Announcements

- Videos: SciVee
  www.rostlab.org

THANKS:
  Tim Karl + Manfred Roos

- NO lectures:
  Tue May 31(!)
  studentische vollversammlung
  Thu Jun 2
  Thu Jun 16

- LAST lecture: Jul 7

- Examen: Jul 12 (?), 10:30 (likely this room)
  • Makeup: likely: October 13 - morning

CONTACT: Marlena Drabik assistant@rostlab.org

Let it go. Let it out.
Let it all unravel.
Let it free and it can be
A path on which to travel.
Protein Prediction
- Part 1: Structure
Structure: from 3D→1D (sec str)
TOC today

- LAST WEEKs
  - amino acids - protein structure - bonds & energies
  - structure comparisons
- THIS WEEK
  - 3D->1D (secondary structure) & prediction methods
- NEXT WEEK
  - secondary structure prediction contd.
Notation: protein structure 1D, 2D, 3D

<table>
<thead>
<tr>
<th>P</th>
<th>PP</th>
<th>P</th>
<th>128</th>
<th>110</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q</td>
<td>QQ</td>
<td>Q</td>
<td>175</td>
<td>97</td>
</tr>
<tr>
<td>I</td>
<td>FF</td>
<td>V</td>
<td>70</td>
<td>60</td>
</tr>
<tr>
<td>T</td>
<td>SS</td>
<td>IV</td>
<td>77</td>
<td>69</td>
</tr>
<tr>
<td>L</td>
<td>LL</td>
<td>STL</td>
<td>120</td>
<td>14</td>
</tr>
<tr>
<td>W</td>
<td>WW</td>
<td>QED</td>
<td>238</td>
<td>81</td>
</tr>
<tr>
<td>Q</td>
<td>RK</td>
<td>QAK</td>
<td>169</td>
<td>97</td>
</tr>
<tr>
<td>R</td>
<td>RRR</td>
<td>PFP</td>
<td>200</td>
<td>62</td>
</tr>
<tr>
<td>P</td>
<td>PPPPPP</td>
<td></td>
<td>24</td>
<td>48</td>
</tr>
<tr>
<td>L</td>
<td>VVTKF</td>
<td>E</td>
<td>71</td>
<td>59</td>
</tr>
<tr>
<td>V</td>
<td>VVLIE</td>
<td></td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>T</td>
<td>TTKEK</td>
<td>E</td>
<td>74</td>
<td>69</td>
</tr>
<tr>
<td>I</td>
<td>AALIV</td>
<td>E</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>K</td>
<td>HYKKF</td>
<td>E</td>
<td>90</td>
<td>73</td>
</tr>
<tr>
<td>I</td>
<td>II</td>
<td>LVI</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>G</td>
<td>EENGG</td>
<td></td>
<td>46</td>
<td>41</td>
</tr>
<tr>
<td>G</td>
<td>GGGT</td>
<td>G</td>
<td>62</td>
<td>53</td>
</tr>
<tr>
<td>Q</td>
<td>QQKRR</td>
<td></td>
<td>68</td>
<td>71</td>
</tr>
<tr>
<td>L</td>
<td>PPLLW</td>
<td>E</td>
<td>118</td>
<td>59</td>
</tr>
<tr>
<td>K</td>
<td>VVFKV</td>
<td>E</td>
<td>31</td>
<td>73</td>
</tr>
<tr>
<td>E</td>
<td>EESKK</td>
<td>E</td>
<td>124</td>
<td>95</td>
</tr>
<tr>
<td>A</td>
<td>VVGLG</td>
<td>E</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>L</td>
<td>LLI</td>
<td>LL</td>
<td>29</td>
<td>0</td>
</tr>
<tr>
<td>L</td>
<td>LLLVV</td>
<td>E</td>
<td>24</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>D</td>
<td>DDDD</td>
<td>49</td>
<td>58</td>
</tr>
<tr>
<td>T</td>
<td>TTTT</td>
<td>T</td>
<td>72</td>
<td>51</td>
</tr>
<tr>
<td>G</td>
<td>GGGG</td>
<td>G</td>
<td>62</td>
<td>30</td>
</tr>
<tr>
<td>A</td>
<td>AAAA</td>
<td>A</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>DDDD</td>
<td>D</td>
<td>102</td>
<td>79</td>
</tr>
<tr>
<td>D</td>
<td>DDAKE</td>
<td>T</td>
<td>69</td>
<td>58</td>
</tr>
<tr>
<td>T</td>
<td>SSSTTV</td>
<td></td>
<td>1</td>
<td>69</td>
</tr>
<tr>
<td>V</td>
<td>IIVIV</td>
<td>E</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>L</td>
<td>VVIVL</td>
<td>E</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

1D 2D 3D

0    -1    -2    -3    -4    -5 kcal/mol

© Burkhard Rost (TU Munich)
3D $\rightarrow$ 1D Secondary structure assignment
3D Structure
3D Structure: cartoon
Secondary structure stabilized by hydrogen bonds
Linus Carl Pauling

- Feb 28, 1901 - Aug 19, 1994
- 2 Nobel prizes (1 of 4)
- Oregon State -> Caltech (PhD X-ray crystallography; 1926: Munich)
- 1st Nobel prize (chemistry) for “Nature of chemical bond”
- discovered sickle cell anemia as “molecular disease”
- 2nd Nobel prize (peace) “Stop Nukes!”
- Vitamin C/megavitamin therapy
Linus Pauling: introduce concept

- L Pauling, RB Corey and HR Branson (1951) The Structure of Proteins: Two Hydrogen-bonded Helical Configurations of the Polypeptide Chain PNAS 37: 205-34
- L Pauling, RB Corey and HR Branson (1951) Two Hydrogen-Bonded Helical Configurations of the Polypeptide Chain PNAS 37: 205-11
- L Pauling and RB Corey (1953) Two Rippled-sheet Configurations of Polypeptide Chains, and a Note About the Pleated Sheets PNAS 39: 253-6
DSSP
Dictionary of Secondary Structure of Proteins
Chris Sander

- Sloan Kettering Cancer Center, NYC

- papers:
  - >770 papers (May 2011)
  - 1 >6,000 citations (May 2011)
  - 6 >1,000 citations (May 2011)
  - 87 over 100
  - H-index 92 (ISI May 2011)

- ISCB Fellow
Secondary Structure Assignment: DSSP

Dictionary of protein Secondary Structure for Proteins


Chris Sander

Wolfgang Kabsch
DSSP: Coulomb

\[ E = f\delta^+\delta^- \left( \frac{1}{r_{NO}} + \frac{1}{r_{HC'}} + \frac{1}{r_{HO}} + \frac{1}{r_{NC'}} \right) \]  

(17.2)

**Figure 17.1.** Distances used to calculate the Coulomb hydrogen bond.

Linus Pauling’s H-bond pattern used in DSSP

DSSP

\[ E_{\text{HB}} < -0.5 \text{ kcal/mol} \]

L Pauling & RB Corey (1953) PNAS 39:247-252
L Pauling, RB Corey & HR Branson (1951) PNAS 37:205-234
Secondary structure assignment

Different evaluation criteria applied:

- **Assignment coverage**: **DEFINE**
  Geometry (fitting ideal sec str segments)

- **Expert assignment**: **STRIDE**

- **Predictability**: **NNass**

- **Enthalpic energy**: **DSSP**
Secondary Structure Assignment: STRIDE

Dimitrij Frishman & Patrick Argos
STRuctural IDEntification method (STRIDE)

Dimitrij Frishman

23:566-79
Secondary Structure Assignment: **STRIDE**

- Dimitrij Frishman & Patrick Argos
  STRuctural IDEntification method (STRIDE)
- H-bond + Ramachandran map (torsion angles)

*Figure 17.2. Angles and distances defining the empirical hydrogen bond. Note: figure similar to the one in Frishman and Argos (1995).*

(Fig. 17.2 after Frishman & Argos 1995)
Mostly methods agree

CAF Andersen & B Rost (2003)
Methods Biochem Anal 44:341-63 Fig. 17.6

CAF Andersen & B Rost (2003) Methods Biochem Anal 44:341-63 Fig. 17.7

Figure 17.7. Comparison of three assignment schemes. The occurrences of three assignment classes (α-helix, β-sheet, and nonregular) by three assignment methods: DSSP, P-Curve, and DEFINE give the 10 displayed categories, if the order is not regarded. When all schemes assign α-helix, this is indicated by (α, α, α), when two assign α-helix and one nonregular, this is indicated by (α, α, c). The distinction between helix and sheet appears clear, since (α, α, β) and (α, β, β) assignments are rare (<0.01%). Data are taken from (Colloc’h et al., 1993).
Issues with secondary structure assignments
Protein motion and function

Calmodulin

Structural superposition of models

Order parameter measuring motion

Measuring motion

Assignment for different NMR models


© Claus AF Andersen
© Burkhard Rost (TU Munich)
Continuous assignment
\[ E = f \delta^+ \delta^- \left( \frac{1}{r_{NO}} + \frac{1}{r_{HC'}} + \frac{1}{r_{HO}} + \frac{1}{r_{NC'}} \right) \]  

(17.2)

Figure 17.1. Distances used to calculate the Coulomb hydrogen bond.

Continuous DSSP

\[ DSSP_{cont} = \sum_{HB \text{ thresholds}} w_{HB} DSSP_{HB} \]
Protein motion influences assignment

Use motion to assign secondary structure

Optimizing the weights

\[
DSSP_{\text{cont}} = \sum_{\text{HB thresholds}} w_{HB} DSSP_{HB}
\]

Global minimum well defined

Weight selection is robust

Agreement between different NMR models
Agreement between continuous assignment

Transitions between similar geometry

(a) X-ray homologues

(b) different NMR models

Continuous sec str correlates with motion


© Claus AF Andersen
© Burkhard Rost (TU Munich)
From 3D to ...
Central dogma of molecular biology

DNA → RNA → Protein

information, code, library, manual

intermediate step

machinery of life
Central dogma of molecular biology

DNA → RNA → Protein

- DNA: information, code, library, manual
- RNA: intermediate step
- Protein: machinery of life
FANTOM 3: Functional annotation of mouse

Science 2005
309:1559-63
The Transcriptional Landscape of the Mammalian Genome

Numbers:
- Total transcripts: 181,047
- New protein-coding transcripts: 16,247
- New proteins: 5,154
- Multiple splice variants: 65%
- 1.35 5’ start sites for each 3’ end
- 1.83 3’ ends for each 5’ end
What from 3D?

Predict function from 3D when function is unknown
Annotation transfer

similar structure $\implies$ similar function

similar function $\implies$ similar structure
Different environment

Eye duck lens /ASL:

- One protein
- One structure
- different tissue -> different function
Structural basis for moonlighting

Todd et al. Structure (2002); 10:1435-1451
Enzyme vs. Non-enzymes

- Number of non-enzymes with an enzyme homolog sharing xx SI
- Cumulative percentage of non-enzymes considered

Todd et al. Structure (2002); 10:1435-1451

Tuesday May 17, 2011
Measuring conservation of enzymatic activity

Enzymes of known function

unique

All proteins
Family size

B Rost 2002 J Mol Biol 318, 595-608
REAL conservation of EC number

Accuracy 1st: bias
Accuracy 1st: unbiased
Accuracy all: bias
Accuracy all: unbiased
Coverage 1st: bias
Coverage 1st: unbiased
Coverage all: bias
Coverage all: unbiased

B Rost 2002 *J Mol Biol* 318, 595-608
REAL conservation of EC number

B Rost 2002 J Mol Biol 318, 595-608
REAL conservation of EC number

bias: 50% found at >90% right

real: 50% found at <15% right!

B Rost 2002 J Mol Biol 318, 595-608
Converged to similar function

Chymotrypsin-Subtilisin:

- Catalytic triad (H-S-D): proteases (cleave peptide bonds)
- Folds: Trypsin-like serine proteases / subtilisin-like
- different fold, different sequence: same catalytic triad/function
Enolase N-terminal domain-like SCOP family

- similar fold
- 25% of seq. id.
- different catalytic residues

Similar function?
Family: Enolase N-terminal domain-like

C-terminal domain is beta/alpha-barrel

Lineage:

1. Root: scop
2. Class: Alpha and beta proteins (a+b) [53931]  
   Mainly antiparallel beta sheets (segregated alpha and beta regions)
3. Fold: Enolase N-terminal domain-like [54825]  
   beta(3)-alpha(3); meander and up-and-down bundle
4. Superfamily: Enolase N-terminal domain-like [54826]
5. Family: Enolase N-terminal domain-like [54827]  
   C-terminal domain is beta/alpha-barrel

Protein Domains:

1. Enolase [54828]
   1. Baker's yeast (Saccharomyces cerevisiae) [54829] (14)  
   2. Lobster (Homarus vulgaris) [54830] (2)
   3. Trypanosoma brucei brucei [89925] (1)
   4. Enterococcus hirae [89926] (1)
   5. Escherichia coli [69710] (1)
   6. Human (Homo sapiens), gamma isoform [110936] (1)
2. D-glucarate dehydratase [54831]
   1. Pseudomonas putida [54832] (1)
   2. Escherichia coli [54833] (6)
3. O-succinylbenzoate synthase [54834]
   1. Escherichia coli [54835] (3)
4. Muconate-lactonizing enzyme (cis muconate cycloisomerase) [54836]
   1. Pseudomonas putida [54837] (5)
5. Mandelate racemase [54838]
   1. Pseudomonas putida [54839] (6)
6. Chlormuconate cycloisomerase [54840]

EC 4.2.1.40  E.C. 5.1.2.2  E.C. 5.5.1.7
Annotation transfer

similar structure  \rightarrow  similar function

similar function  \rightarrow  similar structure
Surface Properties

- Surface amino acid conservation
- Cavities
- Electrostatic potential
Residue conservation on surface

ConSurf

Cavity: but NOT conserved
Cavities
Conserved cavities: different substrates
Difference in electrostatic surface potential

- proteins of similar structure

1ns5

Yibk-methyltransferase

1nxz
Surface comparisons

- Build atlas of all surfaces
  - eF-Site
  - GRASS
    M Nayal et al. (1999) Prot Sci, 8: 676-679

- Interactive comparison of surfaces:
  - pvSOAR
Structural patterns/motifs

- PROCAT

Natively unstructured / disordered
Natively Unstructured (Disordered) Proteins

Natively Unstructured (Disordered) Proteins

Natively Unstructured (Disordered) Proteins

- Missing residues from electron density map
- Non Regular Secondary structure (NORS) regions
- Regions with High B-values
- Proteins with long unstructured regions (characterized by NMR, CD)

**Missing residues from electron density map**

- DISOPRED (Jones)
- PDISORDR (softberry)
- RONN (Esnuof)
- DISpro (Baldi)
- DisEMBL (Russell)

**Non Regular Secondary structure (NORS) regions**

**Regions with High B-values**

- PROFbval, (Rost)
- DisEMBL (Russell)
- GlobPlot (Gibson)

**Proteins with long unstructured regions (characterized by NMR, CD)**

- FoldIndex (Sussman)
- Pred (Simon)
- FoldUnfold (alzitskaya)
- CON (Rost)
Example for what to learn from sequence/structure
Example 1: Target sequence

The MT0146/CbiT sequence

1 MIPDDEFIKNPSVPGPTAMEVRCLIMCLAEPGKNDVAVDVGCCTGGVTELAGRVRVYA
IDRNPEAISTTEMNLQRHGLGDNVTLEMGAPEALCKIPDIDIAVVGSGGELQEILRII
KDKLKPGGRIIVTAILLETKFEAMECLRDLGFVNIETLNIARGRALDRGTMMVSRNPVA
LIYTGVSHEKND 192

Protein structure from structural genomics

Biosynthesis of cobalamin (vitamin B$_{12}$)

AL A $\xrightarrow{\text{HemB}}$ PB G $\xrightarrow{\text{HemC}}$ HMB $\xrightarrow{\text{HemD}}$
Biosynthesis of cobalamin (vitamin B\textsubscript{12})

Known facts about function

Experiments show that CbiE + CbiT do “Methylation + Carboxylation” to Precorrin-6y; unclear which one is which.
Known facts and hypothesis about function

strong sequence similarity with other B12 methyltransferase

Precorrin-6y
Carboxylation
Methyllations

CbiE
CbiT
Known facts and hypothesis about function

Lower sequence similarity with carboxylases

Methylations

Precorrin-6y

Carboxylation

strong sequence similarity with other B12 methyltransferase

© Burkhard Rost (TU Munich)
Known facts and hypothesis about function

- Lower sequence similarity with carboxylases
- Strong sequence similarity with other B12 methyltransferase
- Methyllations
- Precorrin-6y
- Carboxylation

© Burkhard Rost (TU Munich)
ChiT structure

Keller JP, Smith PM, Benach J, Christendat D, deTitta GT, and Hunt JF. *Structure* 2002, 10:1475-87
# 3D comparison

**DALI output**

SEQLENGTH 178  
NALIGN 374  
**WARNING** pairs with Z<2.0 are structurally dissimilar

## SUMMARY: PDB/chain identifiers and structural alignment statistics

<table>
<thead>
<tr>
<th>NR.</th>
<th>STRID1</th>
<th>STRID2</th>
<th>Z</th>
<th>RMSD</th>
<th>LALI</th>
<th>LSEQ2</th>
<th>%IDE</th>
<th>REVERS</th>
<th>PERMUT</th>
<th>NFRAG</th>
<th>TOPO</th>
<th>PROTEIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7768-A</td>
<td>1i9g-A</td>
<td>20.7</td>
<td>2.7</td>
<td>170</td>
<td>264</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>S</td>
<td>Adomet-dependent methyltransfer</td>
</tr>
<tr>
<td>2</td>
<td>7768-A</td>
<td>1dus-A</td>
<td>17.7</td>
<td>2.6</td>
<td>157</td>
<td>194</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>S</td>
<td>mj0882 - methyltransferase</td>
</tr>
<tr>
<td>3</td>
<td>7768-A</td>
<td>1vid</td>
<td>15.7</td>
<td>2.5</td>
<td>159</td>
<td>214</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>S</td>
<td>catechol o-methyltransferase</td>
</tr>
<tr>
<td>4</td>
<td>7768-A</td>
<td>1m6y-A</td>
<td>15.7</td>
<td>2.4</td>
<td>151</td>
<td>289</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>S</td>
<td>s-adenosyl-methyltransferase</td>
</tr>
<tr>
<td>5</td>
<td>7768-A</td>
<td>1im8-A</td>
<td>14.5</td>
<td>2.6</td>
<td>152</td>
<td>225</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>S</td>
<td>yeco (methyltransferase)</td>
</tr>
<tr>
<td>6</td>
<td>7768-A</td>
<td>1fbs-A</td>
<td>14.2</td>
<td>2.9</td>
<td>148</td>
<td>230</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>S</td>
<td>mj fibrillarin homologue</td>
</tr>
<tr>
<td>7</td>
<td>7768-A</td>
<td>1nv8-A</td>
<td>14.0</td>
<td>2.9</td>
<td>151</td>
<td>271</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>S</td>
<td>hemk protein</td>
</tr>
<tr>
<td>8</td>
<td>7768-A</td>
<td>1kh-A</td>
<td>13.9</td>
<td>3.4</td>
<td>151</td>
<td>193</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>S</td>
<td>guanidinoacetate methyltrans</td>
</tr>
<tr>
<td>9</td>
<td>7768-A</td>
<td>1ixk-A</td>
<td>13.5</td>
<td>2.6</td>
<td>148</td>
<td>298</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>S</td>
<td>methyltransferase</td>
</tr>
<tr>
<td>10</td>
<td>7768-A</td>
<td>1jq3-A</td>
<td>13.3</td>
<td>2.9</td>
<td>153</td>
<td>295</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>S</td>
<td>spermidine synthase (putrescine</td>
</tr>
<tr>
<td>11</td>
<td>7768-A</td>
<td>1kr5-A</td>
<td>13.2</td>
<td>2.5</td>
<td>139</td>
<td>218</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>S</td>
<td>l-isoaspartateo-methyltransf</td>
</tr>
<tr>
<td>12</td>
<td>7768-A</td>
<td>1kp9-B</td>
<td>13.1</td>
<td>3.3</td>
<td>156</td>
<td>270</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>S</td>
<td>cyclopropane-fatty-acyl-phosphol</td>
</tr>
<tr>
<td>13</td>
<td>7768-A</td>
<td>1ej0-A</td>
<td>13.1</td>
<td>3.0</td>
<td>144</td>
<td>180</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>S</td>
<td>ftsj (ftsj methyltransferase)</td>
</tr>
<tr>
<td>14</td>
<td>7768-A</td>
<td>2erc-A</td>
<td>12.9</td>
<td>3.2</td>
<td>145</td>
<td>235</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>S</td>
<td>rrna methyl transferase fragment</td>
</tr>
<tr>
<td>15</td>
<td>7768-A</td>
<td>1ri1-A</td>
<td>12.9</td>
<td>3.0</td>
<td>151</td>
<td>252</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>S</td>
<td>mrna capping enzyme - methyltrans</td>
</tr>
</tbody>
</table>

---

**DALI: http://www.ebi.ac.uk/dali/**

© Burkhard Rost (TU Munich) 74 /128

Tuesday May 17, 2011
Visual comparison of target and template structure

Visual comparison with other precorrin methyltransferases

Visual comparison with other precorrin methyltransferases

precorrin methyltransferases

AdoMet/SAM

CysG

CbiF

CbiT

Visual comparison with other precorrin methyltransferases

CbiF

DNA/RNA methyltransferases

AdoMet/SAM

CysG

CbiT

## Analysis of SAM binding sites

<table>
<thead>
<tr>
<th>Motif</th>
<th>M.HhaI</th>
<th>M.TaqI</th>
<th>COMT</th>
<th>role</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>AGxGG</td>
<td>PSxAxGP</td>
<td>GAxxG</td>
<td>H-bonds</td>
</tr>
<tr>
<td>II</td>
<td>E40</td>
<td>E71</td>
<td>E90</td>
<td>H-bond with Ribose hydroxyls</td>
</tr>
<tr>
<td></td>
<td>W41</td>
<td>I72</td>
<td>M91</td>
<td>VdW with adenine</td>
</tr>
<tr>
<td>III</td>
<td>D60</td>
<td>D89</td>
<td>S119</td>
<td>H-bonds to N atoms</td>
</tr>
<tr>
<td>V</td>
<td>L100</td>
<td>L142</td>
<td>W143</td>
<td>VdW contacts with adenine</td>
</tr>
<tr>
<td></td>
<td>F18</td>
<td>F146</td>
<td>H142</td>
<td></td>
</tr>
</tbody>
</table>

# Analysis of SAM binding sites

<table>
<thead>
<tr>
<th>Motif</th>
<th>M.HhaI</th>
<th>M.TaqI</th>
<th>COMT</th>
<th>role</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>AGxGG</td>
<td>PSxAxGP</td>
<td>GAxxG</td>
<td>H-bonds</td>
</tr>
<tr>
<td>II</td>
<td>E40</td>
<td>E71</td>
<td>E90</td>
<td>H-bond with Ribose hydroxyls</td>
</tr>
<tr>
<td></td>
<td>W41</td>
<td>I72</td>
<td>M91</td>
<td>VdW with adenine</td>
</tr>
<tr>
<td>III</td>
<td>D60</td>
<td>D89</td>
<td>S119</td>
<td>H-bonds to N atoms</td>
</tr>
<tr>
<td>V</td>
<td>L100</td>
<td>L142</td>
<td>W143</td>
<td>VdW contacts with adenine</td>
</tr>
<tr>
<td></td>
<td>F18</td>
<td>F146</td>
<td>H142</td>
<td></td>
</tr>
</tbody>
</table>
Analysis of SAM binding sites

CONSURF: http://consurf.tau.ac.il/
CbiT-SAM crystal structure

Analysis of SAM binding sites

Analysis of precorrin binding site

GRASP: http://trantor.bioc.columbia.edu/grasp/

Fig. from JP Keller et al. & JF Hunt (2002) Structure, 10:1475-87
Known facts and hypothesis about function

structural similarity with methyltransferases

CbiE  CbiT

strong sequence similarity with other B12 methyltransferase

Precorrin-6y

Methylation

Carboxylation
Conclusion: hypotheses

- Proposed precorrin carboxylation:
  - Spontaneous after double methylation
  - CbiT protein is also a carboxylase
Announcements

Videos: SciVee
www.rostlab.org

THANKS:
Tim Karl + Manfred Roos

NO lectures:
Tue May 31(!)
studentische vollversammlung
Thu Jun 2
Thu Jun 16

LAST lecture: Jul 7
Examen: Jul 12 (?), 10:30 (likely this room)
• Makeup: likely: October 13 - morning

CONTACT: Marlena Drabik assistant@rostlab.org

Tuesday May 17, 2011