Molecular Descriptors Family on Structure-Activity and Structure-Property Relationships: Results

Lorentz JÄNTSCHI & Sorana BOLBOACĂ

Technical University of Cluj-Napoca, Romania,
http://lori.academicdirect.org
&
„Iuliu Hatieganu“ University of Medicine and Pharmacy, Romania
http://sorana.academicdirect.ro

Purpose

- **To present the results obtained by utilization of an original approach called Molecular Descriptors Family (MDF) on Structure-Property (SPR) and Structure-Activity Relationships (SAR) applied on different classes of chemical compounds and its usefulness as precursors of models elaboration of new chemical compounds with better properties and/or activities.**

Methodology

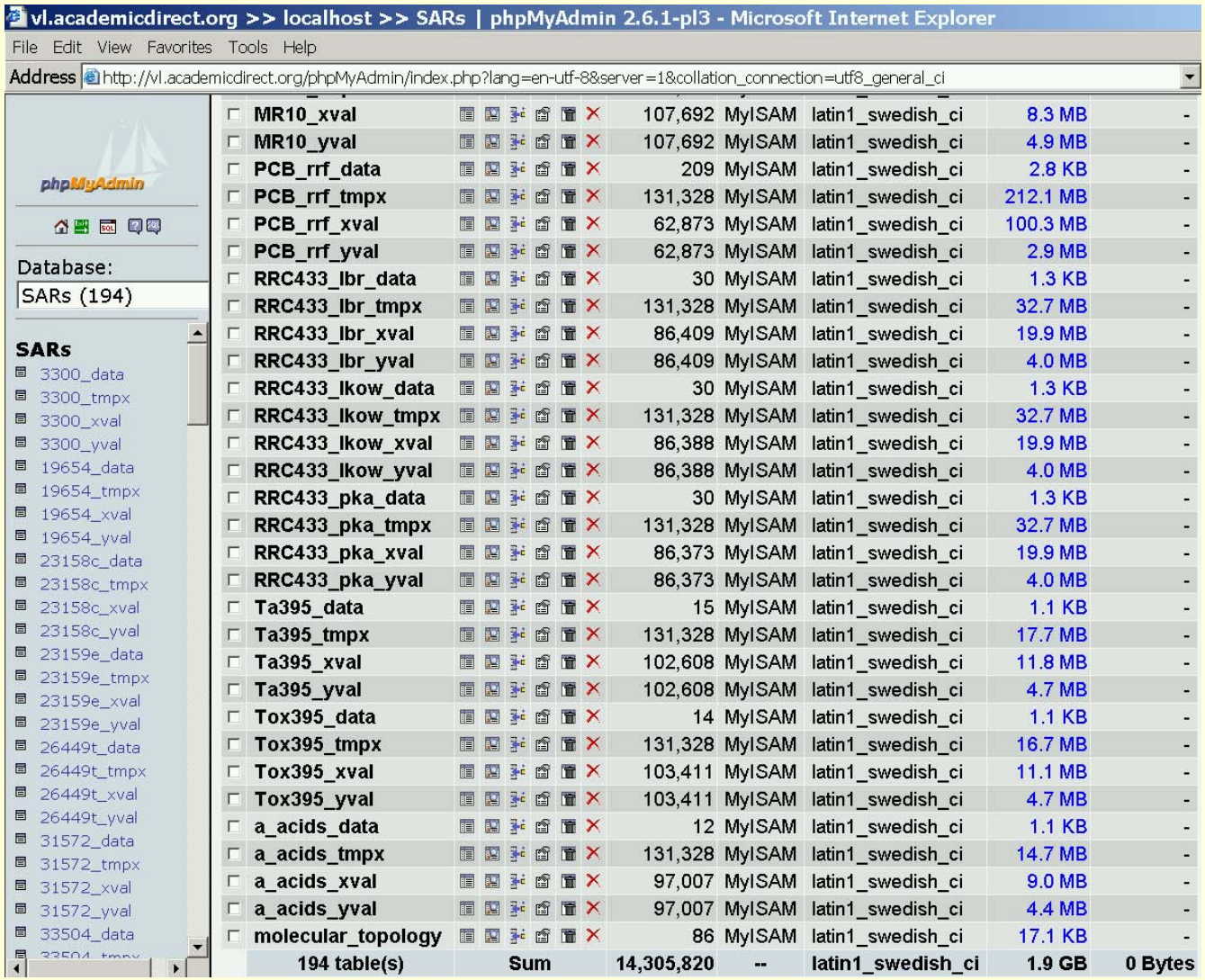
Molecular Descriptors Family

- **Preparing chemical compounds for molecular modeling**
- **Generating the molecular descriptors family**
- **Finding the MDF SPR/SAR models**
- **Validating the MDF SPR/SAR models**
- **Comparing the MDF SPR/SAR models with previous reported model(s)**

Materials

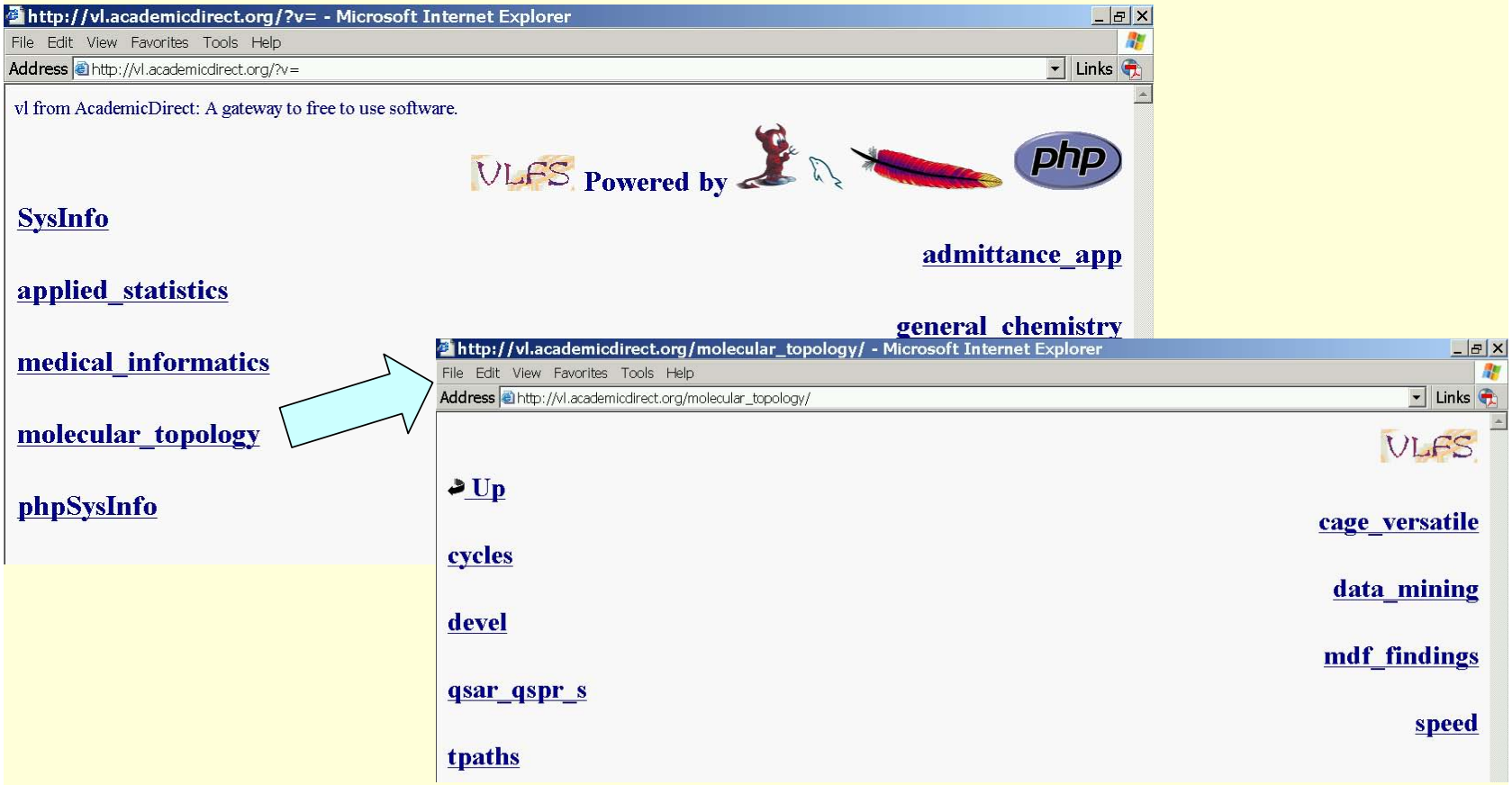
Set name	Observed/Measured Property/Activity
IChr	retention chromatography index
PCB_rrf	relative response factor
23159	octanol/water partition coefficients
23159e	octanol/water partition coefficients
PCB_lkow	octanol/water partition coefficient
36638	water activated carbon adsorption
MR10	molar refraction
Ta395	cytotoxycity
52730	toxicity
Tox395	mutagenicity
41521	insecticidal activity
Triazines	herbicidal activity
52344	antioxidant efficacy
26449	antituberculotic activity
23151	antimalarial activity
22583	anti-HIV-1 potencies

Results: database

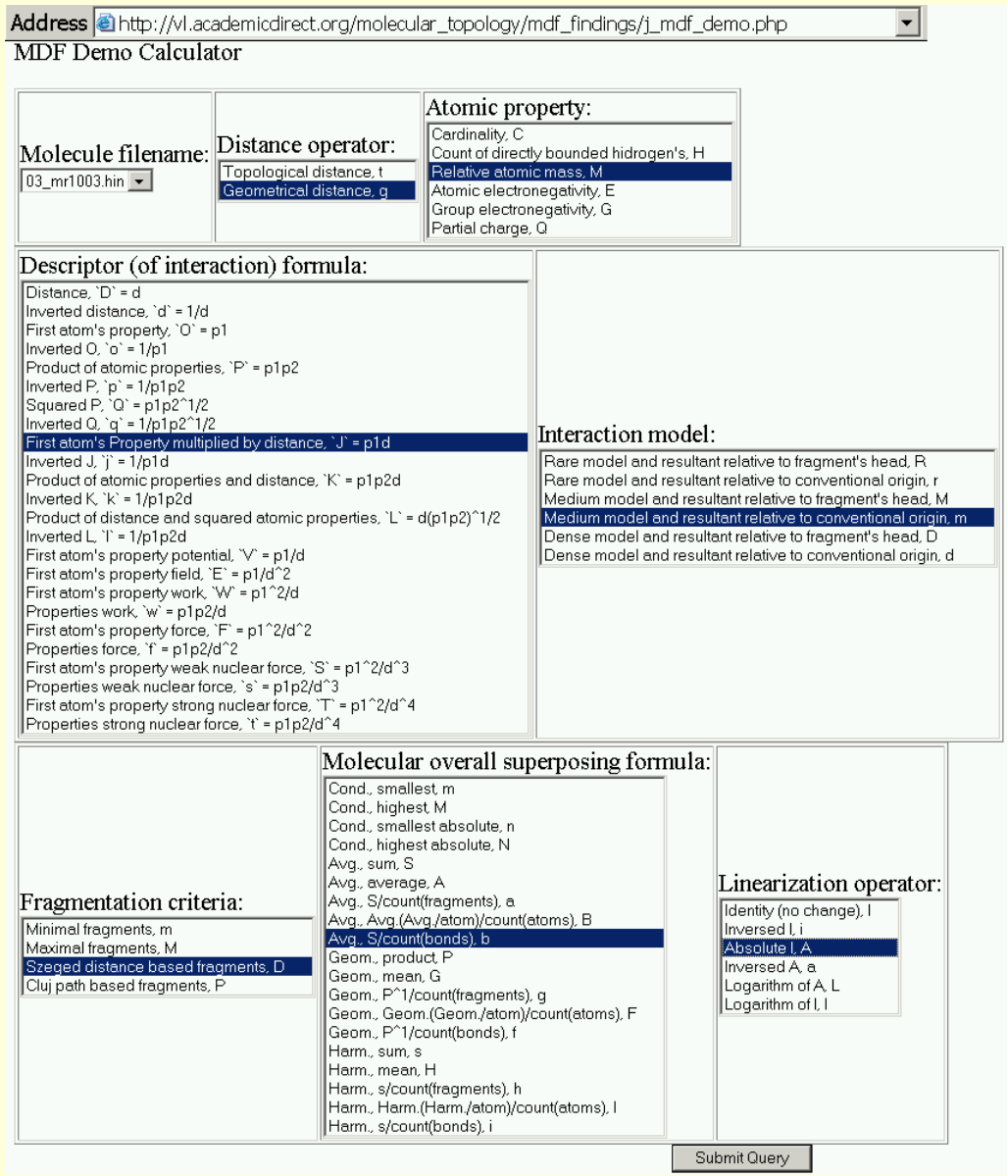


MR10_xval	107,692	MyISAM	latin1_swedish_ci	8.3 MB	-
MR10_yval	107,692	MyISAM	latin1_swedish_ci	4.9 MB	-
PCB_rrf_data	209	MyISAM	latin1_swedish_ci	2.8 KB	-
PCB_rrf_tmpx	131,328	MyISAM	latin1_swedish_ci	212.1 MB	-
PCB_rrf_xval	62,873	MyISAM	latin1_swedish_ci	100.3 MB	-
PCB_rrf_yval	62,873	MyISAM	latin1_swedish_ci	2.9 MB	-
RRC433_lbr_data	30	MyISAM	latin1_swedish_ci	1.3 KB	-
RRC433_lbr_tmpx	131,328	MyISAM	latin1_swedish_ci	32.7 MB	-
RRC433_lbr_xval	86,409	MyISAM	latin1_swedish_ci	19.9 MB	-
RRC433_lbr_yval	86,409	MyISAM	latin1_swedish_ci	4.0 MB	-
RRC433_lkow_data	30	MyISAM	latin1_swedish_ci	1.3 KB	-
RRC433_lkow_tmpx	131,328	MyISAM	latin1_swedish_ci	32.7 MB	-
RRC433_lkow_xval	86,388	MyISAM	latin1_swedish_ci	19.9 MB	-
RRC433_lkow_yval	86,388	MyISAM	latin1_swedish_ci	4.0 MB	-
RRC433_pka_data	30	MyISAM	latin1_swedish_ci	1.3 KB	-
RRC433_pka_tmpx	131,328	MyISAM	latin1_swedish_ci	32.7 MB	-
RRC433_pka_xval	86,373	MyISAM	latin1_swedish_ci	19.9 MB	-
RRC433_pka_yval	86,373	MyISAM	latin1_swedish_ci	4.0 MB	-
Ta395_data	15	MyISAM	latin1_swedish_ci	1.1 KB	-
Ta395_tmpx	131,328	MyISAM	latin1_swedish_ci	17.7 MB	-
Ta395_xval	102,608	MyISAM	latin1_swedish_ci	11.8 MB	-
Ta395_yval	102,608	MyISAM	latin1_swedish_ci	4.7 MB	-
Tox395_data	14	MyISAM	latin1_swedish_ci	1.1 KB	-
Tox395_tmpx	131,328	MyISAM	latin1_swedish_ci	16.7 MB	-
Tox395_xval	103,411	MyISAM	latin1_swedish_ci	11.1 MB	-
Tox395_yval	103,411	MyISAM	latin1_swedish_ci	4.7 MB	-
a_acids_data	12	MyISAM	latin1_swedish_ci	1.1 KB	-
a_acids_tmpx	131,328	MyISAM	latin1_swedish_ci	14.7 MB	-
a_acids_xval	97,007	MyISAM	latin1_swedish_ci	9.0 MB	-
a_acids_yval	97,007	MyISAM	latin1_swedish_ci	4.4 MB	-
molecular_topology	86	MyISAM	latin1_swedish_ci	17.1 KB	-
194 table(s)	Sum	14,305,820	--	latin1_swedish_ci	1.9 GB	0 Bytes

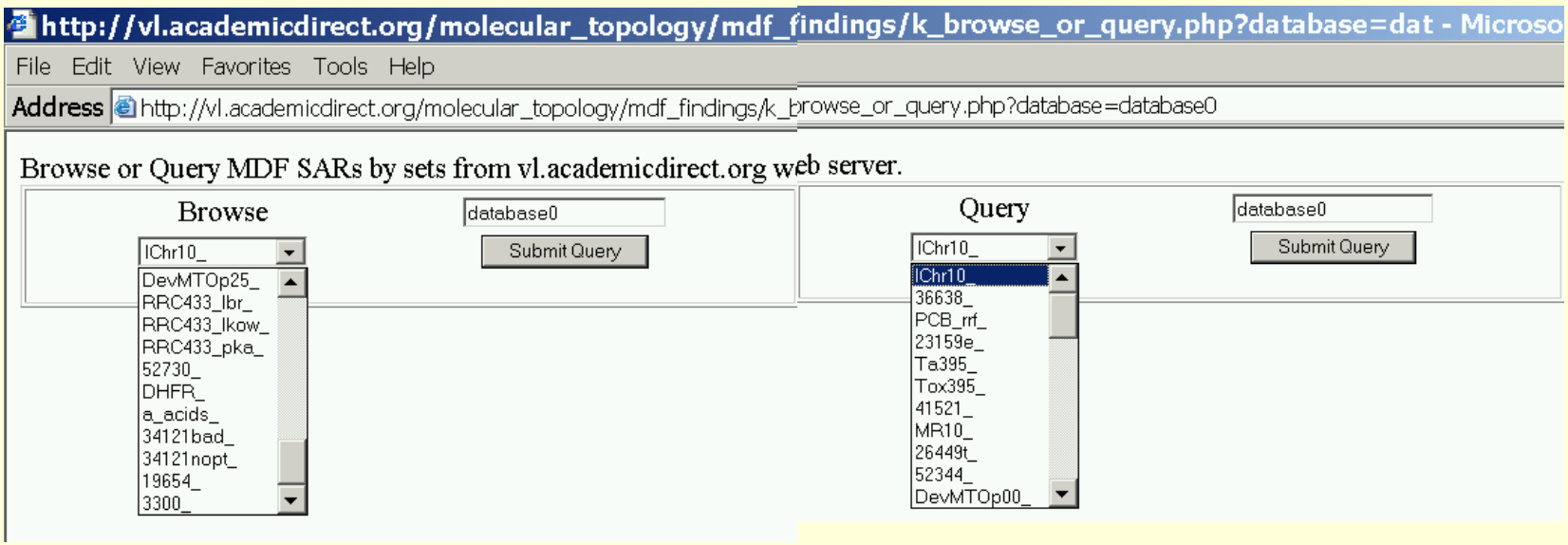
Results: web interface



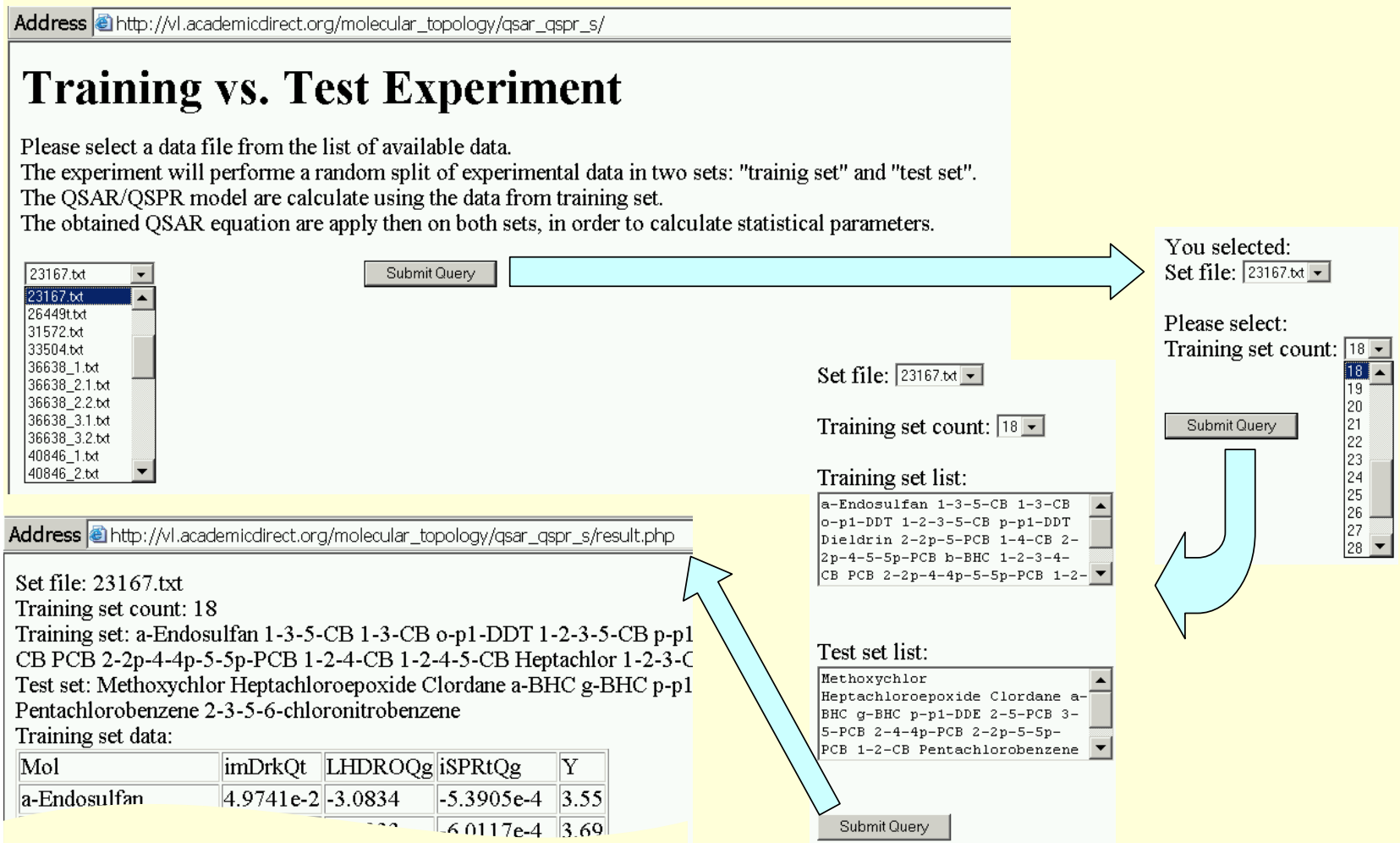
Results: MDF model



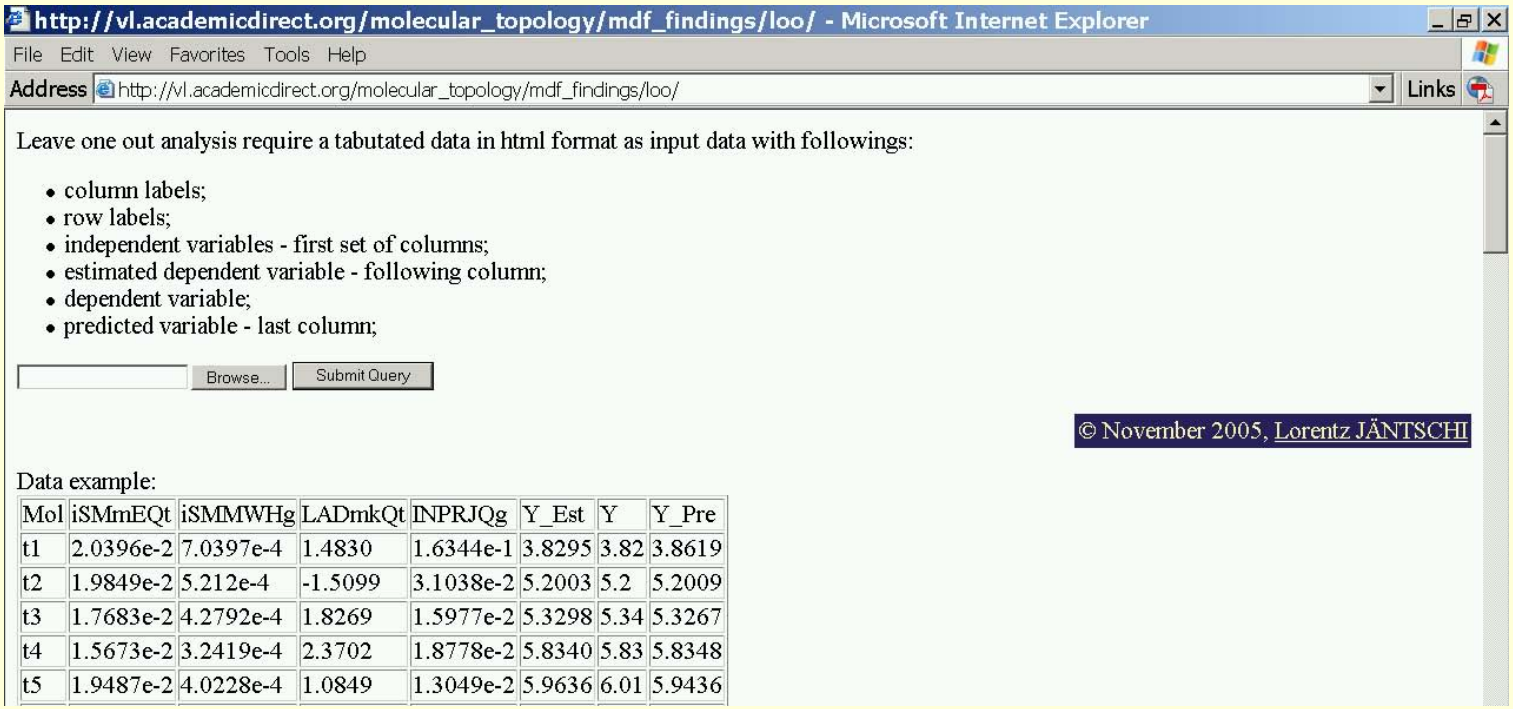
Results: investigated datasets



Results: training vs. test application



Results: leave one out application



Mol	iSMmEQt	iSMMWHg	LADmkQt	INPRJQg	Y_Est	Y	Y_Pre
t1	2.0396e-2	7.0397e-4	1.4830	1.6344e-1	3.8295	3.82	3.8619
t2	1.9849e-2	5.212e-4	-1.5099	3.1038e-2	5.2003	5.2	5.2009
t3	1.7683e-2	4.2792e-4	1.8269	1.5977e-2	5.3298	5.34	5.3267
t4	1.5673e-2	3.2419e-4	2.3702	1.8778e-2	5.8340	5.83	5.8348
t5	1.9487e-2	4.0228e-4	1.0849	1.3049e-2	5.9636	6.01	5.9436

Results: comparisons

Set name	Previous reported SAR			MDF SAR			
	r^2	n	v	r^2	$r^2_{cv(loo)}$	n	v
IChr10	0.900	10	2	0.999	0.999	10	2
PCB_rrf	-	-	-	0.628	0.619	209	1
				0.693	0.682	209	2
				0.737	0.717	209	4
PCB_lkow	-	-	-	0.873	0.87	206	1
				0.89	0.885	206	2
				0.917	0.909	206	4
36638	0.967	16		0.994	0.991	16	3
23159	0.388	18	1	0.755	0.684	18	1
	0.839	18	3	0.982	0.974	18	2
23159e	-	-	-	0.899	0.758	8	1
				0.968	0.898	8	2
Ta395	0.870	13	2	0.977	0.961	15	2
Tox395	0.800	13	2	0.957	0.934	14	2

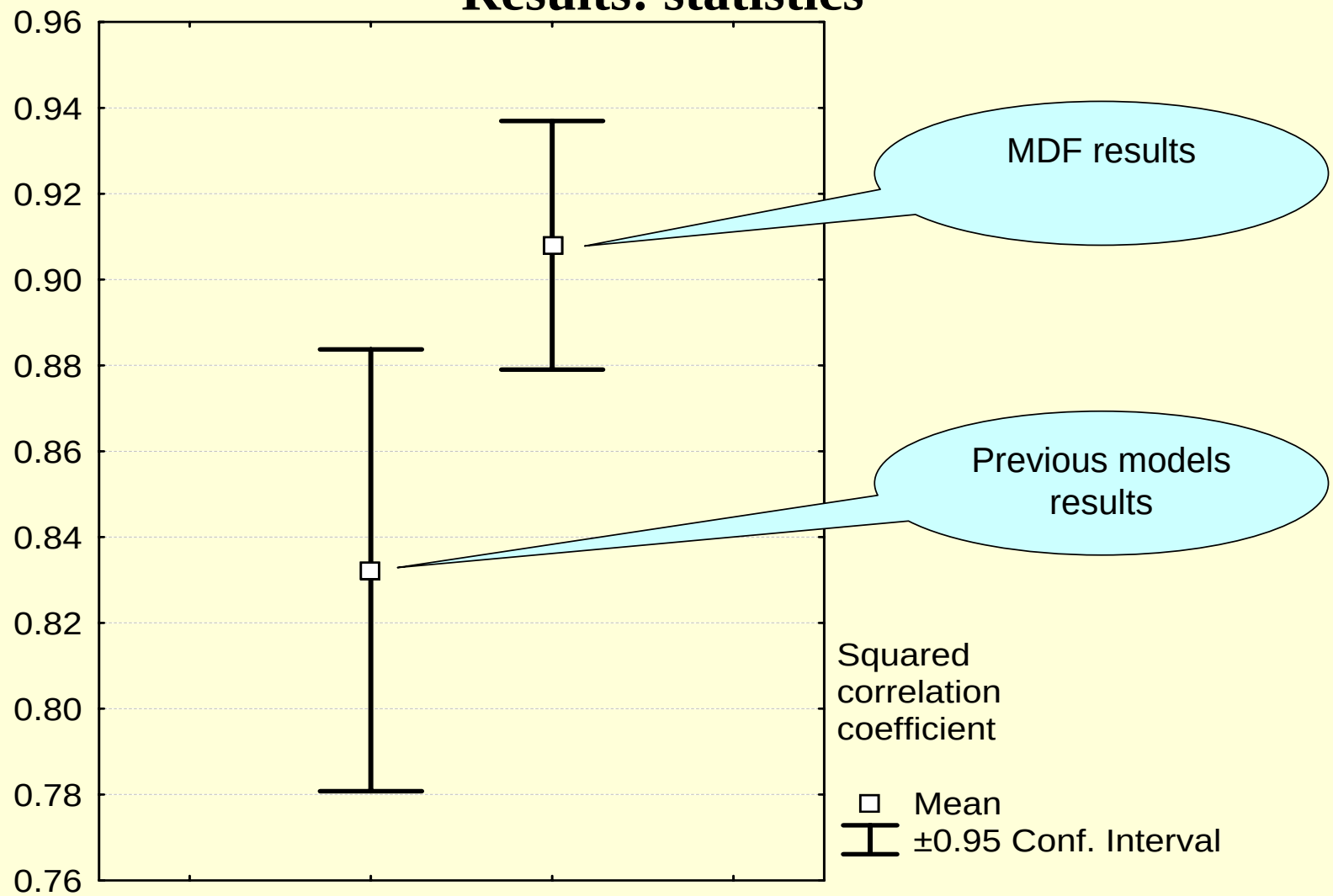
Results: comparisons

Set name	Previous reported SAR			MDF SAR			
	r^2	n	v	r^2	$r^2_{cv(loo)}$	n	v
	0.913	8	3				
41521	0.985	8	5	0.999	0.998	8	2
	0.991		1	0.961	0.954	10	1
	0.998		2	0.99	0.988	10	2
26449	0.993	10	4	0.998	0.997	10	4
MR10	0.976	10	2	0.999	0.999	10	2
	0.741	16	4				
23151	0.985	13	4	0.997	0.995	16	3
	0.780	8	1	0.904	0.832	8	1
	0.710	8	1	0.999	0.999	8	2
	0.810	8	2	0.999	0.999	8	2
52344	0.970	8	4	0.999	0.999	8	2

Results: comparisons

Set name	Previous reported SAR			MDF SAR			
	r^2	n	v	r^2	$r^2_{cv(loo)}$	n	v
52730	-	-	-	0.966	0.947	10	1
				0.998	0.996	10	2
Triazines	0.970	30	3	0.951	0.946	30	1
				0.975	0.971	30	2
				0.983	0.976	30	3
				0.989	0.985	30	4
22583	0.888	37	5	0.783	0.766	57	2
	0.885	20	5	0.835	0.809	57	3
	0.883	57	5	0.9	0.884	57	4
				0.918	0.9	57	5

Results: statistics



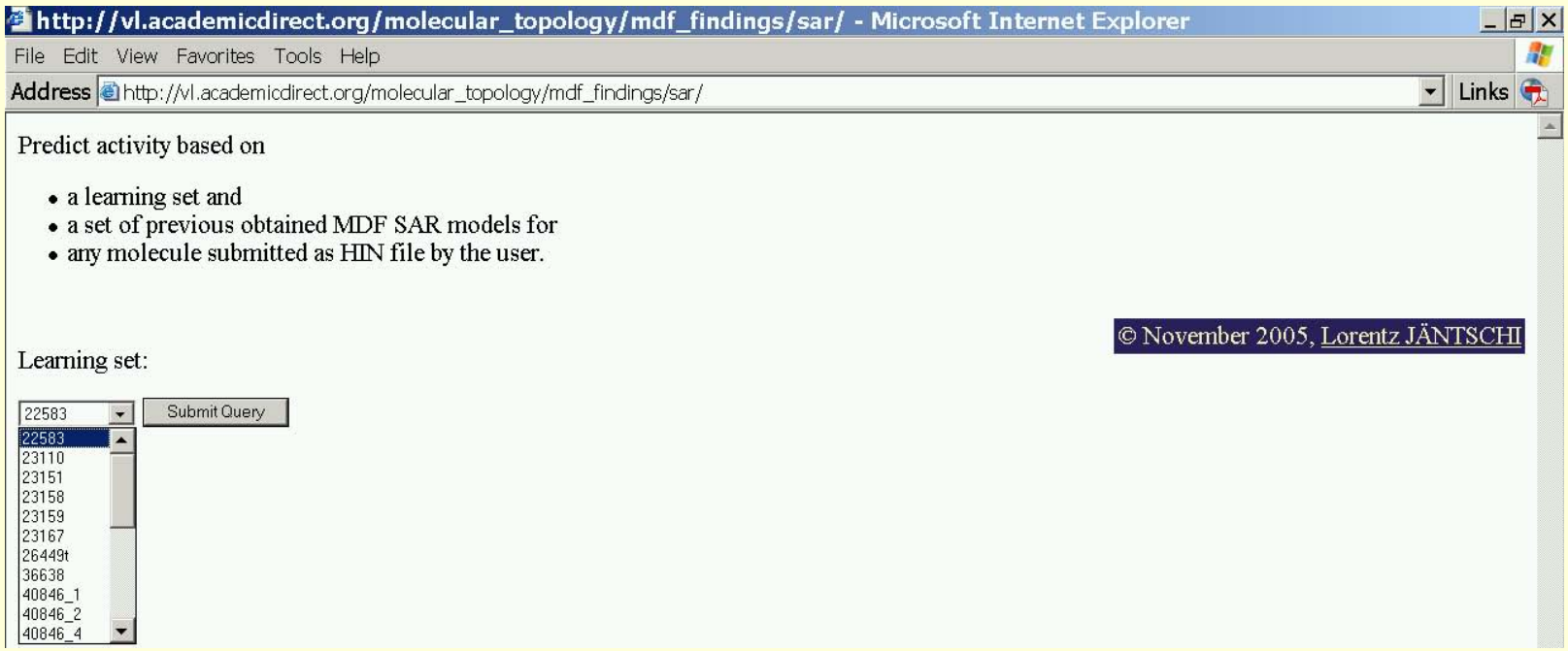
Conclusions: SPR

- **Molecular Descriptors Family on Structure-Property Relationships obtained very good performances in estimation and prediction of compound's properties as are for example retention chromatographic index, molar refraction and water activated carbon organics adsorption.**

Conclusions: SAR

- **Good performances are also obtained by Molecular Descriptors Family on Structure-Activity Relationships in estimation and prediction of compound's activities as are for example toxicity of alkyl metal compounds, insecticidal activity of neonicotinoid compounds, antituberculotic activity of polyhydroxyxanthones, antioxidant efficacy of 3-indolyl derivates, or antimalarial activity of some 2,4-diamino-6-quinazoline sulfonamide derivates.**

Conclusions: new compounds virtual building



Conclusions

- **The MDF SPR/SAR methodology opens a new pathway in:**
  - **understanding the relationships between compounds structure and property/activity,**
  - **characterization,**
  - **investigation,**
  - **development of new compounds,**

  more competitive as costs and property/activity.

Acknowledgments

- **Research was partially supported by the M. Ed. & Res., through ET/36/2005(-2008) project.**
- **AcademicDirect.Org domain is sponsored by Register.Com registrar.**
- **Servers IP's are from T.U. C-N addresses space.**

Thank you for your attention!