A.N. Frumkin
Institute of Physical Chemistry and Electrochemistry
Institution of Russian Academy of Sciences

31 Leninsky prospect, Moscow GSP-1, 119071 Russia

Modelling of Hydrogen Bond Thermodynamics

Vitaly Solov'ev

Laboratory of New Physico-Chemical Problems

Institute of Physical Chemistry and Electrochemistry, Moscow

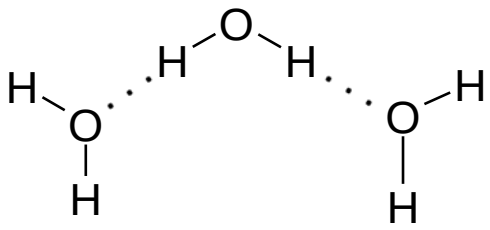
E-mail: solovev-vp@mail.ru

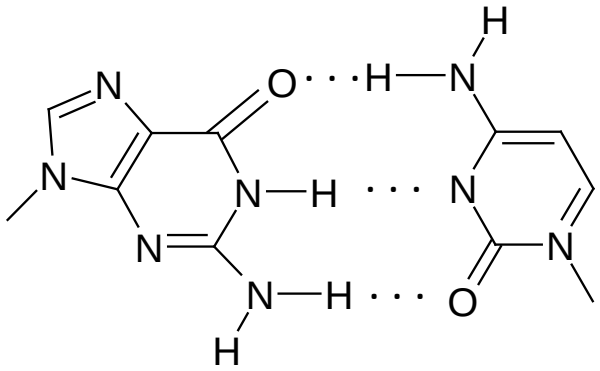
Hydrogen Bonding
Importance

The hydrogen bond is one of the fundamental interactions

—X-H · · · Y—

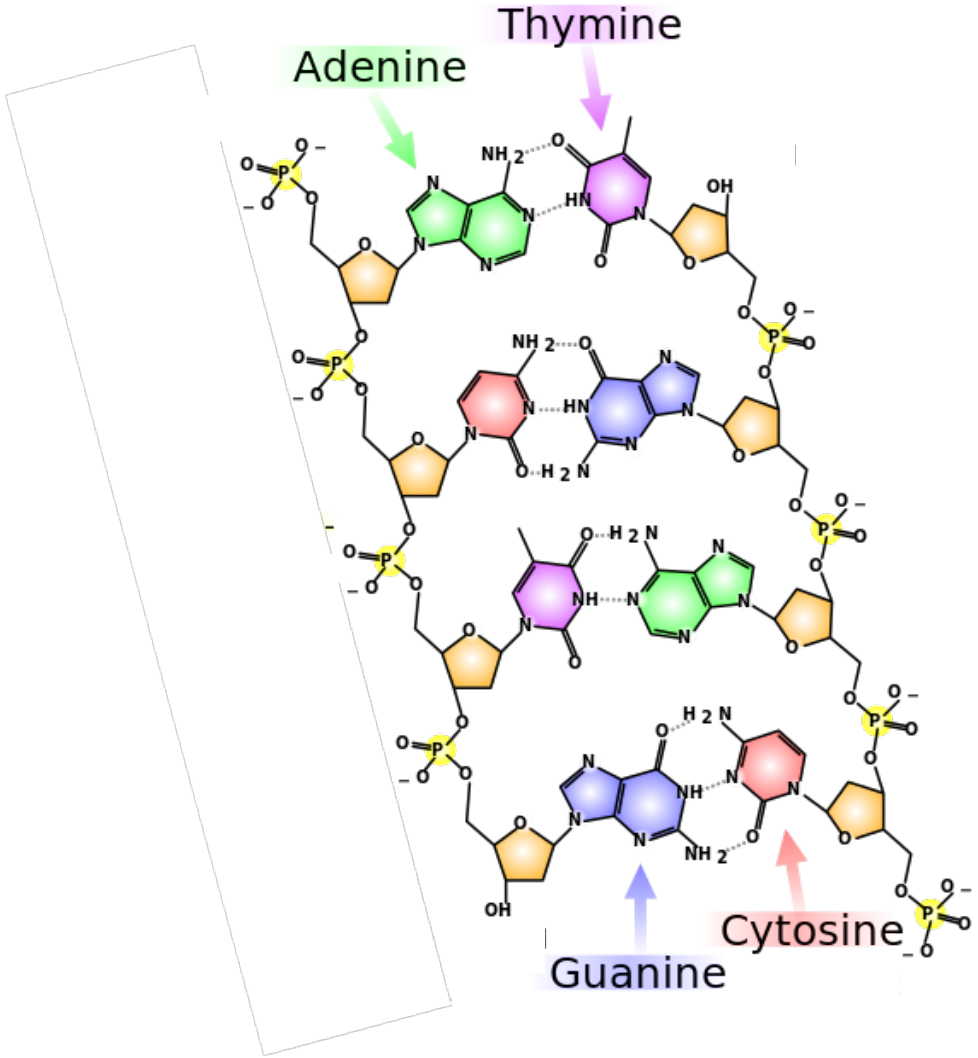
X, Y = O, N, S, F





G C

Hydrogen Bonding
Importance



Thermodynamic Functions of Hydrogen Bonding

Single Hydrogen Bond

$$AH + B = AH\cdots B$$

$K = C_{AH\cdots B}/(C_{AH}C_B)$	*stability constant*
$\Delta G = -RT \ln 10 \log K$	*free energy*
ΔH	*enthalpy*
$\Delta S = (\Delta H - \Delta G) / T$	*entropy*

$\log K$ 1 … 4 $\quad -\Delta H$ 15 … 30 kJ/mol

Modeling and Prediction of H-Bond Thermodynamics

1. Linear free energy relationships (LFER)

2. Quantum chemical (QC) methods

3. Quantitative structure-property relationships (QSPR)

Modeling and Prediction of H-Bond Thermodynamics

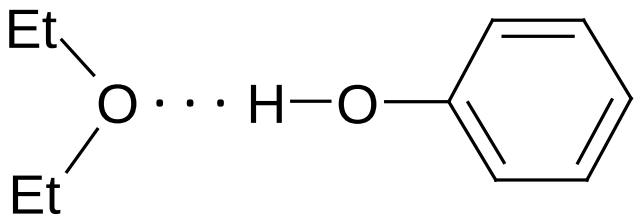
1. Linear free energy relationships

2. Quantum chemical methods

3. Quantitative structure-property relationships

Linear Free Energy Relationships
Modeling and Prediction of H-Bond Thermodynamics

$$\Delta H = k\, h_a\, h_d$$



H-bond acceptor h_a

H-bond donor h_d

1. Sherry A., Purcel K. *J. Am. Chem. Soc.* **1972**, *94*, 1848–1853

2. Iogansen A. *Teor. Eksp. Khim. (Rus.)* **1971**, *7*, 302–311

3. Raevsky O. A., Novikov V. P. *Khim. Farm. Zh. (Rus.)* , **1982**, *16*, 583–586

Linear Free Energy Relationships
Modeling and Prediction of H-Bond Thermodynamics

The 1:1 hydrogen bonding in CCl_4 at 298 K

$$\Delta S = k\, s_a\, s_d$$

$$\Delta G = k\, g_a\, g_d$$

1. Raevsky O. A., Grigor'ev V. Yu., Solov'ev V. P., Martynov I. V. *Doklady Akademii Nauk SSSR (Rus),* **1988**, *299*, 1433-1438

2. Raevsky O. A., Grigor'ev V. Yu., Solov'ev V. P. *Chim.-Pharm. J. (Rus.)*, **1989**, 1294-1300

Linear Free Energy Relationships
Modeling and Prediction of H-Bond Thermodynamics

The 1:1 hydrogen bonding in CCl_4 at 298 K

$$logK = m\ \alpha\ \beta + c$$

α is parameter of H-bond donor

β is parameter of H-bond acceptor

Abraham M. H., Grellier P. L., Prior D. V., Taft R. W., Morris J. J., Taylor P. J., Laurence C., Berthelot M., Doherty R. M. et al.
J. Am. Chem. Soc., **1988**, *110* , 8534–8536

Linear Free Energy Relationships

Modeling and Prediction of H-Bond Thermodynamics

The 1:1 hydrogen bonding in CCl_4 at 298 K

$$\log K_{calc} = 7.35\,\alpha\,\beta - 1.09 \qquad n = 1312,\ R^2 = 0.991,\ s = 0.09$$

Abraham M. H. et al. *J. Am. Chem. Soc.*, **1988**, *110* , 8534–8536

$$\Delta H_{calc} = 0.5 + 0.99\ \Delta H_{exp} \qquad n = 2787,\ R^2 = 0.941,\ s = 2.4$$

$$\Delta G_{calc} = 0.1 + 1.04\ \Delta G_{exp} \qquad n = 3301,\ R^2 = 0.982,\ s = 1.12$$

Raevsky O. A. *J. Phys. Chem.*, **1997**, *10*, 405-413

Modeling and Prediction of H-Bond Thermodynamics

1. Linear free energy relationships

2. Quantum chemical methods

3. Quantitative structure-property relationships

Quantum chemical methods

Modeling and Prediction of H-Bond Thermodynamics

Hydrogen-Bonding of Nitrogen Bases: A Theoretical Evaluation of Stability Constant and Enthalpy

Besseau F., Graton J., Berthelot M. *Chem. Eur. J.* **2008**, *14*, 10656–10669

Hydrogen Bonding of nitrogen bases with 4-F-Phenol

Data Set

The 1:1 hydrogen bonding in CCl_4 at 298 K

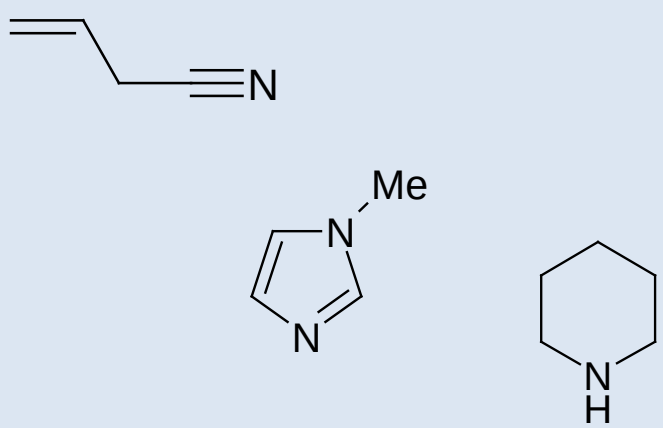
$\log K_{exp}$ ΔH_{exp}

H-bond acceptors: the sp, sp^2 and sp^3 nitrogen bases

H-bond donor: 4-F-phenol

Training set: 59 nitrogen bases

Test set: 99 nitrogen bases



Besseau F., Graton J., Berthelot M. *Chem. Eur. J.* **2008**, *14*, 10656–10669

Hydrogen Bonding of nitrogen bases with 4-F-Phenol

Methods

Density Functional Theory (DFT)

Gaussian 03

Spartan

DFT B3LYP functional for Geometry optimization

6-31+G(d,p) Pople basis set

Calculations:

$V_{s,min}$ - minimal electrostatic potential around nitrogen atom

D_0^{HF} - interaction energy with HF

ΔH^{HF} - theoretical enthalpy of the complexation with HF

Hydrogen Bonding of nitrogen bases with 4-F-Phenol

Results

Training:

$$\Delta H = 0.616\ \Delta H^{HF} + 3.400$$

$$R^2 = 0.984,\ s = 0.85\ \text{kJ/mol},\ n = 59$$

$$\log K = -0.0612\ D_0^{HF} - 0.0102\ V_{s,min} - 2.829$$

$$R^2 = 0.986,\ s = 0.07,\ n = 59$$

Prediction:

$\log K$ $\quad R_{det}^2 = 0.958,\ RMSE = 0.16,\ n = 90$

ΔH $\quad R_{det}^2 = 0.870,\ RMSE = 1.8,\ n = 29$

Besseau F., Graton J., Berthelot M. *Chem. Eur. J.* **2008**, *14*, 10656–10669

Quantum chemical methods

Modeling of H-Bond Thermodynamics

Interpretation of Experimental Hydrogen-bond Enthalpies and Entropies from COSMO Polarization Charge Densities

Klamt A., Reinisch J., Eckert F., Graton J., Le Questel J. *Phys. Chem. Chem. Phys.*, **2013**, *15*, 7147

Hydrogen Bonding of organic bases with 4-F-Phenol

Data Set

The 1:1 hydrogen bonding in CCl_4 or C_2Cl_4 at 298 K

ΔG_{exp} ΔH_{exp}

H-bond acceptors: 309 organic bases

H-bond donor: 4-F-phenol

Klamt A., Reinisch J., Eckert F., Graton J., Le Questel J. *Phys. Chem. Chem. Phys.*, **2013**, *15*, 7147

Hydrogen Bonding of organic bases with 4-F-Phenol

Methods

Density Functional Theory (DFT) and
COSMO continuum solvation model
TURBOMOLE program

Calculations:

$\langle \sigma_{acc} \rangle_{conf}$ – mean value of the COSMO polarization charge density

Ω_{HB} – H-Bond partition function

Hydrogen Bonding of organic bases with 4-F-Phenol

Results

Training:

$$\Delta H = -17.9 \langle \sigma_{acc} \rangle_{conf} + 5.4$$

$R^2 = 0.946$, $s = 2.2$ kJ/mol, $n = 307$

$$\Delta G = -14.1 \langle \sigma_{acc} \rangle_{conf} + 10.1 - RT \ln \Omega_{HB}$$

$R^2 = 0.924$, $s = 2.1$ kJ/mol, $n = 307$

Klamt A., Reinisch J., Eckert F., Graton J., Le Questel J. *Phys. Chem. Chem. Phys.*, **2013**, *15*, 7147

Quantitative Structure-Property Relationships

Prediction of H-Bond Thermodynamics

1. Linear free energy relationships

2. Quantum chemical methods

3. Quantitative structure-property relationships

Quantitative Structure-Property Relationships

Modeling and Prediction of H-Bond Thermodynamics

Hydrogen Bonding of Various Organic Compounds with Phenols

Varnek A., Fourches D., Hoonakker F., Solov'ev V. *J. Comp.-Aided Mol. Design*, **2005**, *19*, 693-703

Hydrogen Bonding of Organic Acceptors with Phenols

Data Set

The 1:1 Complexation in CCl_4 at 298 K

ΔH_{exp} ΔG_{exp}

365 H-bond acceptors: organic compounds

46 H-bond donors: phenols

Training set: 292 H-Bond complexes

Test set: 73 H-Bond complexes

Hydrogen Bonding of Organic Acceptors with Phenols

Data Set: H-Bond Acceptor Groups

acceptor group	group name
>C=O	carbonyl oxygen
>N-	amine nitrogen
-O-	ether oxygen
>C=N-	imine nitrogen
$>N_{ar}$	aromatic nitrogen
-C≡N	nitrile nitrogen
-S-	sulfur of sulfide
>C=S	sulfur of thiocarbonyl group
>C=Se	selen of selenocarbonyl group
>S=O	oxygen of sulfinyl group
O=P←	oxygen of phosphoryl group
S=P←	sulfur of thiophosphoryl group
Se=P←	selen of selenophosphoryl group

Hydrogen Bonding of Organic Acceptors with Phenols
QSPR methods

The ISIDA/QSPR software:

Multiple Linear Regression method: *ISIDA MLR*

The *EdChemS* editor of 2D structures

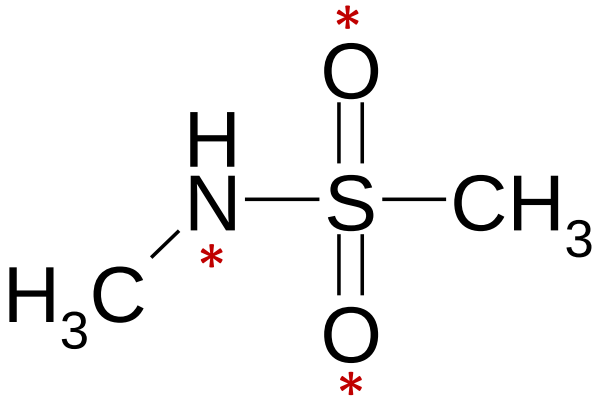
The SDF manager: *EdiSDF*

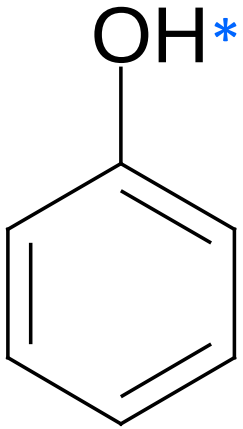
URLs:
http://infochim.u-strasbg.fr/spip.php?rubrique53
http://vpsolovev.ru/programs/

Varnek A., Fourches D., Hoonakker F., Solov'ev V. *J. Comp.-Aided Mol. Design*, **2005**, *19*, 693-703

Hydrogen Bonding of Organic Acceptors with Phenols

Descriptors: Substructural Molecular Fragments





H-bond acceptor

H-bond donor

O=S
C-O=S
C-N-S=O
…

H-O
H-O-C
H-O-C=C
…

counts of shortest topological paths including labeled atoms

Hydrogen Bonding of Organic Acceptors with Phenols
Results

Predictions, TEST SETs, each 5th complex:

$\Delta G_{pred} = 0.10 + 1.00\Delta G_{exp}$ $\quad R^2 = 0.916$, $s = 1.7$, $n = 66$

$\Delta H_{pred} = 2.3 + 0.90\Delta H_{exp}$ $\quad R^2 = 0.868$, $s = 2.2$, $n = 66$

Varnek A., Fourches D., Hoonakker F., Solov'ev V. *J. Comp.-Aided Mol. Design*, **2005**, *19*, 693-703

Individual Hydrogen-Bond Strength QSPR Modelling with ISIDA Local Descriptors: a Step Towards Polyfunctional Molecules

Ruggiu F., Solov'ev V., Marcou G., Horvath D., Graton J., Le Questel J.-Y., Varnek A.
Mol. Inf. **2014**, *33*, 477–487

Hydrogen Bonding of Organic Acceptors with 4-F-Phenol

Data Set

The 1:1 Complexation in CCl_4 at 298 K

$\log K_{exp}$

537 mono- and bi-functional H-bond acceptors: organic compounds

H-bond donor: 4-F-phenol

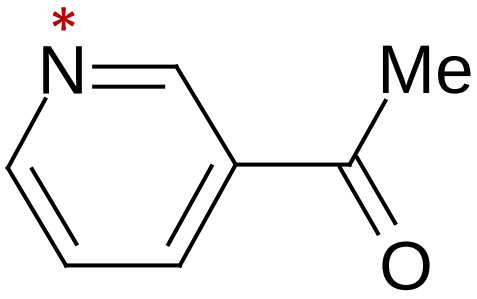
Training and test sets: 5-fold cross validation

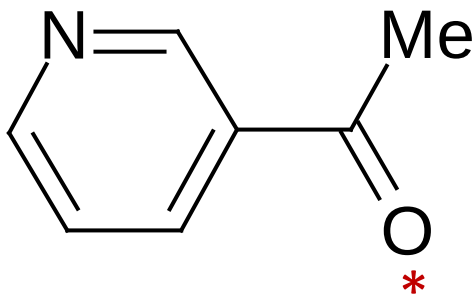
External test set: 451 H-Bond acceptors

Hydrogen Bonding of Organic Acceptors with 4-F-Phenol

Data Set

$\log K_{exp}$ = 1.39





$\log K_{exp}$ = 0.9

Binding atom is established by Fourier transform IR spectroscopy

Database: Laurence C. et al. *J. Med. Chem.* **2009**, *52*, 4073–4086

Hydrogen Bonding of Organic Acceptors with 4-F-Phenol

Data Set: H-Bond Acceptor Groups

acceptor group	group name
>C=O	carbonyl oxygen
>N-	amine nitrogen
-O-	ether oxygen
>C=N-	imine nitrogen
$>N_{ar}$	aromatic nitrogen
-C≡N	nitrile nitrogen
-S-	sulfur of sulfide
>C=S	sulfur of thiocarbonyl group
>S=O	oxygen of sulfinyl group
O=P<	oxygen of phosphoryl group
$-NO_2$	oxygens of nitro group
S=P<	sulfur of thiophosphoryl group

Hydrogen Bonding of Organic Acceptors with Phenols

Descriptors

ISIDA Fragment Descriptors

Descriptors generated by Molecular Operating Environment (MOE)

Ruggiu F., Solov'ev V., Marcou G., Horvath D., Graton J., Le Questel J.-Y., Varnek A. *Mol. Inf.* **2014**, *33*, 477–487

Hydrogen Bonding of Organic Acceptors with Phenols

Descriptors

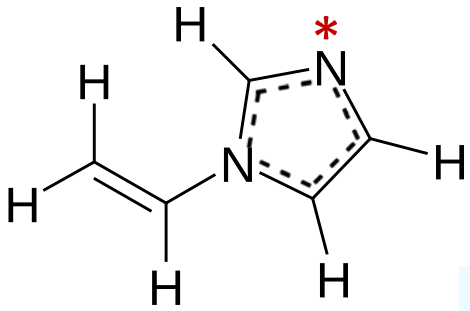
Four Types of ISIDA Fragment Descriptors:

- MA0 – no marked atoms, all generated fragments
- MA1 – only fragments beginning with marked atom
- MA2 – only fragments containing marked atom
- MA3 – all fragments with a special flag on marked atom

Hydrogen Bonding of Organic Acceptors with Phenols

Descriptors

MA3 fragments with a special flag on marked atom



H-C-N*

H-C-N

H-C-N:::C:::N*

...

Hydrogen Bonding of Organic Acceptors with Phenols
Methods

Support vector machines:

The LibSVM package

Ensemble Multiple Linear Regression method:

The ISIDA/QSPR software

URLs:
http://infochim.u-strasbg.fr/spip.php?rubrique53
http://vpsolovev.ru/programs/

Ruggiu F., Solov'ev V., Marcou G., Horvath D., Graton J., Le Questel J.-Y., Varnek A. *Mol. Inf.* **2014**, *33*, 477–487

Hydrogen Bonding of Organic Acceptors with Phenols
Results

Prediction of log*K*, TEST SETs:

5-fold cross validation

$R_{det}^2 = 0.90$, $RMSE = 0.24$, $n = 542$

External test set

$R_{det}^2 = 0.91$, $RMSE = 0.26$, $n = 75$

URL:

The SVM consensus model is publically available:
http://infochim.u-strasbg.fr/webserv/VSEngine.html

Ruggiu F., Solov'ev V., Marcou G., Horvath D., Graton J., Le Questel J.-Y., Varnek A. *Mol. Inf.* **2014**, *33*, 477–487

CONCLUSIONS

- LFER, Quantum chemical and QSPR methods provide predictions of thermodynamic functions of hydrogen bonding with accuracy compatible with the systematic errors in experimental data.

- Known methods are developed for predictions of thermodynamic functions of the complexes with single hydrogen bond.

- The models based on the FTIR data and new ISIDA descriptors are able to predict the H-bond strength (basicity) of the O, N and S atoms in different hybridization states as HBAs of the polyfunctional molecules.

ACKNOWLEDGMENTS

Prof. Alexandre Varnek

Prof. Aslan Tsivadze

Prof. Igor Antipin

Dr. Timur Madzhidov

Dr. Olga Klimchuk

Organizers of The 1st International School-Seminar
“From Empirical to Predictive Chemistry”