1a5n
Evolutionary trace report by report_maker
August 6, 2009

CONTENTS

1 Introduction

2 Chain 1a5nA
   2.1 P18316 overview
   2.2 Multiple sequence alignment for 1a5nA
   2.3 Residue ranking in 1a5nA
   2.4 Top ranking residues in 1a5nA 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.
      2.4.3 Possible novel functional surfaces at 25% coverage.

3 Chain 1a5nC
   3.1 P18314 overview
   3.2 Multiple sequence alignment for 1a5nC
   3.3 Residue ranking in 1a5nC
   3.4 Top ranking residues in 1a5nC 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 Chain 1a5nB
   4.1 P18316 overview
   4.2 Multiple sequence alignment for 1a5nB
   4.3 Residue ranking in 1a5nB
   4.4 Top ranking residues in 1a5nB and their position on the structure
      4.4.1 Clustering of residues at 25% coverage.
      4.4.2 Overlap with known functional surfaces at 25% coverage.
      4.4.3 Possible novel functional surfaces at 25% coverage.

5 Notes on using trace results
   5.1 Coverage
   5.2 Known substitutions
   5.3 Surface
   5.4 Number of contacts
   5.5 Annotation
   5.6 Mutation suggestions

6 Appendix
   6.1 File formats
   6.2 Color schemes used
   6.3 Credits
      6.3.1 Alistat
      6.3.2 CE
      6.3.3 DSSP
      6.3.4 HSSP
      6.3.5 LaTex
      6.3.6 Muscle
      6.3.7 Pymol
   6.4 Note about ET Viewer
   6.5 Citing this work
   6.6 About report_maker
   6.7 Attachments

1 INTRODUCTION

From the original Protein Data Bank entry (PDB id 1a5n):
Title: K217a variant of klebsiella aerogenes urease, chemically rescued by formate and nickel
Compound: Mol id: 1; molecule: urease (gamma subunit); chain: a; ec: 3.5.1.5; engineered: yes; mutation: yes; mol id: 2; molecule: urease (beta subunit); chain: b; ec: 3.5.1.5; engineered: yes; mol id: 3; molecule: urease (alpha subunit); chain: c; ec: 3.5.1.5; engineered: yes
Organism, scientific name: Klebsiella Aerogenes;
1a5n contains unique chains 1a5nA (100 residues), 1a5nC (551 residues), and 1a5nB (101 residues)
2 CHAIN 1A5NA

2.1 P18316 overview

From SwissProt, id P18316, 100% identical to 1a5nA:

Description: Urease gamma subunit (EC 3.5.1.5) (Urea amidohydrolase gamma subunit).

Organism, scientific name: Klebsiella aerogenes.

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

Catalytic activity: Urea + H(2)O = CO(2) + 2 NH(3).

Subunit: (Alpha, beta, gamma)(3).

Subcellular location: Cytoplasmic.

Similarity: Belongs to the urease gamma subunit 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 1a5nA

For the chain 1a5nA, the alignment 1a5nA.msf (attached) with 379 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 1a5nA.msf. Its statistics, from the alistat program are the following:

<table>
<thead>
<tr>
<th>Format:</th>
<th>MSF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of sequences:</td>
<td>379</td>
</tr>
<tr>
<td>Total number of residues:</td>
<td>36762</td>
</tr>
<tr>
<td>Smallest:</td>
<td>24</td>
</tr>
<tr>
<td>Largest:</td>
<td>100</td>
</tr>
<tr>
<td>Average length:</td>
<td>97.0</td>
</tr>
<tr>
<td>Alignment length:</td>
<td>100</td>
</tr>
<tr>
<td>Average identity:</td>
<td>60%</td>
</tr>
<tr>
<td>Most related pair:</td>
<td>99%</td>
</tr>
<tr>
<td>Most unrelated pair:</td>
<td>4%</td>
</tr>
<tr>
<td>Most distant seq:</td>
<td>49%</td>
</tr>
</tbody>
</table>

Furthermore, <1% of residues show as conserved in this alignment.

The alignment consists of 3% eukaryotic (1% fungi, 1% plantae), 20% prokaryotic, and <1% 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 1a5nA.descr.

2.3 Residue ranking in 1a5nA

The 1a5nA sequence is shown in Fig. 1, with each residue colored according to its estimated importance. The full listing of residues in 1a5nA can be found in the file called 1a5nA.ranks,sorted in the attachment.

2.4 Top ranking residues in 1a5nA and their position on the structure

In the following we consider residues ranking among top 25% of residues in the protein. Figure 2 shows residues in 1a5nA 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.

2.4.1 Clustering of residues at 25% coverage

Fig. 3 shows the top 25% of all residues, this time colored according to clusters they belong to. The clusters in Fig.3 are composed of the residues listed in Table 1.

<table>
<thead>
<tr>
<th>cluster color</th>
<th>size</th>
<th>member residues</th>
</tr>
</thead>
<tbody>
<tr>
<td>red</td>
<td>11</td>
<td>53,57,81,83,86,88,89,90,91,92,94</td>
</tr>
<tr>
<td>blue</td>
<td>10</td>
<td>16,20,23,27,30,31,34,69,70,72</td>
</tr>
<tr>
<td>yellow</td>
<td>2</td>
<td>48,50</td>
</tr>
</tbody>
</table>

Table 1. Clusters of top ranking residues in 1a5nA.

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.
Fig. 3. Residues in 1a5nA, 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.

Interface with 1a5nC. Table 2 lists the top 25% of residues at the interface with 1a5nC. The following table (Table 3) suggests possible disruptive replacements for these residues (see Section 5.6).

<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>83</td>
<td>E</td>
<td>E(99)D.</td>
<td>0.03</td>
<td>44/5</td>
<td>2.89</td>
</tr>
<tr>
<td>92</td>
<td>L</td>
<td>L(98).V</td>
<td>0.04</td>
<td>11/0</td>
<td>3.71</td>
</tr>
<tr>
<td>34</td>
<td>E</td>
<td>.(2)</td>
<td>0.05</td>
<td>8/2</td>
<td>3.83</td>
</tr>
<tr>
<td>31</td>
<td>N</td>
<td>N(94)TA</td>
<td>0.15</td>
<td>38/5</td>
<td>2.68</td>
</tr>
<tr>
<td>23</td>
<td>R</td>
<td>R(94)AY</td>
<td>0.16</td>
<td>21/0</td>
<td>2.78</td>
</tr>
<tr>
<td>90</td>
<td>S</td>
<td>T(89)</td>
<td>0.19</td>
<td>1/0</td>
<td>4.56</td>
</tr>
<tr>
<td>16</td>
<td>A</td>
<td>.(3)</td>
<td>0.20</td>
<td>5/3</td>
<td>4.00</td>
</tr>
<tr>
<td>81</td>
<td>Q</td>
<td>G(6)</td>
<td>0.22</td>
<td>49/25</td>
<td>2.84</td>
</tr>
</tbody>
</table>

continued in next column

<table>
<thead>
<tr>
<th>Table 2. continued</th>
</tr>
</thead>
<tbody>
<tr>
<td>res</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>70</td>
</tr>
<tr>
<td>72</td>
</tr>
</tbody>
</table>

Table 2. The top 25% of residues in 1a5nA at the interface with 1a5nC. (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>Table 3. disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>res</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>83</td>
</tr>
<tr>
<td>92</td>
</tr>
<tr>
<td>34</td>
</tr>
<tr>
<td>31</td>
</tr>
<tr>
<td>23</td>
</tr>
<tr>
<td>90</td>
</tr>
<tr>
<td>16</td>
</tr>
<tr>
<td>81</td>
</tr>
<tr>
<td>70</td>
</tr>
<tr>
<td>72</td>
</tr>
</tbody>
</table>

Table 3. List of disruptive mutations for the top 25% of residues in 1a5nA, that are at the interface with 1a5nC.

Figure 4 shows residues in 1a5nA colored by their importance, at the interface with 1a5nC.

Interface with 1a5nC1. Table 4 lists the top 25% of residues at the interface with 1a5nC1. The following table (Table 5) suggests possible disruptive replacements for these residues (see Section 5.6).

<table>
<thead>
<tr>
<th>Table 4.</th>
</tr>
</thead>
<tbody>
<tr>
<td>res</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>89</td>
</tr>
<tr>
<td>86</td>
</tr>
<tr>
<td>83</td>
</tr>
<tr>
<td>92</td>
</tr>
<tr>
<td>88</td>
</tr>
</tbody>
</table>

continued in next column

Table 4. continued

continued in next column
Fig. 4. Residues in 1a5nA at the interface with 1a5nC, colored by their relative importance. 1a5nC is shown in backbone representation (See Appendix for the coloring scheme for the protein chain 1a5nA.)

Table 4. The top 25% of residues in 1a5nA at the interface with 1a5nC1. (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; cvg: contacts realized through backbone atoms given 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 1a5nA, that are at the interface with 1a5nC1.

Figure 5 shows residues in 1a5nA colored by their importance, at the interface with 1a5nC1.  

Interface with 1a5nB. Table 6 lists the top 25% of residues at the interface with 1a5nB. The following table (Table 7) suggests possible disruptive replacements for these residues (see Section 5.6).

Table 6. The top 25% of residues in 1a5nA at the interface with 1a5nB. (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; cvg: contacts realized through backbone atoms given 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.)
Fig. 5. Residues in 1a5nA, at the interface with 1a5nC1, colored by their relative importance. 1a5nC1 is shown in backbone representation (See Appendix for the coloring scheme for the protein chain 1a5nA.)

<table>
<thead>
<tr>
<th>Table 7.</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>72</td>
<td>G</td>
</tr>
</tbody>
</table>

Table 7. List of disruptive mutations for the top 25% of residues in 1a5nA, that are at the interface with 1a5nB.

Figure 6 shows residues in 1a5nA colored by their importance, at the interface with 1a5nB.

**Interface with 1a5nA2.** Table 8 lists the top 25% of residues at the interface with 1a5nA2. The following table (Table 9) suggests possible disruptive replacements for these residues (see Section 5.6).

<table>
<thead>
<tr>
<th>Table 8. continued</th>
</tr>
</thead>
<tbody>
<tr>
<td>res</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>70</td>
</tr>
</tbody>
</table>

Table 8. The top 25% of residues in 1a5nA at the interface with 1a5nA2. (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>
<thead>
<tr>
<th>Table 9.</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>34</td>
<td>E</td>
</tr>
<tr>
<td>30</td>
<td>L</td>
</tr>
<tr>
<td>31</td>
<td>N</td>
</tr>
<tr>
<td>23</td>
<td>R</td>
</tr>
<tr>
<td>16</td>
<td>A</td>
</tr>
<tr>
<td>70</td>
<td>M</td>
</tr>
</tbody>
</table>

Table 9. List of disruptive mutations for the top 25% of residues in 1a5nA, that are at the interface with 1a5nA2.
Figure 7 shows residues in 1a5nA colored by their importance, at the interface with 1a5nA2.

**Interface with 1a5nA1.** Table 10 lists the top 25% of residues at the interface with 1a5nA1. The following table (Table 11) suggests possible disruptive replacements for these residues (see Section 5.6).

### Table 10.
<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>48</td>
<td>R</td>
<td>(2) R (96) CD NW</td>
<td>0.12</td>
<td>68/4</td>
<td>2.88</td>
</tr>
</tbody>
</table>

**Table 10.** The top 25% of residues in 1a5nA at the interface with 1a5nA1. (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; cvg: 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 11.
<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>48</td>
<td>R</td>
<td>(T) (D) (YE) (VCAG)</td>
</tr>
</tbody>
</table>

**Table 11.** List of disruptive mutations for the top 25% of residues in 1a5nA, that are at the interface with 1a5nA1.

2.4.3 Possible novel functional surfaces at 25% coverage. One group of residues is conserved on the 1a5nA surface, away from (or substantially larger than) other functional sites and interfaces recognizable in PDB entry 1a5n. It is shown in Fig. 9. The residues belonging to this surface "patch" are listed in Table 12, while Table

Figure 8 shows residues in 1a5nA colored by their importance, at the interface with 1a5nA1.
13 suggests possible disruptive replacements for these residues (see Section 5.6).

### Table 12.

<table>
<thead>
<tr>
<th>Residue</th>
<th>Type</th>
<th>Substitutions(%)</th>
<th>CvG</th>
</tr>
</thead>
<tbody>
<tr>
<td>34</td>
<td>E</td>
<td>(2) E(97)</td>
<td>0.05</td>
</tr>
<tr>
<td>27</td>
<td>G</td>
<td>(3) G(95) AK</td>
<td>0.09</td>
</tr>
<tr>
<td>69</td>
<td>V</td>
<td>V(90) C(8) IT</td>
<td>0.13</td>
</tr>
<tr>
<td>31</td>
<td>N</td>
<td>(3) N(94) TAS(1)</td>
<td>0.15</td>
</tr>
<tr>
<td>23</td>
<td>R</td>
<td>(3) R(94) AYNH</td>
<td>0.16</td>
</tr>
<tr>
<td>70</td>
<td>M</td>
<td>M(80) L(12) Q(1)</td>
<td>0.23</td>
</tr>
<tr>
<td>72</td>
<td>G</td>
<td>G(91) S(4) DA(2) T</td>
<td>0.24</td>
</tr>
</tbody>
</table>

Table 12. Residues forming surface "patch" in 1a5nA.

### Table 13.

<table>
<thead>
<tr>
<th>Residue</th>
<th>Type</th>
<th>Disruptive Mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>34</td>
<td>E</td>
<td>(FWH)(VCAG)(YR)(T)</td>
</tr>
<tr>
<td>27</td>
<td>G</td>
<td>(E)(R)(FKWHD)(Y)</td>
</tr>
<tr>
<td>69</td>
<td>V</td>
<td>(R)(K)(E)(Y)</td>
</tr>
<tr>
<td>31</td>
<td>N</td>
<td>(Y)(H)(FW)(R)</td>
</tr>
<tr>
<td>23</td>
<td>R</td>
<td>(D)(T)(E)(CG)</td>
</tr>
<tr>
<td>70</td>
<td>M</td>
<td>(Y)(T)(H)(CG)</td>
</tr>
<tr>
<td>72</td>
<td>G</td>
<td>(R)(K)(H)(FW)</td>
</tr>
</tbody>
</table>

Table 13. Disruptive mutations for the surface patch in 1a5nA.

### 3 CHAIN 1A5NC

#### 3.1 P18314 overview

From SwissProt, id P18314, 92% identical to 1a5nC:

**Description:** Urease alpha subunit (EC 3.5.1.5) (urea amidohydrolase).

**Organism, scientific name:** Klebsiella aerogenes.

**Taxonomy:** Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Klebsiella.

**Catalytic activity:** Urea + H(2)O = CO(2) + 2 NH(3).

**Cofactor:** Binds 2 nickel ions per subunit.

**Subunit:** Alpha, beta, gamma(3).

**Ptm:** Lys-217 is carbamylated. The carbamoyl group provides the ligands for the two nickel ions.

**Similarity:** Belongs to the urease 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 1a5nC

For the chain 1a5nC, the alignment 1a5nC.msf (attached) with 352 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 Fig. 10. Residues 2-276 in 1a5nC colored by their relative importance. (See Appendix, Fig. 33, for the coloring scheme.)

To this report, under the name of 1a5nC.msf. Its statistics, from the allstat program are the following:

- Format: MSF
- Number of sequences: 352
- Total number of residues: 192477
- Smallest: 419
- Largest: 551
- Average length: 546.8
- Alignment length: 551
- Average identity: 63%
- Most related pair: 99%
- Most unrelated pair: 36%
- Most distant seq: 47%

Furthermore, 4% of residues show as conserved in this alignment. The alignment consists of 3% eukaryotic (1% fungi, 1% plantae), and 11% prokaryotic sequences. (Descriptions of some sequences were not readily available.) The file containing the sequence descriptions can be found in the attachment, under the name 1a5nC.descr.

#### 3.3 Residue ranking in 1a5nC

The 1a5nC sequence is shown in Figs. 10–11, with each residue colored according to its estimated importance. The full listing of residues in 1a5nC can be found in the file called 1a5nC.ranks sorted in the attachment.

#### 3.4 Top ranking residues in 1a5nC and their position on the structure

In the following we consider residues ranking among top 25% of residues in the protein. Figure 12 shows residues in 1a5nC 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. 13 shows the top 25% of all residues, this time colored according to clusters they belong to. The clusters in Fig.13 are composed of the residues listed...
Fig. 11. Residues 277-567 in 1a5nC colored by their relative importance. (See Appendix, Fig.33, for the coloring scheme.)

Fig. 12. Residues in 1a5nC, colored by their relative importance. Clockwise: front, back, top and bottom views.

in Table 14.

Table 14.

<table>
<thead>
<tr>
<th>cluster color</th>
<th>size</th>
<th>member residues</th>
</tr>
</thead>
</table>

Table 14. continued

<table>
<thead>
<tr>
<th>cluster color</th>
<th>size</th>
<th>member residues</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>blue</td>
<td>9</td>
<td>338, 341, 342, 344, 345, 346, 348</td>
</tr>
<tr>
<td></td>
<td></td>
<td>350, 352, 359, 360, 362, 363, 364</td>
</tr>
<tr>
<td></td>
<td></td>
<td>365, 366, 368, 369, 373, 376, 378</td>
</tr>
<tr>
<td></td>
<td></td>
<td>381, 386, 396, 397, 399, 402, 403</td>
</tr>
<tr>
<td></td>
<td></td>
<td>406, 408, 410, 411, 412, 416, 422</td>
</tr>
<tr>
<td></td>
<td></td>
<td>435, 440, 444</td>
</tr>
<tr>
<td>yellow</td>
<td>9</td>
<td>91, 129, 146, 150, 152, 190, 430</td>
</tr>
<tr>
<td></td>
<td></td>
<td>431, 451</td>
</tr>
<tr>
<td>green</td>
<td>7</td>
<td>14, 15, 16, 18, 19, 35, 83</td>
</tr>
<tr>
<td>purple</td>
<td>3</td>
<td>158, 160, 161</td>
</tr>
<tr>
<td>azure</td>
<td>3</td>
<td>67, 113, 116</td>
</tr>
<tr>
<td>turquoise</td>
<td>2</td>
<td>459, 470</td>
</tr>
<tr>
<td>brown</td>
<td>2</td>
<td>295, 522</td>
</tr>
<tr>
<td>coral</td>
<td>2</td>
<td>40, 42</td>
</tr>
<tr>
<td>magenta</td>
<td>2</td>
<td>77, 80</td>
</tr>
<tr>
<td>LightSalmon</td>
<td>2</td>
<td>93, 427</td>
</tr>
<tr>
<td>SkyBlue</td>
<td>2</td>
<td>233, 266</td>
</tr>
<tr>
<td>violet</td>
<td>2</td>
<td>242, 269</td>
</tr>
<tr>
<td>gold</td>
<td>2</td>
<td>85, 97</td>
</tr>
</tbody>
</table>

Table 14. continued in next column

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.
Formic acid binding site. Table 15 lists the top 25% of residues at the interface with 1a5nFMT999 (formic acid). The following table (Table 16) suggests possible disruptive replacements for these residues (see Section 5.6).

**Table 15.** The top 25% of residues in 1a5nC at the interface with formic acid. (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>subst's</th>
<th>cvg</th>
<th>noc/bb</th>
<th>dist (Å)</th>
<th>antn</th>
</tr>
</thead>
<tbody>
<tr>
<td>219</td>
<td>H</td>
<td>H(100)</td>
<td>0.05</td>
<td>15/0</td>
<td>2.84</td>
<td>site</td>
</tr>
<tr>
<td>272</td>
<td>H</td>
<td>H(100)</td>
<td>0.05</td>
<td>4/0</td>
<td>3.66</td>
<td>site</td>
</tr>
<tr>
<td>134</td>
<td>H</td>
<td>H(99)C</td>
<td>0.07</td>
<td>10/0</td>
<td>3.20</td>
<td>site</td>
</tr>
<tr>
<td>136</td>
<td>H</td>
<td>H(99)W</td>
<td>0.07</td>
<td>8/0</td>
<td>2.83</td>
<td></td>
</tr>
<tr>
<td>360</td>
<td>D</td>
<td>D(99)R</td>
<td>0.07</td>
<td>1/0</td>
<td>4.16</td>
<td>site</td>
</tr>
<tr>
<td>246</td>
<td>H</td>
<td>H(99)Q</td>
<td>0.14</td>
<td>10/1</td>
<td>3.04</td>
<td></td>
</tr>
<tr>
<td>169</td>
<td>T</td>
<td>T(92)</td>
<td>0.22</td>
<td>11/4</td>
<td>3.46</td>
<td></td>
</tr>
</tbody>
</table>

**Table 16.** List of disruptive mutations for the top 25% of residues in 1a5nC, that are at the interface with formic acid.

**Table 17.** The top 25% of residues in 1a5nC at the interface with 1a5nC2. (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 17. 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>220</td>
<td>E</td>
<td>E(100)</td>
<td>0.05</td>
<td>9/8</td>
<td>3.61</td>
</tr>
<tr>
<td>221</td>
<td>D</td>
<td>D(100)</td>
<td>0.05</td>
<td>61/32</td>
<td>3.15</td>
</tr>
<tr>
<td>252</td>
<td>E</td>
<td>E(100)</td>
<td>0.05</td>
<td>12/0</td>
<td>2.82</td>
</tr>
<tr>
<td>315</td>
<td>M</td>
<td>M(100)</td>
<td>0.05</td>
<td>15/3</td>
<td>3.32</td>
</tr>
<tr>
<td>365</td>
<td>G</td>
<td>G(100)</td>
<td>0.05</td>
<td>23/23</td>
<td>3.07</td>
</tr>
<tr>
<td>366</td>
<td>R</td>
<td>R(99)P</td>
<td>0.07</td>
<td>52/14</td>
<td>3.16</td>
</tr>
<tr>
<td>369</td>
<td>E</td>
<td>E(99)R</td>
<td>0.07</td>
<td>12/1</td>
<td>2.91</td>
</tr>
</tbody>
</table>

Fig. 14. Residues in 1a5nC, at the interface with formic acid, colored by their relative importance. The ligand (formic acid) 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 1a5nC.)
contacts realized through backbone atoms given in the bracket; dist: distance of closest approach to the ligand. )

Table 18.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>220</td>
<td>E</td>
<td>(FWH) (YVCARG) (T) (SNKLPI)</td>
</tr>
<tr>
<td>221</td>
<td>D</td>
<td>(R) (FWH) (KYVCAG) (TQM)</td>
</tr>
<tr>
<td>252</td>
<td>E</td>
<td>(FWH) (YVCARG) (T) (SNKLPI)</td>
</tr>
<tr>
<td>315</td>
<td>M</td>
<td>(Y) (TH) (SCRG) (FWD)</td>
</tr>
<tr>
<td>365</td>
<td>G</td>
<td>(KER) (FQMWHD) (NYLPI) (SVA)</td>
</tr>
<tr>
<td>366</td>
<td>R</td>
<td>(T) (YD) (SECG) (VA)</td>
</tr>
<tr>
<td>369</td>
<td>E</td>
<td>(FW) (YVCAHG) (T) (SLPIR)</td>
</tr>
<tr>
<td>160</td>
<td>G</td>
<td>(KER) (FQMWHD) (NLPI) (Y)</td>
</tr>
<tr>
<td>364</td>
<td>M</td>
<td>(Y) (T) (HR) (CG)</td>
</tr>
<tr>
<td>223</td>
<td>G</td>
<td>(KER) (QHD) (FYMW) (N)</td>
</tr>
<tr>
<td>363</td>
<td>A</td>
<td>(KER) (Y) (QHD) (N)</td>
</tr>
<tr>
<td>196</td>
<td>K</td>
<td>(Y) (T) (FW) (SVCAG)</td>
</tr>
<tr>
<td>362</td>
<td>Q</td>
<td>(Y) (TH) (FW) (SCG)</td>
</tr>
<tr>
<td>368</td>
<td>G</td>
<td>(E) (KD) (R) (M)</td>
</tr>
<tr>
<td>222</td>
<td>W</td>
<td>(K) (E) (T) (QD)</td>
</tr>
<tr>
<td>161</td>
<td>P</td>
<td>(R) (Y) (H) (K)</td>
</tr>
<tr>
<td>158</td>
<td>G</td>
<td>(FKWR) (E) (H) (M)</td>
</tr>
</tbody>
</table>

Table 18. List of disruptive mutations for the top 25% of residues in 1a5nC, that are at the interface with 1a5nC2.

Fig. 15. Residues in 1a5nC, at the interface with 1a5nC2, colored by their relative importance. 1a5nC2 is shown in backbone representation (See Appendix for the coloring scheme for the protein chain 1a5nC.)

Table 19.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>subst's</th>
<th>cvg</th>
<th>noc/ bb</th>
<th>dist</th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>R</td>
<td>R (100)</td>
<td>0.05</td>
<td>62/15</td>
<td>2.82</td>
</tr>
<tr>
<td>67</td>
<td>D</td>
<td>D (100)</td>
<td>0.05</td>
<td>30/8</td>
<td>3.31</td>
</tr>
<tr>
<td>116</td>
<td>T</td>
<td>T (99) S</td>
<td>0.09</td>
<td>21/20</td>
<td>2.85</td>
</tr>
<tr>
<td>46</td>
<td>G</td>
<td>G (99) A</td>
<td>0.11</td>
<td>21/21</td>
<td>3.63</td>
</tr>
<tr>
<td>451</td>
<td>G</td>
<td>G (99). P</td>
<td>0.12</td>
<td>17/17</td>
<td>2.69</td>
</tr>
<tr>
<td>47</td>
<td>G</td>
<td>G (95)</td>
<td>0.13</td>
<td>15/15</td>
<td>3.15</td>
</tr>
<tr>
<td>459</td>
<td>G</td>
<td>G (99)</td>
<td>0.15</td>
<td>2/2</td>
<td>4.09</td>
</tr>
<tr>
<td>48</td>
<td>G</td>
<td>G (99) A</td>
<td>0.18</td>
<td>3/3</td>
<td>3.91</td>
</tr>
<tr>
<td>91</td>
<td>G</td>
<td>G (98) AS</td>
<td>0.18</td>
<td>8/8</td>
<td>3.86</td>
</tr>
<tr>
<td>464</td>
<td>S</td>
<td>S (95)</td>
<td>0.20</td>
<td>82/59</td>
<td>2.63</td>
</tr>
<tr>
<td>113</td>
<td>G</td>
<td>AG (97)</td>
<td>0.22</td>
<td>4/4</td>
<td>3.27</td>
</tr>
</tbody>
</table>

Table 19. The top 25% of residues in 1a5nC at the interface with 1a5nC1. (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 20.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>R</td>
<td>(TD) (SYEVCLAPIG) (FYMW) (N)</td>
</tr>
<tr>
<td>67</td>
<td>D</td>
<td>(R) (FWH) (KYVCAG) (TQM)</td>
</tr>
<tr>
<td>116</td>
<td>T</td>
<td>(KR) (QFMWH) (NELPI) (D)</td>
</tr>
<tr>
<td>46</td>
<td>G</td>
<td>(KER) (QHD) (FYMW) (N)</td>
</tr>
<tr>
<td>451</td>
<td>G</td>
<td>(R) (KE) (H) (FWD)</td>
</tr>
<tr>
<td>47</td>
<td>G</td>
<td>(KE) (R) (QD) (MH)</td>
</tr>
<tr>
<td>459</td>
<td>G</td>
<td>(KER) (FQMWHD) (NLPI) (Y)</td>
</tr>
<tr>
<td>48</td>
<td>G</td>
<td>(KER) (QHD) (FYMW) (N)</td>
</tr>
<tr>
<td>91</td>
<td>G</td>
<td>(R) (E) (KH) (FW)</td>
</tr>
<tr>
<td>464</td>
<td>S</td>
<td>(KR) (H) (Q) (FMW)</td>
</tr>
<tr>
<td>113</td>
<td>G</td>
<td>(KR) (E) (Q) (H)</td>
</tr>
</tbody>
</table>

Table 20. List of disruptive mutations for the top 25% of residues in 1a5nC, that are at the interface with 1a5nC1.

Figure 16 shows residues in 1a5nC colored by their importance, at the interface with 1a5nC1.

Nickel (ii) ion binding site. Table 21 lists the top 25% of residues at the interface with 1a5nNI574 (nickel (ii) ion). The following table (Table 22) suggests possible disruptive replacements for these residues (see Section 5.6).

Interface with 1a5nC1. Table 19 lists the top 25% of residues at the interface with 1a5nC1. The following table (Table 20) suggests possible disruptive replacements for these residues (see Section 5.6).

Figure 15 shows residues in 1a5nC colored by their importance, at the interface with 1a5nC2.
Table 21. The top 25% of residues in 1a5nC at the interface with nickel (ii) ion. (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>219</td>
<td>H</td>
<td>H(100)</td>
<td>0.05</td>
<td>4/0</td>
<td>3.36</td>
<td>site</td>
</tr>
<tr>
<td>272</td>
<td>H</td>
<td>H(100)</td>
<td>0.05</td>
<td>5/0</td>
<td>2.51</td>
<td>site</td>
</tr>
<tr>
<td>277</td>
<td>G</td>
<td>G(100)</td>
<td>0.05</td>
<td>1/1</td>
<td>4.10</td>
<td></td>
</tr>
<tr>
<td>134</td>
<td>H</td>
<td>H(99)C</td>
<td>0.07</td>
<td>3/0</td>
<td>3.60</td>
<td>site</td>
</tr>
<tr>
<td>136</td>
<td>H</td>
<td>H(99)W</td>
<td>0.07</td>
<td>1/0</td>
<td>4.94</td>
<td></td>
</tr>
<tr>
<td>360</td>
<td>D</td>
<td>D(99)R</td>
<td>0.07</td>
<td>3/0</td>
<td>4.46</td>
<td>site</td>
</tr>
<tr>
<td>246</td>
<td>H</td>
<td>H(99)Q</td>
<td>0.14</td>
<td>6/0</td>
<td>2.53</td>
<td></td>
</tr>
</tbody>
</table>

Table 22. List of disruptive mutations for the top 25% of residues in 1a5nC, that are at the interface with nickel (ii) ion. (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>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>219</td>
<td>H</td>
<td>(E) (TQMD) (SNKVCAP vacancies) (YR)</td>
</tr>
<tr>
<td>272</td>
<td>H</td>
<td>(E) (TQMD) (SNKVCAP vacancies) (YR)</td>
</tr>
<tr>
<td>277</td>
<td>G</td>
<td>(KER) (FQMWHD) (NYLPI) (SVA)</td>
</tr>
<tr>
<td>134</td>
<td>H</td>
<td>(E) (QMD) (K) (TNLPI)</td>
</tr>
<tr>
<td>136</td>
<td>H</td>
<td>(E) (TQD) (KM) (SNCG)</td>
</tr>
<tr>
<td>360</td>
<td>D</td>
<td>(FW) (YVCAHRG) (T) (KM)</td>
</tr>
<tr>
<td>246</td>
<td>H</td>
<td>(TE) (D) (SVNCAG) (QLPI)</td>
</tr>
</tbody>
</table>

Table 23. The top 25% of residues in 1a5nC at the interface with 1a5nB2. (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>
</tr>
</thead>
<tbody>
<tr>
<td>251</td>
<td>N</td>
<td>N(100)</td>
<td>0.05</td>
<td>41/22</td>
<td>2.79</td>
</tr>
<tr>
<td>252</td>
<td>E</td>
<td>E(100)</td>
<td>0.05</td>
<td>89/32</td>
<td>2.71</td>
</tr>
<tr>
<td>282</td>
<td>P</td>
<td>P(100)</td>
<td>0.05</td>
<td>4/1</td>
<td>4.01</td>
</tr>
<tr>
<td>283</td>
<td>D</td>
<td>D(98)</td>
<td>0.07</td>
<td>13/0</td>
<td>2.95</td>
</tr>
<tr>
<td>335</td>
<td>S</td>
<td>S(97)</td>
<td>0.08</td>
<td>5/0</td>
<td>3.92</td>
</tr>
<tr>
<td>281</td>
<td>A</td>
<td>A(96)</td>
<td>0.13</td>
<td>5/5</td>
<td>4.34</td>
</tr>
<tr>
<td>254</td>
<td>G</td>
<td>G(96)</td>
<td>0.18</td>
<td>21/21</td>
<td>3.19</td>
</tr>
<tr>
<td>332</td>
<td>F</td>
<td>F(98) IM</td>
<td>0.18</td>
<td>15/9</td>
<td>3.06</td>
</tr>
</tbody>
</table>

Figure 17 shows residues in 1a5nC colored by their importance, at the interface with 1a5nB2.
Table 24. List of disruptive mutations for the top 25% of residues in 1a5nC, that are at the interface with 1a5nB2.

Table 25. The top 25% of residues in 1a5nC at the interface with 1a5nA. (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 26. List of disruptive mutations for the top 25% of residues in 1a5nC, that are at the interface with 1a5nA.

Fig. 18. Residues in 1a5nC, at the interface with 1a5nB2, colored by their relative importance. 1a5nB2 is shown in backbone representation (See Appendix for the coloring scheme for the protein chain 1a5nC.)

Figure 18 shows residues in 1a5nC colored by their importance, at the interface with 1a5nB2.

Interface with 1a5nA. Table 25 lists the top 25% of residues at the interface with 1a5nA. The following table (Table 26) suggests possible disruptive replacements for these residues (see Section 5.6).

Table 27 lists the top 25% of residues at the interface with 1a5nNI575 (nickel (ii) ion). The following table (Table 28) suggests possible disruptive replacements for these residues (see Section 5.6).

Nickel (ii) ion binding site. Table 27 lists the top 25% of residues at the interface with 1a5nNI575 (nickel (ii) ion). The following table (Table 28) suggests possible disruptive replacements for these residues (see Section 5.6).
Table 27. The top 25% of residues in 1a5nC at the interface with nickel (ii) ion. (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>subst’s (%)</th>
<th>cvg</th>
<th>noc/ bb</th>
<th>dist (Å)</th>
<th>antn</th>
</tr>
</thead>
<tbody>
<tr>
<td>219</td>
<td>H</td>
<td>H(100)</td>
<td>0.05</td>
<td>1/0</td>
<td>4.74</td>
<td>site</td>
</tr>
<tr>
<td>272</td>
<td>H</td>
<td>H(100)</td>
<td>0.05</td>
<td>2/0</td>
<td>4.73</td>
<td>site</td>
</tr>
<tr>
<td>134</td>
<td>H</td>
<td>H(99)C</td>
<td>0.07</td>
<td>5/0</td>
<td>2.31</td>
<td>site</td>
</tr>
<tr>
<td>136</td>
<td>H</td>
<td>H(99)W</td>
<td>0.07</td>
<td>5/0</td>
<td>2.27</td>
<td>site</td>
</tr>
<tr>
<td>360</td>
<td>D</td>
<td>D(99)R</td>
<td>0.07</td>
<td>5/1</td>
<td>2.03</td>
<td>site</td>
</tr>
<tr>
<td>363</td>
<td>A</td>
<td>A(93)</td>
<td>0.12</td>
<td>2/1</td>
<td>4.11</td>
<td></td>
</tr>
<tr>
<td>169</td>
<td>T</td>
<td>T(92)</td>
<td>0.22</td>
<td>1/0</td>
<td>4.40</td>
<td></td>
</tr>
</tbody>
</table>

Table 28. List of disruptive mutations for the top 25% of residues in 1a5nC, that are at the interface with nickel (ii) ion.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>219</td>
<td>H</td>
<td>(E)(TQMD)(SNKVCAPMG)(YR)</td>
</tr>
<tr>
<td>272</td>
<td>H</td>
<td>(E)(TQMD)(SNKVCAPMG)(YR)</td>
</tr>
<tr>
<td>134</td>
<td>H</td>
<td>(E)(QMD)(K)(TNLPI)</td>
</tr>
<tr>
<td>136</td>
<td>H</td>
<td>(E)(TQD)(KM)(SNCG)</td>
</tr>
<tr>
<td>360</td>
<td>D</td>
<td>(FW)(YVCAHRG)(T)(KM)</td>
</tr>
<tr>
<td>363</td>
<td>A</td>
<td>(KER)(Y)(QHD)(N)</td>
</tr>
<tr>
<td>169</td>
<td>T</td>
<td>(R)(K)(H)(FW)</td>
</tr>
</tbody>
</table>

Table 29. The top 25% of residues in 1a5nC at the interface with 1a5nB. (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>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>G</td>
<td>G(100)</td>
<td>0.05</td>
<td>40/40</td>
<td>3.46</td>
</tr>
<tr>
<td>49</td>
<td>K</td>
<td>K(100)</td>
<td>0.05</td>
<td>21/6</td>
<td>3.21</td>
</tr>
<tr>
<td>52</td>
<td>R</td>
<td>R(100)</td>
<td>0.05</td>
<td>10/0</td>
<td>3.89</td>
</tr>
<tr>
<td>42</td>
<td>E</td>
<td>E(99)Q</td>
<td>0.08</td>
<td>13/7</td>
<td>3.17</td>
</tr>
<tr>
<td>23</td>
<td>L</td>
<td>L(99) .Q</td>
<td>0.15</td>
<td>1/1</td>
<td>4.83</td>
</tr>
<tr>
<td>54</td>
<td>G</td>
<td>G(93)</td>
<td>0.16</td>
<td>15/15</td>
<td>3.02</td>
</tr>
<tr>
<td>48</td>
<td>G</td>
<td>G(99)A</td>
<td>0.18</td>
<td>10/10</td>
<td>3.77</td>
</tr>
<tr>
<td>14</td>
<td>G</td>
<td>G(99).A</td>
<td>0.19</td>
<td>1/1</td>
<td>4.51</td>
</tr>
<tr>
<td>16</td>
<td>T</td>
<td>T(98).</td>
<td>0.19</td>
<td>1/0</td>
<td>4.46</td>
</tr>
</tbody>
</table>

Table 29. 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>18</td>
<td>G</td>
<td>R(1)P</td>
<td>0.19</td>
<td>10/10</td>
<td>4.49</td>
</tr>
<tr>
<td>19</td>
<td>D</td>
<td>G(97).N</td>
<td>0.19</td>
<td>44/8</td>
<td>3.11</td>
</tr>
<tr>
<td>102</td>
<td>P</td>
<td>P(97)</td>
<td>0.23</td>
<td>28/25</td>
<td>2.82</td>
</tr>
<tr>
<td>15</td>
<td>P</td>
<td>P(97).A</td>
<td>0.24</td>
<td>15/3</td>
<td>3.64</td>
</tr>
<tr>
<td>22</td>
<td>R</td>
<td>R(96).K</td>
<td>0.25</td>
<td>78/24</td>
<td>2.69</td>
</tr>
</tbody>
</table>

Figure 20 shows residues in 1a5nC colored by their relative importance, at the interface with nickel (ii) ion.

Interface with 1a5nB. Table 29 lists the top 25% of residues at the interface with 1a5nB. The following table (Table 30) suggests possible disruptive replacements for these residues (see Section 5.6).

Table 30. List of disruptive mutations for the top 25% of residues in 1a5nC at the interface with 1a5nB. (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>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>G</td>
<td>(KER)(FW)WHD)(NYLPI)(SVA)</td>
</tr>
<tr>
<td>49</td>
<td>K</td>
<td>(Y)(FTW)(SVA)AG)(HD)</td>
</tr>
</tbody>
</table>

continued in next column
Table 30. List of disruptive mutations for the top 25% of residues in 1a5nC, that are at the interface with 1a5nB.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>R</td>
<td>(TD) (SYEVCAPGL) (FMW) (N)</td>
</tr>
<tr>
<td>42</td>
<td>E</td>
<td>(FWH) (Y) (VCA) (TR)</td>
</tr>
<tr>
<td>23</td>
<td>L</td>
<td>(Y) (R) (TH) (CG)</td>
</tr>
<tr>
<td>54</td>
<td>G</td>
<td>(E) (K) (FW) (MDR)</td>
</tr>
<tr>
<td>48</td>
<td>G</td>
<td>(KER) (QH) (EYMW) (N)</td>
</tr>
<tr>
<td>14</td>
<td>G</td>
<td>(KER) (HD) (Q) (FMW)</td>
</tr>
<tr>
<td>16</td>
<td>T</td>
<td>(KR) (FWH) (QM) (E)</td>
</tr>
<tr>
<td>18</td>
<td>G</td>
<td>(R) (FKWH) (E) (M)</td>
</tr>
<tr>
<td>19</td>
<td>D</td>
<td>(R) (H) (FW) (Y)</td>
</tr>
<tr>
<td>102</td>
<td>P</td>
<td>(Y) (R) (H) (T)</td>
</tr>
<tr>
<td>15</td>
<td>P</td>
<td>(Y) (R) (H) (T)</td>
</tr>
<tr>
<td>22</td>
<td>R</td>
<td>(T) (D) (YE) (SCG)</td>
</tr>
</tbody>
</table>

Table 31. The top 25% of residues in 1a5nC at the interface with 1a5nA2. (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>
<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>311</td>
<td>E</td>
<td>E (100)</td>
<td>0.05</td>
<td>52/6</td>
<td>3.01</td>
</tr>
<tr>
<td>315</td>
<td>M</td>
<td>M (100)</td>
<td>0.05</td>
<td>6/2</td>
<td>4.11</td>
</tr>
<tr>
<td>366</td>
<td>R</td>
<td>R (99)P</td>
<td>0.07</td>
<td>48/0</td>
<td>2.81</td>
</tr>
<tr>
<td>369</td>
<td>E</td>
<td>E (99)R</td>
<td>0.07</td>
<td>14/0</td>
<td>2.60</td>
</tr>
<tr>
<td>300</td>
<td>P</td>
<td>P (99)T</td>
<td>0.10</td>
<td>1/1</td>
<td>4.91</td>
</tr>
<tr>
<td>373</td>
<td>R</td>
<td>R (98)KC</td>
<td>0.17</td>
<td>1/0</td>
<td>3.01</td>
</tr>
</tbody>
</table>

Table 32. List of disruptive mutations for the top 25% of residues in 1a5nC, that are at the interface with 1a5nA2.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>303</td>
<td>P</td>
<td>(YR) (TH) (SKECG) (FQWD)</td>
</tr>
<tr>
<td>311</td>
<td>E</td>
<td>(FWH) (YVCA) (T) (SNKLPI)</td>
</tr>
<tr>
<td>315</td>
<td>M</td>
<td>(Y) (TH) (SCRG) (FWD)</td>
</tr>
<tr>
<td>366</td>
<td>R</td>
<td>(T) (YD) (SECG) (VA)</td>
</tr>
<tr>
<td>369</td>
<td>E</td>
<td>(FW) (YVCA) (T) (SLPIR)</td>
</tr>
<tr>
<td>300</td>
<td>P</td>
<td>(R) (YH) (R) (E)</td>
</tr>
<tr>
<td>373</td>
<td>R</td>
<td>(D) (Y) (T) (E)</td>
</tr>
<tr>
<td>376</td>
<td>Q</td>
<td>(Y) (T) (FW) (H)</td>
</tr>
<tr>
<td>368</td>
<td>G</td>
<td>(E) (RD) (R) (M)</td>
</tr>
<tr>
<td>307</td>
<td>N</td>
<td>(Y) (R) (H) (FTEW)</td>
</tr>
<tr>
<td>470</td>
<td>P</td>
<td>(YR) (TH) (SCG) (KE)</td>
</tr>
<tr>
<td>305</td>
<td>T</td>
<td>(KR) (QH) (FEMW) (N)</td>
</tr>
<tr>
<td>314</td>
<td>D</td>
<td>(R) (K) (H) (FW)</td>
</tr>
</tbody>
</table>

Figure 21 shows residues in 1a5nC, at the interface with 1a5nB, colored by their importance. 1a5nB is shown in backbone representation (See Appendix for the coloring scheme for the protein chain 1a5nC.)

Figure 22 shows residues in 1a5nC colored by their importance, at the interface with 1a5nA2.
Fig. 22. Residues in 1a5nC, at the interface with 1a5nA2, colored by their relative importance. 1a5nA2 is shown in backbone representation (See Appendix for the coloring scheme for the protein chain 1a5nC.)

3.4.3 Possible novel functional surfaces at 25% coverage. One group of residues is conserved on the 1a5nC surface, away from (or substantially larger than) other functional sites and interfaces recognizable in PDB entry 1a5n. It is shown in Fig. 23. The residues belonging to this surface "patch" are listed in Table 33, while Table 34 suggests possible disruptive replacements for these residues (see Section 5.6).

<table>
<thead>
<tr>
<th>res type</th>
<th>substitutions(%)</th>
<th>cvg</th>
</tr>
</thead>
<tbody>
<tr>
<td>85 D</td>
<td>D(98)NS(1)</td>
<td>0.07</td>
</tr>
<tr>
<td>97 G</td>
<td>G(99)C</td>
<td>0.11</td>
</tr>
<tr>
<td>35 D</td>
<td>D(99)AN</td>
<td>0.13</td>
</tr>
<tr>
<td>83 K</td>
<td>K(99)VA</td>
<td>0.16</td>
</tr>
<tr>
<td>16 T</td>
<td>T(98).R(1)P</td>
<td>0.19</td>
</tr>
<tr>
<td>18 G</td>
<td>G(97).N(1)D</td>
<td>0.19</td>
</tr>
<tr>
<td>19 D</td>
<td>D(99).A</td>
<td>0.19</td>
</tr>
<tr>
<td>15 P</td>
<td>P(97).A(1)VI</td>
<td>0.24</td>
</tr>
</tbody>
</table>

Table 33. Residues forming surface "patch" in 1a5nC.

<table>
<thead>
<tr>
<th>res type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>85 D</td>
<td>(R)(FWH)(Y)(K)</td>
</tr>
<tr>
<td>97 G</td>
<td>(KER)(FQMWHD)(NYLPI)(SVA)</td>
</tr>
<tr>
<td>35 D</td>
<td>(R)(H)(FW)(Y)</td>
</tr>
<tr>
<td>83 K</td>
<td>(Y)(FTW)(H)(D)</td>
</tr>
<tr>
<td>16 T</td>
<td>(KR)(FWH)(QM)(E)</td>
</tr>
<tr>
<td>18 G</td>
<td>(R)(FKWH)(E)(M)</td>
</tr>
<tr>
<td>19 D</td>
<td>(R)(H)(FW)(Y)</td>
</tr>
<tr>
<td>15 P</td>
<td>(R)(Y)(H)(T)</td>
</tr>
</tbody>
</table>

Table 34. Disruptive mutations for the surface patch in 1a5nC.

Another group of surface residues is shown in Fig. 24. The residues belonging to this surface "patch" are listed in Table 35, while Table

Fig. 23. A possible active surface on the chain 1a5nC.

Fig. 24. Another possible active surface on the chain 1a5nC.
Table 35. Residues forming surface "patch" in 1a5nC.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>substitutions(%)</th>
<th>cvg</th>
</tr>
</thead>
<tbody>
<tr>
<td>49</td>
<td>K</td>
<td>K(100)</td>
<td>0.05</td>
</tr>
<tr>
<td>52</td>
<td>R</td>
<td>R(100)</td>
<td>0.05</td>
</tr>
<tr>
<td>46</td>
<td>G</td>
<td>G(99)A</td>
<td>0.11</td>
</tr>
<tr>
<td>47</td>
<td>G</td>
<td>G(95)A(2)CF(1)</td>
<td>0.13</td>
</tr>
<tr>
<td>54</td>
<td>G</td>
<td>G(93)S(6)R</td>
<td>0.16</td>
</tr>
<tr>
<td>48</td>
<td>G</td>
<td>G(99)A</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Table 36. Disruptive mutations for the surface patch in 1a5nC.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>49</td>
<td>K</td>
<td>(Y) (FTW) (SVCAG) (HD)</td>
</tr>
<tr>
<td>52</td>
<td>R</td>
<td>(TD) (SYEVCALAPIG) (FMW) (N)</td>
</tr>
<tr>
<td>46</td>
<td>G</td>
<td>(KER) (QHD) (FYMW) (N)</td>
</tr>
<tr>
<td>47</td>
<td>G</td>
<td>(KE) (R) (QD) (MH)</td>
</tr>
<tr>
<td>54</td>
<td>G</td>
<td>(E) (K) (FW) (MDR)</td>
</tr>
<tr>
<td>48</td>
<td>G</td>
<td>(KER) (QHD) (FYMW) (N)</td>
</tr>
</tbody>
</table>

Table 37. Residues forming surface "patch" in 1a5nC.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>substitutions(%)</th>
<th>cvg</th>
</tr>
</thead>
<tbody>
<tr>
<td>67</td>
<td>D</td>
<td>D(100)</td>
<td>0.05</td>
</tr>
<tr>
<td>116</td>
<td>T</td>
<td>T(99)S</td>
<td>0.09</td>
</tr>
<tr>
<td>427</td>
<td>G</td>
<td>G(99)KPN</td>
<td>0.09</td>
</tr>
<tr>
<td>451</td>
<td>G</td>
<td>G(99)P</td>
<td>0.12</td>
</tr>
<tr>
<td>91</td>
<td>G</td>
<td>G(98)ASNQ</td>
<td>0.18</td>
</tr>
<tr>
<td>93</td>
<td>I</td>
<td>I(98)V(1)</td>
<td>0.24</td>
</tr>
</tbody>
</table>

Table 38. Disruptive mutations for the surface patch in 1a5nC.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>67</td>
<td>D</td>
<td>(R) (FWH) (KYVCAG) (TQM)</td>
</tr>
<tr>
<td>116</td>
<td>T</td>
<td>(KR) (FQMWH) (NELPI) (D)</td>
</tr>
<tr>
<td>427</td>
<td>G</td>
<td>(ER) (FYWH) (KD) (M)</td>
</tr>
<tr>
<td>451</td>
<td>G</td>
<td>(R) (KE) (H) (FWD)</td>
</tr>
<tr>
<td>91</td>
<td>G</td>
<td>(R) (E) (KH) (FW)</td>
</tr>
<tr>
<td>93</td>
<td>I</td>
<td>(YR) (H) (TRE) (SQCIGD)</td>
</tr>
</tbody>
</table>

Another group of surface residues is shown in Fig.25. The residues belonging to this surface "patch" are listed in Table 37, while Table 38 suggests possible disruptive replacements for these residues (see Section 5.6).

Fig. 25. Another possible active surface on the chain 1a5nC.

Another group of surface residues is shown in Fig.26. The right panel shows (in blue) the rest of the larger cluster this surface belongs to.

Fig. 26. Another possible active surface on the chain 1a5nC. The larger cluster it belongs to is shown in blue.

The residues belonging to this surface "patch" are listed in Table 39, while Table 40 suggests possible disruptive replacements for these residues (see Section 5.6).

Table 39.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>substitutions(%)</th>
<th>cvg</th>
<th>antn</th>
</tr>
</thead>
<tbody>
<tr>
<td>219</td>
<td>H</td>
<td>H(100)</td>
<td>0.05</td>
<td>site</td>
</tr>
<tr>
<td>220</td>
<td>E</td>
<td>E(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>221</td>
<td>D</td>
<td>D(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>248</td>
<td>D</td>
<td>D(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>251</td>
<td>N</td>
<td>N(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>252</td>
<td>E</td>
<td>E(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>272</td>
<td>H</td>
<td>H(100)</td>
<td>0.05</td>
<td>site</td>
</tr>
</tbody>
</table>

continued in next column
Table 39. continued

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>substitutions(%)</th>
<th>cvg</th>
<th>antn</th>
</tr>
</thead>
<tbody>
<tr>
<td>274</td>
<td>E</td>
<td>E(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>277</td>
<td>G</td>
<td>G(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>282</td>
<td>P</td>
<td>P(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>298</td>
<td>T</td>
<td>T(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>311</td>
<td>E</td>
<td>E(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>315</td>
<td>M</td>
<td>M(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>336</td>
<td>R</td>
<td>R(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>338</td>
<td>R</td>
<td>R(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>365</td>
<td>G</td>
<td>G(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>369</td>
<td>E</td>
<td>E(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>370</td>
<td>M</td>
<td>M(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>372</td>
<td>T</td>
<td>T(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>374</td>
<td>E</td>
<td>E(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>375</td>
<td>M</td>
<td>M(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>376</td>
<td>R</td>
<td>R(100)</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>377</td>
<td>M</td>
<td>M(100)</td>
<td>0.05</td>
<td></td>
</tr>
</tbody>
</table>

Table 40. continued

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>298</td>
<td>T</td>
<td>(KR)(FQMWHN)(YLM)</td>
</tr>
<tr>
<td>311</td>
<td>E</td>
<td>(FWH)(YVCCARG)(T)(SNKLPI)</td>
</tr>
<tr>
<td>315</td>
<td>M</td>
<td>(Y)(TH)(SCRG)(FWD)</td>
</tr>
<tr>
<td>336</td>
<td>R</td>
<td>(TD)(SYECLAPIC)(FMY)(N)</td>
</tr>
<tr>
<td>338</td>
<td>R</td>
<td>(TD)(SYECLAPIC)(FMY)(N)</td>
</tr>
<tr>
<td>365</td>
<td>G</td>
<td>(KER)(FQMWHD)(NLYPL)(SVL)</td>
</tr>
<tr>
<td>369</td>
<td>E</td>
<td>(FW)(YVCCARG)(T)(SLPRL)</td>
</tr>
<tr>
<td>370</td>
<td>M</td>
<td>(Y)(TM)(SCRG)(FWD)</td>
</tr>
<tr>
<td>372</td>
<td>T</td>
<td>(KR)(FQMWHN)(YLM)</td>
</tr>
<tr>
<td>374</td>
<td>G</td>
<td>(FW)(YVCCARG)(T)(SNKLPI)</td>
</tr>
<tr>
<td>376</td>
<td>Q</td>
<td>(Y)(TH)(FW)(SCG)</td>
</tr>
<tr>
<td>378</td>
<td>P</td>
<td>(Y)(TM)(FW)(H)</td>
</tr>
<tr>
<td>223</td>
<td>G</td>
<td>(R)(K)(E)(H)</td>
</tr>
<tr>
<td>336</td>
<td>A</td>
<td>(K)(E)(T)(R)</td>
</tr>
<tr>
<td>300</td>
<td>N</td>
<td>(Y)(R)(H)(FTEW)</td>
</tr>
<tr>
<td>312</td>
<td>H</td>
<td>(E)(Q)(TKD)(M)</td>
</tr>
<tr>
<td>362</td>
<td>Q</td>
<td>(Y)(TH)(FW)(SCG)</td>
</tr>
<tr>
<td>376</td>
<td>Q</td>
<td>(Y)(TH)(FW)(H)</td>
</tr>
<tr>
<td>254</td>
<td>G</td>
<td>(R)(K)(E)(H)</td>
</tr>
<tr>
<td>332</td>
<td>F</td>
<td>(K)(E)(T)(R)</td>
</tr>
<tr>
<td>299</td>
<td>N</td>
<td>(Y)(FWH)(R)(E)</td>
</tr>
<tr>
<td>368</td>
<td>G</td>
<td>(E)(K)(D)(R)(M)</td>
</tr>
<tr>
<td>370</td>
<td>N</td>
<td>(Y)(R)(H)(FTEW)</td>
</tr>
<tr>
<td>222</td>
<td>W</td>
<td>(K)(E)(T)(QD)</td>
</tr>
<tr>
<td>305</td>
<td>T</td>
<td>(KR)(QH)(FEMW)(N)</td>
</tr>
<tr>
<td>314</td>
<td>D</td>
<td>(R)(K)(H)(FW)</td>
</tr>
<tr>
<td>333</td>
<td>A</td>
<td>(KR)(E)(Y)(QH)</td>
</tr>
<tr>
<td>249</td>
<td>T</td>
<td>(KR)(FQMWHN)(E)(NLM)</td>
</tr>
</tbody>
</table>

Table 39. Residues forming surface "patch" in 1a5nC.

Table 40. Disruptive mutations for the surface patch in 1a5nC.

4 CHAIN 1A5NB

4.1 P18315 overview

From SwissProt, id P18315, 100% identical to 1a5nB:

**Description:** Urease beta subunit (EC 3.5.1.5) (Urea amidohydrolase).

**Organism, scientific name:** Klebsiella aerogenes.

**Taxonomy:** Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Klebsiella.

**Catalytic activity:** Urea + H2O = CO2 + 2 NH3.

**Subunit:** (Alpha, beta, gamma)3.

**Similarity:** Belongs to the urease beta subunit 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.
4.2 Multiple sequence alignment for 1a5nB

For the chain 1a5nB, the alignment 1a5nB.msf (attached) with 383 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 1a5nB.msf. Its statistics, from the alistat program are the following:

- Format: MSF
- Number of sequences: 383
- Total number of residues: 38071
- Smallest: 57
- Largest: 101
- Average length: 99.4
- Alignment length: 101
- Average identity: 56%
- Most related pair: 99%
- Most unrelated pair: 25%
- Most distant seq: 49%

Furthermore, <1% of residues show as conserved in this alignment.

The alignment consists of 3% eukaryotic (1% fungi, 2% plantae), 15% prokaryotic, and 1% 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 1a5nB.descr.

4.3 Residue ranking in 1a5nB

The 1a5nB sequence is shown in Fig. 27, with each residue colored according to its estimated importance. The full listing of residues in 1a5nB can be found in the file called 1a5nB.ranks, sorted in the attachment.

4.4 Top ranking residues in 1a5nB and their position on the structure

In the following we consider residues ranking among top 25% of residues in the protein. Figure 28 shows residues in 1a5nB 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.

4.4.1 Clustering of residues at 25% coverage. Fig. 29 shows the top 25% of all residues, this time colored according to clusters they belong to. The clusters in Fig.29 are composed of the residues listed in Table 41.
Table 41.

<table>
<thead>
<tr>
<th>cluster color</th>
<th>size</th>
<th>member residues</th>
</tr>
</thead>
<tbody>
<tr>
<td>red</td>
<td>23</td>
<td>28, 31, 32, 33, 35, 37, 38, 39, 41, 42, 44, 46, 49, 51, 53, 58, 66, 70, 71, 72, 73, 74, 92</td>
</tr>
</tbody>
</table>

Table 41. Clusters of top ranking residues in 1a5nB.

4.4.2 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 1a5nC. Table 42 lists the top 25% of residues at the interface with 1a5nC. The following table (Table 43) suggests possible disruptive replacements for these residues (see Section 5.6).

Table 42.

<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>38</td>
<td>S</td>
<td>S(99).</td>
<td>0.03</td>
<td>3/1</td>
<td>3.86</td>
</tr>
<tr>
<td>86</td>
<td>G</td>
<td>G(99)P.</td>
<td>0.04</td>
<td>35/35</td>
<td>3.21</td>
</tr>
<tr>
<td>92</td>
<td>G</td>
<td>G(98)A.</td>
<td>0.11</td>
<td>5/5</td>
<td>3.96</td>
</tr>
<tr>
<td>39</td>
<td>H</td>
<td>H(98)E.</td>
<td>0.12</td>
<td>63/11</td>
<td>3.02</td>
</tr>
<tr>
<td>37</td>
<td>G</td>
<td>G(97)S</td>
<td>0.14</td>
<td>7/7</td>
<td>3.77</td>
</tr>
<tr>
<td>66</td>
<td>G</td>
<td>G(95)SA</td>
<td>0.19</td>
<td>32/32</td>
<td>3.21</td>
</tr>
<tr>
<td>4</td>
<td>G</td>
<td>G(96)</td>
<td>0.21</td>
<td>58/58</td>
<td>2.71</td>
</tr>
</tbody>
</table>

Table 42. The top 25% of residues in 1a5nB at the interface with 1a5nC. (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 43.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>S</td>
<td>(KR) (FQMWH) (NLPI) (YE)</td>
</tr>
<tr>
<td>86</td>
<td>G</td>
<td>(R) (KE) (H) (FWD)</td>
</tr>
<tr>
<td>92</td>
<td>G</td>
<td>(E) (RD) (R) (FQMWH)</td>
</tr>
<tr>
<td>39</td>
<td>H</td>
<td>(T) (EVCAG) (SMD) (QLPI)</td>
</tr>
<tr>
<td>37</td>
<td>G</td>
<td>(E) (R) (R) (FWD)</td>
</tr>
<tr>
<td>66</td>
<td>G</td>
<td>(R) (R) (E) (H)</td>
</tr>
<tr>
<td>4</td>
<td>G</td>
<td>(R) (KE) (H) (FWD)</td>
</tr>
</tbody>
</table>

Table 43. List of disruptive mutations for the top 25% of residues in 1a5nB, that are at the interface with 1a5nC.

Figure 30 shows residues in 1a5nB colored by their importance, at the interface with 1a5nC.
Table 44. The top 25% of residues in 1a5nB at the interface with 1a5nC1. (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>
</tr>
</thead>
<tbody>
<tr>
<td>46</td>
<td>N</td>
<td>(Y) (FTWH) (SEVCARG) (MD)</td>
<td>100</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>R</td>
<td>(T) (Y) (CG) (D)</td>
<td>99</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Q</td>
<td>(Y) (T) (FCWHG) (SVA)</td>
<td>99</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>92</td>
<td>G</td>
<td>(E) (KD) (R) (FQMWH)</td>
<td>98</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>E</td>
<td>(FWH) (Y) (CG) (R)</td>
<td>98</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>G</td>
<td>(E) (K) (R) (FW)</td>
<td>97</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>R</td>
<td>(TD) (Y) (SECG) (VLAPI)</td>
<td>97</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>N</td>
<td>(Y) (FTW) (EH) (VCARG)</td>
<td>96</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>D</td>
<td>(R) (H) (FW) (K)</td>
<td>96</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>P</td>
<td>(R) (Y) (H) (K)</td>
<td>95</td>
<td>0.11</td>
<td></td>
</tr>
</tbody>
</table>

Figure 31 shows residues in 1a5nB colored by their importance, at the interface with 1a5nC1.

Table 45. List of disruptive mutations for the top 25% of residues in 1a5nB, that are at the interface with 1a5nC1.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>46</td>
<td>N</td>
<td>N(100)</td>
</tr>
<tr>
<td>70</td>
<td>R</td>
<td>R(99)LE</td>
</tr>
<tr>
<td>35</td>
<td>Q</td>
<td>Q(99)</td>
</tr>
<tr>
<td>92</td>
<td>G</td>
<td>G(98)CSDQ</td>
</tr>
<tr>
<td>41</td>
<td>E</td>
<td>E(98)DG(1)KP</td>
</tr>
<tr>
<td>37</td>
<td>R</td>
<td>R(98)V(1).K</td>
</tr>
<tr>
<td>28</td>
<td>N</td>
<td>N(97)</td>
</tr>
<tr>
<td>31</td>
<td>D</td>
<td>D(96)</td>
</tr>
<tr>
<td>33</td>
<td>P</td>
<td>P(95)</td>
</tr>
</tbody>
</table>

Table 46. Residues forming surface "patch" in 1a5nB.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>substitutions (%)</th>
<th>cvg</th>
</tr>
</thead>
<tbody>
<tr>
<td>46</td>
<td>N</td>
<td>N(100)</td>
<td>0.01</td>
</tr>
<tr>
<td>70</td>
<td>R</td>
<td>R(99)LE</td>
<td>0.02</td>
</tr>
<tr>
<td>35</td>
<td>Q</td>
<td>Q(99)</td>
<td>0.03</td>
</tr>
<tr>
<td>92</td>
<td>G</td>
<td>G(98)CSDQ</td>
<td>0.05</td>
</tr>
<tr>
<td>41</td>
<td>E</td>
<td>E(98)DG(1)KP</td>
<td>0.07</td>
</tr>
<tr>
<td>37</td>
<td>R</td>
<td>R(98)V(1).K</td>
<td>0.10</td>
</tr>
<tr>
<td>28</td>
<td>N</td>
<td>N(97)</td>
<td>0.11</td>
</tr>
<tr>
<td>31</td>
<td>D</td>
<td>D(96)</td>
<td>0.12</td>
</tr>
<tr>
<td>33</td>
<td>P</td>
<td>P(95)</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Table 47. Disruptive mutations for the top 25% of residues in 1a5nB, that are at the interface with 1a5nC1.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>46</td>
<td>N</td>
<td>(Y) (FTWH) (SEVCARG) (MD)</td>
</tr>
<tr>
<td>70</td>
<td>R</td>
<td>(T) (Y) (CG) (D)</td>
</tr>
</tbody>
</table>

4.4.3 Possible novel functional surfaces at 25% coverage. One group of residues is conserved on the 1a5nB surface, away from (or substantially larger than) other functional sites and interfaces recognizable in PDB entry 1a5n. It is shown in Fig. 32. The right panel shows (in blue) the rest of the larger cluster this surface belongs to.

Fig. 32. A possible active surface on the chain 1a5nB. The larger cluster it belongs to is shown in blue.

The residues belonging to this surface "patch" are listed in Table 46, while Table 47 suggests possible disruptive replacements for these residues (see Section 5.6).

Figure 31 shows residues in 1a5nB colored by their importance, at the interface with 1a5nC1.
5 NOTES ON USING TRACE RESULTS

5.1 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.

5.2 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 W, 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%.

5.3 Surface

To detect candidates for novel functional interfaces, first we look for residues that are solvent accessible (according to DSSP program) by at least 10Å², 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 5Å 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.)

5.4 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 5Å.

5.5 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).

5.6 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 [LPNQDEMIK], large [WFYHR], hydrophobic [LPVAMWF], polar [GTCY]; positively [KHR], or negatively [DE] charged, aromatic [WFYH], long aliphatic chain [EKDQM], OH-group possession [SDETY], and NH2 group possession [NQRK]. 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.

6 APPENDIX

6.1 File formats

Files with extension “ranks_sorted” 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
- variability has two subfields:
  1. number of different amino acids appearing in this column of the alignment
  2. their type

Table 47.

<table>
<thead>
<tr>
<th>res</th>
<th>type</th>
<th>disruptive mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>S</td>
<td>(KR) (FQMWH) NLPI (YE)</td>
</tr>
<tr>
<td>74</td>
<td>G</td>
<td>(R) (K) (FQWH) E</td>
</tr>
<tr>
<td>35</td>
<td>Q</td>
<td>(Y) (T) (FCWHG) (SVA)</td>
</tr>
<tr>
<td>41</td>
<td>H</td>
<td>(E) (T) (QD) (SCG)</td>
</tr>
<tr>
<td>73</td>
<td>P</td>
<td>(Y) (R) (TH) (E)</td>
</tr>
<tr>
<td>52</td>
<td>R</td>
<td>(T) (VCAG) (D) (LPI)</td>
</tr>
<tr>
<td>92</td>
<td>G</td>
<td>(E) (KD) (R) (FQMW)</td>
</tr>
<tr>
<td>39</td>
<td>H</td>
<td>(T) (EVCA) (SMI) (QLPI)</td>
</tr>
<tr>
<td>44</td>
<td>E</td>
<td>(FQWH) (Y) (CG) (R)</td>
</tr>
<tr>
<td>37</td>
<td>G</td>
<td>(E) (K) (R) (FW)</td>
</tr>
<tr>
<td>51</td>
<td>F</td>
<td>(TE) (K) (DR) (SCG)</td>
</tr>
<tr>
<td>72</td>
<td>E</td>
<td>(FQWH) (Y) (R) (VCAG)</td>
</tr>
<tr>
<td>32</td>
<td>R</td>
<td>(TD) (Y) (SECG) (VLAPI)</td>
</tr>
<tr>
<td>66</td>
<td>G</td>
<td>(R) (K) (E) (H)</td>
</tr>
<tr>
<td>31</td>
<td>D</td>
<td>(R) (H) (FW) (K)</td>
</tr>
<tr>
<td>33</td>
<td>P</td>
<td>(R) (Y) (H) (K)</td>
</tr>
</tbody>
</table>

Table 47. Disruptive mutations for the surface patch in 1a5nB.
6.3.2 CE To map ligand binding sites from different source structures, report_maker uses the CE program:


6.3.3 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.


http://swift.cmbi.kun.nl/swift/hssp/

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


http://www.drive5.com/muscle/

6.3.7 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.)

6.4 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.

6.5 Citing this work


6.6 About report_maker

report_maker was written in 2006 by Ivana Mihalek. The 1D ranking visualization program was written by Ivica Reš. report_maker is copyrighted by Lichtarge Lab, Baylor College of Medicine, Houston.

6.7 Attachments

The following files should accompany this report:

- 1a5nA.complex.pdb - coordinates of 1a5nA with all of its interacting partners
- 1a5nA.etvx - ET viewer input file for 1a5nA
- 1a5nA.cluster_report.summary - Cluster report summary for 1a5nA
- 1a5nA.ranks - Ranks file in sequence order for 1a5nA
- 1a5nA.clusters - Cluster descriptions for 1a5nA
- 1a5nA.msf - the multiple sequence alignment used for the chain 1a5nA
- 1a5nA.descr - description of sequences used in 1a5nA msf
- 1a5nA.ranks_sorted - full listing of residues and their ranking for 1a5nA
- 1a5nA.1a5nC.if.pml - Pymol script for Figure 4
- 1a5nA.cbcvg - used by other 1a5nA-related pymol scripts
- 1a5nA.1a5nC1.if.pml - Pymol script for Figure 5
- 1a5nA.1a5nB.if.pml - Pymol script for Figure 6
- 1a5nA.1a5nA2.if.pml - Pymol script for Figure 7
- 1a5nA.1a5nA1.if.pml - Pymol script for Figure 8
- 1a5nB.complex.pdb - coordinates of 1a5nB with all of its interacting partners
- 1a5nB.etvx - ET viewer input file for 1a5nB
- 1a5nB.cluster_report.summary - Cluster report summary for 1a5nB
- 1a5nB.ranks - Ranks file in sequence order for 1a5nB
- 1a5nB.clusters - Cluster descriptions for 1a5nB
- 1a5nB.msf - the multiple sequence alignment used for the chain 1a5nB
- 1a5nB.descr - description of sequences used in 1a5nB msf
- 1a5nB.ranks_sorted - full listing of residues and their ranking for 1a5nB
- 1a5nB.1a5nC.if.pml - Pymol script for Figure 30
- 1a5nB.cbcvg - used by other 1a5nB-related pymol scripts
- 1a5nB.1a5nC1.if.pml - Pymol script for Figure 31