Atomic Contact Energies

Max Emil Schön, Adrian Geißler

June 15, 2015

1 Introduction

The free energy functions needed to model Protein structures are made up of multiple components including van-der-Waals energies, electrostatics, dipole interactions, torsions and others. Here, we focus on atomic contact energies (ACE) of Proteins.
Zhang et al. (1997) describe a method to compute contact energies based on atoms rather than residue interactions, as proposed by Miyazawa and Jernigan (1996).

2 Material and Methods

The parsing/filtering of input `.pdb` files as well as the computation of RMSD was done using Biopython's module PDB (Hamelryck and Manderick, 2003).

2.1 Energy computation

We implemented the Computation of atomic contact energies as proposed by Zhang et al. (1997). Only heavy atoms were considered as potential pair members. We scanned every pair of atoms in a peptide chain for possible contact pairs. A contact pair is defined by Zhang et al. (1997) as two atoms that have a distance ≤ 6 Å and are located at least ten bonds of the backbone away from each other. The bond distance criterion is implemented in terms of residues and the atoms' connectivity classes (Appendix in Zhang et al. (1997)).
For every valid contact pair an energy is given by Table 1 in Zhang et al. (1997), which is based on the idea of grouping atoms into atom types with similar contact energy behaviour. As Zhang et al. (1997), we used 18 atom types. To compute the overall energy of a given structure, we summed over all such contact pairs.

2.2 Data

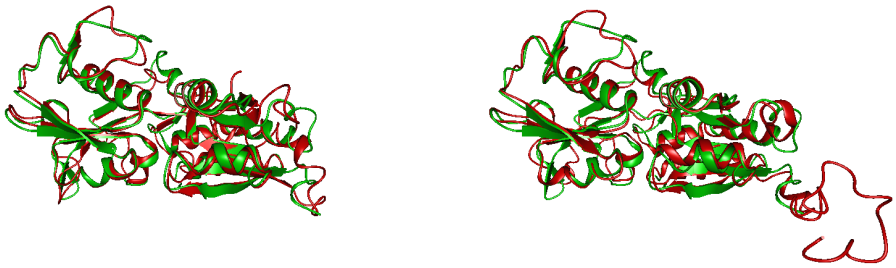
The Critical Assessment of protein Structure Prediction (CASP) experiments[1] provides a variety of protein structure predictions from different researchers worldwide. We used 4 Protein targets (T0762, T0769, T0776, T0784) and the respective predictions from the current CASP11 experiment to test our energy computation from 2.1.

2.3 RMSD

For the evaluation of the predicted structures we aligned the backbone Cα atoms of the prediction with those of the target experimental structure.

3 Results

For the target structure T0762, predictions 008 and 251 yielded the lowest ACE (-262.8469047619) and RMSD (2.0353168441281952), respectively (Fig. 1).



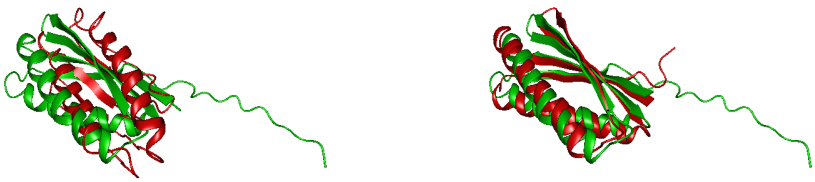
(a) (b)

Figure 1: Best scoring predictions of T0762 evaluated using (a) atomic contact energies (008) and (b) the RMSD between prediction and experimental structure (251). Both predictions were visualized in BALLView (Moll et al., 2006). green: target experimental stucture. red: predicted structure.

4 Discussion

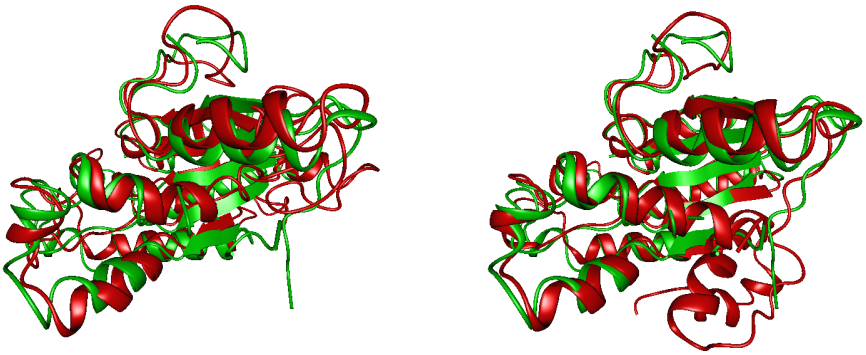
As shown by Fig. 5, RMSD values change little for the predictions which scored highest in the original CASP11 experiment. Together with the fact

[1] `http://www.predictioncenter.org/casp11` (last accessed: June 15, 2015)



(a) (b)

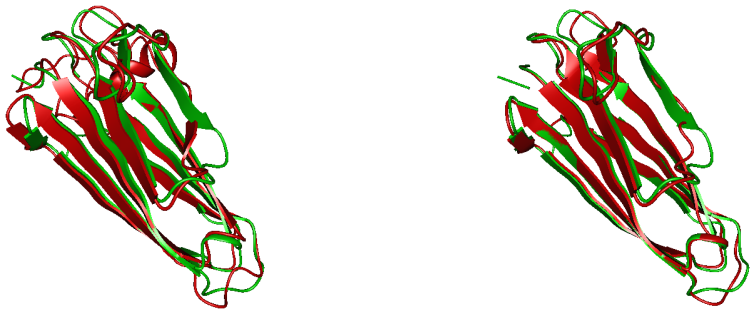
Figure 2: Best scoring predictions of T0769 evaluated using (a) atomic contact energies (442) and (b) the RMSD between prediction and experimental structure (241). Both predictions were visualized in BallView (Moll et al., 2006). green: target experimental stucture. red: predicted structure.



(a) (b)

Figure 3: Best scoring predictions of T0776 evaluated using (a) atomic contact energies (300) and (b) the RMSD between prediction and experimental structure (420). Both predictions were visualized in BallView (Moll et al., 2006). green: target experimental stucture. red: predicted structure.

that the prediction with lowest energy in no case possesed the lowest RMSD as well, this suggests that the RMSD measure is a rather poor assessment for prediction quality on its own. Further it seems that the best scoring predictions are not far away from each other in most cases.



(a) (b)

Figure 4: Best scoring predictions of T0776 evaluated using (a) atomic contact energies (117) and (b) the RMSD between prediction and experimental structure (156). Both predictions were visualized in BallView (Moll et al., 2006). green: target experimental stucture. red: predicted structure.

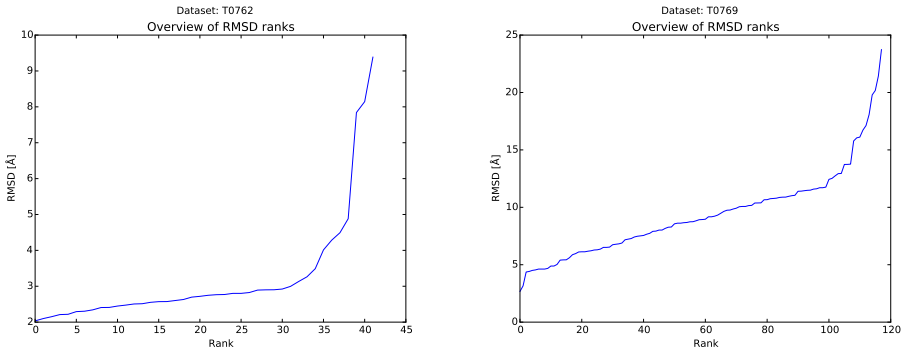
References

T. Hamelryck and B. Manderick. PDB file parser and structure class implemented in Python. *Bioinformatics*, 19(17):2308–2310, 2003.

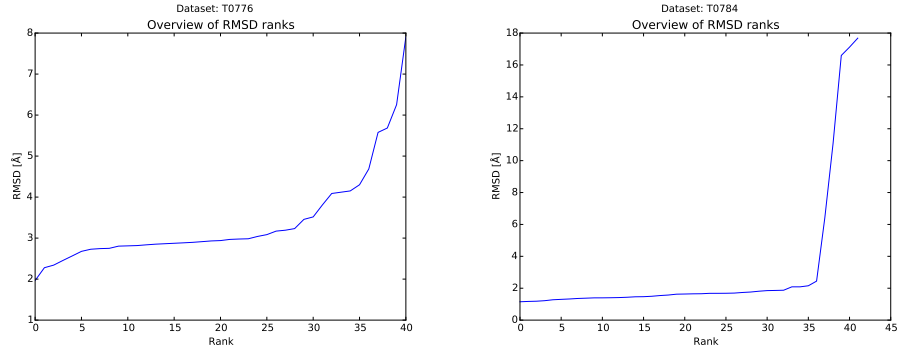
S. Miyazawa and R. L. Jernigan. Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. *Journal of molecular biology*, 256(3):623–644, 1996.

A. Moll, A. Hildebrandt, H.-P. Lenhof, and O. Kohlbacher. BALLView: a tool for research and education in molecular modeling. *Bioinformatics*, 22(3):365–366, 2006.

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.



(a) (b)



(c) (d)

Figure 5: Calculated RMSD values are plotted against the rank of the corresponding prediction in the original CASP11 exeperiment. (a) T0762 (b) T0769 (c) T0776 (d) T0784