Cross-Species Analysis of the Glycolytic Pathway by Comparison of Molecular Interaction Fields

Matthias Stein, Razif R. Gabdoulline, Rebecca C. Wade

Supplementary Information

Table of Contents
1. Choice of protein sequences and structural templates
2. Evaluation of the choice of PDB template on the similarity of electrostatic potentials
3. Visualisation of distance matrices of electrostatic potentials
4. Discussion of means to compare electrostatic potentials
1. Choice of protein sequences and structural templates

**Step 01 Template Used 1V4S**
- Q91753_XENLA
- Q8AYP7_CHICK
- HXK2_DROME
- Q5TW28_ANOGA
- Q6NX09_BRARE
- HXK4_HUMAN
- HXK4_MOUSE
- HXK2_ARATH
- HXK4_RAT
- Q6Z398_ORYSA
- Q5RC71_PONPY

**Step 02 Template Used 1HOX**
- G6PI_PONPY
- G6PI_DROME
- Q6GP33_XENLA
- Q5ZMU3_CHICK
- Q8QFU1_BRARE
- G6PI_MOUSE
- Q6P6V0_RAT
- G6PIA_ORYSA
- G6PI_HUMAN
- Q7PVA2_ANOGA
- G6PI_ARATH

**Step 03 Template Used 4PFK**
- K6PL_HUMAN
- Q5R7V5_PONPY
- K6PL_RAT
- K6PL_MOUSE
- Q90YA3_CHICK
- Q6DD69_XENLA
- Q66HV8_BRARE
- K6PF_DROME
- Q7QJT0_ANOGA
- Q5SNH5_ORYSA
- Q9M0F9_ARATH

**Step 04 Template Used 1FDJ**
- ALDOB_CHICK
- ALDOB_DANRE
- ALDOB_HUMAN
- ALDOB_MOUSE
- ALDOB_PONPY
- ALDOB_RAT
- ALF_ARATH
- ALF_DROME
- ALF_ORYSJ
- Q5XHC6_XENLA
- Q7PG19_ANOGA

**Step 05 Template Used 1R2R**
- Q7PXW5_ANOGA_xp_321467.2
- TPIS_DROME
- Q90XG0_DANRE
- TPIS_CHICK
- Q7ZWNS_XENLA
- TPIS_RAT
- Q5R928_PONPY
TPIS_HUMAN
TPIS_MOUSE
TPIS_ORYSJ
TPIS_ARATH

**Step 06 Template Used 1NQO**
G3PB_ARATH
G3P_MOUSE
G3P_CHICK
Q7Q1U8_ANOGA
G3P2_DROME
Q5RAB4_PONPY
G3P2_HUMAN
G3P_XENLA
Q6NYM9_BRARE
G3PC_ORYSA

**Step 07 Template Used 1HDI**
PGK1_HUMAN
PGK1_PONPY
PGK1_MOUSE
PGK1_RAT
PGK_CHICK
Q7ZV29_BRARE
PGK_XENLA
PGK_ANOGA
PGK_DROME
Q655T1_ORYSA
PGKH_ARATH

**Step 08 Template Used 1YFK**
PGAM1_HUMAN
PGAM1_RAT
PGAM1_MOUSE
PGAM1_PONPY
Q6TNR9_BRARE
PGAM1_CHICK
Q6AZP8_XENLA
Q9VAN7_DROME
Q7PXI5_ANOGA
Q6Z8JQ_ORYSA
Q9LM13_ARATH

**Step 09 Template Used 2ONE**
ENO_DROME
Q7Q3D8_ANOGA
Q6TH14_BRARE
ENOB_HUMAN
ENOA_CHICK
ENOB_MOUSE
Q5R6Y1_PONPY
ENOB_RAT
ENO_XENLA
ENO_ORYSA
ENO_ARATH

**Step 10 Template Used 1LIU**
KPYK_DROME
Q7PPE7_ANOGA
KPYR_HUMAN
KPYR_MOUSE
KPYR_RAT
KPYK_CHICK
2. Evaluation of the choice of PDB template on the similarity of electrostatic potentials

The PGI from rabbit (PDB entry 1HOX) and the protein structure from human autocrine motility factor (PDB entry 1JIQ) were used as the structural templates (the RMSD between the two crystal structures is 0.63 Å for backbone atoms and 0.91 Å for all atoms; using SuperPose1.0 (Maiti R, van Domselaar GH, Zhang H., Wishart DS (2004) SuperPose: a simple server for sophisticated structural superposition. Nucl. Acids Res. 32: W590-W594). We repeated the multiple sequence alignment of the eleven model PGI sequences using ClustalW, used the two PDB entries as the template structures, calculated the electrostatic potentials with UHBD and ranked the PGI enzymes according to their similarity to the human enzyme (Table SI).

The percentage sequence identity between the modelled sequences and the templates is high for both PDB templates (see Table SI). The overall rather high sequence identity for PGIs enables the generation of protein structural models with confidence. The lowest sequence identity to the human enzyme is found for the PGI from rice (47% and 46%, respectively). The Hodgkin similarity indices vary very little whether using the PGI from rabbit or human as protein templates. The largest difference in similarity indices is found for the PGI from rice which varies from 0.34 (1HOX template) to 0.46 (1JIQ template) (see Table SI). This value of about 0.1 may thus be considered as an estimate of the uncertainty in Hodgkin SIs when discussing the relation between two proteins in terms of their electrostatic potentials.

The PGI from mosquito (Anopheles gambiae) is less affected by the choice of protein template structure. The electrostatic potential shows no correlation with that of the human enzyme when using either the rabbit or the human protein as a template structure (Hodgkin SI 0.07 and 0.04, respectively).

**Table SI**: Comparison of the calculated Hodgkin similarity indices of the electrostatic potentials for glucose 6-phosphate isomerases (PGI) using two different template structures: The PGI from Oryctolagus cuniculus (rabbit; PDB entry 1HOX) and Homo sapiens (PDB entry 1JIQ).

<table>
<thead>
<tr>
<th>PDB Template</th>
<th>PDB Template</th>
</tr>
</thead>
<tbody>
<tr>
<td>1HOX (Rabbit)</td>
<td>1JIQ (Human)</td>
</tr>
</tbody>
</table>
3. Visualisation of distance matrices of electrostatic potentials

<table>
<thead>
<tr>
<th>Step</th>
<th>Entire Protein</th>
<th>Active Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- 3. Visualisation of distance matrices of electrostatic potentials
4. Discussion of means to compare electrostatic potentials

The Hodgkin similarity index measures relative similarities of two three-dimensional fields. The scalar product of fields \((2 \cdot a \cdot b)\) is divided by the sum of squares \((a^2 + b^2)\) and thus one obtains a relative measure of the similarities of molecular interaction fields. For example for enolases (step 09) one observes rather large electrostatic potential differences between the model species (Figure 9C), but Hodgkin similarity indices close to unity (Figure 9B) because of the large (and negative) magnitude of the electrostatic potential around the active site.
(Figure 9D). The same argument holds for PGIs (step 02) which exhibit a large (and positive) electrostatic potential around the active site (Figure 9D) and large electrostatic potential differences between the species (Figure 9C) but Hodgkin similarity indices which are between 0.7 and 1.0 (Figure 9B). For triosephosphate isomerases (step 05), the electrostatic potential at the active site is very small in magnitude (Figure 9D). Thus relatively small differences in electrostatic potentials around the active site between the species (Figure 9C) lead to larger differences in Hodgkin similarity indices (Figure 9B) for that region. An absolute electrostatic potential difference (Figure 9C) gives a more quantitative means of comparing enzyme kinetic parameters between different species. It is the absolute difference in electrostatic potentials that is, for example important for stabilizing the transition state, and can therefore be related to species-to-species variations in kinetic parameters.