Local Minima Hopping Along the Protein Energy Surface

Brian Olson[1] and Amarda Shehu[1,2]

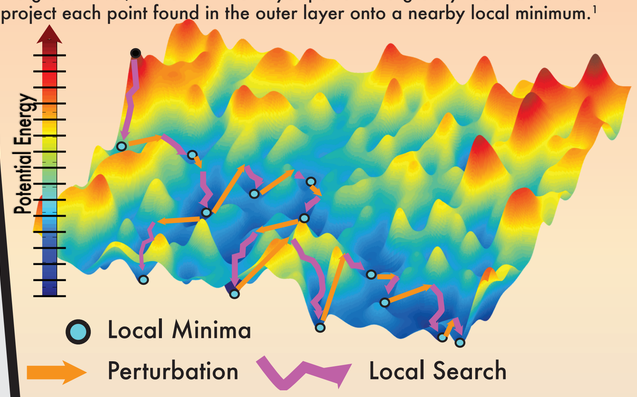
amarda@gmu.edu
http://www.cs.gmu.edu/~ashehu

[1]Department of Computer Science
[2]Department of Bioinformatics and Computational Biology
George Mason University, Fairfax, VA 22030

Protein models visualize the protein energy surface as a funnel with the native structure at the low-energy minimum. Current protein structure prediction algorithms seek the global minimum by searching for low-energy conformations in the hope that some of these reside in local minima near the native structure. The search techniques employed, however, fail to explicitly model these local minima. We propose a memetic algorithm which combines methods from evolutionary computation with cutting-edge structure prediction protocols. The Protein Local Optima Walk (PLOW) algorithm explores the space of local minima by explicitly projecting each move in the conformation space to a nearby local minimum. This allows PLOW to jump over local energy barriers and more effectively sample near-native conformations. Analysis across a broad range of proteins shows that PLOW outperforms an MMC-based method and compares favorably against other published ab-initio structure prediction algorithms[4,5].

Protein Local Optima Walk (PLOW)

PLOW employs a two layer search process to explore the space of local minima. The outer layer (see pseudo code) simulates an MMC search at the global level, while the inner layer performs a greedy local search to project each point found in the outer layer onto a nearby local minimum.[1]

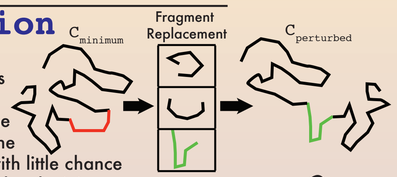


The canonical Iterated Local Search (ILS) framework is shown. This work defines domain-specific implementations of `InitialSelection`, `LocalSearch`, `Perturbation`, and `AcceptanceCriterion`.

```
1: Evalmax ← (Set runtime)
2: Evalcount ← 0
3: H ← InitialSelection()
4: while Evalcount < Evalmax do:
5:   Hnew ← Perturbation(H)
6:   Evalcount ← LocalSearch(Hnew,Evalcount)
7:   H ← AcceptanceCriterion(H, Hnew)
```

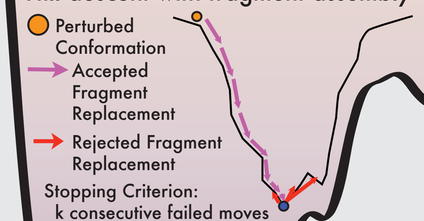
Perturbation

A single random trimer fragment replacement as the `Perturbation` function accomplishes the goal of jumping out of the current local minimum with little chance of returning to the same local minimum.



LocalSearch

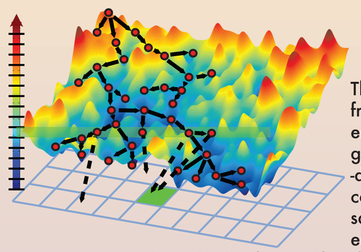
Hill descent with fragment assembly



InitialSelection AcceptanceCriterion

`InitialSelection` initializes H as a fully extended conformation projected onto a nearby local minimum using the `LocalSearch` function. `AcceptanceCriterion` uses the Metropolis Criterion to decide whether or not to move its current state to H_{new} or remain at the current value of `H`.

FeLTr



The FeLTr[3,4] framework seeks to ensure more geometrically -diverse conformational sampling by employing a geometric projection layer. The algorithm grows a search tree in the conformational space by expanding selected conformations with short fixed-length MMC trajectories and maintains a representative ensemble of previously visited conformations in memory. Selection from this ensemble is biased towards low-energy conformations and regions in under explored areas of the conformational space. In this way, FeLTr is able to dynamically redirect computational resources at the global level to ensure a degree of geometric diversity.

Experimental Results

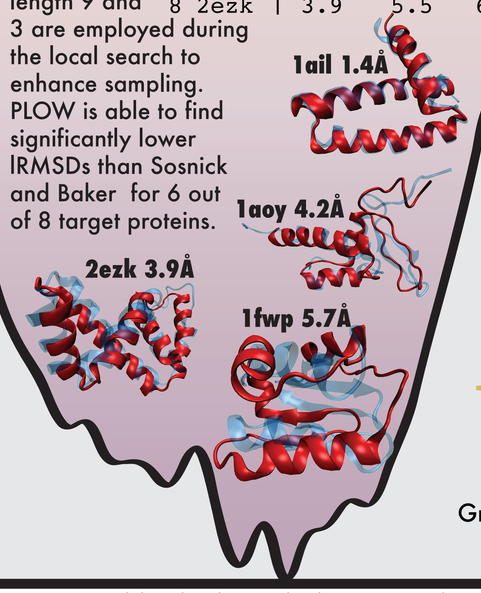
The Table compares the results from PLOW[1] to the FeLTr framework[3,4]. PLOW is able to find a structure more than 0.5Å lRMSD closer to the native structure than FeLTr for 9 out of the 12 target proteins. In the cases of 1aoy, 1csp, and 1fwp both algorithms, on average, find equivalent structures.

	PDBID	Length	fold	lowest lRMSD to native (Å) PLOW	FeLTr
1	1dtd	61	α/β	7.1	7.7
2	1sap	66	α/β	6.5	7.1
3	1csp	67	β	6.4	6.4
4	1wap	68	β	7.2	7.8
5	1fwp	69	α/β	6.5	6.8
6	1ail	70	α	2.7	4.7
7	1aoy	78	α/β	5.4	5.1
8	1cc5	83	α	5.5	6.4
9	1hhp	99	β	10.4	11.1
10	2ezk	93	α	4.6	6.4
11	2hq6	106	α/β	8.9	10.1
12	2h5n	123	α/β	7.0	9.0

Comparison to Other Groups

In this table the results obtained by PLOW are compared to those obtained by the Sosnick[5] and Baker[6] research groups. For these experiments fragments of both length 9 and 3 are employed during the local search to enhance sampling. PLOW is able to find significantly lower lRMSDs than Sosnick and Baker for 6 out of 8 target proteins.

	PDBID	lowest lRMSD to native (Å) PLOW	Sosnick	Baker
1	1dtdB	6.7	6.5	5.7
2	1sap	5.6	4.6	6.6
3	1wapA	7.0	8.0	7.7
4	1fwp	5.7	8.1	7.3
5	1ail	1.4	5.4	6.0
6	1aoy	4.2	5.7	5.7
7	1cc5	5.3	6.5	6.2
8	2ezk	3.9	5.5	6.6

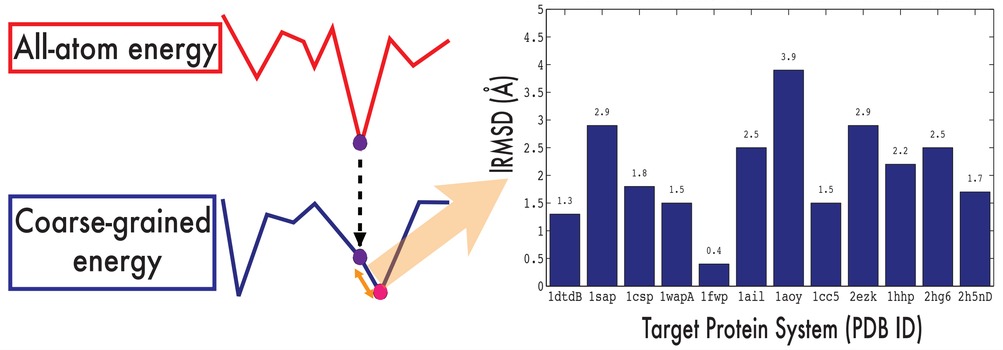


Conclusions

The Protein Local Optima Walk (PLOW) algorithm proposed here is a novel ab-initio structure prediction algorithm for effectively sampling local minima in the protein energy surface. The algorithm works by effectively projecting the search space onto the sub-space of local energy minima. By traversing only these local minima, PLOW is able to more effectively sample near-native conformations which are candidates for all-atom refinement in further studies. PLOW outperforms our previous work on a diverse set of target proteins and performs favorably when compared to published results from other research groups[4,5].

Many studies have demonstrated the effectiveness of an evolutionary framework to optimize intermediate conformations. However, these studies use overly simplified models, focus solely on optimization of an objective function, and fail to compare results with experimentally determined structures. Here we combine cutting-edge stochastic optimization strategies from evolutionary computation with established procedures for assembly of coarse-grained structures and analysis of results. Our results show this approach offers improved sampling at the coarse-grained level, the results of which may be further refined by additional studies.

Analysis of Local Minima



Grant No. 1016995

References

1. B. Olson and A. Shehu, "Populating Local Minima in the Protein Conformational Space." IEEE International Conference on Bioinformatics and Biomedicine (IEEE BIBM), Atlanta, GA, 2011.

2. B. Olson "Local Minima Hopping in the Protein Energy Surface", submitted.

3. A. Shehu and B. Olson, "Guiding the search for native-like protein conformations with an ab-initio tree-based exploration," Int. J. Robot. Res., vol. 29, no. 8, pp. 1106–11 227, 2010.

4. B. Olson, K. Molloy, and A. Shehu, "In search of the protein native state with a probabilistic sampling approach," J. Bioinf. and Comp. Biol., vol. 9, no. 3, pp. 383–398, 2011.

5. J. DeBartolo, A. Colubri, A. K. Jha, J. E. Fitzgerald, K. F. Freed, and T. R. Sosnick, "Mimicking the folding pathway to improve homology-free protein structure predic- tion," Proc. Natl. Acad. Sci. USA, vol. 106, no. 10, pp. 3734–3739, 2009.

6. J. Meiler and D. Baker, "Coupled prediction of protein secondary and tertiary structure," Proceedings of the National Academy of Sciences of the United States of America, vol. 100, no. 21, pp. 12 105–12 110, 2003.