Documentation of SeqRate 1.0

SeqRate 1.0 can be installed on Linux system. We haven't tested it on Windows system.

First, you need to go to http://babbage.cs.missouri.edu/~chengji/pspro1.1.tar.gz
to download and install PSPRO1. It has an installation instruction file in the package;

You may need to go to the ./otherToolPath foler and do "chmod 755 svm_classify" to make that file executable;

How to run the SVMFoldRate:
java -jar SVMFolderRate.jar   seqPath   seqFileName   foldType   PSPROExePath PSPROOutPath   otherToolPathSVMExeFile   tmpOutDir
You need to provide 8 parameters:
1. seqPath: full path of FASTA formated input sequence (contain single sequence in single line without line breaker)
    ex: /home/nick/cross-species/protein_folding/1ADW.fas
2. seqFileName: name of protein sequence file composed by proteinID and file type extension
    ex: 1ADW.fas
3. foldType[two-state, multi-state, unknown]: 3 folding types, two-state, multi-state or unknown
    IF user inputs "two-state" or "multi-state", then program will skip classification step, and go directly to rate prediction step using corresponding model.
    IF user inputs "unknown" type, then program will classify the protein first into either one of fold type, then predict rate accordingly
    ex: unknown
4. PSPROExePath: full path of PSPRO executable folder path, no need to specify the name of executable, it has been specified in the code
    ex: /home/nick/pspro1.1/bin/predict_ss_sa_cm.sh
5. PSPROOutPath: full path of PSPRO output folder, including the name of the output folder
    ex: /home/nick/cross-species/protein_folding/PSPRO_Out
6. otherToolPath: full path of SVM-light executable £»
    ex: /home/nick/cross-species/protein_folding/otherTools
7. SVMExeFile" name of SVM-regression executable file name
    ex: svm_classify
8. tmpOutDir: full path of a temp folder, where the outputs of the SVM regression predictions will be saved, please specify the name of output folder
    ex: /home/nick/cross-species/protein_folding/tmpOut

Example of running command:
java   -jar   ./SVMFoldRate.jar   ./1ADW.fas   1ADW.fas
unknown   ./pspro1.1/bin/predict_ss_sa_cm.sh   ./tmpOut   ./otherTools/
svm_classify ./tmpOut

Output files and temp files are saved at ./tempOut

The SVM output folder will have 3 files generated from codes
     (1) proteidID_predRate.txt: predicted rate for the given proteinID
     (2) SVM_proteinID_input.txt: SVM-light input used for prediction
     (3) predOut: predictions for contact number, contact order and folding rate. You want
to use these results for display.

After running the command, you should get the following outputs in command line:

Making Contact map ... Done
Making folding rate test file ...Done
Making SVM prediction ... Done
Combine results ... Done

Elapsed seconds: 118.0
Done

Then the output files are at ./tmpOut.

In the preOut.txt file, if you get the following outputs, your program is correctly installed
and running correctly too:

folding type: two-state

predicted contact number: 0.7857

predicted contact order: 0.2948

predicted folding rate: 120.8638

predicted folding rate in log based 10 value: 2.0823

predicted folding rate in natural log value: 4.7947