DTU Chemical Engineering
Department of Chemical and Biochemical Engineering

Property Model-Based Chemical Substitution and Chemical Formulation Design

Spardha Jhamb[a]

Supervisors: Xiaodong Liang[a] , Georgios Kontogeorgis[a], Amol Hukkerikar[b], Kim Dam Johansen[a] , Rafiqul Gani[a]

[a]KT Consortium, Department of Chemical and Biochemical Engineering, DTU
[b]Alfa Laval Copenhagen A/S, Maskinvej 3,2860 Søborg

Outline

- Motivation and Objectives
- Role of Property Models in Chemical Substitution and Chemical Formulation Design
  - Proposed Framework for Chemical Substitution
  - Framework for Model-based Chemical Formulation Design
- Group Contribution Models for Pure Component, Primary Property Prediction
- Examples employing the Framework for Chemical Substitution
  - Example 1: Single Molecule Product Substitution
  - Example 2: Substitution of an Additive from a Formulation
- Conclusion and Future Work

Motivation and Objectives

Motivation for Chemical Substitution: Why Substitute?

- Hindrance in Process Operability and/or Process Economics

- Non-compliance with the Regularity Standards

- The Desirable Properties of a Product are Not Achieved

- High Product Price

- Low Product Quality

=> Based on 'Reason for Substitition', the 'Substitution Problem' is categorised into different types

- Objective: Propose a **systematic framework** for **property model-based chemical substitution and** apply it to various **product** and **process-centered** problems

Motivation and Objectives

Motivation for Model-based Chemical Product (Formulation) Design

- Chemical Product: Consumer-oriented, chemical-based, structured products
  - Single-species (molecule) products
  - Multiple-species products (Formulations, Blends)
  - Devices
- Chemical Product Design
  - Perceived to be the third chemical engineering paradigm
- Traditional Chemical Product Design
  - Based on trial and error approach
  - Inefficient: time and resource consuming
  - Final, feasible product design maybe more due to luck than the ability of the designer

=> Speed-up the Design Procedure and Spare the Resources

Objective: Apply the framework for model-based chemical formulation design (by Conte *et. al*) **to design a paint formulation for varied applications**

Role of Property Models in Chemical Substitution and Chemical Formulation Design

Primary Properties: **M&G GC Method**

Secondary and Functional Properties: **Co-relations and Functions**

- Translation of needs to properties
- Screening of alternatives
- Verification of Product

Service Role

- Generation of feasible alternatives/ candidate molecules

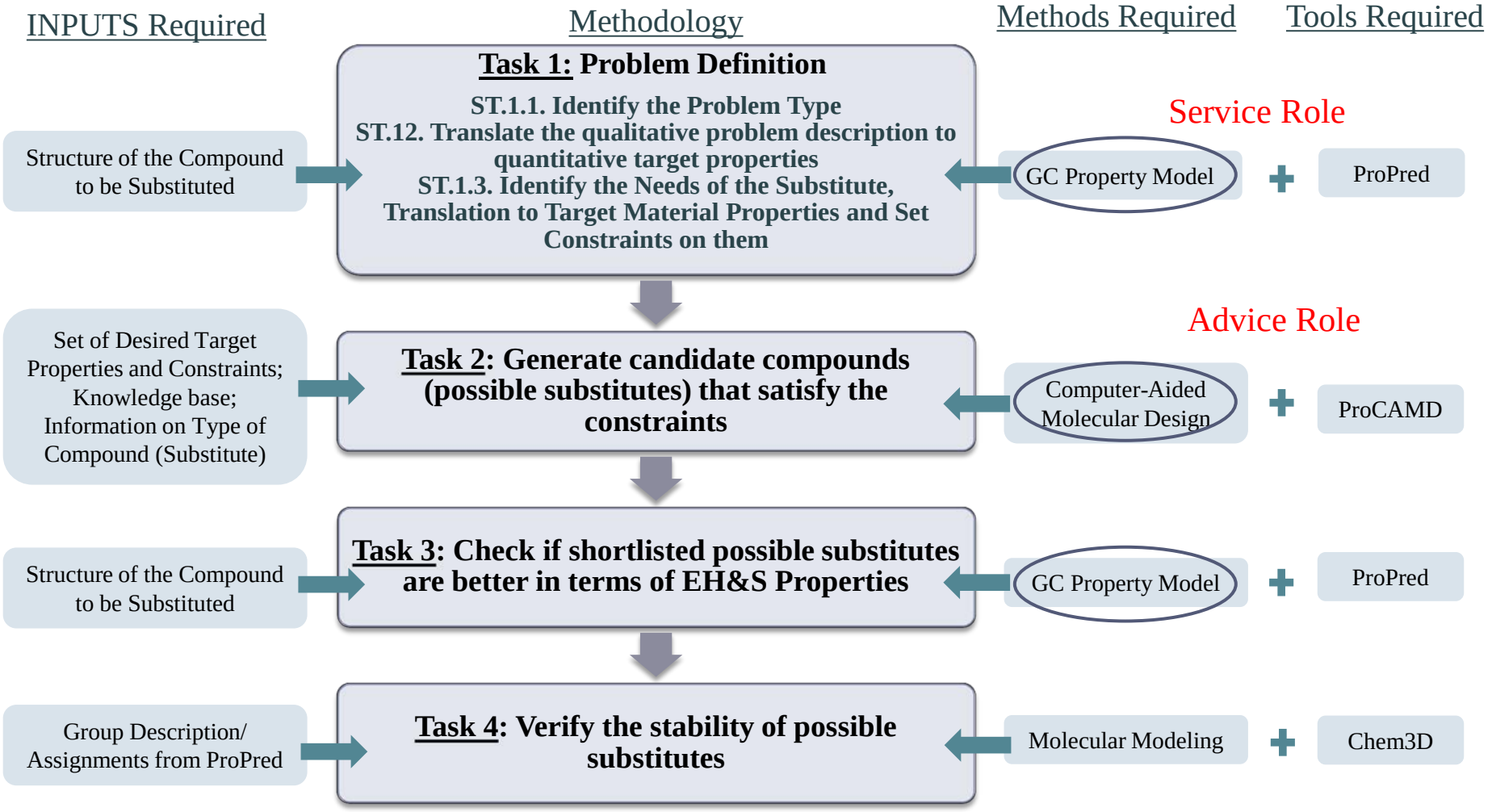
Advice Role

Mixture Properties: **KT-UNIFAC**

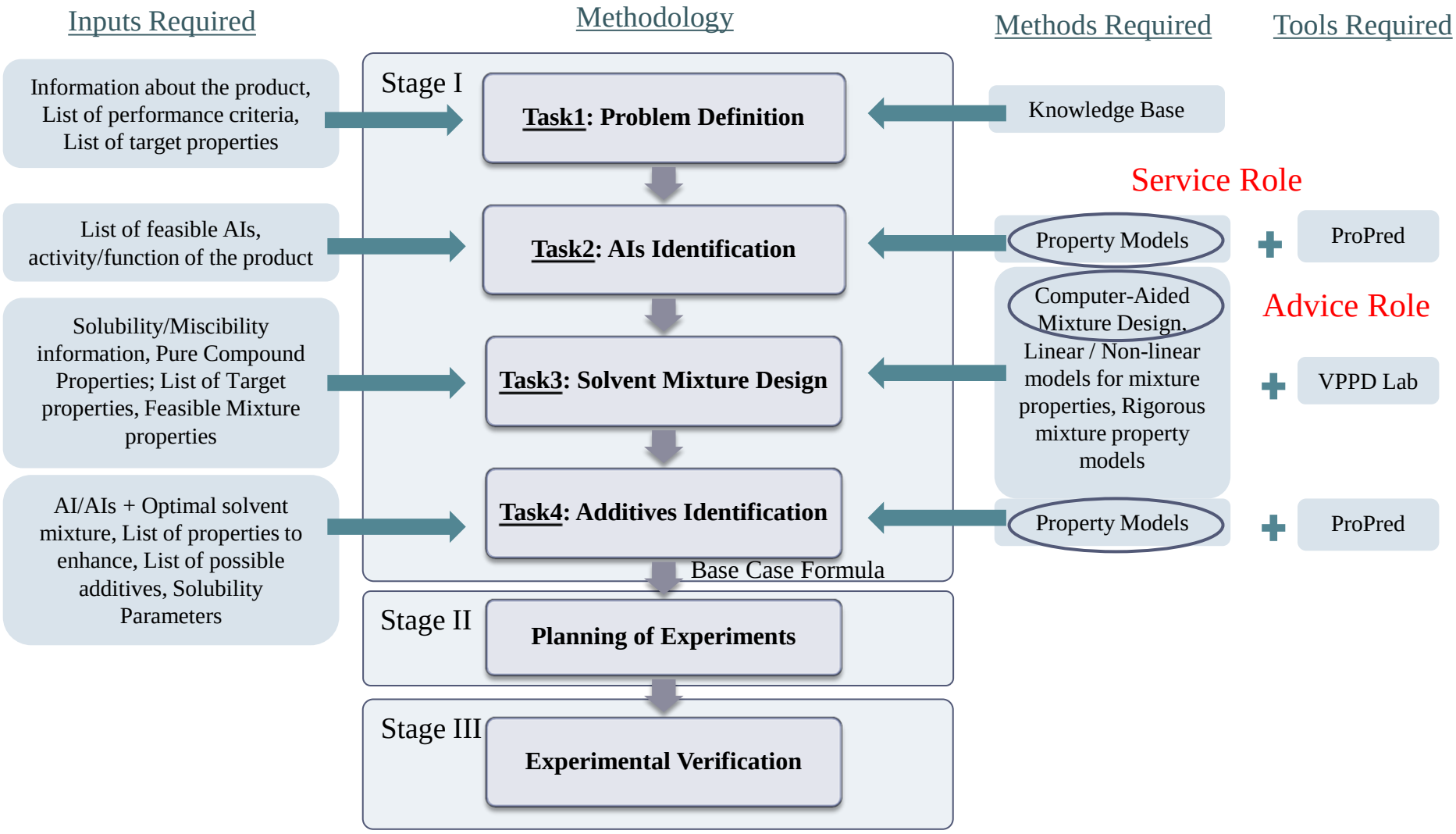
Polymer Models: **By Kontogeorgios**

Proposed Framework for Chemical Substitution

Reason for Substitution: Non-compliance with Regulatory Standards



Framework for Chemical Formulation Design (by Conte *et al.*)



Group Contribution Models for Pure Component, Primary Property Prediction

- Existing GC-Models for Pure Component Property Prediction by Marrero and Gani (MG) Method
  - Properties that can be predicted by 'GC-Model (Marrero and Gani)'
    - Organic compounds (C=3 - 60)

• T_m • T_b	• T_c • P_c • V_c	• ΔH_f^0 • $\Delta H_{fus}{}^0$ • $\Delta H_{vap}{}^0$ • ΔG_f

  - Properties that can be predicted by 'GC-Model (Hukkerikar *et al.*)'

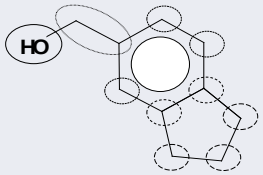
Environmental Impact on Air	Environmental Impact on Water	Environmental Impact on Earth (Soil)
• Oral rat LD50 • Permissible exposure limit (OSHA-TWA) • Photochemical oxidation potential • Global Warming Potential • Ozone Depletion Potential • Acidification Potential • Emission to urban air (carcinogenic and noncarcinogenic)	• Fathead minnow 96-h LC50 • Daphnia magna 48-h LC50 • Aqueous solubility • Bioconcentration Factor • Emission to Continental Fresh Water (carcinogenic and noncarcinogenic) • Emission to Continental Seawater (carcinogenic and noncarcinogenic)	• Emission to Continental Natural Soil (Carcinogenic and Noncarcinogenic) • Emission to Continental Agricultural Soil (Carcinogenic and Noncarcinogenic)

Group Contribution Models for Pure Component, Primary Property Prediction

- Marrero and Gani (MG) Method
  - Primary Property of Pure Compound
    - Function of structurally dependent parameters
    - Determined as a function of the frequency of the groups representing the pure compound and their contributions
  - Molecular Structure: Collection of three types of groups

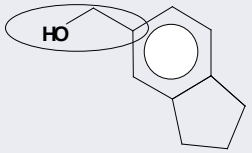
First Order Groups

- Simple and monofunctional groups
- Partially captures the proximity effects



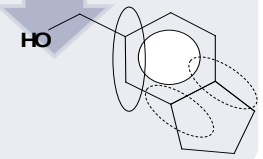
Second Order Groups

- Polyfunctional, polar/nonpolar compounds of medium size (C =3-6); aromatic or cycloaliphatic compounds with only one ring and several substituents
- Differentiation between isomers



Third Order Groups

- Complex heterocyclic and large (C=7-60), polyfunctional acyclic compounds
- For molecular fragments of compounds whose description is insufficient through the first and second level groups



$$f(X) = \sum_i N_i C_i + \sum_j M_j D_j + \sum_k O_k E_k$$

=> Computationally undemanding

Group Contribution Models for Pure Component, Primary Property Prediction

Extensive Use of Amino Acids in various Industrial Sectors

Food Industry
- Aspartyl-phenylalanine methyl ester: Low Calorie Sweetener
- L-Cysteine: Antioxidant for preservation of fruit juices

Cosmetics and personal care
- N-Acylglycinate: Non-ionic surfactant
- L-Glutamine: Neutralizer in soap (alternative acid)

Pharmaceutical
- Dihydroxyphenylalanine: Treatment of Parkinson's disease
- 5-hydroxytryptophan: Treatment of depression, and sleeping disorders

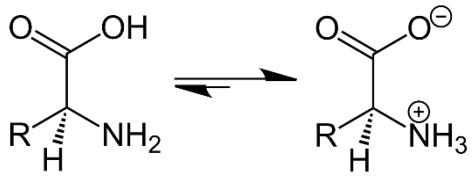
As a Solvent
- Potassium salt solution of L-Lysine: Solvent for CO_2 capture in flue gas

As a Biodegradable polymer
- Polyaspartate: Disposable diapers

- Scarcity of accurate predictive models for their physicochemical properties
- Insufficient description of these molecules through existing groups of GC models

Need for Introduction of New Structural Parameters for Amino Acids

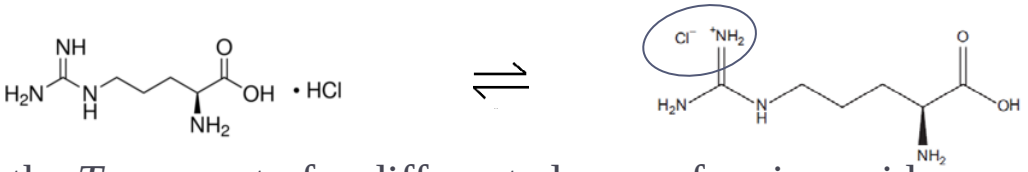
- To account for the complex interactions due to zwitterionic structures



Non-ionic form Zwitterionic form

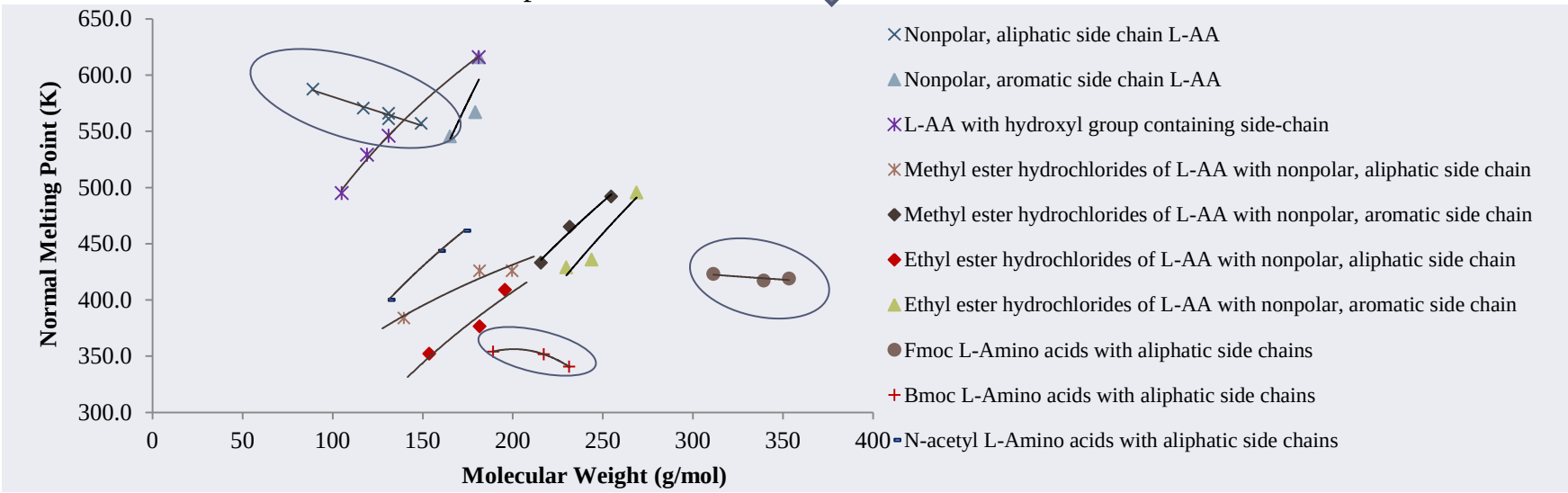
Group Contribution Models for Pure Component, Primary Property Prediction

- Need for Introduction of New Structural Parameters for Amino Acids
  - To account for ionic interactions in amino acid salts



  - Trend in the T_m property for different classes of amino acids
    - 'Nonpolar amino acids with aliphatic side chain'
    - 'Fmoc-L amino acids with aliphatic side chain'
    - 'Bmoc L-amino acids with aliphatic side chain'

Decrease in T_m with increase in MW



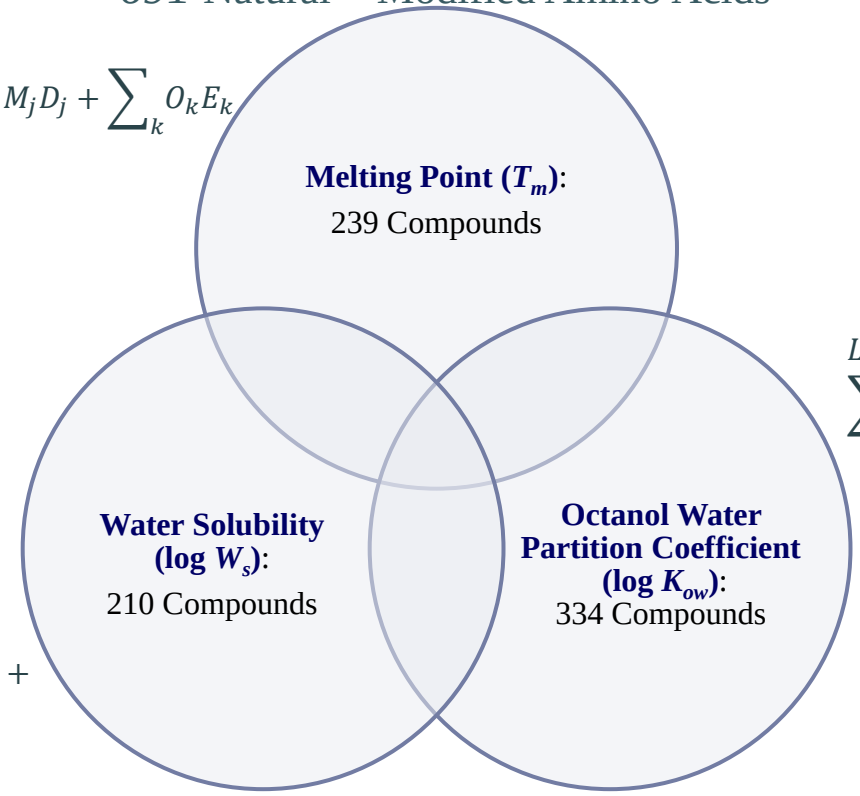
=> requires special treatment to achieve reliable predictions of such molecules

Group Contribution Models for Pure Component, Primary Property Prediction

- Experimental Property Dataset* and Models used for Amino Acid Property Prediction

691 Natural + Modified Amino Acids

$$exp\left(\frac{T_m}{T_{m0} + \alpha\, MW}\right) = \sum_i N_i C_i + \sum_j M_j D_j + \sum_k O_k E_k$$



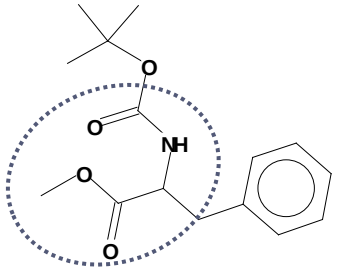
$$Log\ K_{ow} = A_{LogKow} + \sum_i N_i C_i + \sum_j M_j D_j + \sum_k O_k E_k$$

$$Log\ W_s = A_{LogWs} + B_{LogWs}\, MW + \sum_i N_i C_i + \sum_j M_j D_j + \sum_k O_k E_k$$

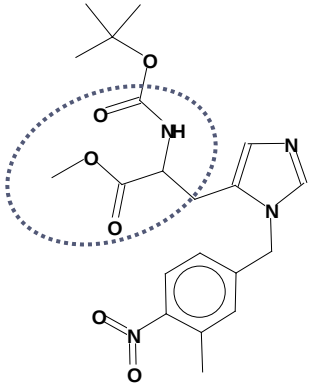
*Collected from Sigma Aldrich and database available at KT-Consortium

Group Contribution Models for Pure Component, Primary Property Prediction

- New Set of Groups to Define Amino Acid Structures (2 out of 36)
  1. O=C-NH-CH-COO-$(CH_n)_m$- (n in 0,…,2) (m in 0,1, 2…)

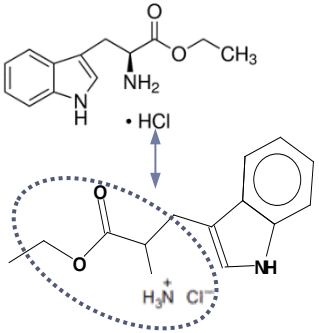


N-(tert-Butoxycarbonyl)-L-phenylalanine methyl ester

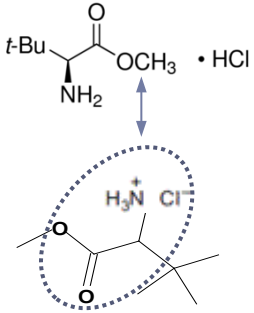


N-Boc-3-(3-methyl-4-nitrobenzyl)-L-histidine methyl ester

  2. $CH_n(NH_3^+Cl^-)$-COO-$(CH_2)_m$ *n* in (1,2); m in (0, 1, 2…)



L-Tryptophan ethyl ester hydrochloride

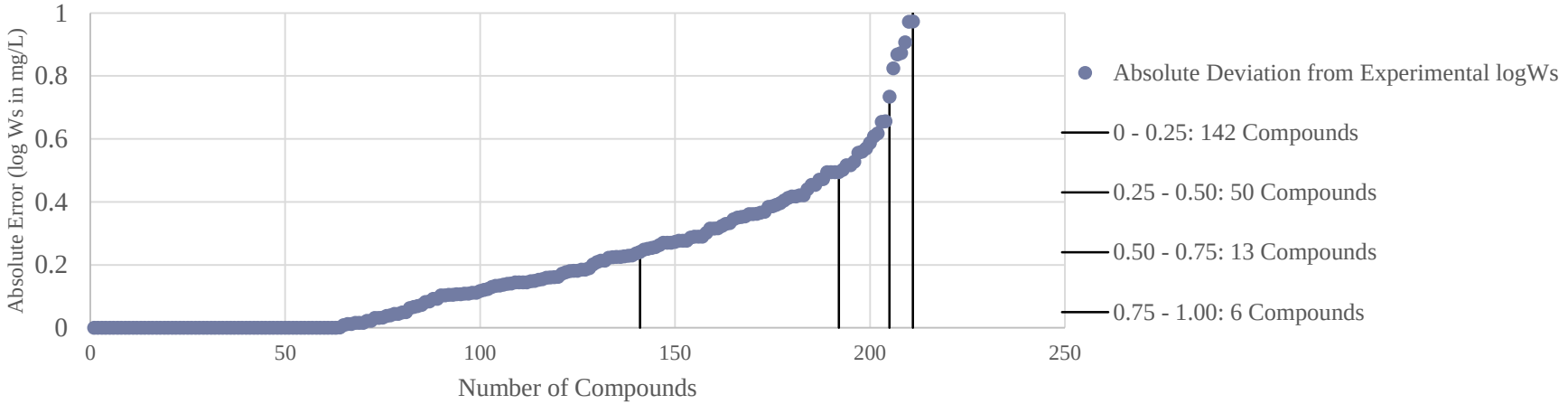


L-tert-Leucine methyl ester hydrochloride

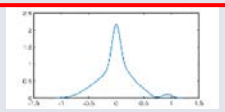
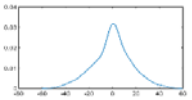
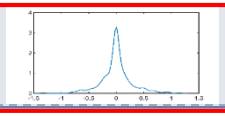
Group Contribution Models for Pure Component, Primary Property Prediction

- Model Performance Statistics for the Developed GC-Models

Absolute Error for Property 3: Melting Point (T_m) to Water Solubility (W_s) to Octanol Water Partition Coefficient (log K_{ow})



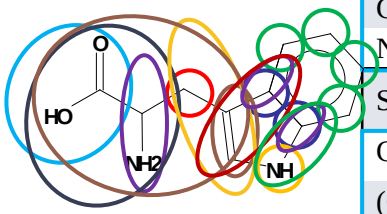
Summary of Statistics for all three properties

Property	N	R^2	Residual distribution plot	SD	AAE	AE_{max}	Main type of compounds giving large deviation
T_m	239	0.94		15.03	10.81	42.82	Ester hydrochlorides of amino acids with aliphatic side chain
Log W_s	210	0.94		0.29	0.19	0.97	Amino benzene sulfonic acids
Log K_{ow}	334	0.99		0.25	0.16	1.18	Aminobenzoic acids

Group Contribution Models for Pure Component, Primary Property Prediction

- Examples to compare model prediction with experimental values

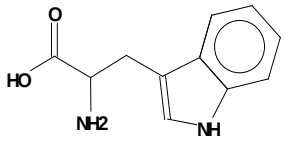
  Property 1: Melting Point (T_m) – Modified Amino Acid

Compound-1	First-order groups	Occurrences (N_i)	Contribution (C_i)
L-Tryptophan **(CAS No. 73-22-3)** 	CH_2	1	-0.0687
	aCH	4	0.4154
	aC fused with non-aromatic ring	2	-0.3506
	CHNH2	1	4.0020
	COOH	1	2.2226
	CH=C (cyclic)	1	0.8447
	NH (cyclic)	1	1.4318
	Second-order groups	Occurrences (M_j)	Contribution (D_j)
	$CH_m(NH_n)$-COOH ($m.n$ in 0..2)	1	3.3531
	$(CH_n=C)_{cyc}$-CH_2 (n in 0..2)	1	0.8447
$T_{m,experimental}$ = 555.7 *K*	Third-order groups	Occurrences (O_k)	Contribution (E_k)
	aC-$(CH_n=CH_m)_{cyc}$ (fused rings) ($n.m$ in 0..1)	1	0.6558
	aC-NH_{ncyc} (fused rings) (n in 0..1)	1	-0.7514
	AROM.FUSED[2]	1	0.3252
	$(CHn=C)_{cyc}$-CH_2-$CH(NH_2)$-COOH (n in 0..2)	1	1.0612
	$\exp\left(\frac{T_m}{T_{m0}+\alpha.MW}\right) = \sum_i N_i C_i + \sum_j M_j D_j + \sum_k N_k C_k \Rightarrow T_m =$ **548.8 *K*** (T_{m0} = 217.1 *K*; α=-0.0623)		

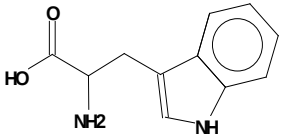
Group Contribution Models for Pure Component, Primary Property Prediction

- Examples to compare model prediction with experimental values

Property 2: Water Solubility (log W_s)

Compound-3	
L-Tryptophan **(CAS No. 73-22-3)**  ***LogW*$_{s,experimental}$ = 4.0569**	$LogW_s = A_{LogW_s} + B_{LogW_s} \times MW + \sum_i N_i C_i + \sum_j M_j D_j + \sum_k O_k E_k$ **= 4.0569** The calculated 95% confidence interval of the predicted propety value = 1.4492

Property 3: Octanol Water Partition Coefficient (log K_{ow})

Compound-4	
L-Tryptophan **(CAS No. 73-22-3)**  ***LogK*$_{ow,experimental}$ = - 1.06**	$LogK_{ow} = A_{LogK_{ow}} + \sum_i N_i C_i + \sum_j M_j D_j + \sum_k O_k E_k$ **= -1.06** The calculated 95% confidence interval of the predicted propety value = 0.8526

Examples employing the Framework for Chemical Substitution

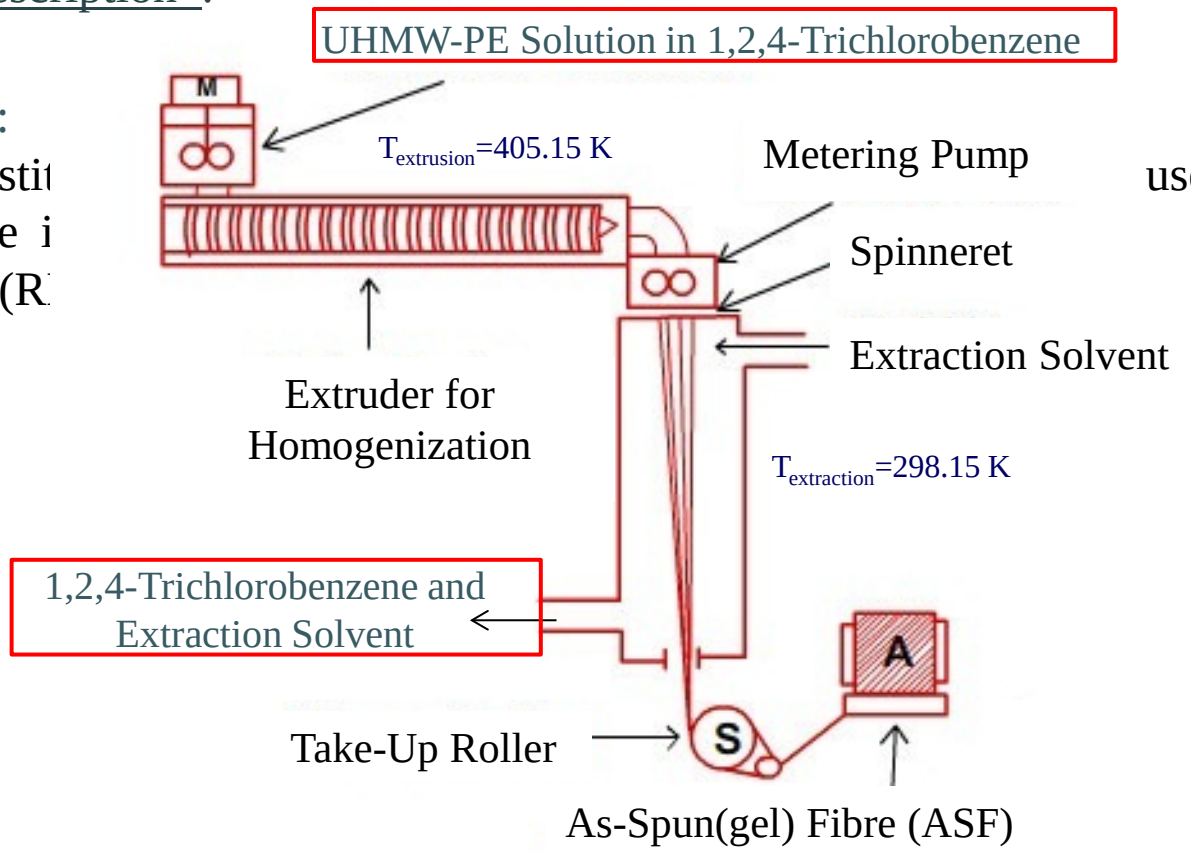
Examples employing the Framework for Chemical Substitution

Example 1: Single Molecule Product Substitution in a Process-Centered Problem

Problem description*:

Objective:

To substit … used in a process because i … ıe regulatory norms (R…



Gel-Spinning Manufacturing Process of UHMW-PE fibers

* Yufeng, Z., Changfa, X., Guangxia, J. and Shulin, A. (1999). Study on gel-spinning process of ultra-high molecular weight polyethylene. *Journal of Applied Polymer Science*, 74(3), pp.670-675.

Examples employing the Framework for Chemical Substitution

Example 1: Single Molecule Product Substitution in a Process-Centered Problem

Procedure:

Methodology	Methods Used	Tools Used	Inputs Required
Task 1: **Problem Definition** ST.1. Identification of Problem Type: ➢ Non-compliance with regulatory standards (i.e. 1,2,4-Trichlorobenzene appears in the Restricted Substances List compiled by ECHA)			
ST.2. Translate the qualitative problem type description to quantitative target property ➢ -log $(LC_{50})_{FM,1,2,4\text{-}Trichlorobenzene}$ = **4.63** (*Log mol/L*) => - log$(LC_{50})_{substitute}$ < 4.63 (*Log mol/L*)	MG GC Property Model	Pro-Pred	Structure of the compound to be substituted
ST.3. Identification of Needs, Translation to Target Material Properties and Setting Constraints on them 1.3.1. Pure Component Target Properties of the substitute			
Need 1: Dissolves UHMW-PE at temperatures 405.15 K ➢ Hildebrand Solubility Parameter of substitute at 405.15 K ($\delta_{substitute,\ 405.15\ K}$) ≈ Hildebrand Solubility Parameter of UHMW-PE at 405.15 K ($\delta_{UHMW\text{-}PE,\ 405.15\ K}$) ⇒ $\delta_{substitute,\ 405.15\ K} \approx 21\ MPa^{1/2}$ ⇒ $\delta_{substitute,\ 298.15\ K} < 21\ MPa^{1/2}$	Knowle-dge Base	--	Solubility Parameter of UHMW-PE at 405.15 K

Examples employing the Framework for Chemical Substitution

Example 1: Single Molecule Product Substitution in a Process-Centered Problem

Procedure:

	Methodology	Methods Used	Tools Used	Inputs Required
Task 1:	ST.3. Identification of Needs, Translation to Target Material Properties and Setting Constraints on them 1.3.1. Pure Component Target Properties of the Substitute			
	Need 2: Should allow easy crystallization of the polymer at 298.15 K ➢ Hildebrand Solubility Parameter of substitute at 298.15 K ($\delta_{substitute,\ 298.15\ K}$) ≠ Hildebrand Solubility Parameter of UHMW-PE at 298.15 K ($\delta_{UHMW\text{-}PE,\ 298.15\ K}$) ⇒ $\delta_{UHMW\text{-}PE,\ 298.15\ K} = 17.2\ MPa^{1/2}$ ⇒ $\delta_{substitute,\ 298.15\ K} > 17.2\ MPa^{1/2}$ OR $< 17.2\ MPa^{1/2}$	Knowle-dge Base	--	Solubility Parameter of UHMW-PE at 298.15 K
	Need 3: Should be removed by an extraction solvent (decalin) post-extrusion ➢ Hildebrand Solubility Parameter of substitute at 298.15 K ($\delta_{substitute,\ 298.15\ K}$) ≈ Hildebrand Solubility Parameter of extraction solvent (decalin) at 298.15 K ($\delta_{decalin,\ 298.15\ K}$) ⇒ $\delta_{decalin,\ 298.15\ K} \approx 17.5\ MPa^{1/2}$ ⇒ $\delta_{substitute,\ 298.15\ K} \approx 17.5\ MPa^{1/2}$	MG GC Property Model	Pro-Pred	Structure of the compound
	From Needs 1, 2 and 3, $17.2\ MPa^{1/2} < \delta_{substitute,\ 298.15\ K} < 21\ MPa^{1/2}$			

Examples employing the Framework for Chemical Substitution

Example 1: Single Molecule Product Substitution in a Process-Centered Problem

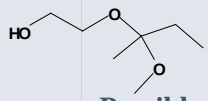
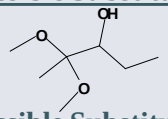
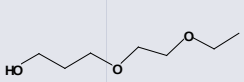
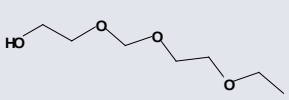
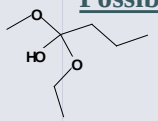
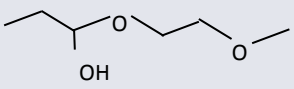
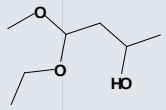
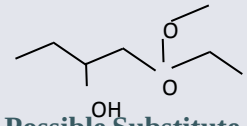
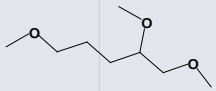
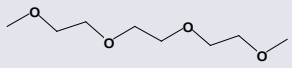
Procedure:

	Methodology	Methods Used	Tools Used	Inputs Required
Task 1:	ST.3. Identification of Needs, Translation to Target Material Properties and Setting Constraints on them 1.3.1. Pure Component Target Properties of the Substitute			
	Need 4: Should be in liquid phase at room as well as operating temperature of extrusion ➢ Melting Point of substitute ($MP_{substitute}$) < 285.15 K ➢ Boiling Point ($BP_{substitute}$) > 485.15 K	MG GC Property Model	Pro-Pred	Structure of the compound to be substituted
	1.3.2. Mixture Target Properties of the Substitute Need 5: Should be miscible with the extraction solvent ➢ Miscibility of the substitute with the extraction solvent (decalin): **Partial to Total Miscibility**			

Examples employing the Framework for Chemical Substitution

Example 1: Single Molecule Product Substitution in a Process-Centered Problem

Procedure:

Methodology	Methods Used	Tools Used	Inputs Required
Task 2: Generate candidate compounds (possible substitutes) satisfying the constraints ➢ Number of Compounds satisfying the Constraints considering Acyclic Alcohols, Ethers and Hydrocarbons = **10** **Possible Substitute-1**  δ = 20.90 $MPa^{1/2}$ T_m = 256.09 K T_b = 496.50 K **Possible Substitute-2**  δ = 20.89 $MPa^{1/2}$ T_m = 265.08 K T_b = 495.80 K **Possible Substitute-3**  δ = 20.68 $MPa^{1/2}$ T_m = 235.65 K T_b = 490.85 K **Possible Substitute-4**  δ = 20.67 $MPa^{1/2}$ T_m = 263.06 K T_b = 503.45 K **Possible Substitute-5**  δ = 20.66 $MPa^{1/2}$ T_m = 262.85 K T_b = 490.13 K **Possible Substitute-6**  δ = 20.64 $MPa^{1/2}$ T_m = 275.44 K T_b = 493.78 K **Possible Substitute-7**  δ = 20.50 $MPa^{1/2}$ T_m = 247.05 K T_b = 490.06 K **Possible Substitute-8**  δ = 20.25 $MPa^{1/2}$ T_m = 268.02 K T_b = 487.26 K **Possible Substitute-9**  δ = 17.87 $MPa^{1/2}$ T_m = 261.35 K T_b = 491.03 K **Possible Substitute-10**  δ = 17.76 $MPa^{1/2}$ T_m = 267.99 K T_b = 498.17 K	Computer-Aided Molecular Design	ProCAMD	Set of Desired Target properties and Constraints; Knowledge base; Information on Type of Compound (Substitute)

=> Only pure component target properties corresponding to Needs 1-4 have been considered here

Examples employing the Framework for Chemical Substitution

Example 1: Single Molecule Product Substitution in a Process-Centered Problem

Procedure:

	Methodology	Methods Used	Tools Used	Inputs Required
Task 3:	Check if shortlisted possible substitutes are better in terms of EH&S Properties ➢ 0.94 *Log mol/L* < -log $(LC_{50})_{FM,substitute-1,\ldots,10}$ < 2.40 *Log mol/L*	GC Property Model	ProPred	Structure of the possible substitute
Task 4:	Verify the stability of possible substitutes	Molecular modeling	Chem3D	Group Description/ Assignments from ProPred

Future Work:

- To Check if Mixture Target Properties of the Substitute are met and select the final substitute

  i.e. Need 5: Should be miscible with the extraction solvent

  - Miscibility of the substitute with the extraction solvent (decalin): **Partial to Total Miscibility**

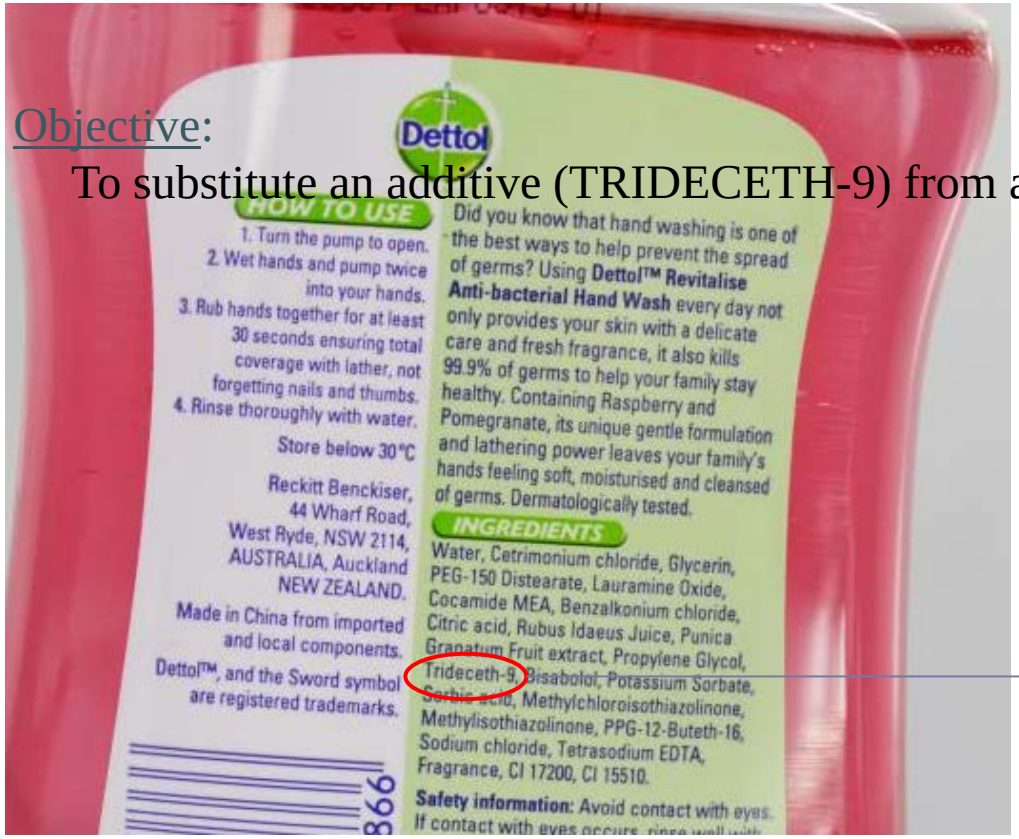
Examples employing the Framework for Chemical Substitution

Example 2: Substitution of Additive from a Chemical Product Formulation

Problem description:

Objective:

To substitute an additive (TRIDECETH-9) from a liquid soap formulation



Liquid Soap Formulation

Emulsifying agent:
TRIDECETH-9

Examples employing the Framework for Chemical Substitution

Example 2: Substitution of Additive from a Chemical Product Formulation

Procedure:

	Methodology	Methods Used	Tools Used	Inputs Required
Task 1:	**Problem Definition** ST.1. Identification of Problem Type: ➢ Non-compliance with regulatory standards (i.e. Classified as '*expected to be toxic*' in 'Environment Canada Domestic Substance List') ST.2. Translate the qualitative problem type description to quantitative target property ➢ $-\log (LC_{50})_{FM,Trideceth\text{-}9} = \mathbf{5.05}$ (Log mol/L) $\Rightarrow -\log (LC_{50})_{substitute} < 5.05$ (Log mol/L) ST.3. Identification of Needs, Translation to Target Material Properties and Setting Constraints on them 1.3.1. Target Properties of the substitute	MG GC Property Model	ProPred	Structure of the compound to be substituted
	Need 1: HLB Parameter of the substitute ($HLB_{substitute}$) should be $12 < HLB_{substitute} < 16$ to form a O/W emulsion and nearly equal to $HLB_{Trideceth\text{-}9}$ ➢ $HLB_{substitute} \approx HLB_{Trideceth\text{-}9}$ $\Rightarrow HLB_{substitute} \approx \mathbf{13.3}$	GC method by Davis-Rideal	--	Group parameters for HLB calculation; Structure of the compound

Examples employing the Framework for Chemical Substitution

Example 2: Substitution of Additive from a Chemical Product Formulation

Procedure:

	Methodology	Methods Used	Tools Used	Inputs Required
Task 2:	**Generate candidate compounds (possible substitutes) satisfying the constraints** ➢ A mixture of Oleic Acid (OA) and N-[3-lauryloxy-2-hydroxypropyl]-L-Arginine L-glutamate (C12HEA-Glu) has been reported* as a candidate with comparable emulsifying potency $\sum_i w_i HLB_i = 13.3 \Rightarrow wC_{12HEA\_Glu} = 31.25\ \%$	--	--	Set of desired Target properties and Constraints; Knowledge base; Information on Type of Compound (Substitute)
Task 3:	**Check if shortlisted possible substitute mixtures are better in terms of EH&S Properties** -log (LC_{50}) $_{C12HEA\_Glu}$ values are not available **However qualitative description of easy biodegradability of C12HEA-Glu has been reported*.**	GC Property Model	ProPred	Structure of the substitute mixture

*KOUCHI, J., TABOHASHI, T., YOKOYAMA, S., HARUSAWA, F., YAMAGUCHI, A., SAKAI, H. and ABE, M. (2001). Emulsifying Potency of New Amino Acid-Type Surfactant. (1). O/W Emulsions. Journal of Oleo Science, 50(11), pp.847-855.

Examples employing the Framework for Chemical Substitution

Example 2: Substitution of Additive from a Chemical Product Formulation

Future Work:

1. To complete this example we are working on:
   - Getting Log LC_{50} values for amino acids (possible also to develop a model)
   - Quantify the biodegradability property of the substitutes

2. To generate and test more candidates that satisfy the problem needs
   - Collect data for HLB parameters of more amino acids

Conclusion and Future Work

- The developed GC- property prediction models perform well
  - Define structures of natural as well as modified amino acids
  - Performance statistics are better in comparison to currently used property prediction models
  - Motivated by these results, current and future work is focussed on:
    - Modelling more properties for amino acids
    - Using them to develop chemical product formulations with amino acid as additive
- A framework for Chemical Substitution is proposed and used in two conceptual examples
  - The framework needs to be expanded to include problem types apart from non-compliance with regulatory norms
- Paint formulation for a wider range of applications needs to be developed using the framework by Conte *et. al*
  - Expansion of the database for AI (different pigments and binders)
  - Study of other target properties of a paint formulation for applications besides home decor
    - Example, ‘resistance to corrosion and abrasion’ incase of marine paints
  - Testing the applicability of rigorous thermodynamic models for property prediction

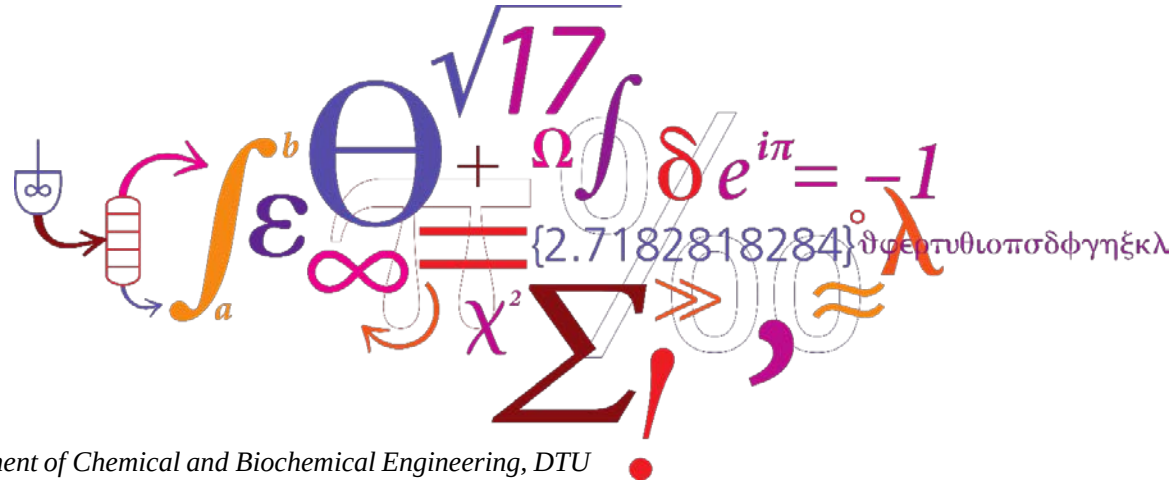
Thank-you for your attention!

Name: Spardha Jhamb

Email ID: spajha@kt.dtu.dk

Supervisors: Xiaodong Liang[a] , Georgios Kontogeorgis[a], Amol Hukkerikar[b], Kim Dam Johansen[a] , Rafiqul Gani[a]

Any Questions?



[a]*KT Consortium, Department of Chemical and Biochemical Engineering, DTU*
[b]*Alfa Laval Copenhagen A/S, Maskinvej 3,2860 Søborg*

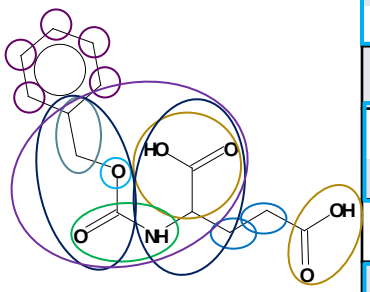
Extra Slides

Group Contribution Models for Pure Component, Primary Property Prediction

- Examples to compare model prediction with experimental values

Property 1: Melting Point (T_m) – Modified Amino Acid

Compound-1	First-order groups	Occurrences (N_i)	Contribution (C_i)
N-(Carbobenzyloxy)-L-glutamic acid (CAS No. 1155-62-0)	CH_2	2	-0.0687
	CH	1	-0.6259
	aCH	5	0.4154
	$aC\text{-}CH_2$	1	-0.1866
	COOH	2	2.2226
	NHCO except as above	1	0.5721
	-O-	1	-0.8632
	Second-order groups	Occurrences (M_j)	Contribution (D_j)
	$CH_m(NH_n)$-COOH ($m.n$ in 0..2)	1	3.3531
	$aC\text{-}CH_n$-OOC (n in 1..2)	1	-0.6218
	Third-order groups	Occurrences (O_k)	Contribution (E_k)
	$aC\text{-}CH_2\text{-}O\text{-}O{=}C\text{-}NH\text{-}CH_n\text{-}COOH$ (n in 1,2)	1	-1.0516



$T_{m,experimental}$ = 389.2 K

$$\exp\left(\frac{T_m}{T_{m0}+\alpha.MW}\right)=\sum_i N_iC_i+\sum_j M_jD_j+\sum_k N_kC_k \Rightarrow T_m = 387.3\ K$$

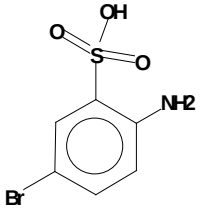
Comparison of other methods with the current work:

Method	Predicted Value (K)	Absolute Error (K)
Present work	384.6	1.8
MPBPWIN (EPI® suite)	462.5	75.2

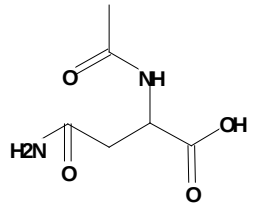
Group Contribution Models for Pure Component, Primary Property Prediction

- Examples to compare model prediction with experimental values

Property 2: Water Solubility (log W_s)

Compound-3	
2-amino 5-bromo benzenesulfonic acid (CAS No. 1576-59-6)  **$LogW_{s,experimental}$ = 2.7745**	$Log\ W_s = A_{LogWs} + B_{LogWs}\ MW + \sum_i N_i C_i + \sum_j M_j D_j + \sum_k O_k E_k$ = **3.3629** The calculated 95% confidence interval of predicted property value = 1.1521

Property 3: Octanol Water Partition Coefficient (log K_{ow})

Compound-4	
N_α-acetyl-L-Asparagine (CAS No. 4033-40-3)  **$LogK_{ow,experimental}$ = - 2.60**	$Log\ K_{ow} = A_{LogKow} + \sum_i N_i C_i + \sum_j M_j D_j + \sum_k O_k E_k$ = **-2.36** The calculated 95% confidence interval of predicted property value = 0.59