JSS CH2 (KEMV992) Determination of Solution Phase Structures

Prof. Petri Pihko

Department of Chemistry

UNIVERSITY OF JYVÄSKYLÄ

The National Doctoral Programme in Nanoscience

NGS-NANO

Doctoral Program of

OCCB

Organic Chemistry and Chemical Biology

Lecturers

Lecturer:
Prof. Samuel H. Gellman
University of Wisconsin-Madison
gellman(at)chem.wisc.edu

Course coordinator
Prof. Petri Pihko
O505
Petri.Pihko(at)jyu.fi

Computational assistance:
Dr. Ádám Madarász, Hungarian Academy of Sciences

Course tutor: Ms. Sanna Yliniemelä-Sipari

Goals of the course

- To emphasize the importance of structural information in predicting activity, reactivity and selectivity
- To promote and provoke the use of tools of structural chemistry in the solution phase
  - NMR methods
  - Computational methods
  - Integration of various methods

Timetable

Wednesday 8.8.
13:00-14:00 Introduction lecture by Petri Pihko
14:15-15:45 Lecture 1 by Sam Gellman (SG): Hydrogen-bonded structures
15:45-17:00 Demonstration and initiation of computational exercise

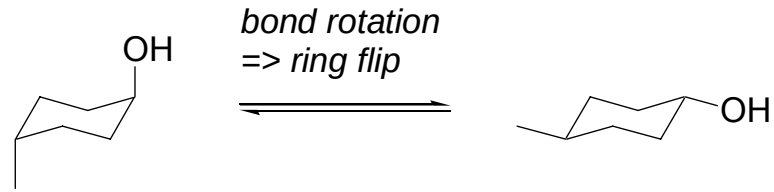
Thursday 9.8.
08:30-11:00 Lectures 2&3 by SG: Aromatic-Aromatic Interactions and beta-Sheets
11:00-12:00 Lunch break
12:00-14:00 Continuation of the computational exercise
14:00-16:00 Lecture 4 by SG: Tertiary Structures

Friday 9.8.
09:00-11:00 Lecture 5 by SG: Foldamers
11:00-12:00 Concluding remarks, discussion of the computational results

What is a ”solution structure”?
And why should we be interested?

- Most (but not all!) biological recognition events, catalytic processes and indeed most chemical reactions take place in the solution phase
- The shapes of the molecules in solution – their conformations – are of course not as ”frozen” as in the solid state
- However, small molecules can display a high level of conformational rigidity in solution if there are sufficient *conformational constraints* (i.e. allylic strain) or attractive interactions (hydrogen bonding, dispersion effects etc.)
- Rigid, conformationally constrained molecules are often useful for
  - Platforms for new catalysts and ligands (enantioselective catalysis)
  - in drug design where the presence of multiple conformations may reduce binding and hence the efficacy of the pharmaceutical candidate
- **A ”solution structure” is an ensemble of lowest energy *conformations* of the molecule**
  - A combination of computational and NMR tools are often used to determine the solution structures

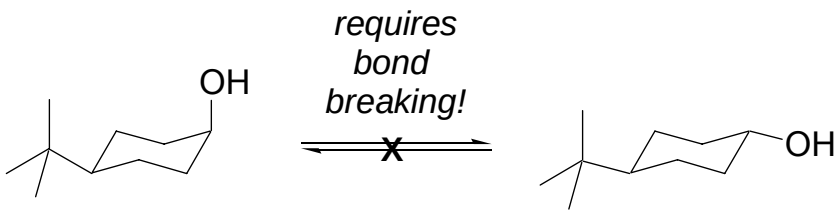
Conformation and configuration



diaxial *conformer*

diequatorial *conformer*

same ***configuration***
different ***conformation***



syn *isomer*

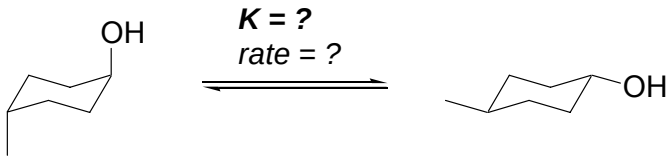
anti *isomer*

different configuration

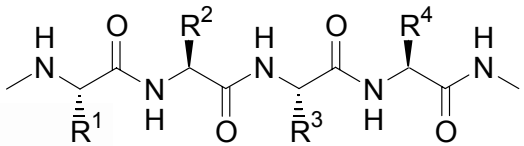
UNIVERSITY OF JYVÄSKYLÄ

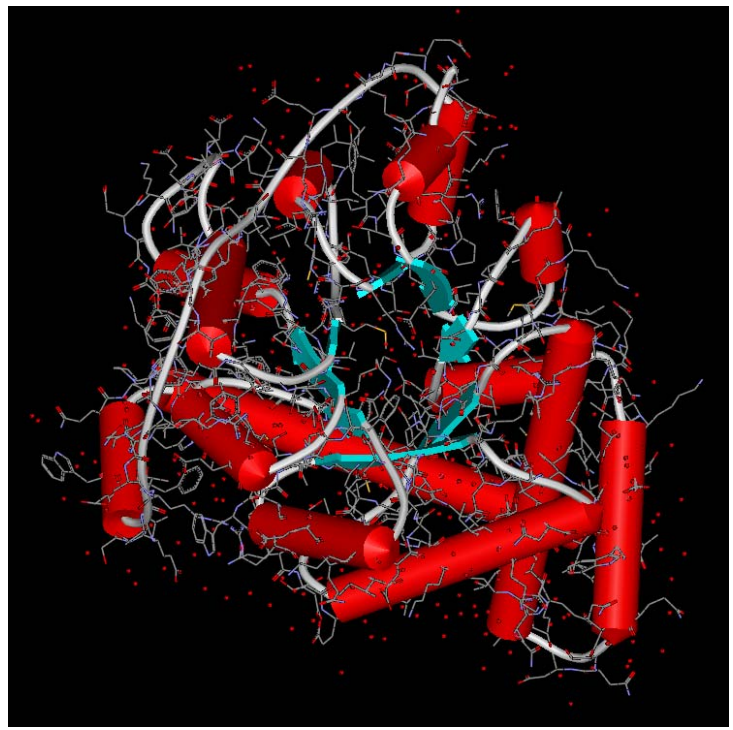
Conformational analysis

• conformational preferences in cyclic and acyclic molecules



• prediction of structural parameters
• predictions on reactivity
• predictions of binding, inhibitory activity, biological activity -> medicinal chemistry
• shapes of proteins

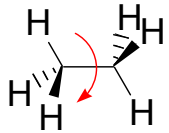




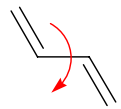
E. Coli transaldolase enzyme

UNIVERSITY OF JYVÄSKYLÄ

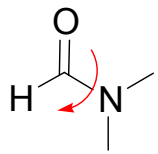
Barriers to rotation around single bonds



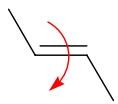
12 kJ/mol



30 kJ/mol

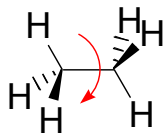


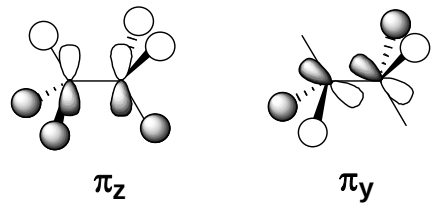
85 kJ/mol



260 kJ/mol

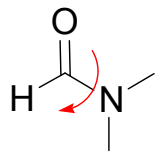
Most C-C σ bonds have a varying degree of π character:

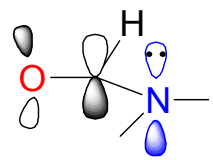




π_z π_y

two molecular orbitals of ethane





Nitrogen lone pair donates e^- density to C=O

Detection of rotational barriers

- The barrier of 20 kJ/mol for butane corresponds to a rate of rotation of two billion (2^x10^9) rotations s^{-1}. The different *conformers* cannot be detected by NMR.

- The 'shutter speed' of any spectroscopic method is given by the equation

$$k = \pi\Delta\nu/\sqrt{2} = 2.22 \times \Delta\nu$$

where k is the fastest exchange rate that allows separation of individual signals and $\Delta\nu$ is the separation of those signals in Hz.

For a 400 MHz NMR spectrometer, two signals separated by 1 ppm are 400 Hz apart => k = 888 s^{-1}. However, in an IR spectrometer, two absorptions separated by 100 cm^{-1} correspond to a wavelength of 0.01 cm or a frequency of $3x10^{12}$ s^{-1}

=> k = ca. $7x10^{12}$ s^{-1}.

=> IR can detect changes happening a lot faster than NMR!

Computational methods

- *ab initio*
  - hydrogen-like orbitals are used to arrive at a self-consistent field
  - fundamentally the most accurate and reliable method
  - separation of individual energy contributions not always easy
  - requires heavy computational resources
  - allows computational treatment of solvation models, transition states etc.
- *Molecular mechanics*
  - empirical force fields based on classical mechanical analogues
  - quick, accurate and reliable (atomic positions ca 0.01 Å, angles 1-2 °, ΔH_f ca 10 kJ/mol)
  - can be done on a PC on fairly large systems, even dynamics quite easy
  - limitations: poor with electronic effects

Computer times and results for propane

	Molecular mechanics (MM2)	Semi-empirical MINDO/3	*ab initio* 3-21G	*ab initio* 6-31G*	Exp.
CPU time	0.83	9.75	550	4702	-
r_{CC}	1.534	1.495	1.541	1.528	1.526
CCC angle °	111.7	121.5	111.6	112.7	112.4
ΔH^0_f (kcal/mol)	-24.8	-26.5	-	-	-25.0

UNIVERSITY OF JYVÄSKYLÄ

Molecular mechanics

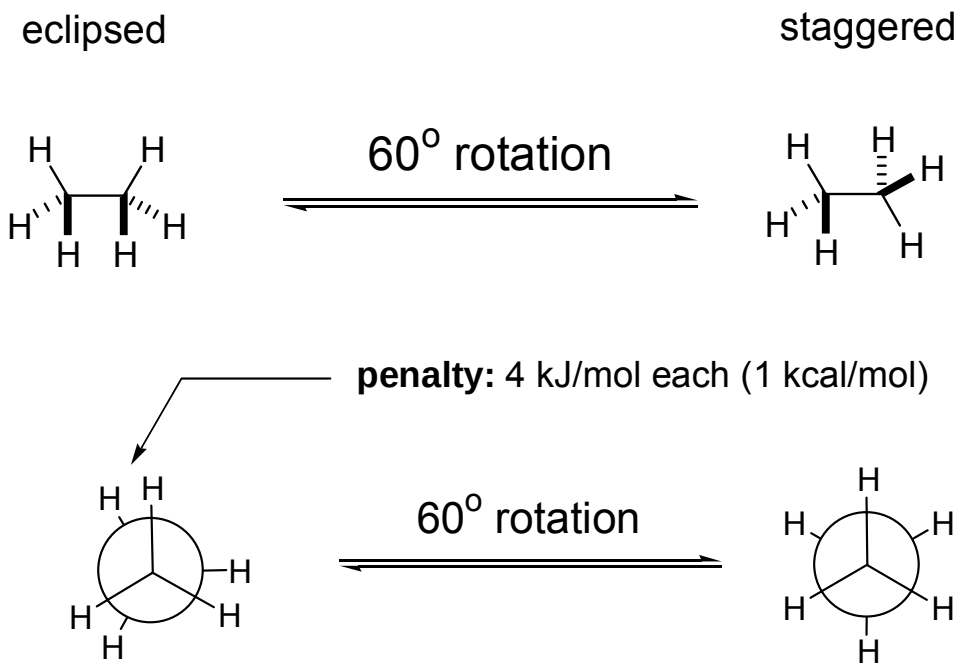
- *Empirical force fields*
  - Total strain energy: bond stretching + bond angle distortion + torsional strain + nonbonded interactions
  - $E_s = E(r) + E(\theta) + E(\phi) + E(d)$
  - Simplified equations:
  - $E(r) = 0.5\ k_r(r-r_0)^2$
  - $E(\theta) = 0.5\ k_q(\Delta\theta)^2$
  - For molecules with a threefold torsional barrier, $E(\phi) = 0.5\ V_0(1 + \cos 3\phi)^2$
  - Nonbonded interactions: van der Waals energies
  - Today several force fields available, most common are MMFF, MM3 and AMBER

Cliff notes on energy minimization

- *Minimization methods:*
  - **Steepest descent (SD):** follows the gradient of the energy function at each step
  - Can lead to backtracking
  - Does not converge easily (gradient becomes smaller!)
  - OK if the conformation is far from minimum
  - **Conjugate gradient (CG):** remembers the gradients calculated from previous steps
  - Reduces backtracking; faster than SD
  - **Newton-Raphson or BFGS:** predicts the location of the minimum and begins a descent towards it
  - Requires most memory

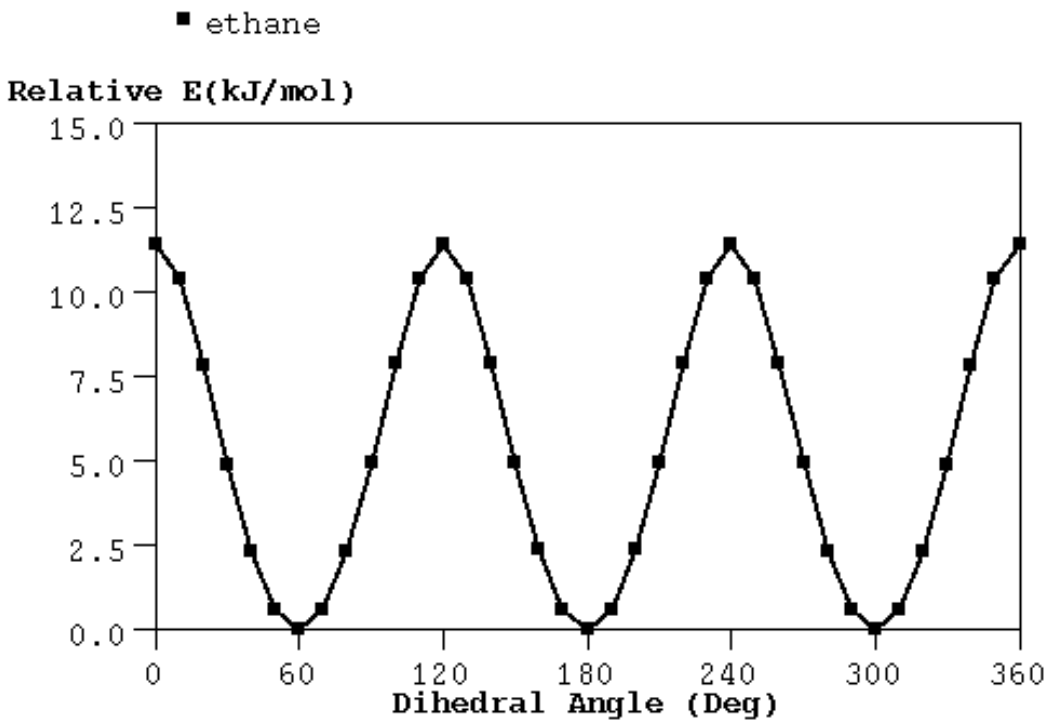
UNIVERSITY OF JYVÄSKYLÄ

Ethane



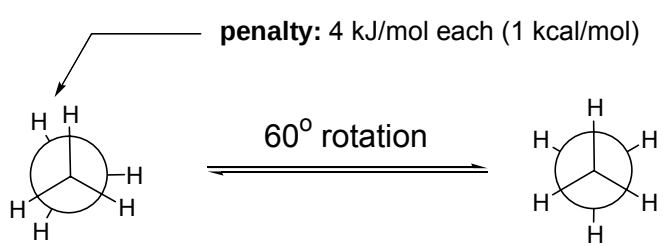
UNIVERSITY OF JYVÄSKYLÄ

Ethane

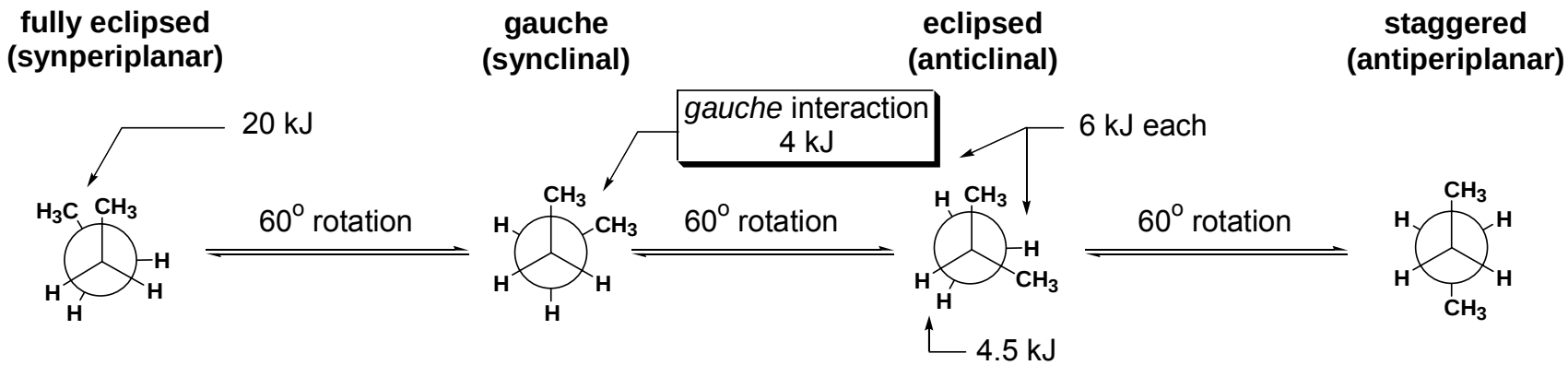


Conclusion: barrier in ethane is ca. 12 kJ/mol ≈ 3 kcal/mol

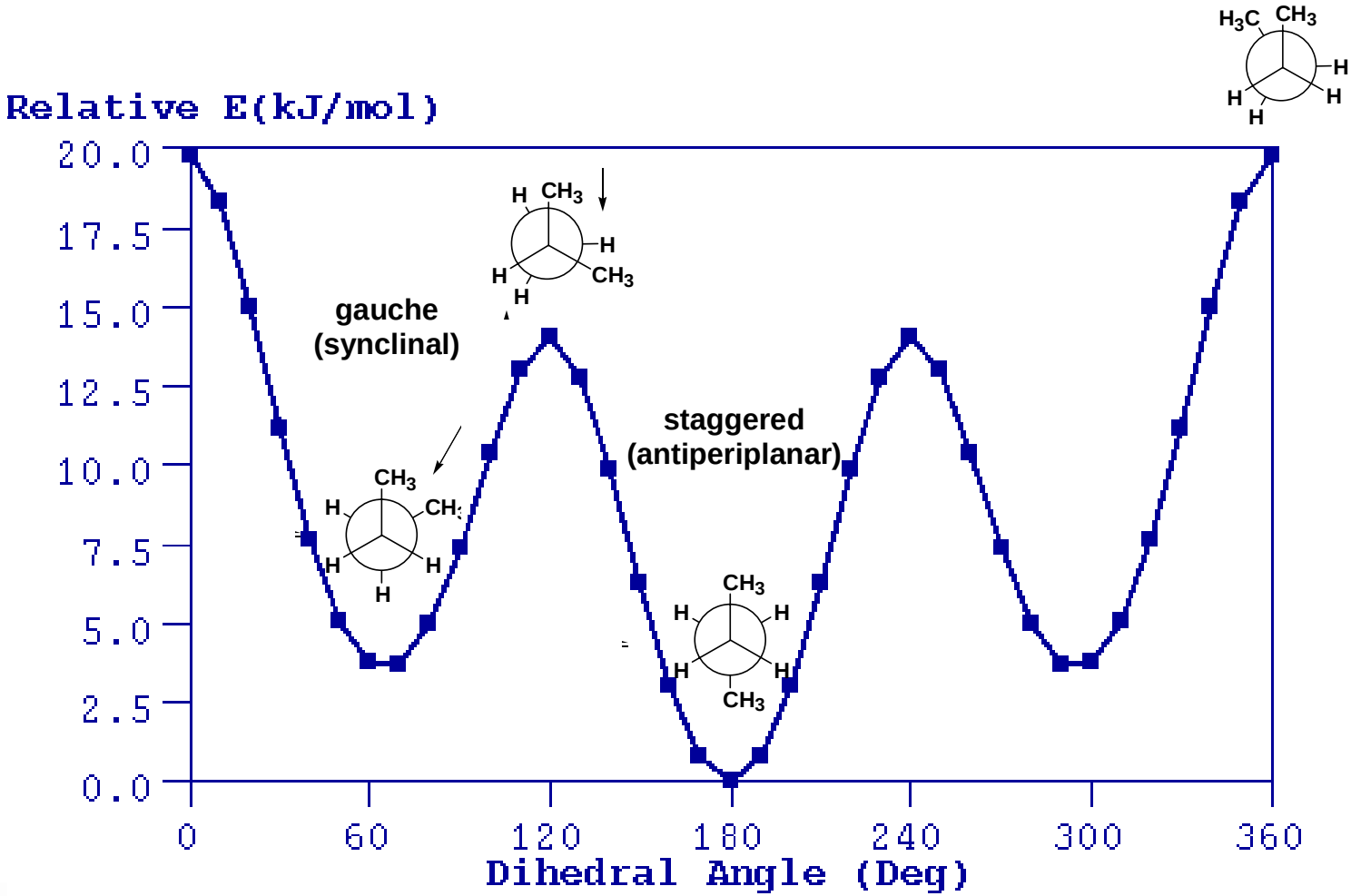
1 kcal/mol per eclipsing H-C-C-H



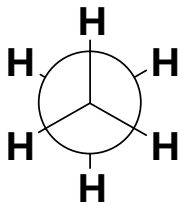
Butane



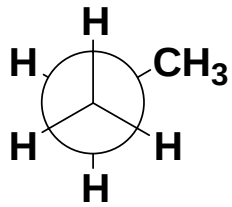
Butane



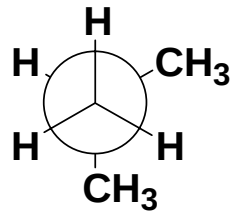
Rotational barriers in other molecules



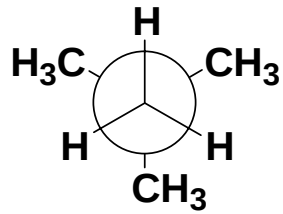
12 kJ/mol



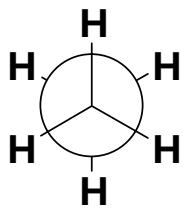
14 kJ/mol



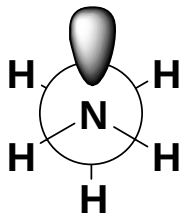
20 kJ/mol



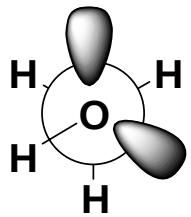
>20 kJ/mol



12 kJ/mol



8 kJ/mol



5 kJ/mol

Note: rotation becomes easier as the number of eclipsing bonds decreases

UNIVERSITY OF JYVÄSKYLÄ

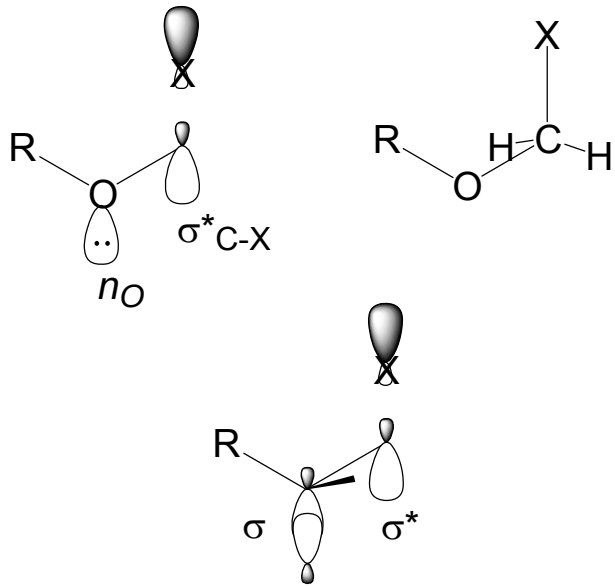
Composition–equilibrium–free energy

$$\Delta G = -RT \ln K$$

More stable Isomer (%)	Equilibrium Constant (K)	Free energy ΔG_{25} (kcal/mol)
50	1	0.0
55	1.22	-0.119
60	1.50	-0.240
65	1.86	-0.367
70	2.33	-0.502
75	3.00	-0.651
80	4.00	-0.821
85	5.67	-1.028
90	9.00	-1.302
95	19.00	-1.744
98	49.00	-2.306
99	99.00	-2.722
99.9	999.0	-4.092

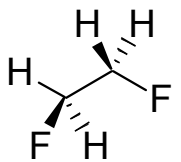
Donors and acceptors

- Anomeric effect**: interaction between a lone pair n and neighboring antibonding σ* orbital**
- **Interaction greatest when the orbitals are *antiperiplanar***
- Generally**: *The most favorable conformations have the best donor (lone pair or bond) antiperiplanar to the best acceptor bond.***

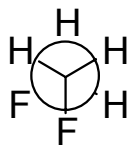


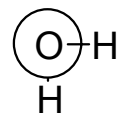
Donors, acceptors and gauche-*effect*

- Donor orbitals: $n_N > n_O > \sigma_{C-C}, \sigma_{C-H} > \sigma_{C-X}$ (X = N > O > S > Hal)
- Acceptor orbitals: $\pi^*_{C=O} > \sigma^*_{C-Hal} > \sigma^*_{C-O} > \sigma^*_{C-C}, \sigma^*_{C-H}$
- *gauche effect*: In systems X-C-C-Y (X = electronegative group) X and Y disfavor antiperiplanar orientation to each other, but prefer antiperiplanarity to C-H or C-C bonds.



≡



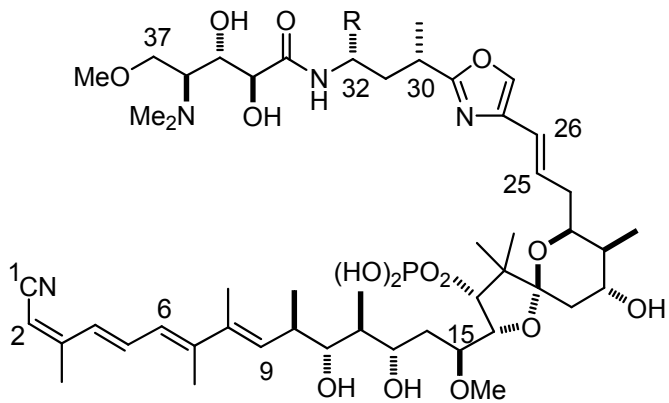


H_2O_2

not stable

Case: Calyculin A

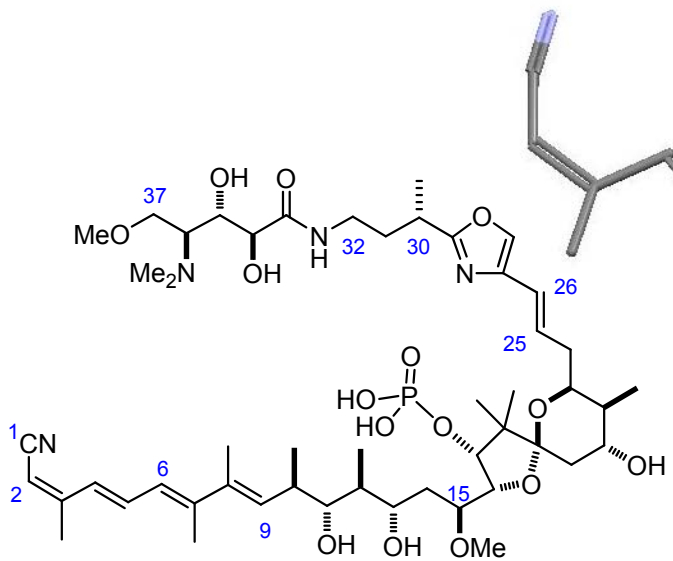
- Inhibitor of protein phosphatases 1 and 2A (PP1/PP2A), key regulatory enzymes that regulate the level of phosphorylation of many proteins
- Other inhibitors of PP1/PP2A include okadaic acid, tautomycin, and the microcystins (from cyanobacterial blooms)



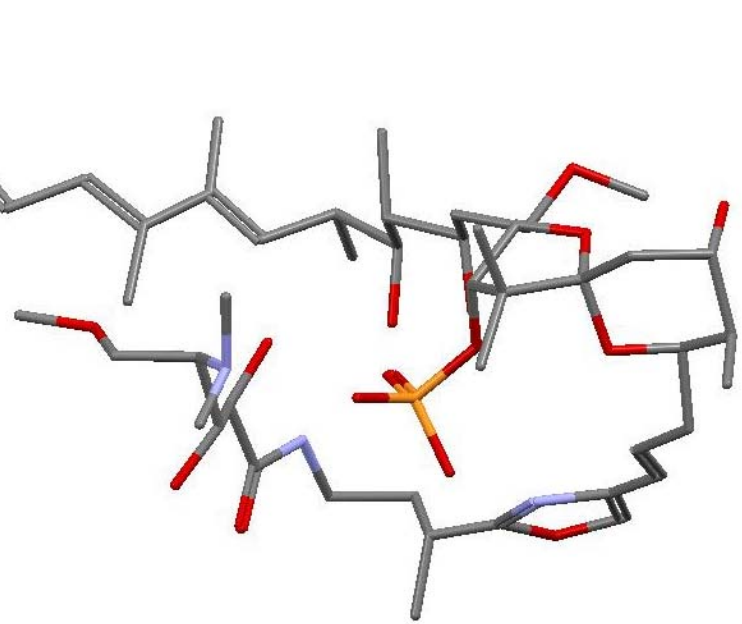
Calyculin A (1): R = H
Calyculin B (2): R = H, (2*E*)-isomer
Calyculin C (3): R = Me
Calyculin D (4): R = Me, (2*E*)-isomer

X-ray structure of calyculin A

- Is this how the compound binds to its target?



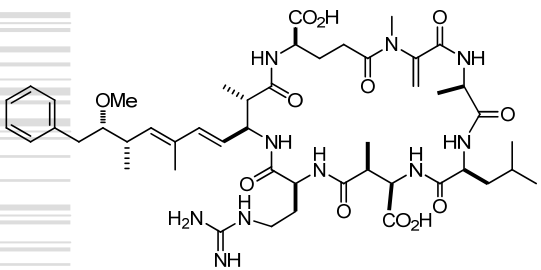
Calyculin A



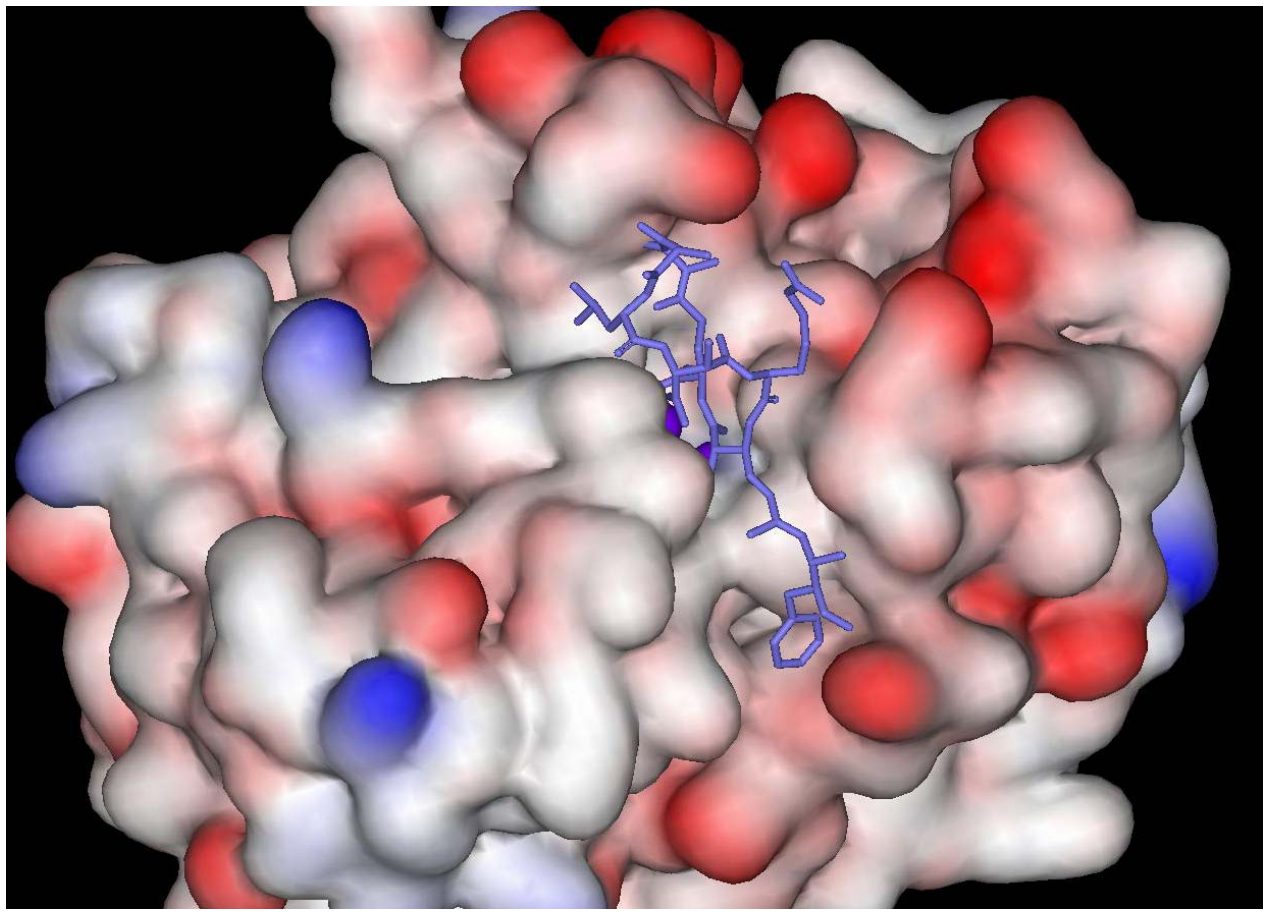
For a comprehensive docking study of multiple inhibitors with PP1 and PP2A, see: Gauss, C.-M.; Sheppeck, J. E.; Nairn, A. C.; Chamberlin, R. *Bioorg. Med. Chem.* **1997,** *5*, 1751-1773.

UNIVERSITY OF JYVÄSKYLÄ

Starting point (1995): X-ray of PP1 with microcystin inhibitor



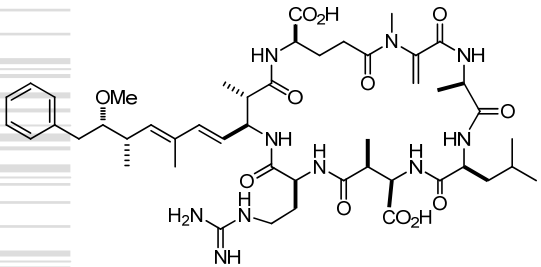
Microcystin LR (18)



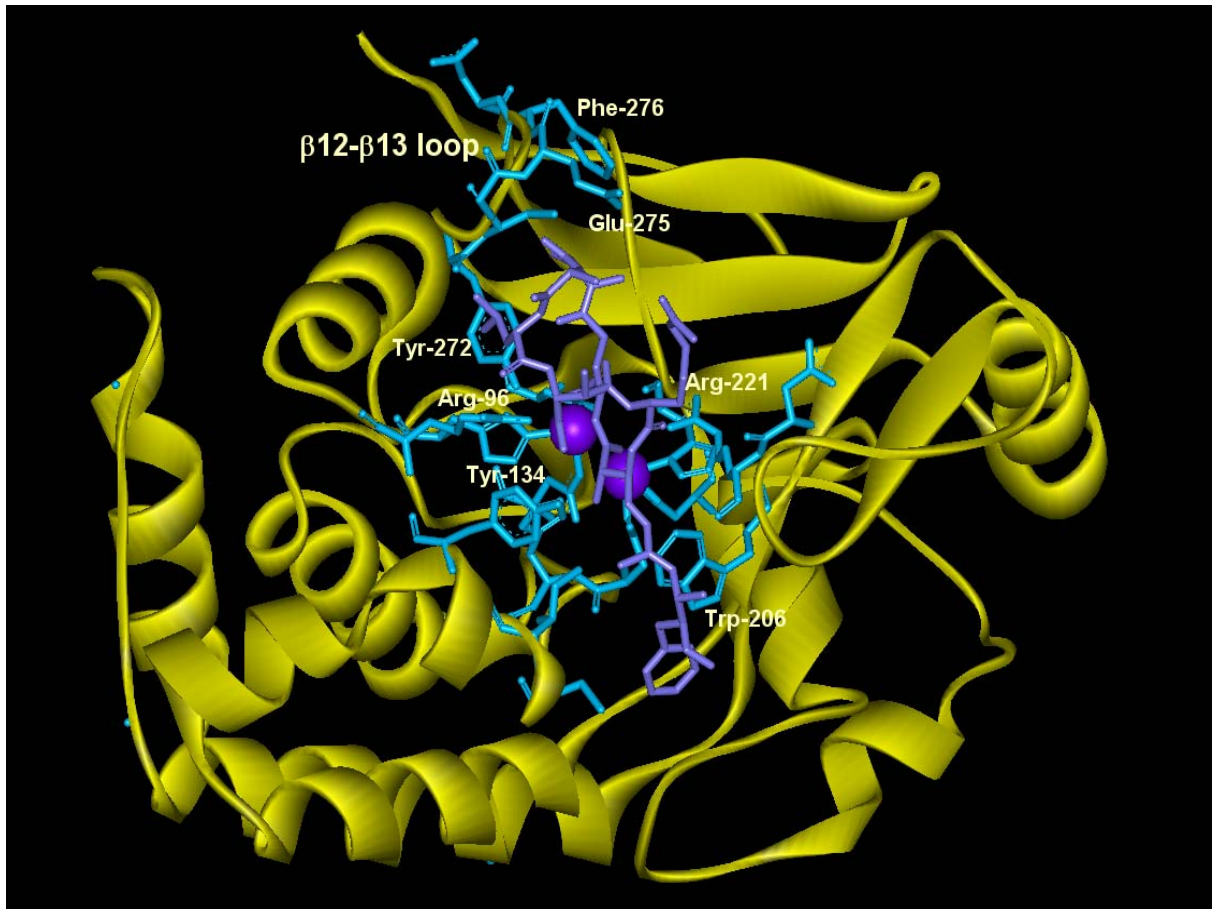
Goldberg, J.; Huang, H.-B..; Kwon, Y.-G.; Greengard, P.; Nairn, A. C.; Kuriyan, J. *Nature* **1995,** *376*, 745-753.

UNIVERSITY OF JYVÄSKYLÄ

Alternative view of PP1-microcystin complex

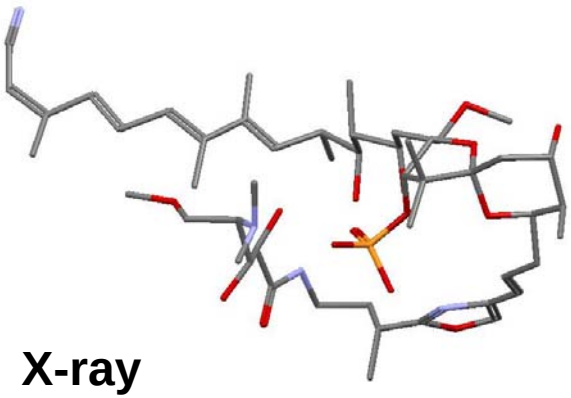


Microcystin LR (18)

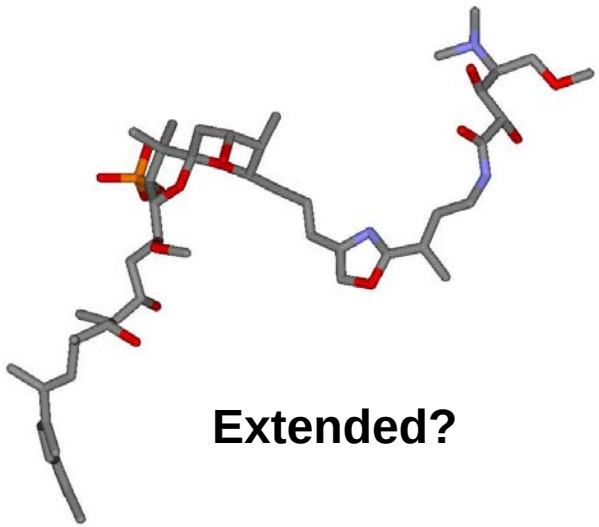


UNIVERSITY OF JYVÄSKYLÄ

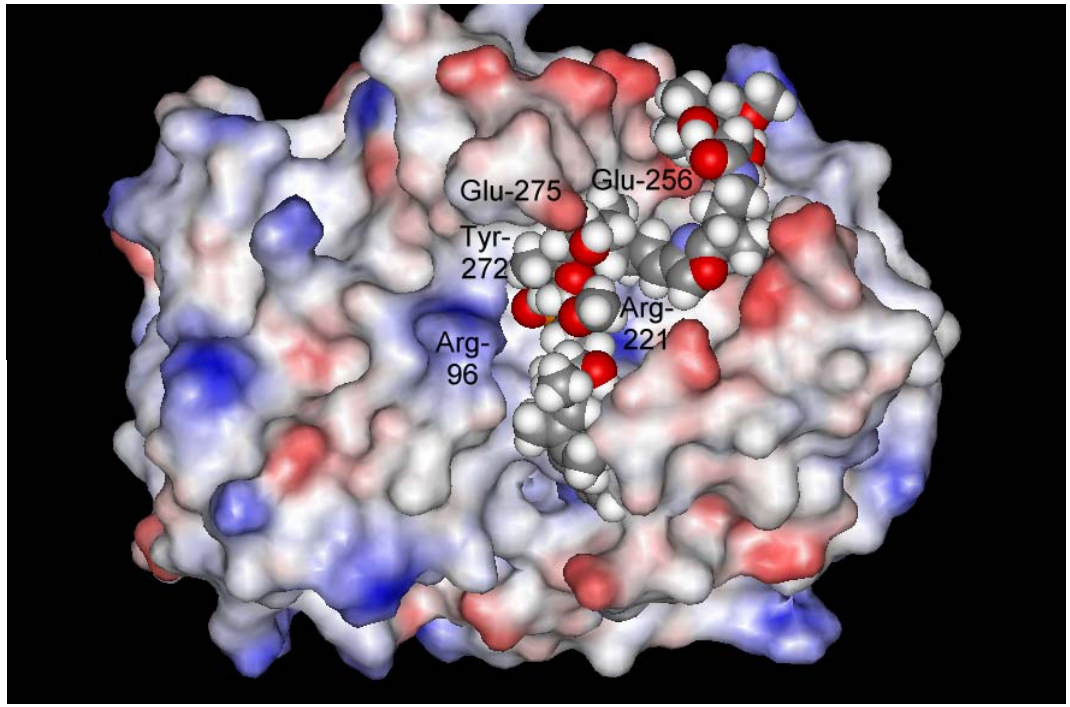
1997: An ”educated guess” – extended conformation for calyculin A?



X-ray



Extended?



A model of calyculin A (space-filling model) bound to the active site of PP1. The solvent-accessible surface of PP1 is displayed. Regions of the surface that have a highly negative electrostatic potential are shown in red, with a smooth variation in color through white (zero) to blue (positive). Selected residues contacting the inhibitor on the surface of the protein are also shown.

Lindvall, M. K.; Pihko, P. M.; Koskinen, A. M. P. ‘*J. Biol. Chem.* **1997**, 272, 23312-23316.

UNIVERSITY OF JYVÄSKYLÄ

Mika Lindvall

1st

Computational Chemist at Novartis

San Francisco Bay Area | Pharmaceuticals

Current	**Computational Chemist** at **Novartis**
Past	Computational Chemist at GlaxoSmithKline
Education	Oulun yliopisto University of California, Berkeley
Connections	**93** connections
Public Profile	http://www.linkedin.com/pub/mika-lindvall/8/697/79b

Share | PDF | Print | vCard | Flag

Summary

Drug discovery scientist applying computational chemistry to advance preclinical research. Experience in two pharmaceutical companies where co-inventor of novel agents in three therapeutic projects entering human clinical trials.

Specialties

Computational and experimental hit finding to identify new chemical starting points for drug discovery projects. Computer-Aided Drug Design. Kinase and other enzyme inhibitors in oncology and inflammation.

Solution structure of calyculin A

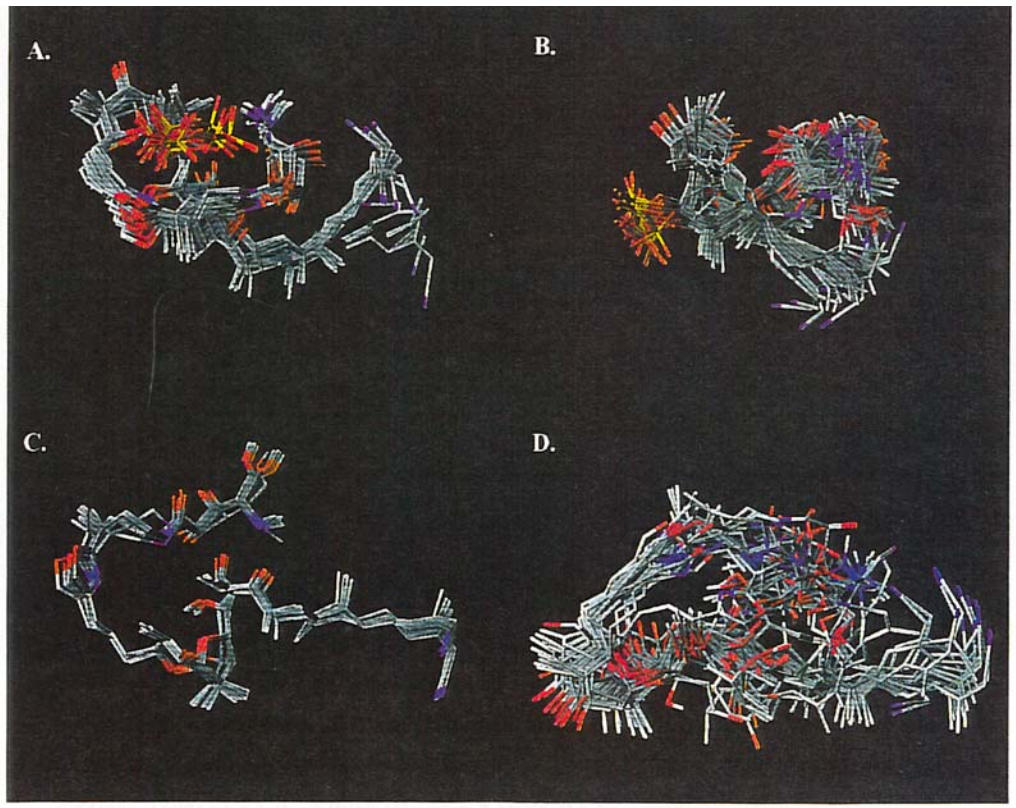


Figure 2: (A) The solution structure of Calyculin A in chloroform. (B) Calyculin A in methanol. (C) Dephosphonocalyculin A in chloroform, and (D) Dephosphonocalyculin A in methanol.

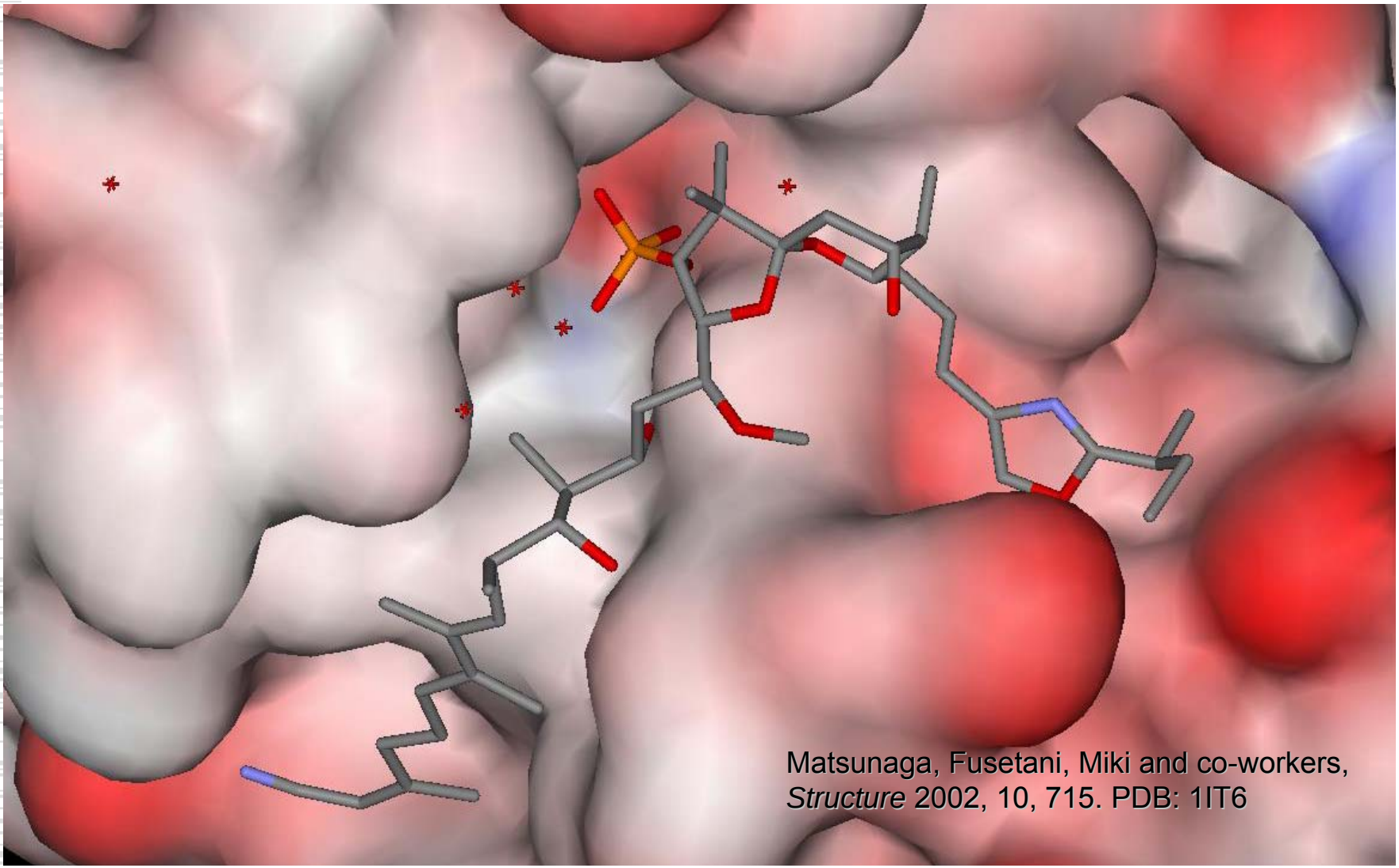
Volter, K. E.; Pierens, G. K.; Quinn, R. J.; Wakimoto, T.; Matsunaga, S.; Fusetani, N. *Bioorg. Med. Chem. Lett.* **1999,** *9*, 717-722

Tools for the solution structure generation (1999)

- NMR spectra of Calyculin A in $CDCl_3$ or in CD_3OD were obtained:
  - COSY and TOCSY for ^1H assignments
  - NOESY and ROESY experiments (mixing times 50 to 250 ms) to generate distance constraints:
    - Cross-peaks were divided into strong (1.8 – 2.5 Å distance constraint), medium (1.8 – 3.5 Å) and weak (1.8 – 5.0 Å)
    - These distance constraints were then used in simulated annealing in *MacroModel 6.0* (MM2* force field) – this involved heating the molecule to 1000 K for 0.2 ms, followed by cooling to 200 K to generate 100 structures. These structures were then minimized (MM2*) using TNCG (Truncated Newton Conjugate Gradient) method.
- This procedure eventually led to an ensemble of structures, and the 20 lowest energy conformers were selected and superimposed
  - Today the procedure would be essentially the same except that the modeling tools have become easier to use and more accurate

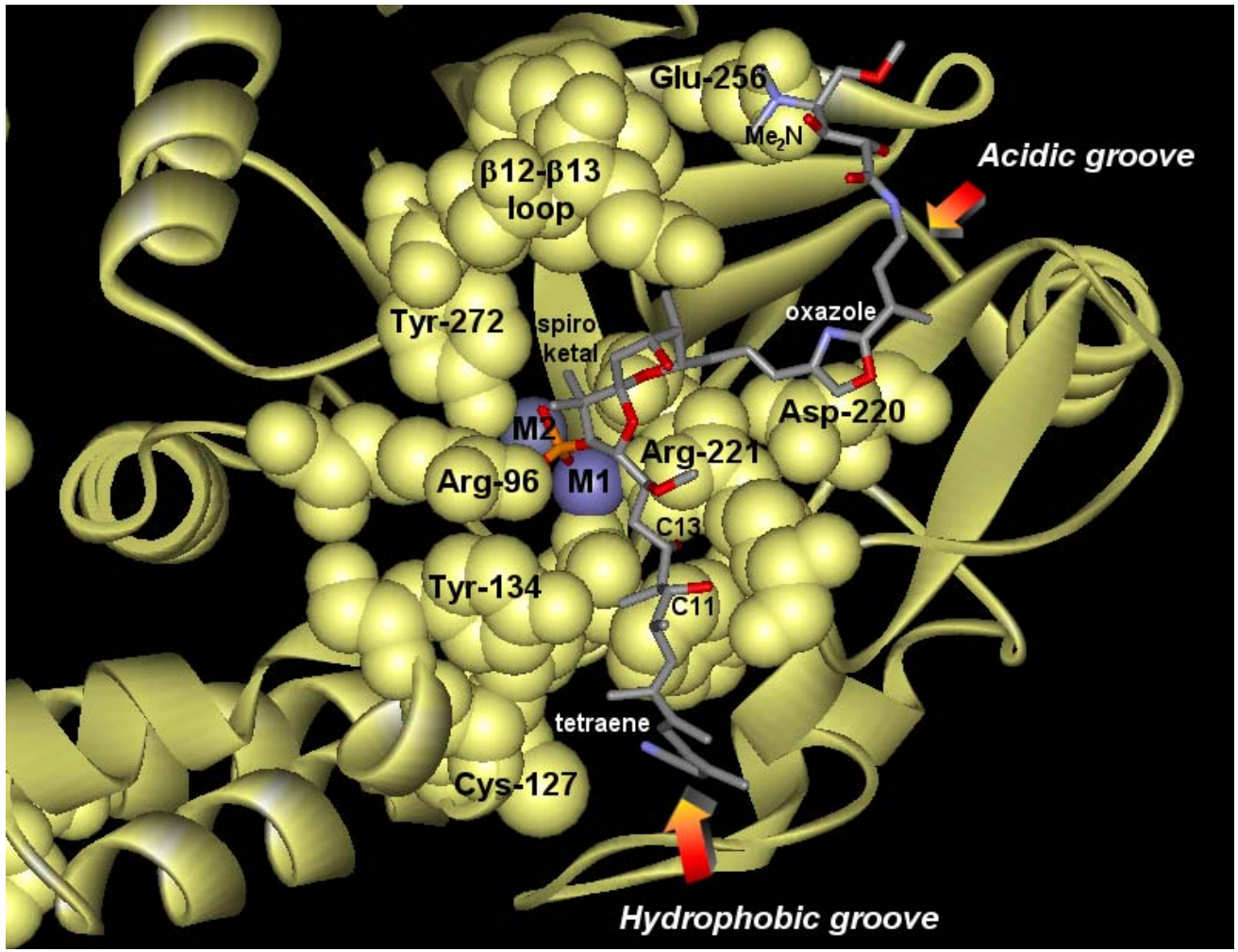
Volter, K. E.; Pierens, G. K.; Quinn, R. J.; Wakimoto, T.; Matsunaga, S.; Fusetani, N. *Bioorg. Med. Chem. Lett.* **1999,** *9*, 717-722

X-ray of calyculin A bound to PP1 – vindication of the extended conformation



Matsunaga, Fusetani, Miki and co-workers, *Structure* 2002, 10, 715. PDB: 1IT6

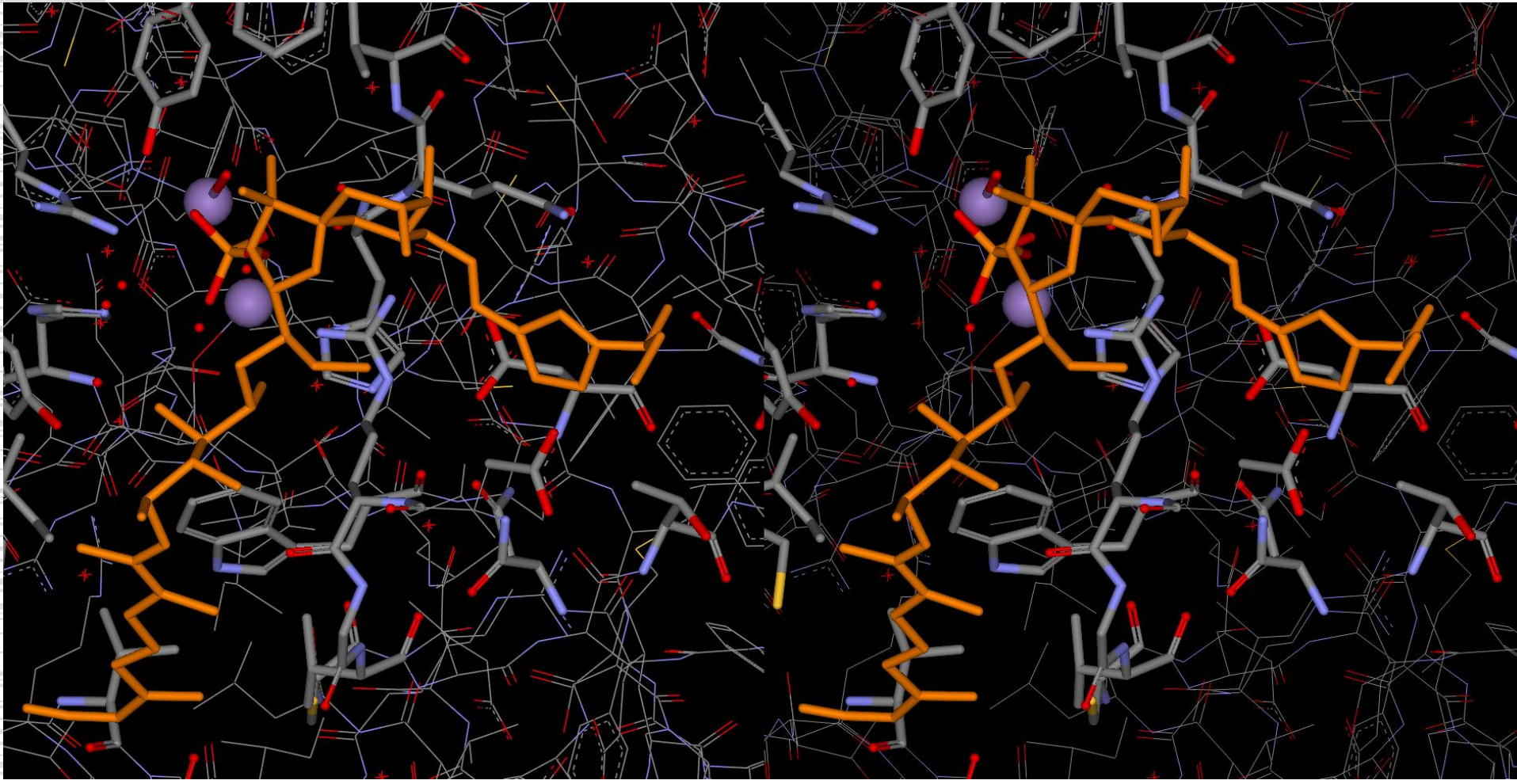
Alternative view



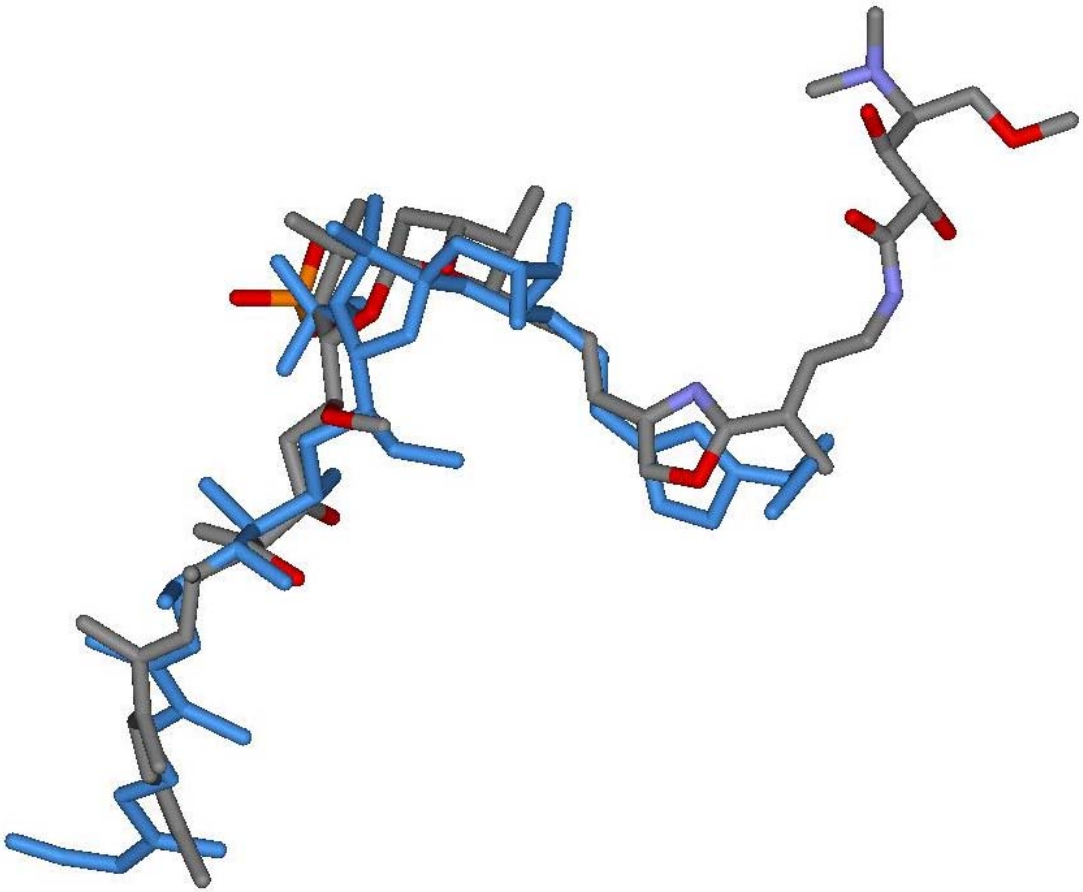
UNIVERSITY OF JYVÄSKYLÄ

Stereo view of calyculin A bound to PP1

Calyculin A: orange, key contacting residues are shown in stick model, and two Mn^{2+} ions in the enzyme are shown as lilac spheres



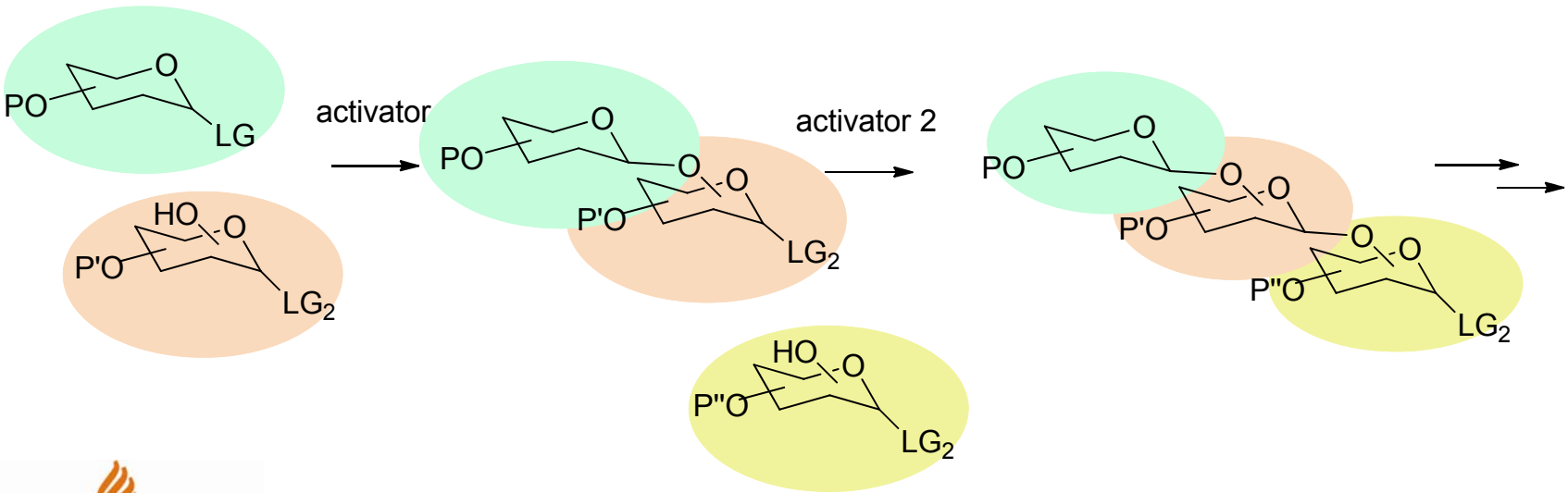
Overlay of predicted and X-ray structure of calyculin A in PP1



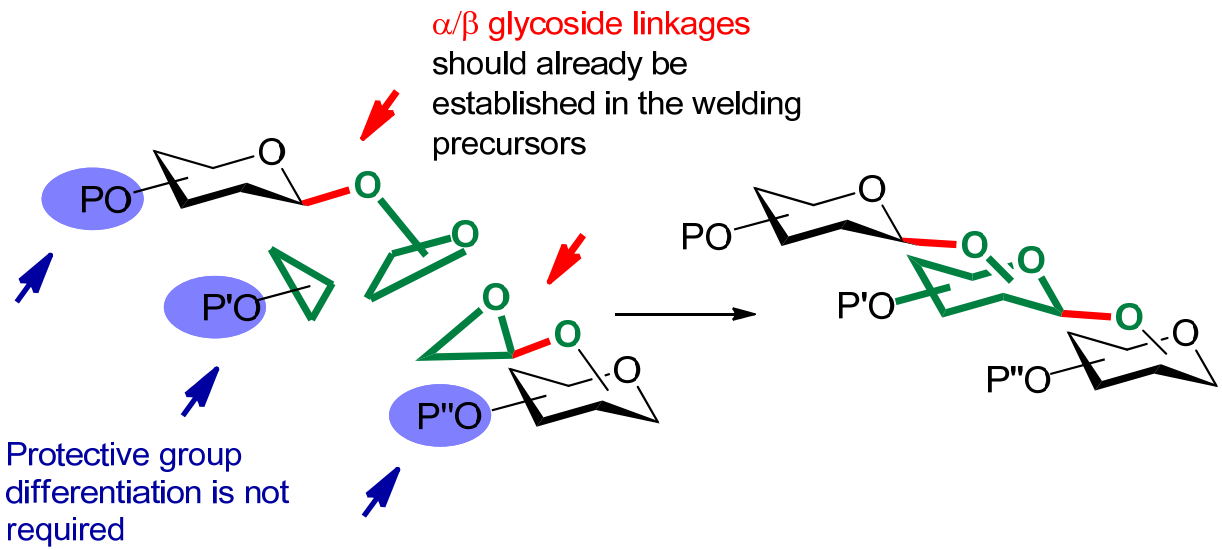
UNIVERSITY OF JYVÄSKYLÄ

Classical Oligosaccharide Synthesis

- Requires the synthesis of differentially protected donors and acceptors
- In the construction of more complex oligosaccharides, very lengthy syntheses can result

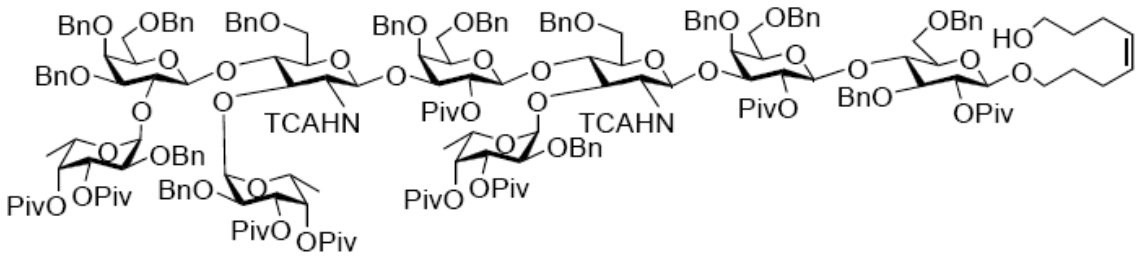


Case 2: de novo synthesis of oligosaccharides

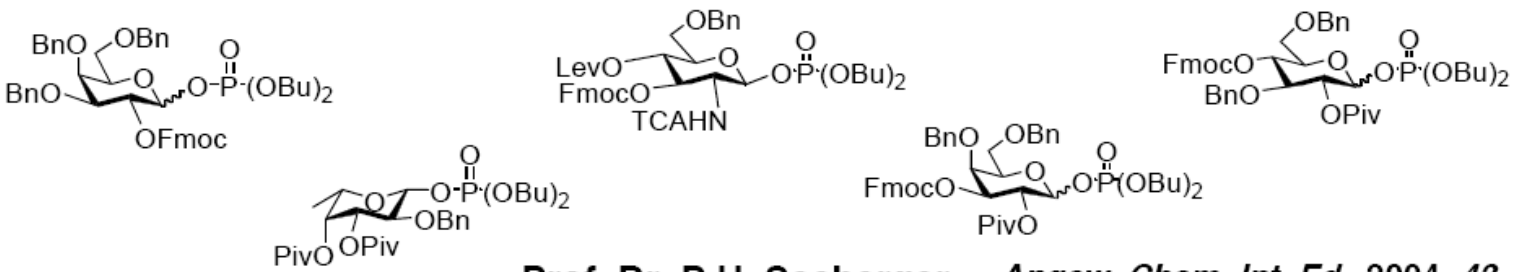


UNIVERSITY OF JYVÄSKYLÄ

Examples of Classical Subunits



Le^yLe^x

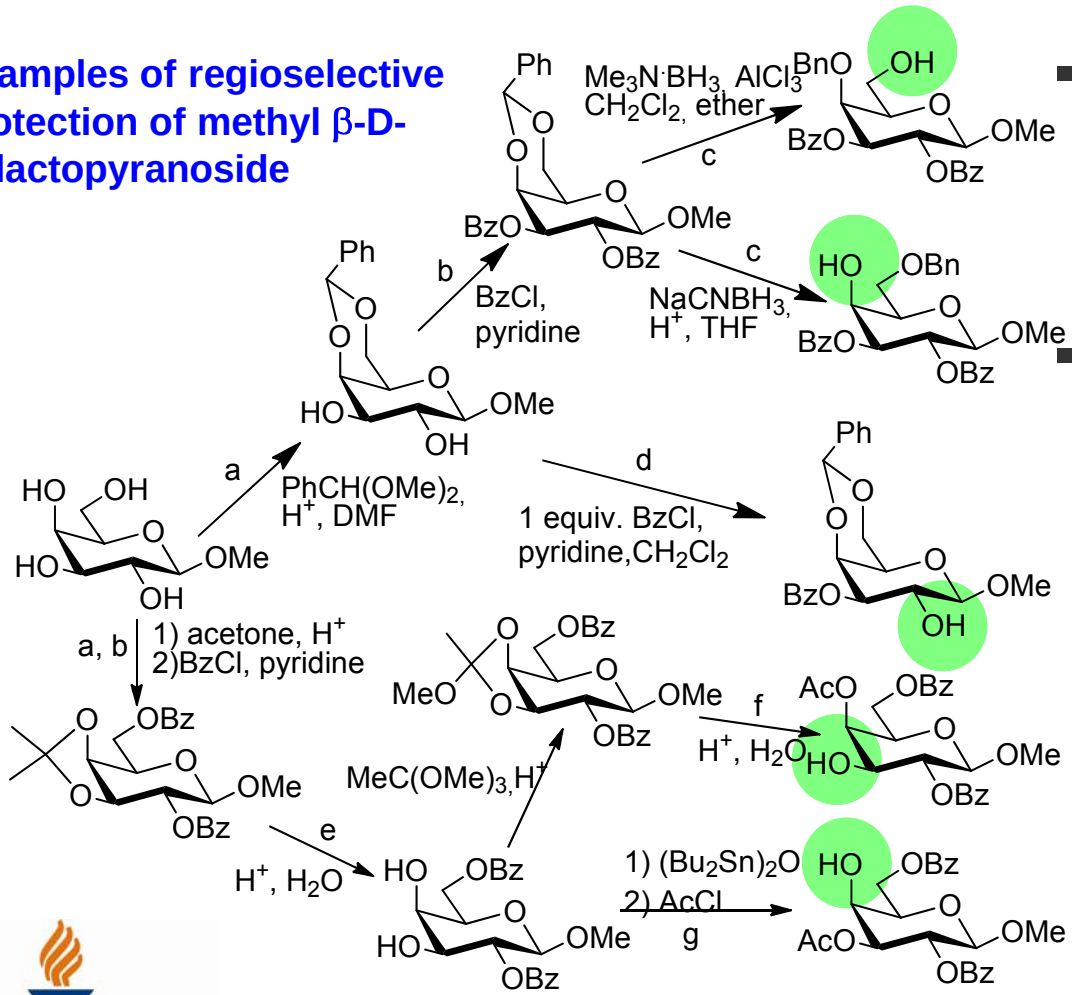


Prof. Dr. P.H. Seeberger ***Angew. Chem. Int. Ed.*** **2004, *43*, 602**

UNIVERSITY OF JYVÄSKYLÄ

Protection Strategy

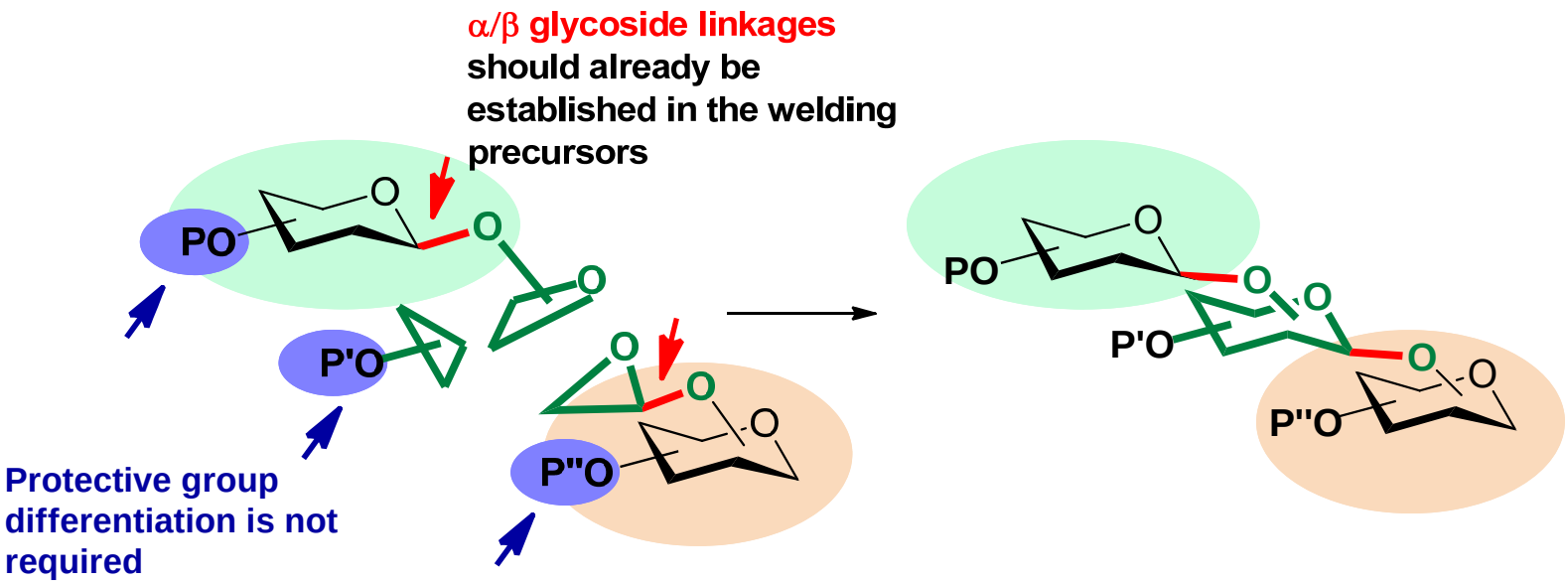
Examples of regioselective protection of methyl β-D-galactopyranoside



- This kind of differential protection must typically be performed to each and every monosaccharide subunit
- And we are not even touching the subject of donor activation or a/b selectivity in the glycosylation step!

- Stefan Oscarson, GGS Glycosynthesis course 25.05.2007

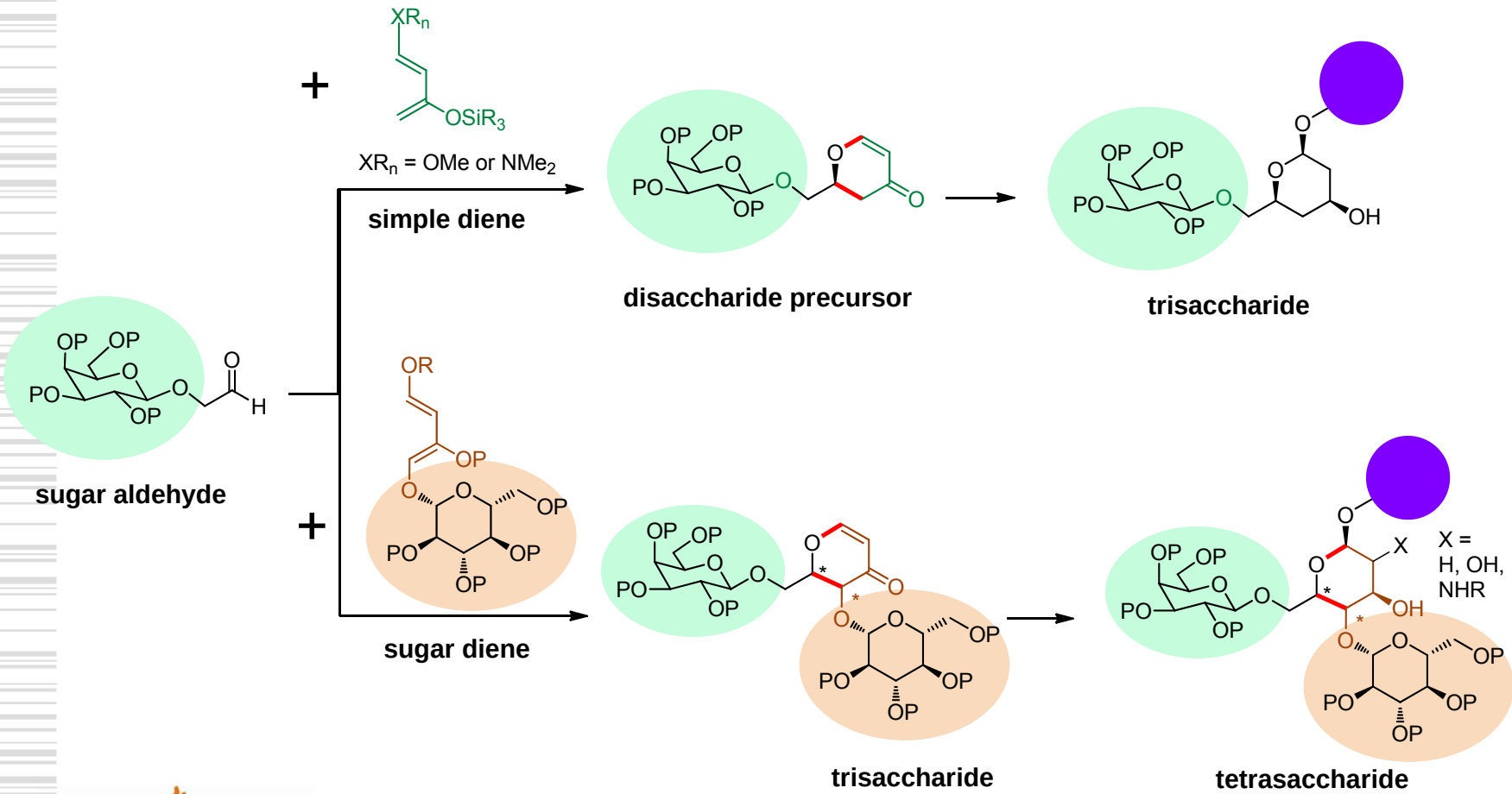
De Novo Saccharide Welding: the Principle



- The saccharide units, already including the glycosidic linkages, are *welded* together to generate a central monosaccharide unit
- No need to worry about glycosidic linkages in the coupling step
- No need for differential protection -> fewer final deprotection steps

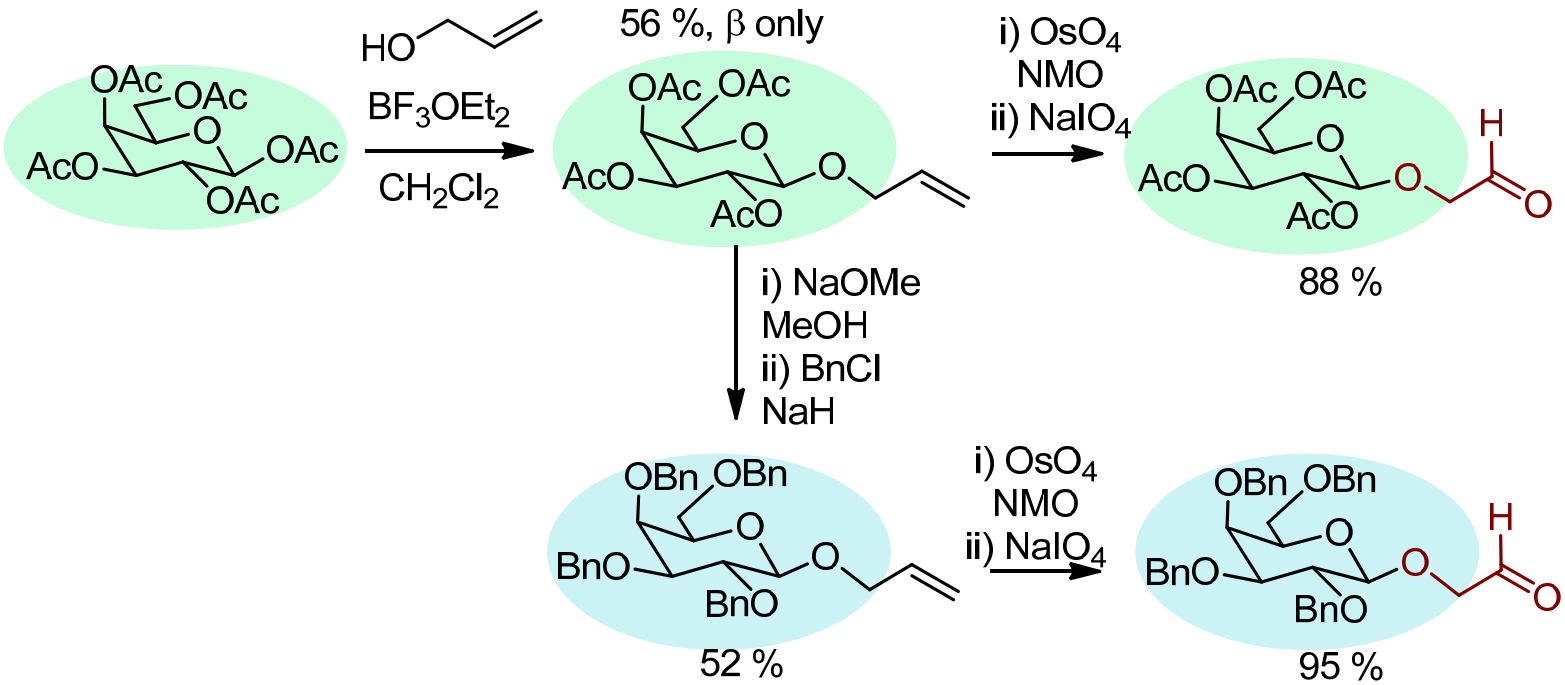
UNIVERSITY OF JYVÄSKYLÄ

Strategy: Hetero-Diels-Alder (HDA)



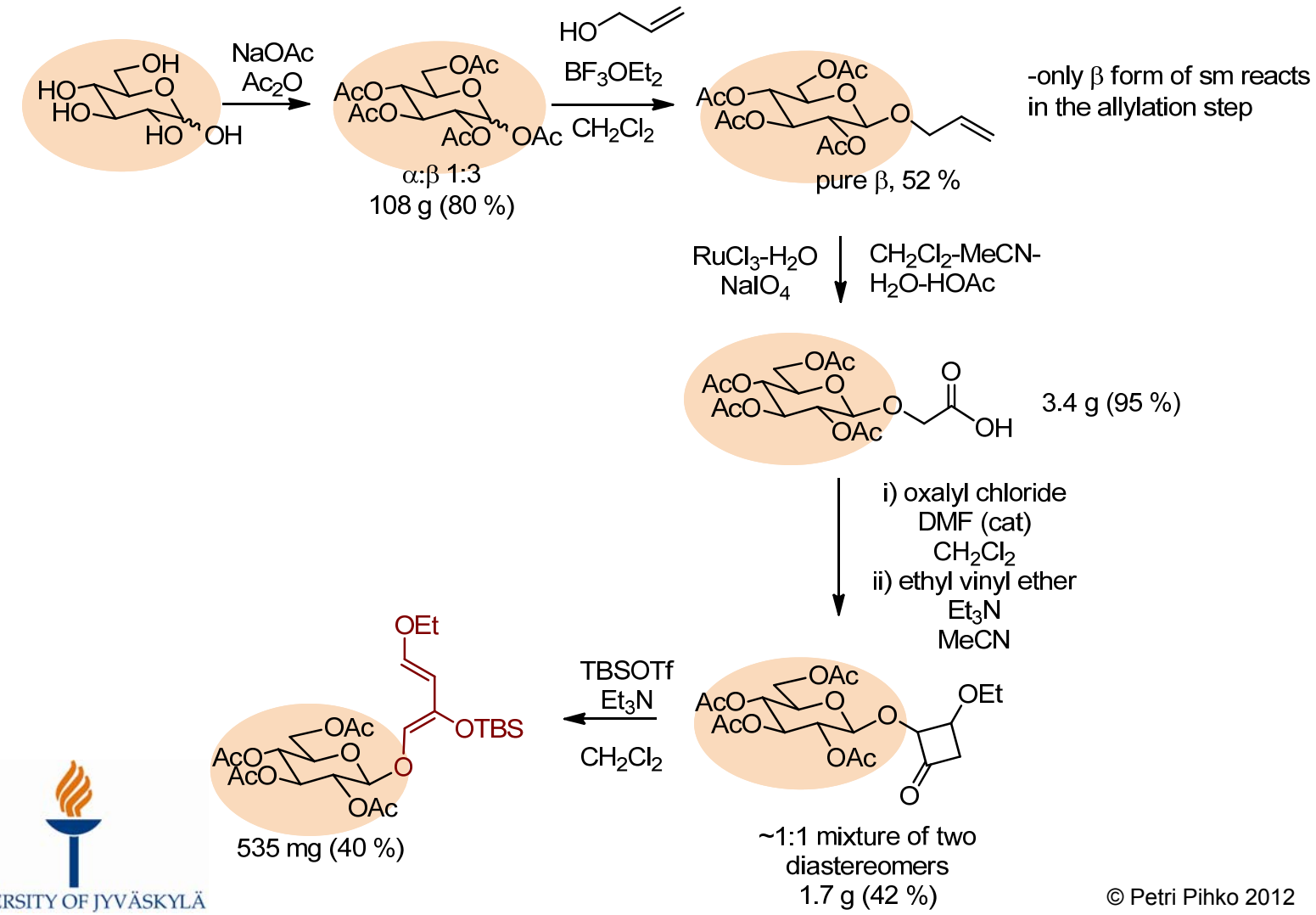
For reviews of HDA methods, see: a) Jørgensen, K. A. *Angew. Chem. Int. Ed.* **2000**, 39, 3558; b) Pellissier, H. *Tetrahedron* **2009**, *65*, 2839.
For precedents in the synthesis of monosaccharides, see e.g. Danishefsky, S.; Maring, C. *J. Am. Chem. Soc.* **1985**, *107*, 1269.

Synthesis of the Starting Materials

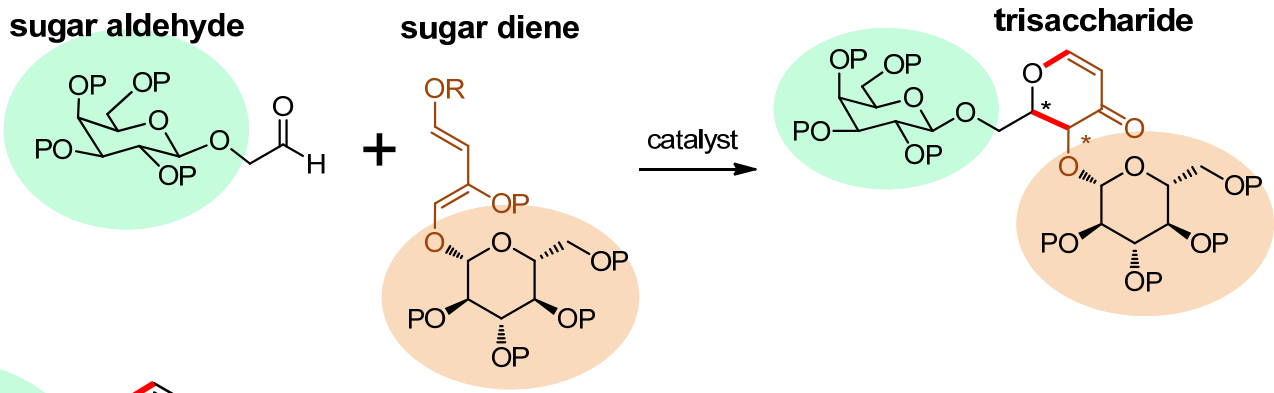


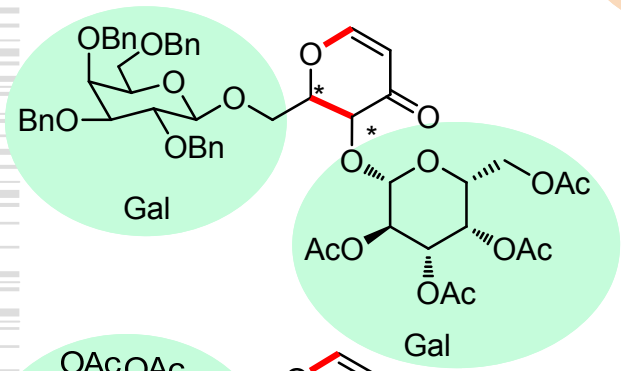
UNIVERSITY OF JYVÄSKYLÄ

Synthesis of the Diene Component



Trisaccharide synthesis

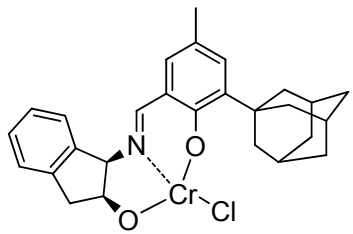




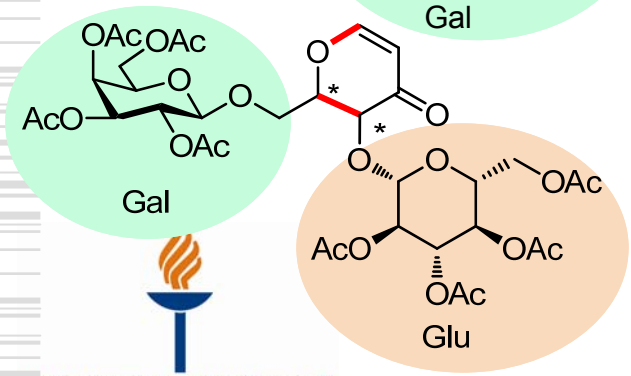
$MgBr_2 \cdot OEt_2$: 57%, 2 different diastereomers

***5 mol% (R,S)*-Cr-cat: 41%, single diastereomer**

***5 mol% (S,R)*-Cr-cat: 50%, single diastereomer**



(*R,S*)-Cr(inda) **2**



no cat: no reaction

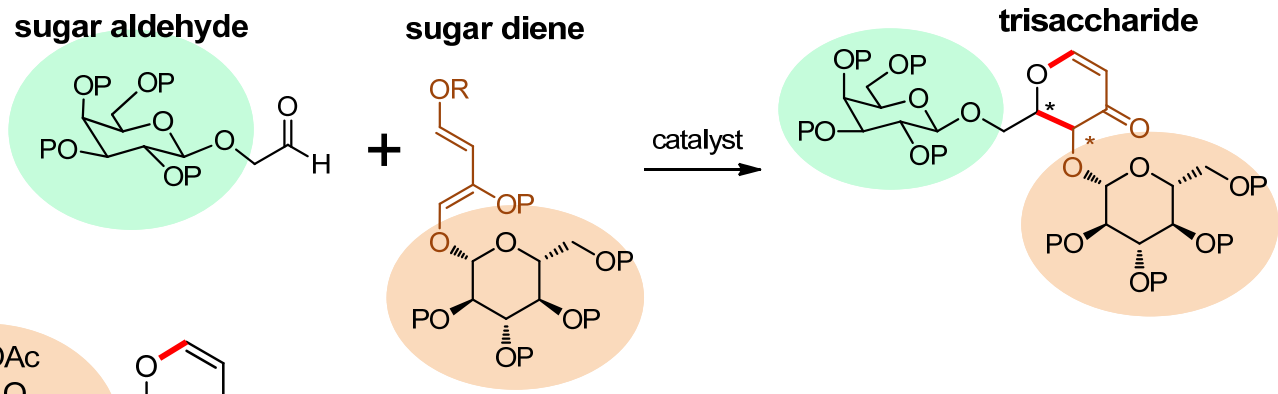
$MgBr_2 \cdot OEt_2$: 47%, 3 different diastereomers

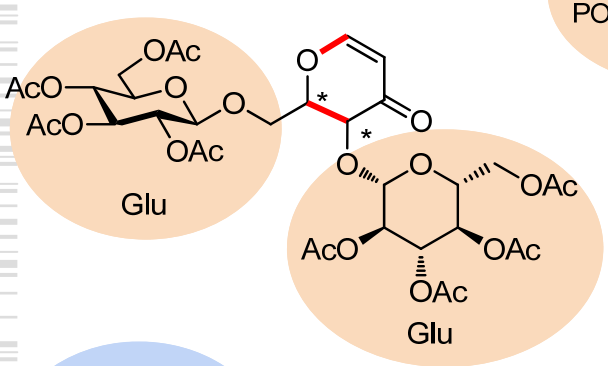
***5 mol% (R,S)*-Cr-cat: 61%, single diastereomer**

***5 mol% (S,R)*-Cr-cat: 68%, single diastereomer**

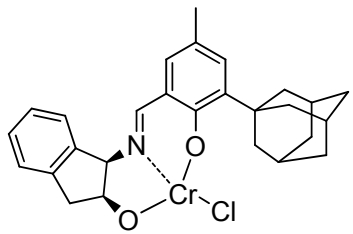
UNIVERSITY OF JYVÄSKYLÄ

Trisaccharide synthesis 2

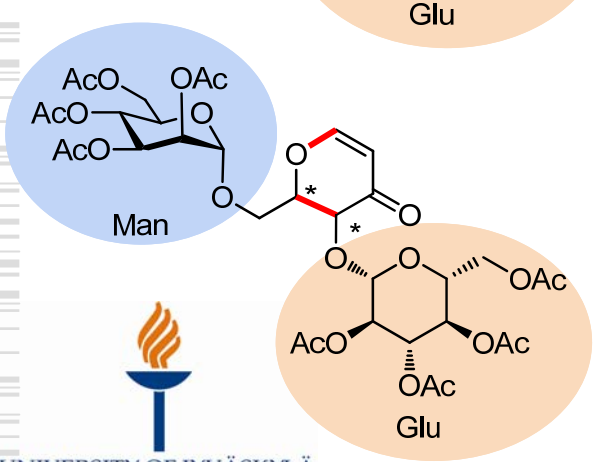




15 mol% (R,S)-Cr-cat: 51%, single diastereomer
5 mol% (S,R)-Cr-cat: 31%, single diastereomer



(*R,S*)-Cr(inda) **2**



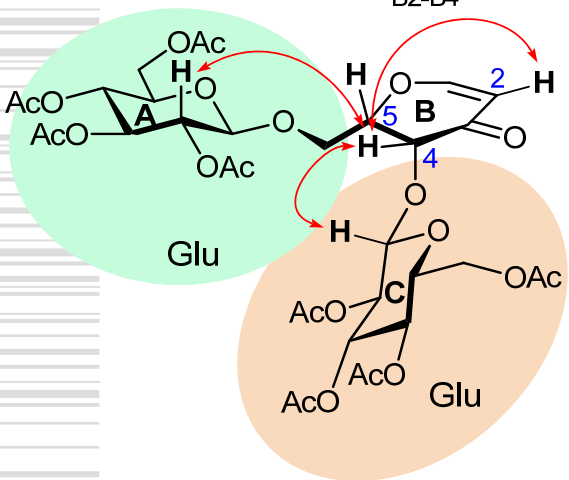
5 mol% (R,S)-Cr-cat: 48%, 92:4:4 ratio of diastereomers
5 mol% (S,R)-Cr-cat: 31%, mixture of diastereomers

UNIVERSITY OF JYVÄSKYLÄ

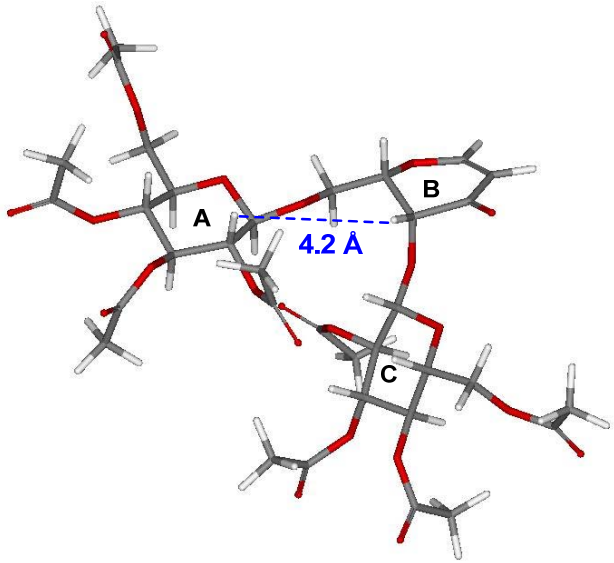
Relative Stereochemistry

NMR studies

J_{B4-B5} = 2.7 Hz
J_{B2-B4} = 1.6 Hz (W coupling)



Conformational analysis



- 1097 conformers generated by Monte Carlo search (MMFFs force field)
- 10 lowest energy conformers of each cluster optimized by DFT B3LYP/6-311G**

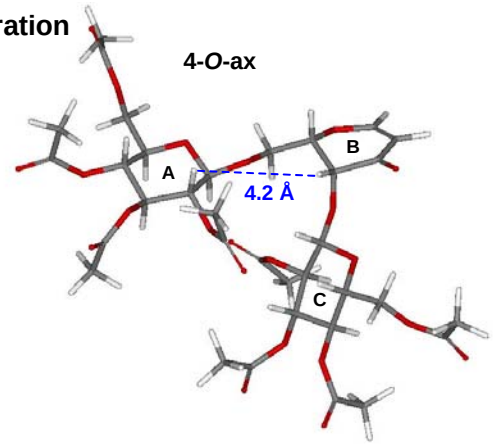
- The relative configuration of the B ring is *cis*
- NMR studies and conformational analysis appears to support the unnatural L configuration of the newly generated B ring
- The coupling constants of the B4 proton in all other products were virtually identical

UNIVERSITY OF JYVÄSKYLÄ

Further evidence for the proposed relative stereochemistry
from the solution structures

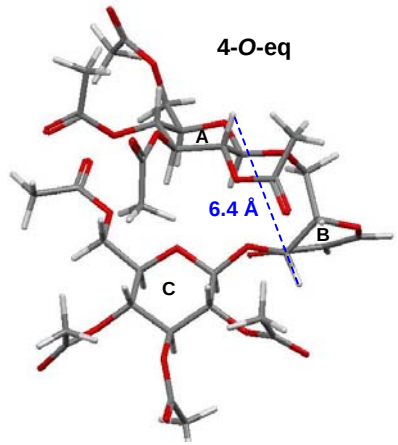
L-configuration

4-*O*-ax



Gas Phase Energy (kJ/mol): 0.0
Solution Phase Energy (kJ/mol): 0.0

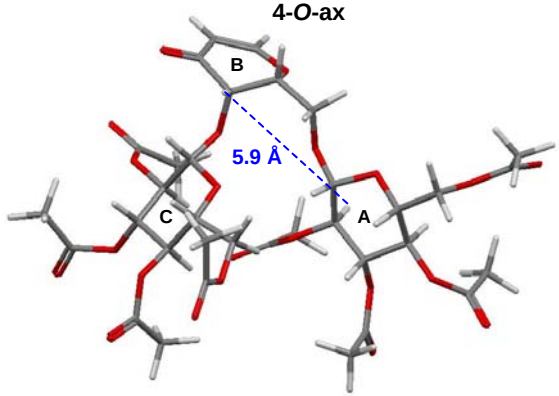
4-*O*-eq



Gas Phase Energy (kJ/mol): 20.7
Solution Phase Energy (kJ/mol): 37.1

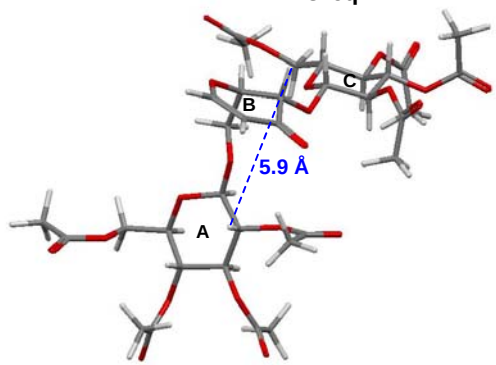
D-configuration

4-*O*-ax



Gas Phase Energy (kJ/mol): 53.3
Solution Phase Energy (kJ/mol): 45.9

4-*O*-eq



Gas Phase Energy (kJ/mol): 33.7
Solution Phase Energy (kJ/mol): 24.5

UNIVERSITY OF JYVÄSKYLÄ

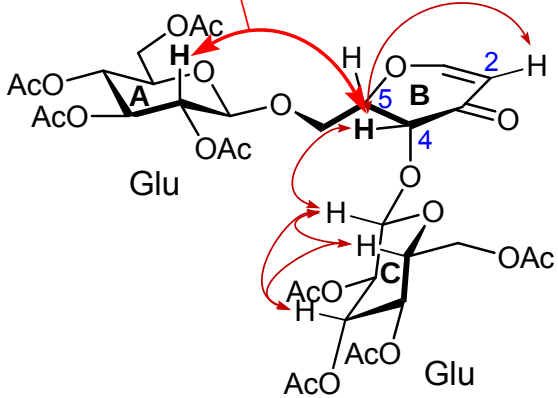
The piece that fits

NMR

key NOE - observed in both 2D NOESY and ^{1}H DPFGSE NOE

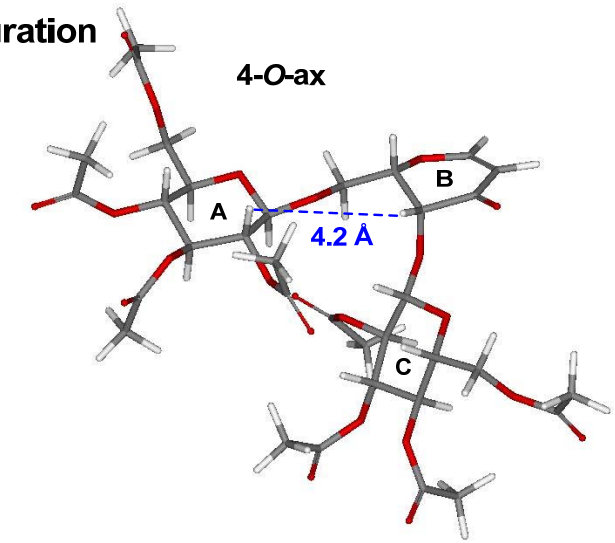
$J_{B4\text{-}B5}$ = 1.5 Hz

$J_{B2\text{-}B4}$ = 1.6 Hz (W coupling)



DFT

L-configuration



Gas Phase Energy (kJ/mol): 0.0
Solution Phase Energy (kJ/mol): 0.0

UNIVERSITY OF JYVÄSKYLÄ