Marek Kochańczyk
kochanczyk@corcoran.if.uj.edu.pl

Institute of Physics, Astronomy and Applied Computer Science,
Jagiellonian University, Kraków, Poland

jEllipse

the geometry-based structure highlighter for polypeptides

Project Homepage
http://bioinformatics.org/kochanczyk/jellipse

Supervisor
prof. dr hab. Irena Roterman-Konieczna
Department of Bioinformatics and Telemedicine,
Medical College, Jagiellonian University,
Kraków, Poland

This brief overview is an Annual Work Report for the Studies of Natural Sciences.
The author is a 4th semester student at the Institute.

Aug 18, 2005

1 Motivation

A unified description of protein structure is essential when structural similarity judgement, motif transitions or general folding are considered. In the applied approach an assumption that all polypeptide chain conformations can be classified as helix-derived coerced invention of a new method based on geometrical parameters – radius of curvature R and dihedral angle between two consecutive peptide bond planes V. Developed basing on the model pentapeptides [Bch1995] turned out to be more useful than traditional ϕ/ψ mapping [Bi2004] [InSB2004].

Within the model, for low-energy backbones the dependence of $ln(R)$ upon the V angle can be expressed as an empirical square function [JThB1995]. In real proteins this dependence is altered specifically for a particular structure resulting in a criterion for judging similarity [Bi2003].

2 Capabilities

In order to facilitate structural analysis, *jEllipse* has features like:

- automated retrieval of crystallographic data by *PDB id* [NAR2004],
- loading a PDB file from the local filesystem or local network,
- chain selection in case of multidomain proteins,
- basic structure integrity verification, automated disulfide bonds recognition,
- calculation of ϕ and ψ torsional angles and plotting Ramachandran map (with the ellipse-path limited conformational sub-space marked),
- calculation of V and $ln(R)$ parameters and plotting them (with empirical parabolic function marked),
- plotting V, $ln(R)$, deviation of measured and theoretical values of V and $ln(R)$ versus parabolic curve and absolute value of the last – along the chain,
- highlighting abovementioned parameters – on the 3D backbone representation,
- various color schemes.

3 Technologies

The key factor while selecting the programming environment was user-centered design paradigm. After thorough reconaissance technologies of choice became

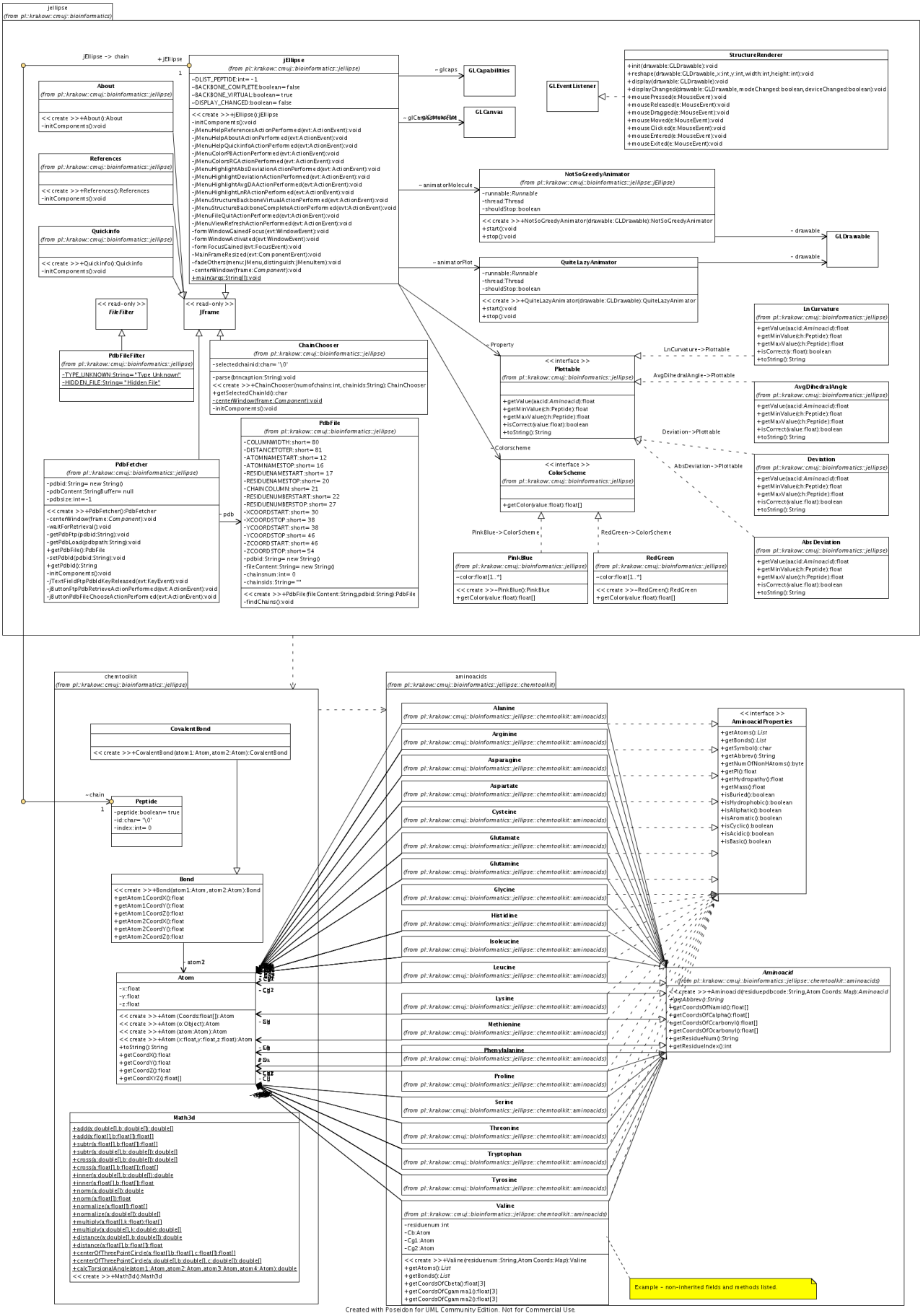
- Java 5.0 Standard Edition platform for
  - portability – a unique Java Web Start technology allowed to launch fully-fledged Java application with a single click in the web browser; graphical user interface was bulid upon the platform--independent Swing widget set,
  - networking – enabling direct connection to the archives of the RCSB Protein Data Bank (and on-the-fly decompression),
  - generic programming – especially useful when hierarchical structural realtions in biopolymers were to be modeled,
  - Java2D – to render high quality diagrams,
  - JavaDoc – that generates excellent documentation, particularly useful as project development was spanned over months.
- OpenGL 1.5 for efficient and interactive 3D rendering of molecules throught the JOGL – a set of Java reference bindings for OpenGL API.
- Poseidon for UML CE 3.1 for preparing draft project and generate initial source code,
- NetBeans IDE 4.1 with Apache Ant and CVS for project automation and general management.

jEllipse does not use any third-part bioinformatics libraries. Project was developed under Debian GNU/Linux, but is tested also for Sun JRE 5.0 on Ms Windows 2000/XP, Mac OS X 10.4 and OpenSolaris 1.

4 Calculations

Geometrical measurements are entirely based on [Bch1995].

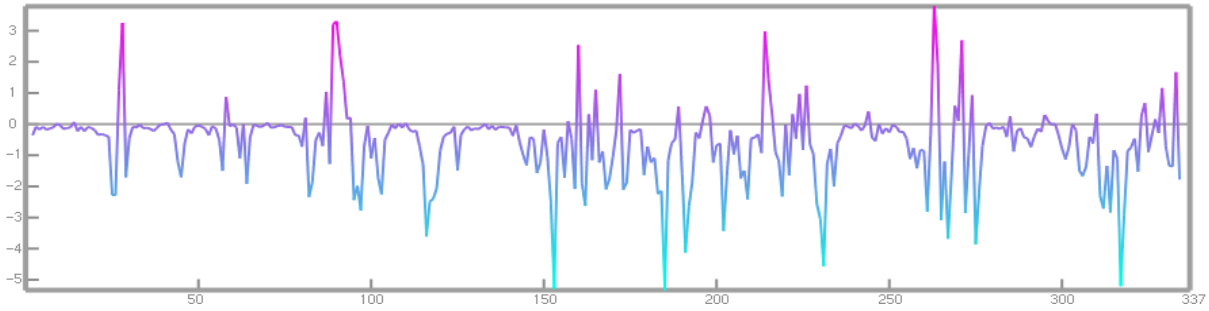
5 Class diagram



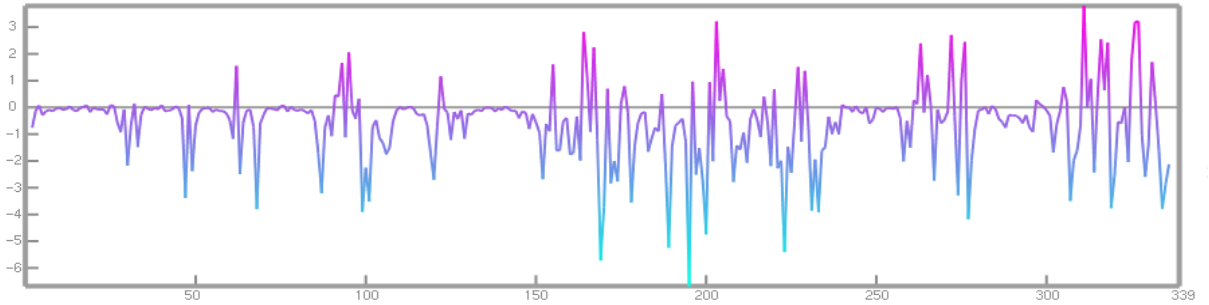
6 Example

Deviation of experimentally observed values of V and $ln(R)$ versus theoretical parabolic curve found based on model pentapaptides as the criterion for estimating structural similarity between serpine family members.

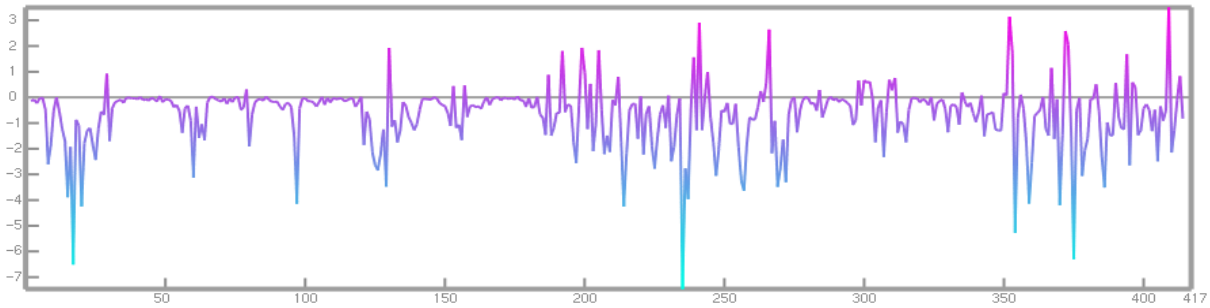
2ACH(A) Human α 1-antichymotrypsin



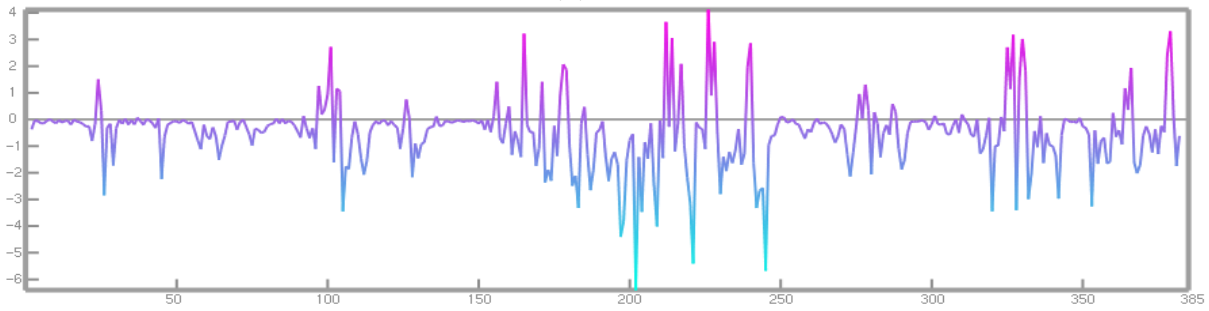
7API(A) The S variant of human α 1-antitrypsin



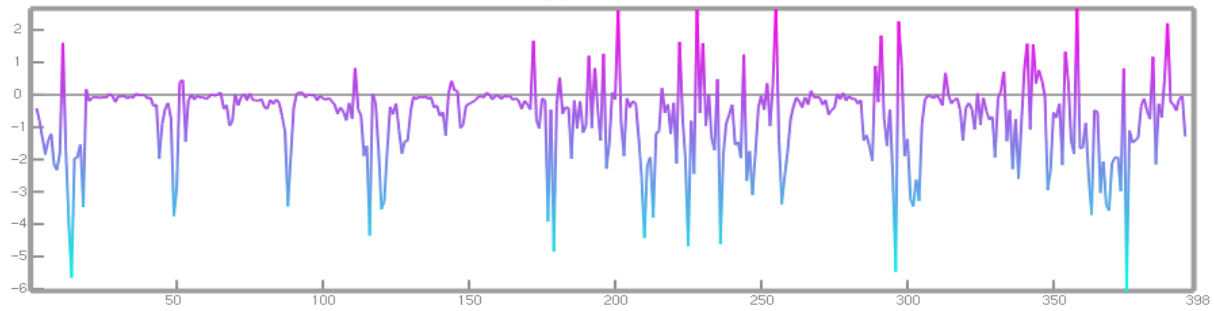
1AZX(I) Antithrombin



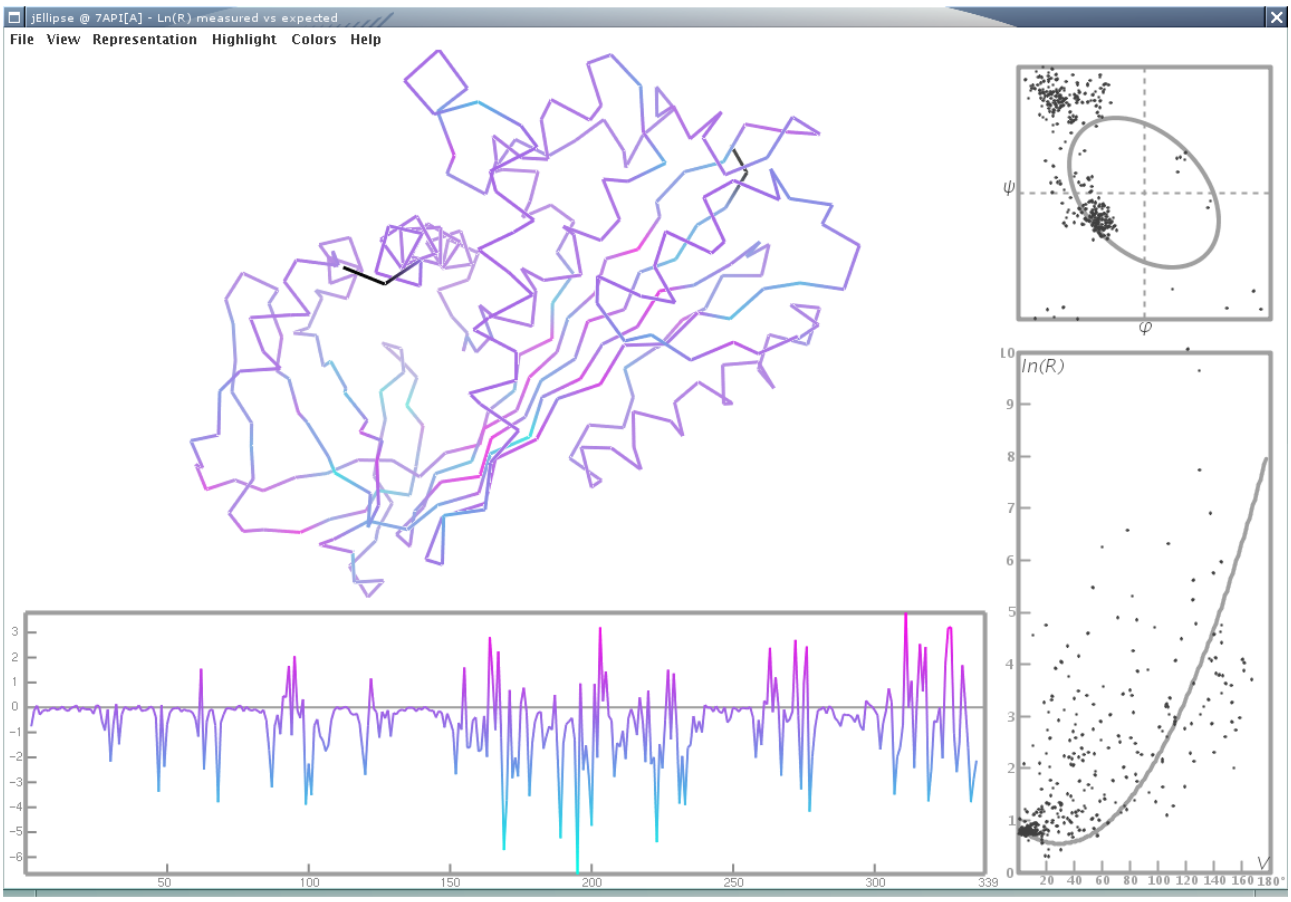
1OVA(A) Ovalbumin



1ANT(L) Antithrombin



7 Screenshot



(colors are inversed)

References

[Bch1995] Roterman I (1995) The geometrical analysis of peptide backbone structure and its local deformations. *Biochimie* **77**, 204-216

[JThB1995] Roterman I (1995) Modelling the optimal simulation path ni the peptide chain folding – Studies based on geometry of alanine heptapeptide. *J Theor Biol* **177**, 283-288

[Bi2003] Leluk J, Konieczny L, Roterman I (2003) Search for structural similarity in proteins. *Bioinformatics* **19**, 117-124

[Bi2004] Bryliński M, Jurkowski W, Konieczny L, Roterman I (2004) Limited conformational space for early-stage folding simulation. *Bioinformatics* **20**, 199-205

[InSB2004] Brylinski M, Konieczny L, Roterman I (2004) SPI - Structure predictability index for protein sequences *In Silico Biol* **5**

[NAR2004] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. *Nucl Acids Res* **28**, 235-242