

Atomic Contact Energies

Bioinformatics II

Adrian Geißler, Max-Emil Schön

June 22, 2015

Tutor: Linus Backert

Structure

Theoretical Background

What does the energy function model?

Methods

Implementation of the function.

Results

Performance Evaluation

Discussion

Theoretical Background

Appetizer



from: <http://thejobmouse.com>

The Free Desolation Energy

- ▶ Energy needed for transferring atoms from the solvent (water) into the protein's interior.
- ▶ One possible measure of protein stability
- ▶ The project:
 - ▶ Implement desolation energy by (Zhang et al., 1997)
 - ▶ Evaluate on CASP11 data

The Free Desolation Energy

- ▶ Energy needed for transferring atoms from the solvent (water) into the protein's interior.
- ▶ One possible measure of protein stability
- ▶ The project:
 - ▶ Implement desolation energy by (Zhang et al., 1997)
 - ▶ Evaluate on CASP11 data

The Free Desolation Energy

- ▶ Energy needed for transferring atoms from the solvent (water) into the protein's interior.
- ▶ One possible measure of protein stability
- ▶ The project:
 - ▶ Implement desolation energy by (Zhang *et al.*, 1997)
 - ▶ Evaluate on *CASP11* data

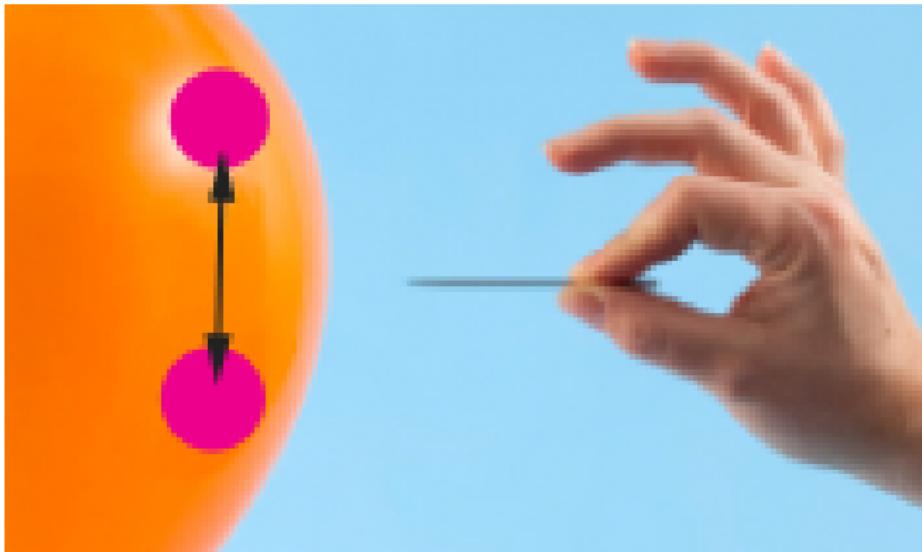
The Free Desolation Energy

- ▶ Energy needed for transferring atoms from the solvent (water) into the protein's interior.
- ▶ One possible measure of protein stability
- ▶ The project:
 - ▶ Implement desolation energy by (Zhang *et al.*, 1997)
 - ▶ Evaluate on *CASP11* data

The Free Desolation Energy

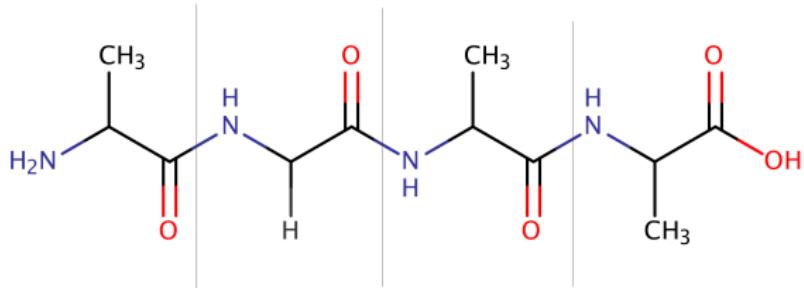
- ▶ Energy needed for transferring atoms from the solvent (water) into the protein's interior.
- ▶ One possible measure of protein stability
- ▶ The project:
 - ▶ Implement desolation energy by (Zhang *et al.*, 1997)
 - ▶ Evaluate on *CASP11* data

Atomic Contacts



Based on a picture from: <http://thejobmouse.com>

Atomic Contacts Pairs



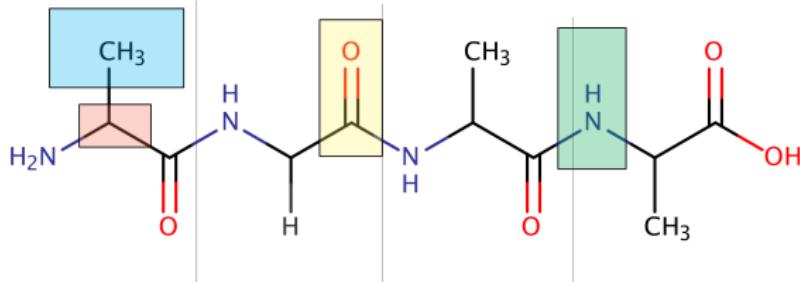
A valid contact pair:

- ▶ Only heavy atoms
- ▶ Distance below 6 Å
- ▶ More than 10 covalent bonds in between

Estimated by connectivity class & residue index differences

Overall energy is a simple sum

Atomic Contacts Pairs

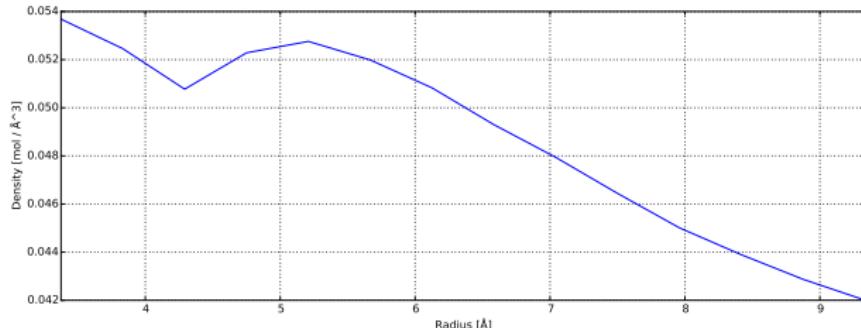


A valid contact pair:

- ▶ Only heavy atoms
- ▶ Distance below 6 Å
- ▶ More than 10 covalent bonds in between
Estimated by connectivity class & residue index differences

Overall energy is a simple sum

Atomic Packaging



- ▶ Number density of interior atoms (SAS = 0)
- ▶ Relative to a sphere of variable radius
- ▶ Evaluated on non-homologous protein set

Methods

Implementation

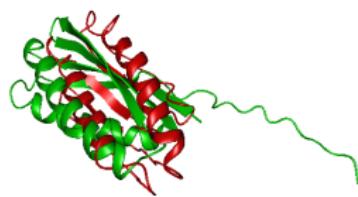
- ▶ Implemented in Python
- ▶ PDB package from BioPython
- ▶ Parameters identical to original version

Evaluation

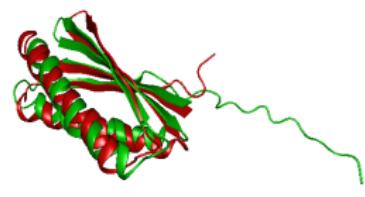
- ▶ Free desolution energy can imply structural stability
- ▶ Possible measure for Quality for predicted structures
- ▶ Evaluated on targets of CASP11 competition
- ▶ Comparison of superimposed $C\alpha$ atoms

Results

Comparison with Reference (1/2)

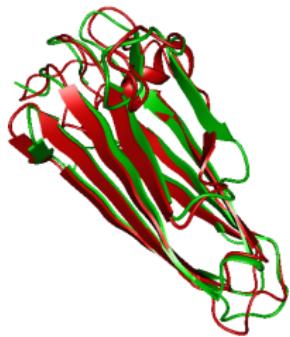


T0769-442



T0769-241

Comparison with Reference (2/2)



T0784-117



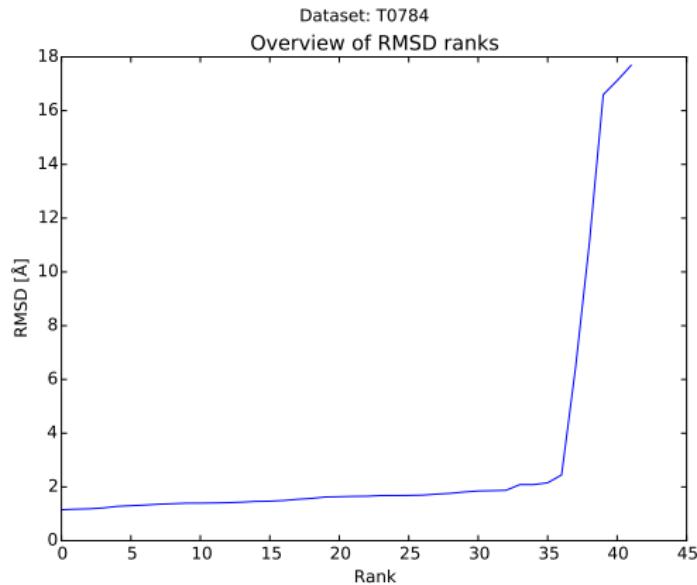
T0784-156

Energy v. RMSD

T0784	Energy in $\frac{kcal}{mol}$	RMSD	T0784	Energy in $\frac{kcal}{mol}$	RMSD
T0784TS156_1	-130.48	1.15	T0784TS117_1	-230.12	1.73
T0784TS420_1	-127.74	1.17	T0784TS008_1	-203.16	1.86
T0784TS499_1	-149.46	1.18	T0784TS251_1	-193.6	1.63
T0784TS237_1	-139.99	1.22	T0784TS038_1	-162.31	1.38
T0784TS268_1	-160.82	1.28	T0784TS268_1	-160.82	1.28

T0769	Energy in $\frac{kcal}{mol}$	RMSD	T0769	Energy in $\frac{kcal}{mol}$	RMSD
T0769TS241_1	-59.34	2.67	T0769TS442_1	-90.73	16.72
T0769TS368_1	-66.75	3.16	T0769TS155_1	-90.62	17.12
T0769TS258_1	-74.39	4.37	T0769TS044_1	-84.32	10.38
T0769TS361_1	-79.04	4.41	T0769TS169_1	-81.02	10.39
T0769TS186_1	-79.97	4.51	T0769TS317_1	-80.61	6.89

Investigation of correlation



Pearson: 0.19 – 0.44
Spearman: 0.15 – 0.33

Conclusion

Conclusion

- ▶ Trends of structure correspondence (except for T0769)
- ▶ Linear Correlation measures of energy and RMSD fail
- ▶ RMSD measure not reliable// ⇒ CASP uses multiple measures
- ▶ Function by (Zhang *et al.*, 1997) is fast! (only sums)
- ▶ But: 10 years old

Conclusion

- ▶ Trends of structure correspondence (except for T0769)
- ▶ Linear Correlation measures of energy and RMSD fail
- ▶ RMSD measure not reliable// ⇒ CASP uses multiple measures
- ▶ Function by (Zhang *et al.*, 1997) is fast! (only sums)
- ▶ But: 10 years old

Conclusion

- ▶ Trends of structure correspondence (except for T0769)
- ▶ Linear Correlation measures of energy and RMSD fail
- ▶ RMSD measure not reliable// ⇒ CASP uses multiple measures
- ▶ Function by (Zhang *et al.*, 1997) is fast! (only sums)
- ▶ But: 10 years old

Conclusion

- ▶ Trends of structure correspondence (except for T0769)
- ▶ Linear Correlation measures of energy and RMSD fail
- ▶ RMSD measure not reliable// \Rightarrow CASP uses multiple measures
- ▶ Function by (Zhang *et al.*, 1997) is fast! (only sums)
- ▶ But: 10 years old

Conclusion

- ▶ Trends of structure correspondence (except for T0769)
- ▶ Linear Correlation measures of energy and RMSD fail
- ▶ RMSD measure not reliable// \Rightarrow CASP uses multiple measures
- ▶ Function by (Zhang *et al.*, 1997) is fast! (only sums)
- ▶ But: 10 years old

Thank you for your attention!

Bibliography



C Zhang, G Vasmatzis, J L Cornette, and C DeLisi.

Determination of atomic desolvation energies from the structures of crystallized proteins.

Journal of molecular biology, 267(3):707–726, 1997.