[4] M. Meuwly, *Chem. Phys. Chem.*, 2006, *7*, 2061.
[5] C. Rovira, B. Schulze et al., *Biophys. J.* 2001, *81*, 435.

MS01 O4

Origin and Evolution of One of the Most Ancient Rossmann Folds. W. L. Duax[a], R. Huether[a], Q. Mao[a], V. Pletnev[a], T. Umland[a] and C. M. Weeks[a], [a]*Hauptman-Woodward Medical Research Institute, Buffalo, NY 14203.* E-mail: duax@hwi.buffalo.edu

Keywords: Evolution, β-k-ACPR, Rossmann Fold

The Rossmann fold is one of the most ancient and commonly encountered protein folds. Members of the short-chain oxidoreductase (SCOR) enzyme family, which contains a Rossmann fold, are present in the genomes of all species sequenced to date. The SCOR family is part of a much larger superfamily of Rossmann folds that use NAD(H) or NADP(H) as cofactors. Crystallographic and biochemical studies have revealed that, in the vast majority of Rossmann folds, NAD or NADP cofactor preference is dictated by the presence of an Asp residue in a specific sequence position in the $\beta_2\alpha_3$ turn or an Arg in the adjacent sequence position, respectively. NAD binding is usually associated with enzyme preference for oxidation, and NADP binding is associated with a reductase preference.

The β-ketoacyl [acyl carrier protein] reductase (β-k-ACPR) enzymes, a 690-member subset of the SCOR family, are essential to fatty acid synthesis in bacteria and plants. By focusing analysis on the sequence and structure of the 690 β-k-ACPRs in the gene bank, we have discovered that (1) the most primitive member of the family was an NADP reductase, (2) that NADP binding was originally contingent upon a Ser or Thr residue in the $\beta_2\alpha_3$ turn (not an Arg), (3) that a specific dimer assembly is stabilized by the stacking of aromatic groups at specific sites on the α_5 and α_6 helices and (4) that a previously undetected GGMYM sequence at the C-terminus, conserved in all species of γ-proteobacteria and most species of β-proteobacteria, stabilizes the functionally required tetramer by multiple hydrogen bonding and aromatic ring stacking crosslinking the four monomers together. Our analysis indicates that the primordial members of the β-k-ACPR family probably arose in the GC-rich γ-proteobacteria and that they are distinguished by the presence of multiple open reading frames (MORFs), an extreme codon bias in their DNA and an amino acid bias in their protein composition. The β-k-ACPRs in α- and β-proteobacteria resemble the γ-proteobacteria in having a high degree of conservation of the 40 residues characteristic of the SCOR folds, the 9 residues that are specific to β-k-ACPRs and the conserved residues that are critical to dimer and tetramer assemblies. All β-proteobacteria for which genomes have been reported are GC rich and have MORF's and GC codon bias. However, α-proteobacteria are AT-rich, do not have MORFs, do not exhibit a significant codon bias and their amino acid composition is more divergent. Further analysis should make it possible to determine if β-k-ACPR genes in α-, β-, and γ-proteobacteria evolved from a common ancestor, if γ-proteobacteria are the ancestor of the others, or if the β-k-ACPR gene was horizontally transferred from GC-rich proteobacteria to AT-rich ones.

We can now characterize fully and unambiguously the β-k-ACPRs in the gene bank according to cofactor preference and mechanism of recognition, catalytic residues, residues that directly or indirectly form the proton wire via bound waters, residues controlling β-face hydride transfer, and substrate-defining residues. We can also describe, in detail, the stereochemistry of the tetrameric form required for activity and the residues that control tetramer formation as well as any conserved residues on the surface of the tetramer that may be related to specific protein/protein interactions of the β-k-ACPR. We can identify conserved sequence differences between β-k-ACPRs in major classes of bacteria that may make it possible to design inhibitors that can selectively block specific families of bacteria without interfering with mammalian biochemistry. We will use similar techniques to characterize all 10,000 SCOR genes in the databank and determine which residues specifically bind each of over 150 probable substrates.

This research was supported by NIH grant No. DK26546.

MS01 O5

Importance of local model quality in Molecular Replacement method Marcin Pawlowski, Janusz M. Bujnicki, *Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology in Warsaw, Poland.* Email:marcinp@genesilico.pl

Keywords: molecular replacement, homology modeling, local model quality

Computational models of protein structure have been shown to be useful as search models in Molecular Replacement (MR), a common method to experimentally determine protein structures by X-ray crystallography. It was shown that the success of MR depends on the high accuracy of the models, a parameter that remains unknown as long as the final structure is not available. Interestingly, during the last two years several methods (termed MQAPs) were developed to predict the local accuracy of theoretical models. We were interested in analyzing whether the application of MQAP can improve the utility of theoretical models in MR.

We focused our analyses on four known protein structures with resolution better than 2 Å, for which we generated theoretical models based on 'ideal' sequence alignments to homologous proteins obtained from the SCOP database. After superposition of models onto the corresponding native target structures we measured the local deviation between each atom in the model and its counterpart in the native structure. We also used MQAPs to 'predict' this value for models, using methods that do not take into account any information about the true difference between the model and the real structure. The known or predicted deviation was then used as the B-factor in MR calculations with experimental structure factor of the target protein. We found that the known or accurately predicted deviation of individual atoms in the search model can have significant impact on success of MR. In particular, we demonstrate that a MQAP-evaluated theoretical model that is relatively diverged from the real structure (GDT-TS score 69.1) can be used to obtained a correct MR solution, while models with unknown /non-estimated local quality typically have to exhibit close similarity (GDT-TS score > 80) to generate a comparable solution. These results indicate that theoretical modeling in combination with accurate prediction of quality of models can provide useful search models for crystallographic structure solution by MR.