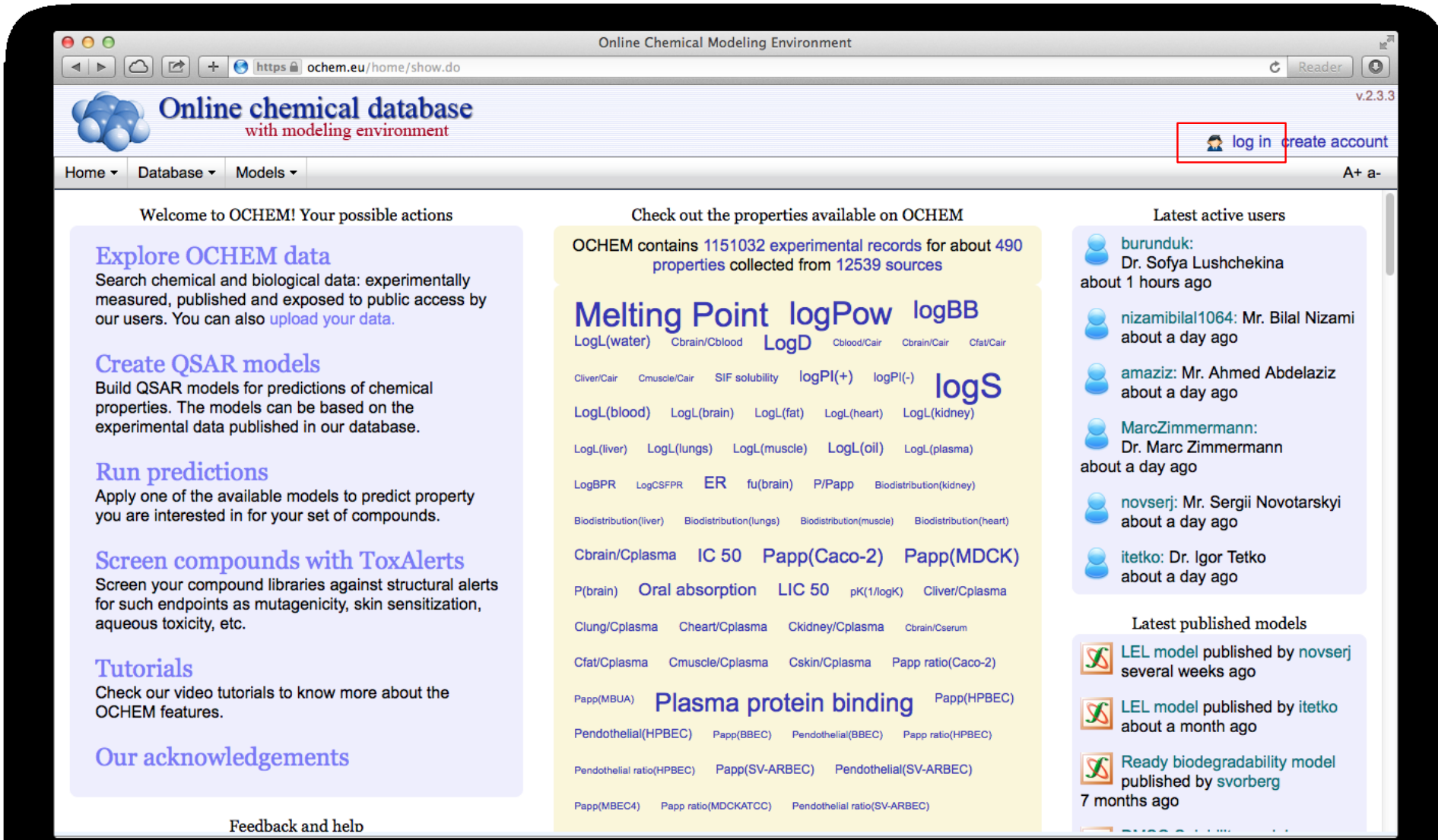
Tutoriál
OCHEM

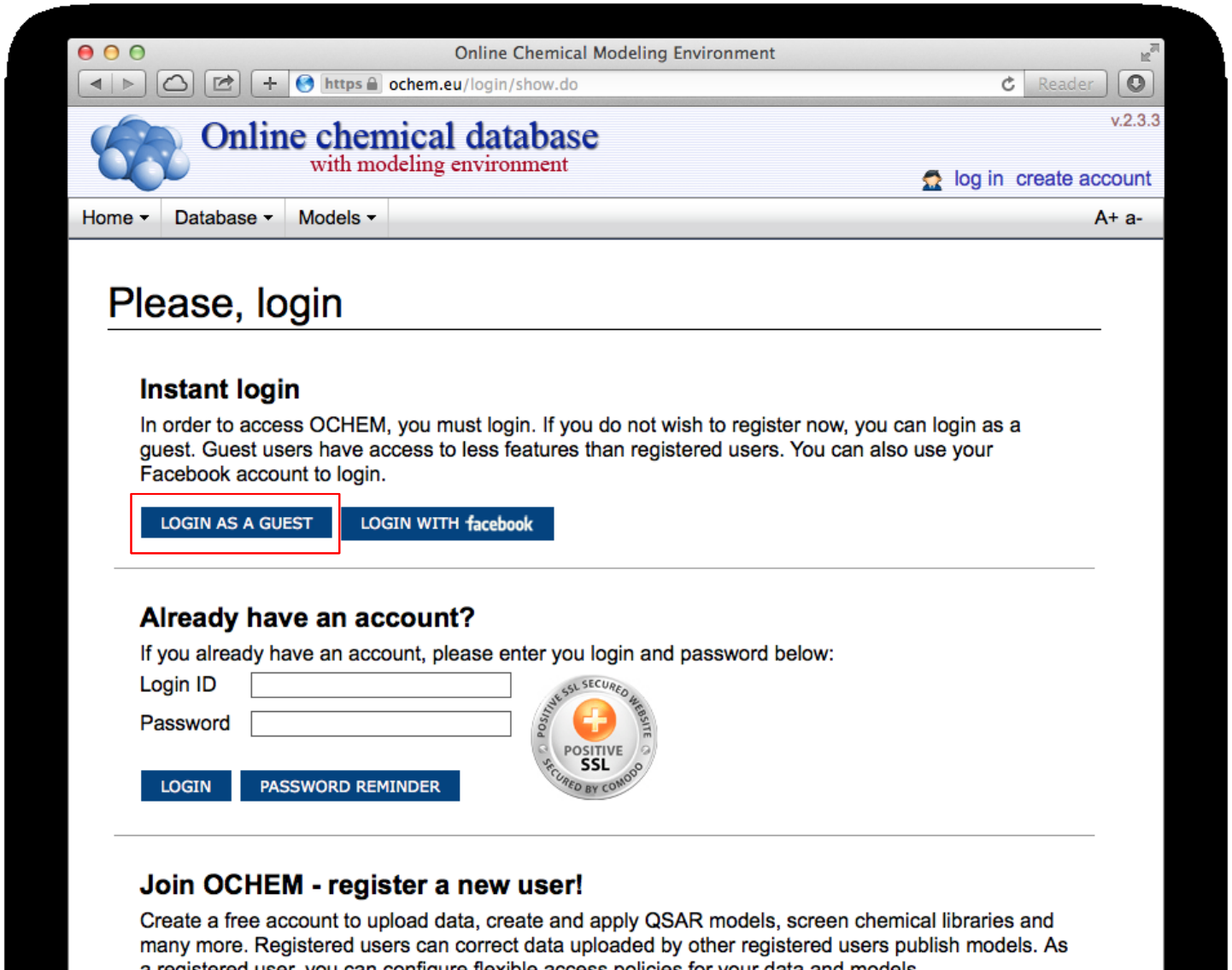
Pokročilá chemoinformatika

cvičení 3

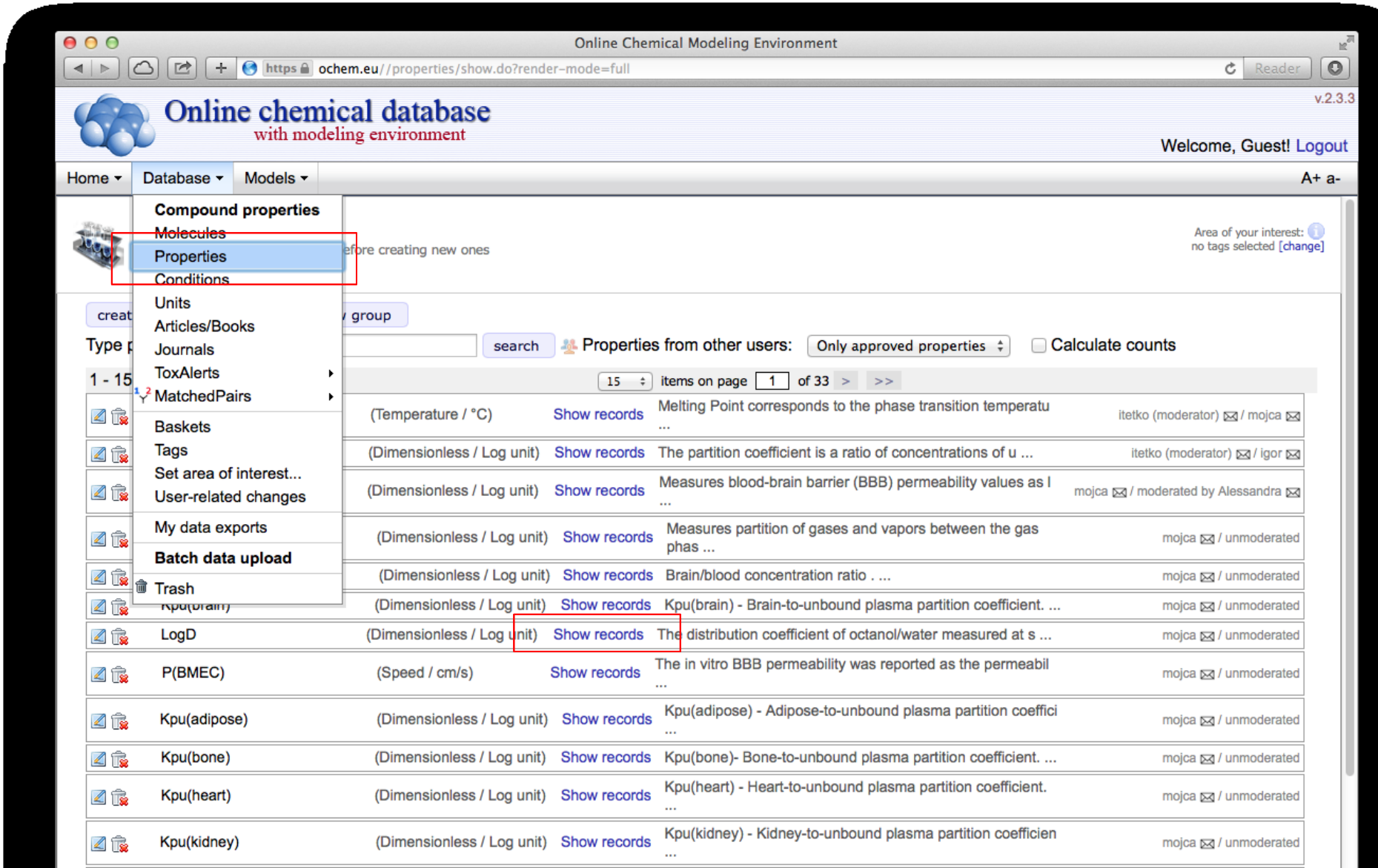
OChem – databáze a modelovací nástroj



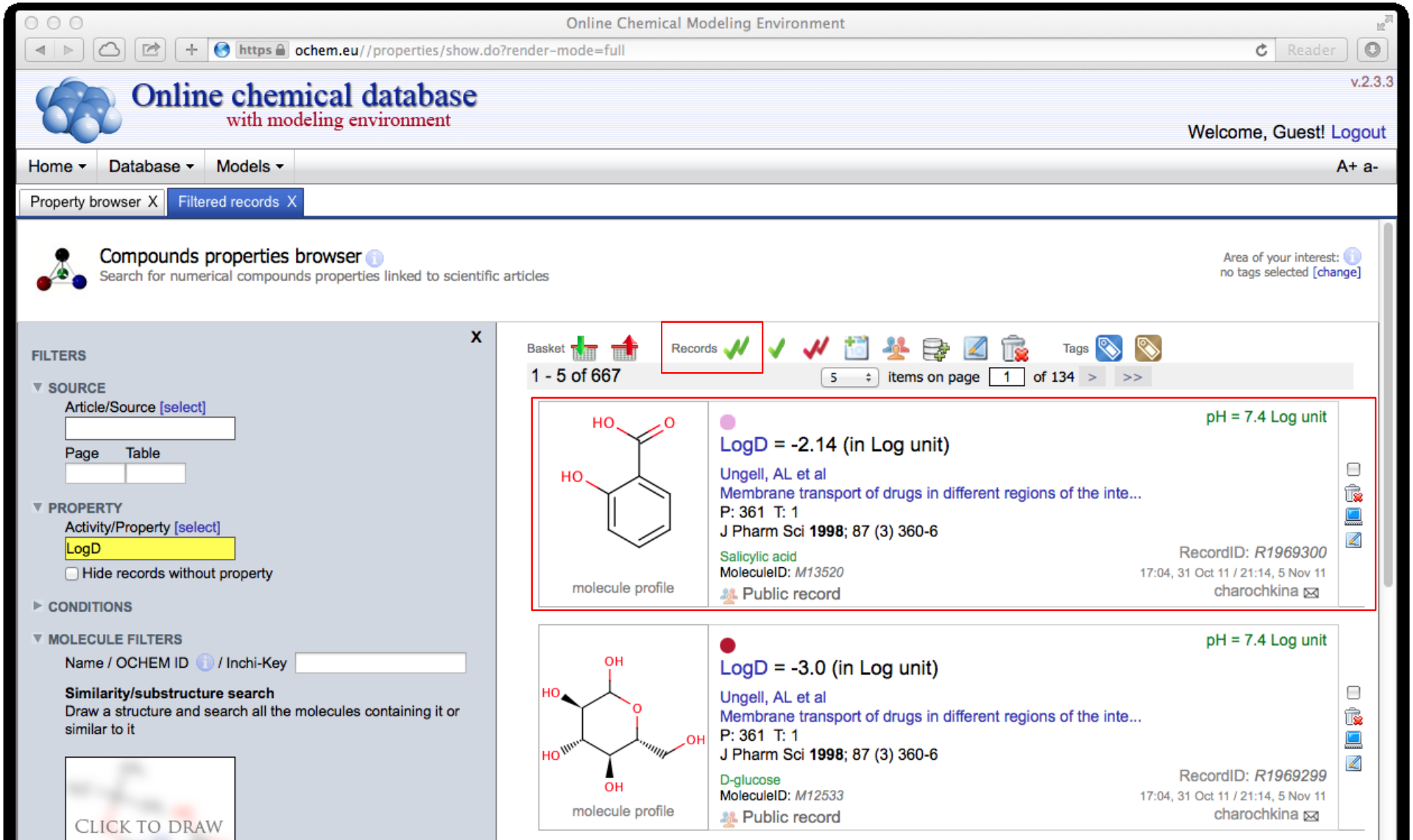
OChem – přihlášení



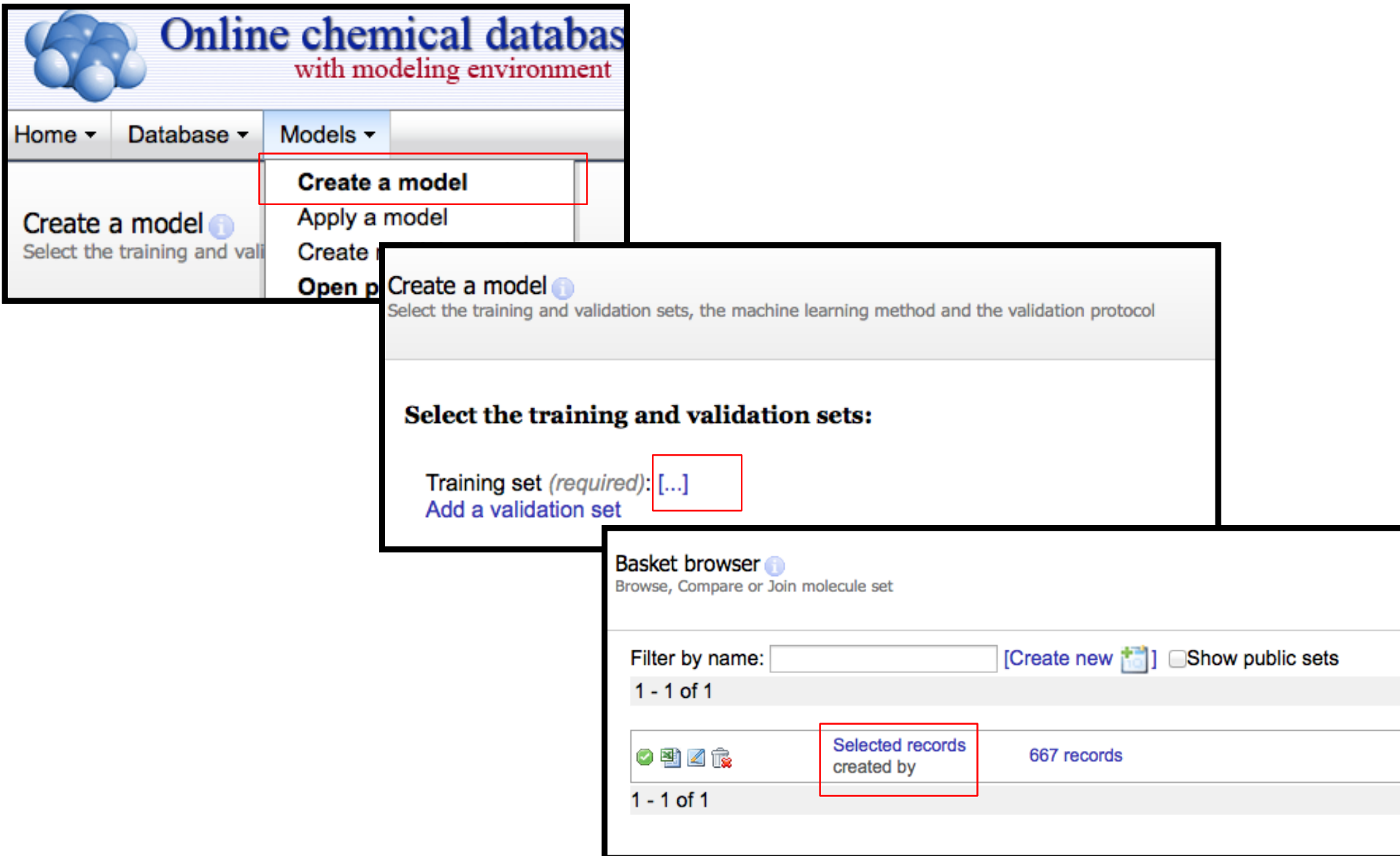
OChem – výběr vlastnosti a tréninkové sady



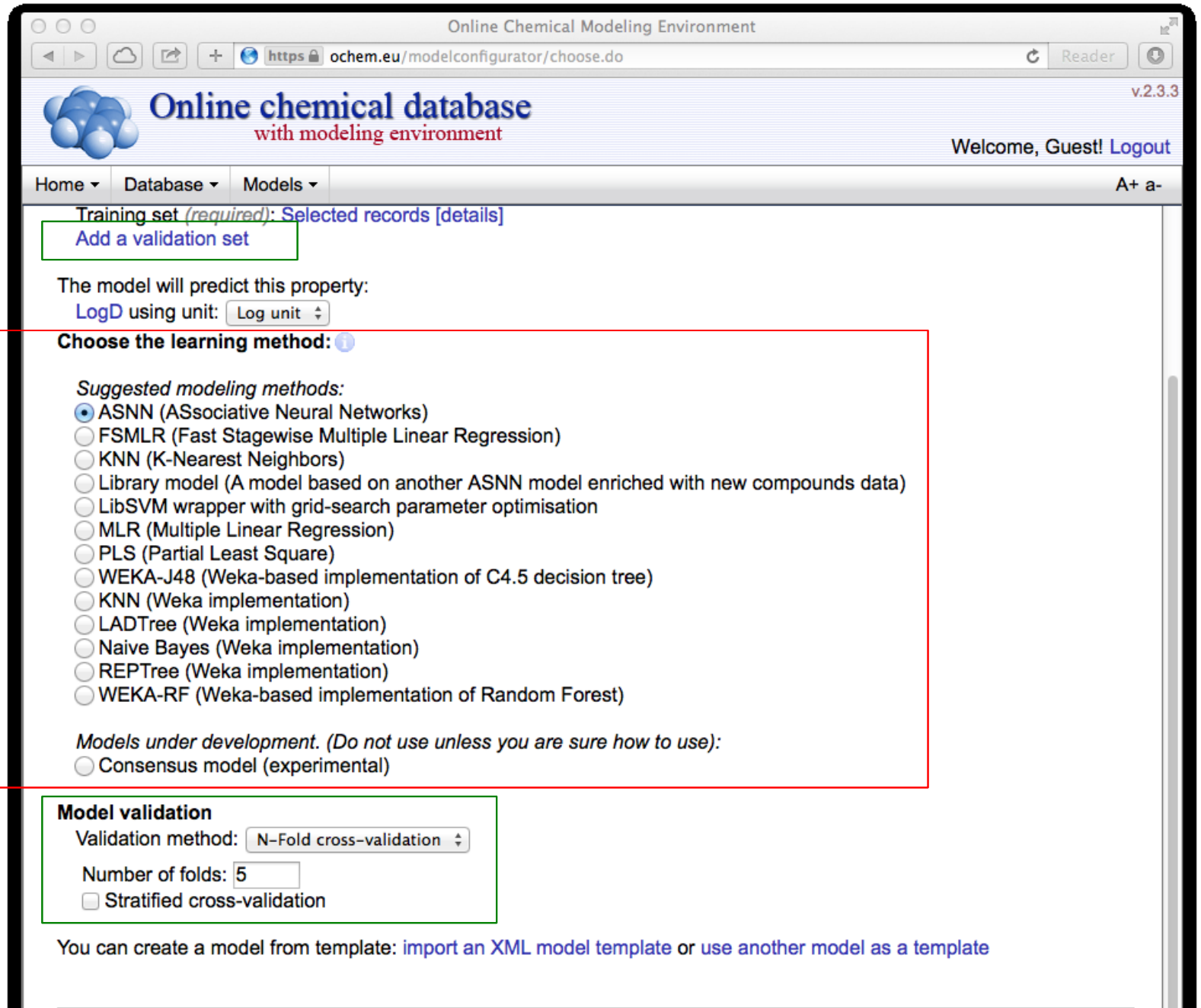
OChem – výběr vlastnosti a tréninkové sady



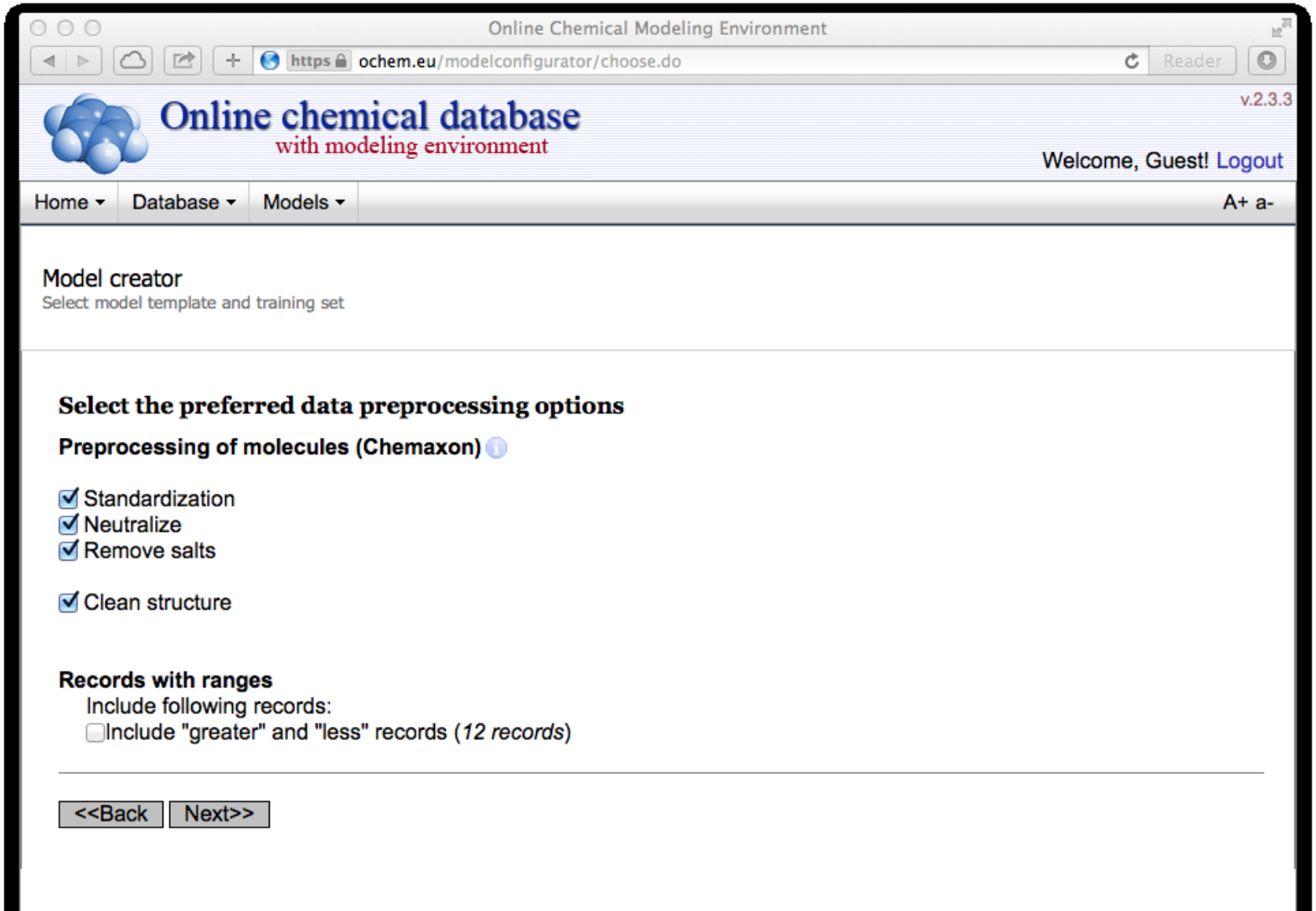
OChem – příprava modelu



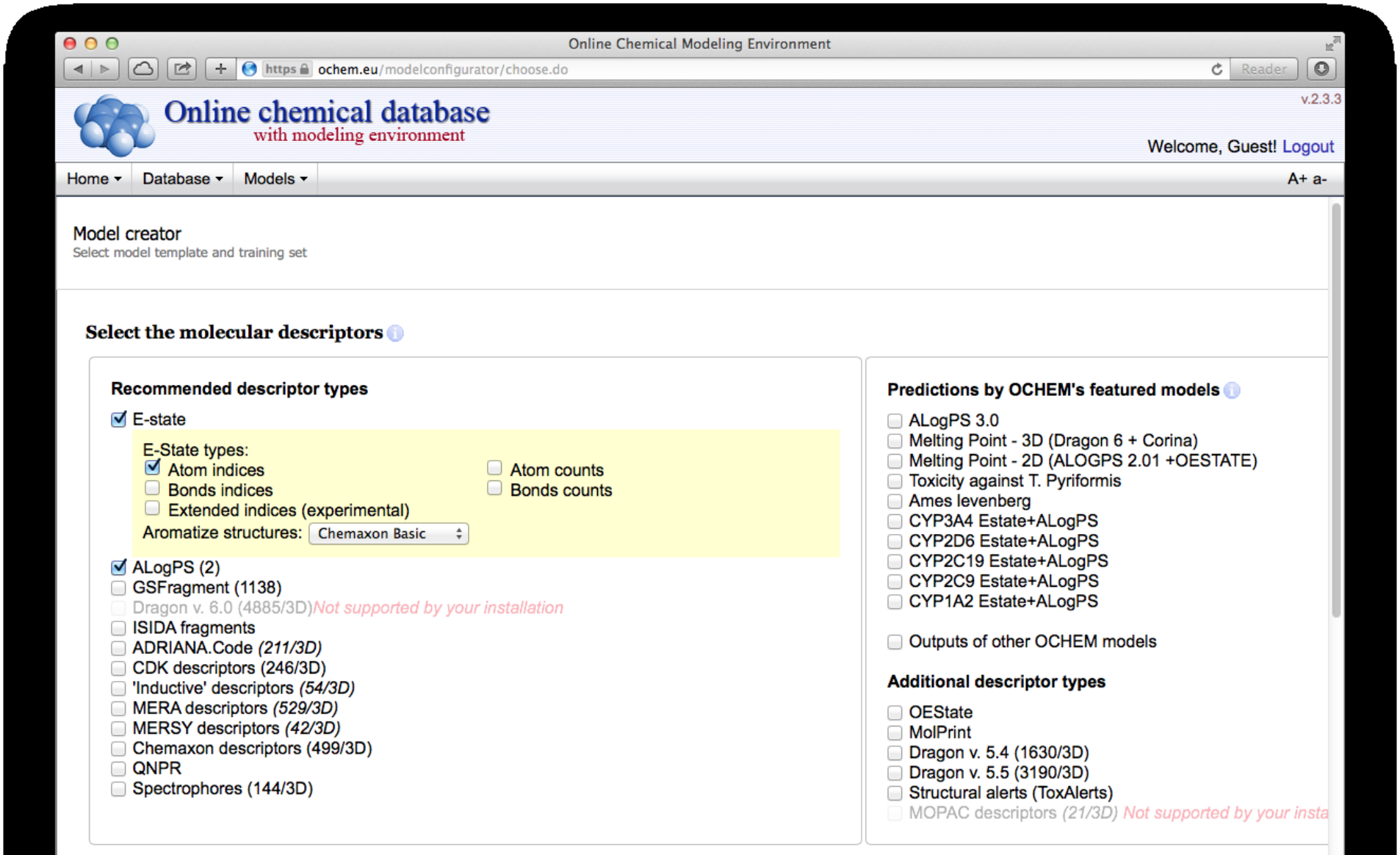
OChem – výběr metody



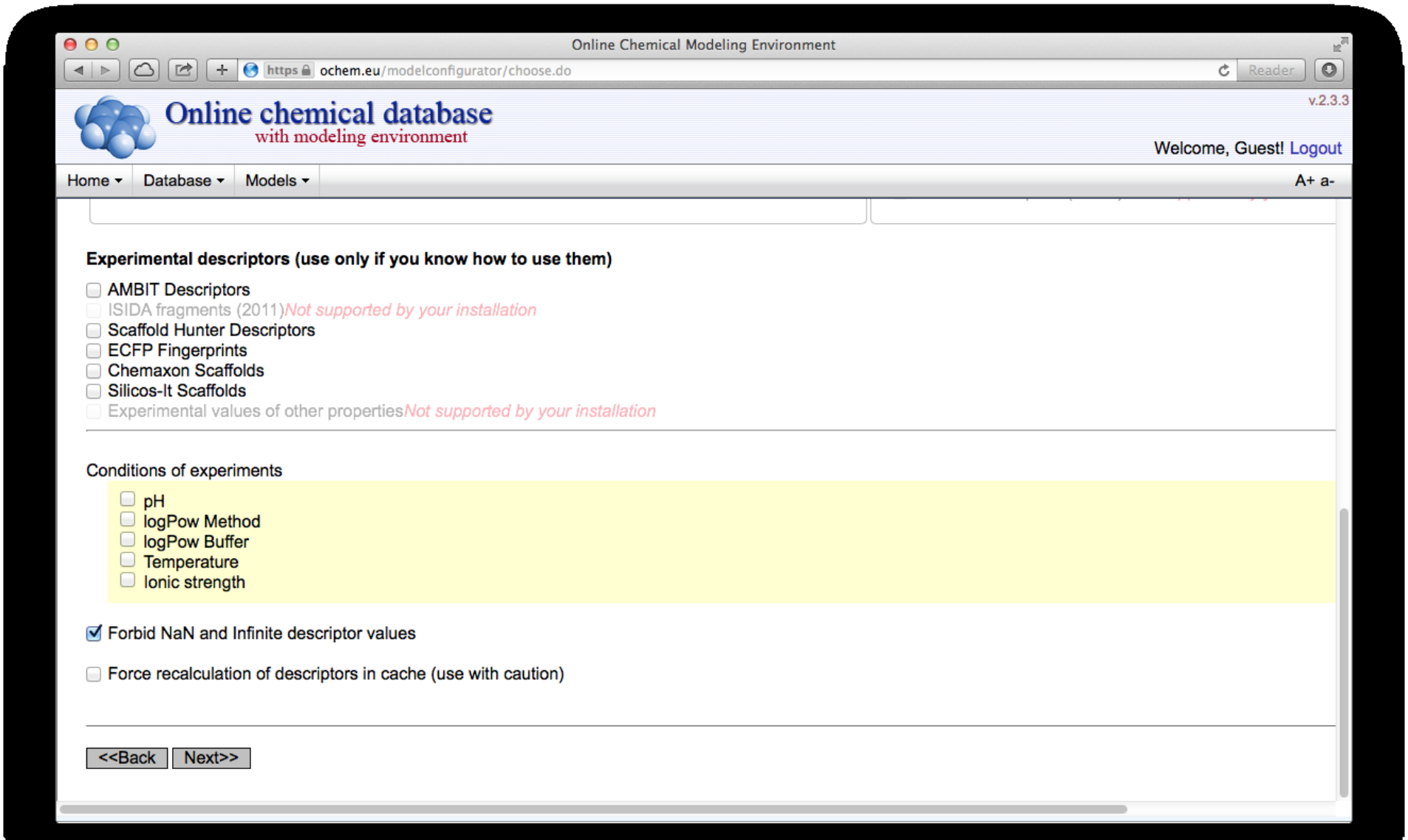
OChem – preprocessing



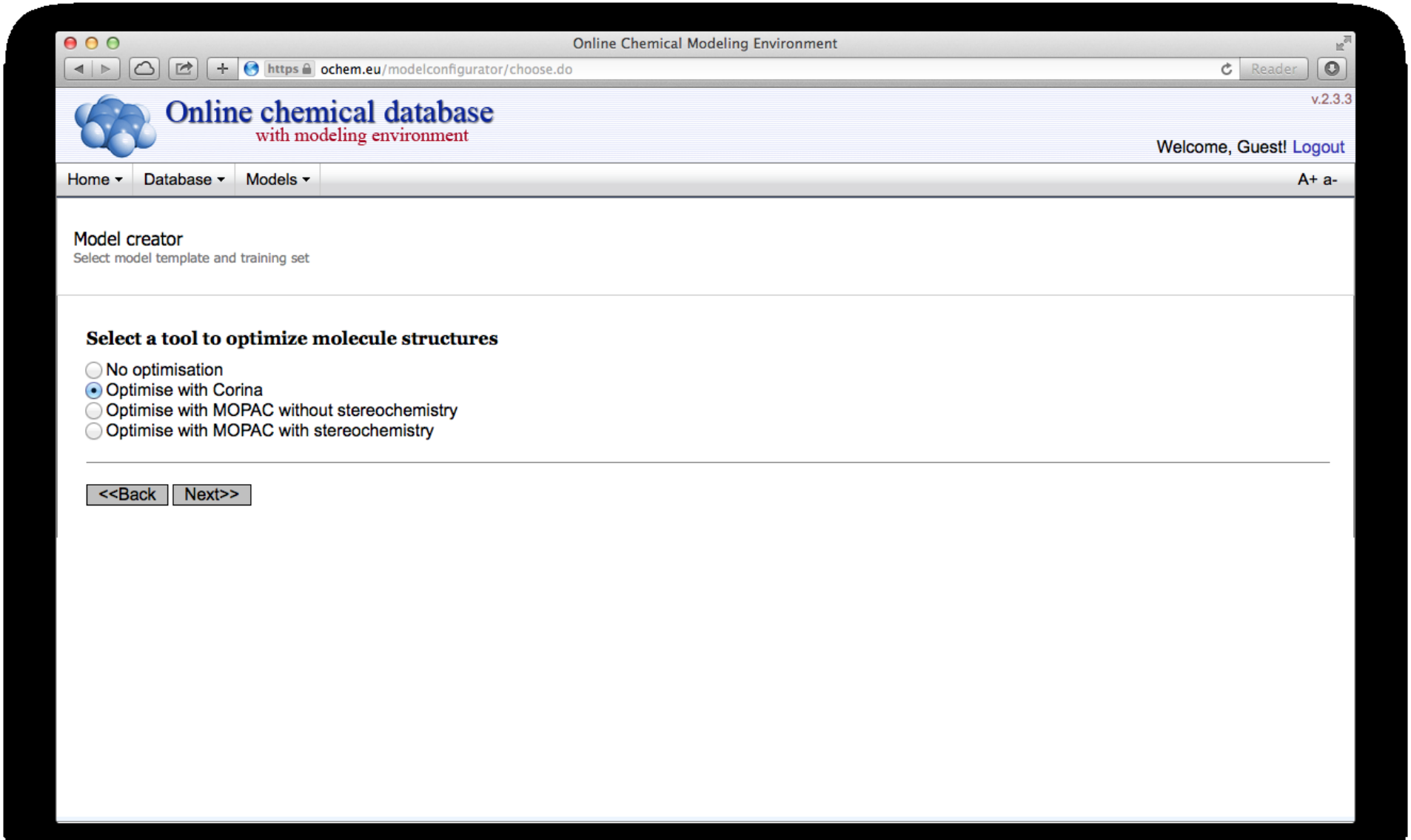
OChem – výběr deskriptorů



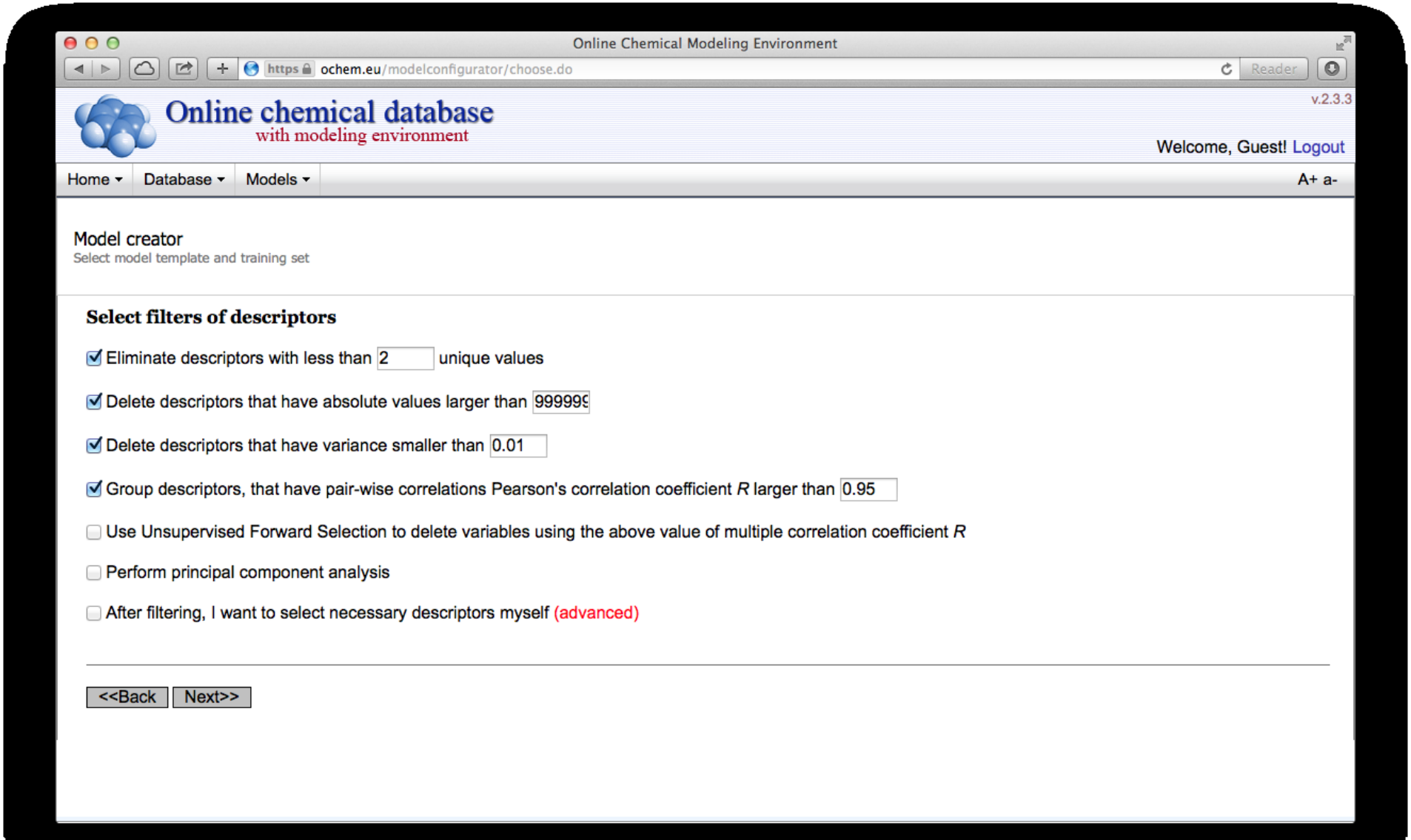
OChem – výběr deskriptorů 2



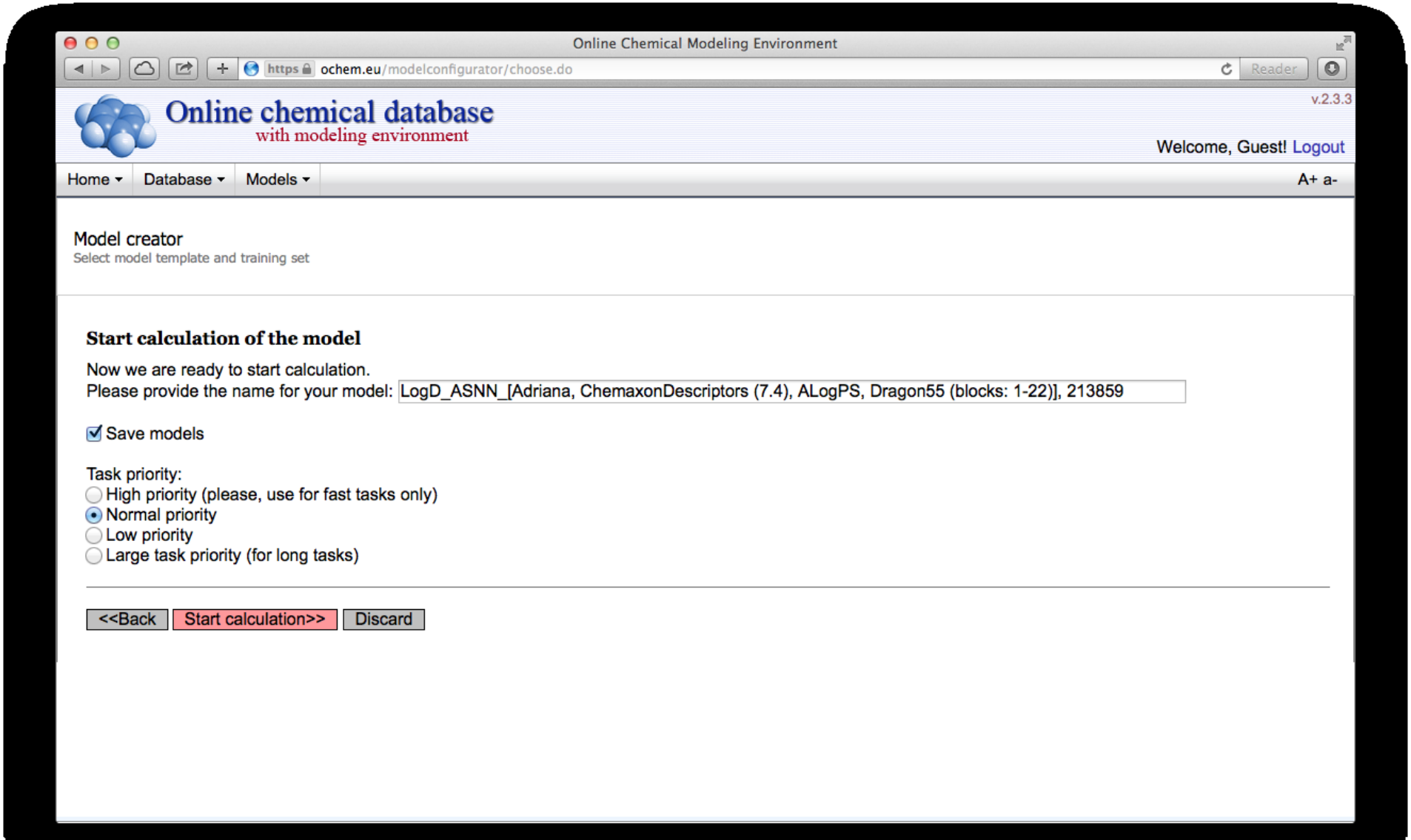
OChem – optimalizace struktur



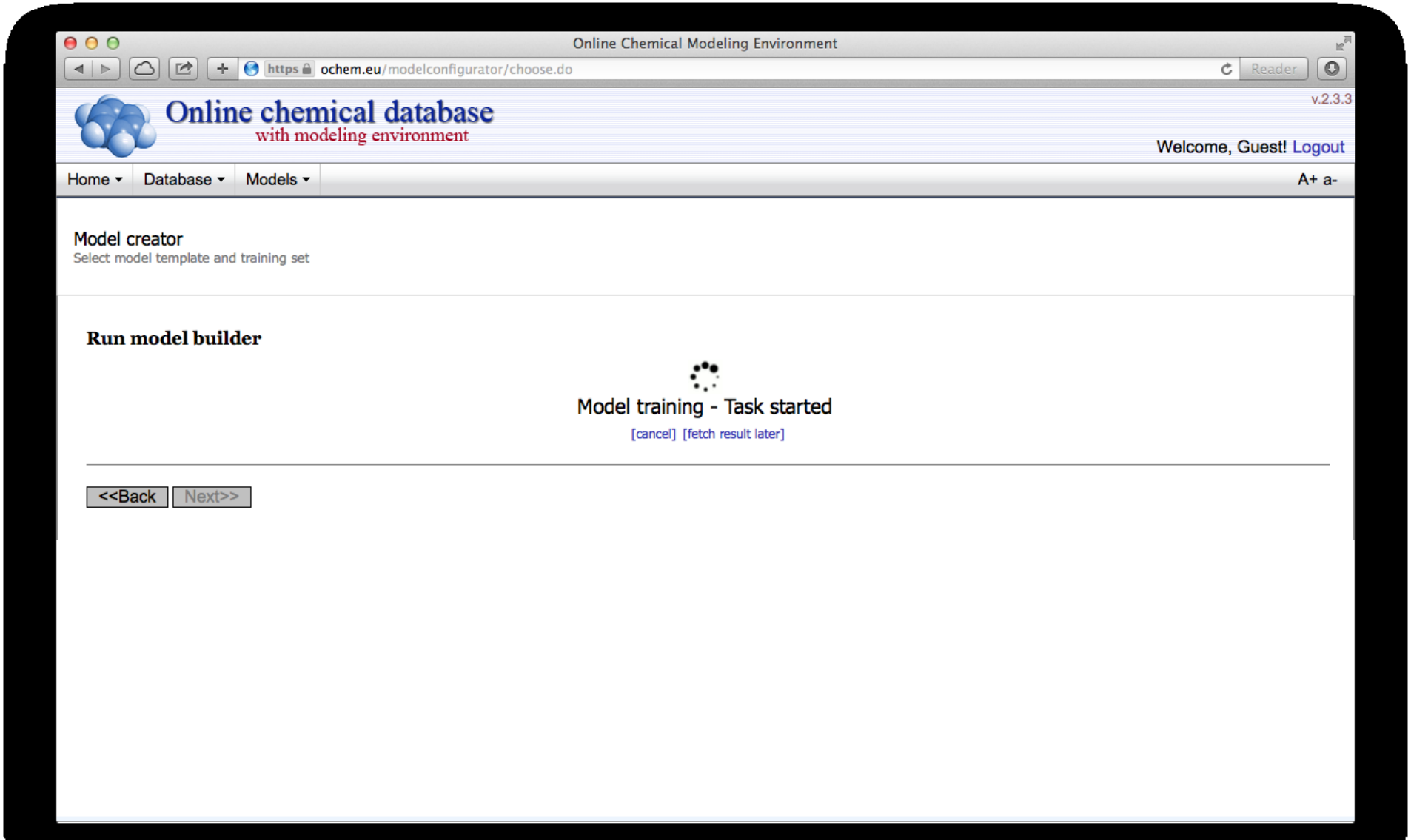
OChem – filtry deskriptorů



OChem – spuštění výpočtu



OChem – …


Online Chemical Modeling Environment
https ochem.eu/modelconfigurator/choose.do
Reader
Online chemical database
with modeling environment
v.2.3.3
Welcome, Guest! Logout
Home
Database
Models
A+ a-
Model creator
Select model template and training set
Run model builder
Model training - Task started
[cancel] [fetch result later]
<<Back
Next>>

OChem – vvsledkv

Online Chemical Modeling Environment

https ochem.eu/modelconfigurator/choose.do

Online chemical database
with modeling environment

v.2.3.3

Welcome, Guest! Logout

Home | Database | Models | A+ a-

Save the model

Please enter your model's name: LogD_ASNN_[Adriana, C

Overview | Applicability domain

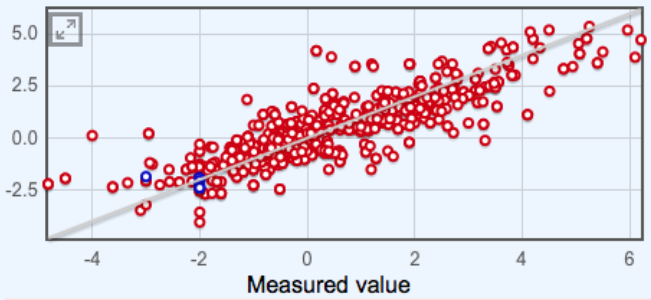
Model name: LogD_ASNN_[Adriana, ChemaxonDescriptors (7.4), ALogPS, Dragon55 (blocks: 1-22)], 213859 [rename]
Private ID is 31481910

Predicted property: **LogD**
Training method: ASNN

[Adriana, ChemaxonDescriptors (7.4), ALogPS, Dragon55 (blocks: 1-22)]
Correl. limit: 0.95 Variance threshold: 0.01, Maximum value: 999999, Supersab, 1000 iterations, 3 neurons ensemble=64 additional param PARTITION=3,SELECTION=2
5-fold cross-validation
-
1221 pre-filtered descriptors
Supersab, 1000 iterations, 3 neurons ensemble=64 k=32 additional param PARTITION=3,SELECTION=2

Calculated in 3186 seconds
Size: 659 Kb

Data Set	#	R^2	q^2	RMSE	MAE
Training set: Selected records	646 records	0.71 ± 0.02	0.7 ± 0.02	1.01 ± 0.04	0.77 ± 0.02
Excluded from training set	12 records	0.2 ± 0.2	0 ± 0	0.4 ± 0.1	0.27 ± 0.09



Number of compounds ignored because of errors in original model = 9

Save | Discard

Úkoly

① Vyberte si jednu metodu přípravy QSAR/QSPR modelu:

   ① ASNN
   ② MLR
   ③ KNN

   a pomocí OChemu připravte podle této metody model predikující teplotu bodu tání. Porovnejte s uloženými modely.

② Připravte model predikující hodnotu pKa a diskutujte kvality modelu pomocí R2, RMSE a MAE.

③ Porovnejte výsledky validace pomocí tréninkové + testovací sady a „cross“ validace.