

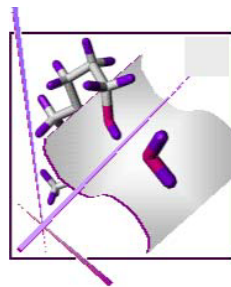
ADME-Tox Predictors Based on Structure-Information Representation

Lowell H. Hall

Department of Chemistry
Eastern Nazarene College
Quincy, MA 02170
hall@enc.edu

L. Mark Hall

Hall Associates Consulting
2 Davis Street
Quincy, MA 02170
hallm@comcast.net



UK QSAR
& Chemoinformatics
Spring Meeting 2005

Our Challenge

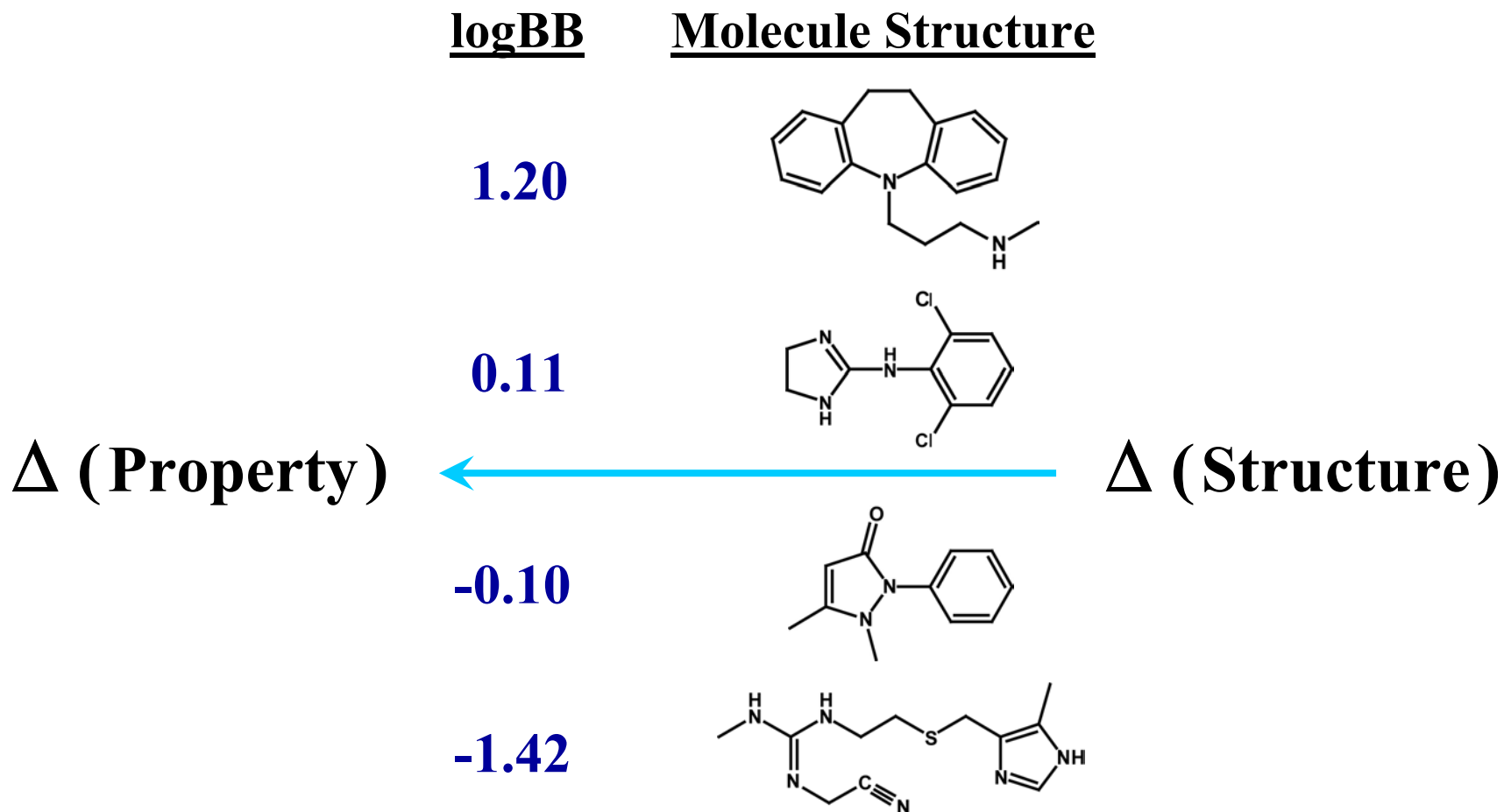
Facing serious issues in drug design . . .

Develop ways to anticipate the quantitative behavior of molecules not yet made . .

With sufficient reliability of prediction . .

that development can proceed based on estimates obtained from the process.

The QSAR Problem



$$\text{Property} = f(\text{Structure})$$



Structure-Information Representation

- Develop **molecule representation** so that **structure information** can be encoded, as necessary for non-covalent interactions

Comment: Mechanism not assumed nor biological/chemical/physical process simulated;
Does not require explicit 3-D geometry information

- Statistical methods yield QSAR model that **captures parallel** between **variation in structure features** and **corresponding variation in property values**

What we know directly



Comment: Structure information in model leads to direct statements about significant structure feature as aid to design



Types of Topological Structure Descriptors

For Non-Covalent Intermolecular Interactions

Electron Accessibility (each atom)

Atom Level E-State Indices, Hydrogen E-State Indices

Atom Types E-State Indices

Group Types E-State Indices

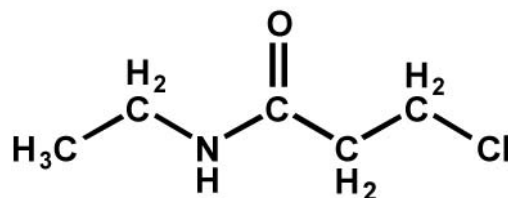
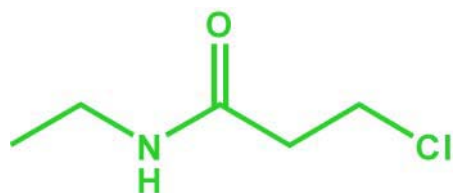
Skeletal Variation / Ramification (whole molecule)

Molecular Connectivity Chi Indices

Kappa Shape Indices



Basic Structure Information



Structure-Information
Representation
based on
Hydrogen-Implied Graph

Visualized Connection Table

provides

- Atom identity (C, N, O, . . .); electron, valence electron count: Z , Z^v
- Atom connections (adjacencies, skeletal bonds)

and

- Leads to characterization of branching pattern, rings, aromaticity . . .
- Leads to characterization of valence state (number of $s + p + n$ electrons)

LH Hall and LB Kier, *J. Chem. Inf. Comput. Sci.*, 35, 1039-1045 (1995)



Basic Structure Information

Structure-Information Representation based on Hydrogen-Implied Graph

Electron Counts, each atom / hydride group: -O-, -OH, =O, -CH₂-, -Cl . .

Simple delta: $\delta = \sigma - h$ sigma electron count

Valence delta: $\delta^v = Z^v - h = \sigma + \pi + n - h$ valence electron count

LH Hall and LB Kier, J. Chem. Inf. Comput. Sci., 35, 1039-1045 (1995)



Derived Structure Information

Delta Values encode significant additional information

For each atom (-O-, -OH, =O, -CH₂-, -Cl, ..)-

Consider: $\delta^v - \delta = \pi + n$ (pi and lone pair electron count)

$$X_{MJ} = 7.58 \frac{(\delta^v - \delta)}{N^2} - 7.19 \quad r^2 = 0.97, s = 0.53, n = 20$$

Mulliken-Jaffe **valence state electronegativity**

Consider: $\delta^v + \delta$

$$\text{Volume} = 17.00 - 6.88 \frac{(\delta^v + \delta)}{N^2} \quad r^2 = 0.98, s = 0.53, n = 21$$

Volume of atom/group (cm³/mol) taken from Bondi

L.B. Kier and L.H. Hall, J. Pharm. Sci., 70, 583 (1981)



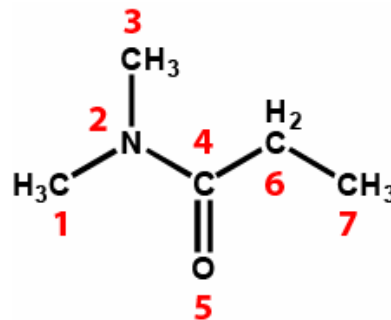
Electrotopological State

Concept: Potential for Intermolecular Interaction

Encodes **electron accessibility** at each atom :
build-up / depletion of electrons at each atom
modified
by **topological accessibility** at the atom

Calculated for:

- Each atom in molecule [Atom Level E-State]
- Atom types, Group Types
-CH₃, -NH-, -OH, =O, -Cl, etc



Based on the Intrinsic State, *I*, of each atom as a ratio:

$I \rightarrow$ *Valence State Electronegativity* / *Local Topology*

$$I \rightarrow (\delta^v - \delta) / \delta$$

Intrinsic State Table

Atom (Skeletal Hydride Group)	Intrinsic State $((2/N)^2 \delta^v + 1) / \delta$
>C<	1.250
>CH-	1.333
-CH ₂ -	1.500
>C=	1.667
-S-	1.833
-CH ₃ , =CH-, >N-	2.000
-I	2.120
≡C-, -NH-	2.500
-Br	2.750
=CH ₂ , =N-	3.000
-SH	3.222
-O-	3.500
≡CH, -NH ₂	4.000
-Cl	4.111
≡N, -OH	6.000
=O	7.000
-F	8.000



Atom Level E-State Index S_i

Electronic

Combines and character of the atom

Topological

Composed of

Intrinsic State of Atom, I_i

plus

Sum of Perturbations $\Delta I_{ij} = (I_i - I_j) / r_{ij}^2$

by all other atoms in molecule

E-State for atom i : $S_i = I_i + \sum_j \Delta I_{ij}$

Both
arise from
electron
distribution

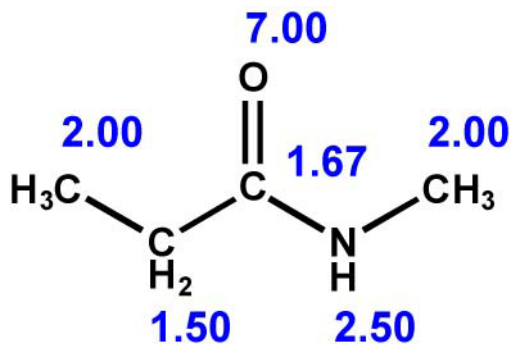
E-State index is an expression of

Electron Accessibility at that atom

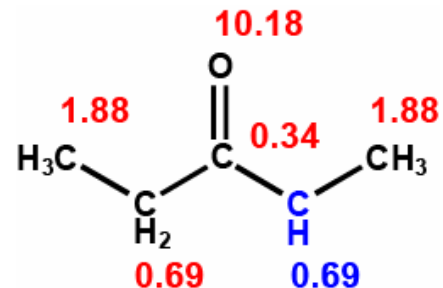
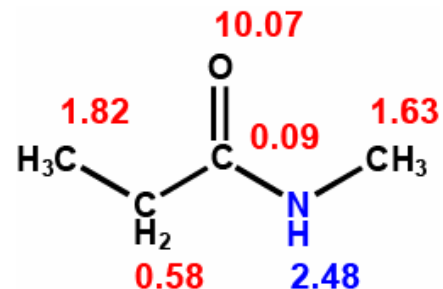
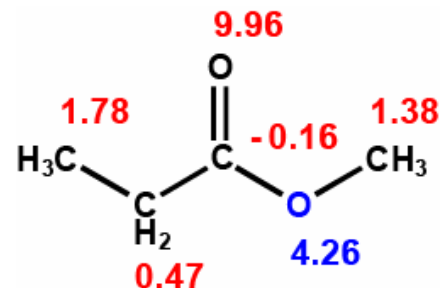
For non-covalent interactions



Example: N-Methylpropanamide



Intrinsic State Values



E-State Index Values

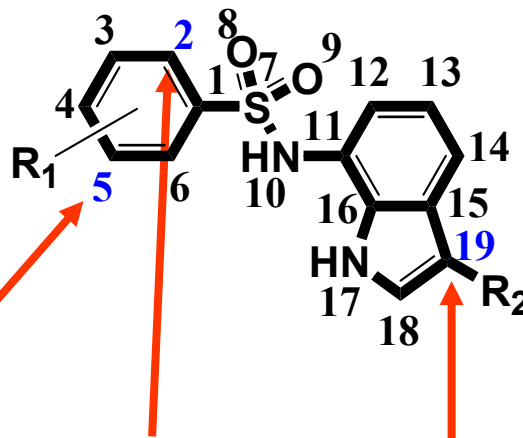


Illustration of Atom-Level E-State Application

Anti-proliferative Activity

Murine colon 38 cell line

$$r^2 = 0.73 \quad pC = -1.91 * S_5 + 3.34 * S_2 - 2.47 * S_{19} + 4.94$$



S_{19} also found important for P388 Leukemia cell line

S_2 found important for KB Nasopharynx cell line

K. Roy, D. K. Pal and C. Sengupta,
“QSAR of Antineoplastics V. Exploration of Receptor Interaction Sites of
Antitumor N-(7-indoyl)benzenesulfonamides Targeting G1 Phase Using
Electrotopological State Atom Index”,
Drug Design and Discovery, 17, 207-218 (2001)



Atom Type E-State Descriptors

Extension of Atom-Level E-State formalism

- **Atom Types/Groups within a molecule** -

for heterogeneous data sets and similarity searching

Sum of E-State Indices for all atoms of a given type

e.g. $S^T(-CH_2-)$, $S^T(-OH)$, $S^T(=O)$, $S^T(=C^H-)$, $S^T(-C^L-)$
 $S^T(HBa)$, $S^T(arom)$

Atom Type E-State Descriptors encode

- **Electron Accessibility** at atoms of same type
- **Presence/Absence** of atom type
- **Count** of atom type present in molecule



Hydrogen E-State Indices

A separate E-State Index is computed for hydrogen atoms in each hydride group: Hs_i

$-CH_3$, $=CH_2$, aromatic CH , $-NH_2$, $-OH$, etc

Hydrogen E-State describes
electron depletion in the X-H bond
or the **hydrogen accessibility**

Atom-Type Hydrogen E-State indices:

eg: $HS^T(-CH_3)$, $HS^T(-OH)$, $HS^T(-NH_2)$, $HS^T(=CH_2)$

Non-polar groups: $HS^T(\text{other})$, $HS^T(\text{arom})$

Polar groups: $HS^T(\text{HBd})$



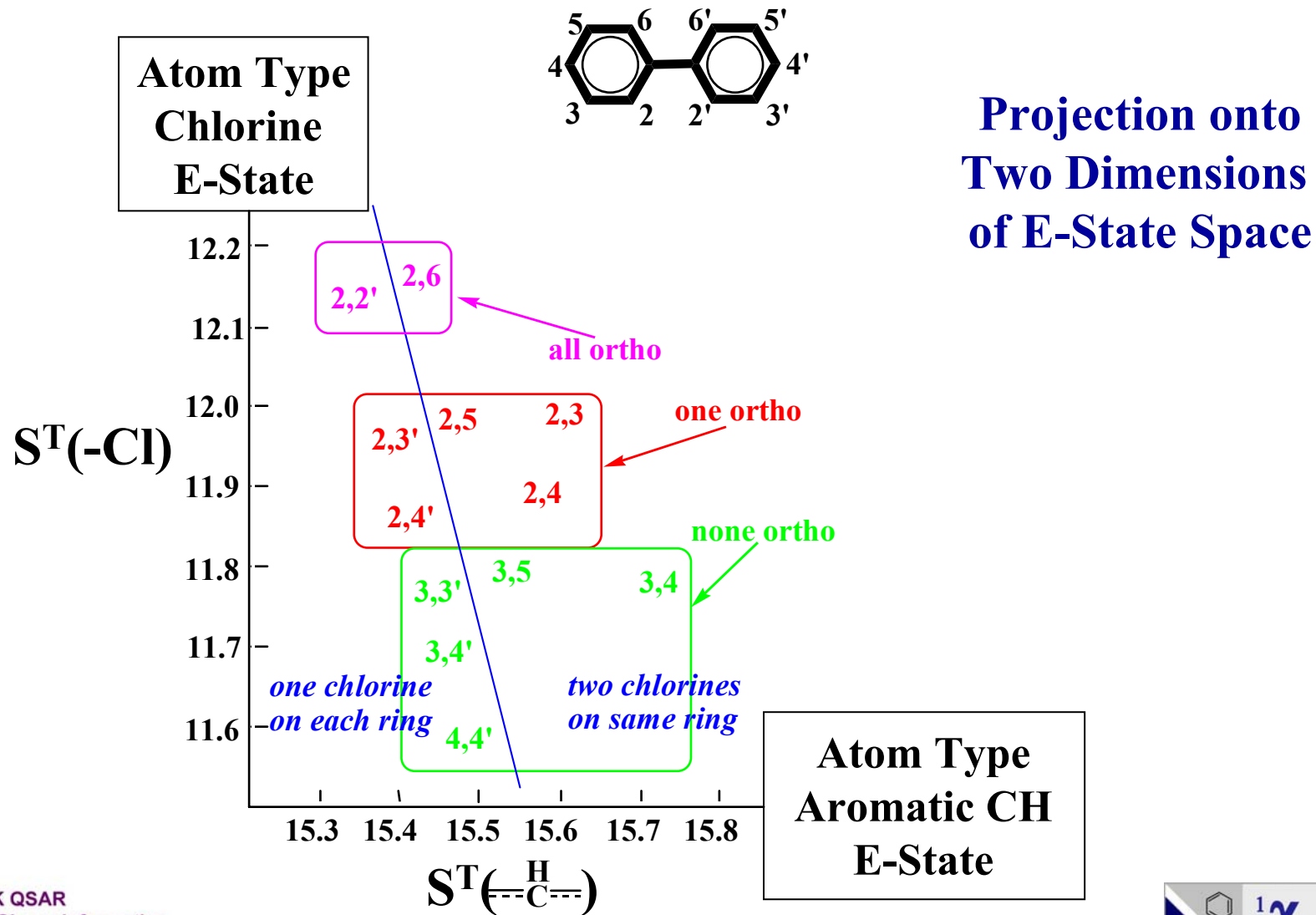
Descriptors as Basis for Structure Space

- **Atom-type E-State, Molecular Connectivity . . .**
 - create a structure space for molecules

Think of the descriptors as a basis set.
- **Creates distribution of molecular structures,**
 - clear intuitive relation to organic structure
- **Similarity searching of databases done quickly**
 - easy relation to chemical structures

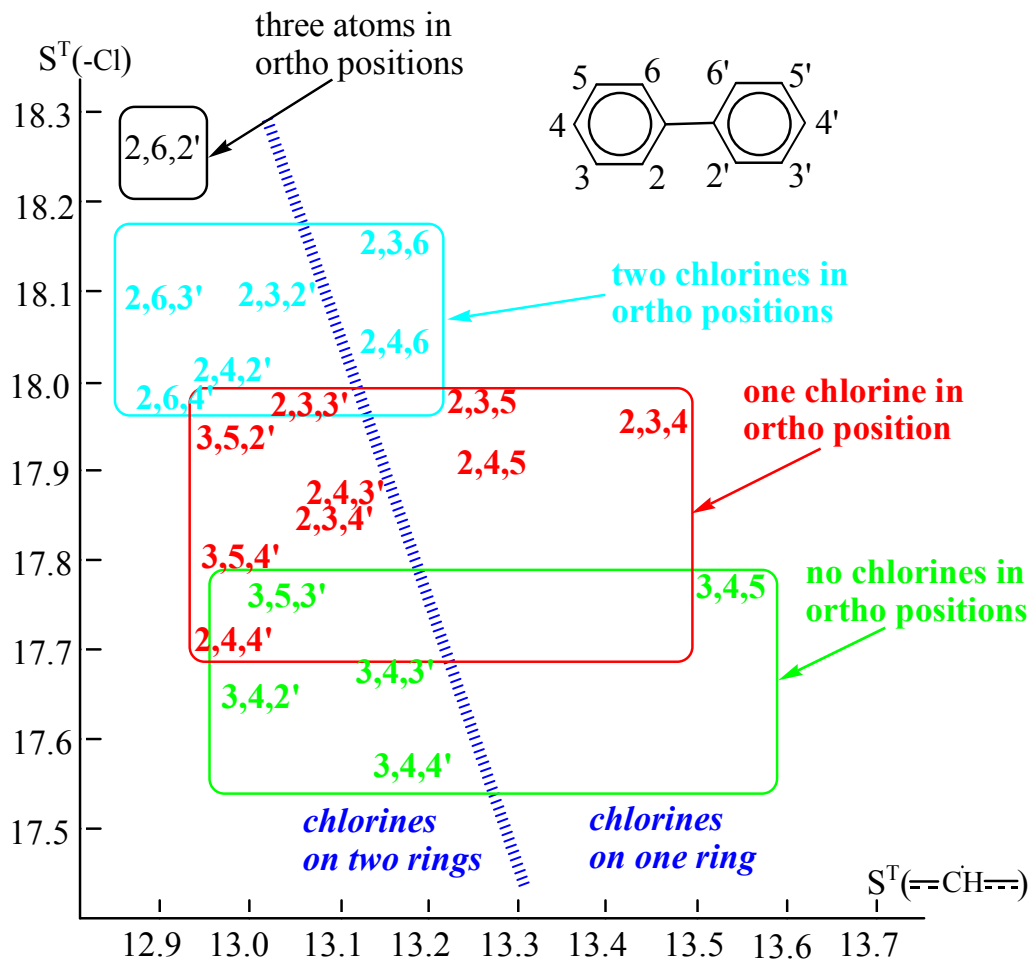


Dichloro-PCBs in E-State Space



Trichloro-PCBs in E-State space

Structurally Meaningful Distribution Projection onto Two Dimensions of E-State Space



Skeletal Variation / Ramification

Whole molecule encoding of

- **Degree of branching**
- **Types of branching**
- **Patterns of branching**

***A Chi Index* is summation of subgraph (fragments) contributions across whole molecule**

Molecular Connectivity Chi Indices:

Simple ${}^m\chi_t$ **order m, type t**

Valence ${}^m\chi_t^v$



Molecular Connectivity First Order Chi Index $^1\chi$


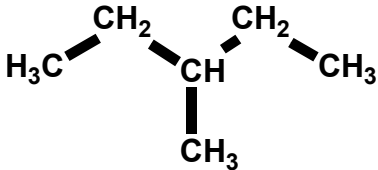
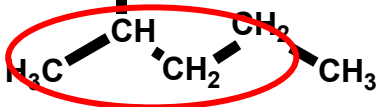
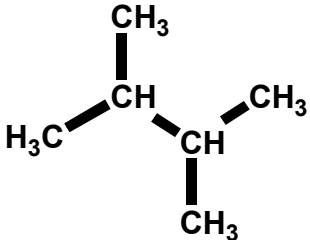
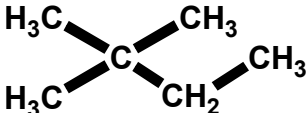
Definition:

$$^1\chi = \sum_k (\delta_i \delta_j)_k^{-1/2}$$
$$= \sum_k c_k \quad \text{sum of terms}$$

one for each skeletal bond

Descriptor variation
with
Structure variation

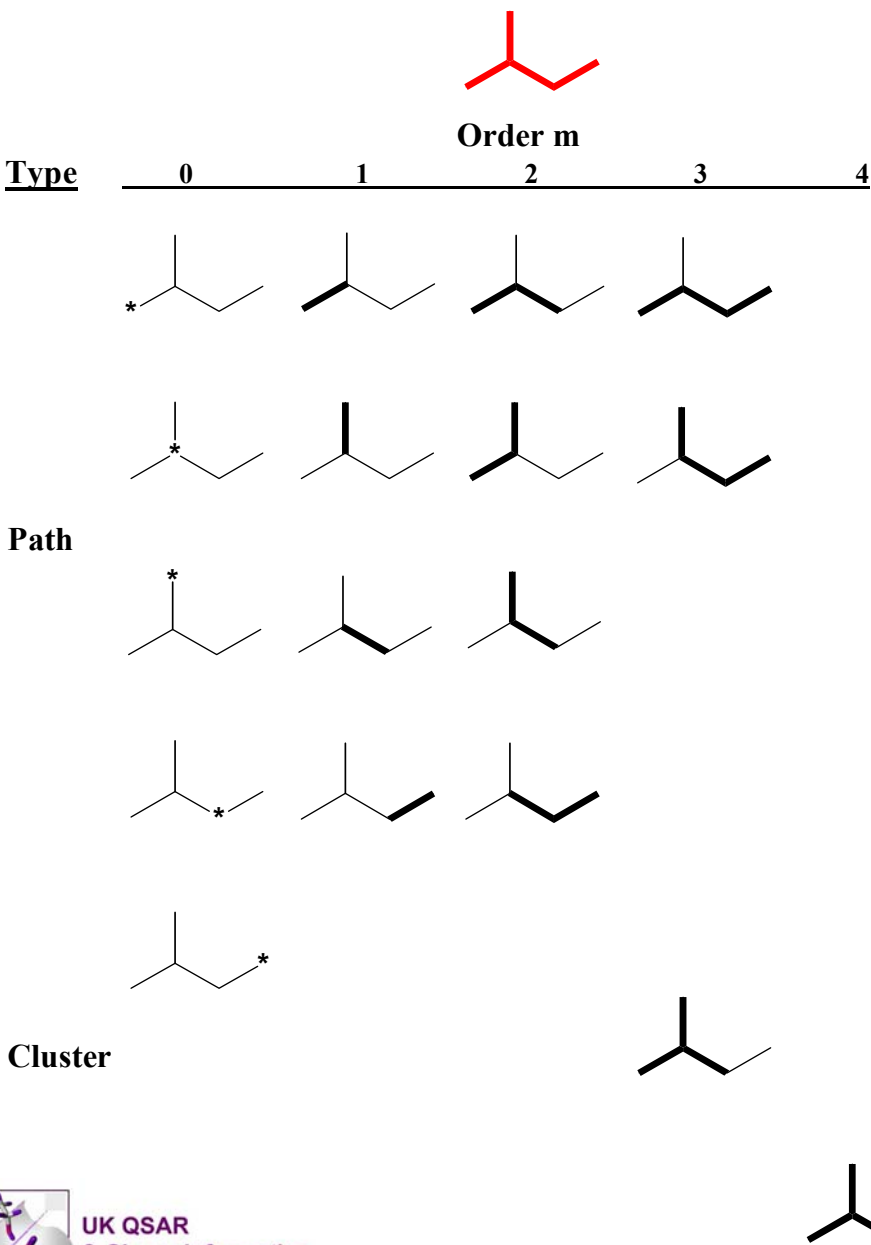
$^1\chi$ encodes amount of
skeletal branching

$^1\chi$	Structure
2.914	
2.808	
2.707	
2.643	
2.561	



Molecular Connectivity Subgraph Orders and Types

Molecular Connectivity Chi Indices



Chi index
for each
subgraph
order, type

$${}^m\chi_t$$

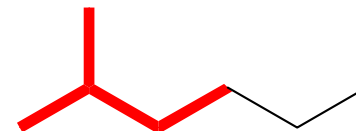
Representing Skeletal
Variation Information

Molecular Connectivity Example

path-cluster-four: ${}^4\chi_{PC}$

${}^4\chi_{PC}$ index encodes skeletal branching pattern

- number of branch points, - degree of branch points
- position of branch points, - adjacency of branch points



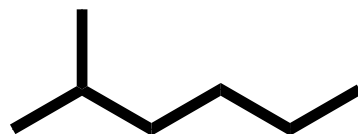
${}^4\chi_{PC}$

Skeleton

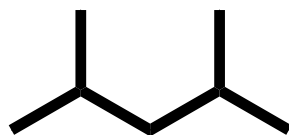
0.000



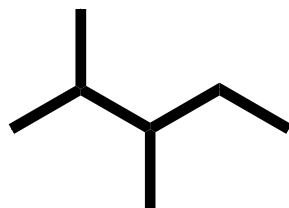
0.289



0.471



1.276



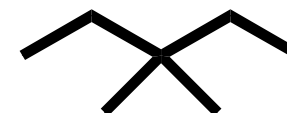
${}^4\chi_{PC}$

Skeleton

0.750



1.707



2.598

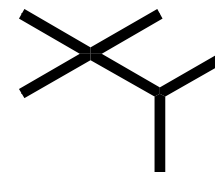
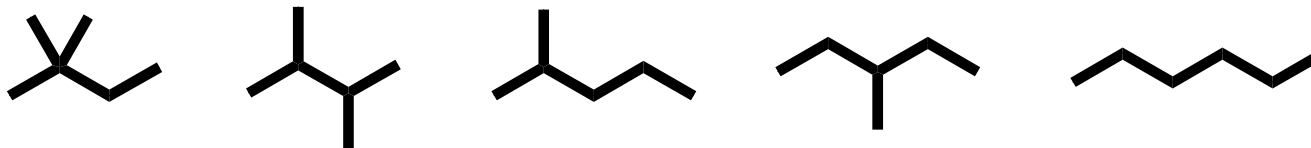


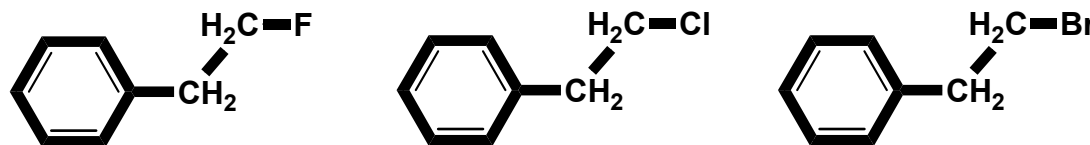
Illustration : Difference Second Order Valence Chi Index

$$d^2\chi^v = {}^2\chi^v - {}^2\chi_N^v$$



$$d^2\chi^v : 1.207 > 0.781 > 0.475 > 0.215 > 0.000$$

greater skeletal branching: larger value of $d^2\chi^v$

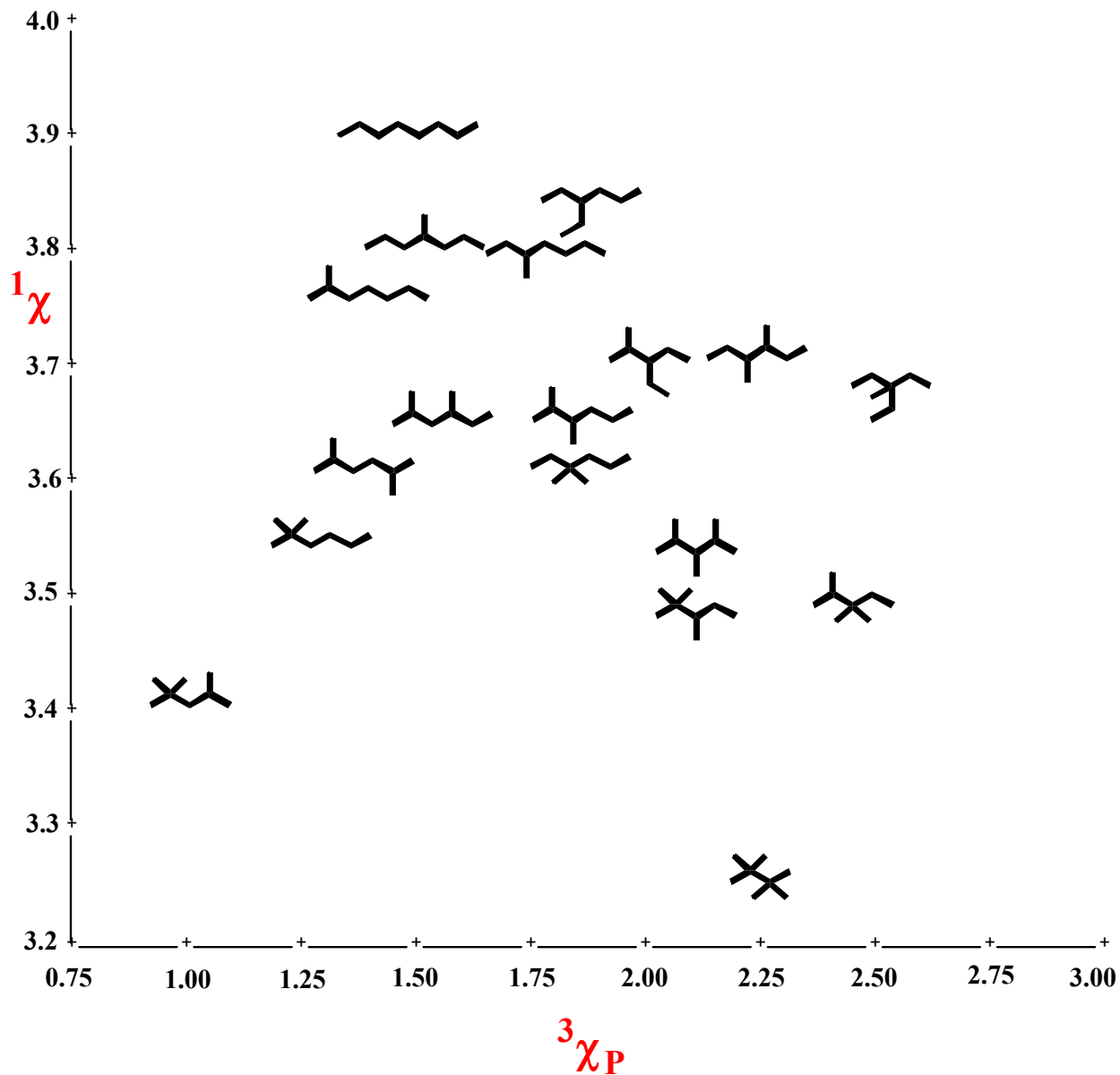


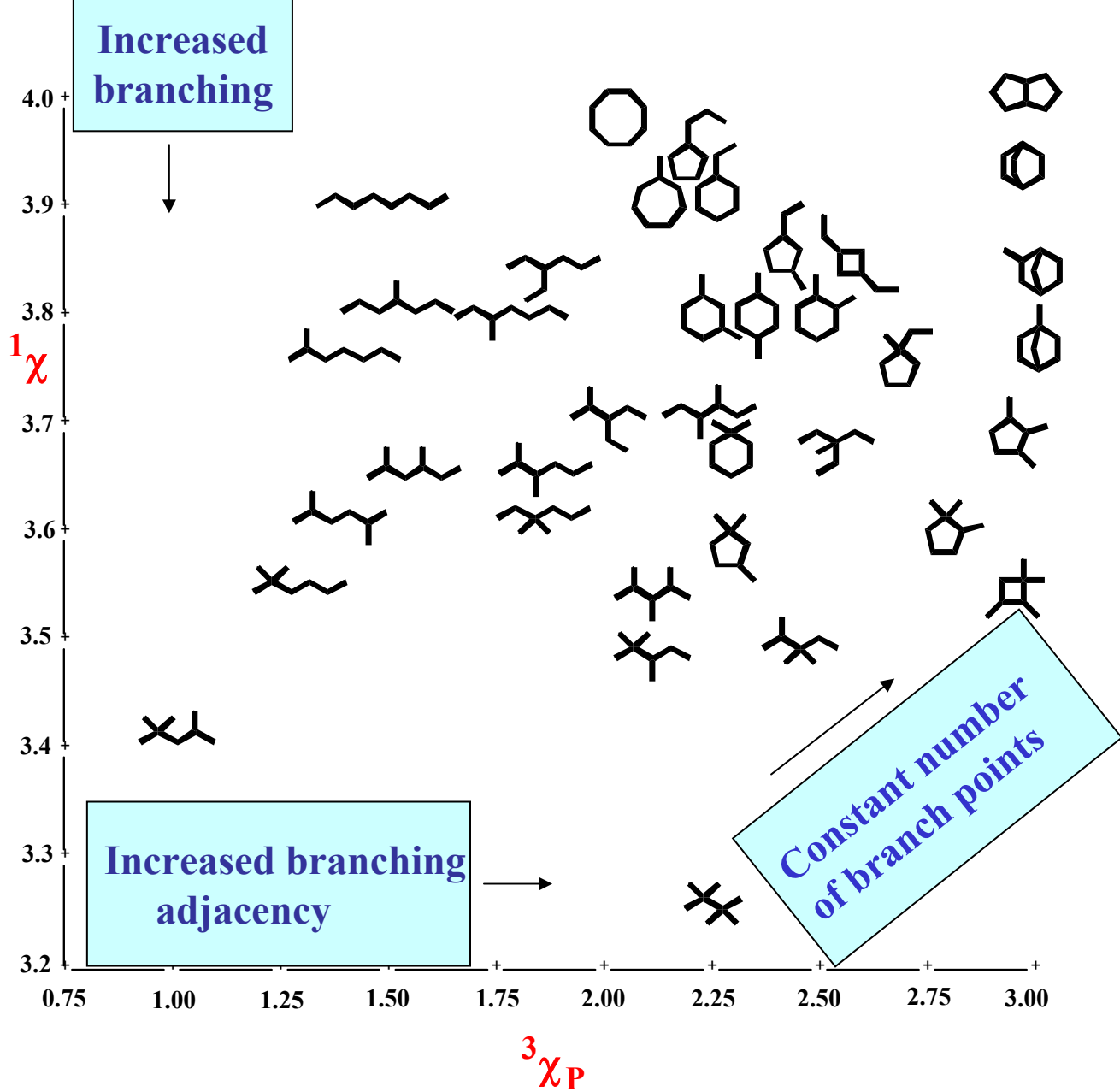
$$d^2\chi^v : -1.131 > -0.484 > -0.264$$

greater electronegativity: larger value is $d^2\chi^v$

Information is Independent of Molecular Size







Illustrations of Method with Examples

- 1 - Protein Binding Beta-lactams (115 cpds); Drugs (data set: 1000 cpds)
- 2 - Blood Brain Barrier Partitioning Training set: 114 cpds, very diverse set
- 3 - Ames Test Genotoxicity 2963 cpds training set w/drugs; 290 validate test set
- 4 - Human Intestinal Absorption 417 training set of drugs; validate set, 195 drugs
- 5 - Fish Toxicity (92 Organic Chemicals)
- 6 - Risk of Human Carcinogenicity Model FDA Data: 1050 cpds; validate set : 50 cpds
- 7 - Hepatotoxicity Model 259 Drugs training set; 33 drugs validate set
- 8 - Human Metabolic Stability Model 693 Drugs train set; 76 cpds in validate test set
- 9 - Molecular Structure Space Similarity Database Searching, Clustering, Classification

List of Papers using *Structure-Information Representation* descriptors is available:

hall@eng.umd.edu



Protein Binding to Human Serum Proteins

Penicillins

$r^2 = 0.80$, $s = 12\%$, $s_{\text{press}} = 13$ (LOO), $n = 74$

Validate Set (n = 13 commercial penicillins) : $q^2 = 0.84$, MAE = 13%

Commercial Drugs (eclectic data set, many sources)

Training set: 800 compounds

Validate Set (200 drugs)

$r^2 = 0.72$, MAE = 12.5, $n = 200$

92% within 15% of experimental value

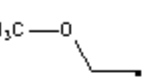
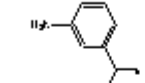
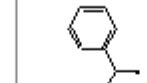
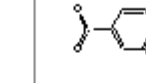
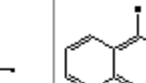
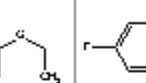
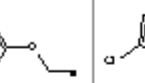
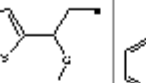
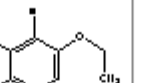
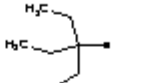
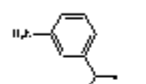
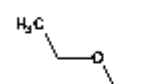
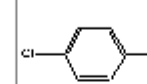
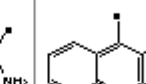
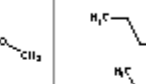
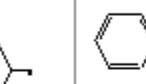
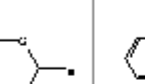
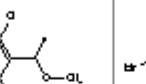
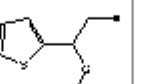
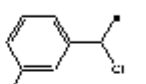
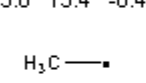
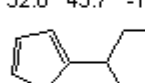
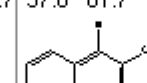
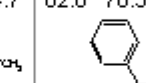
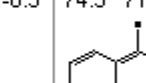
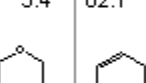
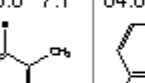
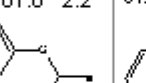
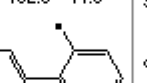
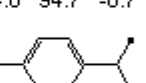
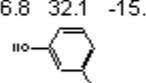
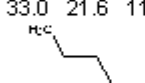
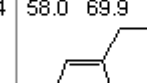
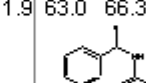
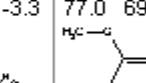
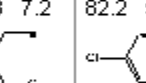
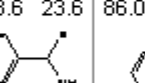
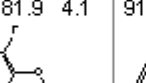
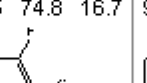
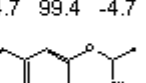
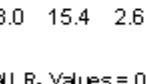
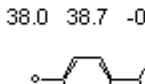
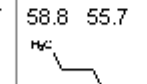
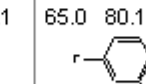
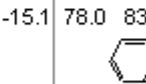
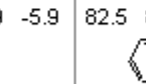
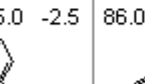
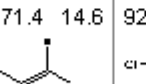
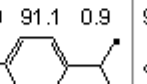
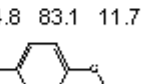
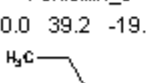
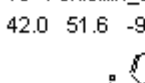
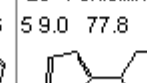
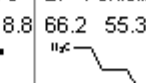
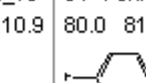
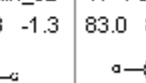
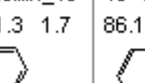
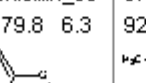
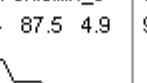
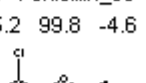
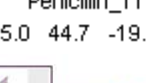
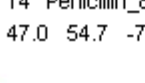
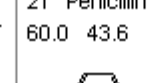
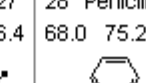
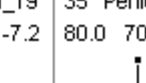
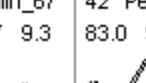
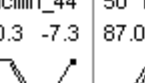
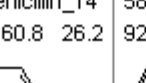
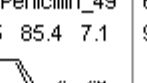
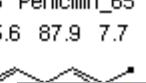
Albumin Binding Affinity (HPLC, immobilized albumin)

$r^2 = 0.77$, $s = 0.29$, $n = 84$, $q^2 = 0.70$, $s(\text{press}) = 0.33$ (LOO)

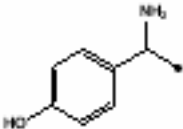
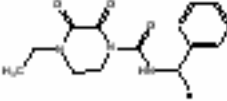
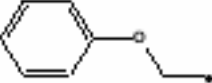
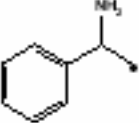
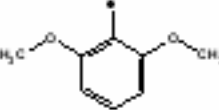
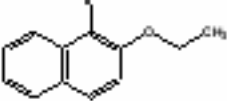
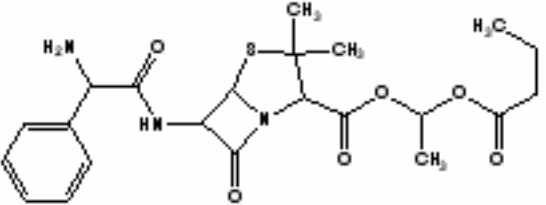
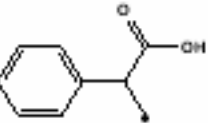
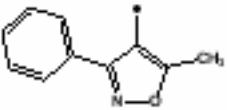
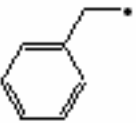
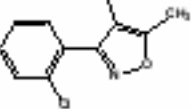
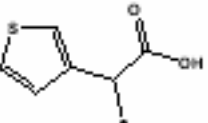
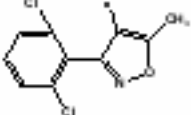
Validate Set (n = 10 drugs) : $q^2 = 0.74$, MAE = 0.31, rms = 0.32



Penicillin Training Set

1 Penicillin_5 7.2 26.7 -19.5 	8 Penicillin_30 26.0 47.5 -21.5 	15 Penicillin_37 53.2 65.5 -12.3 	22 Penicillin_46 60.0 64.8 -4.8 	29 Penicillin_78 69.7 68.3 1.4 	36 Penicillin_54 81.5 75.0 6.5 	44 Penicillin_75 83.6 93.1 -9.5 	52 Penicillin_68 89.3 77.5 11.8 	60 Penicillin_20 93.6 71.0 22.6 	68 Penicillin_57 96.0 92.7 3.3 
2 Penicillin_31 12.0 4.8 7.2 	9 Penicillin_6 28.0 33.5 -5.5 	16 Penicillin_28 55.0 49.3 5.7 	23 Penicillin_77 61.7 64.6 -2.9 	30 Penicillin_10 74.0 68.5 5.5 	37 Penicillin_48 81.5 72.2 9.3 	45 Penicillin_41 84.0 90.8 -6.8 	53 Penicillin_74 89.7 95.7 -6.0 	61 Penicillin_23 94.0 94.8 -0.8 	69 Penicillin_43 96.5 101.0 -4.5 
3 Penicillin_2 15.0 15.4 -0.4 	10 Penicillin_72 32.0 45.7 -13.7 	17 Penicillin_76 57.0 61.7 -4.7 	24 Penicillin_38 62.0 70.3 -8.3 	31 Penicillin_79 74.5 71.1 3.4 	38 Penicillin_66 82.1 75.0 7.1 	46 Penicillin_61 84.0 81.8 2.2 	54 Penicillin_16 91.0 102.0 -11.0 	62 Penicillin_24 94.0 94.7 -0.7 	70 Penicillin_25 97.0 104.6 -7.6 
4 Penicillin_32 16.8 32.1 -15.3 	11 Penicillin_12 33.0 21.6 11.4 	18 Penicillin_71 58.0 69.9 -11.9 	25 Penicillin_36 63.0 66.3 -3.3 	32 Penicillin_18 77.0 69.8 7.2 	39 Penicillin_29 82.2 58.6 23.6 	47 Penicillin_60 86.0 81.9 4.1 	55 Penicillin_52 91.5 74.8 16.7 	63 Penicillin_69 94.7 99.4 -4.7 	71 Penicillin_59 97.0 94.1 2.9 
5 Penicillin_1 18.0 15.4 2.6 All R ₂ Values = 0 	12 Penicillin_34 38.0 38.7 -0.7 	19 Penicillin_7 58.8 55.7 3.1 	26 Penicillin_45 65.0 80.1 -15.1 	33 Penicillin_22 78.0 83.9 -5.9 	40 Penicillin_21 82.5 85.0 -2.5 	48 Penicillin_64 86.0 71.4 14.6 	56 Penicillin_42 92.0 91.1 0.9 	64 Penicillin_56 94.8 83.1 11.7 	72 Penicillin_51 97.2 111.1 -13.9 
6 Penicillin_9 20.0 39.2 -19.2 	13 Penicillin_35 42.0 51.6 -9.6 	20 Penicillin_73 59.0 77.8 -18.8 	27 Penicillin_13 66.2 55.3 10.9 	34 Penicillin_62 80.0 81.3 -1.3 	41 Penicillin_40 83.0 81.3 1.7 	49 Penicillin_50 86.1 79.8 6.3 	57 Penicillin_3 92.4 87.5 4.9 	65 Penicillin_63 95.2 99.8 -4.6 	73 Penicillin_58 97.4 92.9 4.5 
7 Penicillin_11 25.0 44.7 -19.7 	14 Penicillin_8 47.0 54.7 -7.7 	21 Penicillin_27 60.0 43.6 16.4 	28 Penicillin_19 68.0 75.2 -7.2 	35 Penicillin_67 80.0 70.7 9.3 	42 Penicillin_44 83.0 90.3 -7.3 	50 Penicillin_14 87.0 60.8 26.2 	58 Penicillin_49 92.5 85.4 7.1 	66 Penicillin_65 95.6 87.9 7.7 	74 Penicillin_70 97.4 83.2 14.2 

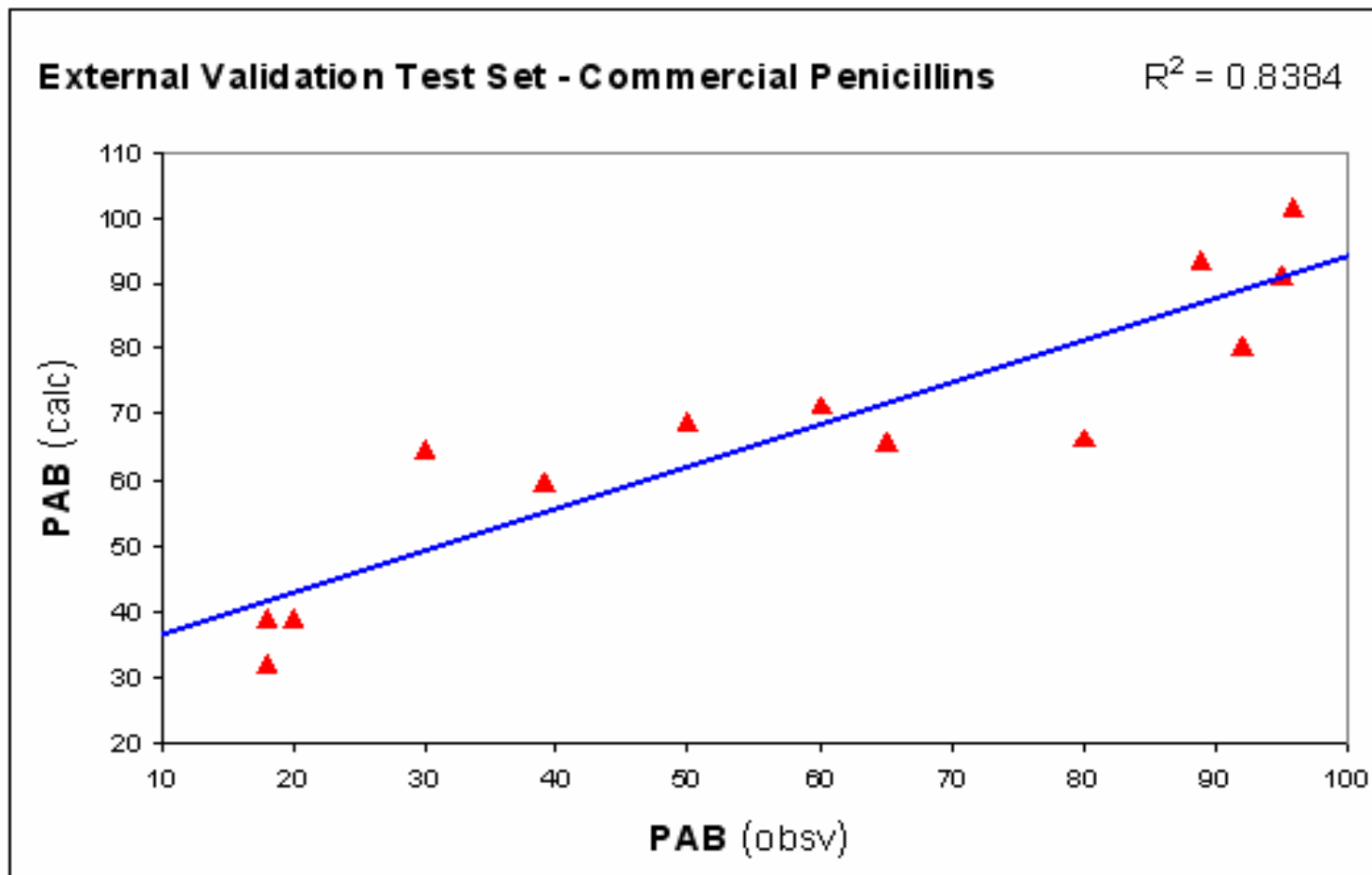
Commercial Penicillin Validation Set

ID	Name	PAB Prd. Res	R ₂	ID	Name	PAB Prd. Res	R ₂	ID	Name	PAB Prd. Res	R ₂
75	Amoxicillin	18.0 32.3 -14.3		78	Piperacillin	30.0 64.7 -34.7		83	Penicillin V	80.0 66.5 13.5	
76	Ampicillin	18.0 38.8 -20.8		79	Methicillin	39.0 59.8 -20.8		84	Nafcillin	89.0 93.3 -4.3	
77	Bacampicillin	20.0 38.8 -18.8		80	Carbenicillin	50.0 68.9 -18.9		85	Oxacillin	92.0 80.5 11.5	
				81	Penicillin G	60.0 71.5 -11.5		86	Cloxacillin	95.0 91.0 4.0	
				82	Ticarcillin	65.0 65.7 -0.7		87	Dicloxacillin	96.0 101.6 -5.6	



Calculated vs Observed PPB

- Commercial Penicillins -



$R^2 = 0.84$, MAE = 12.7%



Structure Interpretation

* Protein Binding Increased by:

➤ increased electron accessibility and number of aromatic/saturated carbons :



➤ increased branching in R_2 :



➤ increased electron accessibility and number of halogens:



* Protein Binding Decreased by:

➤ increased hydrogen accessibility and number of amine groups in R_2 :



Quantitative Calculations Done with Equation Model



Blood-Brain Barrier Partition Model

Training Set

$r^2 = 0.79$, $s = 0.37$, $F = 38.8$, $n = 93$ $q^2 = 0.73$, $s_{\text{press}} = 0.41$ (LOO)

External Validation Test Sets

$r^2 = 0.79$, $\text{rms} = 0.38$, $\text{MAE} = 0.29$, $n = 11$

$r^2 = 0.63$, $\text{rms} = 0.58$, $\text{MAE} = 0.60$, $n = 6$

Descriptors -

Positive factors:

Hydrogen E-State for aromatic CHs

Non-polarity index

Atom Type E-State for -F, -Cl

E-State for Internal H Bonding

Atom Type E-State for >N-

Negative factors:

H E-State Sum for H-Bond donors

Difference chi valence-two index

(skeletal branching)

Atom Type E-State for ==N==

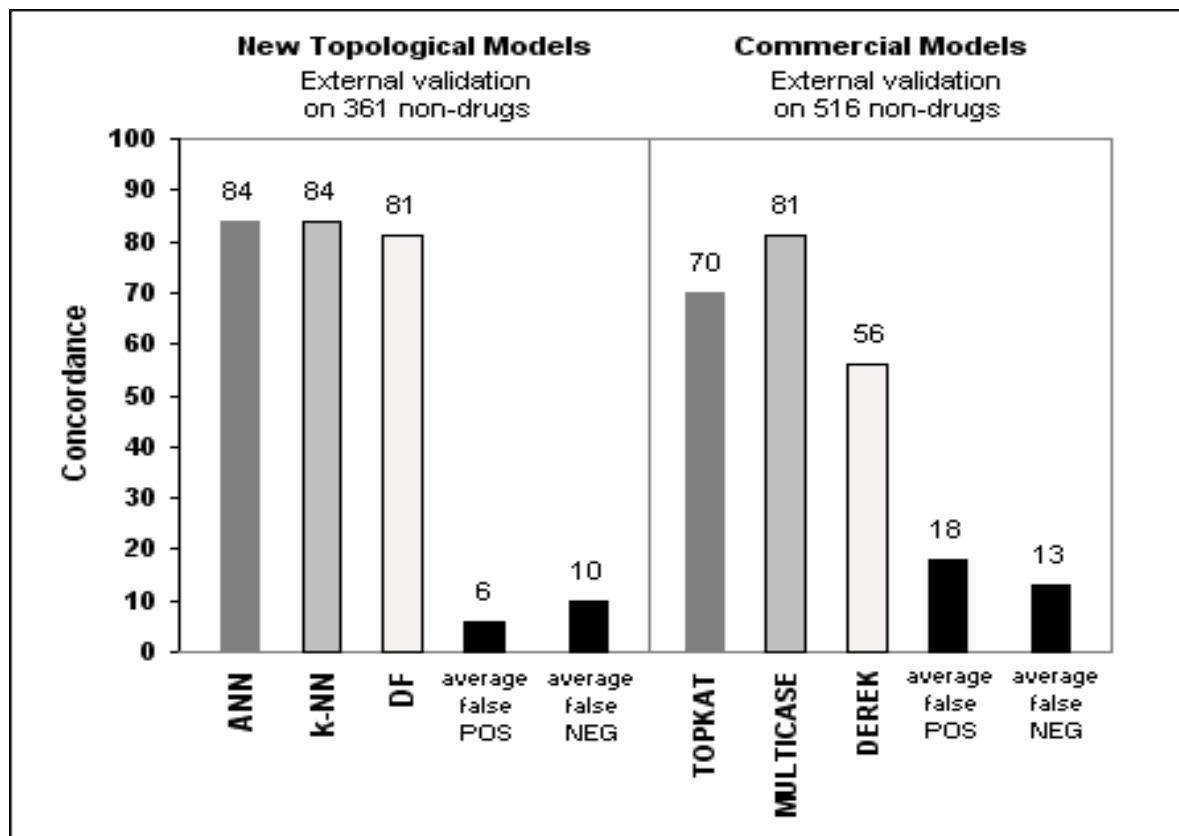


Statistics for Three Models of Genotoxicity

Model	N	M	Concordance (train)	N	Concordance (validate)
ANN	2963	38	89%	400	83%
DF	2963	144	90%	400	82%
KNN	2963	51	84%	400	84%
			Therapeutic Drugs	Therapeutic Drugs	
ANN	290	38	93%	39	92%
DF	290	144	93%	39	85%
KNN	290	51	90%	39	69%



Ames Test GenoToxicity External Validation



Pearl, G.M., Livingston-Carr, S., Durham, S.K. (2001)
Integration of computational analysis as a sentinel tool
in toxicological assessments,
Curr.Topics Med. Chem., 1, 247-255



Top 10 Descriptors – Ames Genotoxicity

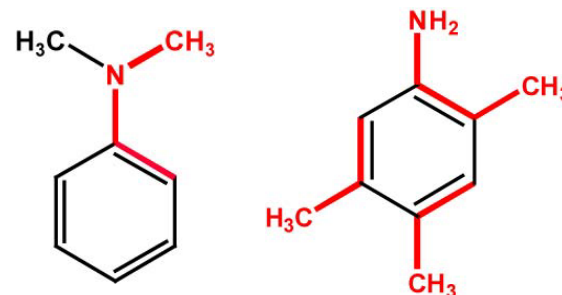
ANN Model		
Variable	Trend	Frequency
ArNH21	+	417
Hmax	+	2963
ArHNH21	+	245
Gmax	variable	2963
SHBint2	variable	665
Qv	-	2963
eaC2C3s	+	1986
SdsN	+	370
Gmin	+	2963
SddsN	+	284

Descriptor

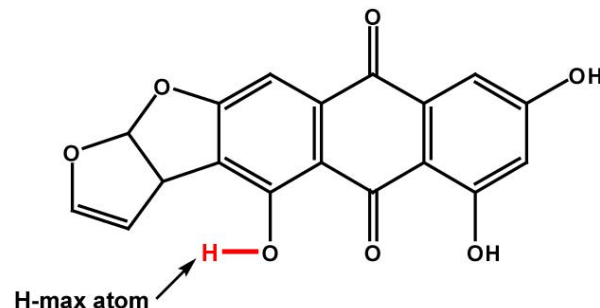
Known Structure Alert

eaC2C3s

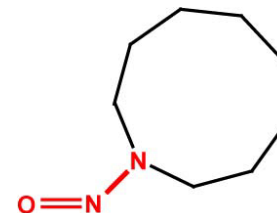
ArNH21



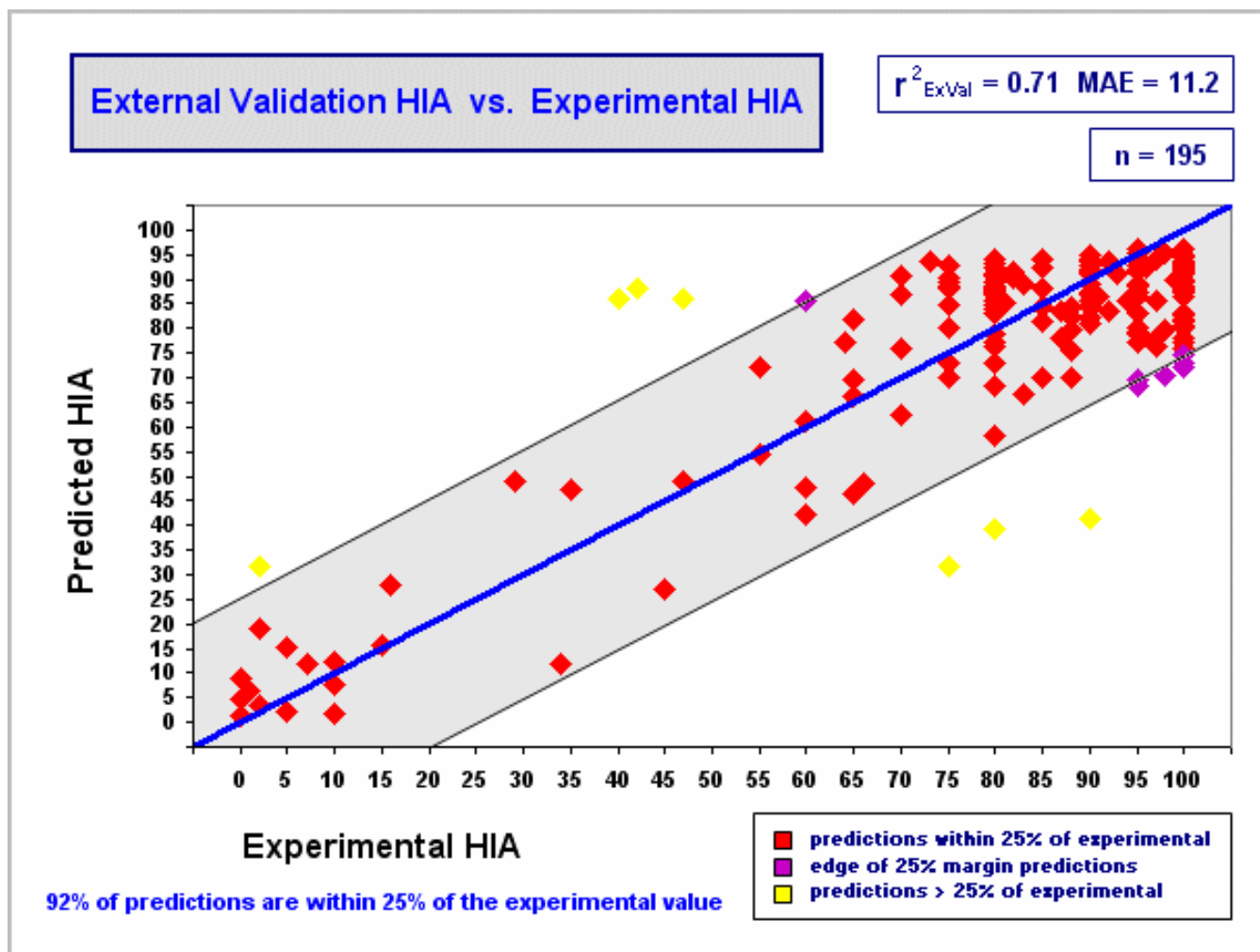
Hmax



SdsN



Oral Absorption External Validation

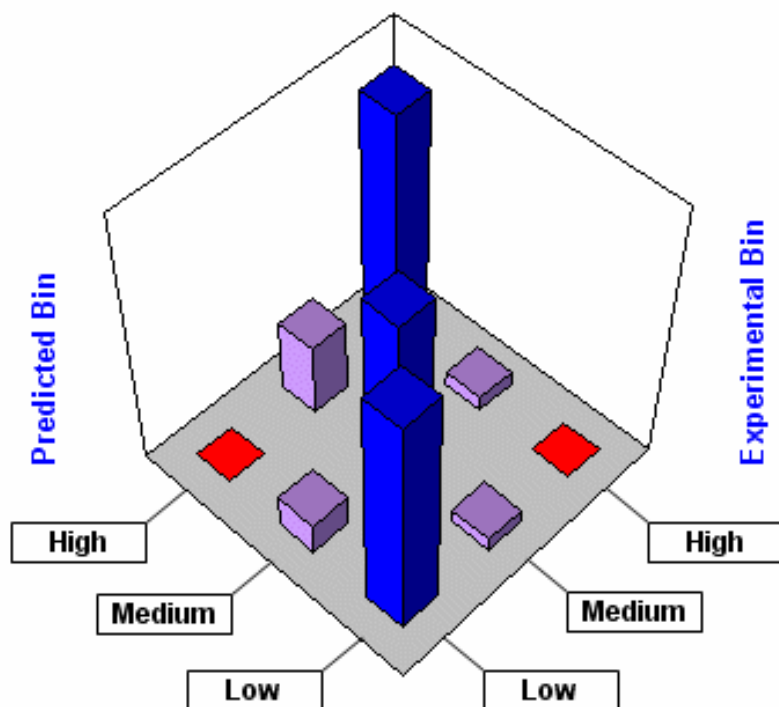


Training set: 417 drugs



Oral Absorption: Bin Assignment

Oral Absorption by Bin Value
External Validation vs Experimental



Predicted Bin	high	0.0% 0/16	28.6% 6/21	94.3% 149/158
	medium	12.5% 2/16	66.7% 14/21	5.7% 9/158
	low	87.5% 14/16	4.8% 1/21	0.0% 0/158

low	medium	high
0-20%	21-69%	70-100%
MAE = 8	MAE = 17	MAE = 11
Experimental Bin		

91% of predictions are placed in the correct experimental bin

- compounds in experimental bin
- compounds + or - 1 bin
- compounds + or - 2 bins



Top 10 Descriptors – Passive Oral Absorption

Human Intestinal Absorption (train)

Variable	Trend	Rank	Frequency
SHBint4	+	1.55	59
ArNH1	+	1.44	52
SHBint2	-	1.33	146
CSLogP	+	1.28	308
SCarOH1	-	1.27	57
e1C3O1d	-	1.26	66
Qv	+	1.25	308
Hmax	-	1.19	308
SssO	-	1.19	147
THBSA	-	1.16	308



Human Carcinogenicity Risk

Design and Validation

➤ FDA Center for Drug Evaluation and Research (CDER)

2 year Rodent Carcinogenicity database and historical assessments used

Structure-based model: long term human carcinogenic risk.

Endpoint: High Risk, Low Risk

Method: Tennant, R. W. (1993). “Stratification of rodent carcinogenicity bioassay results to reflect relative human hazard” *Mutation Res.*, 286, 111-118

➤ Built on 1022 compound training set

- Non-proprietary data
- Discriminant Analysis used to create model
- FDA Center for Drug Evaluation and Research (CDER)

Validation Statistics*

- **n = 50 compounds** (25 positive, 25 negative)
- **Concordance** = $\frac{\text{total found correct}}{\text{total data set}}$ = **80%**
- **Sensitivity** = $\frac{\text{carcinogenic found}}{\text{total known carcinogenic}}$ = **76%**
- **Specificity** = $\frac{\text{non-carcinogenic found}}{\text{total known non-carcinogenic}}$ = **84%**
- **For comparison : MultiCase results, Specificity: ~ 59%**

* **Validation performed by Center for Drug Evaluation and Research (CDER), US Food and Drug Administration**

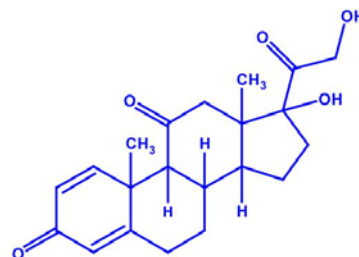
Publication: J. Contrera, P. MacLaughlin, L. Hall, L. Kier,
Current Drug Discovery Technologies, 2005 (in press).



Database Similarity Search: Prednisone

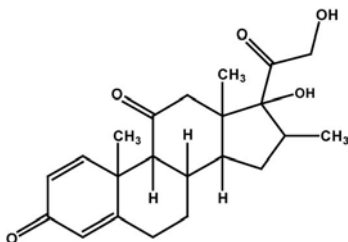
Search Based on all **eight atoms types** in the structure

Prednisone

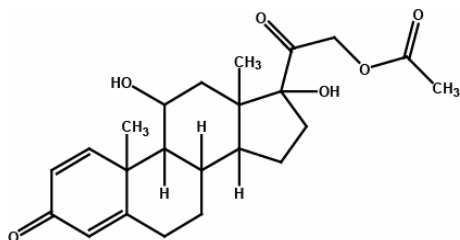


<u>Distance</u>	<u>Cosine</u>
reference	1.000

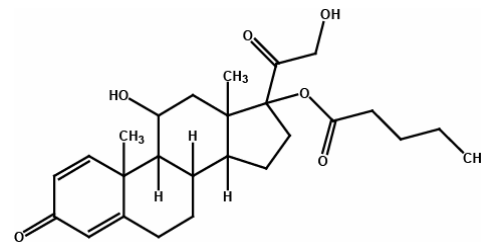
<u>Structure</u>	<u>Distance</u>	<u>Cosine</u>
Meprednisone	0.75	0.989



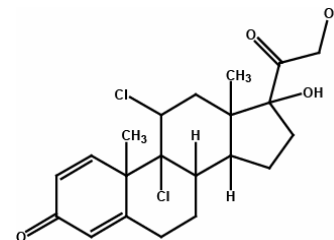
Prednisolone acetate	0.88	0.984
----------------------	------	-------



<u>Structure</u>	<u>Distance</u>	<u>Cosine</u>
Prednival	1.18	0.973



Dichlorisone	1.29	0.964
--------------	------	-------



Conclusions

- **Structure-Information Representation provides basis for structure representation, versatile method, applied to**
 - * **binding**
 - * **inhibition**
 - * **ADME/Tox**
 - * **database similarity searching . . .**
- **Structure-Information Representation applied using standard statistical methods**
- **Structure-Information Representation very fast in calculations**
- **Structure features identified in model useful for drug / molecule design process**



Acknowledgements

Kimberly Rose, MD-PhD program Wake Forest University

* blood-brain barrier, fish toxicity

Hlaing Hlaing Maw, MS program, Tufts University

* fish toxicity, HIV-1 inhibitors

L. Mark Hall, Hall Associates Consulting, ChemSilico LLC

* protein binding, blood-brain barrier, HIA

Borislav Spasov, Senior, Eastern Nazarene College

* dipeptide ACE inhibitors, bitter taste

Lemont B. Kier, Virginia Commonwealth University, ChemSilico LLC

* long time colleague

Marc E. Parham, Joe Votano, ChemSilico LLC

* colleagues

- Northeast Region American Chemical Society, Summer Scholar Program
- Shrader Society Endowment, Eastern Nazarene College

References Available

Handout – list of papers/sources cited in this talk

**List of papers that use various topological descriptors
available electronically**

Email request to hall@enc.edu
