## MS08.P16

*Acta Cryst.* (2011) A**67,** C268

**Perfect alignment of ribosomal protein S3 in creating an evolutionary tree**

Isabel Xu,[a,b] Mark Chee,[a,b] William L. Duax,[a,c] Robert Huether,[c] and David Dziak [a]*Hauptman-Woodward Medical Research Institute, Buffalo NY. (USA).* [b]*Williamsville Central Schools, Williamsville NY. (USA).* [c]*Department of Structural Biology, University at Buffalo. Buffalo NY. (USA).* E-mail: xu@hwi.buffalo.edu

Early attempts to use genomic analysis to define a tree of life used sequence alignment of ribosomal RNA. Subsequent analysis focused on proteins thought to be common to all species, revealed significant divergences from the rRNA based trees and the discrepancies have been attributed to horizontal gene transfer. Because the 52 bacterial ribosomal proteins are unlikely to have been affected by horizontal gene transfer, we examined the S3 ribosomal protein in search of a reliable method to use sequence alignment and divergence to trace species evolution. The amino acid length of bacterial ribosomal protein S3 has changed little throughout evolotion. The search vector MGX XX(20)X(0,20)[DST]X(3)[RK]X(18)XX(0,15)X(3)X[PA]XX$^{G}$X(0,100)[GSTA][KR]X(6)GX[LIVMT]X(2)[NQSCH]X(1,3)[LIVFCA]X(3)[LIV]XX(7)[LMT]X(2)GX(2)[GS]X(0,100) retrieves and aligns over 1800 S3 ribosomal proteins from SwissPro/TrEMBL. Fully conserved positions in the fingerprint and those where occupancy is limited to two amino acids is dominated by Gly, Ala, Arg, and Pro residues (GARP) and the sites in which GARP is 100% conserved act as markers for perfect sequence alignment. Because the output contains more highly conserved residues than the input, the critical importance of conserved glycines is further established. Accuracy of alignment is tested by separating Gram-positive (G+) from Gram-negative (G-) bacteria. Certain amino acid positions within the fingerprint are capable of achieving such a separation, providing insight into the evolutionary history of the ribosomal protein S3. Accurate alignment also reveals homology between the sequences of the S3 ribosomal proteins of cyanobacteria and chloroplasts. The accuracy of a perfect sequence alignment can also be assessed by tracing its divergence through phylum , class, order, family, genus, species and strain. The percent of residues in the S3 protein that are fully conserved rises from 5% in all bacterial phylum to 98% in Escherichia and the number of conserved residues of GARP increases from 5 to 74. The amino acid sequence and three dimensional structure of the S3 proteins of all species and strains of a single genus of bacteria have undergone no significant change over billions of years of DNA replication. Mapping the most highly conserved GARP residues onto the three-dimensional structure of S3 in the ribosome provides insight into the reason for their conservation and the role of S3 in ribosomal function. Support in part by: Mr Roy Carver, Stafford Graduate Fellowship, Caerus Forum Fund, The East Hill Foundation and the generous help of a number of High School students from the Buffalo NY area.

**Keywords: evolution, ribosomal protein, alignment.**

## MS09.P01

*Acta Cryst.* (2011) A**67,** C268

**Underlying nets in three-periodic coordination polymers**

Eugeny V. Alexandrov,[a] Vladislav A. Blatov,[a] Anton V. Kochetkov[a] and Davide M. Proserpio[b] [a]*Inorganic Chemistry Department, Samara State University, Samara (Russia).* [b]*Dipartimento di Chimica Strutturale e Stereochimica Inorganica (DCSSI), Università degli Studi, Milano (Italy).* E-mail: aleksandrov_ev1@mail.ru

We discuss a recently developed approach to formalize the analysis of extended architectures by successive simplifications of a crystal structure perceived as a periodic net. The approach has been implemented into the program package TOPOS that allows one to simplify and classify coordination polymers of any complexity in an automated mode [1]. Using TOPOS, we retrieved 6620 3-periodic coordination polymers from the Cambridge Structural Database and represented them in a standard way as underlying nets. The topological classification of both 975 interpenetrating and 5645 single 3-periodic underlying nets have been performed and compared.

The topological properties of coordination polymers can be formalized with the concept of underlying net. The nodes and edges of the underlying net correspond to structural groups and links between them. To obtain an underlying net for valence-bonded coordination polymers we use two types of their representation. The *standard* representation considers metal atoms and organic ligands as structure units, ignoring extra-framework ions and molecules. In the *cluster* representation the structural units include polynuclear coordination groups. We define *topological type*, the basic taxon to be used in the systematics of coordination polymers, as a set of crystal structures with the same underlying net.

In general, the distributions of interpenetrating 3D motifs on topological types and main interpenetration parameters are similar to the data of 2004 [2]. In particular, the most abundant topologies (**dia**, **pcu**, **srs**, **ths**) remain the same. The typical interpenetrating motifs are 3-, 4-, or 6-coordinated.

For single nets the leaders are basically the same as for interpenetrating nets; 6 nets (**dia**, **pcu**, **srs**, **cds**, **pts**, **nbo**) are among the ten most frequent nets in both samples. The top list of single nets contains a wider range of coordinations: 3-, 4-, 5-, 6-, 8-, or 3,6-coordinated.

The data on the abundance of underlying nets provide wide opportunities to find relations between the chemical composition of a coordination polymers and the topological properties of the corresponding underlying net. The most symmetrical (regular) nets for most widespread coordination figures (triangle, square, tetrahedron, octahedron, cube) are **srs**, **nbo**, **dia**, **pcu** and **bcu**, respectively, and these nets are the most abundant.

We have arranged the ways of the influence of coordination features of metal centers and ligands on the overall topology of coordination polymers. We found that the number of possible overall topologies is essentially restricted by the geometrical-topological properties of coordination groups.

[1] V.A. Blatov, *IUCr CompComm Newsletter* **2006**, *7*, 4-38; http://www.topos.ssu.samara.ru. [2] V.A. Blatov, L. Carlucci, G. Ciani, D.M. Proserpio, *CrystEngComm*, **2004**, *6*, 377-395.

**Keywords: crystal chemistry, crystal design, coordination polymer**

## MS09.P02

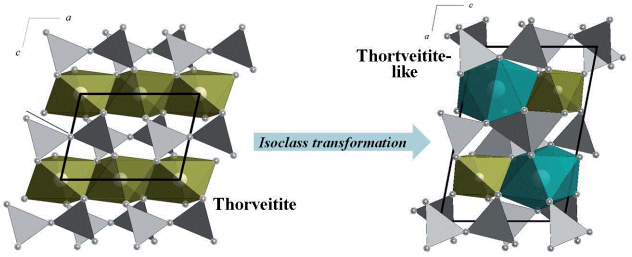
*Acta Cryst.* (2011) A**67,** C268-C269

**Structural relationships in some monoclinic layered compounds**

I.Rosales[a], L.Bucio,[a] E.Orozco[a]. [a]*Instituto de Física, U.N.A.M. Circuito Exterior, C.U. México, D.F. 04510 México.* E-mail: rosales@fisica.unam.mx

The laminar structure of thortveitite $Sc_2Si_2O_7$ (S.G. *C*12/*m*1), was reported by Zachariasen [1]. This structure offers the possibility of have structural modifications changing the scandium and silicon by other cations as (iron/indium) and germanium respectively, keeping the layered character of the crystalline arrangement. Atomic substitution in the thortveitite $FeInGe_2O_7$ (S.G. *C*12/*m*1) reduces its symmetry to thortveitite-like $YFeGe_2O_7$ (S.G. $P12_1/m1$) [2]. The aim

of this work was to determinate the structural relations between the space groups of the basic structure of layered thortveitite (S.G. *C*12/*m*1) and its derivative structure named thortveitite-like (S.G. $P12_1/m1$) when iron is partially substituted by yttrium in the layered thotveitite $FeInGe_2O_7$. The symmetry reduction promoted by the incorporation of yttrium in the formula $Y_xIn_{1-x}FeGe_2O_7$ (x = 0, 0.25, 0.50, 0.75, 0.90 and 1.0), is explained by mean of the crystallographic group-subgroup relationships.




[1] W. H. Zachariasen. *Z. Kristallogr*. **1930**, *73*, 1-6. [2] C. Cascales, M.T. Fernández-Díaz, M.A. Monge, L. Bucio, L. *Chem. Matter*. **2002**, *14*, 1995-2003.

**Keywords: symmetry relations, klassengleiche subgroups, thortveitite**

## MS09.P03

*Acta Cryst.* (2011) A**67,** C269

### Using Lee-Richards Surfaces to calculate close contacts and complex interfaces

Herbert J. Bernstein,[a] Lawrence C. Andrews,[b] [a]*Department of Mathematics and Computer Science, Dowling College, Oakdale, NY 11769 (USA).* [b]*Micro Encoder, Inc., Kirkland, WA 98034 (USA).* E-mail: yaya@dowling.edu

The calculation of geometric interactions among macromolecular fragments, such as the identification of close contacts and complex interface regions, normally involves atom-by-atom comparisons of coordinates. In a brute-force calculation distance search involving molecular fragment *A* with $N_A$ atoms and fragment *B* with $N_B$, $O(N_AN_B)$ distances need to be computed. A calculation using Levinthal cubing, Voronoi diagrams, NearTrees or other partitioning schemes can reduce this time to

$$O(\min(N_A log(N_A) + N_B log(N_A), N_B log(N_B) + N_A log(N_B))) = O((N_A + N_B)(log(min(N_A, N_B))))$$

which is excellent for fragments of very different sizes but problematic for fragments of similar sizes. It is difficult to reduce the time further with any atom-by-atom based scheme. An alternative is to base the analysis of an interface on the electron density or a pseudo-density, which should display significant excess density in an interface, or on a molecular surface, which should tend to bridge an interface. Use of the Pseudo-Gaussian Approximation to Lee-Richards Surfaces (PGALRS) algorithm [2] to compute Lee-Richards surfaces [3] combined with use of a NearTree [1] to partition space allows generation of a molecular surface in linear time and identification of atoms forming a surface (*LRSA*()) in linear time. The atoms in the interface are then

$$(LRSA(A) \cup LRSA(B)) \cap (\neg LRSA(A \cup B))$$

which can be computed in $O(N_A+N_B)$ time. This calculation is sufficient for identification of close contracts and an efficient first approximation for interface identification. We are looking at using atom pairs for more accurate interface identification. Use of only the atoms in A and B within a reasonable distance of $(LRSA(A) \cup LRSA(B)) \cap (\neg LRSA(A \cup B))$ would help to pre-prune the search tree for these calculations.

Work funded in part by grants from NIH/NIGMS and DOE.

[1] L. Andrews, *Journal* **2001,** *19*, 40-49. [2] H.J. Bernstein, P.A. Craig, *J. Appl. Crystallography* **2010**, *43:2*, 356-361. [3] B. Lee, F. M. Richards, *J. Mol. Biol.* **1971**, *55*, 379 – 400.

**Keywords: complex_interface, molecular_surface, contact**

## MS09.P04

*Acta Cryst.* (2011) A**67,** C269-C270

### Odd n-fatty acids $C_nH_{2n}O_2$ and their alloys on the X-ray powder diffraction data

Elena Kotelnikova, Julia Trushkina, *Department of Crystallography, St. Petersburg State University, St. Petersburg (Russia).* E-mail: elena@ek7740.spb.edu

Saturated fatty acids $C_nH_{2n}O_2$ play important role in the life of plants and animals. Some even ones (even number *n*) were specified in the literature including our works. We refined X-ray characteristics of six even acids (*n* = 12, 14, 16, 18, 20, and 22) [1]; determined thermal deformations and polymorphic transformations of three acids (*n* = 12, 14 and 16); synthesized double alloys $C_{12}H_{24}O_2$–$C_{14}H_{28}O_2$ and $C_{14}H_{28}O_2$–$C_{16}H_{32}O_2$ as well as quaternary one $C_{12}H_{24}O_2$–$C_{14}H_{28}O_2$–$C_{16}H_{32}O_2$ in relative binary and ternary systems; and revealed rather limited isomorphism in these mixtures [1, 2].

Oppositely, odd fatty acids (odd *n*) have a weak state of knowledge because of their less spread in the nature [3]. For example, information about odd acids missed in the ICDD bank [4]. In this work we present results of X-ray powder diffraction study of odd fatty acids.

X-ray characteristics (crystal system, polymorph modification, indices *hkl*, elementary cell parameters, space group, etc.) of four odd fatty acids with *n* = 13, 15, 17, and 19 (homological purity 98–99 %) were obtained for the first time. In particular it was found that odd acids could crystallize at room temperature in monoclinic polymorphs of single and double layers.

Limits of solid solutions were established for two binary systems of odd components: $C_{13}H_{26}O_2$–$C_{15}H_{30}O_2$ (studied compositions, mol. % $C_{15}H_{30}O_2$: 10, 33, 50, 67, and 90) and $C_{15}H_{30}O_2$–$C_{17}H_{34}O_2$ (studied compositions, mol. % $C_{17}H_{34}O_2$: 10, 25, 33, 50, 67, 75, and 90). Our experiments showed rather limited isomorphism too, like in binary systems of even components [1, 2]. Thus, rather limited isomorphism specifies n-fatty acids even in systems of one parity. This differs principally n-fatty acids from n-paraffins, which molecules mix together in wide ranges forming solid solutions easily [5].

The diffractograms of the majority of melted binary mixtures of odd n-fatty acids displayed peaks of binary acid compound. Remarkable, the binary compounds are the new acids. They are presented by dimers combined of two molecules of different lengths. Herewith, we identified on these diffractograms also peaks of solid solutions, which compositions are nearly close to those of the excess components.

In binary systems of mixed parity, binary compounds of fatty acids do not apparently form after melting. For example, the diffractogram of the melted mixture (mol.) $C_{18}H_{36}O_2$: $C_{19}H_{38}O_2$ = 1:1 represents peaks of only the original components. In our opinion, the limitation of the isomorphic miscibility should be caused by the dimeric nature of fatty acid molecules.

Apart from this, structural (thermal) deformations and polymorphic transitions of stearinic acid $C_{18}H_{36}O_2$ were studied in view of the variety its polymorphic modifications; high temperature X-ray powder diffraction method was used.