Supporting Information

© Copyright Wiley-VCH Verlag GmbH & Co. KGaA, 69451 Weinheim, 2018

Bimolecular Nucleophilic Substitution Reactions: Predictive Models for Rate Constants and Molecular Reaction Pairs Analysis

Timur Gimadiev, Timur Madzhidov, Igor Tetko, Ramil Nugmanov, Iury Casciuc, Olga Klimchuk, Andrey Bodrov, Pavel Polishchuk, Igor Antipin, and Alexandre Varnek*

T.R. Gimadiev et al. « Bimolecular nucleophilic substitution reactions: predictive models for rate constants and Molecular Reaction Pairs analysis »

Supporting information

Model	Fragmentation type	Kernel type	RMSE, $(\log k)$	Q ²
Ν				
1			0.00.0000	0.00.0.004
1	Sequences of atoms and	Gaussian	0.38 ± 0.008	0.89 ± 0.004
	bonds; length min=1, max=6,			
	explicit charges			
2	Atom contored fragments	Coussian	0.4+0.011	0.88+0.007
2	Atom centered magnetics	Gaussian	0.4± 0.011	0.00± 0.007
	based on sequences of bonds;			
	length min=1, max=3; explicit			
	charges			
3	Sequences of stoms and	Gaussian	0.39 ± 0.006	0.88+0.006
5	bender length min 2 mers (Gaussian	0.39± 0.000	0.88± 0.000
	bonds; length min=2, max=6,			
	explicit charges			
4	Augmented atoms; sphere	Gaussian	0.42 ± 0.014	0.87 ± 0.009
	radius min=1 max=6: explicit			
	charges			
	charges			
5	Sequences of atoms and	Gaussian	0.4 ± 0.007	0.88 ± 0.005
	bonds; length min=2, max=6;			
6	Augmented atoms; sphere	Polynomial	0.4 ± 0.016	0.89 ± 0.009
	radius min=1, max=7;			
7	Sequences of stoms and	Polynomial	0.41 ± 0.014	0.87+0.008
/	sequences of atoms and	Forynonnai	0.41 ± 0.014	0.87±0.008
	bonds, min=3, max=4;			
8	Sequences of atoms and	Gaussian	0.42 ± 0.008	0.87 ± 0.005
	bonds; length min=2, max=6,			
	explicit charges: all possible			
	explicit enarges, an possible			

1. Description of 10 best individual models selected for consensus calculations

	sequences;			
9	Sequences of atoms and bonds; length min=1, max=7, explicit charges	Gaussian	0.41± 0.009	0.87± 0.005
10	Sequences of atoms and bonds; length min=1, max=6, explicit charges	Polynomial	0.42± 0.008	0.87± 0.005

2. Description of Web interface for models' application.

For model publication, in-house client-server application has been developed. A client side description is described below.

The model's application to a query reaction proceeds in 3 steps: structure preparation, setting the condition (solvent, temperature and pressure) and analysis of the results. The first page of predictor contains User manual and tools for interactive building and uploading of chemical structures (the **Add** task, Figure S1).



Figure S1. Interactive preparation of query reaction with Marvin Web.

Then, the one should click the button "**Validate**" (Figure S2) which launches procedures of structure verification, structure standardization and atom-to-atom mapping.

1 Start Choosing a structure for searching	2 Validate 3 Result The choice of model and the correctness of the structure Search results
1. Upload file	+ Add structure > Validate
$\cdot, \bigcup_{i=1}^{n} \to \bigcup_{i=1}^{n} \cdot \bigcup_{i=1}^{n}$	

Figure S2. Structure validation step.

At the next stage, the user is invited to entry the parameters needed to perform the calculations for a query reaction: a particular model (in our case, this is the model for the rate constant of SN2 reaction), temperature, pressure and solvent (see Figure S3)

CIMM News About Us Education +	Search Modeling - D8 Form Timur Madzhidov -
Start Choosing a structure for searching	Validate 3 Result The choice of model and the correctness of the structure Search results
< Back	Actions V Show result(s)
	Models: Please select a model
$c_{i_{0}} \underbrace{ \begin{pmatrix} c_{i_{0}} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ $	Temperature : CFM 200 273 298 400 298
m	Pressure (atm):
	Solvents :
	Please select

CIMM News About Us Education -	Search Modeling - DB Form Timur Madzhidov -
	* Models :
	Substitution nucleophilic bimolecular ×
	Iemperature:
	200 273 298 400
	Pressure (stm) ·
	0 1 3 6
	Solvents:
	n.n-dimethylformamide ×
	nitromethane
	nitrobenzene toluene
	heavy water
	tetrahydrothiophene 1.1-dioxide
	hexamethylphosphoramide *
CIMM News About Us Education +	Search Modeling + DB Form Timur Madzhidov +
CIMM News About Us Education •	Search Modeling + DB Form Timur Madzhidov +
CIMM News About Us Education - Start Choosing a structure for searching	Search Modeling - DB Form Timur Madzhidov - Validate
CIMM News About Us Education -	Search Modeling - DB Form Timur Madzhidov - Validate
CIMM News About Us Education - Start Choosing a structure for searching < Back	Search Modeling - DB Form Timur Madzhidov - Validate
CIMM News About Us Education -	Search Modeling • DB Form Timur Madzhidov • Validate
CIMM News About Us Education -	Search Modeling • DB Form Timur Madzhidov • Validate
CIMM News About Us Education -	Search Modeling • DB Form Timur Madzhidov • Validate 3 Result The choice of model and the correctness of the structure 3 Result Actions > Show result(s) • Models :
CIMM News About Us Education -	Search Modeling • DB Form Timur Madzhidov • Validate 3 Result The choice of model and the correctness of the structure 3 Result Actions • > Show result(s) • Models :
CIMM News About Us Education -	Search Modeling • DB Form Timur Madzhidov • Validate 3 Result The choice of model and the correctness of the structure 3 Result Actions • > Show result(s)
CIMM News About Us Education -	Search Modeling • DB Form Timur Madzhidov • Validate 3 Result The choice of model and the correctness of the structure 3 Result Actions • > Show result(s)
CIMM News About Us Education - Start Choosing a structure for searching < Back $c_{H_{SR}} \leftarrow c_{H_{SR}}^{H_{SR}} + c_{H_{SR}}^{H_{R}} c_{H_{SR}} + c_{H_{SR}}^{C_{H_{SR}}} + c_{H_{SR}}^{C_{H_{SR}$	Search Modeling DB Form Timur Madzhidov • Validate 3 Result The choice of model and the correctness of the structure 3 Result Actions > Show result(s) • Models : Substitution nucleophilic bimolecular × Substitution nucleophilic bimolecular Tatomers equilibria prediction Elimination bimolecular OptiesAider Cycloaddition based on CGR (log K)
CIMM News About Us Education - Start Choosing a structure for searching Choosing	Search Modeling DB Form Timur Madzhidov • Validate 3 Result The choice of model and the correctness of the structure 3 Result Actions > Show result(s) • Models : Substitution nucleophilic bimolecular × Substitution nucleophilic bimolecular Diels-Alder Cycloaddrition Lga A
CIMM News About Us Education + Image: Start	Search Modeling • DB Form Timur Madzhidov • Validate 3 Result The choice of model and the correctness of the structure 3 Result Actions • • Show result(s) • Models: • Substitution nucleophilic bimolecular × Substitution nucleophilic bimolecular • • Imination bimolecular • • Diels-Alder Cycloaddition based on CGR (log K) Cycloaddition (Lg A) Tautomeric equilibrium constant prediction • •
CIMM News About Us Education + Image: Start	Search Modeling • DB Form Timur Madzhidov • Validate 3 Result The choice of model and the correctness of the structure 3 Result Actions • • Show result(s) • Models: • Substitution nucleophilic bimolecular × Substitution nucleophilic bimolecular • Tautomers equilibria prediction Elimination bimolecular Diels-Alder Cycloaddition (Lg A) Tautomeric equilibrium constant prediction • Cycloaddition (pre-exponential factor) •
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	Search Modeling DB Form Timur Madzhidov • Validate 3 Result The choice of model and the correctness of the structure 3 Result Actions > Show result(s) • Models : • Substitution nucleophilic bimolecular × Substitution nucleophilic bimolecular • Tautomers equilibria prediction • Elimination bimolecular • Diels-Alder • Cycloaddition based on CGR (log K) • Cycloaddition (pre-exponential factor) • Solvents: •

Figure S3. Setting modeling parameters for a query reaction: temperature and pressure (*top*), solvent (*middle*) and a particular model (*bottom*)

Clicking the "**Show results**" button launches the calculations. This results in displaying the predicted log*k* value and the trustworthy parameter ("optimal", good", "poor") in a separate window (see Figure S4).

CIMM News About Us Education -	Search Modeling + D8 Form Timur Madzhidov +
Choosing a structure for searching	The choice of model and the correctness of the structure Search results
< Back	D Save
1	Temperature(K): 298
	Pressure(atm): 1
	Additives:
H ₃ C ₁₀ CH ₃	n.n-dimethylformamide : 1
$c_n \leftarrow c_n + \bigcap_{n=1}^{n} \cdots + \bigcap_{n=1}^{n} + c_n$	Substitution nucleophilic bimolecular
	Predicted ± sigma: -7.35 ± 0.44
	Trust of prediction: Optimal
	Distrust reason:
© Kazan Chemoinformatics and Molecular Modeling Laboratory 20	18

Figure S4. Results of calculations.