Evaluation of ParaSurf descriptors and local properties

Cepos User Meeting, Portsmouth, 11.03.2010

Stefan Güssregen
Gerhard Hessler
Hans Matter
Marco Müller
Friedemann Schmidt

sanofi aventis
Because health matters

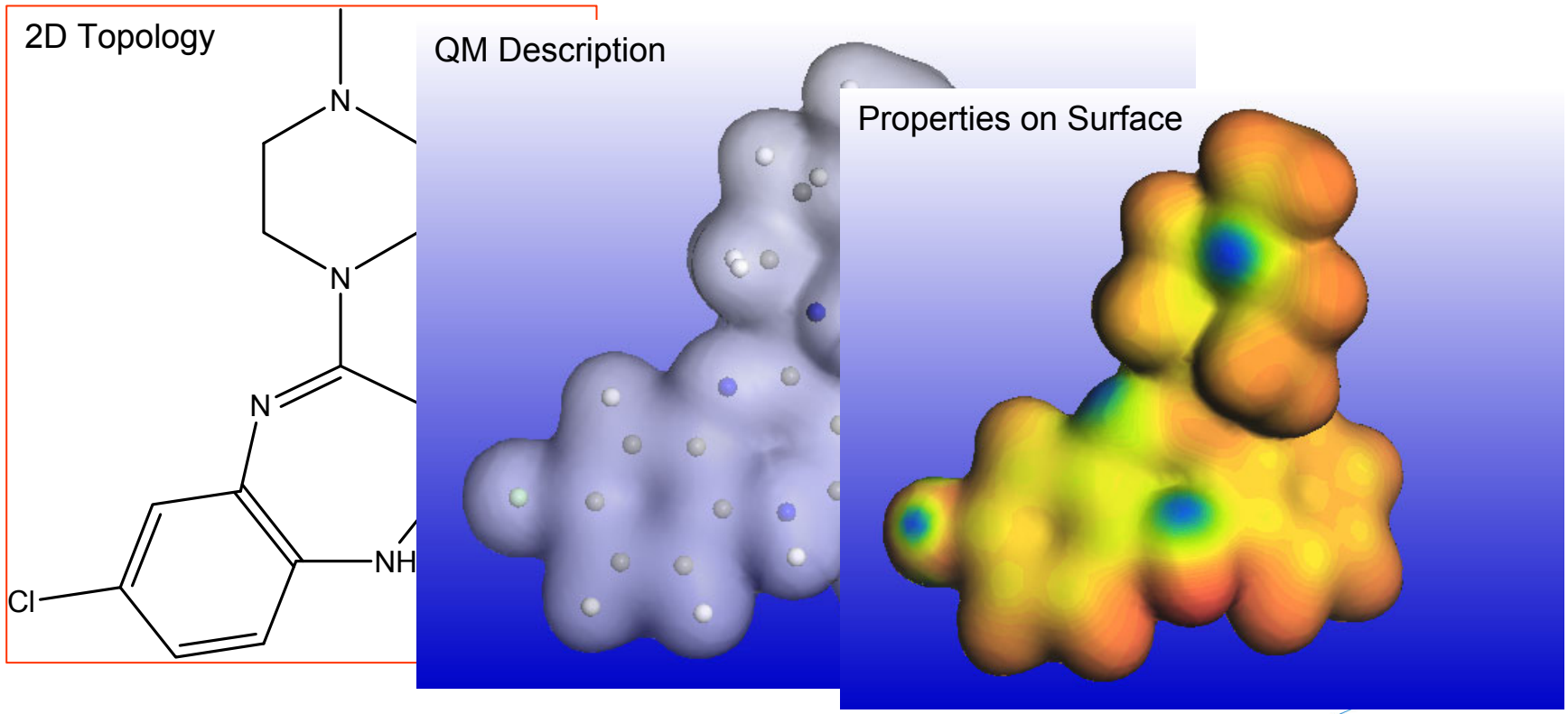
Overview

- Motivation
- Use of Surface-based descriptors for 2D-QSAR
  - Comparison with MOE Descriptors
- Use of local properties for 3D-QSAR
- Conclusions

Motivation

Why ParaSurf?

- **Provide new descriptors that are independent of topology of a molecule**

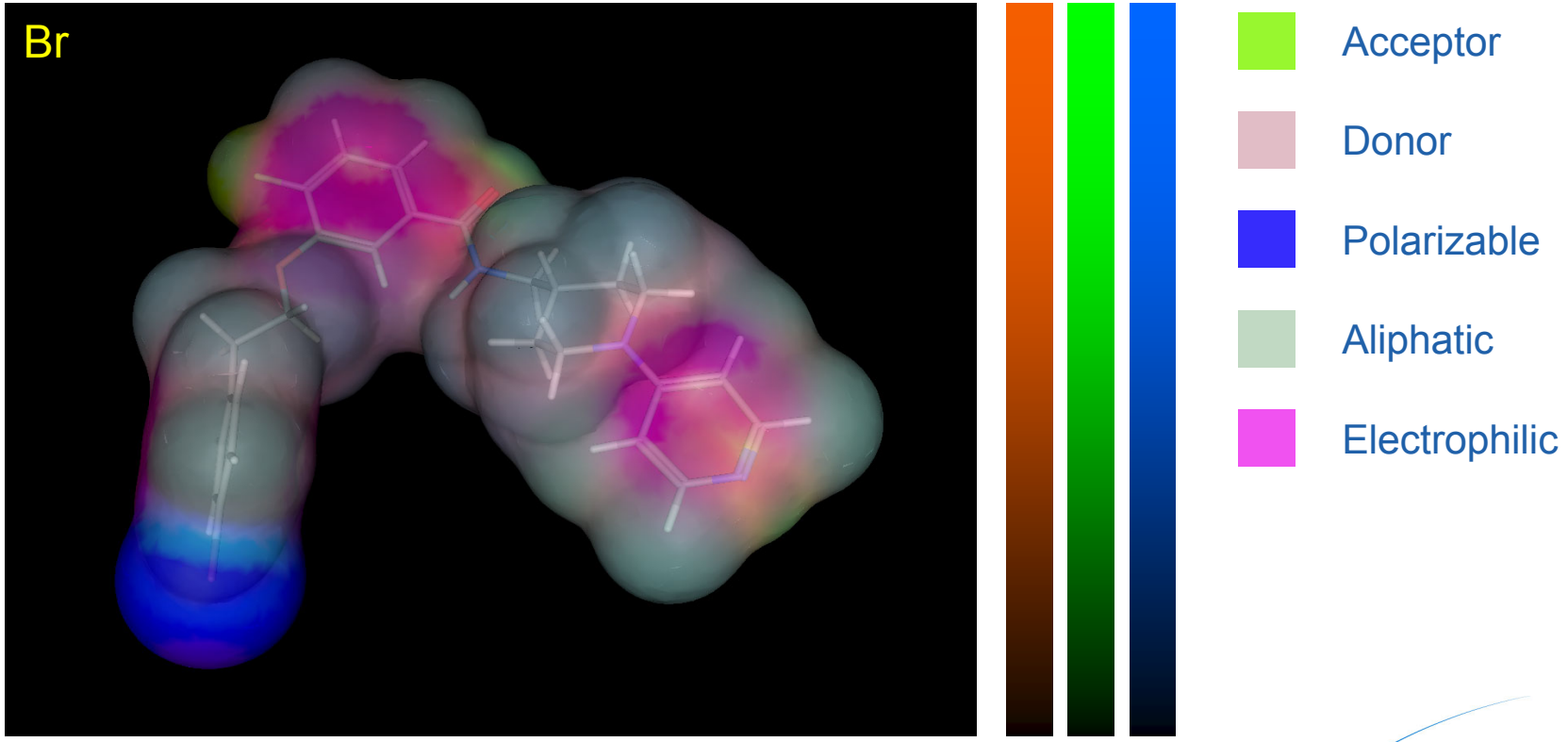


sanofi aventis
Because health matters

Taken from Cepos InSilico Ltd.

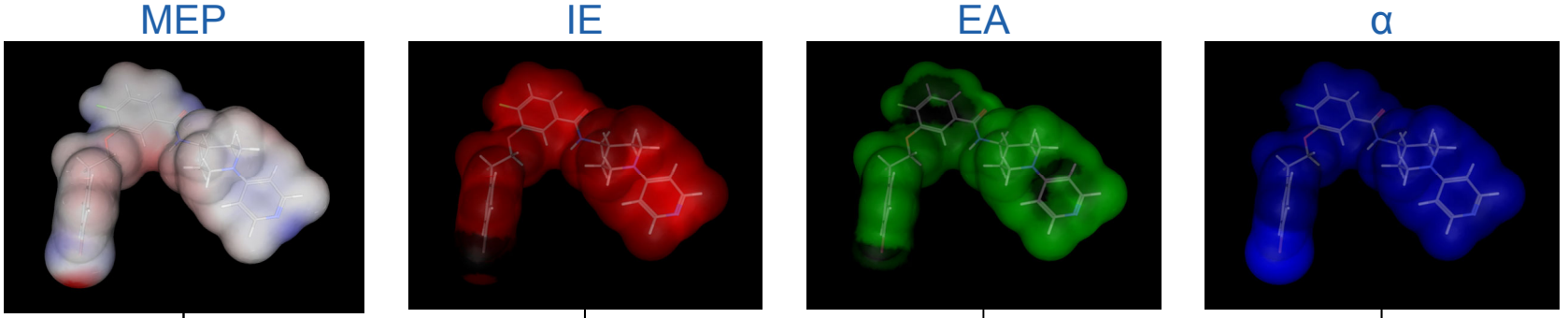
Local Properties –
Combined visualization of IA, EA and α

- Provides a fine differentiated view on the molecule
- Pharmacophoric properties and chemical reactivity are visible



Evaluation of Descriptors for 2D-QSAR

Surface-based Descriptors for 2D-QSAR



Descriptive Statistics

Mean, Max, Min, Variance, Range
Skewness, Kurtosis, Integral, etc.

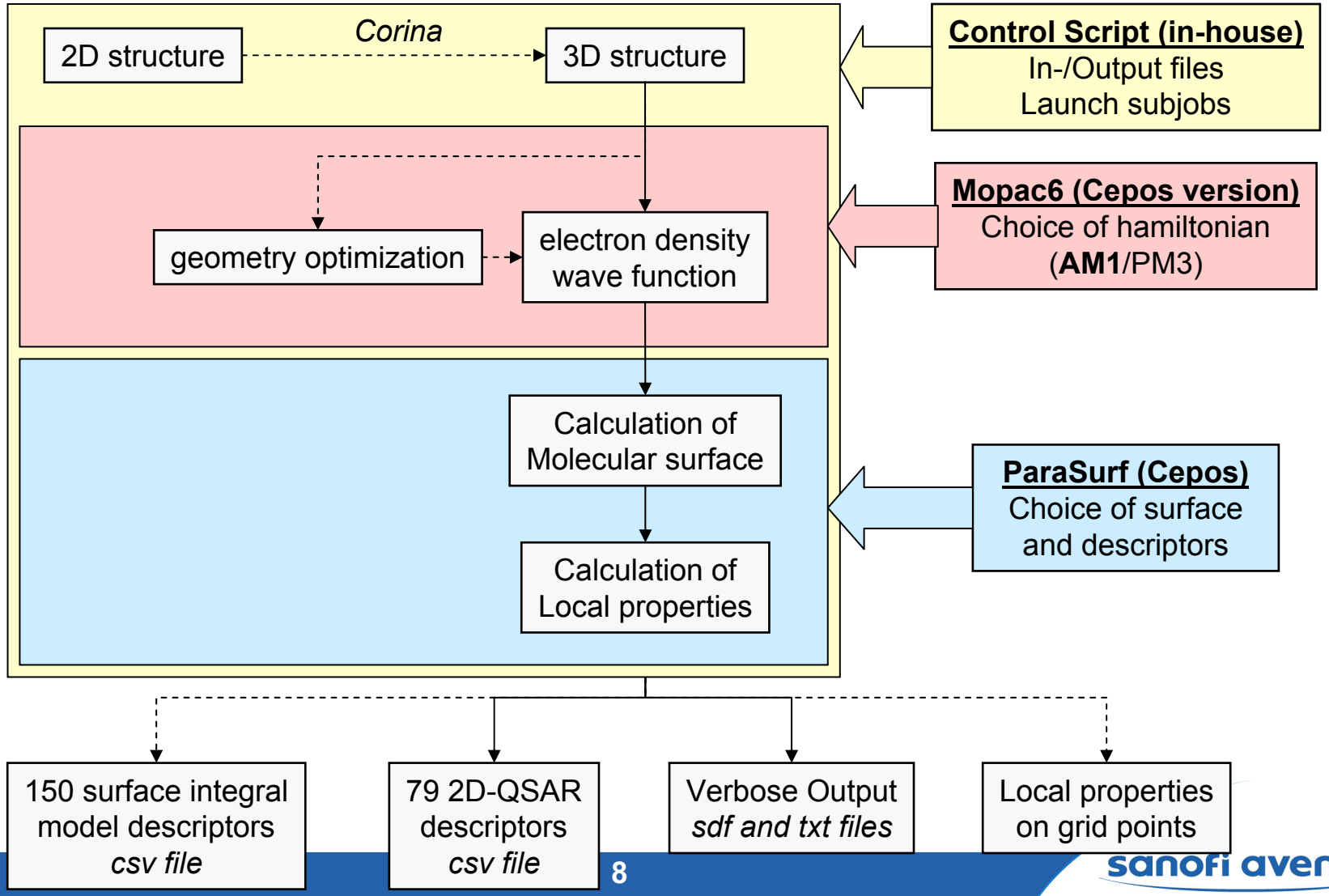
Shannon Entropy Descriptors

79 Descriptors for 2D-QSAR

Whole Molecule Descriptors

Dipole moment, dipolar density, polarizability,molecular weight, globularity, surface area, volume

ParaSurf Workflow

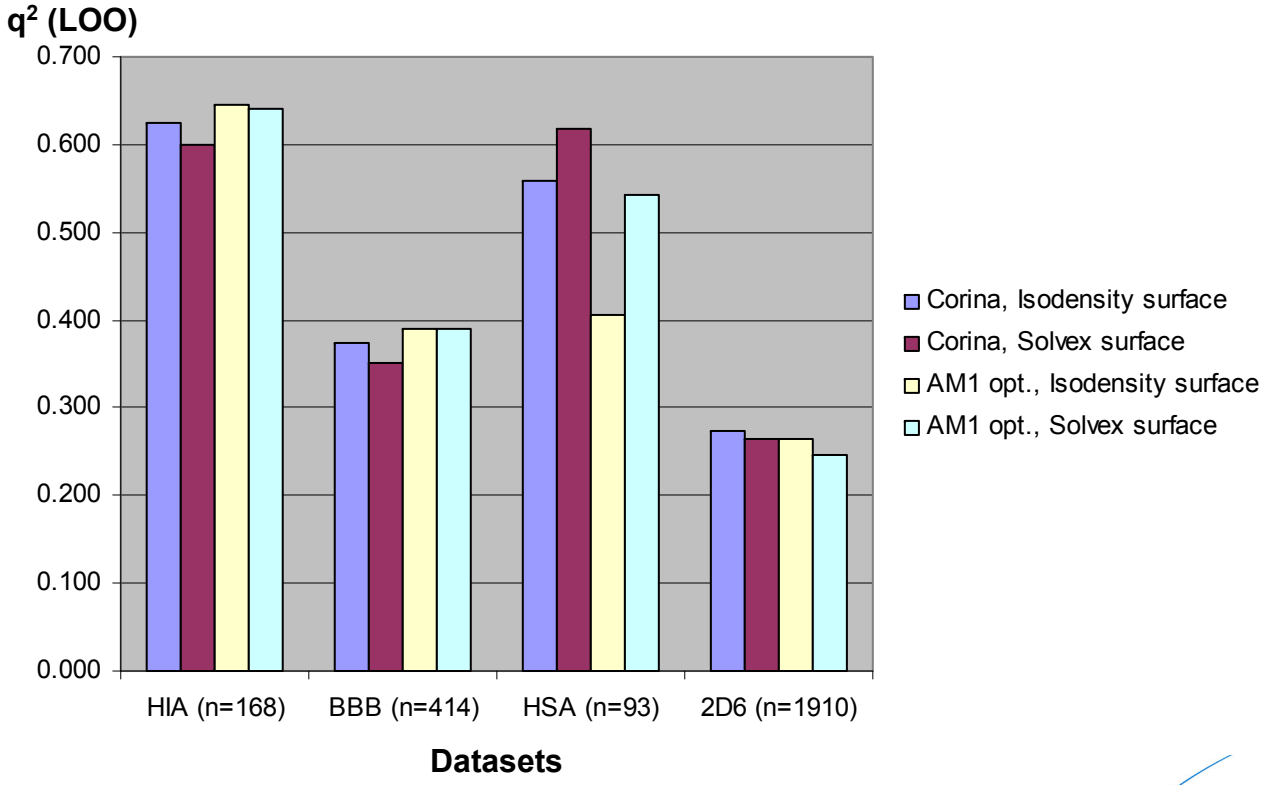


Datasets for Evaluation of 2D Descriptors

- **Human Intestinal Absorption (HIA)**
  - %HIA values for 169 compounds
  - taken from M. Abraham et al., J. Pharm. Sci. (2001), 90, 749
- **Blood-brain barrier permeability (BBB)**
  - logBB values for 419 compounds
  - 289 compounds taken from Huo et al., JCICS (2003), 43, 2137
  - 130 compounds taken from internal projects
- **Human serum albumin binding (HSA)**
  - logK'hsa values for 95 compounds
  - taken from Colmenarejo et al., J. Med. Chem. (2001), 44, 4370
- **Inhibition of human Cytochrom P450 2D6 (2D6 Inh)**
  - 1910 IC_{50} values of recombinant enzyme
  - taken from internal projects

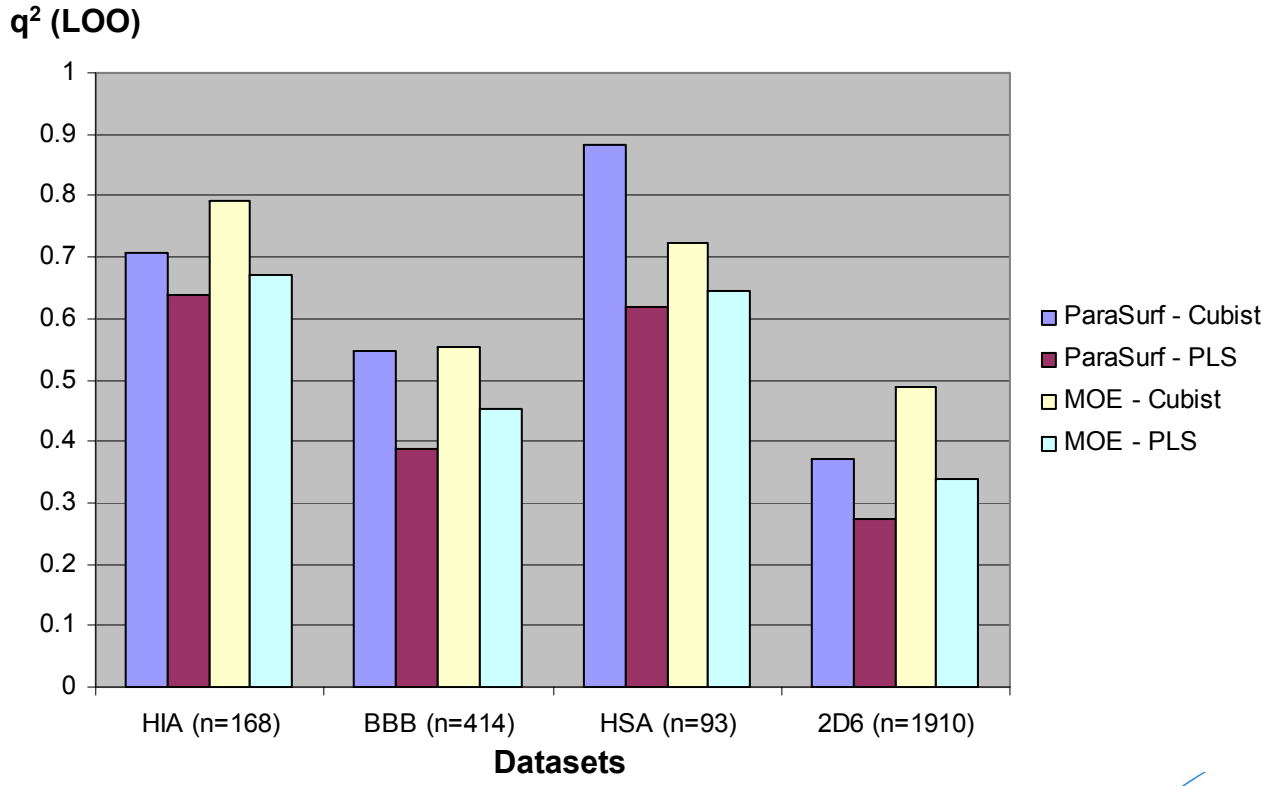
Performance of 2D QSAR Models –
Impact of Parasurf Input Parameters

- The following parameters were investigated
  - Origin of 3D structure (Corina, AM1 optimization)
  - Type of surface (Isodensity surface, solvent excluded surface)



Performance of 2D QSAR Models –
Comparison with MOE descriptors

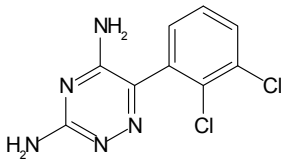
- **Models were built using PLS and Cubist regression trees**
  - **184 MOE 2008 2D descriptors**
  - **78 ParaSurf 2D descriptors**



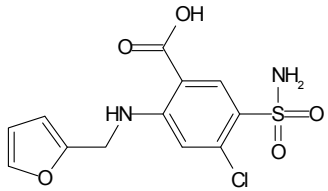
Correlation with MOE Descriptors

Descriptor Correlation ParaSurf / MOE
Datasets

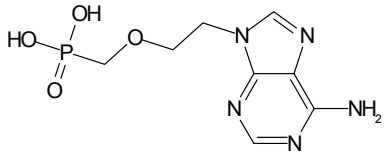
- **Human Intestinal Absorption (HIA)**
  - %HIA values for 169 compounds
  - Diverse Drug Molecules
  - M. Abraham et al., *J. Pharm. Sci.* (2001), 90, 749



Lamotrigine (98%)

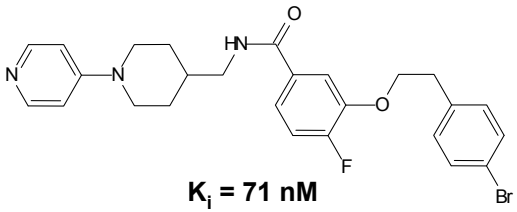


Furosemide (61%)

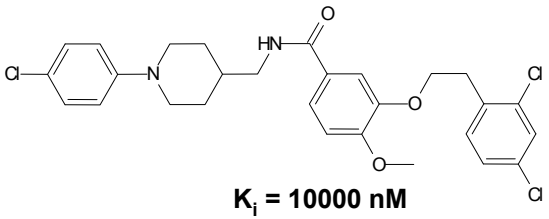


Adefovir (16%)

- **Factor Xa Inhibitors from Oxybenzamide series (FXa)**
  - logK_i values for 152 compounds
  - Highly analogous molecules from lead optimization
  - H. Matter et al., *J. Med. Chem.* (2005), 48, 3290



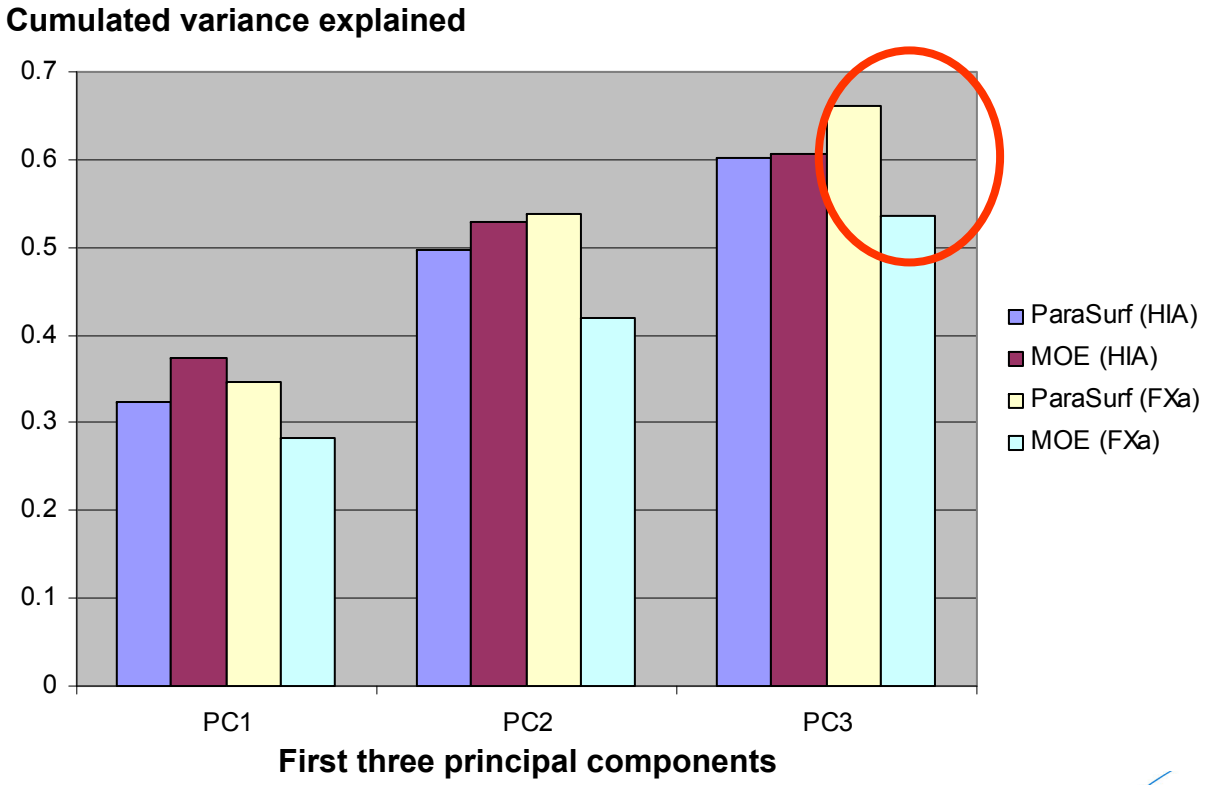
K_i = 71 nM



K_i = 10000 nM

Descriptor Correlation ParaSurf / MOE
Principle Component Analysis

- 78 ParaSurf 2D descriptors
- 184 MOE 2008 2D descriptors



sanofi aventis
Because health matters

Descriptor Correlation ParaSurf / MOE
Descriptor Distance Plots (DDP)

- **Neighborhood plots for distances MOE vs ParaSurf**
  - **Direct measure of descriptor differences**
- **For pairs of compounds, plot ParaSurf distance vs MOE distance**
  - **Autoscaling, mean centering of descriptor set**
  - **Computing pairwise descriptor Euclidean distances (MOE, ParaSurf)**
  - **Computing pairwise biological differences and structure-activity landscape indices (SALI)**

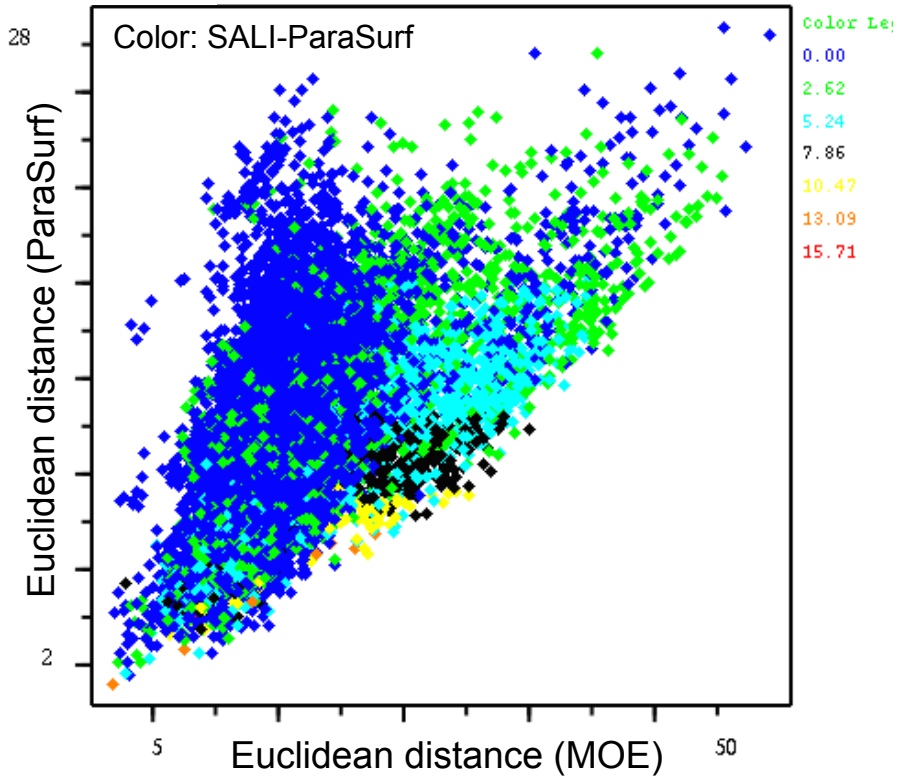
$$SALI_{i,j} = \frac{|A_i - A_j|}{1 - \mathrm{sim}(i,j)}$$

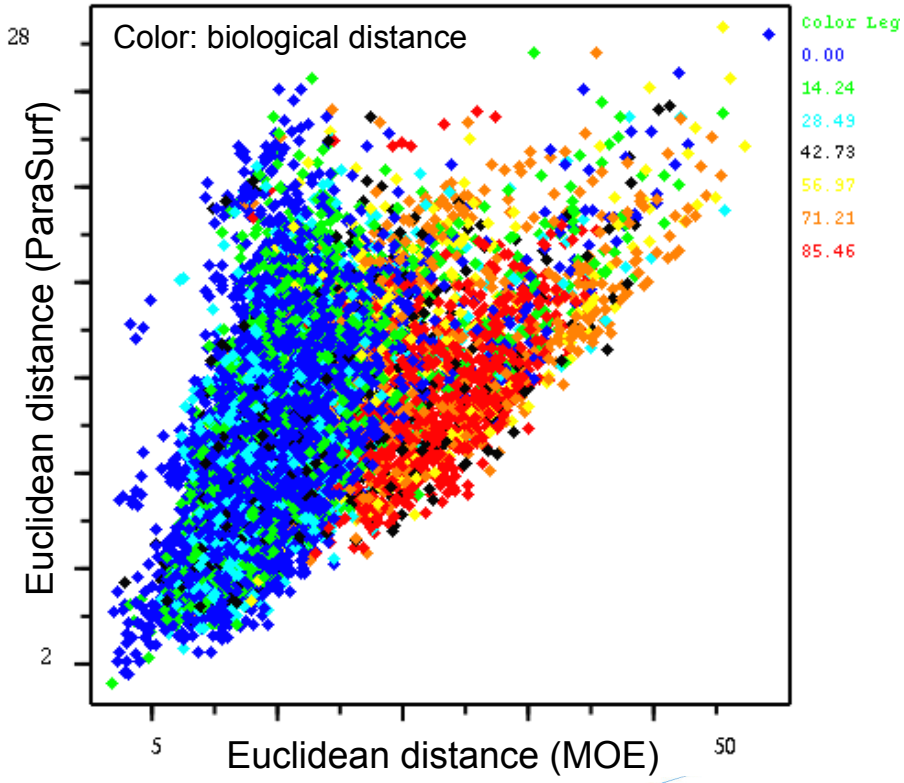
Van Drie et al., JCIM 2008, 48, 646

Descriptor Correlation ParaSurf / MOE

Descriptor Distance Plots on HIA Dataset

Correlation: Each point refers to a pair of n compounds => n * (n-1)/2 pairs
"Off-diagonal elements": Similar in descriptor A, but dissimilar in descriptor B

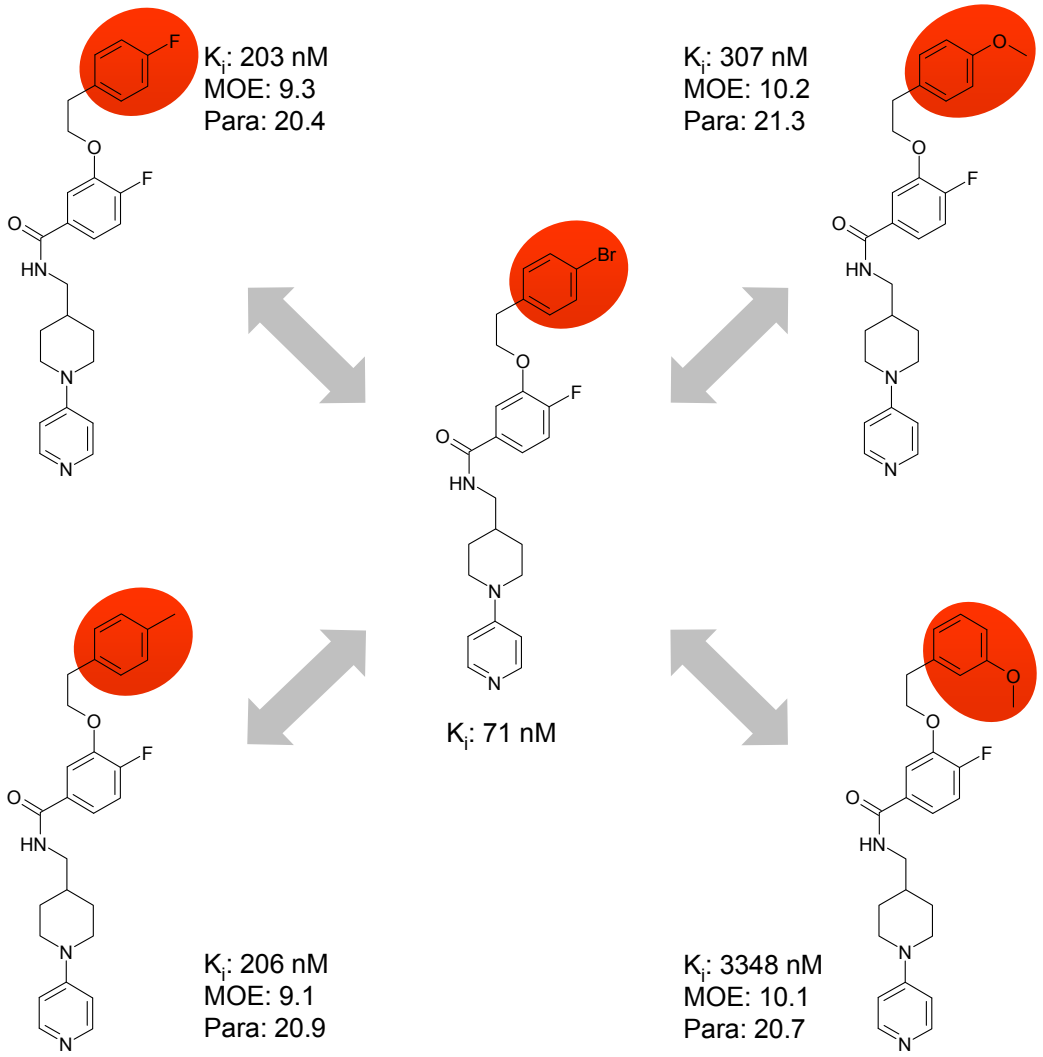


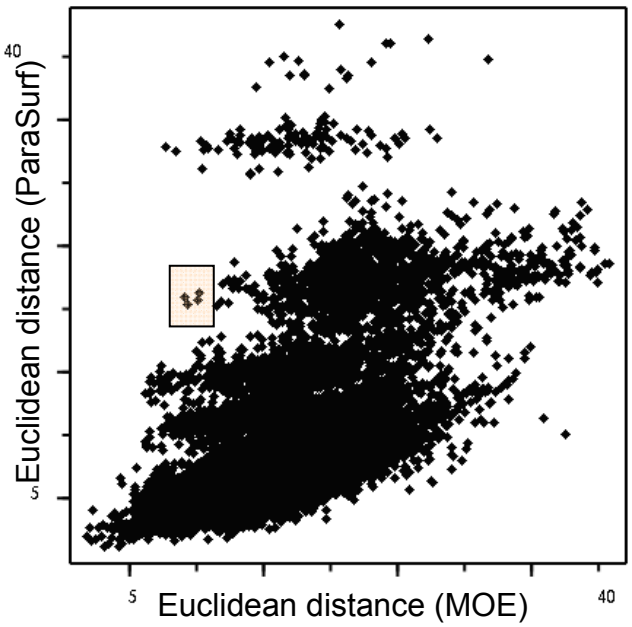


sanofi aventis
Because health matters

Descriptor Correlation ParaSurf / MOE

Descriptor Distance Plots on FXa Dataset





sanofi aventis
Because health matters

Conclusions 2D-QSAR

- Parasurf 2D descriptors provided predictive models
  - Standalone performance comparable to MOE 2D Descriptor set
- ParaSurf 2D descriptors contain different information compared to MOE
  - Provide alternative descriptor metric offering novel insights into SAR

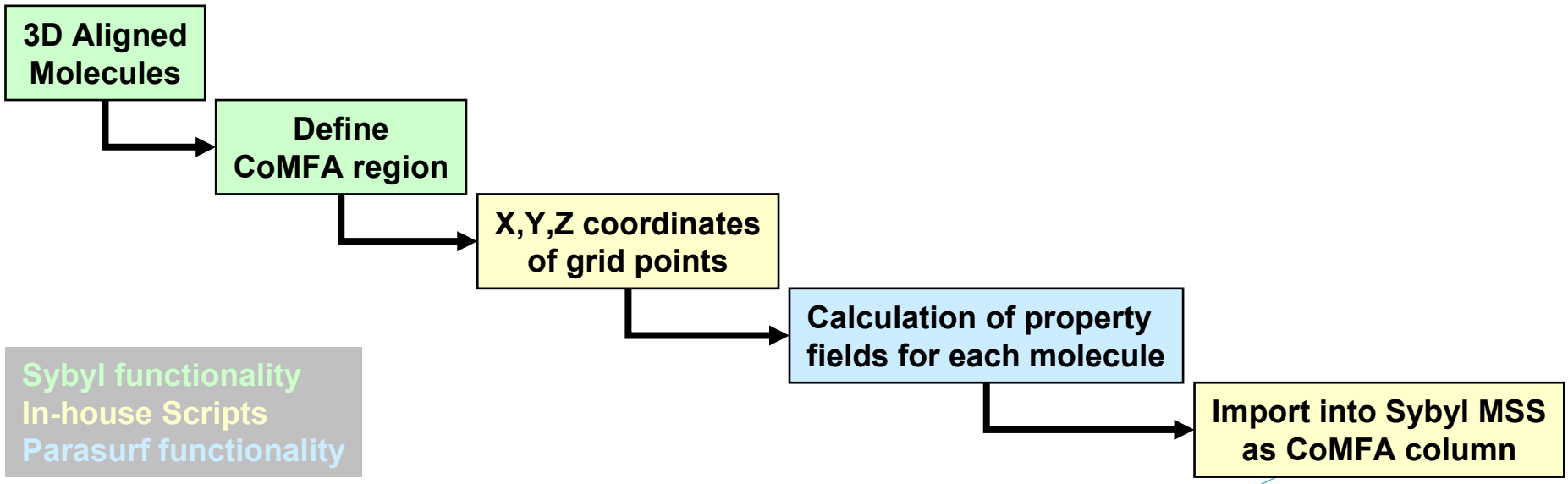
Evaluation of Local Properties for 3D-QSAR

sanofi aventis

Because health matters

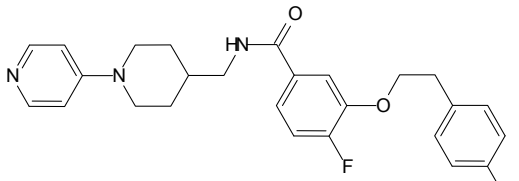
Local Properties on Grid Points

- Local properties can be calculated on any arbitrary point in space
- Can be used as local property fields for 3D-QSAR, e.g. Sybyl

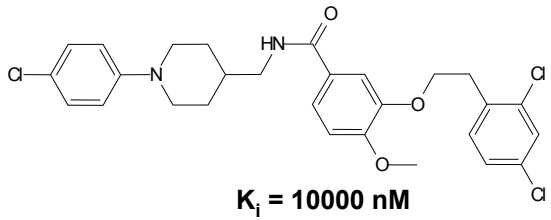


Evaluation of Local Properties
Datasets

- **Structure-based Alignment**
  - Factor Xa Inhibitors from Oxybenzamide series (FXa)
  - logK_i values for 152 compounds
  - H. Matter et al., *J. Med. Chem.* (2005), 48, 3290



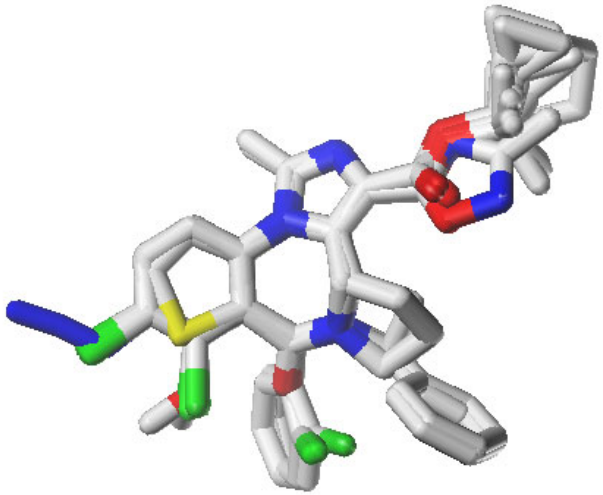
K_i = 71 nM

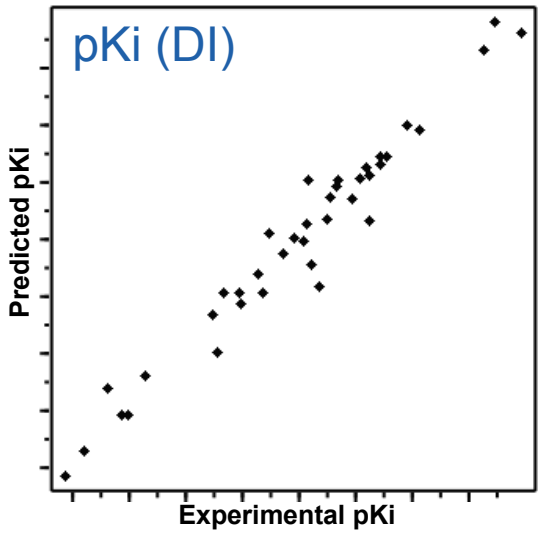


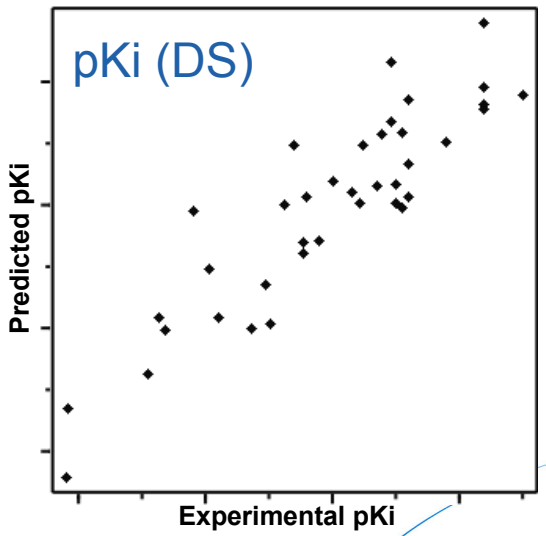
K_i = 10000 nM

- **Ligand-based Alignment**
  - Diazepam-insensitive (DI) and Diazepam-sensitive (DS) subtypes of benzodiazepine-receptor (BzR):$GABA_A$ ligand-gated ion channels
  - logK_i values for 38 1,4-Benzodiazepine-type binders
  - G. Wong, K.F. Koehler, et al. *J. Med. Chem.* (1993), 36, 1820-1830
  - Alignment obtained from the authors

$GABA_A$/Benzodiazepine Receptor Dataset –
CoMFA Model for BzR 1,4-Diazepines (I)







- **CoMFA 3D-QSAR for 38 1,4-Diazepines**
- **Affinity on two BzR subtypes:**
  - **- "diazepam-insensitive" (DI) subtype**
  - **- "diazepam-sensitive" (DS) subtype**
- **Alignment obtained from authors**
- **MMFF charges, corrected atom types and region file**

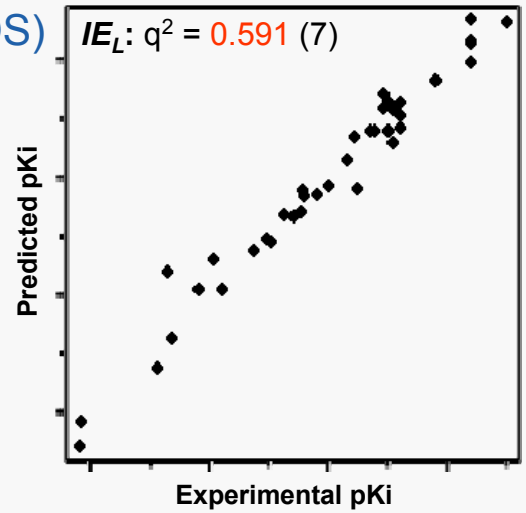
pKi (DS)	q^2 = **0.497** (3),	r^2 = 0.806
pKi (DI)	q^2 = **0.737** (7),	r^2 = 0.958

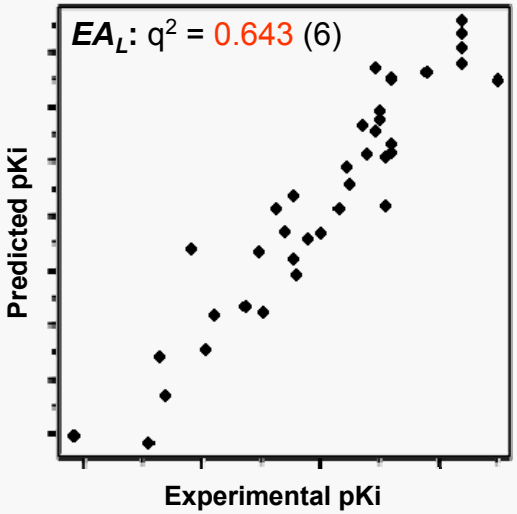
sanofi aventis
Because health matters

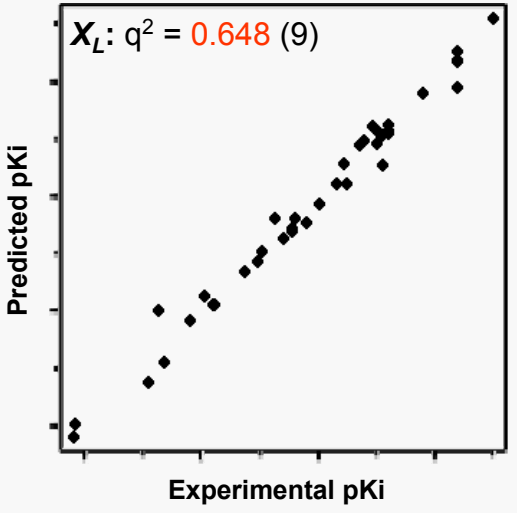
G. Wong, K.F. Koehler, et al. *J. Med. Chem.* (1993), 36, 1820-1830

$GABA_A$/Benzodiazepine Receptor Dataset –
ParaSurf 3D-QSAR Models for BzR 1,4-Diazepines

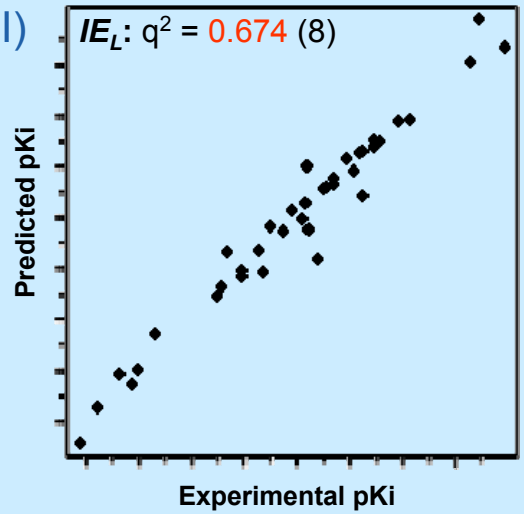
pKi (DS)

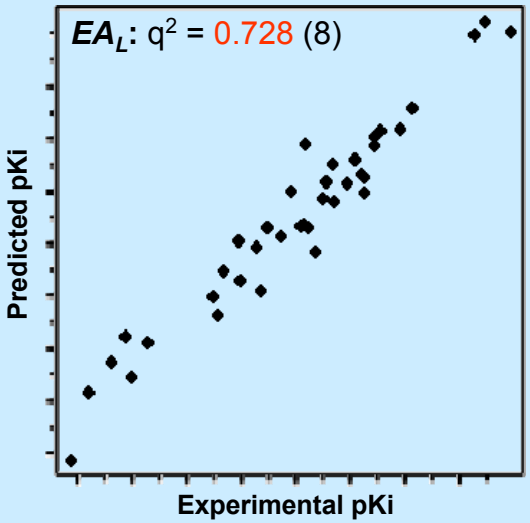


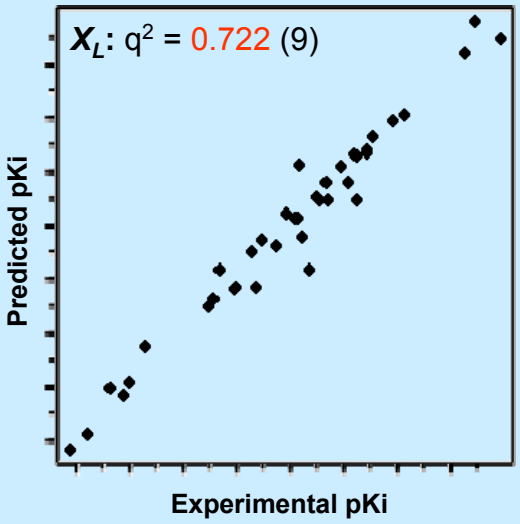




pKi (DI)





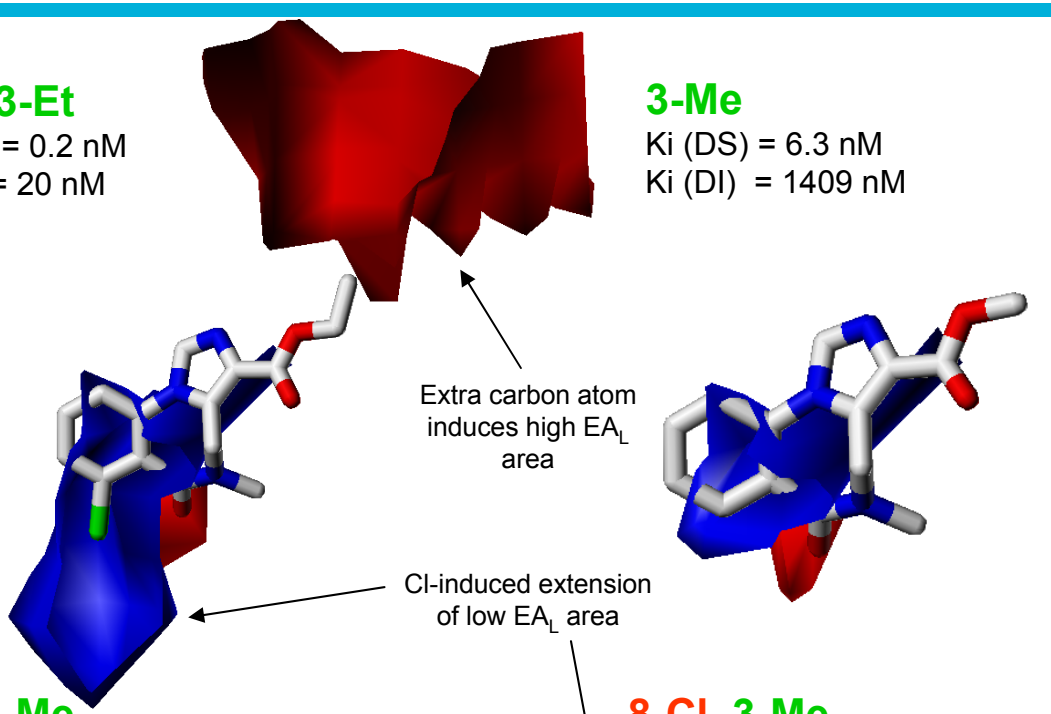


sanofi aventis
Because health matters

$GABA_A$/Benzodiazepine Receptor Dataset –
Comparison of Local Electron Affinity (EA_L) Fields

7-Cl, 3-Et
Ki (DS) = 0.2 nM
Ki (DI) = 20 nM

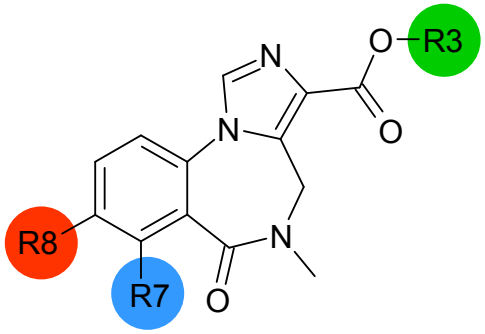
3-Me
Ki (DS) = 6.3 nM
Ki (DI) = 1409 nM



8-F, 3-Me
Ki (DS) = 3.1 nM
Ki (DI) = 239 nM

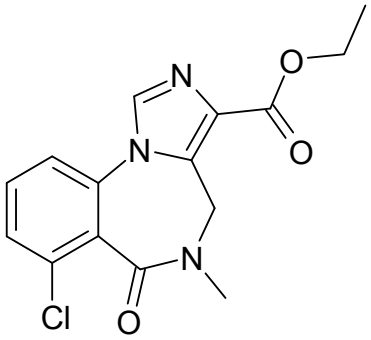
8-Cl, 3-Me
Ki (DS) = 30 nM
Ki (DI) = 123 nM

***EA_L* > -50 kcal/mol (blue)**
low local electron affinity
***EA_L* < -110 kcal/mol (high) (red)**
high local electron affinity



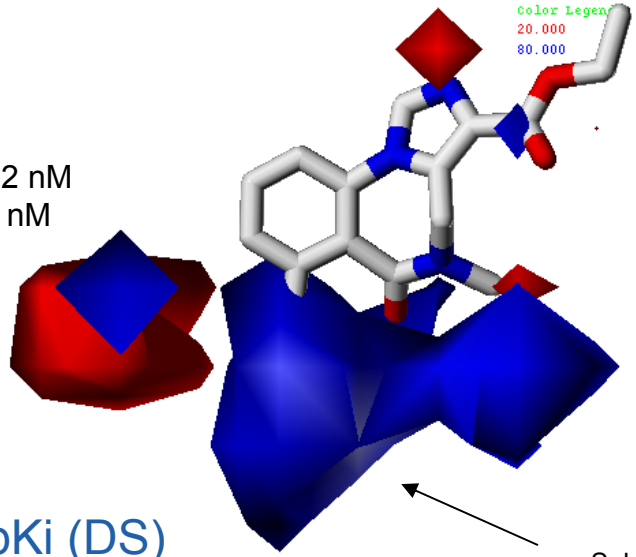
sanofi aventis
Because health matters

$GABA_A$/Benzodiazepine Receptor Dataset –
3D-QSAR using Local Electron Affinity (EA_L) Fields

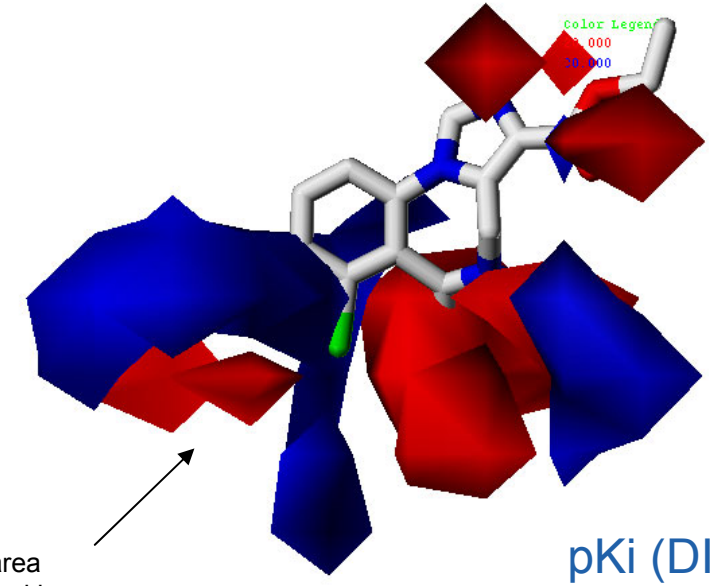


RO-153505
Ki (DS) = 0.2 nM
Ki (DI) = 20 nM

lower EA_L favorable (blue)
higher EA_L favorable (red)



pKi (DS)



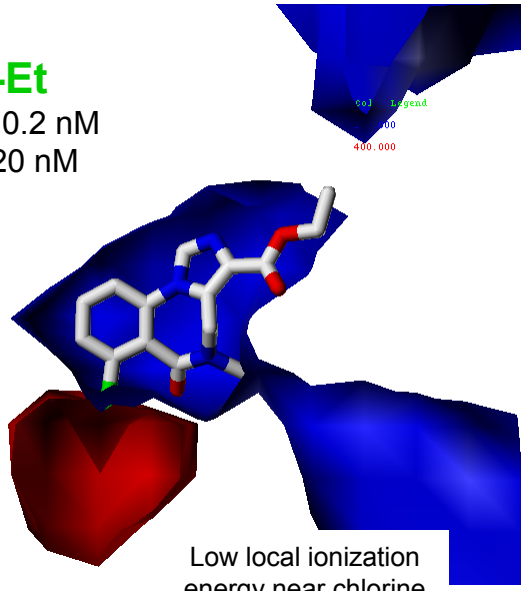
pKi (DI)

Selectivity area clearly reflected in EA_L model

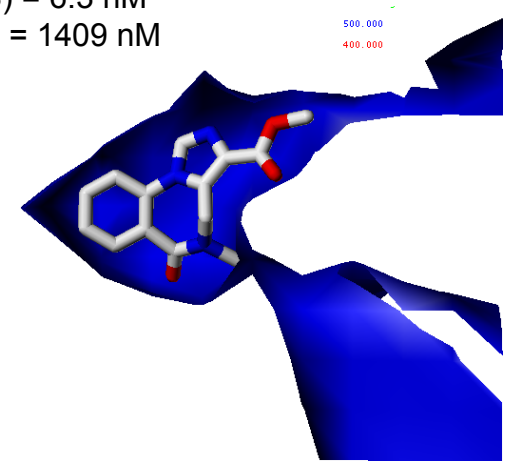
Local electron affinity EA_L
pKi (DS) q^2: **0.643** (6), r^2: 0.902
pKi (DI) q^2: **0.728** (8), r^2: 0.943

$GABA_A$/Benzodiazepine Receptor Dataset $_H$ –
Comparison of Local Ionization Energy (IE_L) Fields

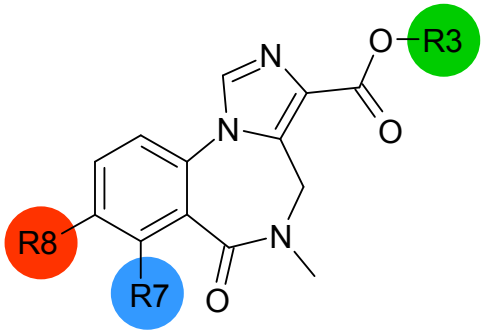
7-Cl, 3-Et
Ki (DS) = 0.2 nM
Ki (DI) = 20 nM



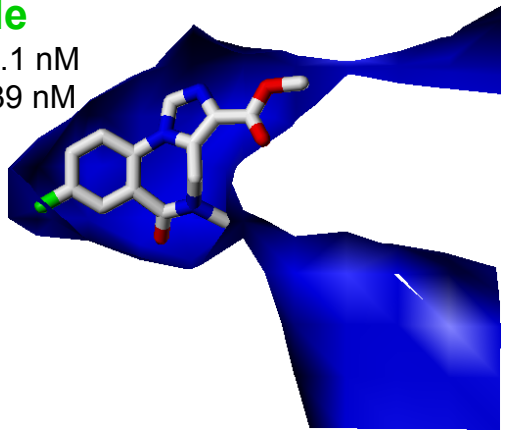
3-Me
Ki (DS) = 6.3 nM
Ki (DI) = 1409 nM



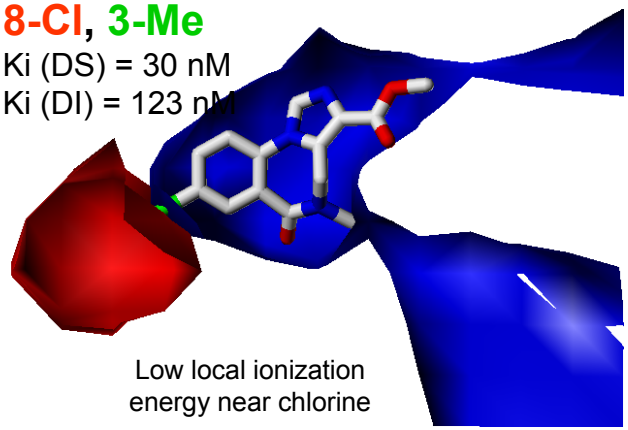
IE_L > 500 kcal/mol (blue)
IE_L < 400 kcal/mol (red)



8-F, 3-Me
Ki (DS) = 3.1 nM
Ki (DI) = 239 nM

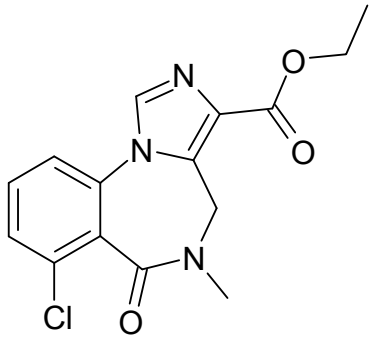


8-Cl, 3-Me
Ki (DS) = 30 nM
Ki (DI) = 123 nM



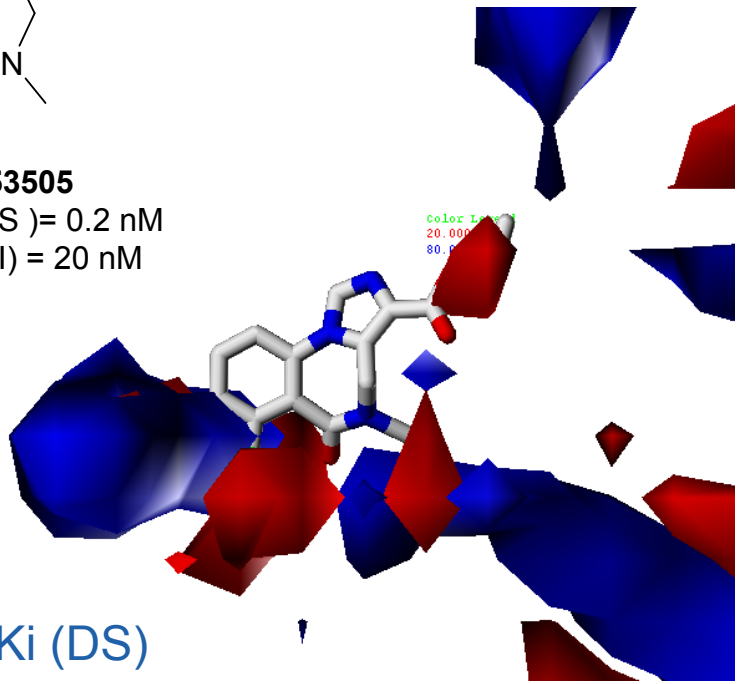
sanofi aventis
Because health matters

$GABA_A$/Benzodiazepine Receptor Dataset$_H$ –
3D-QSAR using Local Ionization Energy (IE_L) Fields

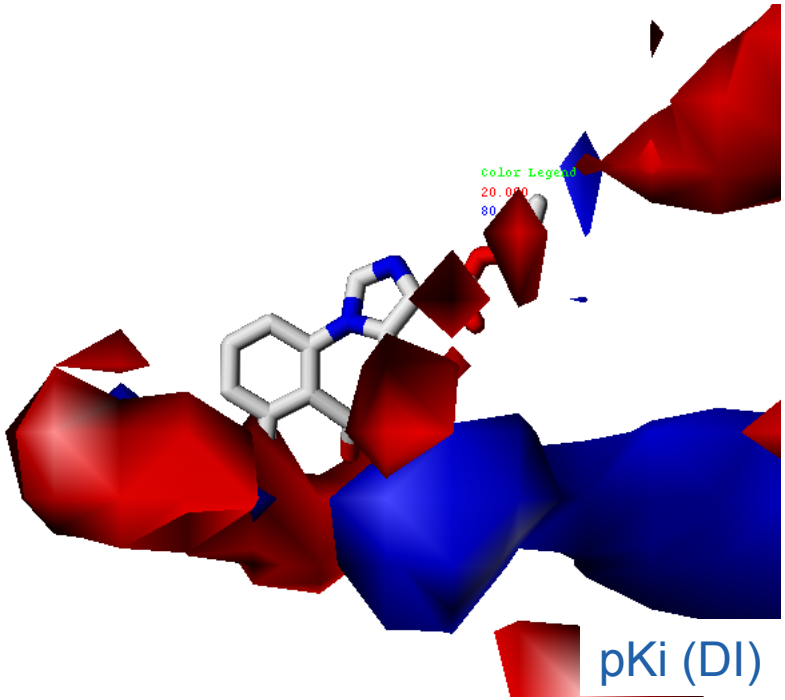


RO-153505
pKi (DS)= 0.2 nM
pKi (DI) = 20 nM

higher IE_L favorable (blue)
lower IE_L favorable (red)



pKi (DS)

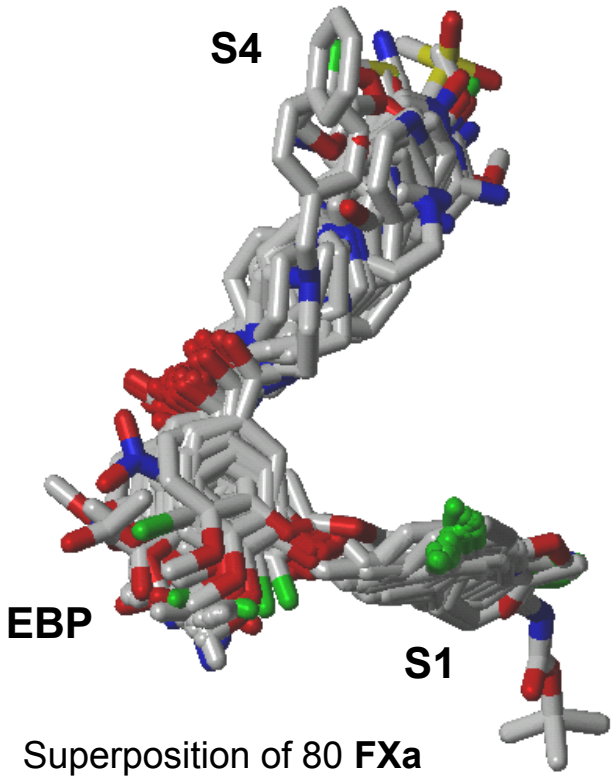


pKi (DI)

Local ionization energy (IE_L):
pKi (DS) q^2: **0.591** (7), r^2: 0.968
pKi (DI) q^2: **0.674** (8), r^2: 0.965

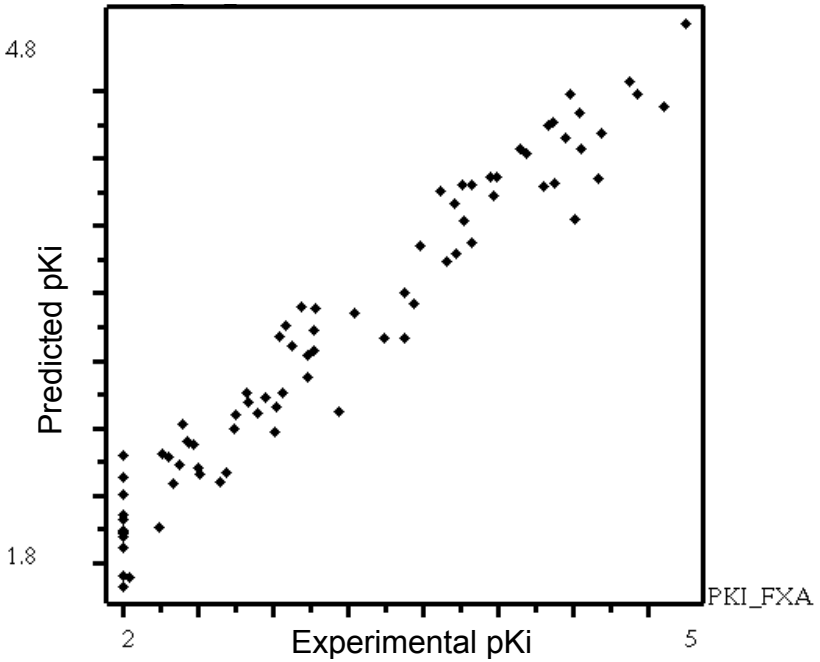
sanofi aventis
Because health matters

FXa Dataset –
3D-QSAR Models for 3-Oxybenzamides



Superposition of 80 **FXa** inhibitors after docking into binding site

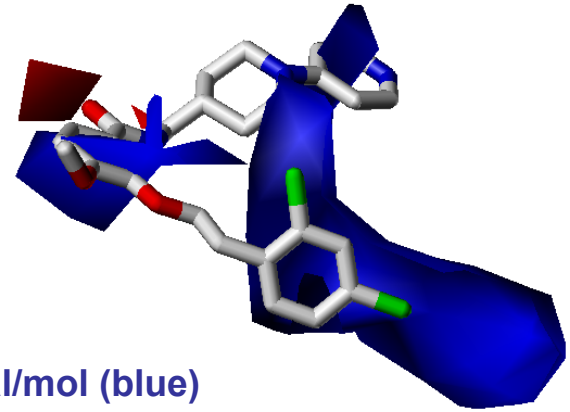
Matter et al., *J. Med. Chem.* **2005,** *48,* 3290-3312.



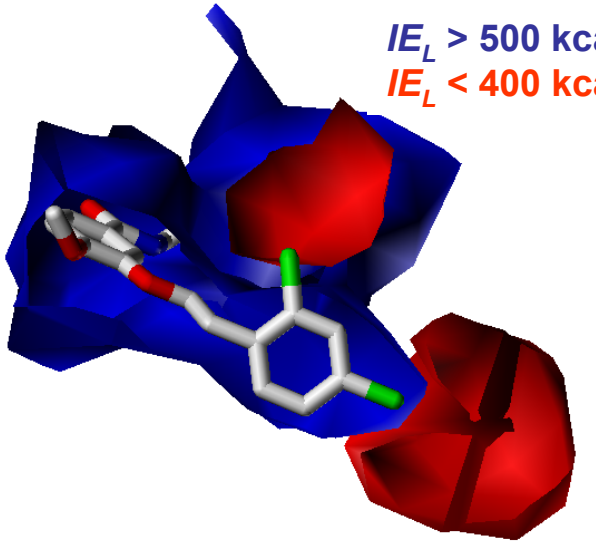
CoMFA (80 cpds):
q^2: **0.741** (6 comp), r^2: 0.947
pred. r^2 0.732 / 0.795 (27 / 72 external cpds)

CoMSIA (80 cpds):
q^2: **0.609** (5 comp), r^2: 0.898
pred. r^2 0.784 / 0.710 (27 / 72 external cpds)

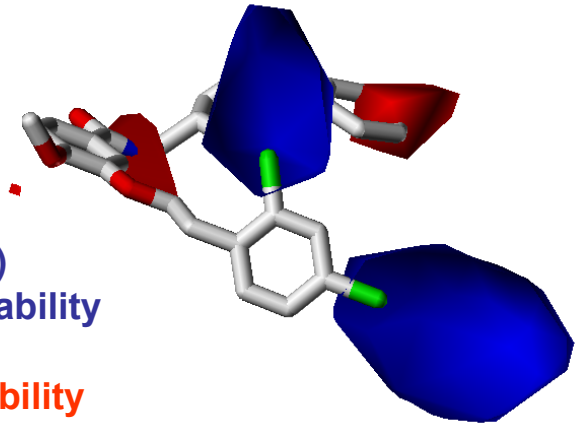
FXa Dataset –
ParaSurf Fields for Oxybenzamide



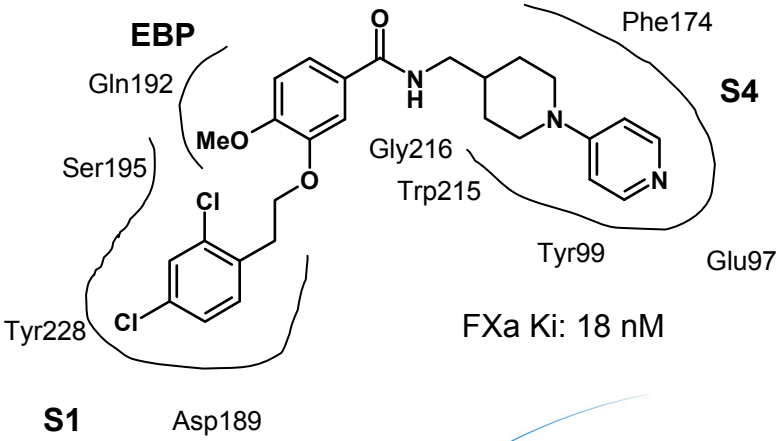
EA_L > -50 kcal/mol (blue)
low local electron affinity
EA_L < -110 kcal/mol (high) (red)
high local electron affinity



IE_L > 500 kcal/mol (blue)
IE_L < 400 kcal/mol (red)



α_L > 0.40 (blue)
high polarizability
α_L < 0.27 (red)
low polarizability

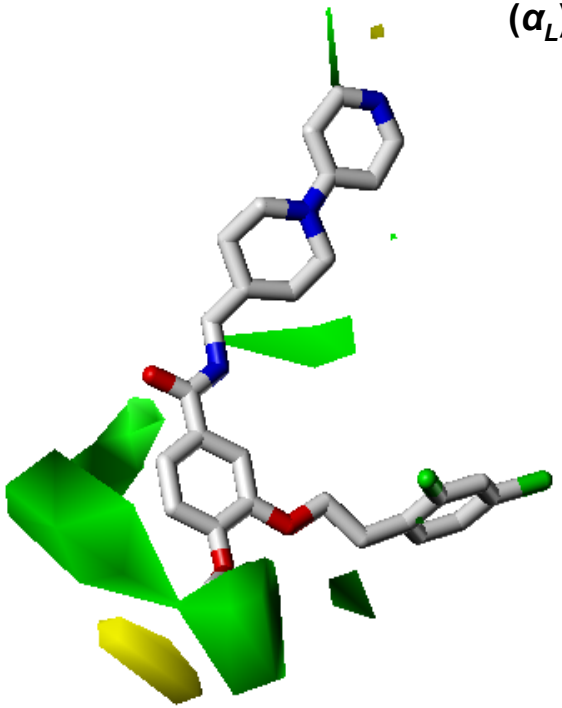


FXa Ki: 18 nM

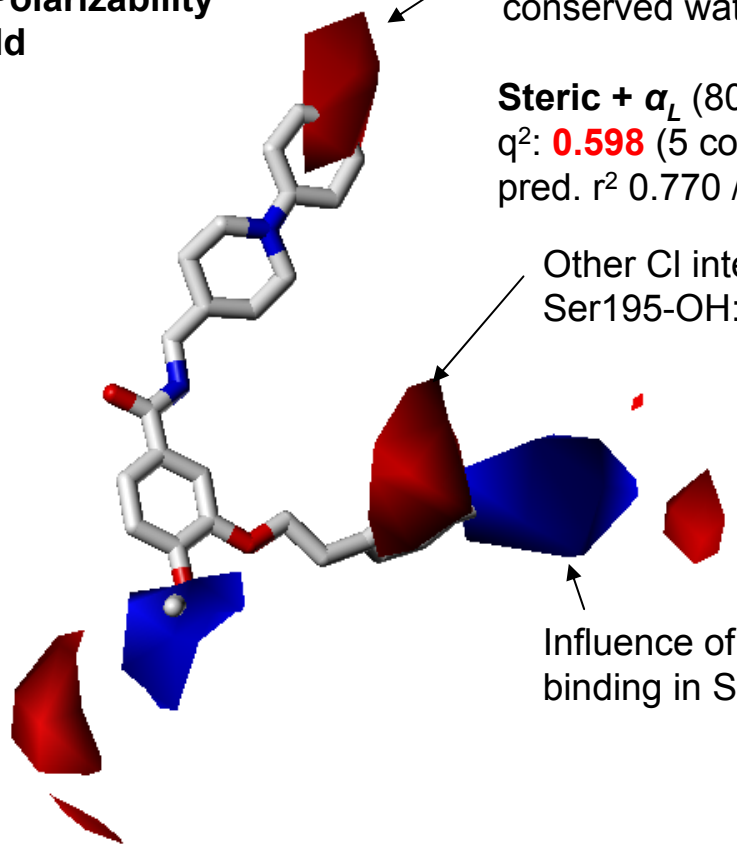
sanofi aventis
Because health matters

FXa Dataset –
Chemical Interpretation of 3D-QSAR Models

CoMFA Steric field



Local Polarizability (α_L) field



Interaction of pyridine N with S4 conserved water: „Hardness“

Steric + α_L (80 cpds):
q^2: **0.598** (5 comp), r^2: 0.827
pred. r^2 0.770 / 0.825 (27/72 cpds)

Other Cl interacts with Ser195-OH: „Hardness“

Influence of polarizability on Cl-binding in S1 („Softness“)

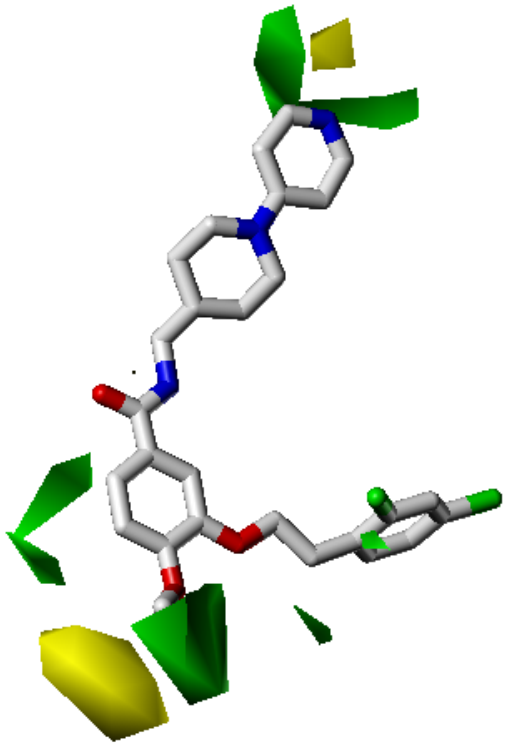
steric bulk favorable (green)
steric bulk detrimental (yellow)

higher α_L (softness) favorable (blue)
lower α_L (hardness) favorable (red)

Matter et al., *J. Med. Chem.* **2005,** *48,* 3290-3312.

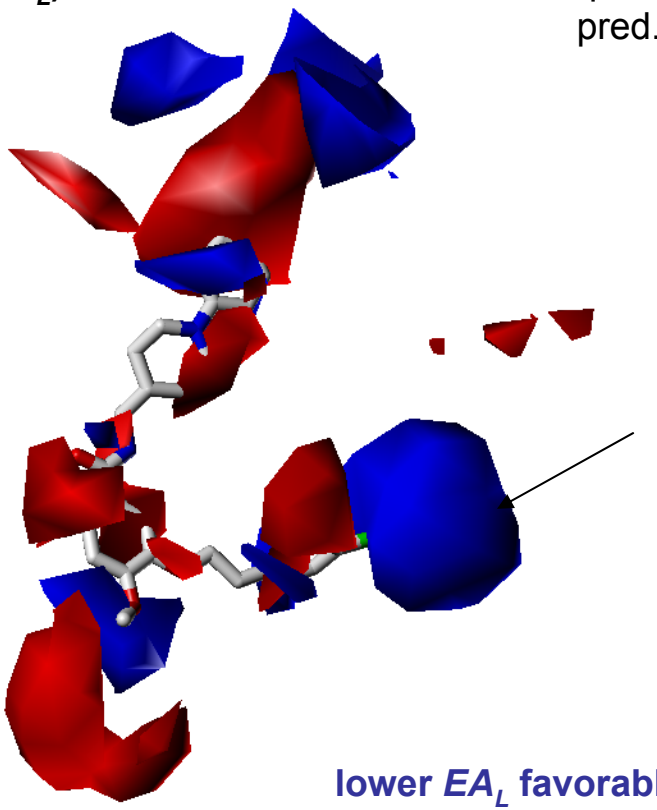
FXa Dataset –
Chemical Interpretation of 3D-QSAR Models

CoMFA steric field



steric bulk favorable (green)
steric bulk detrimental (yellow)

Local electron affinity (EA_L) field



Influence of electron-affinity on Cl-binding by Tyr228

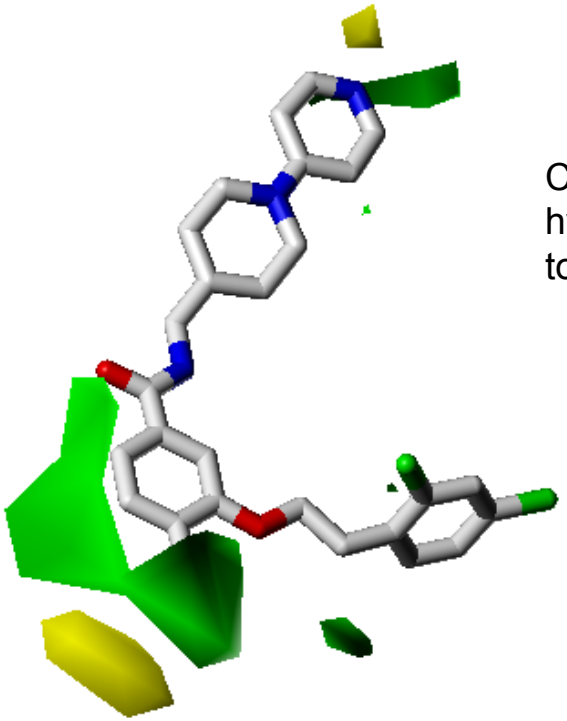
lower EA_L favorable (blue)
higher EA_L favorable (red)

Steric + EA_L (80 cpds):
q^2: **0.609** (6 comp), r^2: 0.910
pred. r^2 0.737 / 0.837 (27/72 cpds)

Matter et al., *J. Med. Chem.* **2005,** *48,* 3290-3312.

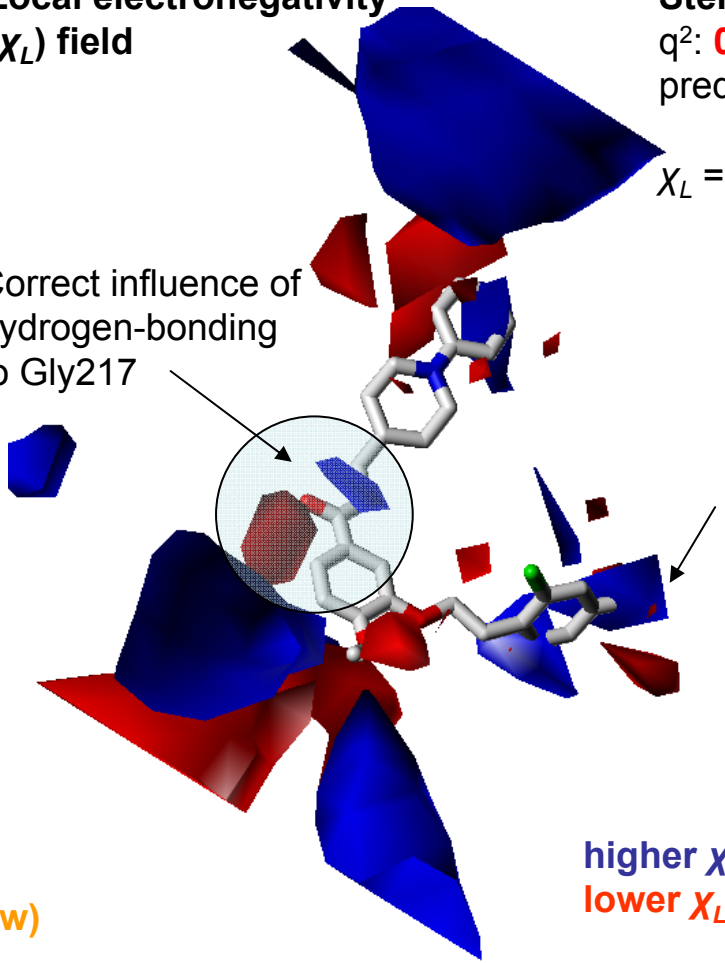
FXa Dataset –
Chemical Interpretation of 3D-QSAR Models

CoMFA steric field



steric bulk favorable (green)
steric bulk detrimental (yellow)

Local electronegativity (χ_L) field



higher χ_L favorable (blue)
lower χ_L favorable (red)

Steric + χ_L (80 cpds):
q^2: **0.638** (4 comp), r^2: 0.853
pred. r^2 0.729 / 0.769 (27/72 cpds)

$\chi_L = (IE_L + EA_L) / 2$

Matter et al., *J. Med. Chem.* **2005,** *48,* 3290-3312.

Conclusions - General

- ParaSurf provides relevant descriptors that are independent of the 2D topology of molecules
- Descriptors are
  - reasonably fast to calculate
  - showing only minor conformational influence
  - Providing a fine differentiated view on a molecule

Conclusions – 2D/3D QSAR

- Parasurf 2D descriptors provided predictive models
  - Standalone performance comparable to MOE 2D Descriptor set
- ParaSurf 2D descriptors contain different information compared to MOE
  - Provide alternative descriptor metric offering novel insights into SAR
- Use of local properties for 3D QSAR provided predictive models
  - Best performance with: Pol, IE and EA
  - Models provide additional insight into the SAR not shown by CoMFA

Acknowledgement

Sanofi Aventis

Karl-Heinz Baringhaus

Andreas Kugelstadt

Cepos InSilico Ltd.

Tim Clark

Ute Seidel

Matthias Hennemann