BIOINFORMATICS

Vol. 00 no. 00 2009
Pages 1

Flexible structural protein alignment by a sequence of local transformations: Supplementary material

Jairo Rocha[a], Joan Segura[a], Richard C. Wilson[b], Swagata Dasgupta[c]

[a]Dept. of Mathematics and Computer Science, University of the Balearic Islands, Palma, E07122, Spain,
[b]Dept. of Computer Science, University of York, YO1 0 5DD, UK,
[c]Dept. of Chemistry, Indian Institute of Technology, Kharagpur, 721 302, India

1 MORE EXAMPLES OF FLEXIBLE ALIGNMENTS

Here, we show two additional examples also suggested in Yuzhen and Godzik, 2003, pag. 253, and other two illustrative examples.

When comparing the L-2-haloacid dehalogenase (PDB code d1zrn_) and the β-phosphoglucomutase (d1lvha_) (Fig. 1), it is obvious that the helices in the lower left region and the upper right are far better aligned when the deformation, represented by the local transformations, is allowed in comparison to the rigid superposition. Both proteins belong to the HAD-like superfamily but from different families based on SCOP classification. However, ProtDeform, as Fatcat does, matches more than 90% of both structures suggesting that the families are close to each other.

We show in Figure 2 the comparison of d1zrna against d1feza_ (phosphonoacetaldehyde hydrolase) also in the HAD-like superfamily. It is apparent from the superposed structures that the alignment is much easier to analyse using the multiple transformation approach. Also, the superposed figures do look good and distinctions are possible between the rigid and deformed transformations clearly indicating that it is desirable to allow for the deformations for a better alignment.

The first of the other two examples is the pair of aldolase class I domains (CATH 3.20.20.130), 1b3oB0 and 1zfjA0, with 476 and 410 amino acids, respectively. ProtDeform aligns 385 pairs, including the whole TIM barrel. Figure 3 shows that the most part of 1b3oB0 is aligned by a single rigid transformation, but a part has a large rotation with respect to the main part. The alignment has a RMSD of 9.1 Å. We use a modified formula for RMSD, that takes into account the local transformations. If in the formula for RMSD we use the local transformation at each matched site instead of the global one, we get what we call the *local RMSD*, lRMSD:

$$\text{lRMSD} = \sqrt{\frac{1}{N_{assig}} \sum_{i \in Dom(f)} \|a_i - T_i^f(f(a_i))\|^2}.$$

In this case, the lRMSD is 2.1. Notice that as the number of neighbours considered grows, lRMSD gets closer to RMSD.

The second example is the pair of Type I PLP-dependent aspartate aminotransferase-like domains 1c4kA2 and 1cj0B2 (CATH 3.40.640.10). ProtDeform aligns 211 of the 313 and 271 sites, respectively, including the 3-layer (aba) sandwich. The alignment has a RMSD of 4.4 and an lRMSD of 2.8. As shown in Figure 4, one part of 1cj0B2 is rotated with respect to the main part.

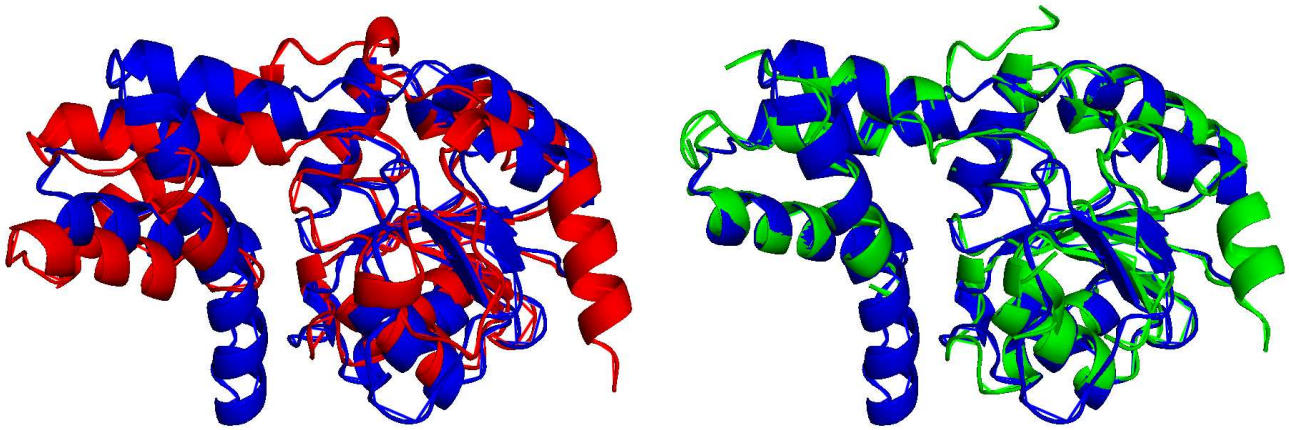

Fig. 1. Comparisons of two proteins from HAD-like superfamily. Left: the best rigid superposition for the matching of the L-2-haloacid dehalogenase (SCOP code d1zrn_, blue ribbons) and the β-phosphoglucomutase (SCOP code d1lvha_, red and green ribbons). Right: a better superposition suggested by the transformations.

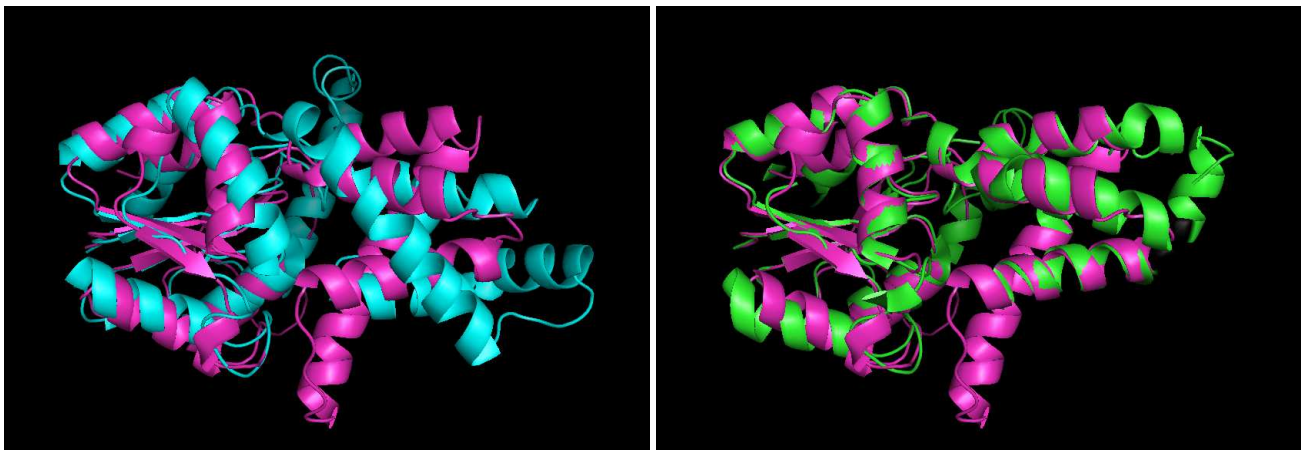

Fig. 2. Comparisons of two proteins from HAD-like superfamily. Left panel: the best rigid superposition for the matching of L-2-haloacid dehalogenase (SCOP code d1zrn_, red ribbons) and phosphonoacetaldehyde hydrolase (SCOP code d1feza, blue ribbons). Right panel: the superposition suggested by the transformations, with the deformed one now in green.

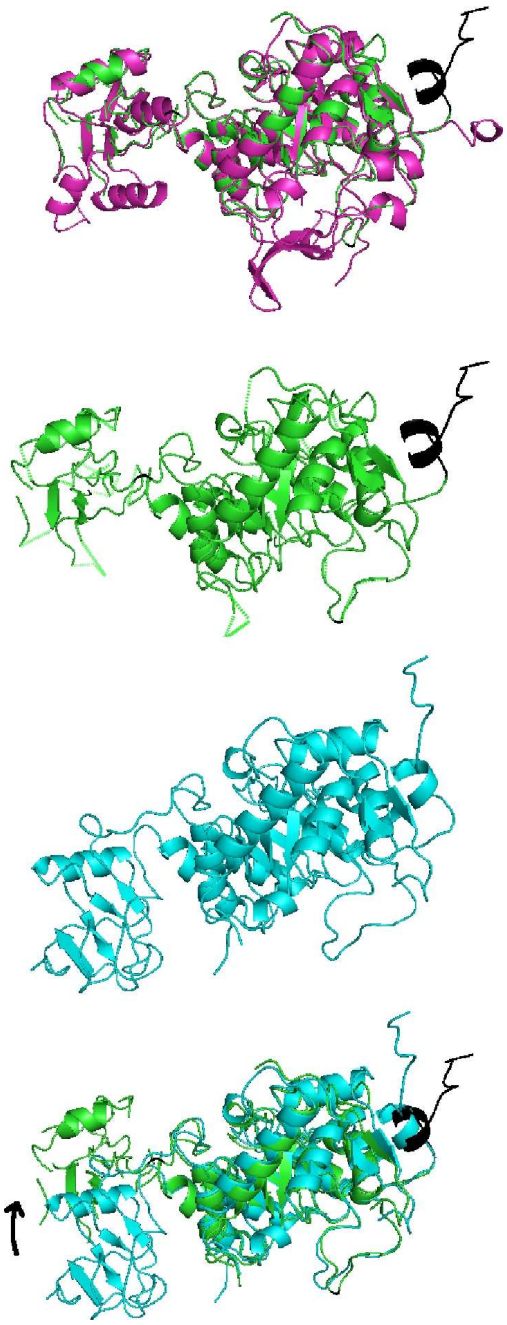

Fig. 3. From top to bottom: First, alignment of the deformed domain 1b3oB0 (green) and 1zfjA0 (red). Second, the same domain 1b3oB0 deformed, where black ribbons are not matched and dotted lines join sites that where originally adjacent. Third, the same domain 1b3oB0 without deformation. Forth, superposition of the rigid (green) and the deformed (blue) domain; the arrow shows the local rotation.

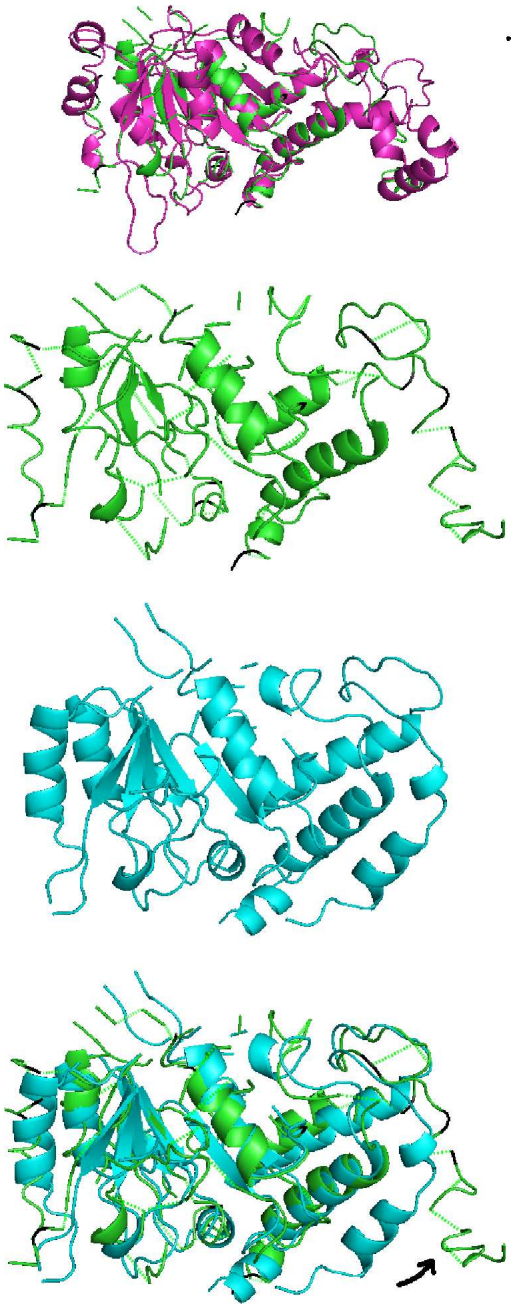

Fig. 4. From top to bottom: First, alignment of the deformed domain 1cj0B2 (green) and 1c4kA2 (red). Second, the same domain 1cj0B2 deformed. Third, the same domain 1cj0B2 without deformation. Forth, superposition of the rigid (green) and the deformed (blue) domains; the arrow shows the local rotation.