been designed on the basis of a probabilistic operator rather than a deterministic one, as used by the other optimization techniques. This means that the molecular growth is completed in a random fashion with priority given to those molecules possessing superior characteristics, and hence, a greater probability of selection. This is achieved by using what is called “Roulette Wheel Selection” [38]. Each of the parent

molecules is scored using a fitness function and the selection pressure for that molecule is determined. As the fitness score of the molecule increases, the molecule has a higher probability to be selected as a parent. Since widely accepted chemical penetration enhancers are used in this work, the selection pressure for the seed molecules in the first generation is identical.

2. Crossover: The crossover operator creates an offspring by recombining the features of parents. Figure 2.3 shows two crossover operators: one-point crossover and two-point crossover. In one-point crossover, each of the parent molecules is cut at one location and the fragments are combined to form offspring with hybrid features. Two-point crossover involves selection of two cut points from each parent and mutual exchange of genetic information to form new molecules. Roulette wheel selection is used to choose between one-point and two-point crossovers in each generation.
3. Mutation: The mutation operator performs random changes in the parent molecule to produce a new offspring. Figure 2.4 presents an example of the various mutation operators used. The functional groups to be mutated and the number of mutations performed are selected randomly in each generation. The extent of the mutation rate determines the diversity of the offspring from the parent molecules.
4. Other operators: Other genetic operators used for molecular generation include functional group addition, functional group deletion and bond substitution. Figure 2.4 presents examples of these operators. The functional groups to be added are selected from a pool of functional groups identified as

being prominent in the currently available CPEs. Thus, the genetic material that is prominent and has potential to yield good CPEs is retained throughout all generations. The functional group to be added or deleted is chosen randomly.

Development of fitness function: Scoring and screening of the generated molecules is a key step in any CAMD technique. A GA-based search technique typically analyzes a few thousand molecules before a suitable candidate is identified. Several techniques have been developed for scoring and screening of generated molecules. One such method is the rejection of candidates that do not satisfy the target property constraints. This method is effective only when the feasible region in the search space is large. All generated molecules are given a fitness score using a fitness function that is tailored to a specific problem. The fitness score can be evaluated in two ways:

1. Assign a score to the molecule based on predicted property values.
2. Specify an acceptable range for each of the properties under consideration.

Each of these methods has advantages. By giving a score to each of the molecules through a set of property models, a minimum score for acceptance can be specified; thus, molecules are not rejected for violating one or more of the properties. For example, if four of the five properties under consideration for the molecule are within the acceptable range, then the molecule is accepted with a cutoff value of 0.8 (= 4/5) on a scale of 0 to 1. This approach ensures that genetic material is not lost completely. Alternatively, by specifying an acceptable range for each of the properties, only a few molecules that satisfy all the conditions are passed to the next generation. We believe a combination of

these two approaches, scoring of the initial generations and specifying a range at the final generations, provides an effective fitness evaluation routine.

2.2.4 Verification

Careful experimental validations for skin permeation and toxicity are conducted on the candidate CPEs that demonstrate the greatest potential. Details on the experimental validation capability of the OSU Thermodynamics Research Group are beyond the scope of this study and are given elsewhere [14, 15].

*2.*3. RESULTS AND DISCUSSION

2.3.1 QSPR models

QSPR models for properties such as skin penetration coefficient, octanol/water partition coefficient, melting point, and skin sensitization were developed to predict the physiochemical properties of the newly-generated molecules. To ensure that the QSPR models have reliable prediction capabilities, molecular databases consisting of chemicals from diverse chemical classes and spanning a wide property range were used for model development. The chemical structures used for modeling are initially optimized using the Chem3D module available in Chem3DUltra [39]. To locate the lowest energy configuration, multiple initializations were used during the structure optimization. AMPAC 6.0 software [40] was then used to further refine the 3-D geometry of the structures. The final optimized structures from AMPAC are provided as inputs to commercial QSPR software to generate over 1200 molecular descriptors. A variety of constitutional, topological, geometrical, thermodynamic, quantum-chemical and electrostatic descriptors are generated using CODESSA [41] and 154 functional group

descriptors are generated using Dragon [42]. The number of descriptors calculated for each molecule depends on the structural complexity of the molecule. Descriptors not calculated for a given molecule were set to zero in subsequent QSPR model development. The descriptor set generated from CODESSA and Dragon is orthogonalized to remove repetitive and insignificant descriptors. This reduced set still contained hundreds of descriptors.

Using non-linear algorithms to find the best set of descriptors from hundreds of descriptors requires large amounts of computational time and is often impractical. Therefore, sequential multiple regression techniques were used to obtain a reduced set of descriptors. To ensure that non-linear relationships between descriptors and properties are not ignored, non-linear transformations of the descriptors were evaluated and an expanded set of descriptors were obtained before beginning the sequential regression analysis. The descriptor set is reduced to 40 descriptors using sequential regression analysis and further reduction is accomplished using the heuristic analysis available in CODESSA. The final set of descriptors is retained for non-linear regression.

Robust ANN algorithms have been developed which are capable of:

1. Finding the optimal network architecture
2. Using cross-validation analysis to avoid over-fitting
3. Dividing the data set systematically into training, validation and testing sub-sets
4. Employing effective normalization techniques
5. Conducting multiple data randomizations and weight initializations
6. Utilizing multiple performance functions for analyzing the network

The network performance was improved by studying networks with multiple transfer functions and numbers of neurons in the hidden layers. Network architectures with one or two hidden layers has proven to be sufficient for non-linear regression and, hence, our algorithm searches for all possible one or two hidden layer architectures that satisfy a degree of freedom ratio (ratio of the number of network connections and the number of data points) lower limit of two [43]. The average of the property values predicted using three independent networks was used in order to nullify the effects of a single favorable network.

2.3.2 CPE design

A database comprised of 160 human skin permeation measurements was used to develop a skin permeation QSPR model. Our QSPR model for skin penetration coefficient is able to predict the penetration data considered within an absolute average percent deviation (%AAD) of 8.0 [44]. Similarly, 2029 octanol/water partition coefficient data [45], 970 melting point data and roughly 900 skin sensitization data [46] were used to develop the respective QSPR models. Data from local lymph node assay (LLNA) experimental procedure, guinea pig maximization test (GPMT) and Federal Institute for Health Protection of Consumers and Veterinary Medicine (BgVV) database were used to develop effective skin sensitization QSPR models. Since the experimental procedure and end-point ranking assigned to molecules by LLNA, GPMT and BgVV are different, three exclusive QSPR models were developed [46]. More details on the prediction networks used are provided in Table 2.1. Properties such as molecular weight, number of hydrogen-bond donors and number of hydrogen bond acceptors were calculated using commercially available software.

As stated earlier, 272 CPEs from the literature were used as input molecules for the first generation. Crossover and mutation operators were assigned equal probabilities of selection in the first generation and monitored in subsequent generations. Roughly 15 functional groups identified to be prominent in literature CPEs were used in functional group addition mutations. After each generation, the offspring molecules are initially monitored to remove any invalid and large molecules. SMILES structure is used to generate the 2-D structure of the offspring molecules using ChemDraw software. The 3-D structures are generated and optimized for minimum energy. Molecular descriptors are generated for property prediction using robust QSPR models already developed. For each property within the acceptable range, the score of the molecule was incremented by a value of 1. Thus a summary numerical score is assigned to each of the molecules generated. Molecules that passed all the screening tests and had a fitness score value of 8 are accepted as potential CPEs. Figure 2.5 summarizes the scoring of the molecules in each generation. The retained molecules are sorted according to their log K_p value and the top 10% of the molecules are added to the parent molecule set to be used in the next generation. Approximately 1000 molecules were generated in each generation run and this procedure was repeated for five generations. Table 2.2 summarizes the results of the five generation runs. In general, slightly less than 20% of the molecules generated were considered candidate for further evaluation as CPEs.

The molecules thus identified are further validated for skin permeation and skin irritation using carefully-designed experiments. The experimental work was done by other members of the OSU Thermodynamics Research Group and is not a part of this dissertation. However, a brief discussion of the experimental work is provided for

completeness. Molecules with a score of 8 and with high log K_p values in each generation were selected for experimental validation. In this work, insulin was the targeted drug to be delivered transdermally and, hence, the CPEs were experimentally validated for penetration enhancement of insulin. The skin permeation experimental procedures were validated by performing permeability measurements on four known CPEs using porcine skin, a Franz Cell, and High-Performance Liquid Chromatography (HPLC). The resistance factor (RF) and the insulin flux obtained using the Franz cell and the HPLC method, respectively, for the experimentally validated molecules are presented in Table 2.3. Chemical compounds with high RF and insulin flux values are considered effective in transdermal penetration. Further, the toxic effects of these enhancers were studied on (a) human foreskin fibroblasts (HFFs) with 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT)-formazan assays at two different concentrations, and (b) porcine abdominal skin using histology and haemotoxylin/eosin (H/E) staining at the end of a 24-hour exposure period. A detailed discussion of the experimental procedures is beyond the scope of this paper and is given elsewhere [14, 15].

One of the major limitations of experimental validation of the newly generated chemicals is the commercial availability of the chemicals. Often molecules with good permeation and fitness scores are not available commercially, and their experimental validation was rendered more difficult. As such, in this study, we elected to validate experimentally the molecules that are available commercially, even though they might not represent the highest fitness score in each generation. This limitation is further amplified by the fact that as the number of generations increases, the crossover and

mutations among molecules becomes extensive, which, in turn, leads to generation of greater numbers of novel molecules not available commercially.

Although a CPE may permeate through the skin, its ability to enhance the permeation of insulin through the skin depends on chemical interaction effects between the CPE and insulin. This explains why some of the virtually designed CPEs were not effective in transporting insulin through the skin. Knowledge of CPE-drug interactions in the pre-design stage is highly desired to enhance the predictive capability of our CAMD algorithms for transdermal drug delivery. From the current CPE CAMD algorithm it has been observed that chemical compounds with hydrogen bonding groups and having log K_{ow} values greater than 2.5 were effective in enhancing insulin permeation through the skin. Acids, alcohols, aldehydes, and ketones are some of the chemical classes found to be effective in enhancing insulin permeation through the skin. Incorporating this knowledge in future insulin CPE CAMD algorithms will further enhance the predictive capabilities of virtual design. Also, using the currently identified insulin CPEs as seed molecules will be effective in developing insulin specific CPEs.

The new molecules evolved in each generation are subjected to a series of steps (e.g., conversion of the SMILES structure of the molecule to 2-D structure, conversion from 2-D to 3-D structure, optimization of the 3-D structure for minimum energy using Chemoffice, re-optimization of the molecule using AMPAC, generation of descriptors, property prediction using the descriptors generated and scoring and screening based on the property values) before passing to the next generation. This process is laborious and becomes very difficult to implement as the number of generations increases. Further, the amount of human involvement required hinders the ability of the GA design program to

run multiple generations and limits the diversity in the population. Although some studies claim to have run multiple generations in their GA program, their discussion was limited concerning the optimization of the newly generated molecules. Hence, a need exists for an automation tool capable of performing multiple generations and achieving more genetic diversity with minimum manual effort.

2.4. CONCLUSIONS

Genetic algorithm-based virtual design of molecules possessing desired properties offers rapid and low-cost development opportunities. Our results indicate that integrating genetic algorithms and non-linear QSPR modeling offers a reliable CAMD algorithm for generation of potential chemical penetration enhancers (CPEs). Further, these results demonstrate the efficacy of this virtual design approach in identifying potential CPEs for transdermal drug delivery of insulin.

The lack of accurate knowledge of the drug-chemical interactions in the pre-design stage represents a limitation in the current methodology. The *a priori* knowledge of drug-chemical interactions would further improve the design ability of our newly-developed algorithms, and thus, potentially reduce the number of experimental validations, which are often expensive and laborious.

A need exists for a computational platform to orchestrate the creation of multiple generations of CPE candidates with greater genetic diversity and minimum manual intervention. Further, synthesis of the chemical compounds identified as effective CPEs would expand the list of insulin enhancers beyond the chemical structures available commercially, and this could potentially lead to identification of superior CPEs.

REFERENCES

[1] V. Venkatasubramanian, K. Chan, and J. M. Caruthers, "Computer-aided molecular design using genetic algorithms," *Computers and Chemical Engineering,* vol. 18, pp. 833–844, 1994.

[2] S. A and V. V, "Parametric sensitivity and search-space characterization studies of genetic algorithms for computer-aided polymer design," *Journal of Chemical Information and Computer Sciences,* vol. 38, pp. 1177–1191, 1998.

[3] D. James, *Genetic Algorithms in Computer-Aided Molecular Design. In Genetic Algorithms in Molecular Modeling*. New York: Academic Press, 1996.

[4] V. Venkatasubramanian, K. Chan, and J. M. Caruthers, "Evolutionary design of molecules with desired properties using the genetic algorithm," *Journal of Chemical Information and Computer Sciences,* vol. 35, pp. 188-195, 1995.

[5] P. M. Harper, R. Gani, P. Kolar, and T. Ishikawa, "Computer-aided molecular design with combined molecular modeling and group contribution," *Fluid Phase Equilibria,* vol. 158, pp. 337-347, 1999.

[6] R. B. Nachbar, "Molecular evolution: A hierarchical representation for chemical topology and its automated manipulation," *Proceedings of the Third Annual Genetic Programming Conference,* pp. 246-253, 1998.

[7] R. Leardi, "Genetic algorithms in chemometrics and chemistry: A review," *Journal of Chemometrics,* vol. 15, pp. 559-569, 2001.

[8] S. S. Godavarthy, "Design of improved solvents for extractive distillation," Ph. D. Dissertation, Oklahoma State University, Stillwater, OK, 2004.

[9] A. Sundaram and V. Venkatasubramanian, "Parametric sensitivity and search-space characterization studies of genetic algorithms for computer-aided polymer design," *Journal of Chemical Information and Computer Sciences,* vol. 38, pp. 1177-1191, 1998.

[10] J. Li, "CAMD in modern drug discovery," *Drug Discovery Today,* vol. 1, pp. 311-312, 1996.

[11] P. M. Harper, R. Gani, P. Kolar, and T. Ishikawa, "Computer-aided molecular design with combined molecular modeling and group contribution," *Fluid Phase Equilibria,* vol. 158, pp. 337-347, 1999.

[12] J. Greene, S. Kahn, H. Savoj, P. Sprague, and S. Teig, "Chemical function queries for 3D database search," *Journal of Chemical Information and Computer Sciences,* vol. 34, pp. 1297-1308, 1994.

[13] R. Zhao, H. Cabezas, and S. R. Nishtala, "Computer-aided solvent design for the environment," in *Process Design Tools for the Environment*: Taylor & Francis, 2001.

[14] V. K. Rachakonda, K. M. Yerramsetty, S. V. Madihally, R. L. Robinson, and K. A. M. Gasem, "Screening of chemical penetration enhancers for transdermal drug delivery using electrical resistance of skin," *Pharmaceutical Research, Under review*.

[15] K. M. Yerramsetty, V. K. Rachakonda, S. V. Madihally, R. L. Robinson, and K. A. M. Gasem, "Influence of enhancer characteristics and drug hydrophobicity on permeation enhancement," *Journal of Controlled Release, Under review*.

[16] P. M. Harper and R. Gani, "A multi-step and multi-level approach for computer aided molecular design," *Computers and Chemical Engineering,* vol. 24, pp. 677-683, 2000.

[17] H. Cabezas, "Designing green solvents," *Chemical Engineering,* vol. 107, pp. 107-109, 2000.

[18] E. J. Pretel, P. A. Lopez, S. B. Bottini, and E. A. Brignole, "Computer-aided molecular design of solvents for separation processes," *AIChE Journal,* vol. 40, pp. 1349-1360, 1994.

[19] L. Constantinou, K. Bagherpour, R. Gani, J. A. Klein, and D. T. Wu, "Computer aided product design: Problem formulations, methodology and applications," *Computers and Chemical Engineering,* vol. 20, pp. 685-702, 1996.

[20] C. J. Schult, "Design of solvents for extractive distillation," Ph. D. Thesis, Oklahoma State University, Stillwater, OK, 2000.

[21] K. A. M. Gasem, R. L. Robinson Jr., C. J. Schult, and B. A. Todd, "Separation of hydrocarbons by extractive distillation," US Patent 6,392,115, 2002.

[22] L. E. K. Achenie, R. Gani, and V. Venkatasubramanian, *Computer Aided Molecular Design: Theory and Practice*, Elsevier Science, 2003.

[23] M. Hostrup, P. M. Harper, and R. Gani, "Design of environmentally benign processes: integration of solvent design and separation process synthesis," *Computers and Chemical Engineering,* vol. 23, pp. 1395-1414, 1999.

[24] C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney, "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings," *Advanced Drug Delivery Reviews,* vol. 46, pp. 3-26, 2001.

[25] B. C. Finnin and T. M. Morgan, "Transdermal penetration enhancers: Applications, limitations, and potential," *Journal of Pharmaceutical Sciences,* vol. 88, pp. 955-958, 1999.

[26] R. Kumar and A. Philip, "Modified Transdermal Technologies: Breaking the Barriers of Drug Permeation via the Skin," *Tropical Journal of Pharmaceutical Research,* vol. 6, pp. 633-644, 2007.

[27] L. Brown and R. Langer, "Transdermal Delivery of Drugs," *Annual Review of Medicine,* vol. 39, pp. 221-229, 1988.

[28] D. W. Osborne and J. J. Henke, "Skin penetration enhancers cited in the technical literature," *Pharmaceutical Technology,* vol. 21, pp. 58-66, 1997.

[29] H. Y. Thong, H. Zhai, and H. I. Maibach, "Percutaneous penetration enhancers: An overview," *Logo,* vol. 20, 2007.

[30] K. C. V Venkatasubramanian, and J. M. Caruthers, "Evolutionary design of molecules with desired properties using the genetic algorithm," *Journal of Chemical Information and Computer Sciences,* vol. 35, pp. 188-195, 1995.

[31] L. Weber, "Evolutionary combinatorial chemistry: application of genetic algorithms," *Drug Discovery Today,* vol. 3, pp. 379-385, 1998.

[32] J. H. Holland, *Adaptation in natural and artificial systems*: University of Michigan Press, Ann Arbor, MI, 1975.

[33] S. Kamphausen, N. Höltge, F. Wirsching, C. Morys-Wortmann, D. Riester, R. Goetz, M. Thürk, and A. Schwienhorst, "Genetic algorithm for the design of molecules with desired properties," *Journal of Computer-Aided Molecular Design,* vol. 16, pp. 551-567, 2002.

[34] E. W. Lameijer, J. N. Kok, T. Bäck, and A. P. Ijzerman, "The molecule Evoluator. An interactive evolutionary algorithm for the design of drug-like molecules," *Journal of Chemical Information and Modeling,* vol. 46, pp. 545-552, 2006.

[35] E. T. D. Douguet, G. Grassy, "A genetic algorithm for the automated generation of small organic molecules: Drug design using an evolutionary algorithm," *Journal of Computer Aided Molecular Design,* vol. 14, pp. 449–466, 2000.

[36] R. G. Bone, "SMILES extensions for pattern matching and molecular transformations: Applications in chemoinformatics," *Journal of Chemical Information and Computer Sciences,* vol. 39, pp. 846-860, 1999.

[37] C. Morley, "OpenBabelGUI," 2.0.2 ed, 2006.

[38] G. Jones, "Genetic and evolutionary algorithms," *Encyclopedia of Computational Chemistry. John Wiley & Sons, Ltd., September,* 1998.

[39] ChemBioOffice 11.0, Cambridge Software, 2008.

[40] AMPAC 6.0., Semichem Inc., 1997.

[41] CODESSA 2.7.8., Semichem Inc., 2007.

[42] Talete srl, DRAGON 5.4 for Windows (Software for molecular descriptor calculations), 2006.

[43] M. Hagan, Personal communication, School of Electrical and Computer Engineering, Oklahoma State University, 2007.

[44] B. J. Neely, S. Golla, E. Whitebay, S. Madihally, R. L. Robinson, and K. A. M. Gasem, "Quantitative structure-property modeling of skin permeation coefficient," *In preparation*.

[45] E. Whitebay, B. J. Neely, S. Golla, S. Madihally, R. L. Robinson, and K. A. M. Gasem, "QSPR modeling of octanol/water partition coefficient for a generalized structure set," *In preparation*.

[46] S. Golla, B. J. Neely, E. Whitebay, S. Madihally, R. L. Robinson, and K. A. M. Gasem, "Quantitative structure-property relationship modeling of skin sensitization: A quantitative prediction," *In preparation*.

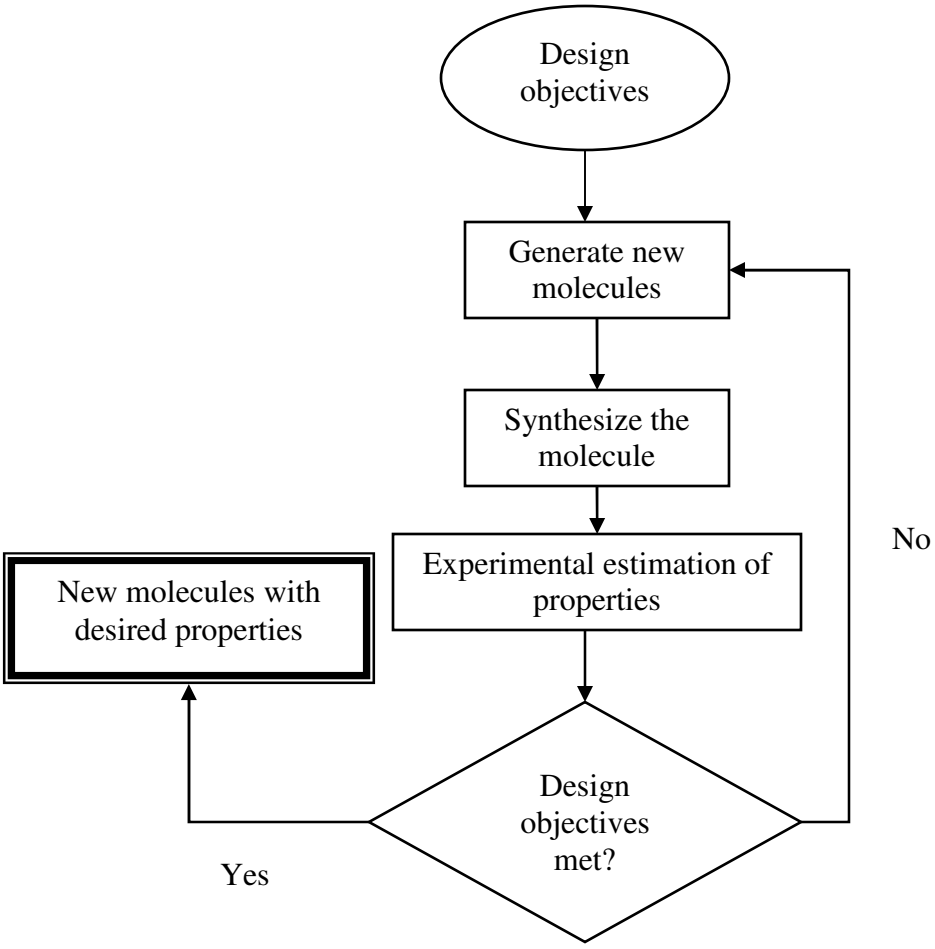


Figure 2.1. Basic steps of CAMD

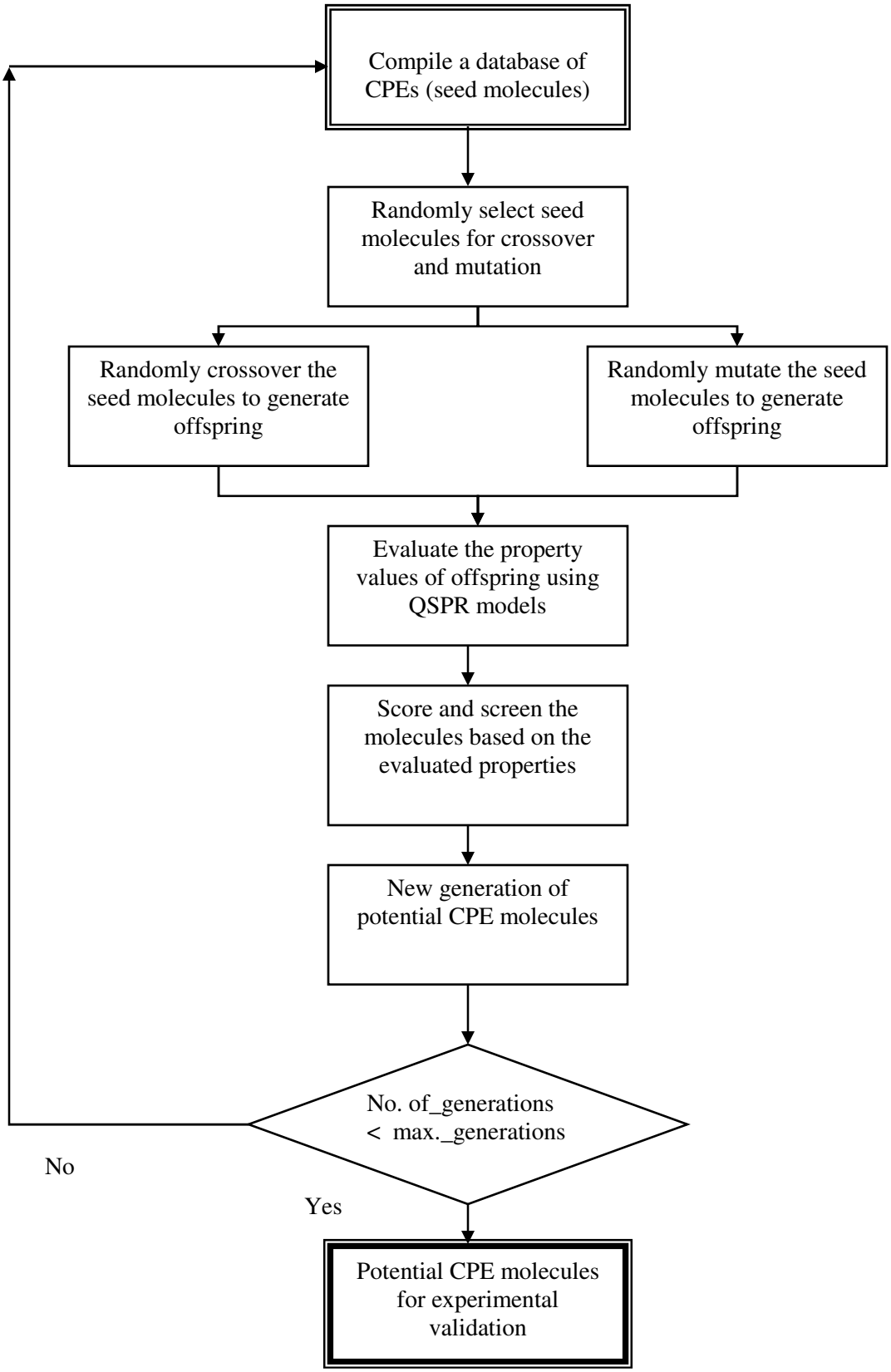
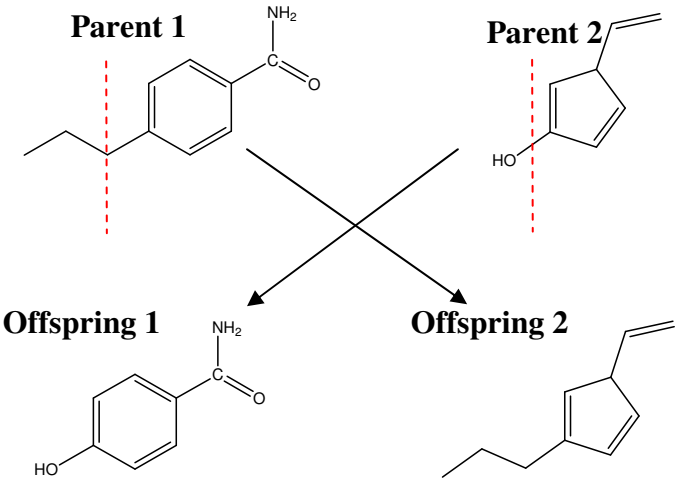
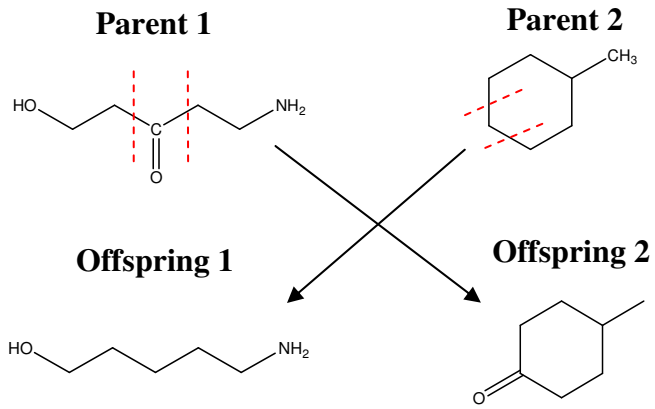


Figure 2.2. Virtual design of CPEs: Flow diagram

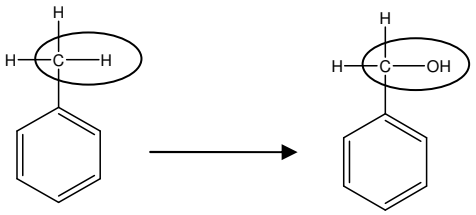


One point crossover operation

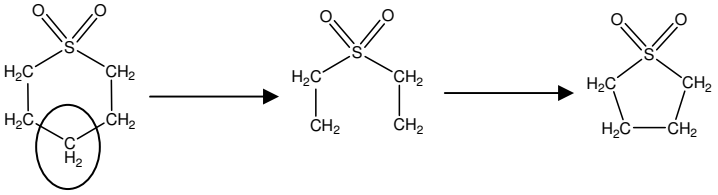


Two point crossover operation

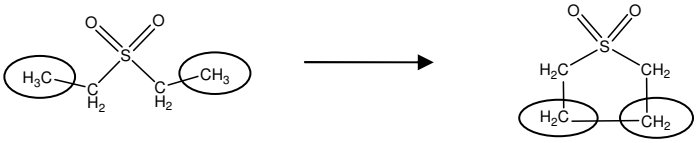
Figure 2.3. Crossover operators



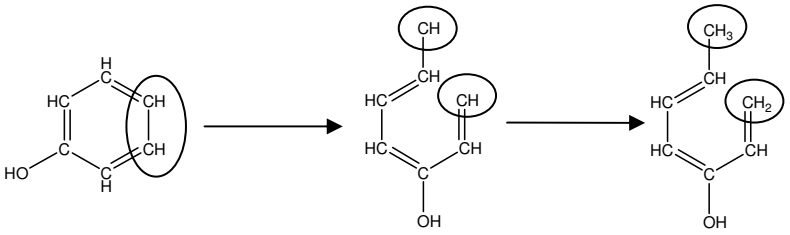
Functional group addition



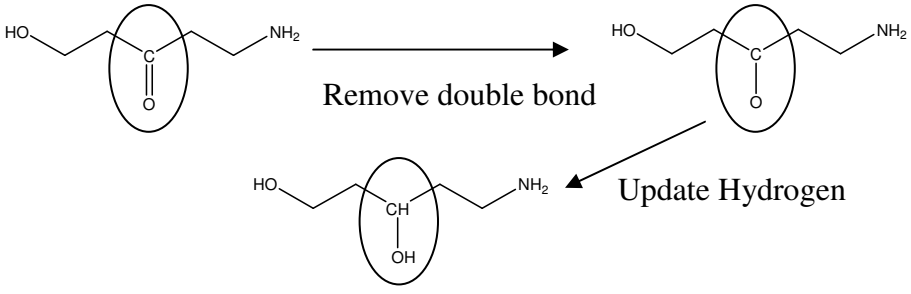
Functional group deletion



Ring addition



Ring deletion



Double bond removal

Figure 2.4. Mutation and other operators

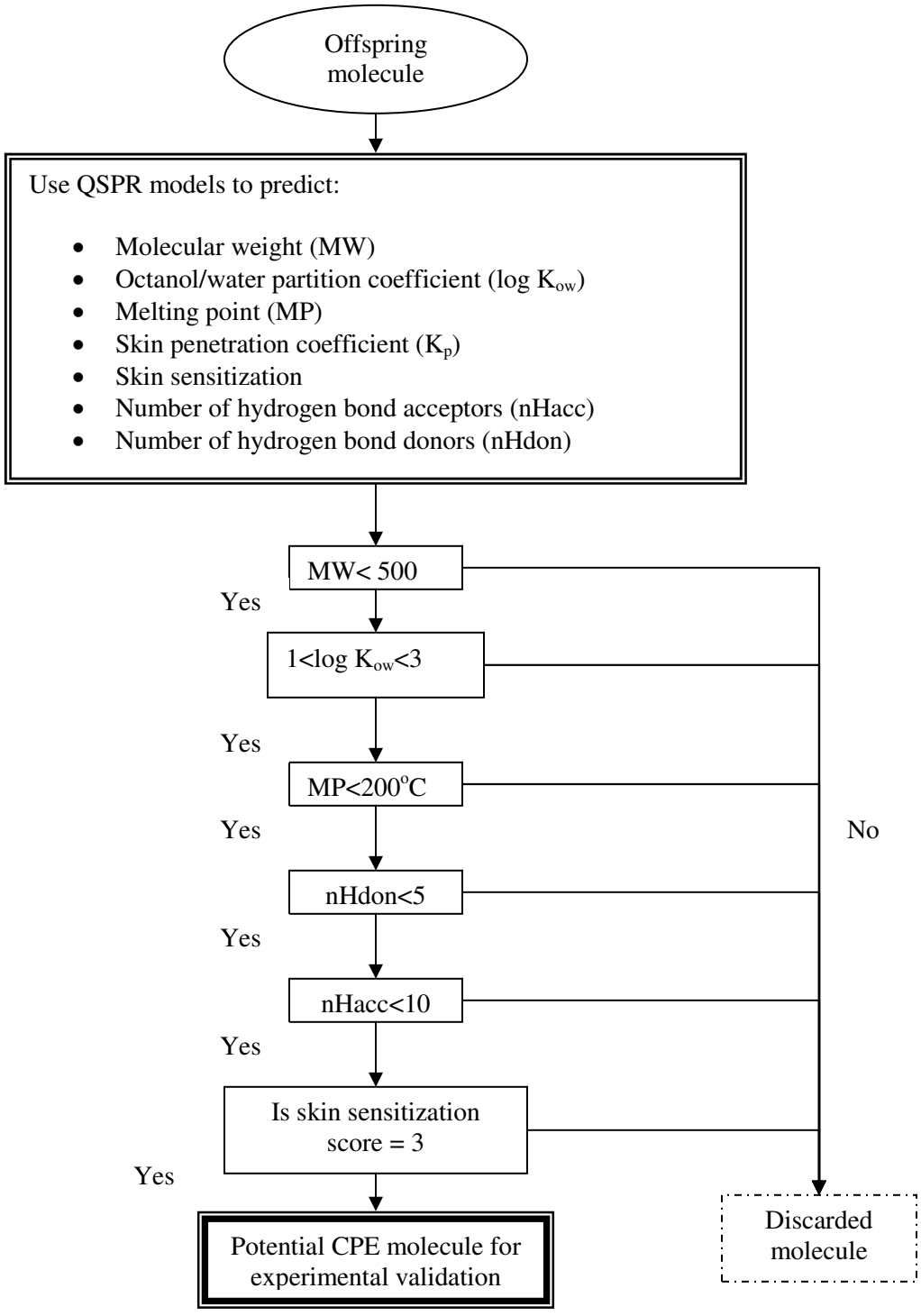


Figure 2.5: Scoring and screening of an offspring molecule

Table 2.1. Summary of QSPR models used for property prediction

Property	No. of datapoints used for modeling	Range of property values	No. of descriptors	Neural network architecture	R^2	RMSE	% Accuracy
Melting point	965	14 – 586 K	20	20-14-7-1	0.9	25 K	-
Octanol/water partition coefficient (log K_{ow})	2029	(-12) - 9.4	9	9-30-6-1	0.91	0.7	-
Skin penetration coefficient (log K_p)	160	(-5.6) - (-1.0)	10	10-5-3-1	0.9	0.36	-
Skin sensitization							
LLNA	358	0 - 1	25	25-4-11-1		-	90
GPMT	307	0 - 1	25	25-3-6-1		-	95
BgVV	251	0 - 1	24	24-4-1		-	93

* RMSE = root-mean-squared error in property predictions
% Accuracy = percentage of correct classisfications

Table 2.2. Genetic algorithm results from each generation

Generation number	Number of seed molecules	Total number of new molecules generated	Accepted molecules: score of 8	% of accepted molecules
1	249	943	120	12.7
2	269	978	155	15.9
3	290	995	269	27.0
4	311	1009	193	19.1
5	331	909	156	17.2
All	1450	4834	893	18.5

Table 2.3. Experimental results and predicted property values of CPEs

Generation	Chemical	log K_{ow}	MW	nHacc	nHdon	MP	log K_p	BgVV score	GPMT score	LLNA score	RF	INSULIN FLUX (10^{-4}*IU/m^2)
1	OSU1	2.1	120.0	1.0	1.0	291.1	-1.5	0.2	0.1	0.2	1.4 ± 0.1	1.8
	OSU2	2.8	128.0	1.0	1.0	242.9	-1.7	0.1	0.3	0.3	28 ± 7	8.1
	OSU3	2.8	128.0	1.0	1.0	264.2	-1.7	0.2	0.1	0.1	17 ± 2	5.5
	OSU4	2.7	72.1	0.0	0.0	124.2	-1.8	0.5	0.2	0.0	2.1 ± 0.1	1.9
	OSU5	1.7	100.0	1.0	1.0	200.4	-1.9	0.1	0.3	0.3	2.1 ± 0.3	4.5
	OSU6	2.0	130.0	0.0	0.0	199.0	-2.2	0.1	0.0	0.2	7 ± 1	6.8
	OSU7	2.4	70.1	0.0	0.0	121.8	-2.2	0.1	0.1	0.1	7 ± 3	-
	OSU8	1.0	86.1	1.0	1.0	193.4	-2.2	0.1	0.3	0.3	-	2.6
	OSU9	2.2	144.0	0.0	0.0	208.4	-2.5	0.1	0.1	0.2	3 ± 1	4.5
2	OSU10	2.9	168.3	1.0	1.0	233.1	-1.6	0.2	0.0	0.3	53 ± 6	10.5
	OSU11	1.3	116.2	1.0	1.0	264.9	-2.3	0.1	0.1	0.0	5 ± 2	2.0
	OSU12	1.1	87.2	1.0	1.0	131.5	-2.6	0.5	0.0	0.1	60 ± 8	6.6
3	OSU13	1.6	116.2	0.0	0.0	175.8	-2.0	0.2	0.1	0.0	2 ± 1	3.0
	OSU14	1.8	100.2	1.0	1.0	187.7	-2.1	0.1	0.1	0.3	14 ± 2	3.8
	OSU15	2.5	129.2	1.0	1.0	262.2	-2.2	0.2	0.2	0.2	76 ± 8	10.3
4	OSU16	2.1	114.2	1.0	1.0	221.8	-1.5	0.2	0.1	0.1	4 ± 2	3.1
	OSU17	1.5	112.2	1.0	1.0	203.4	-1.9	0.2	0.4	0.2	4 ± 2	3.4
5	OSU18	2.3	114.2	1.0	1.0	254.1	-1.4	0.1	0.1	0.1	5.0 ± 0.6	-

CHAPTER 3

QUANTITATIVE STRUCTURE-PROPERTY RELATIONSHIPS MODELING OF SKIN SENSITIZATION

3.1. INTRODUCTION

The discovery of numerous new chemicals for various scientific applications involving humans creates the need for reliable assessment of their toxic effects. Extensive efforts have been made to identify test methods for evaluating toxicity of various chemical compounds [1, 2]. Among the toxicity-related issues, skin sensitization due to exposure to toxic chemical compounds has been a major work-related problem. In fact, it comprises up to 95% of the occupational contact dermatitis cases [3]. Skin sensitization is considered a human health risk that can be caused by skin contact with a wide range of chemicals, including those employed in cosmetics. Skin sensitization occurs when a foreign, low-molecular weight substance that acts as an allergen penetrates the skin and combines with skin proteins to produce an immune response. The initial exposure is called the sensitization phase and has no clinical symptoms. The delayed skin response from a later exposure to the allergen is called the elicitation phase [4]. Clinical symptoms include erythema (redness), vesicles, papules, scaling, and pruritus (itching). Common chemicals that cause of this type of allergy include metals (notably nickel), epoxy and acrylate chemicals, fragrances, preservatives, and many other natural chemicals [5].

Studies that focused on understanding sensitization mechanisms suggest a number of requirements for a chemical to cause skin sensitization. The chemical (a) should be able to penetrate into the epidermis across the stratum corneum, (b) has to react with the protein, (c) has to induce local trauma, and (d) must be recognized by the immune system [4]. In practice, however, the process may be more complex when metabolic transformations of the chemical are involved. Often, the reaction between the chemical and protein is believed to be covalent in nature. Therefore, skin sensitization is underpinned by mechanisms based on chemical reactivity, where the chemical behaves as an electrophile and the protein behaves as a nucleophile [6]. There are various types of electrophile-nucleophile reactions encountered in skin sensitization, including Michael-type reactions; SN2 reactions; SNAr reactions; acylation reactions and Schiff-base formation. A detailed description of each of these reactions is given elsewhere [7].

For many years, the guinea pig maximization test (GPMT) and the Buehler test, as described in the Organization for Economic Co-operation and Development (OECD) Guideline 406, were used as effective methods for assessing skin sensitization [8]. In GPMT, preliminary tests are performed to identify the maximum non-irritating dose. The chemical is then tested for skin sensitization using 15 to 20 animals for about five weeks; the tests including an induction phase, a rest period and a 24-hour topical challenge. The chemicals tested are grouped into categories based on the extent of positive response: strong sensitizers (70-100% positive), moderate sensitizers (30-70% positive), weak sensitizers (0-30% positive), and non-sensitizers (negative). Molecules that belong to the strong and moderate classes are classified as skin sensitizers [9]. Although the presence of readily available GPMT data for many available chemicals facilitates comparative

interpretation, the GPMT method is often not recommended because it uses one concentration and it does not provide any information about the thresholds. Open Epicutaneous Test (OET) is another potency investigating procedure that uses guinea pigs. In addition to determining whether a chemical causes skin sensitization, this procedure gives information about the dose responses in the induction and challenge phases. This procedure is found to be superior to the GPMT and Buehler test because it uses a multiple dose regime to determine the dose responses and thresholds. However, guinea pig tests are expensive and often require a large number of animals to be killed [5]. About $6000 to $7000 and 24 to 32 guinea pigs are required to test a chemical using GPMT procedures [10]. Therefore, a strong incentive exists to find alternate methods that are more cost effective and require fewer animals.

Recently, local lymph node assay (LLNA) described in the OECD Guideline 429 has been accepted as a valid test method for assessing skin sensitization. This test provides both qualitative and quantitative measures of sensitization potency [11]. The costs associated with LLNA have been estimated to be $6000 along with the use of 16 to 30 mice per chemical. In comparison, the amount of chemical required was found to be much less than that used in the GPMT test, resulting in an additional cost reduction [10]. In LLNA, the classification is based on the chemical concentration necessary to induce a threefold or greater increase in lymph node cell proliferation activity in treated groups as compared to the control [6]. This concentration, known as the EC3 value, is estimated by linear interpolation of skin sensitization factors above and below the value of three on the LLNA dose response plot. A close association between the EC3 values and the relative skin sensitizing potential of chemicals among humans has been observed [5]. Thus, based

on the EC3 results obtained, a chemical is classified as being extreme, strong, moderate, weak and non-sensitizing [12].

Although reliable test procedures for skin sensitization exist, their application is limited either by the time consumed or cost involved. Hence, computational techniques that reduce the effort and cost and ensure animal welfare are needed. Non-testing procedures such as quantitative structure-property relationships (QSPR) represent effective methods for *in vitro* prediction of physical properties of chemical compounds. QSPR models offer an attractive alternative since they have the potential to provide reliable property estimates based solely on chemical structure information. If structural information can be successfully decoded, the properties can be determined directly from the known chemical structure. QSPR studies have received a significant boost with the advent of high-speed computers. This has not only led to the development of new and more complex molecular descriptors and also has facilitated the application of QSPR models to properties that were previously infeasible due to computational intensity. QSPR models are now well established and are used to correlate varied, and often complex, physiochemical properties of molecules. The QSPR approach has been applied in different areas, and a detailed review of their applications can be found in one of our earlier publications [13].

Numerous QSPR models in the literature predict skin sensitization of chemical compounds with reasonable accuracy; however, most of these models are developed for a specific class of compounds. Hence, their general applicability is limited. A detail review of available literature models is provided in the next section.
In the present work, an effort has been made to develop a QSPR skin sensitization model

with wider applicability. An extensive database comprised of test results from three test procedures were used for QSPR model development. This work focuses on the following objectives:

(a) Developing QSPR models to predict skin sensitization, utilizing experimental procedures and the end-point rankings from the LLNA, GPMT and Federal Institute for Health Protection of Consumers and Veterinary Medicine (BgVV) databases. (To do so, three separate QSPR models were developed.)

(b) Improving the predictive capability of the QSPR models using a combination of literature, functional group and structural descriptors.

3.2. SKIN SENSITIZATION MODELS

Estimation of skin sensitization for chemical compounds requires reliable generalized predictive models. Skin sensitization of a chemical compound is often found to be dependent on three factors: the chemical reactivity, molecular size, and skin permeability [14]. The mechanistic pathway to cause skin sensitization is often class-specific, multiple class-specific QSPR models are believed to be required for the separate classes. Hence, earlier QSPR models for skin sensitization were targeted for specific classes of chemicals. However, these class-specific models had limited predictive ability and were not universally applicable. QSPR models capable of predicting toxicity for diverse datasets were developed to overcome this limitation. Comprehensive reviews of QSPR skin sensitization models are provided by Rodford and coworkers [15], Pease and coworkers [16], and Patlewicz and coworkers [14, 15].

The models for prediction of skin sensitization can be classified into class/mechanism based QSPR models, statistical QSPR models and expert systems that are either statistically or knowledge based or both. The mechanistic models group chemicals according to their protein-chemical (nucleophile-electrophile) reaction mechanism to develop robust models to predict skin sensitization potential and potency. Some of the reaction mechanisms considered in such models include Michael-type reactions, SN2 reactions, SNAr reactions, acylation reactions, and Schiff-base formation. A detailed description of all the available mechanistic models has been provided by Patlewicz and coworkers [17]. However, the robustness of these models relies on the availability of a comprehensive reactivity database. Also, the mechanistic domain studies fail when a chemical compound cannot be classified accurately. Therefore studies that encompass a range of QSPR computational strategies involving physiochemical descriptors, structural alerts and statistically determined descriptors taken from a large pool of structural descriptors have been developed [18].

Experimental data from LLNA, QSPR or BgVV have been used extensively, either exclusively or in combination, to develop a number of QSPR models utilizing structural descriptors. Although mechanistic interpretation of all the descriptors is not possible, a few studies have tried to address size and reactivity descriptors. Some of the models used a combination of descriptors and structural alerts to model skin sensitization.

Cronin and Basketter [19], using a database of organic compounds from the guinea pig maximization test, developed a QSPR model that predicted 82.6% of the chemical compounds correctly. They used a combination of molecular descriptors and 27 structural alerts to develop a linear regression model. Some of the significant parameters

found in the study include the HOMO-LUMO energy gap which accounts for the stability and reactivity of the molecule, Shannon index which accounts for the molecular size, and 12 structural alerts that relate to the sites of chemical reactivity causing skin sensitization. A nonlinear analysis using artificial neural networks was later developed by Devillers [20] using the same parameters. Accuracies of 90% and 87% were observed for the training set and validation set, respectively.

After LLNA was accepted as a valid test method for assessing skin sensitization, various QSPR models were published based on LLNA data or a combination of LLNA and GPMT data. In 1995, Ashby and coworkers [21] studied the structural activity relationship (SAR) of 106 LLNA tested chemical compounds and categorized sensitizing and non-sensitizing agents. Fedorowicz and co-workers [3] developed a QSPR model using LLNA data for 54 chemical compounds. An accuracy of 83% and 79% was observed for the training and validation sets using four molecular descriptors. Patlewicz and co-workers [22] evaluated two groups of fragrance chemicals – saturated aldehydes and unsaturated aldehydes using QSPR models relating the EC3 values derived from LLNA to physiochemical properties (reactivity and lipophilicity). In 2005, Dimitrov and co-workers [23] developed a SAR/QSPR model by integrating data from LLNA, GPMT and BgVV. Although the end-point ranking for each of these methods is different, data were reassessed using a unified three-category scale. Thus, a larger dataset containing 634 chemicals with wide chemical diversity was obtained. The SAR/QSAR model classified correctly about 80% of the chemicals with significant sensitizing effect and 72% of non-sensitizing chemicals. External validation was done using a set of 96 chemicals, and 87% correct predictions were obtained. Recently, statistical QSPR models

utilizing 4-D finger print descriptors have been found to be effective for predicting skin sensitization. Logistic regression techniques were used to obtain QSPR models with accuracies close to 70% using LLNA data of 218 compounds [24].

A number of expert systems that combine the structure-toxicity relationships, knowledge of mechanisms and skin metabolism, or statistical analysis have been developed. Expert systems such as Derek for Windows (DfW), TOxicity Prediction Komputer-Assisted Technology (TOPKAT), Multi-CASE, Hazard Expert and TIssue MEtabolism Simulator for Skin Sensitization (TIMES-SS) have been found to be effective for predicting skin sensitization. However, their application is limited because the algorithms used to develop the expert systems are not fully described [14]. Local QSAR models using Relative Alkylation Index (RAI) approach relating the skin sensitization potency of a chemical compound to the dosage amounts, rate nucleophile-electrophile rate constants and the octanol/water partition coefficients were developed for small sets of compounds. These models are effective in predicting skin sensitization for the targeted chemicals only [17].

3.3. QSPR METHODOLOGY

The development of a QSPR model for skin sensitization involves several distinct steps: (a) compilation of a data set, (b) generation and optimization of 3-D molecular structures, (c) calculation of molecular descriptors, (d) reduction of the number of descriptors, and (e) development of a regression model. To begin, reliable data for a variety of molecules must be assembled. Once this data set is characterized, the 3-D structures of the molecule are generated using commercial molecular visualization

software. Structure optimization and descriptor generation are performed using commercial QSPR software. The model development phase usually begins with linear analysis, where algorithms such as multiple-linear regression analysis, principal component analysis, heuristic analysis or partial least squares may be used. The best set of descriptors is then used to develop the non-linear QSPR model using artificial neural network algorithms.

3.3.1 Dataset Compilation

High quality data generated through good laboratory practices and complying with the OECD test guidelines have been used for modeling. The majority of the LLNA data (211 molecules) used originated from a study by Gerebrick and co-workers [25]. Other sources of LLNA data include publications by Langton and co-workers [26], Patlewicz and co-workers [14], and NIH publication no. 99-4494 [10]. The GPMT data were compiled from Cronin and co-workers [19] and Devillers and co-workers [20]. An expert group from Federal Institute for Health Protection of Consumers and Veterinary Medicine (BgVV) collected and evaluated data from the literature on substances with documented contact allergenic properties in human and animal experiments. The BgVV data of 244 chemicals with contact allergenic properties published by Schlede and co-workers [27] were used in this study. In-house data of Unilever shared by Patlewicz were also added to the LLNA, GPMT and BgVV databases [28]. Potency categories described in the above mentioned sources were used in the present modeling effort.

Oklahoma State University (OSU) LLNA database: A database containing experimental LLNA data for 392 molecules was compiled from the above-mentioned sources. In terms of chemical diversity, the database contains experimental data pertaining to a number of

chemical classes, including aldehydes, ketones, aromatic amines, quinones and acrylates. It also includes compounds exhibiting various reaction mechanisms. The molecules are classified as non-sensitizers, weak, moderate, strong and extreme sensitizers, based on experimental EC3 values. The molecules were scored for their potency on a scale of 0-1.

Since the LLNA data available in the literature are limited to a few hundred compounds, efforts were made to utilize all reliable experimental data for modeling. Of the 392 molecules, 96 are classified as either sensitizing or non-sensitizing (positive or negative). For these molecules, a score of 0 is assigned to non-sensitizers and a value of 0.625 (average of weak potency and extreme potency) is assigned to sensitizers. Table 3.1 summarizes the scoring of the LLNA data used.

Of the 392 molecules, only 358 structures were considered for modeling. A total of 34 molecules were rejected for one the following reasons: (a) structure could not be found, (b) structure could not be optimized using ChemDraw 3D [29], or (c) molecular descriptors could not be generated for the structure. The molecular weight distribution of the 358 chemical compounds used for modeling is provided in Figure 3.1. All these molecules have molecular weights (<500), to be consistent with the chemicals known to be good skin permeants.

In addition to chemical diversity, for the model to be able to predict skin sensitization effectively, the range of potencies used to train the model should be adequate. As indicated by the EC3 values, the chemicals compiled in the database display a wide range of potencies. Specifically, the LLNA EC3 values show a range of potencies from weak allergens to extreme allergens. Accordingly, the chemicals are classified as weak, moderate, strong and extreme sensitizers. Figure 3.2 shows the

potency category distribution of the molecules considered. Non-sensitizers and weak sensitizers are classified as non-toxic and moderate, strong and extreme sensitizers are classified to be toxic. Hence, based on the scores assigned to each of the potency categories, all molecules that are in the range 0-0.375 are non-toxic, and the ones in the range of 0.375-1 are toxic.

OSU GPMT database: A database containing experimental GPMT data for 334 chemical compounds was used for modeling skin sensitization. The molecules were categorized as being non-sensitizers, weak, moderate and strong sensitizers, and toxicity scores of 0, 0.33, 0.66 and 1 were assigned, respectively. Molecules obtained from Unilever that are classified as being strong were given a score of 0.83 (average of moderate and strong) because the database classifies both moderate and strong sensitizers as being strong. Table 3.1 summarizes the scoring of GPMT molecules.

Of the 334 molecules compiled, only 307 of them were used for modeling. The other 27 molecules were rejected for reasons discussed above. To develop a model applicable to a wide range of chemical compounds, we used training data with wide distributions of molecular weights and potency categories. The molecular weight distribution and the potency distribution of the final set of compounds are shown in Figures 3.3 and 3.4, respectively. Adequate representations of chemical compounds from different potency categories were used. Non-sensitizers and weak sensitizers were classified as being non-toxic, and moderate and strong sensitizers were classified as being toxic. On a scale of 0-1, molecules with toxicity scores in the range of 0-0.5 were classified as being non-toxic, and the ones in the range of 0.5-1 were classified as being toxic.

OSU BgVV database: The BgVV database containing allergic potency data for 272 molecules was used for modeling. All the molecules are classified into three categories: A, B, and C, where A refers to molecules that are strong allergens, B refers to weak allergens, and C refers to molecules with very low or no allergic properties. Toxicity scores of 1, 0.5, and 0 were assigned to A, B, and C categories, respectively, as shown in Table 3.1. Of the 272 molecules, 21 molecules are not included in the QSPR model for reasons discussed earlier. The final dataset included 251 molecules. The molecular weight distribution and potency category distributions are shown in the Figures 3.5 and 3.6, respectively. Molecules belonging to Category A are classified as being toxic, and molecules belonging to Categories B and C are classified as being non-toxic. Hence, on a scale of 0-1, molecules with toxicity score 0-0.75 are non-toxic and 0.75-1 are classified as toxic.

Data validation studies: Validation of the experimental procedures outlined above as test methods for assessing skin sensitization is critical before adopting such data for model development. A committee formed in 1999 compared LLNA data for 209 chemicals to the available GPMT and human skin sensitization data. The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) with support from the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) sponsored this independent scientific peer review to validate the LLNA test [10]. The statistics obtained from the study are shown in Tables 3.2 and 3.3.

Tests comparing LLNA and GPMT show an accuracy of 89% for the available 97 chemical compounds with both GPMT and LLNA data. Comparing LLNA data to human

test data showed an accuracy of 72% for 74 molecules. While comparing the GPMT to human tests yielded an accuracy of 72% for 57 molecules. In terms of accuracy, sensitivity, specificity, and positive and negative predictability, the performance of the LLNA was similar to that of the GPMT. Equally important, the performance of the LLNA and the GPMT was similar when each was compared to human test data.

An assessment of reliability was conducted using data for 2,4-dinitrochlorobenzene (DNCB) and hexylcinnamic aldehyde (HCA). The two data sets consisted of EC3 values for DNCB tested twice in each of five laboratories, and HCA tested six times in each of two laboratories. This reliability analysis calculated the within-laboratory consistency statistics and the between-laboratory consistency statistics. The results indicated agreement within 95% confidence limits [30].

3.3.2 Structure Generation and Optimization

The first step in any QSPR modeling effort is the generation of chemical structures. Various chemical representations have been proposed in the literature. For example, OpenBabel software [31] has around 80 different representations for a given molecular structure. The most common way of representing a chemical is a two-dimensional (2-D) sketch. However, using a 2-D representation does not provide a complete description of the molecule and lacks both shape and surface distribution information of the molecule. To have an effective QSPR model, the representation of a molecule should provide all the necessary structural information. This requires information about the atoms present, along with three dimensional (3-D) coordinates that provide a full spatial depiction of the molecule. A commercial package, ChemDraw [29] was used to generate the 2-D structures of the molecules. The 3-D structures were

generated for these molecules using Chem3DUltra [29]. Since more than one set of 3-D coordinates can satisfy the structural constraints (bond length and bond angle) for any given molecule, the conformation with the lowest energy must be located. The structures were initially optimized using the Chem3D module available in Chem3DUltra. To locate the lowest energy configuration multiple initializations were used during the structure optimization. AMPAC [32] was then used to further refine the 3-D geometry of the structures. Output files from AMPAC were used to calculate various descriptors.

3.3.3 Descriptor Generation

The final optimized structures from AMPAC were provided as inputs to commercial QSPR software to generate over 1200 molecular descriptors. A variety of constitutional, topological, geometrical, thermodynamic, quantum-chemical and electrostatic descriptors were generated using CODESSA [33], and 154 functional group descriptors were generated using Dragon [34]. The number of descriptors calculated for each molecule depends on the structural complexity of the molecule. Descriptors that were not calculated for a given molecule were set to zero in the subsequent QSPR model development.

3.3.4 Descriptor Reduction

All the descriptors generated for a specific molecule are not significant in modeling. The use of all available descriptors in the model development causes dimensionality problems and diminishes the performance of a QSPR model, especially when non-linear algorithms are used in model development [35]. Descriptor reduction is the process of automating the identification of the most relevant set of descriptors for model development and is among the critical steps in QSPR modeling efforts. Different

methods for reduction are available in the literature [36]. The most widely used techniques are the principal-component analysis (PCA), partial least-squares (PLS), genetic algorithms (GA), and neural networks (NNs) [35]. Most QSPR models have been based on multiple linear regression correlations, which requires *a priori* assumption of the form of the mathematical correlation model. However, linear regression analysis ignores the possibility of non-linear relationships between the descriptors and the property being modeled. The use of such linear approaches often leads to loss of critical information and results in models with poor predictive abilities [37].

To ensure that the non-linear relationships are accounted for in the QSPR models, non-linear transformations of all the descriptors were calculated, and an expanded set of descriptors was generated. The expanded set of descriptors was used to find the best set of descriptors through sequential selection. The best descriptor from the expanded set is selected and combined sequentially with the remaining descriptors. Then the best two-descriptor combination is retained and combined with each of the remaining descriptors. The sequence continues until a set of 40 descriptors has been identified. Heuristic regression available in CODESSA was used to further reduce the number of descriptors. The optimal set of descriptors was retained for artificial neural network (ANN) analysis.

3.3.5 Literature Descriptors

For a valid and reliable predictive model, some mechanistic interpretation of the relationship between the property and descriptors would appear possible [38]. However, such interpretation is often not possible for all the descriptors due to the complexities of the property-descriptor relation. A detailed review of the literature indicates that chemical reactivity, molecular size, and skin permeability are important determinants of skin

sensitization. The chemical reactivity of a molecule can be related to the HOMO and LUMO energies. Molecules with low HOMO-LUMO energy-gaps have low kinetic stability and are favorable to addition or removal of electrons, thus forming activated complexes [39]. The frontier orbital approximations state that a molecule preferably reacts with the molecule whose frontier orbitals are closest in energy. Thus, a nucleophile (protein) tends to react with a electrophile (molecule) having the lowest LUMO energy [40]. This approximation was validated by analyzing the LUMO energies of the chemical compounds for three data sets used. The LUMO energies were observed to be relatively lower for compounds that were tested as positive for skin sensitization in the animal tests.

Skin penetration ability is an important factor for the risk assessment of chemicals. The chemical compound should be able to permeate into the skin to cause skin sensitization. Therefore, the permeability of the chemical in the skin becomes an important factor to model skin sensitization. However, due to the limited amount of experimental data available on skin permeability, robust QSPR models for its prediction do not exist. Nevertheless, Barratt and coworkers [9] did find that the molecular volume and the octanol/water partition coefficient are important determinants of skin permeability.

3.4. QSPR MODEL DEVELOPMENT

This work sought to (a) demonstrate the ability of non-linear QSPR models to predict skin sensitization of a chemical compound, (b) examine the efficacy of using data from three different skin sensitization test procedures (LLNA, GPMT, and BgVV) for predicting skin sensitization, (c) utilize descriptors from multiple QSPR software to

identify the most significant descriptors for the prediction of skin sensitization, and (d) determine the efficiency of using a combination of literature, functional group and structural descriptors. The following specific case studies were conducted to achieve these objectives:

Case Study 1: Using literature descriptors to develop a QSPR model for skin sensitization.

Case Study 2: Improve the results obtained in Case Study 1 by including functional group and structural descriptors in the QSPR model.

Case Study 3: Develop three separate non-linear QSPR models to describe data from the LLNA, GPM T and BgVV datasets.

Each case study provided valuable guidance to the development of the final QSPR model.

3.4.1 Linear Model Development

Choosing the best set of descriptors that encode the property of interest is a challenging optimization problem. Fast and reliable regression techniques are required to obtain a reduced set of descriptors, based on discarding repetitive and insignificant descriptors through orthogonalization and non-linear sequential reduction. Multi-linear techniques available in commercial software packages are used to obtain linear regression models. CODESSA includes linear regression analysis techniques that include (a) multi-linear regression, (b) principal-component regression, (c) partial least-squares regression, and (d) heuristic regression. In the search for the best multi-linear correlation equation for a large set of descriptors, the major problem is connected with the mutual collinearity of descriptors, which leads to instability of regression coefficients, overestimated standard errors, and critical loss of predictive information. One way to avoid this effect is by

reducing the non-orthogonal set of the natural descriptors into a set of orthogonal factors. The latter, being linear combinations of the natural scales, preserves the information content whereas the negative effects of the scales multi-collinearity is removed in the subsequent regressions. In this work, heuristic regression analysis technique was used to obtain the best linear regression model after orthogonalization and non-linear sequential reduction of descriptors.

3.4.2 Non-Linear Model Development

Linear regression analysis ignores any possible non-linear relationships between the property and its descriptors. Thus, the use of linear approaches can lead to loss of critical information, and the resultant models may have poor predictive abilities. To ensure that the non-linear relationships are accounted for in the QSPR models, the optimal descriptor set from heuristic analysis was retained for ANN analysis.

ANN Analysis: The efficacy and efficiency of supervised learning in multilayer neural networks strongly depends on the network topology, the transfer functions, the learning rule, and the initial values of the weights. Optimal values for these items are usually unknown *a priori* because they depend on the particular training set to be considered and on the nature of the solution. A feed forward back-propagation neural network model with improved network architecture, consistency, data randomization, allocation of training/validation data, and automated network initialization was used for the non-linear model development [41]. The model searches for all possible combinations of hidden layer units that result in a degree of freedom ratio (ratio of the number of network connections and the number of data points) value greater than two.

The input dataset was divided into training, validation and testing sets containing 70%, 15% and 15% of the data, respectively. Overtraining of the network, which results in poor predictive capability, is avoided by use of the cross-validation set with an early-stopping method. The training set is used for computing the gradient and updating the network weights and biases. The validation set chosen should be representative of all points in the training set for optimal performance. The error in predictions for the validation set is monitored during the training process. The validation error normally decreases during the initial phase of training, as does the training set error. However, when the network begins to over fit the data, the error for the validation set begins to rise. When the validation error increases for a specified number of iterations, training is stopped, and the weights and biases at the minimum of the validation error are retained.

By training the network starting from several different initial conditions, the robustness of the network performance can be verified. Multiple randomizations of the data and initializations of the weights are used to obtain the best network, as suggested by Iyer and Rhinehart [42]. The inputs and targets are normalized to have zero mean and unity standard deviation, which ensures that exceptionally large-valued descriptors do not bias the network. The Nguyen-Widrow algorithm is used to initialize weights and biases, which are updated using a Levenberg-Marquardt optimization technique. The transfer functions and the performance function of the network are tailored to find the best possible network. The final network is further evaluated using sum of squared errors, average-absolute deviations, weighted average-absolute deviations, root-mean-squared error, number of wrong classifications and correlation coefficient, when applicable.

3.5. RESULTS AND DISCUSSION

Table 3.4 and 3.5 contain a brief summary of the modeling results. Literature descriptors such as the molecular volume accounting for the size of the molecule, HOMO-LUMO energy gap accounting for the reactivity of the molecule and octanol/water partition coefficient (log K_{ow}) accounting for the skin penetration ability were forced into the final QSPR models. Initially, the QSPR models were developed using only literature descriptors (Case 1). Although these models were comparable to the models cited in the literature, the errors exceeded the desired level. Therefore, around 1200 structural descriptors and 154 functional descriptors were analyzed to obtain a QSPR model with the best predictive ability. Non-linear transformations of the descriptors were evaluated to identify any non-linear relationships during descriptor reduction. Although the inclusion of these descriptors improved the model predictions, additional improvements were desired, and hence, non-linear neural network based models were developed.

The results obtained using literature descriptors are outlined in Table 3.4. Accuracies of 74%, 80% and 73% were obtained for the LLNA, GPMT and BgVV datasets, respectively. In addition to low accuracies, the literature-descriptor models had large deviations from the experimental values. Also, using only these descriptors to predict skin sensitization of diverse chemical compounds may lead to large errors since all the effects may not be captured by these descriptors alone. Therefore, molecular descriptors that can account for other skin sensitization effects in addition to those accounted by the literature descriptors were included to obtain accurate predictions.

Skin sensitization of a compound is often associated with the presence of alerting groups that react covalently with the protein. In this work, we have attempted to identify functional groups that show significant correlation with skin sensitization along with other molecular descriptors. For this purpose, 154 functional group descriptors were generated using Dragon software. These descriptors were added to the molecular descriptor set during the descriptor reduction process. Non-linear transformations of these descriptors were evaluated and a sequential regression analysis technique was used to reduce the set of descriptors for model development. The network performance improved with the inclusion of more descriptors, but the stability of the network was reduced. Hence, deciding on the number of descriptors is a trade-off between the performance and stability of the network.

Our main objective was to develop a QSPR model that can predict accurately the skin sensitization values for a diverse set of compounds. Therefore, careful analysis was made to determine the optimal set of descriptors for the best predictive network. The total number of significant descriptors was found to be 25 for the LLNA and GPMT QSPR models and 22 for the BgVV QSPR model. The final sets of descriptors used in the models are listed in Table 3.6. Descriptors accounting for molecular size and reactivity were found to be significant, along with other functional group descriptors. Further analysis using the descriptor sets was conducted using ANNs. The ANN model results are summarized in Table 3.5. The predictions obtained for each of the three QSPR models improved significantly when structural descriptors were included in the model development. Accuracies of 90%, 95% and 90% were obtained using 25-4-11-1, 25-3-6-

1, and 22-4-1 network architectures for the LLNA, GPMT and BgVV QSPR models, respectively. Figures 3.7-3.9 present the error distribution for each of the models.

To develop the optimum network architecture, multiple neural network runs with mean-absolute error, mean-squared error, sum-squared error, and mean-squared error with regularization performance functions were conducted. The results obtained using mean-squared error performance function was found to give the best results for skin sensitization modeling. Similarly, the network performance was improved by studying networks with multiple transfer functions and number of neurons in the hidden layers.

The current model for skin sensitization improves on other similar literature models in several respects, including (a) use of a large dataset consisting of diverse chemical classes and potency categories; (b) use of a combination of literature and structural descriptors; (c) use of descriptors from multiple QSPR software to assure model superiority and stability; (d) use of non-linear transformations during descriptor reduction to obtain the most suitable set of descriptors; (e) examining the efficacy of both linear and non-linear QSPR models; and (f) use of robust non-linear neural networks with multiple randomizations and initializations to ensure network stability. Our model is also capable of predicting the skin sensitization potency of a molecule on a scale of 0-1. Thus, a reasonable gradation for the level of potency is provided for the chemical compound using the LLNA, GPMT and BgVV QSPR models. In addition, an estimate of the dosage levels can be obtained using the LLNA skin sensitization score.

3.6. CONCLUSIONS

Following are the conclusions drawn based on this study:

1. Our new structure-based, non-linear models are capable of predicting the skin sensitization of a chemical compound with 90% accuracy for the three datasets considered.
2. Descriptors that account for size, reactivity and skin penetration were confirmed to be significant in modeling skin sensitization.
3. The results of this study indicate that the use of structural descriptors coupled with previously identified descriptors provide improved estimates of skin sensitization.
4. The use of three QSPR models to predict skin sensitization of a chemical compound is effective since the end-point ranking system is different in each of these models.
5. Our new approach of identifying non-linear relationships between descriptors and physical property through non-linear transforms during descriptor reduction reduces the drawbacks associated with linear reduction techniques.
6. Using multiple data randomizations, multiple weight initializations, and non-linear descriptor reduction techniques proved to be effective in developing stable non-linear regression models.

REFERENCES

[1] M. S. Roberts and K. A. Walters, *Dermal absorption and toxicity assessment*: Informa Healthcare, 2007.

[2] T. S. Poet and J. N. McDougal, "Skin absorption and human risk assessment," *Chemico-Biological Interactions,* vol. 140, pp. 19-34, 2002.

[3] A. Fedorowicz, L. Zheng, H. Singh, and E. Demchuk, "QSAR study of skin sensitization using local lymph node assay data," *International Journal of Molecular Sciences,* vol. 5, pp. 56-66, 2004.

[4] I. Kimber, D. A. Basketter, G. F. Gerberick, and R. J. Dearman, "Allergic contact dermatitis," *International Immunopharmacology,* vol. 2, pp. 201-211, 2002.

[5] I. Kimber, D. A. Basketter, K. Berthold, M. Butler, J. L. Garrigue, L. Lea, C. Newsome, R. Roggeband, W. Steiling, and G. Stropp, "Skin sensitization testing in potency and risk assessment," *Toxicological Sciences,* vol. 59, pp. 198-208, 2001.

[6] G. F. Gerberick, J. D. Vassallo, L. M. Foertsch, B. B. Price, J. G. Chaney, and J. P. Lepoittevin, "Quantification of chemical peptide reactivity for screening contact allergens: A classification tree model approach," *Toxicological Sciences,* vol. 97, pp. 417-427, 2007.

[7] G. Patlewicz, A. G. Saliner, M. Pavan, A. Worth, R. Benigni, A. Aptula, A. Bassan, C. Bossa, A. F. Filipsson, V. Gillet, N. Jeliazkova, A. McDougal, J. Mestres, I. Munro, T. Netzeva, B. Safford, B. S. Hettich, I. Tsakovska, M. Wallén, and C. Yang, "Chemical similarity and threshold of toxicological concern (TTC) approaches," in *ECB Workshop Ispra*, 2005.

[8] "Guideline for testing of chemicals, #406: Skin sensitization," *Organisation for Economic Cooperation and Development, Paris,* 1992.

[9] M. D. Barratt, D. A. Basketter, M. Chamberlain, M. P. Payne, G. D. Admans, and J. J. Langowski, "Development of an expert system rulebase for identifying contact allergens," *Toxicology In Vitro,* vol. 8, pp. 837-839, 1994.

[10] "The murine local lymph node assay. The results of an independent peer review evaluation coordinated by the interagency coordinating committee on the validation of alternative methods (ICCVAM) and the national toxicology program center for the evaluation of alternative toxicological methods (NICEATM)," *NIH Publication no. 99-4494,* 1999.

[11] O. D. T. Guideline, "429: Skin sensitisation: Local lymph node assay," June, 2001.

[12] G. F. Gerberick, C. A. Ryan, R. J. Dearman, and I. Kimber, "Local lymph node assay (LLNA) for detection of sensitization capacity of chemicals," *Methods,* vol. 41, pp. 54-60, 2007.

[13] S. S. Godavarthy, R. L. Robinson, and K. A. M. Gasem, "Improved structure–property relationship models for prediction of critical properties," *Fluid Phase Equilibria,* vol. 264, pp. 122-136, 2008.

[14] G. Patlewicz, A. O. Aptula, D. W. Roberts, and E. Uriarte, "A minireview of available skin sensitization (Q) SARs/expert systems," *QSAR and Combinatorial Science,* vol. 27, pp. 60-76, 2008.

[15] R. Rodford, G. Patlewicz, J. D. Walker, and M. P. Payne, "Quantitative structure-activity relationships for predicting skin and respiratory sensitization," *Environmental Toxicology and Chemistry,* vol. 22, pp. 1855-1861, 2003.

[16] C. K. S. Pease, D. A. Basketter, and G. Y. Patlewicz, "Contact allergy: The role of skin chemistry and metabolism," *Clinical & Experimental Dermatology,* vol. 28, pp. 177-183, 2003.

[17] G. Patlewicz, S. D. Dimitrov, L. K. Low, P. S. Kern, G. D. Dimitrova, M. I. H. Comber, A. O. Aptula, R. D. Phillips, J. Niemelä, and C. Madsen, "TIMES-SS—A promising tool for the assessment of skin sensitization hazard. A characterization with respect to the OECD validation principles for (Q) SARs and an external evaluation for predictivity," *Regulatory Toxicology and Pharmacology,* vol. 48, pp. 225-239, 2007.

[18] A. O. Aptula and D. W. Roberts, "Mechanistic applicability domains for nonanimal-based prediction of toxicological end points: General principles and application to reactive toxicity," *Chemical Research in Toxicology,* vol. 19, pp. 1097-1105, 2006.

[19] M. T. D. Cronin and D. A. Basketter, "Multivariate QSAR analysis of a skin sensitization database," *SAR and QSAR in Environmental Research,* vol. 2, pp. 159-179, 1994.

[20] J. Devillers, "A neural network SAR model for allergic contact dermatitis," *Toxicology Mechanisms and Methods,* vol. 10, pp. 181-193, 2000.

[21] J. Ashby, D. A. Basketter, D. Paton, and I. Kimber, "Structure activity relationships in skin sensitization using the murine local lymph node assay," *Toxicology,* vol. 103, pp. 177-194, 1995.

[22] G. Y. Patlewicz, D. A. Basketter, C. K. S. Pease, K. Wilson, Z. M. Wright, D. W. Roberts, G. Bernard, E. G. Nezarnau, and J.-P. Lepoitteven, "Further evaluation of quantitative structure–activity relationship models for the prediction of the skin sensitization potency of selected fragrance allergens," *Contact Dermatitis,* vol. 50, pp. 91–97, 2004.

[23] S. D. Dimitrov, L. K. Low, G. Y. Patlewicz, P. S. Kern, G. D. Dimitrova, M. H. I. Comber, R. D. Phillips, J. Niemela, P. T. Bailey, and O. G. Mekenyan, "Skin sensitization: Modeling based on skin metabolism simulation and formation of

protein conjugates," *International Journal of Toxicology,* vol. 24, pp. 189-204, 2005.

[24] Y. Li, D. Pan, J. Liu, P. S. Kern, G. F. Gerberick, A. J. Hopfinger, and Y. J. Tseng, "Categorical QSAR models for skin sensitization based upon local lymph node assay classification measures part 2: 4D-fingerprint 3-state and two-2-state logistic regression models," *Toxicological Sciences,* vol. 99, pp. 532-544, 2007.

[25] G. F. Gerberick, C. A. Ryan, P. S. Kern, H. Schlatter, R. J. Dearman, I. Kimber, G. Y. Patlewicz, and D. A. Basketter, "Compilation of historical local lymph node data for evaluation of skin sensitization alternative methods," *Dermatitis,* pp. 157-202, 2005.

[26] K. Langton, G. Y. Patlewicz, A. Long, C. A. Marchant, and D. A. Basketter, "Structureactivity relationships for skin sensitization: recent improvements to Derek for Windows," *Contact Dermatitis,* vol. 55, pp. 342-347, 2006.

[27] E. Schlede, W. Aberer, T. Fuchs, I. Gerner, H. Lessmann, T. Maurer, R. Rossbacher, G. Stropp, E. Wagner, and D. Kayser, "Chemical substances and contact allergy—244 substances ranked according to allergenic potency," *Toxicology,* vol. 193, pp. 219-259, 2003.

[28] G. Patlewicz, Personal communication, UniLever, UK, 2007.

[29] ChemBioOffice 11.0, Cambridge Software, 2008.

[30] K. E. Haneke, R. R. Tice, B. L. Carson, B. H. Margolin, and W. S. Stokes, "ICCVAM evaluation of the murine local lymph node assay III. Data analyses completed by the national toxicology program interagency center for the evaluation of alternative toxicological methods," *Regulatory Toxicology and Pharmacology,* vol. 34, pp. 274-286, 2001.

[31] C. Morley, "OpenBabelGUI 2.0.2," 2006.

[32] AMPAC 6.0., Semichem Inc., 1997.

[33] CODESSA 2.7.8., Semichem Inc., 2007.

[34] Talete srl, DRAGON 5.4 for Windows (Software for molecular descriptor calculations), 2006.

[35] S. S. Godavarthy, "Design of improved solvents for extractive distillation," Ph. D. Dissertation, Oklahoma State University, Stillwater, OK, 2004.

[36] A. Jegadeesan, "Structure-based generalized models for selected pure-fluid saturation properties," Master Thesis, Oklahoma State University, Stillwater, OK, 2006.

[37] S. S. Godavarthy, R. L. Robinson, and K. A. M. Gasem, "Improved structure–property relationship models for prediction of critical properties," *Fluid Phase Equilibria,* vol. 264, pp. 122-136, 2007.

[38] M. D. Barratt, "Prediction of toxicity from chemical structure," *Cell Biology and Toxicology,* vol. 16, pp. 1-13, 2000.

[39] J. Aihara, "Correlation found between the HOMO–LUMO energy separation and the chemical reactivity at the most reactive site for isolated-pentagon isomers of fullerenes," *Physical Chemistry Chemical Physics,* vol. 2, pp. 3121-3125, 2000.

[40] A. Nguyên Trong, "What can we do with frontier orbitals?," in *Frontier Orbitals*, 2007, pp. 1-4.

[41] H. Demuth, M. Beale, and M. Hagan, "Neural Network Toolbox™ 6," MathWorks Inc., 2008.

[42] M. S. Iyer and R. R. Rhinehart, "A method to determine the required number of neural-network training repetitions," *IEEE Transactions on Neural Networks,* vol. 10, pp. 427-432, 1999.

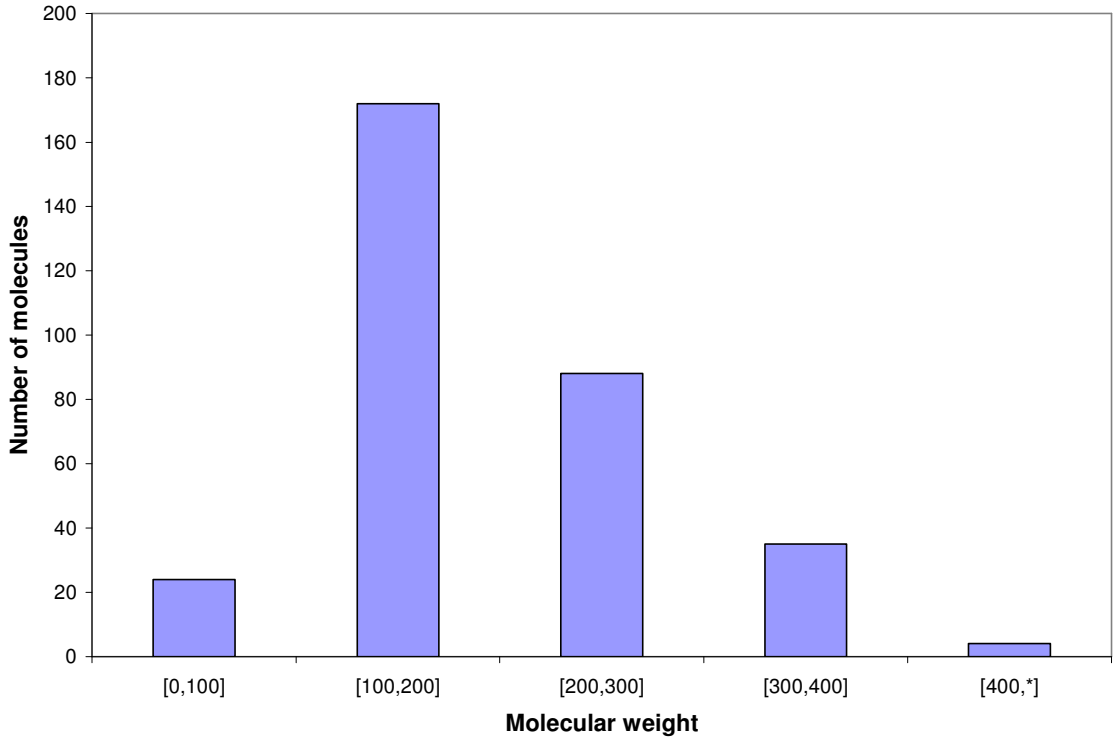


Figure 3.1. Molecular weight distribution of the LLNA chemical compounds

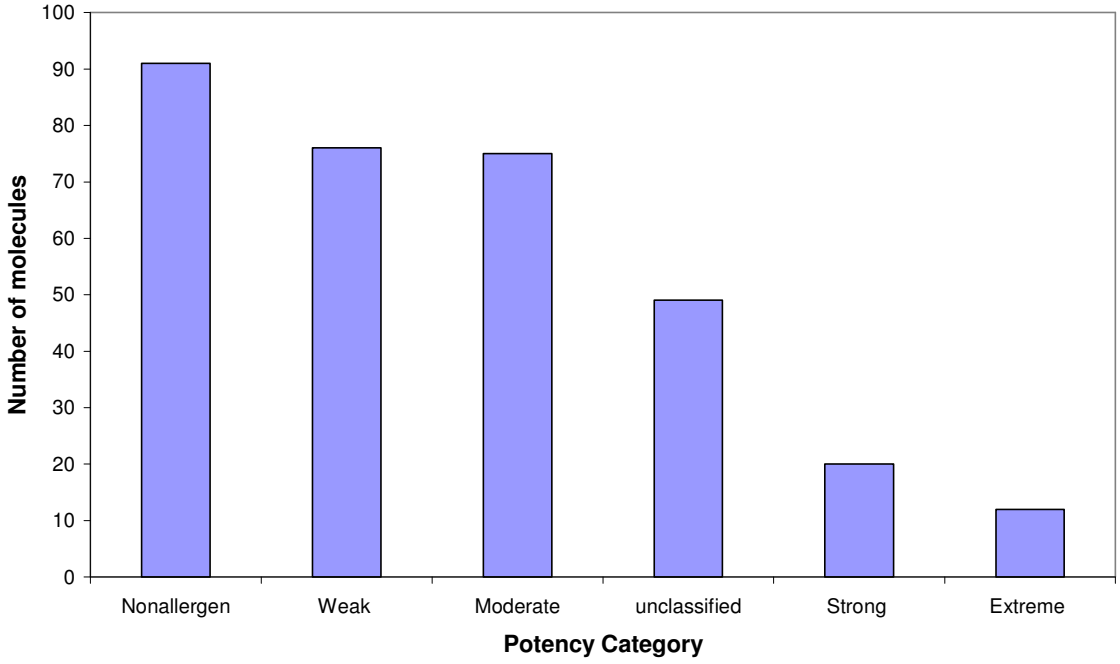


Figure 3.2. Potency distribution of the LLNA chemical compounds

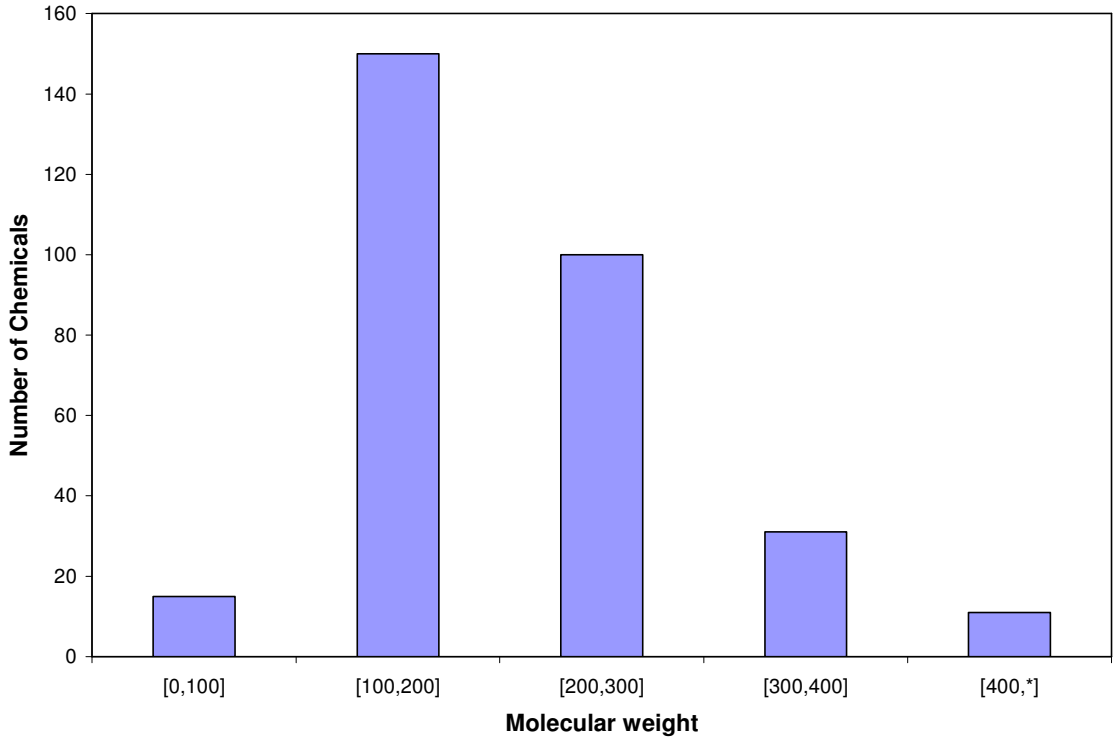


Figure 3.3. Molecular weight distribution of the GPMT chemical compounds

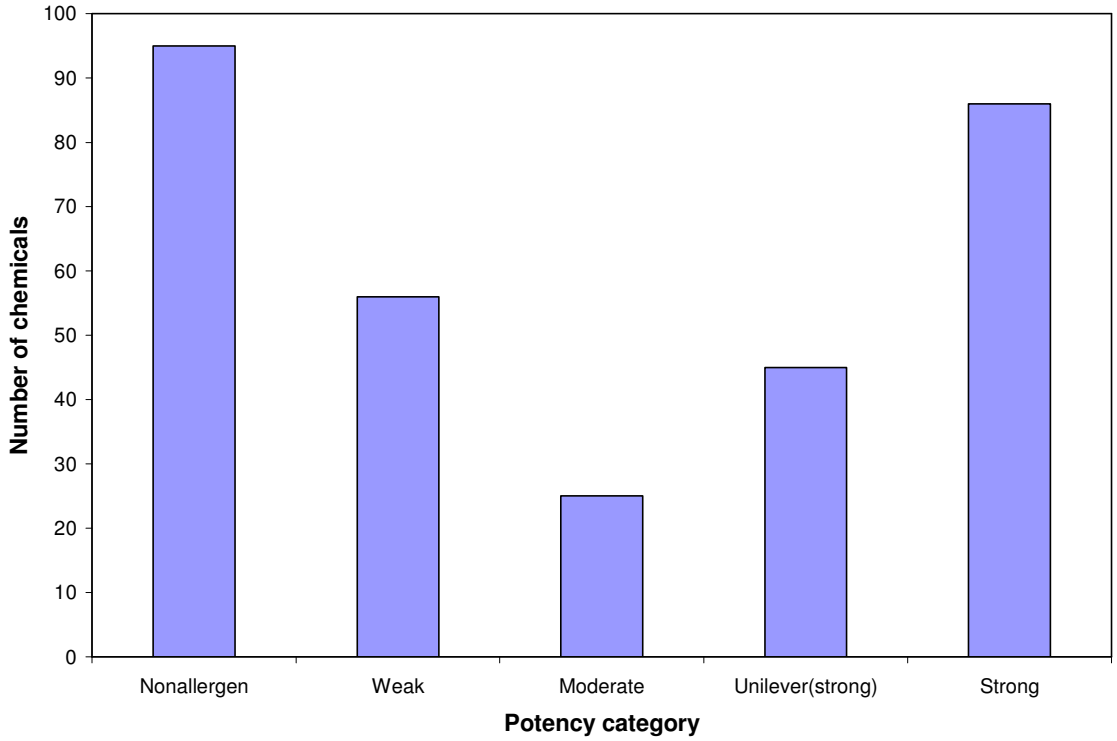


Figure 3.4. Potency distribution of the GPMT chemical compounds

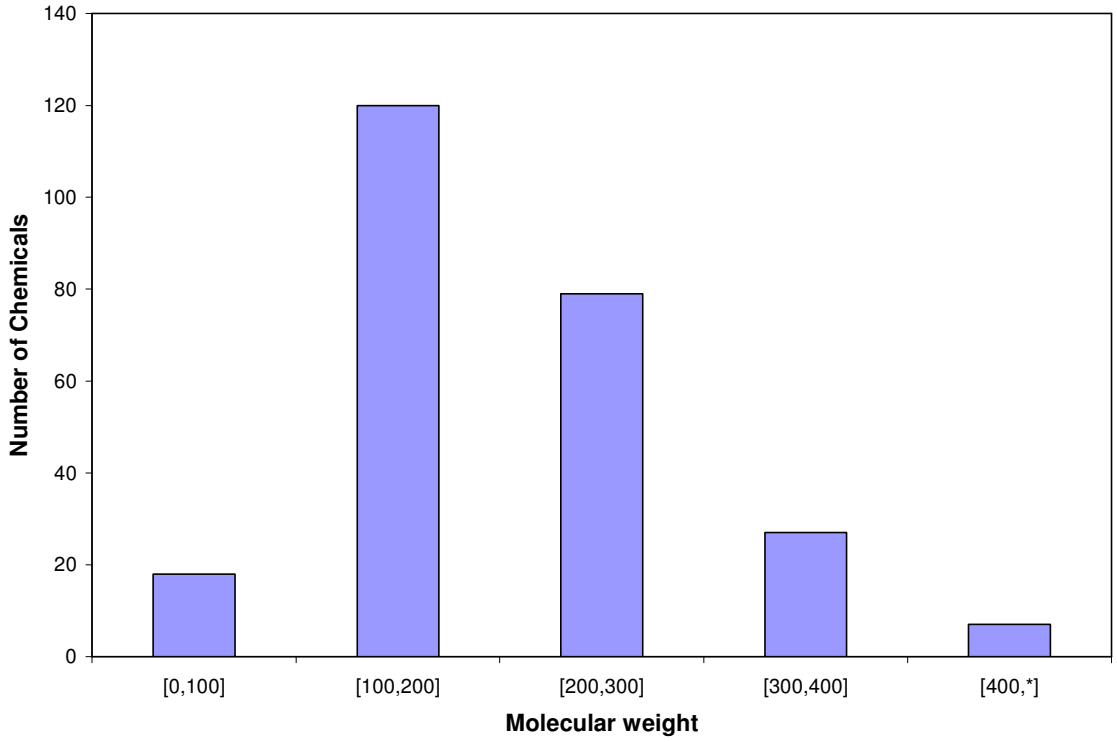


Figure 3.5. Molecular weight distribution of the BgVV chemical compounds

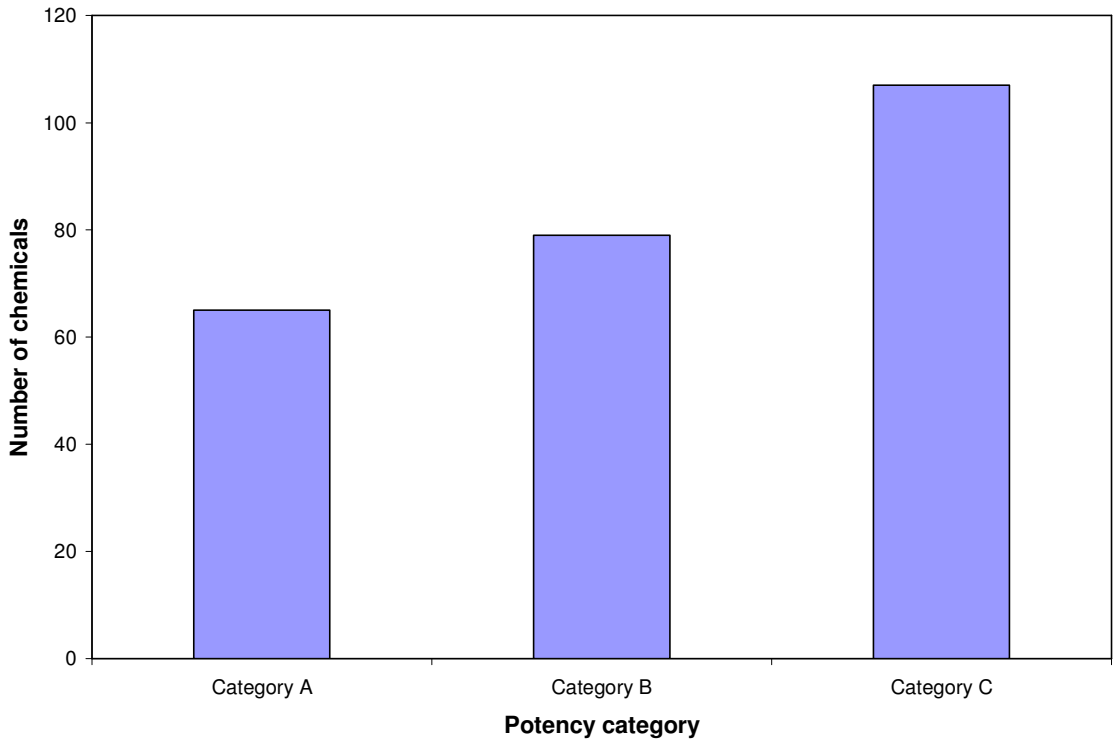


Figure 3.6. Potency distribution of the BgVV chemical compounds

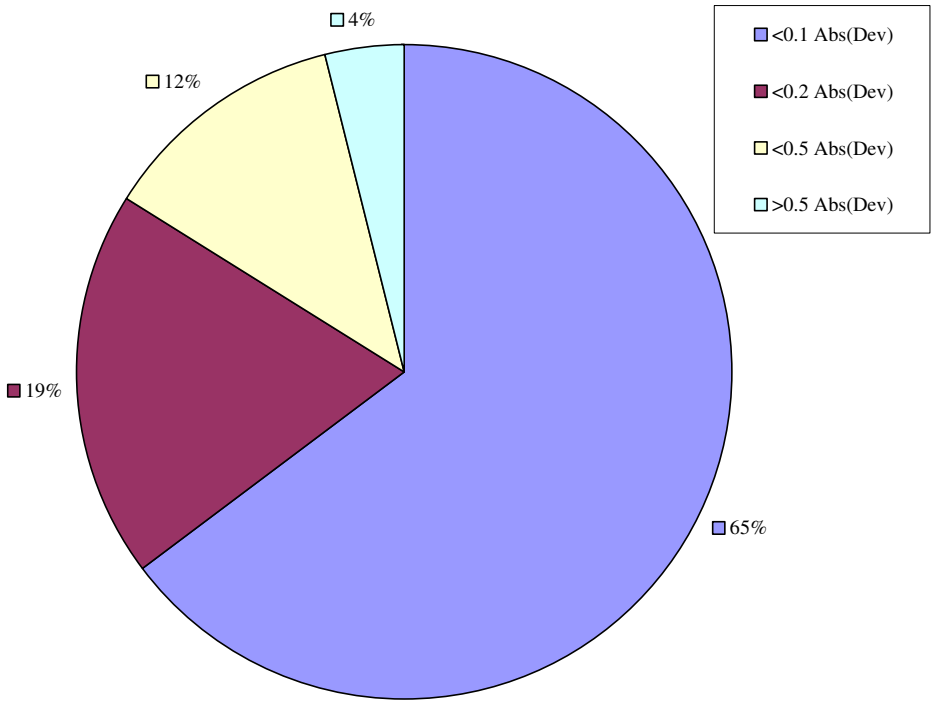


Figure 3.7. Distribution of absolute deviation in predicted LLNA skin sensitization score

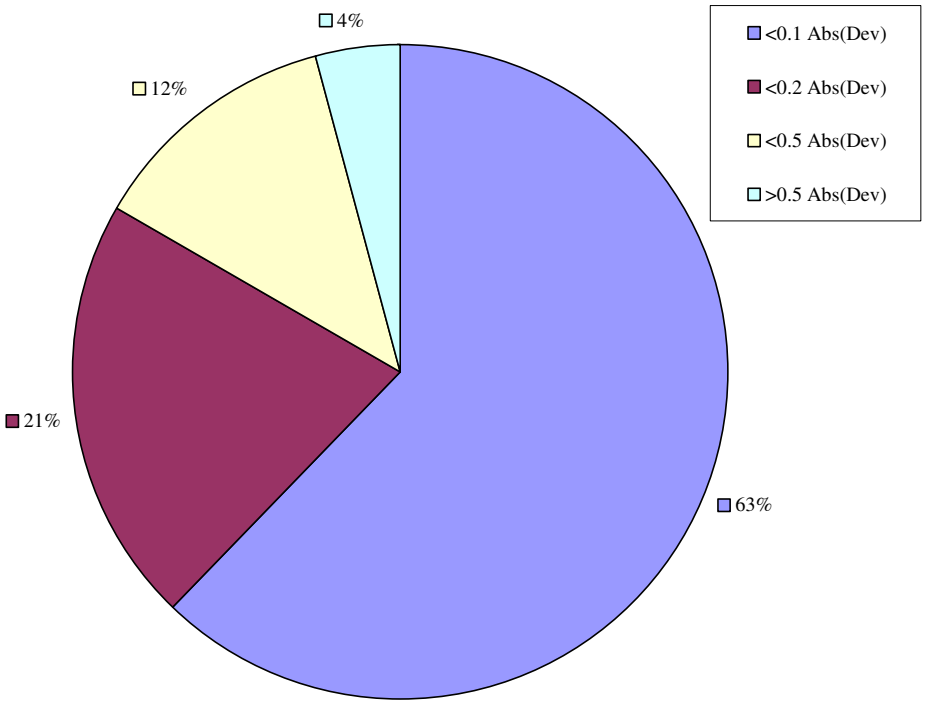


Figure 3.8. Distribution of absolute deviation in predicted GPMT skin sensitization score

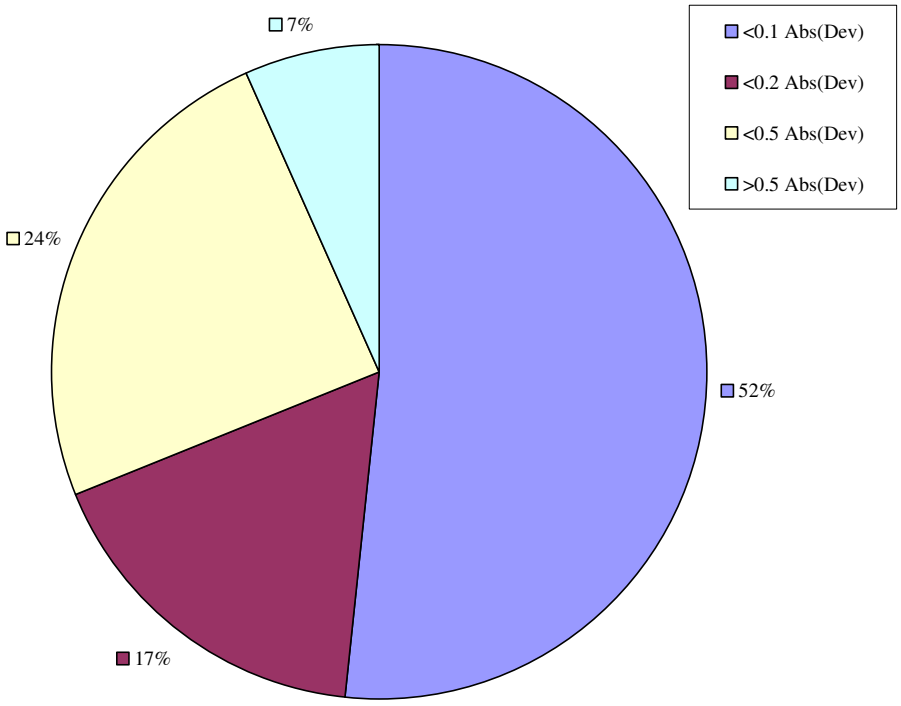


Figure 3.9. Distribution of absolute deviation in predicted BgVV skin sensitization score

Table 3.1. Skin sensitization scores of chemical compounds based on classification

LLNA		GPMT		BgVV	
Classification	**Toxicity score**	**Classification**	**Toxicity score**	**Category**	**Toxicity score**
Non-sensitizer	0	Non-sensitizer	0	C	0
Weak sensitizer	0.25	Weak sensitizer	0.33	B	0.5
Moderate sensitizer	0.5	Moderate sensitizer	0.66	A	1
Sensitizers(unclassified)	0.625	Strong-sensitizers(Unilever)	0.83		
Strong sensitizer	0.75	Strong sensitizer	1		
Extreme sensitizer	1				

Table 3.2. Comparative evaluation of the LLNA database

Comparison	Number	Sensitivity[1]		Specificity[2]		Positive predictivity[3]		Negative predictivity[4]		Accuracy[5]	
		%	Number	%	Number	%	Number	%	Number	%	Number
LLNA vs. GPMT	97	91%	(62/68)	83%	(24/29)	93%	(62/67)	80%	(24/30)	89%	(86/97)
LLNA vs. human	74	72%	(49/68)	67%	(4/6)	96%	(49/51)	17%	(4/23)	72%	(53/74)
GPMT vs. human	57	70%	(38/54)	100%	(3/3)	100%	(38/38)	16%	(3/19)	72%	(41/57)

1 Sensitivity: The proportion of all positive chemicals that are correctly classified as positive in a test.
2 Specificity: The proportion of all negative chemicals that are correctly classified as negative in a test.
3 Positive predictivity: The proportion of correct positive responses among materials testing positive.
4 Negative predictivity: The proportion of correct negative responses among materials testing negative.
5 Accuracy: The proportion of correct outcomes.

Table 3.3. Comparative evaluation of the LLNA database limited to compounds with LLNA, guinea pig, and human data

Comparison	Number	Sensitivity		Specificity		Positive predictivity		Negative predictivity		Accuracy	
		%	Number	%	Number	%	Number	%	Number	%	Number
LLNA vs. human	57	72%	(39/54)	67%	(2/3)	98%	(39/40)	12%	(2/17)	72%	(41/57)
GPMT vs. human	57	70%	(38/54)	100%	(3/3)	100%	(38/38)	17%	(3/19)	72%	(41/57)
LLNA vs. human	62	73%	(43/59)	67%	(2/3)	98%	(43/44)	11%	(2/18)	73%	(45/62)

Table 3.4. Performance of QSPR models using literature descriptors

Comparison	Data Points	Sensitivity		Specificity		Accuracy	
		%	Number	%	Number	%	Number
LLNA	358	77%	(134/175)	72%	(132/183)	74%	(266/358)
GPMT	307	90%	(141/156)	70%	(107/151)	80%	(248/307)
BgVV	251	66%	(71/107)	83%	(120/144)	73%	(191/251)

Table 3.5. Performance of QSPR models using a combination of literature and structural descriptors

Comparison	Data points	Sensitivity		Specificity		Accuracy	
		%	Number	%	Number	%	Number
LLNA	358	92%	(154/175)	88%	(168/183)	90%	(322/358)
GPMT	307	95%	(148/156)	95%	(144/151)	95%	(292/307)
BgVV	251	82%	(88/107)	96%	(138/144)	90%	(223/251)

Table 3.6. Final set of descriptors used in QSPR models

LLNA	
Relative number of N atoms	nArCOOH
Number of benzene rings	nRCOOR
Average Structural Information content (order 0)	nRCONHR
Average Information content (order 2)	nArCONHR
Molecular volume	nRCOX
HOMO - LUMO energy gap	nArCHO
Min net atomic charge for a C atom	nArCO
HASA-1/TMSA [Quantum-Chemical PC]	nROH
Max valency of a C atom	kow
Min resonance energy for a C-C bond	Max electroph. reaction index for a S atom
Zero point vibrational energy	Max total interaction for a Br-C bond
nCs	Max electroph. reaction index for a F atom
nR=Cs	
GPMT	
Molecular volume	nCbH
DPSA-1 Difference in CPSAs (PPSA1-PNSA1) [Zefirov's PC]	nArCONHR
HOMO - LUMO energy gap	nArOCON
Min electroph. reaction index for a C atom	nRCOX
Max electroph. reaction index for a C atom	nRCHO
WNSA-1 Weighted PNSA (PNSA1*TMSA/1000) [Quantum-Chemical PC]	nArNH2
FPSA-3 Fractional PPSA (PPSA-3/TMSA) [Quantum-Chemical PC]	nRNHR
FNSA-3 Fractional PNSA (PNSA-3/TMSA) [Quantum-Chemical PC]	nOHs
RPCS Relative positive charged SA (SAMPOS*RPCG) [Quantum-Chemical PC]	kow
RNCS Relative negative charged SA (SAMNEG*RNCG) [Quantum-Chemical PC]	Min 1-electron reaction index for a Cl atom
Min (>0.1) bond order of a C atom	Min 1-electron reaction index for a Br atom
Max bond order of a H atom	Max total interaction for a C-S bond
nCrt	

Table 3.6 (cont'd). Final set of descriptors used in QSPR models

BgVV	
Molecular volume	WNSA-3 Weighted PNSA (PNSA3*TMSA/1000) [Zefirov's PC]
LUMO energy	Relative number of N atoms
Kow	1X BETA polarizability (DIP)
RPCG Relative positive charge (QMPOS/QTPLUS) [Zefirov's PC]	nArOR
Average Bonding Information content (order 0)	nOxiranes
DPSA-2 Difference in CPSAs (PPSA2-PNSA2) [Zefirov's PC]	Max partial charge for a O atom [Zefirov's PC]
Min partial charge for a C atom [Zefirov's PC]	nArNO2
Number of bonds	Molecular surface area
Max SIGMA-PI bond order	Avg 1-electron reaction index for a N atom
HA dependent HDCA-2/TMSA [Quantum-Chemical PC]	Max 1-electron reaction index for a N atom
HA dependent HDSA-1 [Quantum-Chemical PC]	nROCON

CHAPTER 4

QUANTITATIVE STRUCTURE-PROPERTY RELATIONSHIPS MODELING OF SKIN IRRITATION

4.1. INTRODUCTION

Skin diseases and injuries are the most common job-related problems in industries such as manufacturing, food production, construction, machine tool operation, printing, metal plating, leather processing, engine service, landscaping, farming, or forestry. All industrially important chemicals are assessed for their skin irritation or corrosive ability and the results are listed in Material Safety Data Sheets (MSDS) to ensure the safety of the workers. According to the Organization for Economic Co-operation and Development (OECD) test guideline 404 [1], skin irritation is defined as "the production of reversible damage of the skin following the application of a test substance for up to 4 hours." For many years, the Draize rabbit skin test has been widely accepted for assessing the skin irritation potential of chemicals. In this test, 0.5 mL of the chemical of interest is applied to albino rabbit skin for four hours and kept under clinical observation for 14 days for signs of erythema and edema [2]. Although the Draize test provides reliable estimates of skin irritation, it is often criticized as being cruel to the test animals; hence, the incentive for developing predictive procedures for estimating skin irritation potential of chemicals has been increasing [3].

In the last few decades, various *in vitro* and *in silico* skin irritation test methods have been proposed as replacements for *in vivo* tests. Promising *in vitro* methods that

have been evaluated include EpiDerm, EPISKIN, PREDISKIN, the non-perfused pig ear model, and the mouse skin integrity function test (SIFT) [3]. However, none of the alternate test procedures has been accepted as a valid replacement to the Draize test by OECD [4]. *In silico* methods for predicting skin irritation include expert systems and structure-activity models that express the skin irritation potential of a chemical as a function of a set of physiochemical properties and structural descriptors. The application of these *in silico* models is limited by the availability of the necessary input physiochemical properties such as octanol/water partition coefficient, melting point, lipid solubility, aqueous solubility, surface tension and vapor pressure. Often the properties of the novel chemicals are not readily available, rendering these *in silico* models inapplicable. The limited applicability of these procedures calls for development of novel prediction methods that use only structural descriptors to estimate the skin irritation potential.

Quantitative structure-property relationships (QSPR) modeling offers an attractive alternative because successful models have the potential to provide reliable property estimates based solely on chemical structure information. That is, if structural information can be successfully decoded, the properties can be determined simply from the chemical structure.

QSPR studies have gained impetus with the advent of high-speed computers. This has not only led to the development of new and more complex molecular descriptors, but it has also been instrumental in the application of QSPR models to properties that were previously infeasible due to computational intensity. QSPR models are now well established and are used to correlate varied, and often complex, physiochemical

properties of molecules. The QSPR approach has been applied in different areas and a detailed review of its applications can be found in one of our earlier publications [5, 6].

As outlined in the next section, various QSPR models have been proposed for predicting skin irritation potential of molecules. However, these models are restricted to a limited number of classes of molecules, thus lacking in their general applicability.

In this work, we have developed a skin irritation QSPR model using rabbit Draize test data for 189 compounds consisting of chemicals from various classes. We have evaluated the predictive ability of several QSPR models based on literature descriptors, group contribution descriptors and structural descriptors, used alone or in combination. A QSPR model utilizing only structural descriptors was developed first, then its predictive ability was improved by using a combination of the literature, functional group and structural descriptors. Robust artificial neural network models with superior capabilities were used for non-linear model development. The physical significance of the final set of descriptors was studied and external validation of the QSPR model was performed using data from human-patch tests.

4.2. SKIN IRRITATION MODELS

Although *in vivo* tests for skin irritation cannot be replaced entirely, they can be reduced considerably by initial screening of chemical compounds using *in vitro* and *in silico* methods. While a number of *in vitro* methods have been proposed in the literature, to date, there are no validated *in vitro* tests that can replace the Draize test. A detailed review of the currently available *in vitro* techniques is discussed elsewhere [7, 8]. The *in silico* techniques offer an attractive alternative due to their ease of use and low cost in

comparison to the *in vitro* techniques. In this section, a review of the *in silico* techniques available in the literature is presented. A comprehensive review of the available *in silico* techniques for predicting skin irritation potential of chemical has been reported by Saliner *et al.* [9].

Physiochemical properties that have been found to be significant for modeling skin irritation include molecular volume, dipole moment, molecular shape, the over all number of atoms, steric effects, molar refractivity, acidic dissociation constant (pK_a), absolute hardness, and the octanol/water partition coefficient (log K_{ow}) [9]. Among these properties, log K_{ow} has been found to be most significant since it provides a quantitative measure of partitioning between aqueous and lipid phases [10]. One of the earliest attempts to model skin irritation using structure-activity relationships (SARs) was done by Enslein *et al.* [11]. Separate models were developed for aromatic and aliphatic compounds using molecular connectivity indices, sub-structural keys and molecular length parameters. Since the database used did not provide numerical scores, the models use skin irritation severity ratings. Smith *et al.* [12] developed a SAR model to discriminate skin irritant esters from non-irritant esters using nineteen physiochemical parameters that represent the transport, electronic and steric properties. Human skin irritation data of 42 esters were used to generate ten sub-models using multiple random sampling of the database. The sensitivity and specificity values ranged from 0.846 to 0.923 and 0.615 to 0.923, respectively. The results indicate that physiochemical parameters of esters relate to their skin irritation effects in humans, and chemical partitioning and intermolecular reactions are important components of the response. Smith *et al.* [13] developed an iterative SAR model for human skin irritation. The model

predictions were validated experimentally and these test results were incorporated into the database to refine the model. A total of 34 irritants from the rabbit test were selected of which 16 were predicted by the model to be positive and 18 to be negative. These chemicals were further tested experimentally using human patch test and the results were incorporated into the database to refine the model. However, as the SAR models were based on limited data, the accuracy of the models was not satisfactory, thus emphasizing the need for experimental validation of models and their further refinement as new data become available.

Hayashi *et al.* [14] studied QSAR for skin irritation potential using 24 phenols. Absolute hardness, LUMO and log K_{ow} were used to fit a regression function to the skin irritation scores obtained from a rabbit Draize test study. An R value of 0.85 was obtained. The model predictions were further validated using a set of six additional phenols, and good correlations with the expected skin irritation scores were observed. Berner *et al.* [15] studied the influence of pK_a on skin irritation and found the two to be highly correlated for the chemical compounds examined. Kodithala *et al.* [16] used a membrane interaction QSAR (MI-QSAR) technique to predict the skin irritation potential of 20 hydroxy organic compounds. The MI-QSAR skin irritation predictions were compared to the traditional 2-D-QSAR predictions to prove the superiority of MI-QSAR approach.

In addition to the QSAR models, a number of expert systems that use a knowledge base to determine the skin irritation potential of a chemical have been developed. These expert systems find the skin irritation potential of a chemical either by relating the existing physiochemical properties of chemicals or through pattern

recognition algorithms. A number of papers describing the currently available expert systems have been published [3, 9, 10, 17]. Some of the widely used expert systems are German Federal Institute for Risk Assessment Decision Support System (BfR-DSS), DEREK for Windows (DEREKfW), HazardExpert, MULTICASE, Computer-Aided Chemistry (CAChe) program, Substructure-Based Computerized Chemical Selection Expert System (SuCCSES), OASIS, and TOxicity Prediction Komputer-Assisted Technology (TOPKAT). Saliner *et al.* [9] have provided a detailed review of each of these expert systems. The expert systems are applicable to a wide range of chemical structures in contrast to the currently available QSAR models. However, the application of these models to predict skin irritation requires as input several physiochemical properties such as octanol/water partition coefficient, melting point, lipid solubility, aqueous solubility, surface tension and vapor pressure. Thus, prediction of skin irritation potential using these models often requires other models capable of estimating the required input physiochemical properties.

4.3. QSPR METHODOLOGY

The development of a QSPR model for skin irritation involves several distinct steps: (a) compilation of an experimental data set, (b) generation and optimization of 3-D molecular structures, (c) calculation of molecular descriptors, (d) reduction of the number of descriptors, and (e) development of a regression model. To begin, a data set of molecules chosen from reliable sources must be assembled. Once this dataset is analyzed using target property distribution and chemical diversity, the 3-D structures of the molecule are generated using commercial molecular visualization software. Structure

optimization and descriptor generation are performed using commercial QSPR software. Model development usually begins with linear analysis and algorithms such as multiple-linear regression analysis, principal component analysis, heuristic analysis or partial least squares may be used. In the optimal scenario, the best set of descriptors is then used to develop a non-linear QSPR model using artificial neural network algorithms.

4.3.1 Dataset Compilation

High quality data generated through good laboratory practice and complying with the OECD test guidelines has been used for modeling. The majority of the data are acquired from Bagley *et al.* [18], who cite Draize skin irritation test data for 176 chemicals. However, since 29 structures were tested more than once, a total of 215 data points are obtained. Other sources of Draize test data include publications by Hayashi *et al.* [14] and Kodithala *et al.* [16], who report skin irritation data for 30 phenols and 22 hydroxy organic compounds, respectively. The primary irritation index (PII) scores reported in these sources were used for modeling.

Oklahoma State University (OSU) Draize test database: A database containing experimental Draize skin irritation test data of 205 molecules has been compiled from the above sources. The database contains experimental data including the chemical classes represented by acids, acrylates, alcohols, aldehydes, amides, amines, brominated derivatives, chlorinated solvents, esters, ethers, fatty acids, halogenated aromatics, hydrocarbons, inorganics, ketones, nitrils, phenols, sulphur containing compounds, and triglycerides. The skin irritation effects of the chemicals have been graded according to the scale proposed by OECD test guideline 404 [1, 2] shown in Table 4.1. The chemicals

are graded for erythema and edema using experimental data and a PII is calculated from these grades using the following equation:

$$\text{Primary Irritation Index (PII)} = \frac{\sum \text{erythema grades at 24/48/72 hr} + \sum \text{edema grades at 24/48/72 hr}}{3 \times \text{number of animals}}$$

The value of PII ranges from 0 to 8, where chemicals with 0 PII value have no erythema and edema formation, and those with a value of 8 have severe effects. Chemicals for which there are multiple PII test data were averaged to get a unique skin irritation score. Since the Draize test data available in the open literature were limited to a few hundred compounds, efforts were made to utilize all reliable experimental data for modeling. Molecules for which the PII could not be determined due to the severity of effects were given a score of 8, the maximum possible PII value. Experimental data for mixtures and fragrance oils were not included in the database.

Of the 205 molecules, only 189 structures were considered for modeling. A total of 16 molecules were rejected due to one the following reasons: (a) structure could not be found, (b) structure could not be optimized using ChemDraw [19], or (c) molecular descriptors could not be generated for the structure.

In addition to chemical diversity, for the model to be able to predict skin irritation efficiently, the range of potencies used to train the model should be adequate. Figure 4.1 shows the PII distribution of the data considered in the present modeling.

4.3.2 Structure Generation and Optimization

The first step in QSPR modeling is the generation of chemical structure for each molecule included in the modeling effort. Various schemes for chemical representation have been proposed in the literature. For example, Open Babel software [20] includes about 80 different representations for a given molecular structure. The most common way

of representing a chemical is a two-dimensional (2-D) sketch. However, using a 2-D representation does not provide a complete description of the molecule and lacks shape and surface distribution information of the molecule. To develop an effective QSPR model, the representations of the molecules should provide all the necessary structural information. This requires information about the atoms present, along with three dimensional (3-D) coordinates that provide a full spatial depiction of the molecule. A commercial package, ChemDraw [19] was used to generate the 2-D structures of the molecules. The 3-D structures were generated for these molecules using Chem3DUltra [19]. Since more than one set of 3-D coordinates that satisfy the structural constraints (bond length and bond angle) can be generated for a given molecule, the conformation with the lowest energy must be located. The structures were initially optimized using the Chem3D module available in Chem3DUltra. To locate the lowest energy configuration, multiple initializations were used during the structure optimization. AMPAC 6.0 [21] was then used to further refine the 3-D geometry of the structures. The output files from AMPAC were used to calculate various descriptors.

4.3.3 Descriptor Generation

The final optimized structure from AMPAC was used as input to commercial QSPR software to generate over 1200 molecular descriptors for a given molecule. A variety of constitutional, topological, geometrical, thermodynamic, quantum-chemical and electrostatic descriptors are generated by CODESSA [22] and 154 functional group descriptors are generated by Dragon [23]. The number of descriptors calculated for each molecule depends on the structural complexity of the molecule. Descriptors that were not

pertinent for a given molecule were set to zero in the subsequent QSPR model development.

4.3.4 Descriptor Reduction

Not all the descriptors generated for a molecule are significant in modeling. The use of all available descriptors in the model development effort causes dimensionality problems and diminishes the performance of a QSPR model, especially when non-linear algorithms are used in model development [6]. Descriptor reduction is the process of automating the discovery of potentially useful correlations from large sets of descriptor data. This process involves the identification of the most relevant set of descriptors for model development and is an important step in QSPR modeling [24]. Different methods for reduction are available in the literature. The most widely used techniques are the principal-component analysis (PCA), partial least-squares (PLS), genetic algorithms (GA), and neural networks (NNs) [6]. Most QSPR models developed are multiple-linear regression correlations, which require *a priori* assumption of the form of the mathematical correlation between the property and its descriptors. However, linear regression analysis ignores any non-linear relationships between the descriptors and properties. The use of linear approaches often leads to loss of critical information and results in models with poor predictive abilities [25].

To ensure that the non-linear relationships are accounted for in the QSPR models, non-linear transformations of all the descriptors were calculated, and an expanded set of descriptors was generated. The expanded set of descriptors was used to find the best set of descriptors through sequential selection. The best descriptor from the expanded set was selected and combined sequentially with the remaining descriptors. Then, the best

two-descriptor combination was retained and combined with each of the remaining descriptors. The sequence continued until a set of 40 descriptors was identified. Heuristic regression available in CODESSA was used to further reduce the number of descriptors. The optimal set of descriptors was retained for artificial neural network (ANN) analysis.

4.3.5 Literature Descriptors

If a mechanistic interpretation of the relationship between the property and its descriptor set can be formulated, it provides increased confidence in the accuracy and validity of the model [26]. However, such interpretation is often not possible for all the descriptors due to the complexities involved. A detailed review of the literature indicates that the chemical reactivity, molecular size, and skin permeability are important determinants of skin irritation. The chemical reactivity of a molecule can be determined using the HOMO and LUMO energies. Molecules with low HOMO-LUMO energy gaps have low kinetic stability and are favorably inclined to add or remove electrons, thus forming activated complexes [27]. Although skin permeability is significant in modeling skin irritation, the absence of readily available skin permeation prediction models often leads to the use of log K_{ow} values to represent the skin permeation ability of a chemical.

4.4. QSPR MODEL DEVELOPMENT

The present work aimed at (a) demonstrating the ability of non-linear QSPR modeling to predict the skin irritation potential of a chemical compounds, (b) examining the efficacy of using a combination of literature, functional group and structural descriptors for model development, (c) using descriptors from multiple QSPR software to identify the most significant descriptors, and (d) validating the final QSPR model using

an external dataset. The following specific case studies were conducted to achieve these objectives:

Case 1: Developing a QSPR model using literature descriptors.

Case 2: Identifying any significant functional group contributions to skin irritation using functional group descriptors from Dragon.

Case 3: Developing a QSPR model to predict skin irritation using only structural descriptors from CODESSA.

Case 4: Developing a final robust QSPR model using a combination of literature, functional group and structural descriptors.

Case 5: Validating the final QSPR model using skin irritation data from other experimental tests.

The case studies conducted in the sequence above provided valuable guidance to the development of the final QSPR model.

4.4.1 Linear Model Development

Selecting the set of descriptors that best encodes information on the property of interest is a difficult optimization problem. Fast and reliable regression techniques are required to obtain a reduced set of descriptors after discarding the correlated and insignificant descriptors through orthogonalization and non-linear sequential reduction [6]. Multi-linear techniques available in commercial software packages are used to obtain linear regression models. The following linear regression analysis techniques are provided in CODESSA: (a) multi-linear regression, (b) principal-component regression, (c) partial-least squares regression, and (d) heuristic regression. In the search for the best multi-linear correlation equation for a large set of descriptors, the major problem is

connected with mutual collinearity of descriptors which leads to instability of regression coefficients, overestimated standard errors, and critical loss of predictive information [6]. One way to avoid this effect is by transforming the non-orthogonal set of the natural descriptors into a set of orthogonal factors. The latter, being linear combinations of the natural descriptors, preserves the information content while removing the negative effects of the scales multi-collinearity from the subsequent regressions. In this work, the heuristic regression analysis technique available in CODESSA was used to obtain the best linear regression model after orthogonalization and non-linear sequential reduction of descriptors.

4.4.2 Non-Linear Model Development

Linear regression analysis ignores the non-linear relationships between property and descriptors. To ensure that the non-linear relationships are accounted for in the QSPR models, the optimal descriptor set from the heuristic analysis was retained for ANN analysis.

ANN Analysis: The efficacy and efficiency of supervised learning in multilayer neural networks strongly depends on the network topology, the transfer functions, the learning rule, and the initial values of the weights. Optimal instances for these items are usually unknown *a priori* because they depend mainly on the particular training set to be considered and on the nature of the solution. A feed forward back-propagation neural network model with an improved network architecture, consistency, data randomization, allocation of training/validation data, and automated network initialization is used for the non-linear model development [28]. The model searches for all possible combinations of hidden layer units that result in a degree of freedom ratio (ratio of the number of network

connections and the number of data points) value greater than two. The input dataset is divided into training, validation and testing sets with 70% of data assigned to training set, 15% to validation set and 15% to testing set. Overtraining of the network, which results in poor predictive capability, is avoided by use of the cross-validation set with an early-stopping method. The training set is used for computing the gradient and updating the network weights and biases. The validation set chosen should be representative of all molecular types in the training set for optimal performance. The error on the validation set is monitored during the training process. The validation error normally decreases during the initial phase of training, as does the training set error. However, when the network begins to over-fit the data, the error in the validation set typically begins to rise. When the validation error increases for a specified number of iterations, the training is stopped, and the weights and biases at the minimum validation error are retained.

By training the network starting from several different initial conditions the robustness of the network performance can be verified. Multiple randomizations of the data and initializations of the weights are used to obtain the best network, as suggested by Iyer and Rhinehart [29]. The inputs and targets are normalized to have zero mean and unity standard deviation, which ensures that exceptionally large-valued descriptors do not bias the network. The Nguyen-Widrow algorithm is used to initialize weights and biases, which are updated using a Levenberg-Marquardt optimization technique. The transfer functions and the performance function of the network are tailored to find the best possible network.

The final network is evaluated further using correlation coefficient (R^2), sum-of-squared errors, average-absolute deviations, weighted-average-absolute deviations, root-

mean-squared error (RMSE), number of wrong classifications and correlation coefficient, when applicable.

4.5. RESULTS AND DISCUSSION

Schultz *et al.* [30] suggested that an ideal QSPR model should: (1) consider an adequate number of molecules for sufficient statistical representation, (2) have a wide range of quantified toxic potency, and (3) yield to mechanistic interpretation. In this work, skin irritation data for 189 molecules of diverse classes with a wide range of potency distribution have been considered. Further, to interpret mechanistically the final set of descriptors, the correlation of the descriptors with the molecular size, reactivity and penetration ability (factors that are believed to highly influence the skin irritation potential) was studied. Four case studies were conducted to build an effective skin irritation QSPR model. Each case study, analyzed in sequence, provided valuable guidance to the development of the final QSPR model.

Table 4.2 presents a brief summary of the results obtained for the Cases 1-4. Initially, models were developed using only the literature descriptors (Case 1). Although the performance of these models was comparable to the models cited in the literature, the error in these models exceeded the desired error level. Therefore, around 1200 structural descriptors and 154 functional descriptors were analyzed to obtain a QSPR model with the best predictive ability. Non-linear transformations of the descriptors were evaluated to identify any non-linear relationships during descriptor reduction. Although the inclusion of these descriptors improved the model predictions, additional improvements were desired, and hence, non-linear neural network based models were developed.

Our first non-linear QSPR model used only descriptors identified in the literature. Figure 4.2 illustrates the property predictions obtained in Case 1. The symbols in the plots denote the property estimate obtained from the model while the solid line represents perfect predictions (45° line). R^2 and RMSE values of 0.41 and 1.66 were obtained, respectively. The results show that using only the literature descriptors for modeling skin irritation is inadequate, and additional structural descriptors are required to account for skin irritation effects of diverse chemical classes.

The functional group contributions to the property predictions were studied using 154 functional group descriptors evaluated using Dragon (Case 2). The significant functional groups identified in this case study are listed in Table 4.3, and Figure 4.3 presents the property predictions. More details on the functional groups used for model development are illustrated in the Dragon manual. To further improve the property predictions, 1200 structural descriptors were generated using CODESSA and a non-linear prediction model was developed (Case 3). The descriptors used and the model predictions are illustrated in Table 4.3 and Figure 4.4, respectively. R^2 and RMSE values of 0.69 and 1.19 were obtained, respectively. This model used only structural descriptors; hence, application of this model to any new structure requires no *a priori* estimation of physiochemical properties of chemicals.

A final model, combining the literature, functional and structural descriptors was developed to obtain the best predictive ability (Case 4). Three molecules that were assigned a PII score of 8 were found to be outliers and subsequently removed. A neural network model with 13-3-8-1 architecture utilizing 13 descriptors was found to give the best predictions for the remaining 186 chemical compounds. Table 4.3 presents the final

set of descriptors used for model development. The final model for skin irritation predicted with an R^2 of 0.78. Further, the RMSE value obtained was 1.05, showing the efficacy of using a combined set of descriptors. Figure 4.5 shows the property estimates obtained from the final QSPR model. An ideal QSPR model should yield to mechanistic interpretation of the input parameters used for model predictions. For this, the descriptors obtained were further investigated to identify any correlation with molecular size, reactivity and skin permeation ability using CODESSA. The values of log K_{ow} and log K_p for the chemical compounds were obtained using QSPR prediction models developed by our group. Figure 4.6 illustrates that the descriptors in the final model are indeed highly correlated with these properties.

To further validate the predictive capability of the QSPR model, an external data set from the Tornier *et al.* [31] that was not included in the model development was used. The external dataset contained 22 chemical compounds. A list of the molecules used for the external validation is provided in Table 4.4. This dataset differed from the one that was used for model development in that it primarily contained skin irritation classification values from European Union (EU), human patch, SkinEthic direct application and SkinEthic patch test [31]. Figure 4.7 shows a comparison of the classification system used in the EU and human patch test. More details on the classification used are provided elsewhere [32]. The predictions obtained from the model were compared to the experimental results reported in the external dataset as illustrated in the Table 4.4. Classifying the molecules with PII in the range 0-2 as non-irritants (NI) and 2-8 as irritants (I), the model predictions were found to be in good agreement with the experimental data for a large part of the dataset. Of the 22 molecules validated, lactic

acid is the only molecule that was not consistent with any of the test procedures. Although the secondary validation process provides valuable insights regarding the predictive ability of the developed QSPR model, only a guarded judgment can be made based on this study since the test data reported in the external dataset mostly come from different experimental technique.

Table 4.5 illustrates a comparison of the current QSPR model results with literature skin irritation prediction models. The current model for skin irritation improves on other similar literature models in several aspects, including (a) the use of a large dataset consisting of a wide array of functional groups of significance to most chemical processes; (b) use of descriptors from multiple QSPR software to assure model superiority and stability; (c) use of non-linear transformations during descriptor reduction to obtain the most suitable set of descriptors; (d) mechanistic interpretation of the final non-linear QSPR models; and (e) extensive validation of models to assure robustness and predictive ability. As evidenced by the results, the QSPR model is capable of predicting the skin irritation potential of a diverse set of molecules with varying structural complexities.

4.6. CONCLUSIONS

Following are the conclusions drawn based on this study:

1. The QSPR model was able to predict the skin irritation potential of diverse chemical compounds successfully with an R^2 of 0.78.
2. The results of this study indicate that using a combination of literature, functional group, and structural descriptors are effective in QSPR modeling of skin irritation.

3. The final set of descriptors obtained showed good correlation with the molecular size, reactivity and skin penetration characteristics of chemical compounds, thus accounting for mechanistic interpretation.
4. The final QSPR model was effective in estimating the skin irritation potential of diverse chemical compounds for an external dataset containing 22 compounds.
5. Although the database employed contained experimental data for diverse compounds, expansion of the database as additional experimental data become available would facilitate the further development of the current model.

REFERENCES

[1] OECD, *Guideline for Testing of Chemicals No. 404: Acute dermal irritation/corrosion, Paris, France,* 1992.

[2] J. H. Draize, G. Woodard, and H. O. Calvery, "Methods for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes," *Journal of Pharmacology and Experimental Therapeutics,* vol. 82, pp. 377-390, 1944.

[3] P. A. Botham, L. K. Earl, J. H. Fentem, R. Roguet, and J. J. M. van de Sandt, "Alternative methods for skin irritation testing: the current status," *Alternatives to Laboratory Animals,* vol. 22, pp. 195-211, 1998.

[4] I. Gerner and E. Schlede, "Introduction of in vitro data into local irritation/corrosion testing strategies by means of SAR considerations: Assessment of chemicals," *Toxicology Letters,* vol. 127, pp. 169-175, 2002.

[5] S. S. Godavarthy, R. L. Robinson, and K. A. M. Gasem, "Improved structure–property relationship models for prediction of critical properties," *Fluid Phase Equilibria,* vol. 264, pp. 122-136, 2008.

[6] S. S. Godavarthy, "Design of improved solvents for extractive distillation," Ph. D. Dissertation, Oklahoma State University, Stillwater, OK, 2004.

[7] T. Welss, D. A. Basketter, and K. R. Schröder, "In vitro skin irritation: facts and future. State of the art review of mechanisms and models," *Toxicology In Vitro,* vol. 18, pp. 231-243, 2004.

[8] A. L. Chew and H. I. Maibach, "*In vitro* methods to predict skin irritation," *Irritant Dermatitis,* 2006.

[9] A. G. Saliner, G. Patlewicz, and A. P. Worth, "A review of (Q) SAR models for skin and eye irritation and corrosion," *QSAR & Combinatorial Science,* vol. 27, pp. 49-59, 2008.

[10] R. D. Combes and P. Judson, "The use of artificial intelligence systems for predicting toxicity," *Pesticide Science,* vol. 45, pp. 179-194, 1995.

[11] K. Enslein, H. H. Borgstedt, B. W. Blake, and J. B. Hart, "Prediction of rabbit skin irritation severity by structure-activity relationships," *In Vitro Toxicology,* vol. 1, pp. 129-147, 1987.

[12] J. S. Smith, O. T. Macina, N. B. Sussman, M. I. Luster, and M. H. Karol, "A robust structure-activity relationship (SAR) model for esters that cause skin irritation in humans," *Toxicological Sciences,* vol. 55, pp. 215-222, 2000.

[13] J. S. Smith, O. T. Macina, N. B. Sussman, M. H. Karol, and H. I. Maibach, "Experimental validation of a structure–activity relationship model of skin irritation by esters," *Quantitative Structure-Activity Relationship,* vol. 19, pp. 467-474, 2000.

[14] M. Hayashi, Y. Nakamura, K. Higashi, H. Kato, F. Kishida, and H. Kaneko, "A quantitative structure–activity relationship study of the skin irritation potential of phenols," *Toxicology In Vitro,* vol. 13, pp. 915-922, 1999.

[15] B. Berner, A. Nangia, and H. I. Maibach, "Influence of pKa on skin irritancy potential of chemicals," *In Vitro Skin Toxicology, Irritation, Phototoxicity, Sensitization,* pp. 29–36, 1994.

[16] K. Kodithala, A. J. Hopfinger, E. D. Thompson, and M. K. Robinson, "Prediction of Skin Irritation from Organic Chemicals Using Membrane-Interaction QSAR Analysis," *Toxicological Sciences,* vol. 66, pp. 336-346, 2002.

[17] J. C. Dearden, M. D. Barratt, R. Benigni, D. W. Bristol, R. D. Combes, M. T. D. Cronin, P. N. Judson, M. P. Payne, A. M. Richard, and M. Tichy, "The development and validation of expert systems for predicting toxicity," *Alternatives to Laboratory Animals,* vol. 25, pp. 223-252, 1997.

[18] D. M. Bagley, J. R. Gardner, G. Holland, R. W. Lewis, J. F. Regnier, D. A. Stringer, and A. P. Walker, "Skin irritation: Reference chemicals data bank," *Toxicology In Vitro,* vol. 10, pp. 1-6, 1996.

[19] ChemBioOffice 11.0, Cambridge Software, 2008.

[20] C. Morley, "OpenBabelGUI 2.0.2," 2006.

[21] AMPAC 6.0., Semichem Inc., 1997.

[22] CODESSA 2.7.8., Semichem Inc., 2007.

[23] Talete srl, DRAGON 5.4 for Windows (Software for molecular descriptor calculations), 2006.

[24] A. Jegadeesan, "Structure-based generalized models for selected pure-fluid saturation properties," Master Thesis, Oklahoma State University, Stillwater, OK, 2006.

[25] S. S. Godavarthy, R. L. Robinson, and K. A. M. Gasem, "Improved structure–property relationship models for prediction of critical properties," *Fluid Phase Equilibria,* vol. 264, pp. 122-136, 2007.

[26] M. D. Barratt, "Prediction of toxicity from chemical structure," *Cell Biology and Toxicology,* vol. 16, pp. 1-13, 2000.

[27] J. Aihara, "Correlation found between the HOMO–LUMO energy separation and the chemical reactivity at the most reactive site for isolated-pentagon isomers of fullerenes," *Physical Chemistry Chemical Physics,* vol. 2, pp. 3121-3125, 2000.

[28] H. Demuth, M. Beale, and M. Hagan, "Neural Network Toolbox™ 6," MathWorks Inc., 2008.

[29] M. S. Iyer and R. R. Rhinehart, "A method to determine the required number of neural-networktraining repetitions," *IEEE Transactions on Neural Networks,* vol. 10, pp. 427-432, 1999.

[30] T. W. Schultz, M. T. D. Cronin, J. D. Walker, and A. O. Aptula, "Quantitative structure–activity relationships (QSARs) in toxicology: a historical perspective," *Journal of Molecular Structure: THEOCHEM,* vol. 622, pp. 1-22, 2003.

[31] C. Tornier, M. Rosdy, and H. I. Maibach, "In vitro skin irritation testing on reconstituted human epidermis: Reproducibility for 50 chemicals tested with two protocols," *Toxicology In Vitro,* vol. 20, pp. 401-416, 2006.

[32] D. A. Basketter, M. Chamberlain, H. A. Griffiths, M. Rowson, E. Whittle, and M. York, "The classification of skin irritants by human patch test," *Food and Chemical Toxicology,* vol. 35, pp. 845-852, 1997.

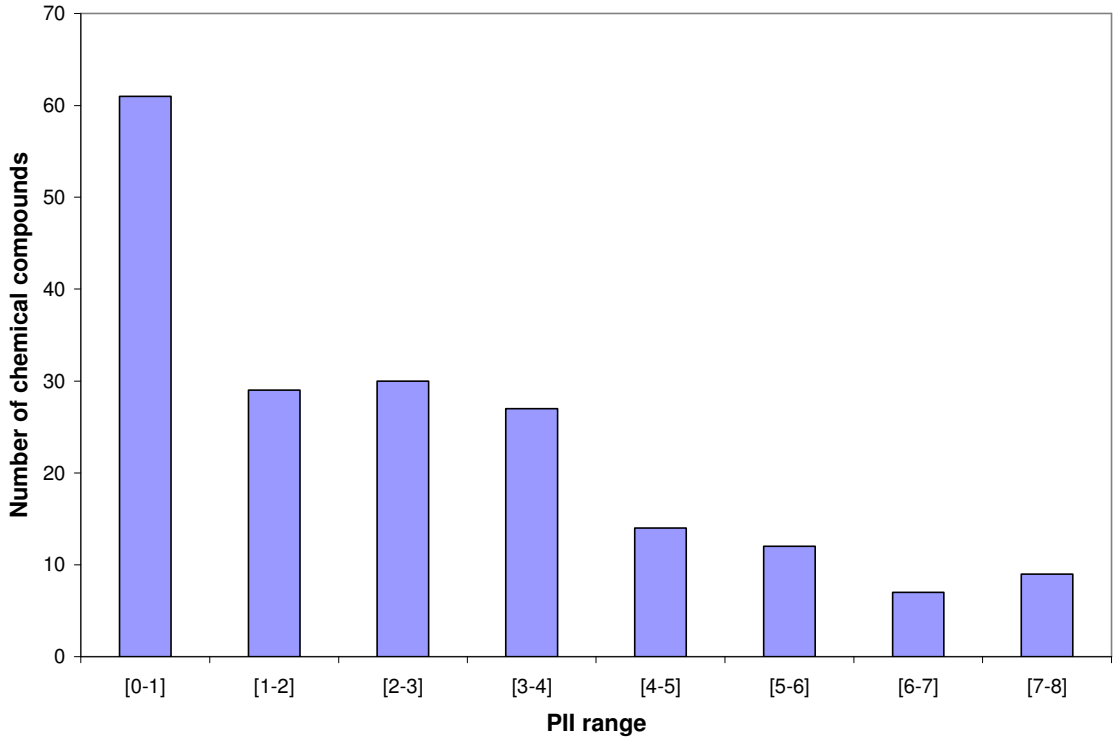


Figure 4.1. PII distribution of the data used for QSPR modeling

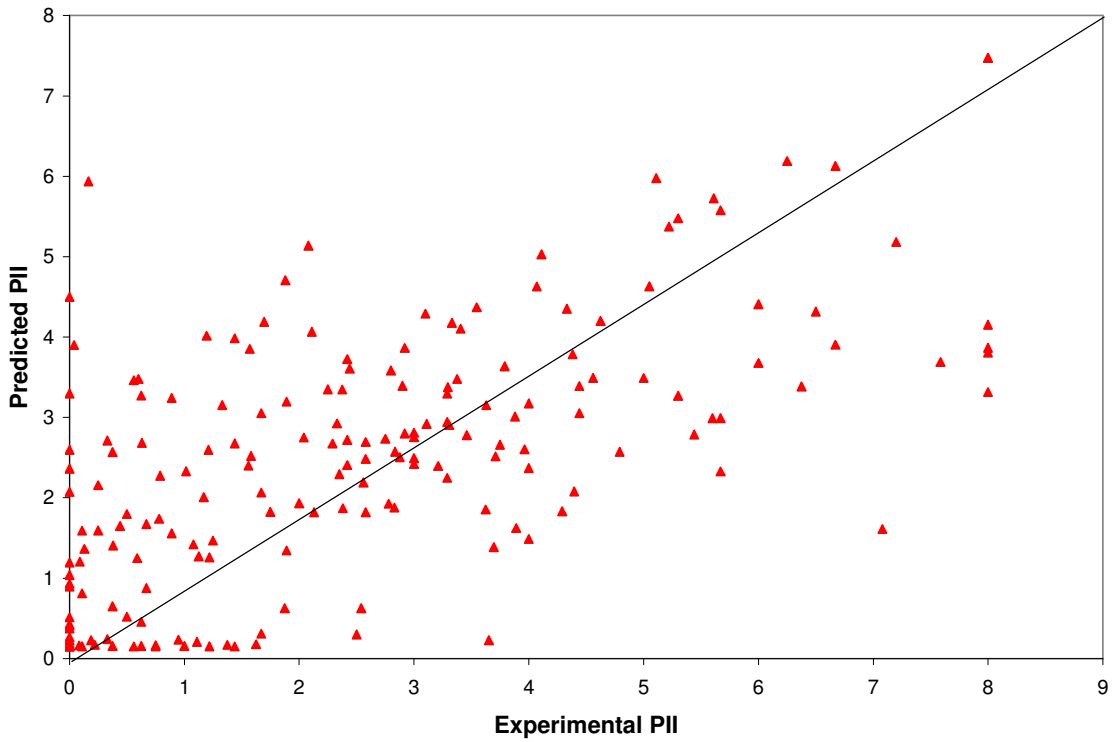


Figure 4.2. QSPR model predictions using literature descriptors (Case 1)

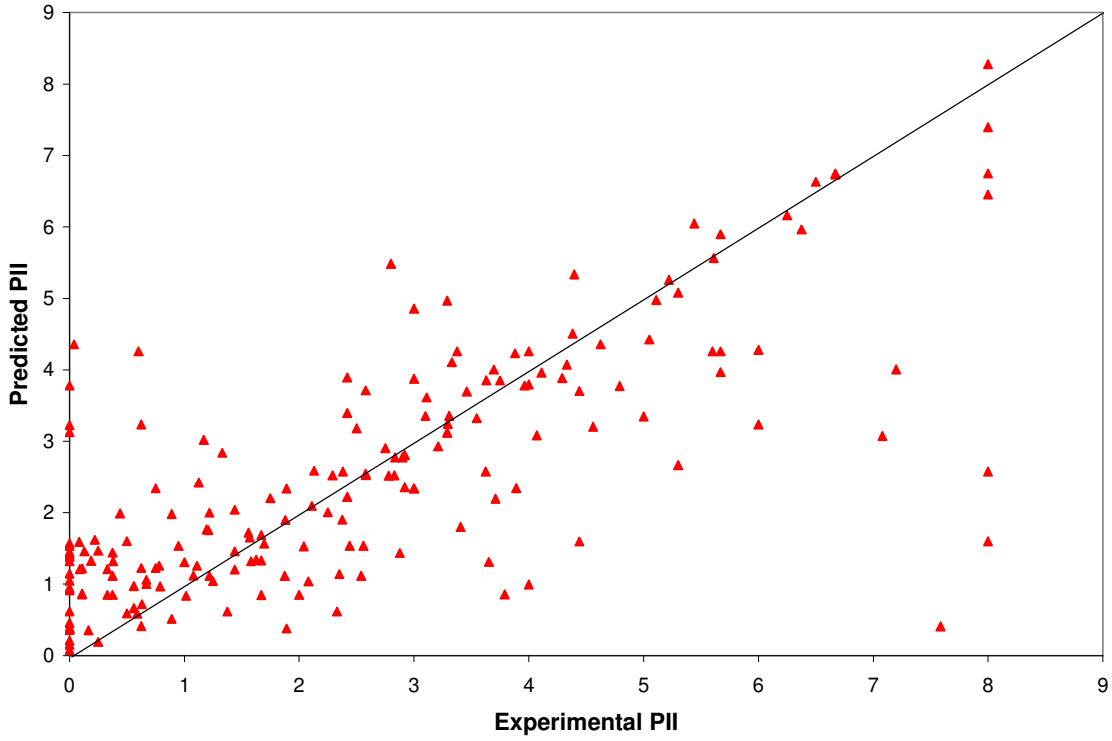


Figure 4.3. QSPR model predictions using functional group descriptors (Case 2)

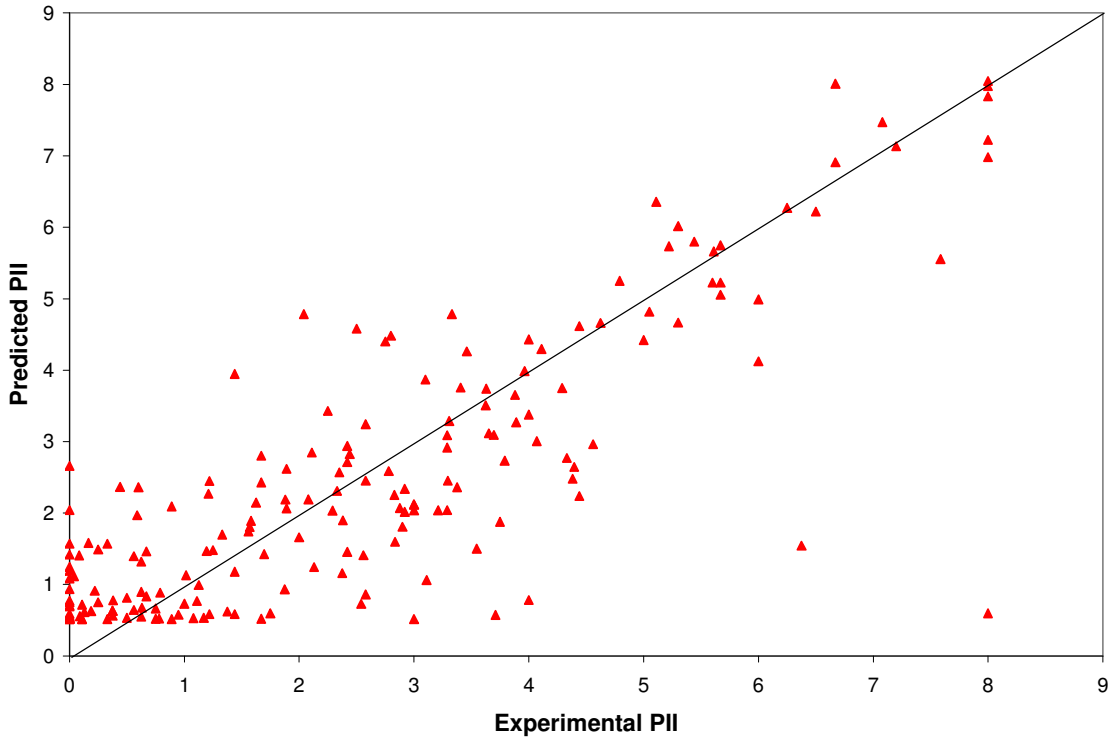


Figure 4.4. QSPR model predictions using structural descriptors from CODESSA (Case 3)

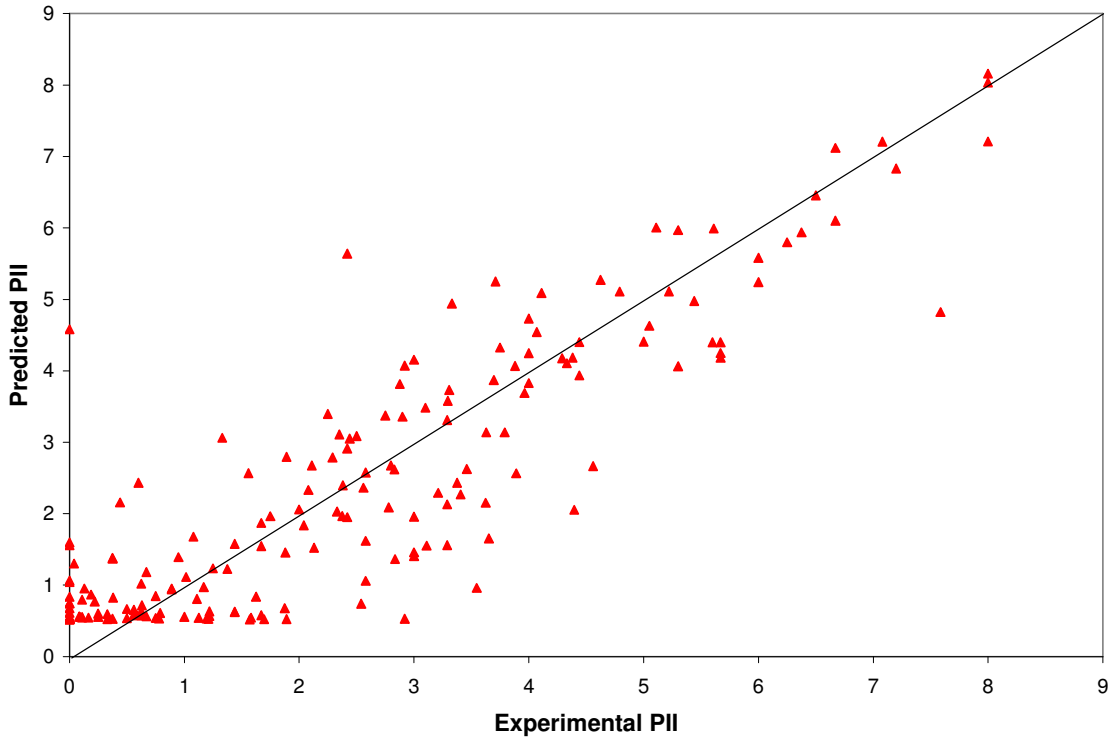


Figure 4.5. QSPR model predictions using literature, functional group and structural descriptors (Case 4)

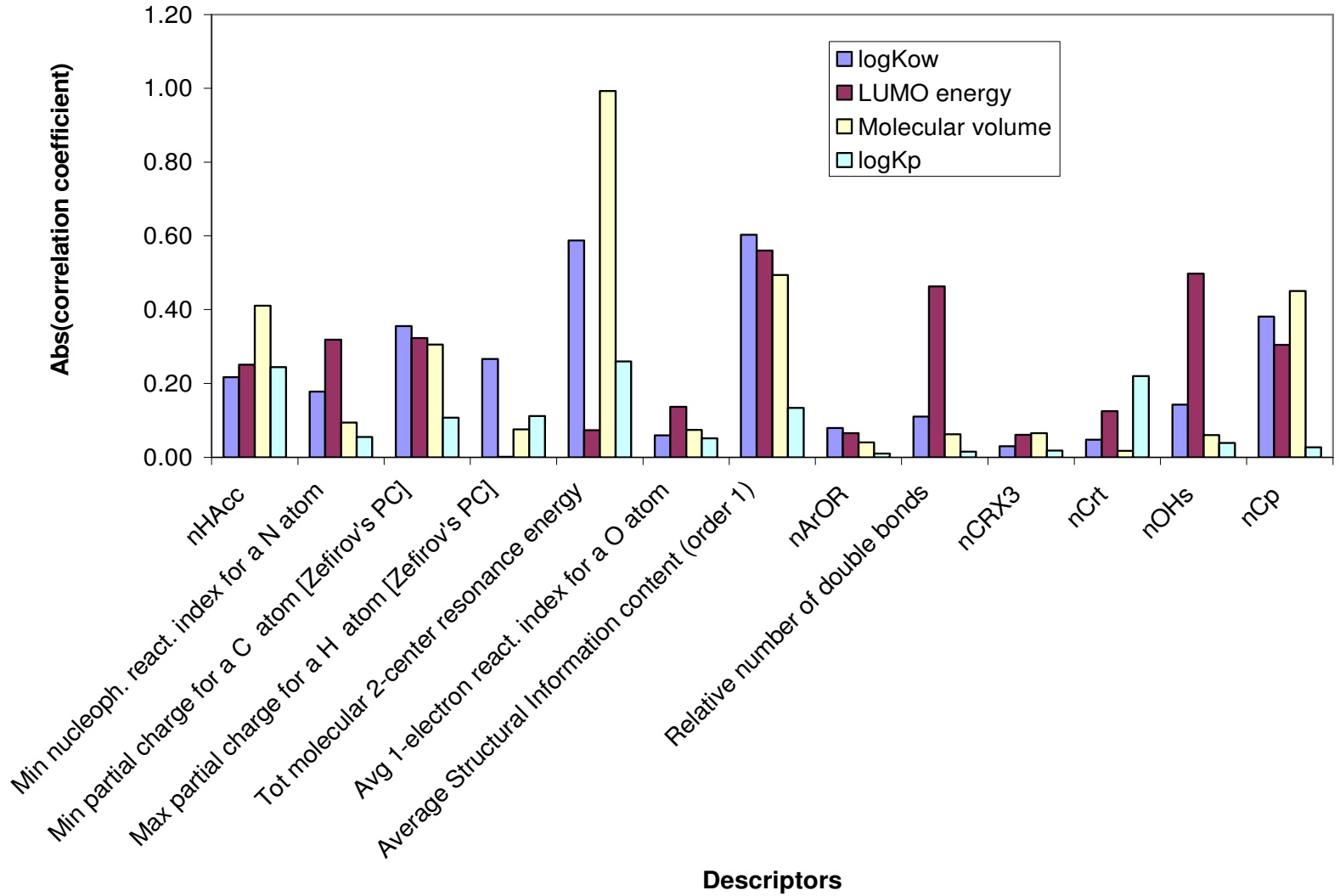


Figure 4.6. Correlation of the final set of descriptors with the molecular size, reactivity and skin penetration ability

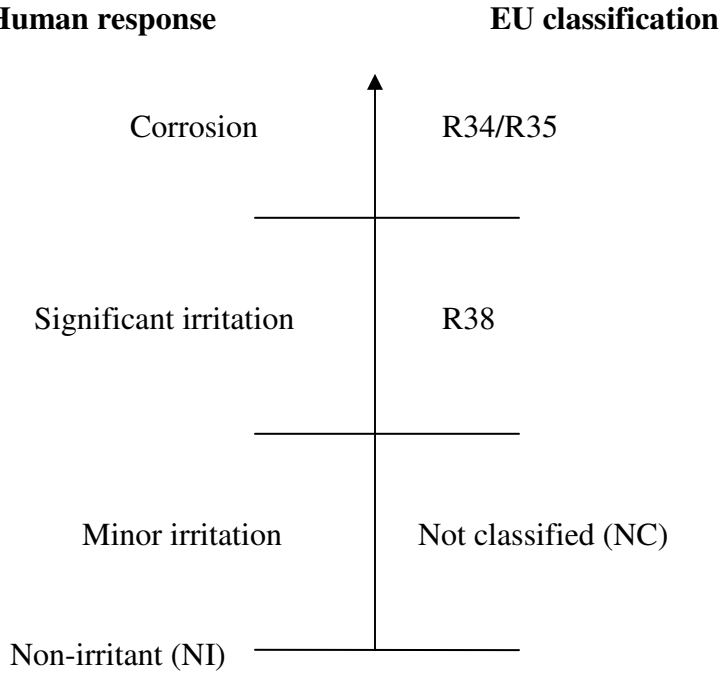


Figure 4.7. Classification of skin irritation hazard [32]

Table 4.1. Grading scale for skin reactions [1]

Erythema and eschar formation grade	Scale
No erythema	0
Very slight erythema (barely perceptible)	1
Well defined erythema	2
Moderate to severe erythema	3
Severe erythema (beet redness) to slight eschar formation (injuries in depth)	4
Edema formation grade	
No edema	0
Very slight edema (barely perceptible)	1
Slight edema (edges of area well defined by definite raising)	2
Moderate edema (raised approximately 1 mm)	3
Severe edema (raised more than 1 mm and extending beyond the area of exposure)	4

Table 4.2. Summary of results obtained for modeling of skin irritation

Case	**Descriptors used**	**N***	**Nd***	**Neural network architecture**	**RMSE**	**R^2**
1	Literature	189	4	4-7-1-1	1.7	0.41
2	Functional group	189	16	16-5-1-1	1.3	0.62
3	Structural	189	19	19-4-1-1	1.2	0.69
4	Literature, functional, and structural	186	13	13-3-8-1	1.1	0.78

* N = Number of molecules
Nd = Number of descriptors

Table 4.3. Descriptors used for model development

Case 1	Case 2	Case 3	Case 4
HOMO energy	nHAcc	FNSA-2 Fractional PNSA (PNSA-2/TMSA) [Quantum-Chemical PC]	nHAcc
LUMO energy	nRNH2	Avg 1-electron react. index for a O atom	Min nucleoph. react. index for a N atom
log Kow	nRNR2	Tot molecular 2-center resonance energy	Min partial charge for a C atom [Zefirov's PC]
Molecular weight	nCH2RX	Avg nucleoph. react. index for a N atom	Max partial charge for a H atom [Zefirov's PC]
	nArX	Moment of inertia C	Tot molecular 2-center resonance energy
	nRCN	RNCS Relative negative charged SA (SAMNEG*RNCG) [Zefirov's PC]	Avg 1-electron react. index for a O atom
	nCrt	Min electroph. react. index for a Cl atom	Average Structural Information content (order 1)
	nR=CX2	DPSA-1 Difference in CPSAs (PPSA1-PNSA1) [Quantum-Chemical PC]	nArOR
	nCconj	Kier shape index (order 3)	Relative number of double bonds
	nCrq	Tot dipole of the molecule	nCRX3
	nCRX3	Kier shape index (order 2)	nCrt
	nROH	PNSA-2 Total charge weighted PNSA [Quantum-Chemical PC]	nOHs
	nCp	Min 1-electron react. index for a S atom	nCp
	nCs	Min total interaction for a C-C bond	
	nArOH	ALFA polarizability (DIP)	
		Max 1-electron react. index for a N atom	
		Min 1-electron react. index for a O atom	
		Max net atomic charge for a F atom	
		DPSA-1 Difference in CPSAs (PPSA1-PNSA1) [Zefirov's PC]	

Table 4.4. External validation of the model and corresponding experimental data

Chemical compound	EU class	Human patch class	SkinEthic direct application class	SkinEthic patch class	Predicted PII	Predicted class
2-Propanol	NC	NC	I	NI	0.6	NI
Isopropyl palmitate	NC	NC	NI	NI	3.9	I
Dimethyl sulphoxide	NC	R38	I	NI/I	5.1	I
Lactic acid	NC	R38	I	I	0.5	NI
Triethanolamine	NC	NC	I	NI	1.1	NI
Dodecanol	NC	NC	NI	NI	3.3	I
Tween 80	NC	NC	NI	NI	5.9	I
Propylene glycol	NC	NC	NI	NI	0.5	NI
Ethanol	NC	NC	I	NI	2.2	I
Octanoic acid	R34	R38	I	I	3.6	I
Heptanoic acid	R34	R38	I	I	4.8	I
1-Decanol	R38	NC	I	I	3.9	I
Decanoic acid	R38	R38	I	I	2.6	I
Dodecanoic acid	R38	NC	I	I	2.4	I
N,N-Dimethyl-N-dodecylaminobetaine	R38	R38	I	NI/I	5.8	I
Acetic acid	R38	NC	I	I	2.8	I
Hydrochloric acid	R38	NC	I	NI	3.2	I
Benzalkonium chloride	R38	R38	I	I	2.9	I
Octanol	R38	NC	I	I	4.9	I
Geraniol	R38	NC	I	I	2.7	I
Linalyl acetate	R38	NC	I	I	4.7	I
Hexanol	R38	NC	I	I	6.8	I

Table 4.5. Comparison of current QSPR model with similar literature models

Author [reference]	No. of compounds	Chemical class	No. of inputs to the model	Model type	R^2
Hayashi *et al.* [14]	13	Phenols	2	Linear	0.67
	11	Phenols	2	Linear	0.52
Kodithala *et al.* [16]	20	Hydroxy alcohols	1	Linear	0.54
	20	Hydroxy alcohols	2	Linear	0.62
	20	Hydroxy alcohols	3	Linear	0.76
	20	Hydroxy alcohols	4	Linear	0.86
	13	Aliphatic alcohols	3	Linear	0.95
	9	Phenols	4	Linear	0.94
This work	186	Diverse classes	13	Non-linear	0.78

VITA

Sharath Golla

Candidate for the Degree of

Master of Science

Thesis: VIRTUAL DESIGN OF CHEMICAL PENETRATION ENHANCERS

Major Field: Chemical Engineering

Biographical:

Personal Data:
Born in Chittoor, India, October 20, 1983, the son of Damodaram Golla and Manjula Golla.

Education:
Received the Bachelor of Engineering (Honors) degree in
Chemical Engineering from Birla Institute of Technology and Science (BITS), Pilani, India, in May 2005. Completed the requirements for the Master of Science degree in Chemical Engineering from Oklahoma State University (OSU), Stillwater, in July 2008.

Experience:
Teaching assistant, Oklahoma State University, Stillwater (Aug 2006-Dec 2006)
Research assistant, Oklahoma State University, Stillwater (Aug 2006-Jul 2008)

Name: Sharath Golla **Date of Degree:** July, 2008

Institution: Oklahoma State University **Location:** Stillwater, Oklahoma

Title of Study: VIRTUAL DESIGN OF CHEMICAL PENETRATION ENHANCERS

Pages in Study: 111 Candidate for the Degree of Master of Science

Major Field: Chemical Engineering

Scope and Method of Study:

This study focused on identifying new potential chemical penetration enhancers (CPEs) for transdermal drug delivery. A computer-aided molecular design (CAMD) algorithm was developed by integrating a new genetic algorithm and non-linear QSPR models to develop a reliable virtual screening algorithm for generation of potential CPEs. Structure-based predictive models for prediction of skin sensitization and skin irritation were developed using reliable experimental data for a wide range of molecular species to estimate the toxic potential of the generated chemical compounds. Non-linear neural network algorithms with superior capabilities were used for model development.

Findings and Conclusions:

The innovative integration of non-linear, QSPR modeling and robust GAs removed some of the existing barriers to the use of computational chemistry in CPE design, and yielded novel chemical structures that are capable of enhancing permeation of insulin through skin. The structure-based models were capable of predicting the skin sensitization and skin irritation potential of a chemical compound with 90% accuracy and an R^2 of 0.78, respectively. This scientific knowledge gained in CAMD of CPEs will be significant to the both drug development and improved therapeutic agent delivery industries.

ADVISER'S APPROVAL: Dr. KHALED A. M. GASEM