1 Introduction

From the original Protein Data Bank entry (PDB id 1at1):

Title: Crystal structures of phosphonoacetamide ligated t and phosphonoacetamide and malonate ligated r states of aspartate carbamoyltransferase at 2.8-angstroms resolution and neutral pH

Compound: Mol id: 1; molecule: aspartate carbamoyltransferase, catalytic chain; chain: a, c; ec: 2.1.3.2; engineered: yes; mol id: 2; molecule: aspartate carbamoyltransferase regulatory chain; chain: b, d; engineered: yes

Organism, scientific name: Escherichia Coli

1at1 contains unique chains 1at1A (310 residues) and 1at1B (146 residues) 1at1C is a homologue of chain 1at1A. 1at1D is a homologue of chain 1at1B.

2 Chain 1at1A

2.1 P0A789 overview

2.2 Multiple sequence alignment for 1at1A

2.3 Residue ranking in 1at1A

2.4 Top ranking residues in 1at1A and their position on the structure

2.4.1 Clustering of residues at 25% coverage.

2.4.2 Overlap with known functional surfaces at 25% coverage.

3 Chain 1at1B

3.1 P0A789 overview

3.2 Multiple sequence alignment for 1at1B

3.3 Residue ranking in 1at1B

3.4 Top ranking residues in 1at1B and their position on the structure

3.4.1 Clustering of residues at 25% coverage.

3.4.2 Overlap with known functional surfaces at 25% coverage.

3.4.3 Possible novel functional surfaces at 25% coverage.

4 Notes on using trace results

4.1 Coverage

4.2 Known substitutions

4.3 Surface

4.4 Number of contacts

4.5 Annotation

4.6 Mutation suggestions

5 Appendix

5.1 File formats

5.2 Color schemes used

5.3 Credits

5.3.1 Alistat

5.3.2 CE

5.3.3 DSSP

5.3.4 HSSP

5.3.5 LaTex

5.3.6 Muscle

5.3.7 Pymol

5.4 Note about ET Viewer

5.5 Citing this work

5.6 About report_maker

5.7 Attachments

Lichtarge lab 2006
Similarity: Belongs to the ATCase/OTCase family.

About: This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

2.2 Multiple sequence alignment for 1at1A

For the chain 1at1A, the alignment 1at1A.msf (attached) with 220 sequences was used. The alignment was downloaded from the HSSP database, and fragments shorter than 75% of the query as well as duplicate sequences were removed. It can be found in the attachment to this report, under the name of 1at1A.msf. Its statistics, from the alistat program are the following:

| Format: MSF |
| Number of sequences: 220 |
| Total number of residues: 64871 |
| Smallest: 258 |
| Largest: 310 |
| Average length: 294.9 |
| Alignment length: 310 |
| Average identity: 37% |
| Most related pair: 99% |
| Most unrelated pair: 18% |
| Most distant seq: 40% |

Furthermore, 2% of residues show as conserved in this alignment. The alignment consists of 21% eukaryotic (2% vertebrata, 1% arthropoda, 6% fungi, 3% plantae), 68% prokaryotic, 10% archaean, and <1% viral sequences. (Descriptions of some sequences were not readily available.) The file containing the sequence descriptions can be found in the attachment, under the name 1at1A.descr.

2.3 Residue ranking in 1at1A

The 1at1A sequence is shown in Figs. 1–2, with each residue colored according to its estimated importance. The full listing of residues in 1at1A can be found in the file called 1at1A.ranks, sorted in the attachment.

2.4 Top ranking residues in 1at1A and their position on the structure

In the following we consider residues ranking among top 25% of residues in the protein. Figure 3 shows residues in 1at1A colored by their importance: bright red and yellow indicate more conserved/important residues (see Appendix for the coloring scheme). A Pymol script for producing this figure can be found in the attachment.
Fig. 4. Residues in 1at1A, colored according to the cluster they belong to: red, followed by blue and yellow are the largest clusters (see Appendix for the coloring scheme). Clockwise: front, back, top and bottom views. The corresponding Pymol script is attached.

Table 1. continued

<table>
<thead>
<tr>
<th>cluster color</th>
<th>size</th>
<th>member residues</th>
</tr>
</thead>
<tbody>
<tr>
<td>red</td>
<td>127, 128, 129, 130, 131, 134, 135, 136</td>
<td></td>
</tr>
<tr>
<td></td>
<td>137, 138, 139, 140, 141, 144, 146, 161</td>
<td></td>
</tr>
<tr>
<td></td>
<td>162, 165, 166, 167, 168, 169, 170</td>
<td></td>
</tr>
<tr>
<td></td>
<td>171, 172, 176, 226, 228, 229, 231</td>
<td></td>
</tr>
<tr>
<td></td>
<td>233, 234, 235, 236, 265, 266, 267</td>
<td></td>
</tr>
<tr>
<td></td>
<td>268, 269, 272, 273, 278, 285, 288</td>
<td></td>
</tr>
<tr>
<td></td>
<td>289, 292, 296, 297, 298, 300</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Clusters of top ranking residues in 1at1A.

2.4.2 Overlap with known functional surfaces at 25% coverage.
The name of the ligand is composed of the source PDB identifier and the heteroatom name used in that file.

Interface with 1at1A2. Table 2 lists the top 25% of residues at the interface with 1at1A2. The following table (Table 3) suggests possible disruptive replacements for these residues (see Section 4.6).

Table 2. The top 25% of residues in 1at1A at the interface with 1at1A2.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>subst's (%)</th>
<th>cvg</th>
<th>noc/ bb</th>
<th>dist (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>S</td>
<td>S(92)I</td>
<td>0.12</td>
<td>36/14</td>
<td>2.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>L(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>T(3)G</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>T</td>
<td>T(88)</td>
<td>0.14</td>
<td>1/0</td>
<td>4.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S(8)AMI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>V(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>D</td>
<td>N(5)</td>
<td>0.21</td>
<td>12/1</td>
<td>2.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D(86)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>R(1)T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>S(1)Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>E(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Q(1)HA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>L</td>
<td>S(70)</td>
<td>0.24</td>
<td>1/1</td>
<td>4.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F(1)H</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>L(9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Q(4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>A(4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>V(3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>I(2)T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M(1)K</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>Y</td>
<td>I(8)</td>
<td>0.25</td>
<td>77/15</td>
<td>3.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Y(53)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M(23)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>L(4)K</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>H(3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>A(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>V(1)S</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. continued

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>subst’s (%)</th>
<th>cvg</th>
<th>noc/ bb</th>
<th>dist (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>S</td>
<td>S(92)I</td>
<td>0.12</td>
<td>36/14</td>
<td>2.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>L(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>T(3)G</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>T</td>
<td>T(88)</td>
<td>0.14</td>
<td>1/0</td>
<td>4.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S(8)AMI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>V(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>D</td>
<td>N(5)</td>
<td>0.21</td>
<td>12/1</td>
<td>2.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D(86)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>R(1)T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>S(1)Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>E(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Q(1)HA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>L</td>
<td>S(70)</td>
<td>0.24</td>
<td>1/1</td>
<td>4.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F(1)H</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>L(9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Q(4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>A(4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>V(3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>I(2)T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M(1)K</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>Y</td>
<td>I(8)</td>
<td>0.25</td>
<td>77/15</td>
<td>3.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Y(53)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M(23)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>L(4)K</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>H(3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>A(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>V(1)S</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. continued

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>subst’s (%)</th>
<th>cvg</th>
<th>noc/ bb</th>
<th>dist (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>S</td>
<td>S(92)I</td>
<td>0.12</td>
<td>36/14</td>
<td>2.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>L(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>T(3)G</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>T</td>
<td>T(88)</td>
<td>0.14</td>
<td>1/0</td>
<td>4.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S(8)AMI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>V(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>D</td>
<td>N(5)</td>
<td>0.21</td>
<td>12/1</td>
<td>2.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D(86)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>R(1)T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>S(1)Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>E(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Q(1)HA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>L</td>
<td>S(70)</td>
<td>0.24</td>
<td>1/1</td>
<td>4.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F(1)H</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>L(9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Q(4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>A(4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>V(3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>I(2)T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M(1)K</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>Y</td>
<td>I(8)</td>
<td>0.25</td>
<td>77/15</td>
<td>3.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Y(53)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M(23)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>L(4)K</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>H(3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>A(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>V(1)S</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>86</td>
<td>E</td>
<td>(FWH) (R) (YVA) (KCG)</td>
</tr>
<tr>
<td>90</td>
<td>D</td>
<td>(R) (FWH) (K) (YVA)</td>
</tr>
<tr>
<td>84</td>
<td>K</td>
<td>(Y) (FW) (T) (D)</td>
</tr>
<tr>
<td>85</td>
<td>G</td>
<td>(FW) (HR) (Y) (K)</td>
</tr>
<tr>
<td>80</td>
<td>S</td>
<td>(R) (K) (H) (FW)</td>
</tr>
<tr>
<td>91</td>
<td>T</td>
<td>(R) (K) (H) (Q)</td>
</tr>
<tr>
<td>100</td>
<td>D</td>
<td>(R) (FW) (H) (Y)</td>
</tr>
<tr>
<td>81</td>
<td>L</td>
<td>(Y) (R) (H) (T)</td>
</tr>
<tr>
<td>98</td>
<td>Y</td>
<td>(K) (QR) (E) (M)</td>
</tr>
</tbody>
</table>

Table 3. List of disruptive mutations for the top 25% of residues in 1at1A, that are at the interface with 1at1A2.
**Table 4. continued**

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>subst's (%)</th>
<th>cvg</th>
<th>noc/ bb</th>
<th>dist (Å)</th>
<th>antn</th>
</tr>
</thead>
<tbody>
<tr>
<td>233</td>
<td>E</td>
<td>MEAG(2)</td>
<td>0.14</td>
<td>13/0</td>
<td>3.61</td>
<td></td>
</tr>
<tr>
<td>269</td>
<td>R</td>
<td>R(71)</td>
<td>0.18</td>
<td>17/3</td>
<td>2.62</td>
<td></td>
</tr>
<tr>
<td>289</td>
<td>A</td>
<td>M(21)</td>
<td>0.21</td>
<td>13/7</td>
<td>3.94</td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>R</td>
<td>K(5)</td>
<td>0.23</td>
<td>10/0</td>
<td>2.83</td>
<td></td>
</tr>
</tbody>
</table>

**Table 4.** The top 25% of residues in 1at1A at the interface with 1at1A1. (Field names: res: residue number in the PDB entry; type: amino acid type; subst: substitutions seen in the alignment; with the percentage of each type in the bracket; noc(bb): number of contacts with the ligand, with the number of contacts realized through backbone atoms given in the bracket; dist: distance of closest approach to the ligand.)

---

**Table 5.** List of disruptive mutations for the top 25% of residues in 1at1A, that are at the interface with 1at1A1.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>S</td>
<td>(KR) (FQMWH) (NYELPI) (D)</td>
</tr>
<tr>
<td>54</td>
<td>R</td>
<td>(TD) (SYEVCLAPig) (FMW) (N)</td>
</tr>
<tr>
<td>53</td>
<td>T</td>
<td>(KR) (FQMWH) (NELPI) (D)</td>
</tr>
<tr>
<td>50</td>
<td>E</td>
<td>(FW) (H) (Y) (VCAG)</td>
</tr>
<tr>
<td>58</td>
<td>S</td>
<td>(KR) (QH) (FMW) (E)</td>
</tr>
<tr>
<td>268</td>
<td>P</td>
<td>(Y) (R) (T) (E)</td>
</tr>
<tr>
<td>229</td>
<td>R</td>
<td>(D) (E) (TY) (LPi)</td>
</tr>
<tr>
<td>267</td>
<td>L</td>
<td>(R) (Y) (H) (K)</td>
</tr>
<tr>
<td>234</td>
<td>R</td>
<td>(Y) (T) (D) (S)</td>
</tr>
<tr>
<td>233</td>
<td>E</td>
<td>(H) (FYw) (R) (CG)</td>
</tr>
<tr>
<td>269</td>
<td>R</td>
<td>(TY) (D) (E) (SCG)</td>
</tr>
<tr>
<td>289</td>
<td>A</td>
<td>(R) (KY) (E) (H)</td>
</tr>
<tr>
<td>56</td>
<td>R</td>
<td>(D) (T) (YE) (S)</td>
</tr>
</tbody>
</table>

**Table 5.** List of disruptive mutations for the top 25% of residues in 1at1A, that are at the interface with 1at1A1.

---

Figure 5 shows residues in 1at1A colored by their importance, at the interface with 1at1A2. 1at1A2 is shown in backbone representation (See Appendix for the coloring scheme for the protein chain 1at1A.)

**Interface with 1at1A1.** Table 4 lists the top 25% of residues at the interface with 1at1A1. The following table (Table 5) suggests possible disruptive replacements for these residues (see Section 4.6).

---

Figure 6 shows residues in 1at1A colored by their importance, at the interface with 1at1A1.
Maltose binding site. Table 6 lists the top 25% of residues at the interface with 1at1AMAL312 (maltose). The following table (Table 7) suggests possible disruptive replacements for these residues (see Section 4.6).

**Table 6.** The top 25% of residues in 1at1A at the interface with maltose. (Field names: res: residue number in the PDB entry; type: amino acid type; subst's: substitutions seen in the alignment; with the percentage of each type in the bracket; noc/bb: number of contacts with the ligand, with the number of contacts realized through backbone atoms given in the bracket; dist: distance of closest approach to the ligand.)

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>subst's</th>
<th>cvg</th>
<th>noc/bb</th>
<th>dist (Å)</th>
<th>antn</th>
</tr>
</thead>
<tbody>
<tr>
<td>134</td>
<td>H</td>
<td>H(100)</td>
<td>0.03</td>
<td>9/0</td>
<td>3.18</td>
<td>site</td>
</tr>
<tr>
<td>137</td>
<td>Q</td>
<td>Q(100)</td>
<td>0.03</td>
<td>1/0</td>
<td>4.99</td>
<td>site</td>
</tr>
<tr>
<td>167</td>
<td>P</td>
<td>P(95)</td>
<td>0.07</td>
<td>21/0</td>
<td>3.14</td>
<td>site</td>
</tr>
<tr>
<td>268</td>
<td>R</td>
<td>R(96)H</td>
<td>0.07</td>
<td>5/1</td>
<td>3.91</td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>R</td>
<td>R(95)VI</td>
<td>0.08</td>
<td>3/0</td>
<td>4.06</td>
<td>site</td>
</tr>
<tr>
<td>229</td>
<td>R</td>
<td>R(94)</td>
<td>0.12</td>
<td>15/0</td>
<td>3.13</td>
<td>site</td>
</tr>
<tr>
<td>267</td>
<td>L</td>
<td>L(10)</td>
<td>0.12</td>
<td>4/4</td>
<td>3.82</td>
<td>site</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A(10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>L(52)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G(32)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>S(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>V(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>234</td>
<td>R</td>
<td>R(93)</td>
<td>0.13</td>
<td>1/0</td>
<td>4.45</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>KQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MEAG(2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>168</td>
<td>T</td>
<td>V(48)</td>
<td>0.15</td>
<td>6/1</td>
<td>3.41</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>T(45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>P(3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 7.** List of disruptive mutations for the top 25% of residues in 1at1A, that are at the interface with maltose.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>134</td>
<td>H</td>
<td>(E) (TQMD) (SNKVCILA) (PIG) (YR)</td>
</tr>
<tr>
<td>137</td>
<td>Q</td>
<td>(Y) (FTWH) (SVCA) (G)</td>
</tr>
<tr>
<td>167</td>
<td>R</td>
<td>(TY) (D) (E) (SCG)</td>
</tr>
<tr>
<td>268</td>
<td>R</td>
<td>(Y) (R) (T) (E)</td>
</tr>
<tr>
<td>105</td>
<td>R</td>
<td>(TYD) (E) (SCG) (FVLAWPI)</td>
</tr>
<tr>
<td>229</td>
<td>R</td>
<td>(D) (E) (TY) (LPI)</td>
</tr>
<tr>
<td>267</td>
<td>L</td>
<td>(R) (Y) (H) (K)</td>
</tr>
<tr>
<td>234</td>
<td>R</td>
<td>(Y) (T) (D) (S)</td>
</tr>
<tr>
<td>168</td>
<td>T</td>
<td>(R) (K) (H) (FEQW)</td>
</tr>
<tr>
<td>231</td>
<td>Q</td>
<td>(Y) (H) (FW) (T)</td>
</tr>
</tbody>
</table>

**Phosphonoacetamide binding site.** Table 8 lists the top 25% of residues at the interface with 1at1APCT311 (phosphonoacetamide). The following table (Table 9) suggests possible disruptive replacements for these residues (see Section 4.6).

**Table 8.**

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>subst's (%)</th>
<th>cvg</th>
<th>noc/bb</th>
<th>dist (Å)</th>
<th>antn</th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>S</td>
<td>S(100)</td>
<td>0.03</td>
<td>17/9</td>
<td>2.68</td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>R</td>
<td>R(100)</td>
<td>0.03</td>
<td>38/11</td>
<td>2.87</td>
<td>site</td>
</tr>
<tr>
<td>55</td>
<td>T</td>
<td>T(100)</td>
<td>0.03</td>
<td>21/9</td>
<td>2.72</td>
<td>site</td>
</tr>
<tr>
<td>134</td>
<td>H</td>
<td>H(100)</td>
<td>0.03</td>
<td>11/0</td>
<td>2.91</td>
<td></td>
</tr>
<tr>
<td>137</td>
<td>Q</td>
<td>Q(100)</td>
<td>0.03</td>
<td>9/0</td>
<td>2.67</td>
<td>site</td>
</tr>
<tr>
<td>53</td>
<td>T</td>
<td>T(96)</td>
<td>0.04</td>
<td>15/11</td>
<td>3.07</td>
<td></td>
</tr>
<tr>
<td>134</td>
<td>H</td>
<td>H(100)</td>
<td>0.03</td>
<td>11/0</td>
<td>2.91</td>
<td></td>
</tr>
<tr>
<td>137</td>
<td>Q</td>
<td>Q(100)</td>
<td>0.03</td>
<td>9/0</td>
<td>2.67</td>
<td>site</td>
</tr>
<tr>
<td>53</td>
<td>T</td>
<td>T(96)</td>
<td>0.04</td>
<td>15/11</td>
<td>3.07</td>
<td></td>
</tr>
<tr>
<td>266</td>
<td>P</td>
<td>P(94)</td>
<td>0.06</td>
<td>7/6</td>
<td>3.46</td>
<td>site</td>
</tr>
<tr>
<td>268</td>
<td>P</td>
<td>P(96)</td>
<td>0.07</td>
<td>5/1</td>
<td>4.37</td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>R</td>
<td>R(99)VI</td>
<td>0.08</td>
<td>15/0</td>
<td>2.57</td>
<td>site</td>
</tr>
</tbody>
</table>

Figure 7 shows residues in 1at1A colored by their importance, at the interface with 1at1AMAL312.
Fig. 7. Residues in 1at1A, at the interface with maltose, colored by their relative importance. The ligand (maltose) is colored green. Atoms further than 30 Å away from the geometric center of the ligand, as well as on the line of sight to the ligand were removed. (See Appendix for the coloring scheme for the protein chain 1at1A.)

Table 8. The top 25% of residues in 1at1A at the interface with phosphonoacetamide. (Field names: res: residue number in the PDB entry; type: amino acid type; subst's: substitutions seen in the alignment; with the percentage of each type in the bracket; noc/bb: number of contacts with the ligand, with the number of contacts realized through backbone atoms given in the bracket; dist: distance of closest approach to the ligand.)

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>subst’s (%)</th>
<th>cvg</th>
<th>noc/bb</th>
<th>dist (Å)</th>
<th>antn</th>
</tr>
</thead>
<tbody>
<tr>
<td>267</td>
<td>L</td>
<td>A(10)</td>
<td>0.12</td>
<td>8/8</td>
<td>3.12</td>
<td>site</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L(52)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G(32)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>S(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>V(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>168</td>
<td>T</td>
<td>V(48)</td>
<td>0.15</td>
<td>1/0</td>
<td>4.70</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>T(45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>P(3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>N(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>A(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>K(5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>R(60)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>S(9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>H(6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>F(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>P(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Q(5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>L(5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>R</td>
<td>A(1)</td>
<td>0.23</td>
<td>1/1</td>
<td>4.74</td>
<td></td>
</tr>
</tbody>
</table>

Table 9. List of disruptive mutations for the top 25% of residues in 1at1A, that are at the interface with phosphonoacetamide.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>S</td>
<td>(KR) (FQMW) (NYLPI) (D)</td>
</tr>
<tr>
<td>54</td>
<td>R</td>
<td>(TD) (SYEVCLAPIG) (FMW) (N)</td>
</tr>
<tr>
<td>55</td>
<td>T</td>
<td>(KR) (FQMW) (NELPI) (D)</td>
</tr>
<tr>
<td>134</td>
<td>H</td>
<td>(E) (TQMD) (SNKCLAPIG) (YR)</td>
</tr>
<tr>
<td>137</td>
<td>Q</td>
<td>(Y) (FTWH) (SVCAG) (D)</td>
</tr>
<tr>
<td>53</td>
<td>T</td>
<td>(KR) (FQMW) (NELPI) (D)</td>
</tr>
<tr>
<td>266</td>
<td>P</td>
<td>(R) (Y) (H) (KE)</td>
</tr>
<tr>
<td>268</td>
<td>P</td>
<td>(Y) (R) (T) (E)</td>
</tr>
<tr>
<td>105</td>
<td>R</td>
<td>(TYD) (E) (SCG) (FVLAWPI)</td>
</tr>
<tr>
<td>267</td>
<td>L</td>
<td>(R) (Y) (H) (K)</td>
</tr>
<tr>
<td>168</td>
<td>T</td>
<td>(R) (K) (H) (FEQW)</td>
</tr>
<tr>
<td>56</td>
<td>R</td>
<td>(D) (T) (YE) (S)</td>
</tr>
</tbody>
</table>

Table 10 lists the top 25% of residues at the interface with 1at1B. The following table (Table 11) suggests possible disruptive replacements for these residues (see Section 4.6).

Figure 8 shows residues in 1at1A colored by their importance, at the interface with 1at1APCT311.

Interface with 1at1B. Table 10 lists the top 25% of residues at the interface with 1at1B. The following table (Table 11) suggests possible disruptive replacements for these residues (see Section 4.6).
Table 10. The top 25% of residues in 1at1A at the interface with 1at1B. (Field names: res: residue number in the PDB entry; type: amino acid type; substs: substitutions seen in the alignment; with the percentage of each type in the bracket; noc/bb: number of contacts with the ligand, with the number of contacts realized through backbone atoms given in the bracket; dist: distance of closest approach to the ligand.)

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>substs’s (%)</th>
<th>cvg</th>
<th>noc/ bb</th>
<th>dist (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>130</td>
<td>G</td>
<td>G(94) D(4)C</td>
<td>0.04</td>
<td>1/1</td>
<td>4.64</td>
</tr>
<tr>
<td>129</td>
<td>D</td>
<td>S(4) N(12)D(79)CA T(2)</td>
<td>0.14</td>
<td>2/0</td>
<td>4.53</td>
</tr>
<tr>
<td>106</td>
<td>H</td>
<td>H(87) D(2)V S(2) T(4) A(1)RQ</td>
<td>0.17</td>
<td>3/0</td>
<td>4.15</td>
</tr>
<tr>
<td>88</td>
<td>L</td>
<td>L(80) V(3) A(1) I(10)WF MY(1)T</td>
<td>0.22</td>
<td>22/11</td>
<td>2.97</td>
</tr>
</tbody>
</table>

Table 11. List of disruptive mutations for the top 25% of residues in 1at1A, that are at the interface with 1at1B.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptives mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>130</td>
<td>G</td>
<td>(R) (K) (FEWH) (QM)</td>
</tr>
<tr>
<td>129</td>
<td>D</td>
<td>(R) (H) (FW) (K)</td>
</tr>
<tr>
<td>106</td>
<td>H</td>
<td>(E) (QMD) (TK) (NLPI)</td>
</tr>
<tr>
<td>88</td>
<td>L</td>
<td>(R) (Y) (K) (E)</td>
</tr>
</tbody>
</table>

Figure 9 shows residues in 1at1A colored by their importance, at the interface with 1at1B.

3 CHAIN 1AT1B

3.1 P0A7F3 overview

From SwissProt, id P0A7F3, 100% identical to 1at1B:

Description: Aspartate carbamoyltransferase regulatory chain.

Organism, scientific name: Escherichia coli.

Taxonomy: Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.

Function: Involved in allosteric regulation of aspartate carbamoyltransferase.

Cofactor: Binds 1 zinc ion per subunit.

Subunit: Heterododecamer (2C3:3R2) of six catalytic pyrB chains organized as two trimers (C3), and six regulatory pyrI chains organized as three dimers (R2).

Similarity: Belongs to the pyrI family.

About: This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the
EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

3.2 Multiple sequence alignment for 1at1B

For the chain 1at1B, the alignment 1at1B.msf (attached) with 49 sequences was used. The alignment was downloaded from the HSSP database, and fragments shorter than 75% of the query as well as duplicate sequences were removed. It can be found in the attachment to this report, under the name of 1at1B.msf. Its statistics, from the alistat program are the following:

- Format: MSF
- Number of sequences: 49
- Total number of residues: 6972
- Smallest: 134
- Largest: 146
- Average length: 142.3
- Alignment length: 146
- Average identity: 45%
- Most related pair: 99%
- Most unrelated pair: 28%
- Most distant seq: 41%

Furthermore, 8% of residues show as conserved in this alignment. The alignment consists of 2% eukaryotic (2% arthropoda), 59% prokaryotic, and 42% archaean sequences. (Descriptions of some sequences were not readily available.) The file containing the sequence descriptions can be found in the attachment, under the name 1at1B.descr.

3.3 Residue ranking in 1at1B

The 1at1B sequence is shown in Fig. 10, with each residue colored according to its estimated importance. The full listing of residues in 1at1B can be found in the file called 1at1B.ranks_sorted in the attachment.

3.4 Top ranking residues in 1at1B and their position on the structure

In the following we consider residues ranking among top 25% of residues in the protein. Figure 11 shows residues in 1at1B colored by their importance: bright red and yellow indicate more conserved/important residues (see Appendix for the coloring scheme). A Pymol script for producing this figure can be found in the attachment.

3.4.1 Clustering of residues at 25% coverage

Fig. 12 shows the top 25% of all residues, this time colored according to clusters they belong to. The clusters in Fig.12 are composed of the residues listed in Table 12.
Table 12.

<table>
<thead>
<tr>
<th>color</th>
<th>size</th>
<th>member residues</th>
</tr>
</thead>
<tbody>
<tr>
<td>red</td>
<td>28</td>
<td>9,12,15,16,17,18,19,20,21,33,47,48,50,56,57,60,61,62,70,74,78,79,81,82,83,91,94</td>
</tr>
<tr>
<td>blue</td>
<td>8</td>
<td>109,111,114,115,119,123,138,140,141</td>
</tr>
</tbody>
</table>

Table 12. Clusters of top ranking residues in 1at1B.

3.4.2 Overlap with known functional surfaces at 25% coverage.
The name of the ligand is composed of the source PDB identifier and the heteroatom name used in that file.

Interface with 1at1D. Table 13 lists the top 25% of residues at the interface with 1at1D. The following table (Table 14) suggests possible disruptive replacements for these residues (see Section 4.6).

Table 13.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>subst's (%)</th>
<th>cvg</th>
<th>noc/ dist (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>47</td>
<td>N</td>
<td>N (89) H (4) M (4) R (2)</td>
<td>0.16</td>
<td>64/13 3.03</td>
</tr>
<tr>
<td>48</td>
<td>L</td>
<td>L (53) V (34) N (4) M (4) A (4)</td>
<td>0.24</td>
<td>1/0 4.77</td>
</tr>
<tr>
<td>9</td>
<td>V</td>
<td>V (87) I (12)</td>
<td>0.25</td>
<td>11/11 3.04</td>
</tr>
</tbody>
</table>

Table 13. The top 25% of residues in 1at1B at the interface with 1at1D. (Field names: res: residue number in the PDB entry; type: amino acid type; subst's: substitutions seen in the alignment; with the percentage of each type in the bracket; noc/bb: number of contacts with the ligand, with the number of contacts realized through backbone atoms given in the bracket; dist: distance of closest approach to the ligand.)

Table 14.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>47</td>
<td>N</td>
<td>(Y) (T) (S) (C) (W) (V)</td>
</tr>
<tr>
<td>48</td>
<td>L</td>
<td>(Y) (R) (H) (T)</td>
</tr>
<tr>
<td>9</td>
<td>V</td>
<td>(YR) (KE) (H) (QD)</td>
</tr>
</tbody>
</table>

Table 14. List of disruptive mutations for the top 25% of residues in 1at1B, that are at the interface with 1at1D.

Figure 13 shows residues in 1at1B colored by their importance, at the interface with 1at1D.

Interface with 1at1A. Table 15 lists the top 25% of residues at the interface with 1at1A. The following table (Table 16) suggests possible disruptive replacements for these residues (see Section 4.6).
Fig. 13. Residues in 1at1B, at the interface with 1at1D, colored by their relative importance. 1at1D is shown in backbone representation (See Appendix for the coloring scheme for the protein chain 1at1B.)

Table 15.
The top 25% of residues in 1at1B at the interface with 1at1A.

<table>
<thead>
<tr>
<th>Res</th>
<th>Type</th>
<th>Substs (%)</th>
<th>Cvg</th>
<th>Noc/BB</th>
<th>Dist (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>111</td>
<td>N</td>
<td>N(100)</td>
<td>0.08</td>
<td>9/0</td>
<td>2.93</td>
</tr>
<tr>
<td>114</td>
<td>C</td>
<td>C(100)</td>
<td>0.08</td>
<td>16/14</td>
<td>3.50</td>
</tr>
<tr>
<td>140</td>
<td>Y</td>
<td>Y(100)</td>
<td>0.08</td>
<td>66/21</td>
<td>3.27</td>
</tr>
<tr>
<td>141</td>
<td>C</td>
<td>C(100)</td>
<td>0.08</td>
<td>33/26</td>
<td>3.09</td>
</tr>
<tr>
<td>119</td>
<td>E</td>
<td>E(87) D(8)</td>
<td>0.16</td>
<td>56/0</td>
<td>2.97</td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>I(83)</td>
<td>0.23</td>
<td>62/11</td>
<td>2.96</td>
</tr>
</tbody>
</table>

Table 16. List of disruptive mutations for the top 25% of residues in 1at1B, that are at the interface with 1at1A.

<table>
<thead>
<tr>
<th>Res</th>
<th>Type</th>
<th>Disruptive Mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>111</td>
<td>N</td>
<td>(Y) (FTWH) (SEVCARG) (MD)</td>
</tr>
<tr>
<td>114</td>
<td>C</td>
<td>(KER) (FQMWHD) (NYLPI) (SVA)</td>
</tr>
<tr>
<td>140</td>
<td>Y</td>
<td>(K) (QM) (NEVLAPIR) (D)</td>
</tr>
<tr>
<td>141</td>
<td>C</td>
<td>(KER) (FQMWHD) (NYLPI) (SVA)</td>
</tr>
<tr>
<td>119</td>
<td>E</td>
<td>(FW) (VCARHGR) (Y) (T)</td>
</tr>
</tbody>
</table>

Table 17. The top 25% of residues in 1at1B at the interface with zinc ion.

<table>
<thead>
<tr>
<th>Res</th>
<th>Type</th>
<th>Substs (%)</th>
<th>Cvg</th>
<th>Noc/BB</th>
<th>Dist (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>109</td>
<td>C</td>
<td>C(100)</td>
<td>0.08</td>
<td>3/1</td>
<td>2.34</td>
</tr>
<tr>
<td>111</td>
<td>N</td>
<td>N(100)</td>
<td>0.08</td>
<td>1/0</td>
<td>4.60</td>
</tr>
<tr>
<td>114</td>
<td>C</td>
<td>C(100)</td>
<td>0.08</td>
<td>3/1</td>
<td>2.37</td>
</tr>
<tr>
<td>138</td>
<td>C</td>
<td>C(100)</td>
<td>0.08</td>
<td>3/1</td>
<td>2.34</td>
</tr>
<tr>
<td>140</td>
<td>Y</td>
<td>Y(100)</td>
<td>0.08</td>
<td>2/1</td>
<td>4.65</td>
</tr>
<tr>
<td>141</td>
<td>C</td>
<td>C(100)</td>
<td>0.08</td>
<td>5/3</td>
<td>2.34</td>
</tr>
</tbody>
</table>

Zinc ion binding site. Table 17 lists the top 25% of residues at the interface with 1at1BZN109 (zinc ion). The following table (Table 18) suggests possible disruptive replacements for these residues (see Section 4.6).

Figure 14 shows residues in 1at1B colored by their importance, at the interface with 1at1A.

Zinc ion binding site. Table 17 lists the top 25% of residues at the interface with 1at1BZN109 (zinc ion). The following table (Table 18) suggests possible disruptive replacements for these residues (see Section 4.6).
Table 18. List of disruptive mutations for the top 25% of residues in 1at1B, that are at the interface with zinc ion.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>109</td>
<td>C</td>
<td>(KER) (FQMWHD) (NYLPI) (SVA)</td>
</tr>
<tr>
<td>111</td>
<td>N</td>
<td>(Y) (FTWH) (SEVCARG) (MD)</td>
</tr>
<tr>
<td>114</td>
<td>C</td>
<td>(KER) (FQMWHD) (NYLPI) (SVA)</td>
</tr>
<tr>
<td>138</td>
<td>C</td>
<td>(KER) (FQMWHD) (NYLPI) (SVA)</td>
</tr>
<tr>
<td>140</td>
<td>Y</td>
<td>(K) (QM) (NEVLAPIR) (D)</td>
</tr>
<tr>
<td>141</td>
<td>C</td>
<td>(KER) (FQMWHD) (NYLPI) (SVA)</td>
</tr>
</tbody>
</table>

Fig. 15. Residues in 1at1B, at the interface with zinc ion, colored by their relative importance. The ligand (zinc ion) is colored green. Atoms further than 30 Å away from the geometric center of the ligand, as well as on the line of sight to the ligand were removed. (See Appendix for the coloring scheme for the protein chain 1at1B.)

Figure 15 shows residues in 1at1B colored by their importance, at the interface with 1at1BZN109.

3.4.3 Possible novel functional surfaces at 25% coverage. One group of residues is conserved on the 1at1B surface, away from (or substantially larger than) other functional sites and interfaces recognizable in PDB entry 1at1. It is shown in Fig. 16. The right panel shows (in blue) the rest of the larger cluster this surface belongs to. The residues belonging to this surface "patch" are listed in Table 19, while Table 20 suggests possible disruptive replacements for these residues (see Section 4.6).

Table 19. Residues forming surface "patch" in 1at1B.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>substitutions(%)</th>
<th>cvg</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>D</td>
<td>D(100)</td>
<td>0.08</td>
</tr>
<tr>
<td>20</td>
<td>H</td>
<td>H(100)</td>
<td>0.08</td>
</tr>
<tr>
<td>56</td>
<td>K</td>
<td>K(100)</td>
<td>0.08</td>
</tr>
<tr>
<td>60</td>
<td>K</td>
<td>K(100)</td>
<td>0.08</td>
</tr>
<tr>
<td>94</td>
<td>K</td>
<td>K(100)</td>
<td>0.08</td>
</tr>
<tr>
<td>50</td>
<td>S</td>
<td>S(95) P(4)</td>
<td>0.09</td>
</tr>
<tr>
<td>18</td>
<td>I</td>
<td>I(97) L(2)</td>
<td>0.10</td>
</tr>
<tr>
<td>57</td>
<td>D</td>
<td>D(93) G(6)</td>
<td>0.11</td>
</tr>
<tr>
<td>17</td>
<td>V</td>
<td>V(97) I(2)</td>
<td>0.12</td>
</tr>
<tr>
<td>79</td>
<td>P</td>
<td>P(97) R(2)</td>
<td>0.12</td>
</tr>
<tr>
<td>33</td>
<td>F</td>
<td>F(46) L(53)</td>
<td>0.13</td>
</tr>
<tr>
<td>12</td>
<td>I</td>
<td>I(93) L(6)</td>
<td>0.14</td>
</tr>
<tr>
<td>74</td>
<td>L</td>
<td>L(57) I(40) V(2)</td>
<td>0.14</td>
</tr>
<tr>
<td>84</td>
<td>N</td>
<td>N(91) S(6) D(2)</td>
<td>0.15</td>
</tr>
<tr>
<td>47</td>
<td>N</td>
<td>N(89) H(4) M(4)</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R(2)</td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>T</td>
<td>T(87) V(6) K(4)</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R(2)</td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>E</td>
<td>E(87) K(6) A(6)</td>
<td>0.19</td>
</tr>
<tr>
<td>61</td>
<td>I</td>
<td>V(16) I(81) L(2)</td>
<td>0.20</td>
</tr>
<tr>
<td>91</td>
<td>V</td>
<td>V(79) L(4) I(16)</td>
<td>0.21</td>
</tr>
<tr>
<td>83</td>
<td>V</td>
<td>V(59) I(32) L(6)</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F(2)</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>L</td>
<td>L(53) V(34) N(4)</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M(4) A(4)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>V</td>
<td>V(87) I(12)</td>
<td>0.25</td>
</tr>
<tr>
<td>70</td>
<td>Q</td>
<td>Q(40) E(40) A(6)</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D(10)</td>
<td></td>
</tr>
</tbody>
</table>

Table 20. Residues forming surface "patch" in 1at1B.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>D</td>
<td>(R) (FWH) (KYVCAG) (TQM)</td>
</tr>
<tr>
<td>20</td>
<td>H</td>
<td>(E) (TQMD) (SNKVLAPIG) (YR)</td>
</tr>
<tr>
<td>56</td>
<td>K</td>
<td>(Y) (FTW) (SVCA) (HD)</td>
</tr>
<tr>
<td>60</td>
<td>K</td>
<td>(Y) (FTW) (SVCA) (HD)</td>
</tr>
</tbody>
</table>

continued in next column
Table 20. Disruptive mutations for the surface patch in 1at1B.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>94</td>
<td>K</td>
<td>(Y) (FTW) (SVCAG) (HD)</td>
</tr>
<tr>
<td>50</td>
<td>S</td>
<td>(R) (K) (H) (FYGQ)</td>
</tr>
<tr>
<td>18</td>
<td>I</td>
<td>(YR) (TH) (SKECG) (FQWD)</td>
</tr>
<tr>
<td>57</td>
<td>D</td>
<td>(R) (FWH) (K) (Y)</td>
</tr>
<tr>
<td>17</td>
<td>V</td>
<td>(YR) (KE) (H) (QD)</td>
</tr>
<tr>
<td>79</td>
<td>P</td>
<td>(Y) (T) (SECHRG) (D)</td>
</tr>
<tr>
<td>33</td>
<td>F</td>
<td>(KE) (T) (QDR) (SCG)</td>
</tr>
<tr>
<td>12</td>
<td>I</td>
<td>(YR) (TH) (SKECG) (FQWD)</td>
</tr>
<tr>
<td>74</td>
<td>L</td>
<td>(YR) (H) (T) (KE)</td>
</tr>
<tr>
<td>84</td>
<td>N</td>
<td>(Y) (FWH) (R) (T)</td>
</tr>
<tr>
<td>47</td>
<td>N</td>
<td>(Y) (T) (SFECWG) (VLA)</td>
</tr>
<tr>
<td>82</td>
<td>T</td>
<td>(FW) (KRR) (EM) (Q)</td>
</tr>
<tr>
<td>62</td>
<td>E</td>
<td>(FYWH) (CRG) (VIA) (T)</td>
</tr>
<tr>
<td>61</td>
<td>I</td>
<td>(YR) (H) (T) (KE)</td>
</tr>
<tr>
<td>91</td>
<td>V</td>
<td>(YR) (KE) (H) (QD)</td>
</tr>
<tr>
<td>83</td>
<td>V</td>
<td>(R) (KYE) (QHD) (T)</td>
</tr>
<tr>
<td>48</td>
<td>L</td>
<td>(YR) (R) (H) (T)</td>
</tr>
<tr>
<td>9</td>
<td>V</td>
<td>(YR) (KE) (H) (QD)</td>
</tr>
<tr>
<td>70</td>
<td>Q</td>
<td>(Y) (H) (FW) (T)</td>
</tr>
</tbody>
</table>

Another group of surface residues is shown in Fig.17. The residues belonging to this surface "patch" are listed in Table 21, while Table 22 suggests possible disruptive replacements for these residues (see Section 4.6).

Table 21. Residues forming surface "patch" in 1at1B.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>substitutions(%)</th>
<th>cvg</th>
<th>antn</th>
</tr>
</thead>
<tbody>
<tr>
<td>109</td>
<td>C</td>
<td>C(100)</td>
<td>0.08</td>
<td>site</td>
</tr>
<tr>
<td>111</td>
<td>N</td>
<td>N(100)</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>114</td>
<td>C</td>
<td>C(100)</td>
<td>0.08</td>
<td>site</td>
</tr>
<tr>
<td>138</td>
<td>C</td>
<td>C(100)</td>
<td>0.08</td>
<td>site</td>
</tr>
<tr>
<td>140</td>
<td>Y</td>
<td>Y(100)</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>141</td>
<td>C</td>
<td>C(100)</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>119</td>
<td>E</td>
<td>E(87) D(8) H(4)</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>I</td>
<td>I(83) V(8) A(8)</td>
<td>0.23</td>
<td></td>
</tr>
</tbody>
</table>

Table 22. Disruptive mutations for the surface patch in 1at1B.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>109</td>
<td>C</td>
<td>(KER) (FQMWHD) (NYLPI) (SVA)</td>
</tr>
<tr>
<td>111</td>
<td>N</td>
<td>(Y) (FTWH) (SEVCARG) (MD)</td>
</tr>
<tr>
<td>114</td>
<td>C</td>
<td>(KER) (FQMWHD) (NYLPI) (SVA)</td>
</tr>
<tr>
<td>138</td>
<td>C</td>
<td>(KER) (FQMWHD) (NYLPI) (SVA)</td>
</tr>
<tr>
<td>140</td>
<td>Y</td>
<td>(K) (QM) (NEVLAPIR) (D)</td>
</tr>
<tr>
<td>141</td>
<td>C</td>
<td>(KER) (FQMWHD) (NYLPI) (SVA)</td>
</tr>
<tr>
<td>119</td>
<td>E</td>
<td>(FW) (VCAHRG) (Y) (T)</td>
</tr>
<tr>
<td>115</td>
<td>I</td>
<td>(YR) (KE) (E) (T)</td>
</tr>
</tbody>
</table>

4 NOTIONS ON USING TRACE RESULTS

4.1 Coverage

Trace results are commonly expressed in terms of coverage: the residue is important if its “coverage” is small - that is if it belongs to some small top percentage of residues [100% is all of the residues in a chain], according to trace. The ET results are presented in the form of a table, usually limited to top 25% percent of residues (or to some nearby percentage), sorted by the strength of the presumed evolutionary pressure. (I.e., the smaller the coverage, the stronger the pressure on the residue.) Starting from the top of that list, mutating a couple of residues should affect the protein somehow, with the exact effects to be determined experimentally.

4.2 Known substitutions

One of the table columns is “substitutions” - other amino acid types seen at the same position in the alignment. These amino acid types may be interchangeable at that position in the protein, so if one wants to affect the protein by a point mutation, they should be avoided. For example if the substitutions are “RVK” and the original protein has an R at that position, it is advisable to try anything, but RVK. Conversely, when looking for substitutions which will not affect the protein, one may try replacing, R with K, or (perhaps more surprisingly), with V. The percentage of times the substitution appears in the alignment is given in the immediately following bracket. No percentage is given in the cases when it is smaller than 1%. This is meant to be a rough guide - due to rounding errors these percentages often do not add up to 100%.
4.3 Surface

To detect candidates for novel functional interfaces, first we look for residues that are solvent accessible (according to DSSP program) by at least 10 Å², which is roughly the area needed for one water molecule to come in the contact with the residue. Furthermore, we require that these residues form a “cluster” of residues which have neighbor within 5 Å from any of their heavy atoms.

Note, however, that, if our picture of protein evolution is correct, the neighboring residues which are not surface accessible might be equally important in maintaining the interaction specificity - they should not be automatically dropped from consideration when choosing the set for mutagenesis. (Especially if they form a cluster with the surface residues.)

4.4 Number of contacts

Another column worth noting is denoted “noc/bb”; it tells the number of contacts heavy atoms of the residue in question make across the interface, as well as how many of them are realized through the backbone atoms (if all or most contacts are through the backbone, mutation presumably won’t have strong impact). Two heavy atoms are considered to be “in contact” if their centers are closer than 5 Å.

4.5 Annotation

If the residue annotation is available (either from the pdb file or from other sources), another column, with the header “annotation” appears. Annotations carried over from PDB are the following: site (indicating existence of related site record in PDB), S-S (disulfide bond forming residue), hb (hydrogen bond forming residue), jb (james bond forming residue), and sb (for salt bridge forming residue).

4.6 Mutation suggestions

Mutation suggestions are completely heuristic and based on complementarity with the substitutions found in the alignment. Note that they are meant to be disruptive to the interaction of the protein with its ligand. The attempt is made to complement the following properties: small [AVGSTC], medium [LPN QDE MIK], large [WFYHR], hydrophobic [LPVAMWF], polar [GTCY]; positively [KHR], or negatively [DE] charged, aromatic [WFYH], long aliphatic chain [ERKQM], OH-group possession [SDETY], and NH2 group possession [NQHR]. The suggestions are listed according to how different they appear to be from the original amino acid, and they are grouped in round brackets if they appear equally disruptive. From left to right, each bracketed group of amino acid types resembles more strongly the original (i.e. is, presumably, less disruptive) These suggestions are tentative - they might prove disruptive to the fold rather than to the interaction. Many researcher will choose, however, the straightforward alanine mutations, especially in the beginning stages of their investigation.

5 APPENDIX

5.1 File formats

Files with extension “ranks_sort” are the actual trace results. The fields in the table in this file:

- **alignment#** number of the position in the alignment
- **residue#** residue number in the PDB file
- **type** amino acid type
- **rank** rank of the position according to older version of ET

5.2 Color schemes used

The following color scheme is used in figures with residues colored by cluster size: black is a single-residue cluster; clusters composed of more than one residue colored according to this hierarchy (ordered by descending size): red, blue, yellow, green, purple, azure, turquoise, brown, coral, magenta, LightSalmon, SkyBlue, violet, gold, bisque, LightSlateBlue, orchid, RosyBrown, MediumAquamarine, DarkOliveGreen, CornflowerBlue, grey55, burlywood, LimeGreen, tan, DarkOrange, DeepPink, maroon, BlanchedAlmond.

The colors used to distinguish the residues by the estimated evolutionary pressure they experience can be seen in Fig. 18.

5.3 Credits

5.3.1 Alistat alistat reads a multiple sequence alignment from the file and shows a number of simple statistics about it. These statistics include the format, the number of sequences, the total number of residues, the average and range of the sequence lengths, and the alignment length (e.g. including gap characters). Also shown are some percent identities. A percent pairwise alignment identity is defined as (idents / MIN(len1, len2)) where idents is the number of exact identities and len1, len2 are the unaligned lengths of the two sequences. The “average percent identity”, “most related pair”, and “most unrelated pair” of the alignment are the average, maximum, and minimum of all (N(N-1)/2) pairs, respectively. The “most distant seq” is calculated by finding the maximum pairwise identity (best relative) for all N sequences, then finding the minimum of these N numbers (hence, the most outlying sequence). alistat is copyrighted.
by HHMI/Washington University School of Medicine, 1992-2001, and freely distributed under the GNU General Public License.

5.3.2 CE To map ligand binding sites from different source structures, report_maker uses the CE program: http://cl.sdsc.edu/. Shindyalov IN, Bourne PE (1998) “Protein structure alignment by incremental combinatorial extension (CE) of the optimal path”. Protein Engineering 11(9) 739-747.

5.3.3 DSSP In this work a residue is considered solvent accessible if the DSSP program finds it exposed to water by at least 10Å², which is roughly the area needed for one water molecule to come in the contact with the residue. DSSP is copyrighted by W. Kabsch, C. Sander and MPI-MF, 1983, 1985, 1988, 1994 1995, CMBI version by Elmar.Krieger@cmbi.kun.nl November 18,2002, http://www.cmbi.kun.nl/gv/dssp/descrip.html.


5.3.5 LaTeX The text for this report was processed using L\LaTeX; Leslie Lamport, “LaTeX: A Document Preparation System Addison-Wesley,” Reading, Mass. (1986).


5.3.7 Pymol The figures in this report were produced using Pymol. The scripts can be found in the attachment. Pymol is an open-source application copyrighted by DeLano Scientific LLC (2005). For more information about Pymol see http://pymol.sourceforge.net/. (Note for Windows users: the attached package needs to be unzipped for Pymol to read the scripts and launch the viewer.)

5.4 Note about ET Viewer Dan Morgan from the Lichtarge lab has developed a visualization tool specifically for viewing trace results. If you are interested, please visit: http://mammoth.bcm.tmc.edu/traceview/

The viewer is self-unpacking and self-installing. Input files to be used with ETV (extension .etvx) can be found in the attachment to the main report.


5.6 About report_maker report_maker was written in 2006 by Ivana Mihalek. The ID ranking visualization program was written by Ivica Reš. report_maker is copyrighted by Lichtarge Lab, Baylor College of Medicine, Houston.

5.7 Attachments The following files should accompany this report:

- 1at1A.complex.pdb - coordinates of 1at1A with all of its interacting partners
- 1at1A.etvx - ET viewer input file for 1at1A
- 1at1A.cluster_report.summary - Cluster report summary for 1at1A
- 1at1A.ranks - Ranks file in sequence order for 1at1A
- 1at1A.clusters - Cluster descriptions for 1at1A
- 1at1A.msf - the multiple sequence alignment used for the chain 1at1A
- 1at1A.descr - description of sequences used in 1at1A msf
- 1at1A.ranks_sorted - full listing of residues and their ranking for 1at1A
- 1at1A.1at1A2.if.pml - Pymol script for Figure 5
- 1at1A.1at1A1.if.pml - Pymol script for Figure 6
- 1at1A.1at1AMAL312.if.pml - Pymol script for Figure 7
- 1at1A.1at1APCT311.if.pml - Pymol script for Figure 8
- 1at1A.1at1AZN109.if.pml - Pymol script for Figure 10
- 1at1A.complex.pdb - coordinates of 1at1A with all of its interacting partners
- 1at1A.etvx - ET viewer input file for 1at1B
- 1at1B.cluster_report.summary - Cluster report summary for 1at1B
- 1at1B.ranks - Ranks file in sequence order for 1at1B
- 1at1B.clusters - Cluster descriptions for 1at1B
- 1at1B.msf - the multiple sequence alignment used for the chain 1at1B
- 1at1B.descr - description of sequences used in 1at1B msf
- 1at1B.ranks_sorted - full listing of residues and their ranking for 1at1B
- 1at1B.1at1D.if.pml - Pymol script for Figure 13
- 1at1B.cbcvg - used by other 1at1B – related pymol scripts
- 1at1B.1at1A.if.pml - Pymol script for Figure 14
- 1at1B.1at1A2.if.pml - Pymol script for Figure 15
- 1at1B.cluster - Cluster report summary for 1at1B