Marta Filipa Calvinho Rosa

Licenciatura em Ciências da Engenharia Química e Bioquímica

Liquid-Phase Peptide Synthesis Optimisation: Selection of Quenching Reagents and Peptide Solubility Estimation

Dissertação para obtenção do Grau de Mestre em

Engenharia Química e Bioquímica

Orientador: Professor Andrew G. Livingston, Imperial College London

Júri:

Presidente: Doutor Mário Fernando José Eusébio, Professor Auxiliar da Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa

Arguentes: Doutora Carla Alexandra Moreira Portugal, Investigadora Auxiliar, Unidade de Investigação LAQV da Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa

Vogais: Professor Doutor João Paulo Serejo Goulão Crespo, Professor Catedrático da Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa

Dezembro 2020

FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Marta Filipa Calvinho Rosa

Bachelor's degree in Science of Chemical and Biochemical Engineering

Liquid-Phase Peptide Synthesis Optimisation: Selection of Quenching Reagents and Peptide Solubility Estimation

A thesis submitted for the degree of Master of Chemical and Biochemical Engineering

Supervisor: Professor Andrew G. Livingston, Imperial College London

December 2020

Acknowledgements

First of all, I want to express my sincere gratitude to Professor Andrew G. Livingston for providing this opportunity. It was a real privilege to learn about the research world and develop my knowledge in contribution of peptide synthesis optimisation process.

I want also to address my acknowledgements to my co-supervisors Dr. Ludmila Peeva and Jet Yeo for providing guidance and feedback throughout this project. Their encouragement, patient, and support were essential to finish the project.

A special thanks to Professor Dimitar Peshev, that allowed me to use their *software* and for the availability that always presented to help me during the project.

To Professor João Paulo Crespo and Professor Mário Eusébio I want to express my special appreciation and to thank for they constant support, availability, and constructive suggestions.

Last, but not least, I want to recognize my family and friends and express my gratitude for all their unconditional support and encouragement, not only during this project, but all my academic journey. You are always there for me.

Abstract

With the increasing demand on peptides market, the peptide synthesis process must be kept upgraded. Therefore, the focus of this project is the optimisation of liquid-phase peptide synthesis. During peptide synthesis, the by-products and the amino acid excess must be kept controlled to not interfere with the main reactions and, consequentially, with the final product yield and purity. In order to achieve this, a quenching reagent is added to inactivate the amino acid in excess.

An optimisation study was performed for the quenching reaction, where it was found that 2 or 3 amino acids equivalents reacting with 2.1 quenching reagent equivalents allowed the best conditions for a controlled reaction. From there, with the conditions established, three reagents – piperidine, aniline, thiomalic acid – were used in the quenching reaction, where the quenching efficiency was evaluated by the quenching rate constant (k_q). The k_q values for piperidine, thiomalic acid and aniline we deduced to 2.962, 1.849, and 0.020 min^{-1} respectively. Piperidine was then determined as the best quenching reagent, followed by thiomalic acid and aniline, respectively.

The second problem of peptide synthesis that this project approached, was the determination of the peptide solubility. Currently, the solubility of peptides is measured experimentally, which is a time-consuming and wasteful process once it needs to be measured in different solvents, thus increasing the process cost.

The second part of the project investigated two theoretical approaches – Solubility Parameters and COSMO-RS – to predict the best and worst solvents for free peptides and peptides attached to a hub (wang and rink amide nanostar). DMSO was found the best solvent for most of the peptides, followed by NFM, which also presented itself as the best green solvent. The validation of the COSMO-RS method was tested and exhibited a RMSE of between 1.10 – 0.74 logarithmic units. Therefore, this model is presented as a tool to be implemented in determining the solubility of the amino acids.

Keywords: Peptide Synthesis; Process Optimisation; Quenching Reaction; Solubility.

x

Resumo

O aumento da procura de mercado por peptídeos obriga a constante otimização do seu processo de síntese. Assim, o foco deste projeto é a otimização da síntese de peptídeos em fase líquida. Durante o processo, os *by-products* e os aminoácidos em excesso devem ser controlados, de forma a não afetar o rendimento e a pureza do produto final. Para tal, é necessário proceder à sua inativação com a adição de um reagente.

A primeira parte deste projeto apresenta um estudo das condições ótimas para a reação de inativação, que determinou-se sendo 2 ou 3 equivalentes de aminoácido a reagir com 2.1 equivalentes de reagente. Com base nas condições estabelecidas, três reagentes – piperidina, anilina e ácido tiomálico - foram testados e avaliados pelas constante de inativação (k_q). Os valores de k_q obtidos foram 2.962, 1.849, e 0.020 min^{-1} respetivamente à piperidina, ácido tiomálico e anilina. Visto que a piperidina apresentou os melhores resultados, a mesma foi reconhecida como o melhor reagente para a inativação dos aminoácidos.

O segundo e principal problema da síntese de peptídeos abordado neste projeto é a determinação da solubilidade dos peptídeos. Atualmente, a solubilidade dos peptídeos é determinada experimentalmente, sendo um processo demorado e com elevado desperdício, uma vez que os peptídeos precisam de ser avaliados em diferentes solventes, tendo por consequência o aumento do custo do processo.

Este projeto apresenta duas abordagens teóricas – Parâmetros de Solubilidade e COSMO-RS – que preveem quais os melhores e piores solventes para os peptídeos e para os peptídeos ligados à âncora, quer por ligantes *wang* ou *rink amide*. No geral, o DMSO mostrou-se como o melhor solvente, seguido pelo NFM, que também se apresentou como o melhor solvente verde. A validade do método COSMO-RS foi testada e exibiu um erro médio quadrático de entre 1.10 – 0.74 unidades logarítmicas. Deste modo, este modelo apresenta-se como uma ferramenta a implementar na determinação da solubilidade dos aminoácidos.

Palavras-chave: Síntese de Peptídeos; Otimização do Processo; Inativação Aminoácidos; Solubilidade.

Table of Contents

List of Figures

List of Tables

Abbreviations

AA	Amino Acid
AE	Active Ester
Arg	Arginine
CAGR	Compound annual growth rate
COSMO-RS	COnductor like Screening MOdel for Real Solvents
DCC	N, N'-Dicyclohexylcarbodiimide
DIC	N, N'-Diisopropylcarbodiimide
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
Fmoc	9-fluorenylmethoxycarbonyl
Gly	Glycine
GSK	GlaxoSmithKline
GVL	γ-Valerolactone
HOBt	Hydroxybenzotriazole
HPLC	High-Performance Liquid Chromatography
HSPs	Hansen Solubility Parameters
LPPS	Liquid-Phase Peptide Synthesis
MeCN	Acetonitrile
MEPS	Membrane Enhanced Peptide Synthesis
MeTHF	2-Methyltetrahydrofuran
NBP	*N*-Butylpyrrolidine
NFM	*N*-Formylmorpholine
NMP	*N*-Methyl-2-pyrrolidine
NST	Nanostar Sieving Technologie
Pbf	2,2,4,6,7-Pentamethyldihydrobenzofuran-5-sulfonyl
PC	Propylene carbonate
Phe	Phenylalanine
QS	Quenched Specie
RMSE	Root Mean Square Error
S	Solubility
Ser	Serine
SPPS	Solid-Phase Peptide Synthesis
tBu	Tert-butyl
TFA	Trifluoracetic acid

THF	Tetrahydrofuran
TPSA	Topological polar surface area
Tyr	Tyrosine
UHLPC	Ultra-High-Performance Liquid Chromatography
UV	Ultraviolet

1. Introduction

1.1. Background and motivation

In recent years, the interest in peptides as drug candidates have increased. The main disease areas driving the use of peptide drugs are metabolic disease and oncology. Despite peptide (chains up to fifty amino acids) present low oral bioavailability and propensity to be rapidly metabolized, pharmaceutical companies have been developing new strategies to improve productivity, reduce the metabolism of the peptide, and working on alternative routes of administration, proving the value of peptides as therapeutics. A SWOT analysis of peptides as therapeutics is presented in Figure 1.1[1]. Nowadays, 60 to 70 peptide drugs have been approved in the USA, Europe, and Japan, more than 150 are in active clinical development and 260 having been tested in clinical trials[2]. This translates into a peptide therapeutics market of US $ 25.6 billion in 2019[3], which is expected to grow with a CAGR of nearly 8.9%[4] over 2020 – 2025 and worth US$ 47 billion by 2025[5]. One of the major factors for that market growth is the increasing metabolic disorders and cancers, as well as the demand for efficient and low-cost drugs and rising investments in research and development of new therapeutics. Also, the technological advancements in the peptide manufacturing process that have been done helped to boost the market growth.

The production of the synthetic therapeutic peptide has become possible for the industry with the development of the common synthesis methods: solid-phase peptide synthesis (SPPS) and liquid-phase peptide synthesis (LPPS). Both methods are explained in detail in Chapter 2, section 2.3. Over the years, several improvements have been developed and applied in those, however, still some steps can be optimised. The next big optimisation is the total processes becoming more environmentally friendly. The synthesis of peptides is one of the most wasteful and least green chemical processes, once most of that solvents used were classified as hazardous materials by several guides for greener chemistry.

A report accomplished by the American Chemical Society Green Chemistry Institute Pharmaceutical Roundtable (ACS GCI PR) concluded that solvents constitute 56% of the material used to make active pharmaceutical ingredient (API)[6]. Peptide synthesis is not an exception, since a huge amount of solvents is required during preparation and purification steps. Reported by MacMillan *et al.* 83% of amide bond formation employed dichloromethane (DCM) (36%) or N,N-dimethylformamide (DMF) (47%) as solvents, and only 0.04% employed greener solvents[7]. In SPPS, the issues are amplified with the excesses of reagents used to wash the resins. N-

methyl-2-pyrrolidine (NMP), is also one of the most used solvents in peptide synthesis, and such as DMF, present a big concern under EU Registration, Evaluation and Authorisation of Chemicals (REACh) for known to be reprotoxic and cause serious environmental problems when present in the waste. Alternatives as acetonitrile and tetrahydrofuran (THF) have been used, but neither is considered to be green solvent[8].

Strengths (S)	Weaknesses (W)
o Good efficacy, safety, and tolerability o High selectivity and potency o Predictable metabolism o Shorter time to market o Standard synthetic protocols	o Chemically and physically instable o Prone to hydrolysis and oxidation o Short half-life and fast elimination o Low orally bioavailable o Low membrane permeability
Opportunities (O)	Threats (T)
o Discovery of new peptides, including protein fragmentation o Focused libraries and optimised designed sequences o Formulation development o Alternative application routes besides oral o Multifunctional peptides and conjugates	o Immunogenicity o New advancements in genomics, proteomics, and personalized medicine o Significant number of patent expiries o Price and reimbursement environment o Increasing safety and efficacy requirements for novel drug

Figure 1.1 - Analysis of the strengths, weaknesses, opportunities, and threats (*SWOT*) of the use of peptides as therapeutics[1].

A solvent is evaluated employing the environmental health and safety (EHS) assessment method or by life-cycle assessment (LCA) method in order to determine their greenness. The focuses of these approaches are: '*'various solvent characteristics, namely sustainability, biodegradability, ease of incineration, recyclability, ozone depletion impact, volatile organic compounds (VOC) and their emission potential, impact on water and on air, risk to human health, and safety hazard'*'[9]. GlaxoSmithKline (GSK) present a solvent guide, several times updated with new data and methods of assessment, ranking the solvents relative to each other based on their inherent waste disposal, environmental, health and safety issues, by a scoring system resulting in a final colour assignment, Figure 1.2[10].
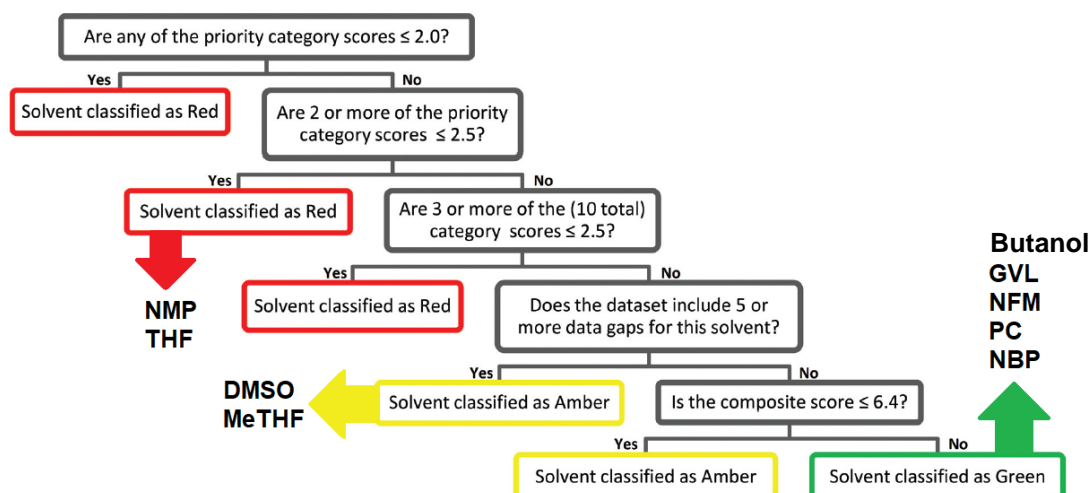
Figure 1.2 - GSK's decision tree for assignment of composite colours (*Adapted from Alder et al.*)[10].

Although relevant advance made in the field of greener solvents, green peptide chemistry is still a relatively small research field. Also, is important to refer that the greener approach, besides to be better for the environment, is more efficient in terms of timelines and cost[11].

1.2. Objectives and contributions

1.2.1. LPPS optimisation

With an increasing application of LPPS in industry is important to keep the processes up-dated and optimised, to contest the growing demand. This project fits in liquid-phase peptide synthesis optimisation and will address two-approaches for that. A project framework is presented in Figure 1.3 and in Figure 1.4.

The first part of the project consists in the optimisation of quenching reagent. In here, three different reagents will be performing the amino acid quenching. The results will be compared and the reagent that provides the best kinetics will be the chosen one.
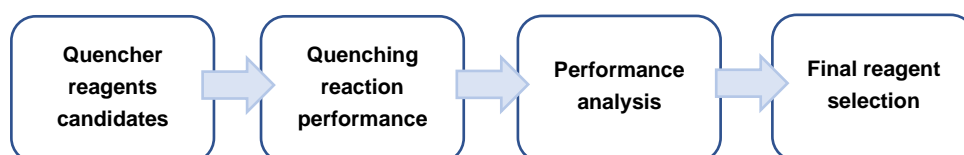
Figure 1.3 - Framework of the Part A of the project.

The second part of the project consists in the estimation of products solubility. Until now, the solubility of the peptides is found by experimental process, what takes more time and is material consuming, reflecting also in a more expensive process. Also, it is a wasteful process since different solvents need to be tested. If the estimates are validated, a big improvement in LPPS is done.
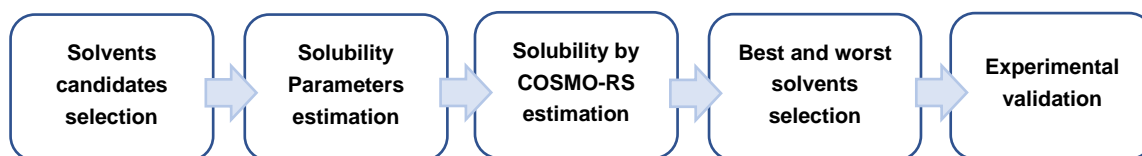
Figure 1.4 - Framework of the Part B of the project.

To improve the peptide synthesis in a more sustainable process requires improvement in different areas: classical synthetic methodologies and novel peptides synthesis methodologies[11]. This project can present a contribution to the sustainability challenge, namely in the area of classical synthetic methodologies and solvent optimisation, Figure 1.5. If the solubility estimation methods are validated, they can be used for a green solvents approach. Indeed, some of the solvents used in this project are green solvents.
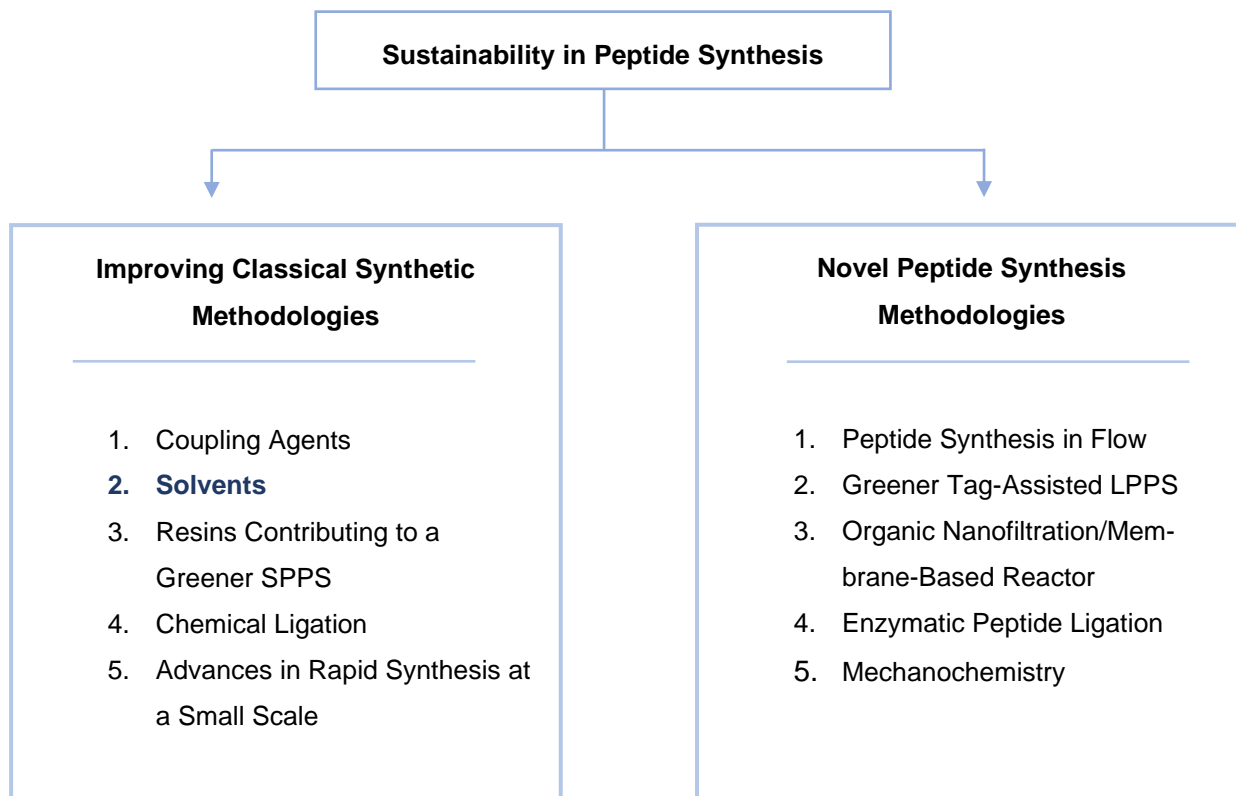
Figure 1.5 - Scheme of sustainability improving areas in peptide synthesis[11].

1.2.2. Project aim

The major project aim is the optimisation of LPPS and is expected to be accomplished through the completion of the following minor objectives:

1. Evaluate the quenching performance of each of the reagents.
2. Choose the reagent that provides the best kinetic for LPPS.
3. Estimation of solubility and solubility parameters for different amino acids and free peptides.
4. Estimation of solubility and solubility parameters for different peptides attached to the nanostar hub with a wang linker.
5. Estimation of solubility and solubility parameters for different peptides attached to the nanostar hub with a rink amide linker.
6. Choose the most and less suitable solvent for each peptide.
7. Validate the estimations methods via experimental solubility analysis.

Due to the current pandemic situation of COVID-19, it wasn't possible for me to realize the experimental solubility analysis for the peptides, so I counted on the help from Dr. Ludmila Peeva and from Jet Yeo for carrying out the experiments.

1.3. Thesis organization

This thesis is composed of 6 main chapters. Chapter 1 is the introductory chapter where are presented an overview and the aim of the project. Chapter 2 describes the previous work related to peptide synthesis, that is applied to this project. The contributions of the work on solubility and solubility parameters are also presented in this chapter. The materials and methods used in the realization of the project are provided in Chapter 3. In a row are Chapter 4 and Chapter 5, which discuss the result for quenching reagents optimisation and the results for validation of the methods of peptide solubility estimation, respectively. Lastly, Chapter 6 expose the final reflections and future considerations, as well as a final summary of the project.

2. Literature Review

2.1. Introduction

In the past years, peptide synthesis already runs a long way of new procedures and improvements. In this chapter, a detailed analysis of the peptide chemistry and the production strategies are presented, including the main and the side reactions that can affect the efficacy of the production.

Also, is presented previous studies about peptides synthesis optimisation, that are beneficial to this study, mainly related to solvents and solubility. To conclude, a description of the solubility determination approaches is also introduced in this chapter, since is one of the focus of this project.

2.2. Peptide Synthesis

A peptide is represented by a chain of amino acids linked to each other through peptide bonds, established between an amino group, N-terminal, of one amino acid and a carboxyl group of another, C-terminal, or vice-versa. Peptides have numerous functions being that factors as the side chain, R, of the amino acid, the sequence of the peptides, and any intra- or inter-chain connections are crucial to specify which function the peptide will have[12].
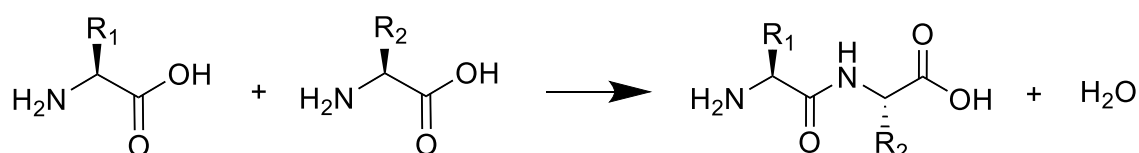
Figure 2.1 - Coupling reaction of two amino acids.

When a reaction occurs between a C-terminal and an N-terminal of two distinct amino acids, a peptide bond is created among the two amino acids and a water molecule is released, Figure 2.1.

To start the peptide synthesis, the first step is the **protection** of some functional groups in the amino acids, that means, the reactivity of the functional groups that would not be part into the peptide bond will be suppressed, avoiding unwanted reactions. Then, the so-called **coupling reaction** is executed, where occurs the formation of the peptide bond and the water molecule released. The following step is the release of the functional group that will complete the second peptide bond, without affecting the other groups' protection – **selective deprotection**. Until the

desired peptide chain is obtained, the last two operations are repeated. When the desired peptide is found, the **final deprotection** (or global deprotection) happens in one or two steps to remove the remain protecting groups[12]. The scheme of this general approach is represented in Figure 2.2.
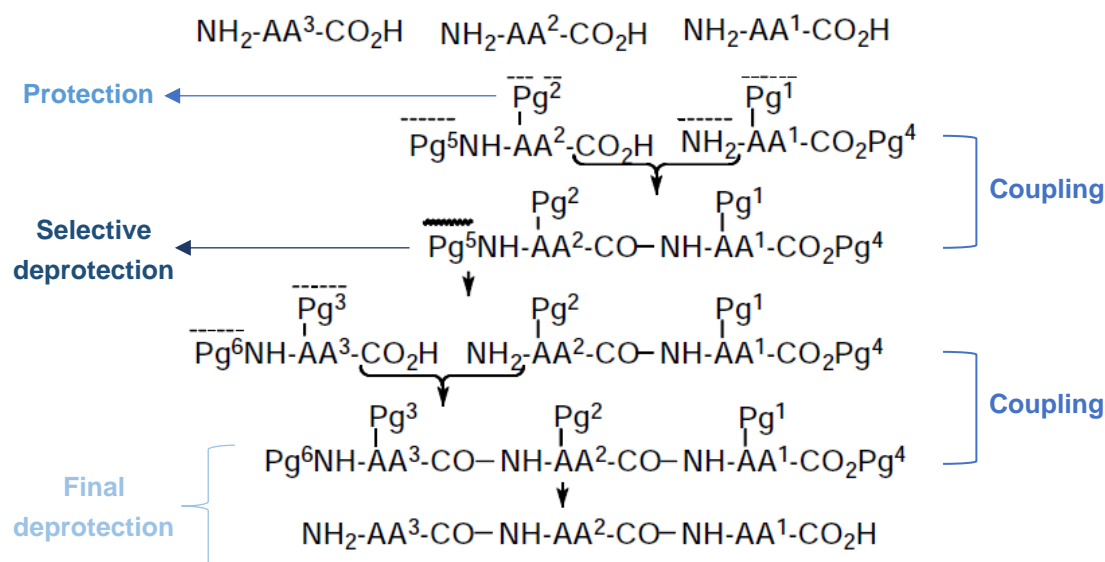
Figure 2.2 - Scheme of the general approach to peptide synthesis (*Adapted from Benoiton*)[12].

Although, the peptide synthesis can be performed in both directions, usually, it starts at C-terminus with selective deprotection at the amino group. When started at N-terminus, selective deprotection is at the C-terminus of the growing chain.

2.2.1. Activation and coupling reaction

In order to two amino acids form a peptide bond, the carboxyl group from those amino acids involved in the bond must be activated, by introducing an electron-withdrawing group Y. The activation of the carboxyl group may be carried in the presence or in the absence of the amino group, according to the necessity. The need to perform activation comes from that at ambient temperature, the carboxylic acids with amines would perform an acid-base reaction and obtained a salt. Despite heating could make the salts couple and form a peptide bond, it has negative effects on the functional side chain, on the structural transformation, and on racemization[12].

Thereby, a new single compound, the coupling reagent, needs to be added to the mixture of the two reactants for the activation and the coupling to happen. The concept of coupling reagents was presented by Sheehan and Hess in 1995, and introduce a carbodiimide compound, N,N'-dicyclohexylcarbodiimide (DCC). The carbodiimide method is the most popular for forming

peptide bonds and it was the method used in this study. The other methods are Aminium/uronium and phophonium salts and propanephosphonic acid anhydride. Figure 2.3 represent the carbodiimide method mechanism[5].
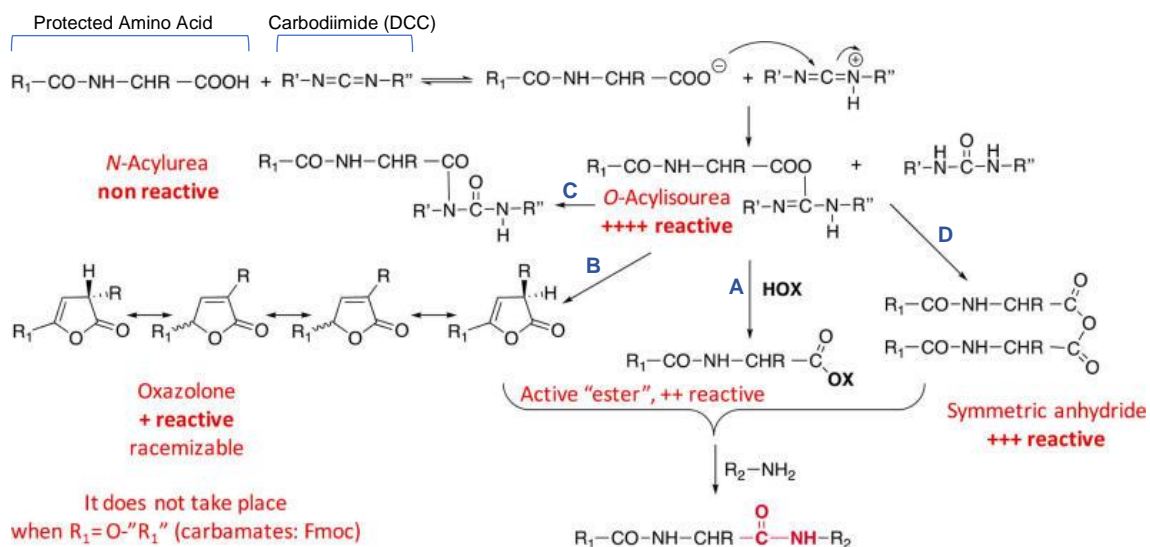
Figure 2.3 - Mechanism of carbodiimide-based activation of carboxylic acids (*Adapted from Albericio et al.)*[5].

The O-acylisourea, a very reactive active specie, is formed after adding the carbodiimide coupling reagent to the carboxyl group. Some activated forms are more stable than others. O-acylisourea with the aid of coupling additive (HOX) forms the active ester (AE) that is then aminolysed to give the peptide (path A). Three side reactions occurred that lower the coupling efficiency and product purity.

The formation of oxazolone specie is much likely to happen (path B) when the amino acid is in the form of an amide or carbamate moiety. Oxazolone is less reactive but it can undergo racemization. The tendency for path B is so strong, that special attention was devoted to trying minimizing the occurrence[5]. A totally inactive molecule, N-Acylurea, can also be formed by the rearrangement of the O-acylisourea, path C. The last competing intermolecular reaction, path D, occurs due to the presence of a second equivalent of the carboxylic acid that origins the symmetric anhydride. It is a highly reactive species and often leads to a double incorporation of the protected amino acid – a ''double hit''. This path is not known to occur[12].

It was found that by adding ''additives'' for carbodiimide-mediated reactions, the side reactions described above will be suppressed by converting the activated species into activated esters before they have time to undergo secondary reactions, what significantly improves the efficiency of coupling. The most common additives used are 1-hydroxybenzotriazole (HOBt)[5] and one of its derivate 1-hydroxy-7-azabenzotriazole (HOAt)[13], represented in Figure 2.4.
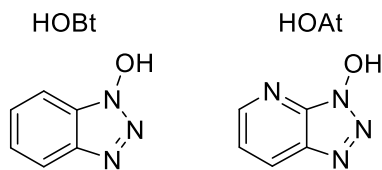
Figure 2.4 - Most common additives derivates used as coupling additives.

The three most popular carbodiimides are DCC, ethyl(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC), and diisopropylcarbodiimide (DIC). Each one was its unique characteristics. DCC generates a very bulky N,N'-dialylurea that is insoluble in organic solvents and it may interfere with the mixing. On the other hand, in DIC, both the reagent and the respective urea are soluble in organic solvents, thus there is no bulky precipitate to resist with. Despite the urea cannot be removed from an organic solution by aqueous extraction, it is soluble enough in the water that final traces can be removed from a precipitated peptide by washing. For EDC, both the reagent and the respective ureas are soluble in water, whereby is employed in amide-bond-forming reactions in partially aqueous mixtures[12].

2.2.2. Protectors and selective deprotection reaction

To avoid uncontrolled multiple coupling or coupling at the wrong terminals, it is necessary to proceed to the protection of the functional groups not involved in the peptide bond. That is achieved by adding another compound, the protecting group, to those functional groups. To be considered a good protector, the protecting group must be very soluble and removable, preferably with a mechanism without side reactions. The common protecting groups, Figure 2.5, used in peptide synthesis are derived from a very limited number of alcohols: Boc and Fmoc[12].
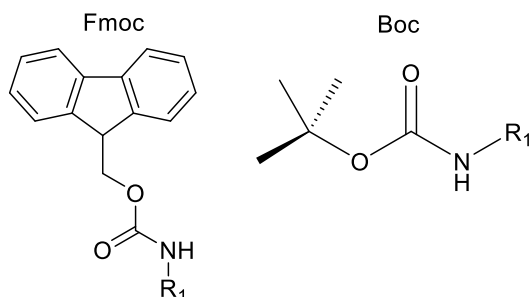
Figure 2.5 - Most common protectors used in peptide synthesis.

Boc is an abbreviation for *tert*-butoxycarbonyl and Fmoc for 9-fluorenylmethoxylcarbonyl. Both protector groups are stable, however, when compared Fmoc group is weakly attached to N-terminus of the peptide chain. So, the critical feature of each one is how it is removed. In Boc

approach, is required the use of hard acids as hydrofluoric acid (HF) or trifluoracetic acid (TFA), which can deteriorate the resultant peptides and it can result in acid-catalysed side reactions after repetitive deprotection reactions. It also affects the final deprotection. The Fmoc approach is then the most used since it allows deprotection by different mechanisms and does not require the use of corrosive acids[14].

After a coupling reaction, it is needed to deprotect the next functional groups that will be in the next coupling, with the removal of the protection group. Without the deprotection reaction, the next coupling reaction would not happen, and the peptide chain would not grow as desired. By other words, a poor deprotection efficiency, the yield and quality of peptide will decrease. The use of cyclic secondary amines (piperidine) is the most convenient method due to their nucleophilicity. The major problem reported by the use of piperidine is the formation of aspartimide[15], Figure 2.6.
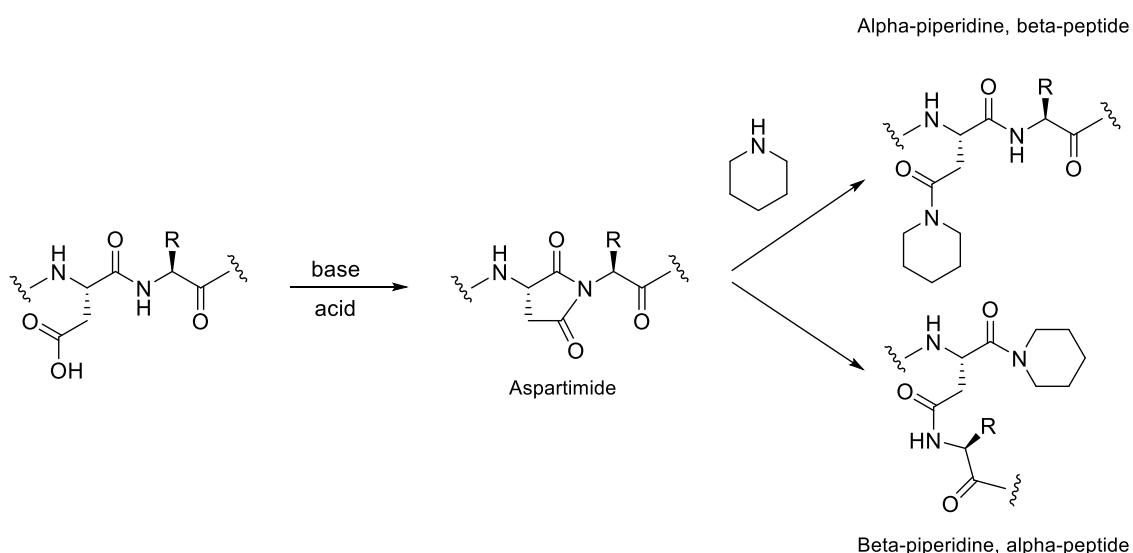
Figure 2.6 - Formation of aspartimide in the presence of piperidine[16].

The aspartimide reaction with the piperidine, leads to the formation of α– and β–piperidines, inhibiting the deprotection reaction. This side reaction is often in the presence of aspartic acid on the growing peptide chain[17]. As described above, also this side reaction can be suppressed by the use of additives, such as HOBt,

2.2.3. Final deprotection and cleavage

TFA is widely used to remove the final side chain protecting group. Mainly in SPPS, TFA it used because simultaneous allows the cleavage of the peptide from the resin. The cleavage cocktail is not an easy reaction, it can generate a succession of competitive reactions and lead to unwanted by-products[14]. In order to minimize and prevent the side reactions additional

substances (such as phenol and thioanisole), called scavengers, are added into the cleavage cocktail[17].

2.3. Peptide Synthesis Processes

The synthesis of peptides has been improved over the past years. Different strategies were adopted, being that the most conventional methods are the solid-phase peptide synthesis (SPPS) and the liquid-phase peptide synthesis (LPPS).

2.3.1. Solid-Phase Peptide Synthesis (SPPS)

SPPS is the leading strategy in peptide synthesis due to capacity to produce long peptide chains, up to 50 mer (e.g. AA-AA is dimer and AA-AA-AA-AA-AA is 5 mer). The growing chain was anchored on an insoluble polymer support, usually called resin, containing reactive sites. The separation is performed in one step, microfiltration (MF) washing, after each synthetic reaction to remove excess reactants and by-products, which allows shorter time cycles[18]. A schematic representation is illustrated in Figure 2.7.
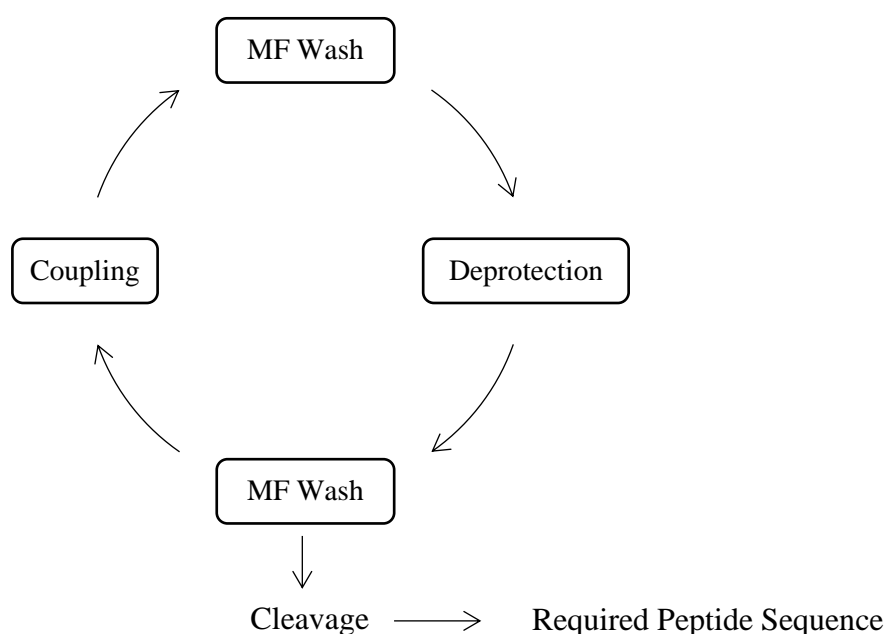
Figure 2.7 - Schematic representation of SPPS[19].

However, this strategy presents some weaknesses, as to produce coupling steps that may not be quantitative, that consequently create wrong amino acid sequences. Also, due to use solid support, it confers mass transfer limitations that slow down the kinetics of the synthesis reactions, meaning lower rate reactions and large excess of reagent[20].

2.3.2. Liquid-Phase Peptide Synthesis (LPPS)

LPPS contains a soluble and linear (or not) polymer, an anchor, serving as the C-terminal protecting group for the peptide which is to be synthesized. The synthesis reactions are all carried out under homogeneous conditions and occurs the dissolution in the mixture of the anchor bounded with the linker-peptide building block. After the reaction is complete, an anti-solvent is added to precipitate the peptide building block. Then, by microfiltration washing, the peptide product is removed[21]. Usually, it is used to produce short peptides. A schematic representation is illustrated in Figure 2.8.
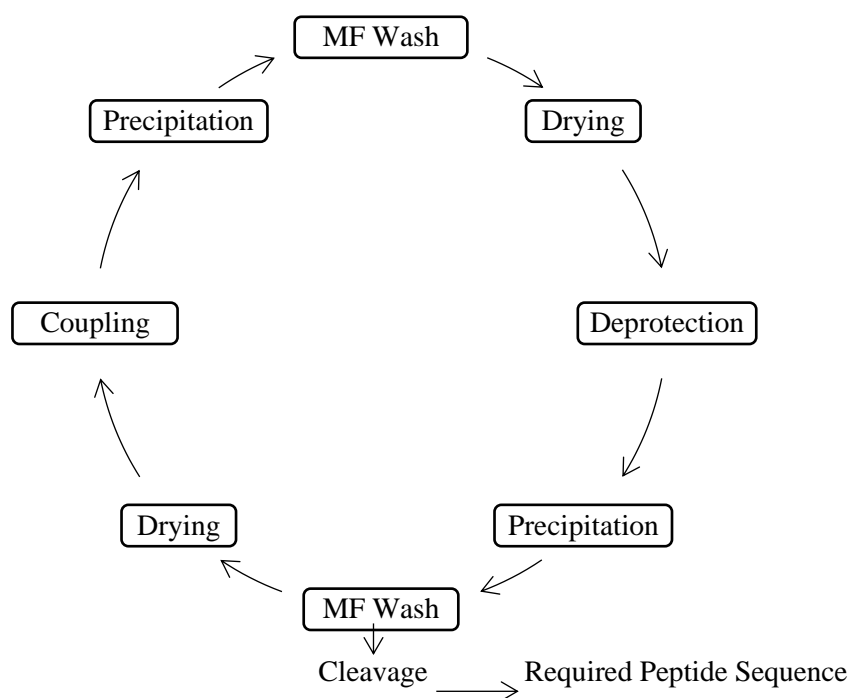
Figure 2.8 - Schematic representation of LPPS[19].

By being carried out as homogeneous reactions it has the advantages of having fast reaction rates, not have mass transfers limitations, not being effected by hindrance and accessibility problems and not having polymer solvation constrains. However, it presents three separation stages requiring a long cycle time[19].

Recently, a new platform was presented for peptide synthesis that can be applied in LPPS approach: Nanostar Sieving Technology (NST). The concept of nanostar hub molecule consists in a three-armed, star-shaped molecule to grow peptides, expanding the loading of building blocks onto the anchor. Sieving efficiency increase by the raising of three-fold amino acid every cycle and it is unimolecular weight[22]. The hub linkers utilised in this project were the wang linker and the rink amide linker,

Figure 2.9.

A:

B:

● Connection to hub
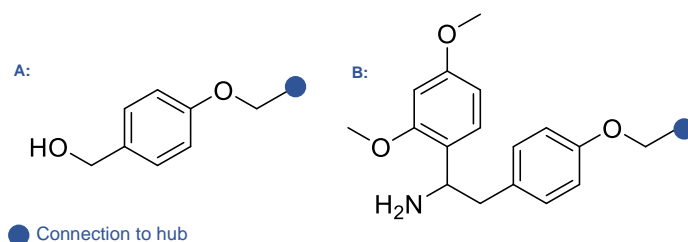
Figure 2.9 - Common linkers to the hub. A: Wang Linker. B: Rink Amide Linker[14].

2.3.3. Membrane Enhanced Peptide Synthesis – MEPS

A different approach, Membrane Enhanced Peptide Synthesis (MEPS), was introduced in 2010 by So *et al.* that increases the peptide purity and overcomes the purification difficulties. Peptides were built on a soluble anchor and the synthesis reactions happen in solution-phase, providing the same benefits as LPPS. The technology of molecular separation in organic solvents via nanofiltration (OSN) is then applied, allowing the in-cycle purification of growing peptides from excess coupling reagents and by-products, Figure 2.10.

Overall, it overcomes the limitations of LPPS and SPPS, and demonstrates admirable purity and yield from the final peptide, showing a promising alternative for industrial scale up[20]. MEPS process can also make use of NST. The aim of this project can be also beneficial and introduced in MEPS.
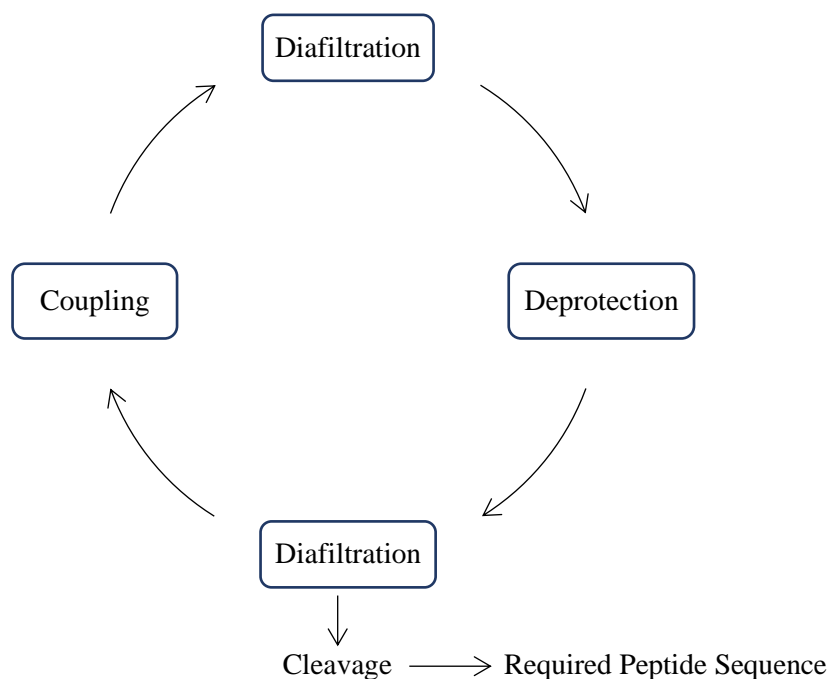
Figure 2.10 - Schematic representation of MEPS.

2.4. Solubility and Solubility Parameters

For many years, the solubility parameters have been used to select solvents, predict the compatibility of polymers, chemical resistance, permeation rates, kinetics control, etc. The recurrent use of this concept is due to it is based on well-defined and correct principles. A lot of improvements have been made to solubility parameters concept and with the general use of computer techniques, it permitted the optimisation of the solvent choice respecting the cost, workplace environment, solvency, evaporation rate, flash point, among others. New developed software allows to predict the solubility parameters, without the need to do it experimentally, and in this way predict how, for example, a polymer will dissolve in a given pure solvent or in a mixture of two solvents. The software used in this study will be addressed in Chapter 3.

The basic principle of the solubility parameters is ''like dissolves like'', what means that two liquids with similar values from the solubility parameters will be miscible and polymers will dissolve in solvents with solubility parameters values similar to theirs.

There is some limitation regarded the solubility parameters that need to be carried carefully. The first is the temperature dependence because the solubility parameters will change to different temperatures. Another point to consider is the size of the molecules since the size of the solvent and solute molecules is significant for solubility, permeation, diffusion, and chemical resistance phenomena. A smaller molecule tends to be more quickly soluble than a large molecule. Also, smaller molar volume solvents tend to be better than those with larger molar volumes, even though they may have identical solubility parameters.

The parameters approached in this chapter are the Hildebrand and the Hansen parameters.

2.4.1. Hildebrand Parameter

In 1936, Joel. H Hildebrand proposed a numerical value that reflects the relative solvency behaviour for a specific solvent or solute, that only in 1950 was recognised as a solubility parameter and represented by the symbol δ [23]. The Hildebrand solubility parameter, δ, is defined as the square root of the cohesive energy density, ced, which in turn is derived from latent energy of vaporization, ΔE_v, dived by molar volume, V[24], represented in Eq. 2.1.

$$\delta = (ced)^{\frac{1}{2}} = \left(\frac{\Delta E_v}{V}\right)^{\frac{1}{2}} \qquad (Eq.\,2.1)$$

$$\Delta E_v = \Delta H_v - RT \qquad (Eq.\,2.2)$$

The latent energy of vaporization is defined by Eq. 2.2, where ΔH_v is the latent heat of vaporization, R is the universal gas constant, and T is the absolute temperature.

The cohesive energy density of a liquid can be considered a reflection of the degree of Van der Waals forces that hold the molecules of the liquid together. Since it is derived from latent energy of vaporization, shows a correlation between vaporization and Van der Waals forces. This correlation translates into a new correlation between vaporization and solubility behaviour. When two liquids are mixed, the molecules of both liquids must be physically separated by the molecules of the other liquid. During vaporization, the same process happens, and the molecules of the liquid must be physically separated. In both cases, the same intermolecular Van der Waals forces are needed to be overcome. Subsequently, if two components only exhibit good solubility behaviour when the intermolecular attractive forces are similar, it can be expected that components with similar cohesive energy density exhibit a good solubility behaviour[25].

The standard international units (SI units) is mega-pascal, MPa$^{1/2}$, since is derived from cohesive pressures and the SI units for pressure is pascal, Pa. Otherwise, the original units and still usually used are (cal/cm^3)$^{1/2}$[24], Eq. 2.3.

$$1 \ cal^{\frac{1}{2}}cm^{-\frac{3}{2}} = 0{,}48888 * MPa^{\frac{1}{2}} \qquad (Eq.\ 2.3)$$

It is possible to create a solvent spectrum by ranking solvents according to their solubility parameter. That means that solvents with similar values and comparable force will be close to each other. Theoretically, just a delimited group of solvents of the spectrum, with Hildebrand parameter values similar to the value of a specific component, will dissolve that specific component and the rest of the solvents will not[25].

In view of identifying suitable solvents and non-solvents for polymers and considering the notion of ''like dissolves like'', quantitative models were developed. The Hildebrand Model states that solvents with δ within $\pm$ 2 MPa$^{1/2}$ from the value of the δ of the polymer are good solvents; solvents with a difference more than 2 MPa$^{1/2}$ from the value of the δ of the polymer are considered non-solvents. The factor 2 MPa$^{1/2}$ was determine based on empirical considerations[26]. Figure 2.11 represent the Hildebrand model criterion.
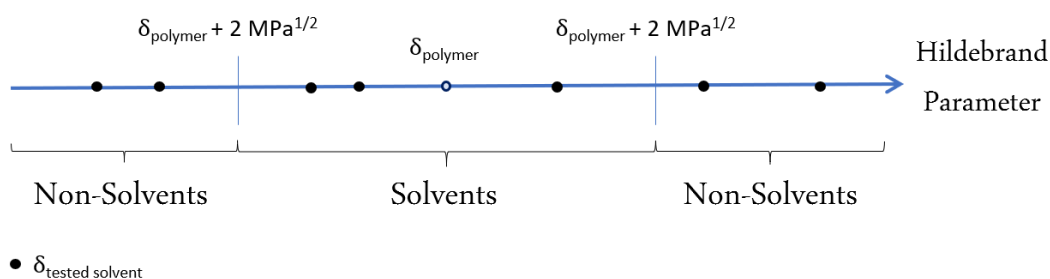
Figure 2.11 - Solubility criteria of the Hildebrand solubility parameter[26].

This model showed an accuracy of 60% for solvents and 70% for non-solvents in previous studies. Consequently, the Hildebrand Model is a potential candidate to predict solvents and non-solvents for polymers. However, it also demonstrated better capability to non-polar (apolar) polymers (accuracy of 70-75%) than for polar polymers (accuracy of 57%)[26].

2.4.2. Hansen Parameter

A new approach to Hildebrand work was published in 1967 by Charles M. Hasen. This extension of the Hildebrand parameter supports that the total energy of vaporization is divided into several individual parts – dispersion forces (atomic); permanent dipole-permanent dipole forces (molecular); and hydrogen bonding (electron exchange)[27]. The basis of the Hansen approach is that all type of physical bonds must be broken during evaporation, including the nonpolar, polar and hydrogen bonding[24]. Since it shows a good representation of the experimental data, the geometric mean is used to estimate the interaction between different molecules in their mixture[27].

The Hildebrand solubility parameter is now divided by three parameters – the Hansen solubility parameters, Eq.2.4. The Hansen solubility parameters, also called HSPs, are δ_D, δ_P, and δ_H, and it represents the dispersion forces (D), the permanent dipole-permanent dipole (polar) forces (P), and the hydrogen bonding (H), respectively[23].

$$\delta^2 = {\delta_D}^2 + {\delta_P}^2 + {\delta_H}^2 \qquad (Eq.\,2.4)$$

The calculation if the dispersion solubility parameter, δ_D, is according to the Blanks and Prausnitz procedures, that use corresponding states principles at 25°C. δ_P, the polar solubility parameter, is calculated from the dipole moment and molar volume. Lastly, the hydrogen bonding solubility parameter, δ_H, is found by subtracting the dispersion and polar energies of vaporization from the total energy of vaporization or by group contributions[27]. Such as Hildebrand Parameter, also the HSPs follows the notion ''like dissolves like''. So, it is expected that two different components with similar parameters values show a good solubility while with very distinct parameter values a poor solubility. To compare two materials, it is calculated the ''distance'' between those materials, Ra, by the following Eq. 2.5, where the subscript 1 and 2 refer to material 1 and material 2[24]:

$$Ra^2 = 4(\delta_{D1} - \delta_{D2})^2 + (\delta_{P1} - \delta_{P2})^2 + (\delta_{H1} - \delta_{H2})^2 \qquad (Eq.\,2.5)$$

When the δ_D, δ_P, and δ_H HSPs from all solvents and material test are plotted three-dimensionally, and a computer program locates a 'sphere' centred in HSPs of the material test, the solvents included in the spere are the 'good' solvents and the excluded are the 'bad' solvents. The radius, Ro, of that sphere describes how small or larger the interaction range is, since the 'good' solvents are those how interact strongly, and the 'bad' solvents do not interact[28]. By other

words, Ro is the largest value acceptable to Ra where solubility is allowed. Frequently, Ro is called the radius of Hansen solubility parameter sphere. To quantify the distances of Ra relative to Ro it is used the Relative Energy Difference (RED) number, Eq. 2.6:

$$RED = \frac{Ra}{Ro} \qquad (Eq.\,2.6)$$

If RED is less than 1 indicates high affinity and is considered a good solvent. If RED is equal to 1 is a boundary condition. The progressive higher RED value indicates progressive lower affinities that consequential means progressive poorer solvents[28], Figure 2.12.
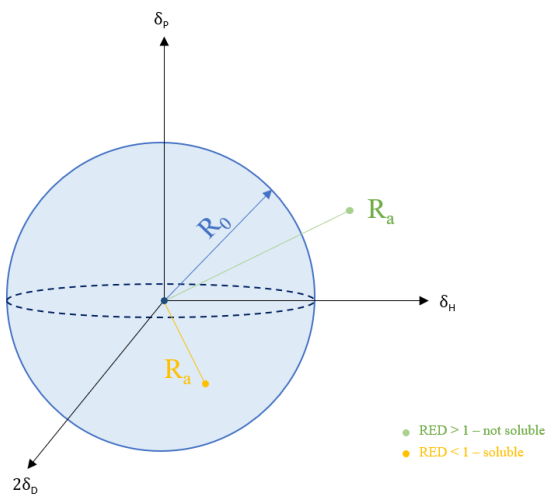
Figure 2.12 - Schematic representation of the Hansen space[29].

Since the Ro is determined experimentally, is impossible to use the RED number to identify suitable solvents just through the estimated HSPs. The Hansen Model affirms that the Ro for the polymers is 8 MPa$^{1/2}$, Figure 2.13. So, the solvents that the Ra to the polymers are within 8 MPa$^{1/2}$ are considered good solvents and those solvents that Ra is bigger than 8MPa$^{1/2}$ are considered non-solvents. The factor 8 MPa$^{1/2}$ was determined based on empirical considerations[26].

The Hansen Solubility Model presented by Venkatram *et al.* has some limitations since for only 25 polymers the values of the HSPs were available, comparable with a data set of 75 polymers to the Hildebrand solubility model. The predictive capability exhibited an accuracy of 67% for solvents and 76% for non-solvents[26].
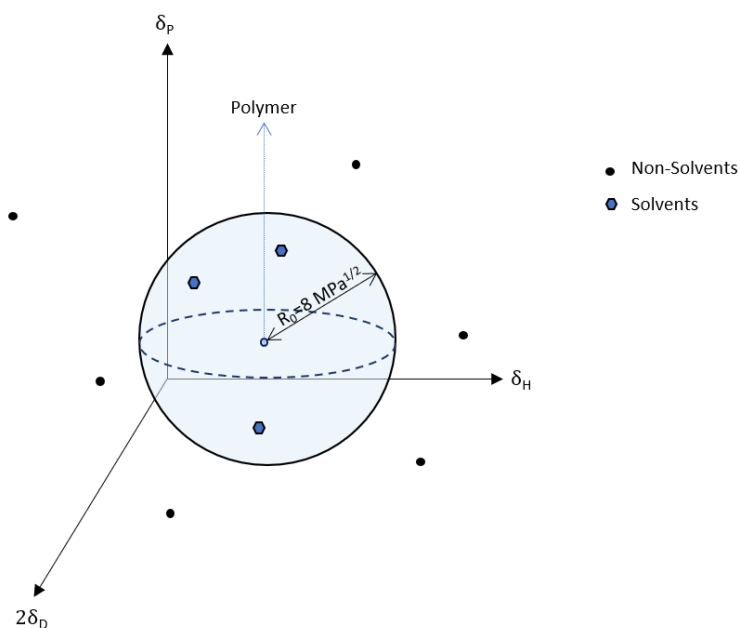
Figure 2.13 - Solubility criteria of the Hansen solubility parameter[26].

2.4.3. COSMO-RS

To study the properties of molecules in solution, it were developed numerous theoretical and computational methods, being that the COnductor-like Screening MOdel (COSMO) is one of them[30]. COSMO belongs to continuum solvation models since it incorporates solvation effects into quantum chemical calculations in order to describe the solvent effects[31]. Therefore, by a dielectric continuum - polarization charges of the continuum caused by the polarity of the solvent, from a scaled-conductor approximation - surrounding the solute molecule outside of a molecular cavity is possible to calculate the surface charge. Some limitations were presented to this model, as for example, the inability to distinguish two solvents with different properties but with essentially identical dielectric constants (as the case of the cyclohexane and benzene or methoxyphenol and heptanone)[32], which means that they are pointless in terms of real solubility[33].

Published in 1995, a method developed by Klamt combine the quantum chemical calculations with statistical thermodynamics, designated COSMO-RS - the COnductor-like Screening Model for Real Solvents – provide the prediction of thermodynamic properties without experimental data[34] and overcoming the COSMO approach limitations. Since then, almost all works were focused in COSMO-RS. This approach by treating the solute and solvent as an equal instead than considering the solvent as just a dielectric field allows consistent mixture thermodynamics at variable temperatures[32].

This approach is divided into two main steps. The first step consists in to obtain the screening charge densities (σ) on the molecular surface, Figure 2.14. For that, is needed to induce a

polarization charge density on the surface by introducing that molecule into a virtual conductor – ''dielectric continuum''. The molecule will converge to the optimal energetically state, considering the electron density and energy, after running the quantum chemical COSMO calculations, creating all around the molecule a realistic surface polarization charge.
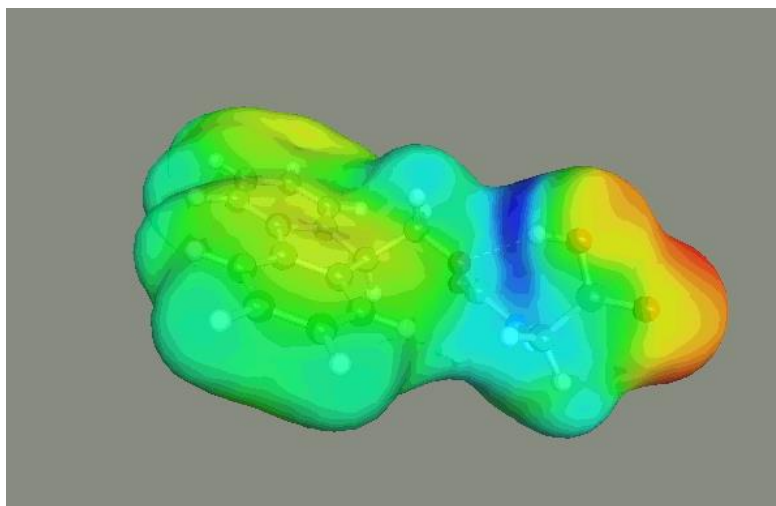
Figure 2.14 - Sigma surface of protected glycine amino acid.

The red colour means a negative charge, the blue colour a positive charge, the green a positive equilibrium layer (partially electropositive) and the yellow a negative equilibrium layer (partially electronegative). Although the 3D charges provide visual clues (a positive charge is rather to match with a negative one), for the solubility calculations the σ-profiles are used.

The σ-profile is outlined on a histogram of the charge density, providing information about molecular polarity distribution, as function of the probability distribution the sigma profile, $p(\sigma)$. This is obtained after the second step, that consists into quantifying the interaction energy of the pairwise interacting surface segments related to molecular interaction modes (i.e. hydrogen bonding), realised by statistical thermodynamic calculations. It is important to refer that a negative partial charge of an atom cause positive screening charge density and vice versa[35]. Also, the σ-potential is found that translate the affinity of the compound to interact with the solvents with polarity and hydrogen bonds[36].

The σ-profile of the amino acid glycine (protected and unprotected), Figure 2.15, showed for both amino acids two peaks in both polar regions, meaning that glycine possesses both hydrogen bonding donor and acceptor capacity, corresponding to NH_3^+ and $CHOO^-$, respectively. However, the protected glycine presents a narrow distribution of the charge densities around zero, reflecting in being the least polar compound.

The correspondent σ-potentials, Figure 2.16, presents higher values around zero for both amino acids, which translates in an unfavourable interaction with nonpolar surface, reflecting on stronger hydrophilicity[37]. Also, the smaller values of μ(σ) on the right side of the graph for the unprotected glycine, show a higher affinity for hydrogen bond donor than the protected amino acid[38]. The Figure 2.14 - Figure 2.16 were made using *COSMOtherm* software, explained in Chapter 3.
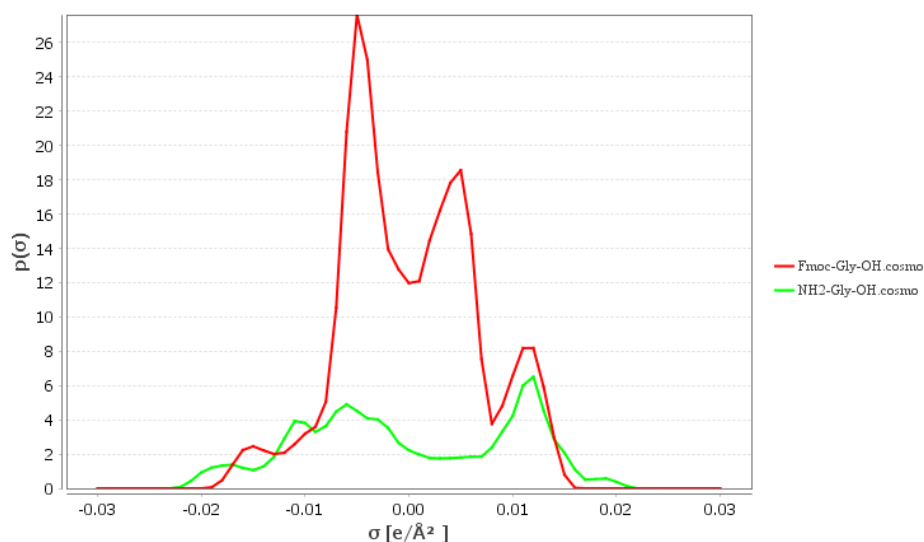
Figure 2.15 - Sigma Profile of glycine (protected and unprotected) amino acid.
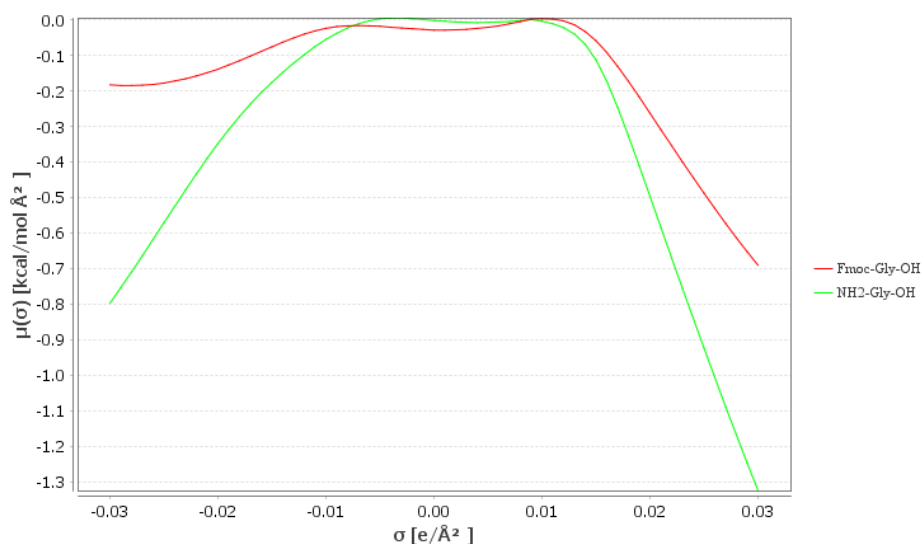
Figure 2.16 - Sigma Potential of glycine (protected and unprotected) amino acid.

This method is usually used to estimate the solubility (mol/L) between the solute and solvent and other properties such as partition coefficients, vapor pressures and activity coefficients[39].

2.5. Optimisation of peptide synthesis

Being peptide synthesis a process of constant improvement, in the aim of this project, two main areas will be focused on optimisation: quenching reagents and peptide solubility estimation.

2.5.1. Quenching reagents

After coupling reaction, residual amino acids and excess of coupling reagent are a potential cause of undesirable side reactions. Typically, in SPPS they are removed via microfiltration wash, however this procedure requires excessive use of washing solvents. As alternative option to minimize the risk of undesired reactions, the residual amino acids need to be quenched (inactivated) before Fmoc deprotection. Then, before the next coupling reaction is essential to proceed to the removal of those by-products[40]. Three different reagents were tested using Fmoc-Phe-OH:

1. Piperidine

Piperidine is the reagent frequently used and it shows good purity yields. The reaction is generally complete within 10 minutes, but for a safe deprotection is recommended 20 minutes. In order to also optimise the use of piperidine, different concentrations have been tested. From a green optic, piperidine is considered as a hazardous substance[9].

2. Thiomalic Acid

Daisuke *et al.* showed that when Fmoc deprotection is carried out by thiomalic acid, the excess AE of the amino acids are converted into acidic species and it can be removed by a simple basic aqueous solution washing, providing no interferences from these by-products in next condensation. This reagent does not derive any impurity and presents a purity yield of 84% after global deprotection[40].

3. Aniline

Aniline is an organic compound, mainly used for the production of methylenedianiline and related compounds by condensation with formaldehyde. Previous studies demonstrated that aniline possesses a slow kinetic related to Fmoc deprotection.

2.5.2. Solubility in peptide synthesis

The presence of solid particles during the process will increase the mass resistance which will affect the reactions kinetic. Consequently, more reagent will be needed, and the operation cost will also increase, besides the yield and the purity of the peptide will be affected. In the production of new peptides, the common way to evaluate the peptide solubility is experimentally.

The experimental process is effective, however, it takes a lot of time and increases the cost and the reagent waste, since the analyses need to be realized in different solvents in order to choose the best one. In this project, a different approach will be taken to overcome those issues. The solvents will be selected by solubility prediction methods and verified experimentally. This analysis includes common solvents used in the peptide synthesis and, in order to turn the process greener, also some green solvents:

1. Dimethyl sulfoxide (DMSO) and *N*-Methyl-2-pyrrolidine (NMP)

DMSO and NMP are two colourless polar aprotic solvents commonly used in peptide synthesis. As polar aprotic solvents, they present excellent results in Fmoc deprotection (Figure 2.18), however, relatively to coupling performances they present poor results (Figure 2.17)[41]. Both solvents show good solubility capacity, standing out the DMSO, that show the best performance to dissolving coupled products, deprotected products and by-products[42]. In GSK solvent sustainability guide, DMSO classified as amber and NMP is classified as red, whereby both represent a risk.
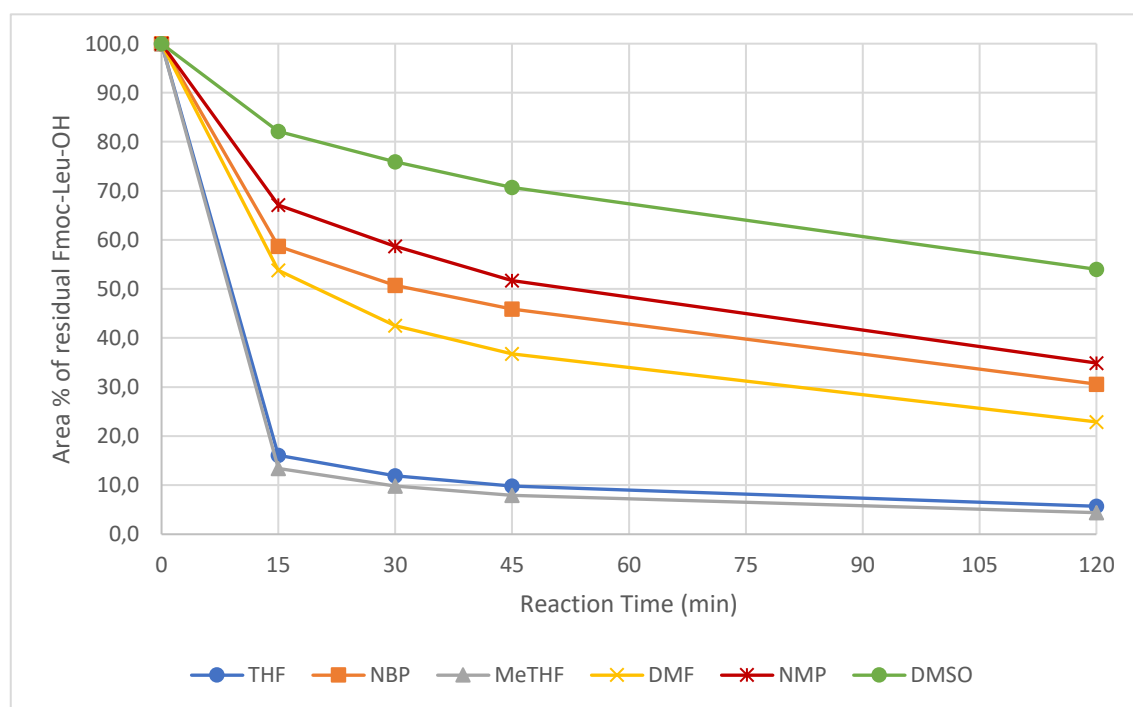
Figure 2.17 - Relative area % of residua Fmoc-Leu-OH during coupling reaction[41].
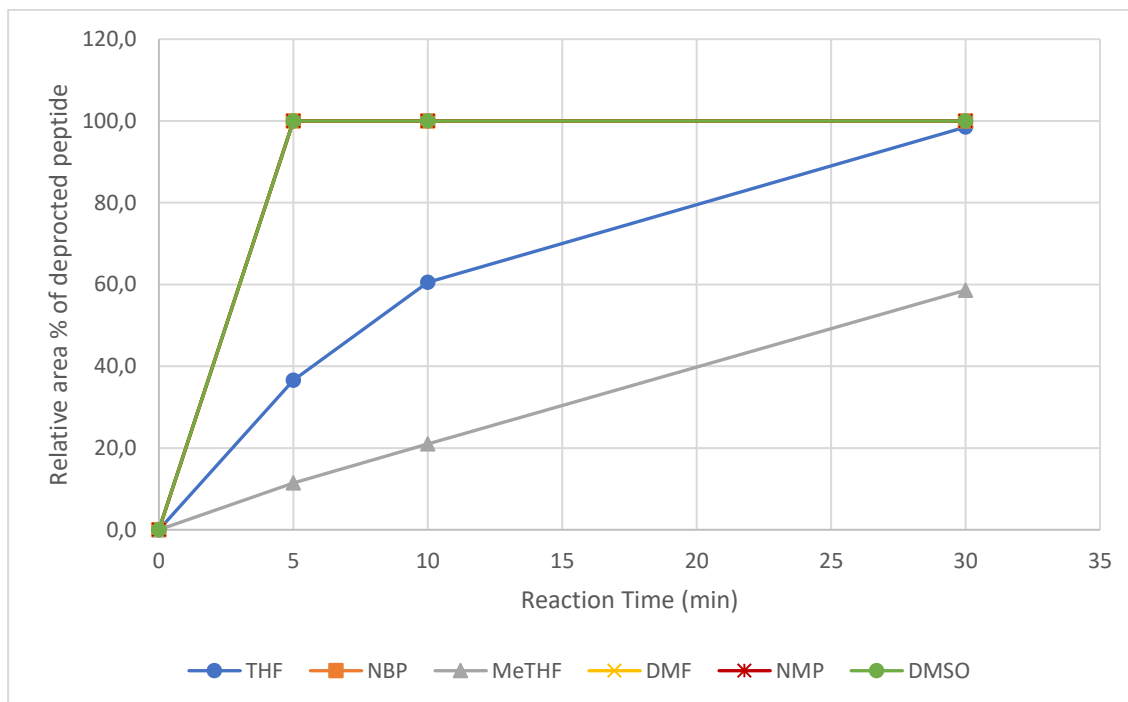
Figure 2.18 - Relative area % of deprotected peptide during deprotection reaction[41].

2. Tetrahydrofuran (THF) and 2-Methyltetrahydrofuran (MeTHF)

THF and MeTHF are two colourless aprotic solvents. THF is largely used in the synthesis of peptides, since it shows good coupling performance[41] (Figure 2.17) and presents good solvation capacity[42]. Regarding deprotection performance, THF is very slow and after the time limit of thirty minutes, the cleavage reaction was not complete (98.5%)[41]. THF represents major issues as being classified as red in the GSK guide. One the other hand, MeTHF is classified as amber (there is a low level of confidence in the solvents placement due to assessments for it includes 4 or more data gaps[10]), presenting as a more eco-friendly alternative to THF. In terms of the coupling step, MeTHF offers better performance than THF (Figure 2.17). However, it shows problems in the deprotection reaction (Figure 2.18) and in the solubility capacity[42].

3. γ-Valerolactone (GVL) and *N*-Formylmorpholine (NFM)

GVL and NFM are two colourless aprotic solvents, classified as greener than THF and NMP. Ashish *et al.* found that in SPPS, GVL and NFM showed solubility good results for the amino acids. The coupling efficiency was evaluated by synthesizing the peptide H-Tyr-Aib-Aib-Phe-Leu-NH2, that by containing two Aib in a row increase the possibility of side-product. The results were promising, showing excellent coupling efficiencies, and the GVL show a peptide purity of 99.2%, being higher than the 97.8% purity showed by DMF (highly hazardous solvent use in peptide synthesis)[43].

4. *N*-Butylpyrrolidine (NBP) and Propylene carbonate (PC)

NBP and PC are two aprotic solvents, both classified as green solvents. NBP showed excellent results in deprotection reaction but is not so effective in coupling reaction (Figure 2.18 and Figure 2.17, respectively). Due to their high viscosity, the transfers of the solution in an automatic synthesizer is slower. The purity of crude peptide generated by NBP (80%) is lower compared to DMF (86%), but the impurity profile is similar[41]. PC showed good results in coupling and deprotection reaction, with a good yield of the peptide, for both LPPS and SPPS using acid and base amine protecting groups, respectively[8].

As learning tool and considered a green solvent by GSK solvent sustainability guide[10], also butanol was added to solvents set.

2.6. Summary

In summary, various studies as So *et al.*[20] and Lopez *et al.*[41], has been made over time to improve the overall peptide synthesis process, that will be applied in this project. In the optimisation of the quenching reagent, the conditions adopted will be in order to suppress the side reactions presented above and to improve the reagent performance. Therefore, Fmoc will be used as protecting groups, DIC as coupling reagent and HOBt as additive to suppress side reactions. This chapter also allowed to identify two others promising quenching reagents, aniline and thiomalic acid, as possible piperidine replacement. In relation to the solubility of the peptide's estimation, two different approaches will be implemented, the Solubility Parameters approach, that include the Hildebrand Parameter Model and the Hansen Parameter Model, and the COSMO-RS approach. This both approaches will be employed in free peptides and in peptides attached to the hub (nanostar), by wang linker and by rink amide linker. Finally, promising common solvents (DMSO, THF, and NMP) and green solvents (MeTHF, GVL, NFM, NBP, PC) with good results in peptide synthesis were identified and selected to be applied in the solubility study, that is, applied in the solubility models and experimentally tested.

3. Materials and Methods

3.1. Introduction

In this chapter, all methods, procedures, and software employed to achieve the goals and complete the project will be specified, as well as all the materials and data used. The results obtained from here will be presented in the next chapters.

3.2. Materials

3.2.1.   Amino Acids

Table 3.1 presents the information of the amino acids used in the global experimental activity and Figure 3.1 presents the structures of the amino acids.

Table 3.1 - Information related to the amino acids used.

Chemicals	CAS number	Supplier	Molecular Weight
Fmoc-Phe-OH	35661-40-6	Sigma-Aldrich	387.43
Fmoc-Gly-OH	29022-11-5	Sigma-Aldrich	297.31
Fmoc-Ser(tBu)-OH	71989-33-8	Sigma-Aldrich	383.44
Fmoc-Arg(Pbf)-OH	154444-77-9	Sigma-Aldrich	648.77
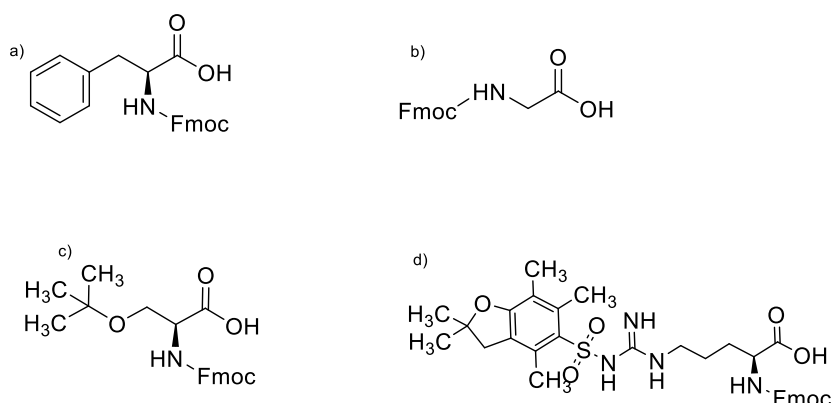
Figure 3.1 - Amino Acids structures. a) Fmoc-Phe-OH; b) Fmoc-Gly-OH; c) Fmoc-Ser(tBu)-OH; d) Fmoc-Arg(Pbf)-OH.

3.2.2. Solvents

Table 3.2 presents the information of the solvents set used in the global experimental activity and Figure 3.2 presents the structures of the solvents. Must be noted that, not all solvents used in the predictions models were experimentally tested.

Table 3.2 - Information related to the solvents used.

Chemicals	CAS number	Supplier	Molecular Weight
NMP	872-50-4	VWR International	99.13
THF	109-99-9	VWR International	72.11
Butanol	71-36-3	VWR International	74.12
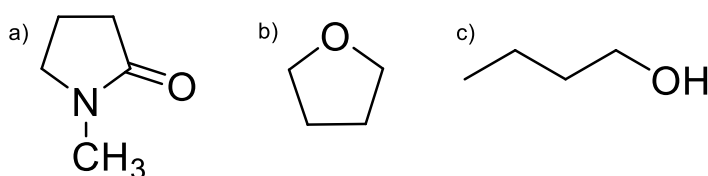
Figure 3.2 - Solvents structures. a) NMP; b) THF; c) Butanol.

3.2.3. Others

Table 3.2 presents the information of other chemicals used in the global experimental activity and Figure 3.3 presents the structures of the quenching reagents.

Table 3.3 - Information related to other chemicals used.

Chemicals	CAS number	Supplier	Molecular Weight
1-Hydroxybenzotriazole hydrate	123333-53-9	Sigma-Aldrich	135.12
DIC	693-13-0	Sigma-Aldrich	126.20
Piperidine	110-89-4	Sigma-Aldrich	85.15
Aniline	62-53-3	Sigma-Aldrich	93.13
Thiomalic Acid	70-49-5	Sigma-Aldrich	150.15
Acetonitrile	75-05-8	Sigma-Aldrich	41.05
Water	7732-18-5	Sigma-Aldrich	18.015
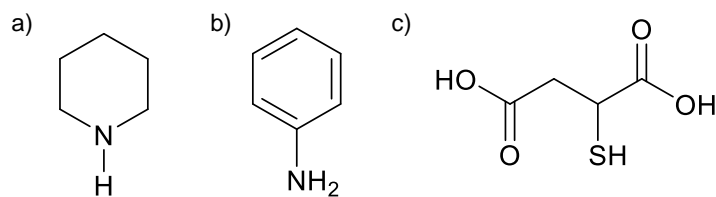
Figure 3.3 - Quenching reagents structures. a) Piperidine; b) Aniline; c) Thiomalic Acid.

3.2.4. Analytical Equipment

Ultra High-Performance Liquid Chromatography (UHPLC) - To control the concentrations during the experimental procedure of the quenching amino acid analysis it was used an UHPLC. Table 3.4 collected the details of the UHPLC method adopted and Table 3.5 summarised the details of the column used.

Table 3.4 - UHPLC method used in the amino acid quenching analysis.

Product Name	1260 Infinity ll UHPLC System			
Manufacturer	Agilent Technologies			
Detection Wavelength (nm)	220			
Solvent A	5 µM ammonium Acetate in deionised water			
Solvent B	Acetonitrile			
Gradient	Time (min)	Flow rate (mL/min)	%A	%B
	0	0.3	90	10
	1	0.3	90	10
	8	0.3	10	90
	9	0.3	90	10

Table 3.5 - Details of the column used in UHPLC.

Product Name	Acquity UPLC Protein BEH C4 Column
Pore side	300 Å
Particle Size (dp)	1.7 mm
ID x Length	2.1 mm x 100 mm
Column Temperature	60°C

Mass Spectrometry - For compound identification in the samples, it was used the Agilent 6130 Single Quadrupole hardware and software. Table 3.6 summarised the details of the mass spectrometry used.

Table 3.6 - Details of the Mass Spectrometry used in UHPLC.

Gas Flow rate	8.0 L/min
Nebulizer Pressure	35 psig
Gas Temperature	350 °C
Capillary Voltage	4000 V

High-Performance Liquid Chromatography (HPLC) - To determine the solubility of the amino acids it was used a HPLC. Table 3.7 collected the details of the UHPLC method adopted and Table 3.8 summarised the details of the column used.

Table 3.7 - HPLC method used in the amino acids solubility determination.

Product Name	1100 HPLC System			
Manufacturer	Agilent Technologies			
Detection Wavelength (nm)	270			
Solvent A	5 µM ammonium Acetate in deionised water			
Solvent B	MeCN/MeOH 4:1			
Gradient	Time (min)	Flow rate (mL/min)	%A	%B
	0	1	80	20
	20	1	5	95
	25	1	80	20

Table 3.8 - Details of the column used in HPLC.

Product Name	Reversed phase C18 Column
Manufacturer	ACE Hichrom
ID x Length	250 mm x 4.6 mm
Column Temperature	60°C

3.3. Amino Acid Quenching Analysis

3.3.1. Experimental Method

In a test tube, it was added 0.2830 mmol of amino acid Fmoc-Phe-OH (3 equivalents of start material), 0.3538 mmol of HOBt (3.75 start material equivalents) and filled with 10 mL of NMP: THF 65:35 solvent. After dissolved, 0.2830 mmol of DIC was added and the activation of the amino acid ran for 10 minutes. At the end of the 10 minutes, the first sample (star material) was taken, 10 µL of the reaction mixture, and transferred to a UHPLC vial containing 1 mL of acetonitrile (MeCN). Then, 0.1981 mmol of the quenching reagent (piperidine, aniline or thiomalic acid) was added to the tube (2.1 start material equivalents). At 0.5; 2; 5; 10; 30; 60; 120 minutes after adding the quenching reagent, 10 µL of the reaction mixture is sampled and transferred to a UHPLC vial containing 1 mL of MeCN, where the all species concentration will be measured. The process was repeated to all quenching reagents, wherein just the initial concentration was modified, by changing the star material equivalents. In the second experiment, the respective quantities were: 0.1887 mmol of amino acid Fmoc-Phe-OH (2 equivalents); 0.2358 mmol of HOBt (2.5 equivalents); and, 0.1887 of DIC (2 equivalents). The quenching reagent amount remains constant.

These equivalents were obtained after an equivalent study, started with 5 equivalents of amino acid and 2.4 equivalents of quenching reagent (piperidine). The process explained above was optimised until found the best conditions – 2 or 3 amino acid equivalents and 2.1 quenching reagent equivalents.

Factors as temperature and stirring speed were controlled since they can affect the experiment. In this analysis, the temperature was kept at 30°C and the stirring speed was kept at strength 6. Despite the fact that the pressure was not controlled during the performance, it was assumed to be constant at ambient temperature.

It should be noted that the molecular weights of amino acids, quenching reagents, and other materials needed to calculate the amount used in this experiment, were collected non-experimental by Sigma-Aldrich (*https://www.sigmaaldrich.com*) products database.

The performance of amino acid quenching was carried out in the Carousel 12 reaction station (Radleys).

3.3.2. Analytical Methods
3.3.2.1. Determination of concentration change form UHPLC chromatogram

It is assumed that the species concentration is linearly proportional to the area under the peak in a chromatogram. So, by controlling the area changing over the reaction time equals to monitor the specie concentration.

The UV chromatogram shows different main peaks, that correspond to different species, with corresponding retention times. Then, in the mass spectrum the species were recognized by their retention time, that generally suffers a lag of 0.5 seconds of the retention time of every peak from the UV chromatogram (since the sample passes in the UV detector in HUPLC first and second on the mass spectrum).

On each mass spectrum peak is possible to find the molecular mass of the specie, that corresponds to the most intense signal. This way, each specie is identified as an amino acid, active ester or quenched specie and their concentrations evolution during the reaction are monitored by converting the area under the peak into % residual of each specie.

3.3.2.2. Kinetic modelling and rate constant of the amino acid reduction

After all chromatogram areas from each sample were converted into % of residual amino acid, they were plotted in a graph of % of residual amino acid against the quenching reaction time. The amino acid with DIC reaction is described in Reaction 3.1.:

$$\text{Fmoc-AA} + \text{DIC} \rightarrow \text{O-Acylisourea} \qquad (Reaction\ 3.1)$$

The kinetic assumed for the amino acid consumption was the 2nd order reaction, where the rate constant is presented in Eq. 3.1, where [AA] is the concentration of the protected amino acid (Fmoc-AA) and [DIC] is the concentration of DIC.

$$Rate = k_a[AA][DIC] = k_a[AA]^2 \qquad (Eq.\ 3.1)$$

Since the initial concentrations of amino acid (Fmoc-Phe-OH) and DIC were the same and they react one to one molar ration, the Eq. 3.1 can be simplified as Eq. 3.2 and, posteriorly, integrated as Eq. 3.3:

$$rate = -\frac{d[AA]}{dt} = k_a[AA]^2 \qquad (Eq.\ 3.2)$$

$$\frac{1}{[AA]} = k_a t + \frac{1}{[AA]_0} \qquad (Eq.\ 3.3)$$

Once again, [AA] is the concentration of the amino acid (Fmoc-protected) at a time, $[AA]_0$ is the initial concentration of the amino acid, t is the reaction time and k_a is the consumption of

32

amino acid rate constant. Therefore, after determined the first-rate constant k_a, the theoretical concentration of the amino acid was calculated by Eq. 3.4:

$$[AA] = \frac{1}{\frac{1}{[AA]_0} + k_a t} \qquad (Eq.\,3.4)$$

Guessed all theoretical values for amino acid concentration, the square of the difference between experimental and theoretical concentration (x^2) for each sample was calculated, and the global error (X^2) founded by summing all x^2. With the aim of minimizing the global error, the Excel Solver function was used, and the k_a was adjusted. Lastly, the new theoretical amino acid concentrations were founded and converted in % of residual amino acids to be plotted in the kinetic models. This kinetic model was applied to the three reagents and to the two different concentrations.

3.3.2.3. Kinetic modelling and rate constant of the quenched specie formation

According to 3.3.2.1., the chromatogram areas (that correspond to the concentrations) for each sample of the quenched specie were converted into % quenched specie (product) formation by Eq. 3.5:

$$\% \; quenched \; specie \; formation_t = \frac{QS \; Area_t}{QS \; Area_{max}} \times 100 \qquad (Eq.\,3.5)$$

Where $QS \; Area_t$ is the chromatogram area of the quenched specie at t hours and $QS \; Area_{max}$ is the maximum chromatogram area of the quenched specie. Ideally, the maximum concentration (100%) of the quenched specie is at 2 hours final, however, that does not always happen and at the end of the reaction the concentration decline. It is suspected that occurred due to uncontrolled deprotection reaction. The % of product formation was then plotted against reaction time. The reaction is described in Reaction 3.2:

$$Fmoc\text{-}AE + Piperidine \; \rightarrow Quenched \; specie \qquad (Reaction\,3.2)$$

The kinetic assumed for the quenched specie formation was of pseudo-1st order instead of 2nd- order reaction, since the piperidine was added in excess (in relation to AE). The equation used for estimating the rate constant was proposed by Lagergren[44], Eq. 3.6, and, posteriorly, integrated as Eq. 3.7:

$$\frac{d[QS]}{dt} = k_q([QS]_{max} - [QS]_t) \qquad\qquad (Eq.\,3.6)$$

$$ln([QS]_{max} - [QS]_t) = \ln[QS]_f - k_q t \qquad\qquad (Eq.\,3.7)$$

Where $[QS]_{max}$ is the maximum quenched specie concentration (as referred, ideally is the final concentration at 2 hours), $[QS]_t$ is the quenched specie concentration at t hours, k_q is the product formation rate constant and t is reaction time. After calculated the rate constant k_q, the theoretical concentration of the product was calculated by Eq. 3.8:

$$[QS]_t = [QS]_f(1 - e^{-k_q \cdot t}) \qquad\qquad (Eq.\,3.8)$$

Like to the kinetic model of amino acid reduction, the global error (X^2) between experimental and theoretical concentrations was founded and minimized by applying the Excel Solver function. The rate constant k_q and the theoretical concentration of the quenched specie were adjusted, and the % of quenched specie plotted in the kinetic model. Again, this kinetic model was applied to the three reagents and to the two different concentrations.

3.3.2.4. Half Time Estimation

As in the previous points, the chromatograph areas of each sample of the active ester were converted in % residual of active ester, in order to calculate the half-time reaction. The half-time is the required time to complete 50% of the quenching reaction and it is used to quantify the reaction rate.

To calculate each $t_{1/2}$, the % of residual active ester was plotted against the reaction time, and the time was the % of the residual active ester is 50% was read.

3.3.2.5. Correlations of k_q and the reagents properties (pKa, LogP, TPSA and molecular weight)

The correlations of k_q and the quenching reagents properties were investigated by plotting a graph of k_q against the individual reagent properties (pKa, LogP, TPSA and molecular weight). A line of the best fit for each correlation was plotted and R^2 calculated to verify the veracity of the correlations.

3.4. Solubility Estimation Methods

In order to estimate the solubility and the solubility parameters of the peptides, three softwares were used. Before all estimations, the amino acids and respective peptides (dimer, 5 mer, 10 mer, 20 mer) were produced in *ChemDraw*. *ChemDraw* is a chemical drawing tool developed in 1985. Different features were included, standing out the option to copy the molecule as a SMILES (simplified molecular-input line-entry system) code. The amino acids and peptides structures are presented in Appendix II.

3.4.1. Hildebrand Solubility Parameter

The Hildebrand solubility parameters were computed with the Polymer Genome (*https://www.polymergenome.org/*), where was needed the SMILES string of each peptide (previously obtained with *ChemDraw*). Polymer Genome is an online platform, that uses machine learning algorithms, capable of predicting a variety of important properties to known and new polymers (within the same chemical class as the parent) and provide uncertainties underlying the predictions[45]. The values obtained appears in Chapter 5, section 5.2.

3.4.2. Hansen Solubility Parameter

To calculate the Hansen solubility parameters was used the program COSMO*quick*. COSMO*quick* offers a large database of quantum chemically calculated σ-profiles and uses, for new molecules, an approximation by molecular fragmentation. This program allows to predict the solubility parameter via QSPR - *quantitative structure–property relationships* (empirical, fitted to published Hansen values) method or COSMO-RS (similar to the experimental procedure, but using COSMO-RS solubilities instead of experimental ones). The QSPR method was the chosen one and it uses a patented algorithm and a procedure that employ sigma moments in combination with an artificial neural network. This method reached high prediction accuracy, while the predictions of δ_D by COSMO-RS did not reach the significance level, as described by Niederquell *et al.*[46].

Despite the peptides can be introduced in COSMO*quick* by SMILES code (if belong to the database) or by a *cosmo* file, when tried the *cosmo* files they didn't work. That way, the Hansen solubility parameter was calculated just for the amino acids and some free peptides. It was not possible to estimate Hansen parameters for peptides attached to the hub.

3.4.3. Solubility estimation

The solubility of peptides in each solvent was calculated with COSMO*therm* software. COSMO*therm* is a central tool that combines quantum chemistry with thermodynamics to predict compounds properties, being even able to predict the properties as a function of concentration

and temperature[30]. The amino acids, peptides and solvents were inserted in the software as *cosmo* files or taken from the database. Using the ''*BP_SVP_AM1_20.ctd*'' parameterization, the solubility of the peptide was calculated, by selecting the ''*Multiple Solvents*'' propriety and defining the temperature to 25°C.

The solubility is calculated from the chemical potentials of pure compound j, μ_j^{pure}, and the chemical potentials at infinite dilution, $\mu_j^{solvent}$, illustrated in Eq. 3.9[35]:

$$log_{10}(x_j) = log_{10}\left[\exp\left(\frac{\mu_j^{pure} - \mu_j^{solvent} - \Delta G_{j,fusion}}{RT}\right)\right] \qquad (Eq.\,3.9)$$

Due to the fact that they are solid compounds, the free energy of fusion ($\Delta G_{j,fusion}$) was to be considered. For that, it was necessary to choose ''*Use ΔG$_{fus}$ estimate* '' option, that use QSPR methods to estimate $\Delta G_{j,fusion}$.

3.4.3.1. Tmolex

To the peptides that are not present in the database, it was necessary to resort to TmoleX software to create new *cosmo* files. After inserted the SMILES code in ''*Geometry*'' section, the molecule was pre-optimised with the semi-empirical ''*GNF2-xTB*'' method. Then, the file was created using ''*COSMO-AM1-SVP*'' job template.

3.4.4. Analytical Methods

3.4.4.1. Protecting group (Fmoc) effect on solubility

The effect of Fmoc in the amino acids and peptides solubility was investigated in relation to the two approaches of solubility estimation.

First was COSMO-RS by plotting a graph of the amino acids logarithmic predicted solubility with Fmoc against the logarithmic predicted solubility without Fmoc.

Second was the Solubility Parameters by plotting a graph of the protected amino acids and peptides Hildebrand and Hansen Parameters against the unprotected amino acids and peptides Hildebrand and Hansen Parameters, respectively.

3.4.4.2. Hub (wang and rink amide nanostar) effect on solubility

The effect of the hubs in the amino acids and peptides solubility was investigated by plotting a graph of the peptides attached to hubs (wang and rink amide nanostar) Hildebrand Parameter

against the free peptides Hildebrand Parameter. The results were exanimated and compared to COSMO-RS results.

3.5. Experimental solubility analysis

3.5.1. Experimental Method

For each solvent and the respective amino acid tested, a calibration curve was created in HPLC-UV, that corresponds to the peaks of the following solubilities: 0 mg/mL; 0.01 mg/mL; 0.1 mg/mL; 0.5 mg/mL. To obtain each calibration point, it was added 1 mL of solvent and the respective amount of the protected amino acid (0.01, 0.1, 0.5 mg) in a test tube. Then, the sample was collected and analysed directly by HPLC. The process was repeated for all calibration points.

A saturated solution was created for the respective protected amino acid and solvent analysis, based on the predicted solubility value. In a test tube was added 10 mL of the solvent, followed by the protected amino acid, that was kept adding until stopped dissolving. The solution was left overnight in a water bath at 25°C with stirring on. After that, the supernatant was collected and diluted to 10 000 times or 500 times. Finally, the sample was collected and analysed directly by HPLC.

This analysis was performed on protected glycine in THF and in butanol, protected serine in butanol, and protected arginine in THF and butanol. Table 3.9 presents the amino acids tested and the respective saturated theoretically estimated solubilities.

Table 3.9 - Amino Acids base solubility for the experimental procedure.

Protected Amino Acid	Fmoc-Gly-OH	Fmoc-Gly-OH	Fmoc-Ser(tBu)-OH	Fmoc-Arg(Pbf)-OH	Fmoc-Arg(Pbf)-OH
Solvent	Butanol	THF	Butanol	THF	Butanol
Theoretically estimated solubility of saturated solution (g/mL)	0.01	0.30	0.04	1.01	$1.18 * 10^{-03}$

3.5.2. Analytical Method

3.5.2.1. Determination of solubility from UHPLC-UV chromatogram

It is assumed that the solubility of the amino acid is linearly proportional to the area under the peak in a chromatogram. So, for the initial samples with the solubility know, the peaks showed by the UV chromatogram were collected and a calibration curved created (*peak versus solubility*).

The peak of the saturated sample was measured and through the calibration curve created the solubility was measured.

3.5.2.2. Predicted absolute solubility validation

The deviation from experimental solubilities to the predicted absolute solubility values was examined by plotting a graph of the logarithmic experimental solubility against the logarithmic predicted absolute solubility and an absolute error bigger than 1 investigated. To measure the differences between the values predicted and the values observed it was used the Root Mean Square Error (RMSE), Eq. 3.10.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}} \qquad (Eq. 3.10)$$

Where $\hat{y}_i$ are the predicted values, y_i are the observed values and n is the number of observations.

4. Results and discussion: Amino Acid Quenching

4.1. Introduction

The quenching of the amino acid excess is an essential step since it avoids the wrong coupling and prevents side reactions, which will affect the product purity and process efficiency. Also, an effective quenching performance will contribute to the reduction time of peptide synthesis. Piperidine showed to be a good quencher, however, still has some negative effects, as mention in Chapter 2, section 2.5. Aniline and thiomalic acid will be tested as quenching reagents and compared to the piperidine. The right amount of amino acid and reagent quencher is indispensable for a controlled analysis.

Therefore, the amino acid study will start with the determination of the best quenching performance conditions, this is, the right amount of amino acid and quenching reagent. Consequently, to determine which the best reagent, the quenching rate constant (k_q) and the activation rate constant (k_a) will be calculated, and the factors, as pKa, polarity, hydrophobicity, and molecular weight will be examined. The study was carried out with the assistance of a UHPLC to control the species evolution during the reaction and the equations described in Chapter 3 (section 3.3) used to determine k_q and k_a.

To summarize, the main objectives of the amino acid quenching study are:
1. Determine the ideal amino acid quenching performance condition.
2. Determine k_q and k_a under the different reagents.
3. Identify the factors affecting the quenching constant rate.
4. Choose the best quenching reagent.

4.2. Equivalents study

In order to obtain a controlled reaction, the amount of amino acid, HOBt, DIC and quencher reagent need to be controlled.

At the beginning of the study, it was adopted a strategy of **5** equivalents for amino acid (Fmoc-Phe-OH and Fmoc-Ser(tBu)-OH) and DIC and of **2.4** equivalents for piperidine, the first quencher testing. The results of the reaction evolution are presented in the following graphs, Figure 4.1 and Figure 4.2.
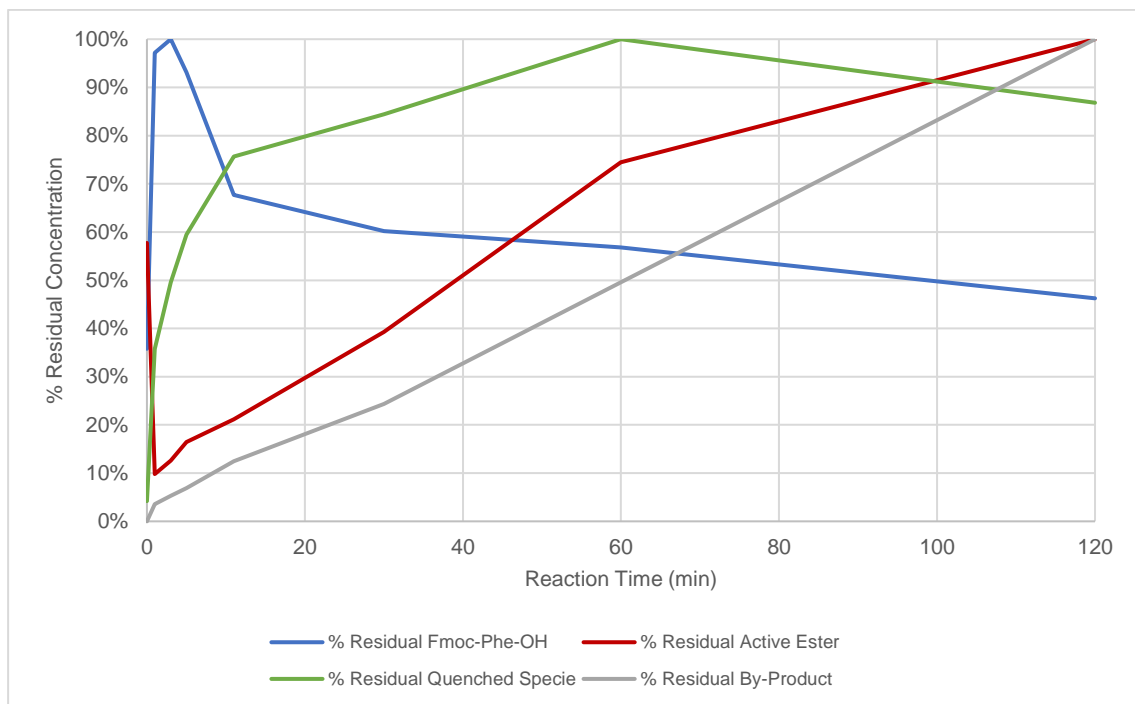
Figure 4.1 – Graph of the evolution of the concentrations of the species (5 equiv. Fmoc-Phe-OH and 2.4 equiv. piperidine) during quenching performance.
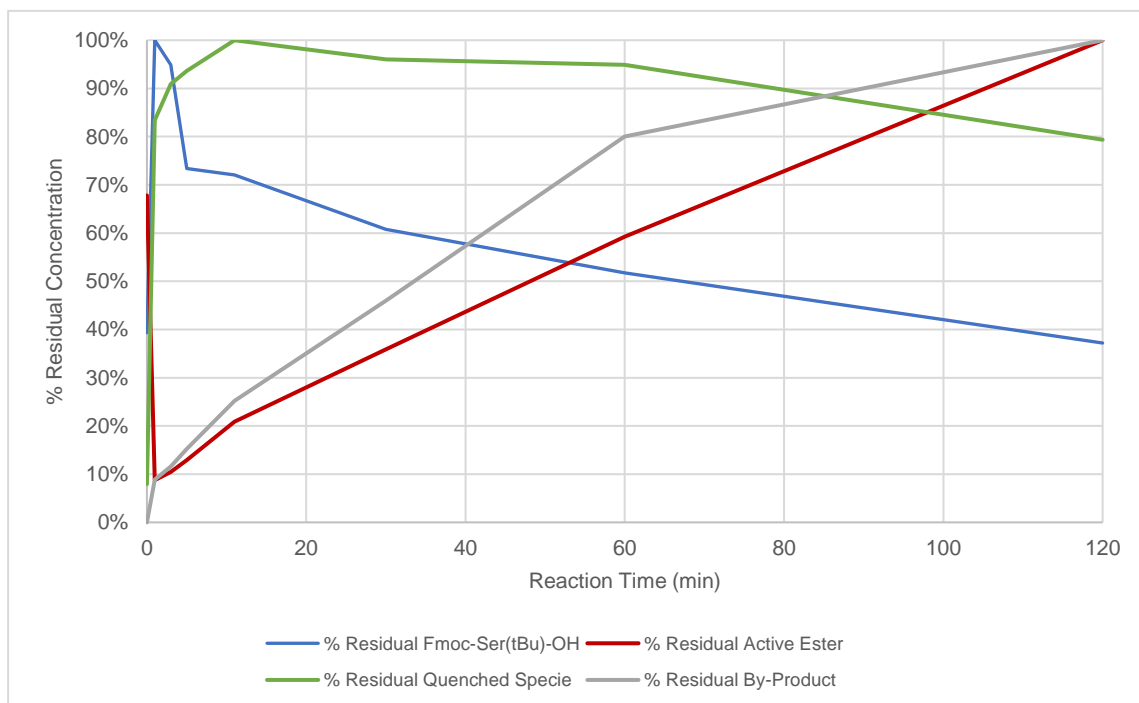
Figure 4.2 - Graph of the evolution of the concentrations of the species (5 equiv. Fmoc-Ser(tBu)-OH and 2.4 equiv. piperidine) during quenching performance.

Ideally, in a controlled reaction, the start material sample would show a maximum of active ester and none quenched specie. During the reaction, the amino acid and active ester concentration would initial decrease and then, stabilize. The quenched specie should increase over time until stabilize. The Figure 4.1 and Figure 4.2 shows clearly an uncontrolled reaction. A fourth specie was detected in the UHPLC, with a molecular weight verified in the mass spectrum of respectability 687.2 g/mol and 611.2 g/mol for Fmoc-Phe-OH and Fmoc-Ser(tBu)-OH, respectively. As expected, the concentration of the unreacted amino acid decreased over time, however, the concentration of the active ester shows an unexpected behaviour, with the increasing concentration over the reaction time.

To mitigate these errors, new amounts were tested, being chosen according to the literature. Carpino and Han demonstrated that one of the key factors to suppress side reactions is an excessive amount of piperidine[47]. So, the piperidine equivalents were increased to **6** (exceeding the amino acids).

The results obtained are presented in Figure 4.3 and Figure 4.4. Unlike the previous one, by-products were not detected, whereby the side reactions were more controlled. However, the concentration evolution of the active ester still shows a not expected behaviour, of increasing over the reaction time instead of decrease (as expected). Also, the quenched specie presents a decrease more accented than the projected, meaning that this specie is, probably, reacting.
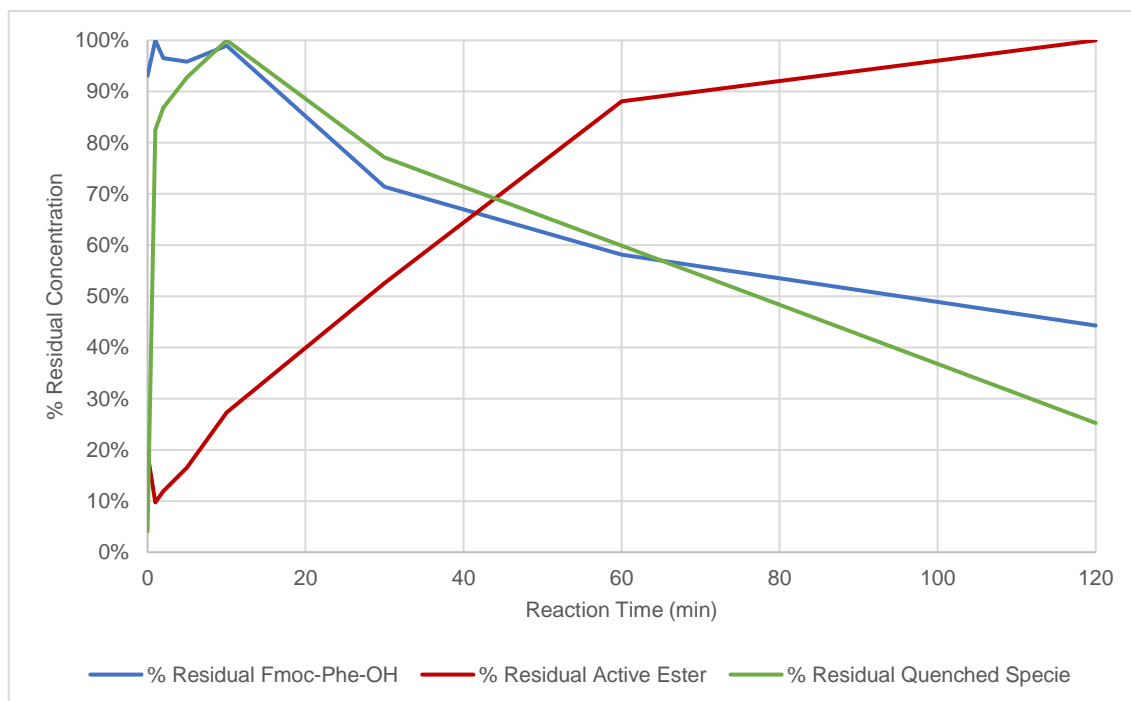
Figure 4.3 - Graph of the evolution of the concentrations of the species (5 equiv. Fmoc-Phe-OH and 6 equiv. piperidine) during quenching performance.
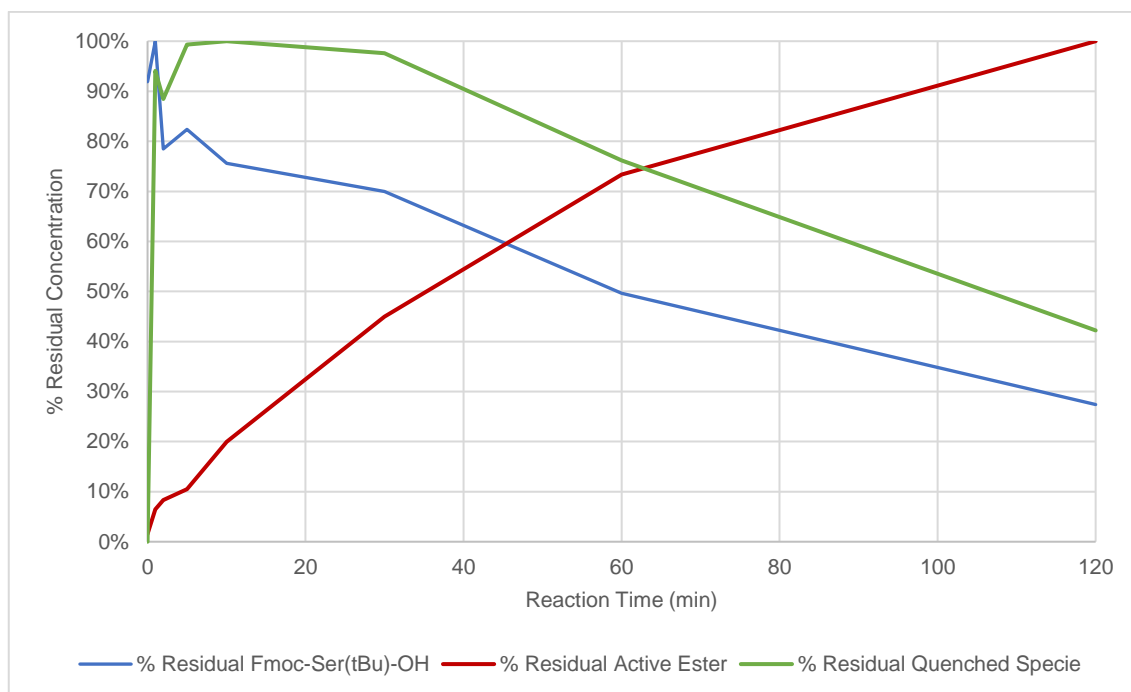
Figure 4.4 - Graph of the evolution of the concentrations of the species (5 equiv. Fmoc-Ser(tBu)-OH and 6 equiv. piperidine) during quenching performance.

To eradicate the faults presented in the previous experiment, the amount of amino acid, HOBt, DIC and quencher reagent were reduced, by reducing the equivalents of the same. The **3** and **2** equivalents of amino acid were tested with **2.1** piperidine equivalents and it was only tested in the Fmoc-Phe-OH amino acid. The results are presented in Figure 4.5 and Figure 4.6. The evolution of the amount of amino acid, HOBt, DIC and piperidine during this study is summarised in Table 4.1.

Table 4.1 - Quenching performance conditions evolution during analysis.

	First		Second		Third		Fourth	
	Equiv.	mmol	Equiv.	mmol	Equiv.	mmol	Equiv.	mmol
Amino Acid	5	0.472	5	0.472	3	0.283	2	0.189
HOBt	6.25	0.590	6.25	0.590	3.75	0.354	2.5	0.236
DIC	5	0.472	5	0.472	3	0.283	2	0.189
Piperidine	2.4	0.226	6	0.566	2.1*	0.198	2.1*	0.198

*It should be noted that the amount of active ester formed remained within ~50% from the AA thus the quenching reagent is still in excess toward the active ester. It should be also noted that a big excess of quenching base would lead to AA deprotection thus the quenching reagent concentration window is limited.
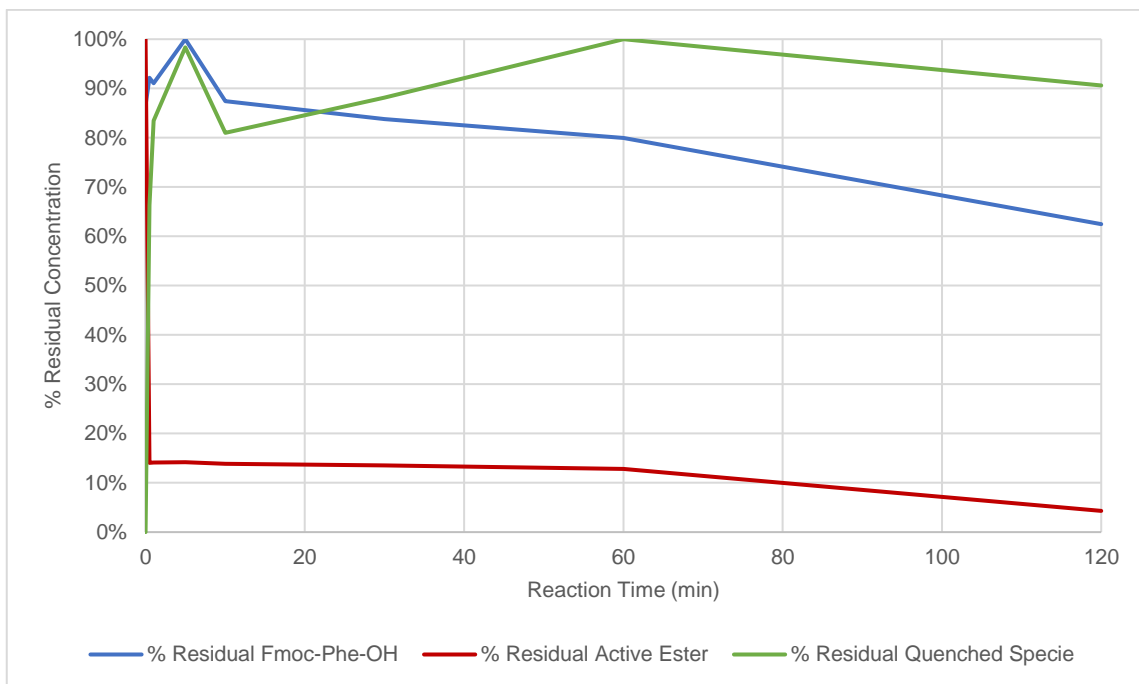
Figure 4.5 - Graph of the evolution of the concentrations of the species (3 equiv. Fmoc-Phe-OH and 2.1 equiv. piperidine) during quenching performance.
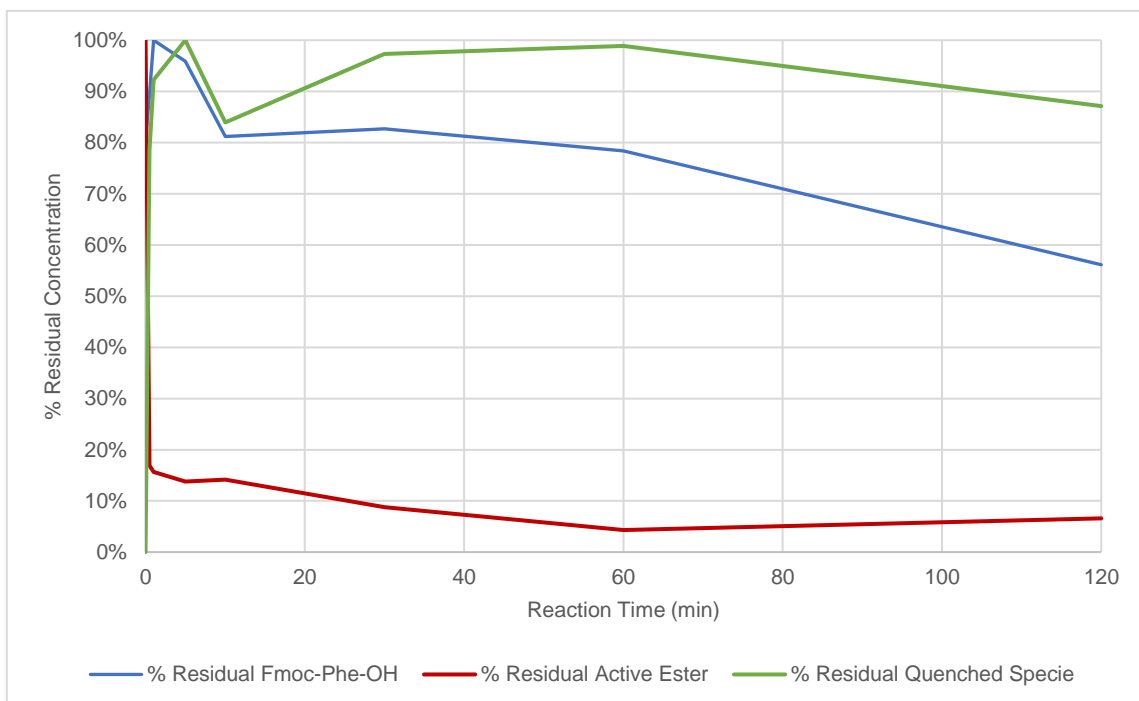
Figure 4.6 - Graph of the evolution of the concentrations of the species (2 equiv. Fmoc-Phe-OH and 2.1 equiv. piperidine) during quenching performance.

As it can be verified in Figure 4.5 and Figure 4.6, the last two equivalent amounts presented good results: the decreasing of the amino acid over the reaction time; the maximum of active ester in start material (0 minutes, before piperidine), a decrease and, lastly, stable over the reaction time; the increasing of the quenched specie (even than in the end it present a lightly decrease). The side reactions were suppressed, and none by-product was detected.

After determining the optimal conditions of the quenching performance (2 and 3 amino acid equivalents for 2.1 equivalents of quenching reagent) the reactions were replicated with aniline and with thiomalic acid, Figure 4.7 - Figure 4.10.

The aniline behaviour illustrated in Figure 4.7 and Figure 4.8 can be explained by it presented to be a very slower quencher. Therefore, the quenched specie takes more time to achieve their maximum value and the active ester consumption is slower.

The thiomalic acid, Figure 4.9 and Figure 4.10, presents a similar behaviour to piperidine, a decreasing followed by a stabilization of the active ester and a very fast increase of the quenched specie. Indeed, the half-time quenching reaction was calculated to these two coupling reagents (piperidine and thiomalic acid).
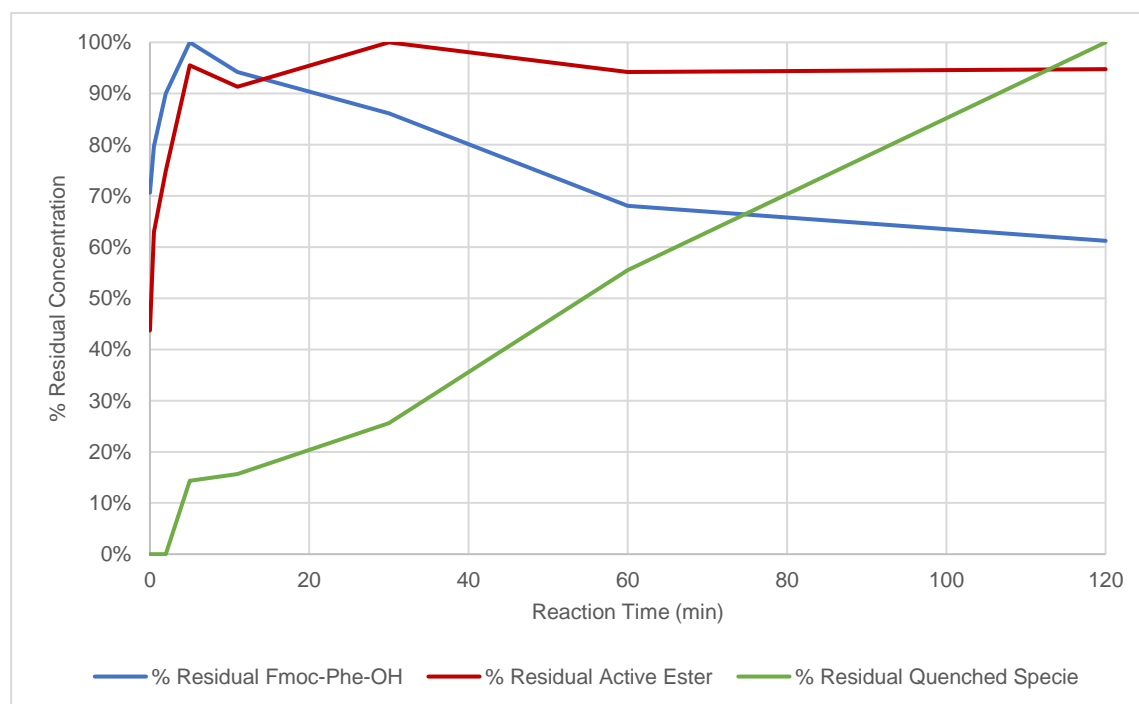
Figure 4.7 - Graph of the evolution of the concentrations of the species (3 equiv. Fmoc-Phe-OH and 2.1 equiv. aniline) during quenching performance.
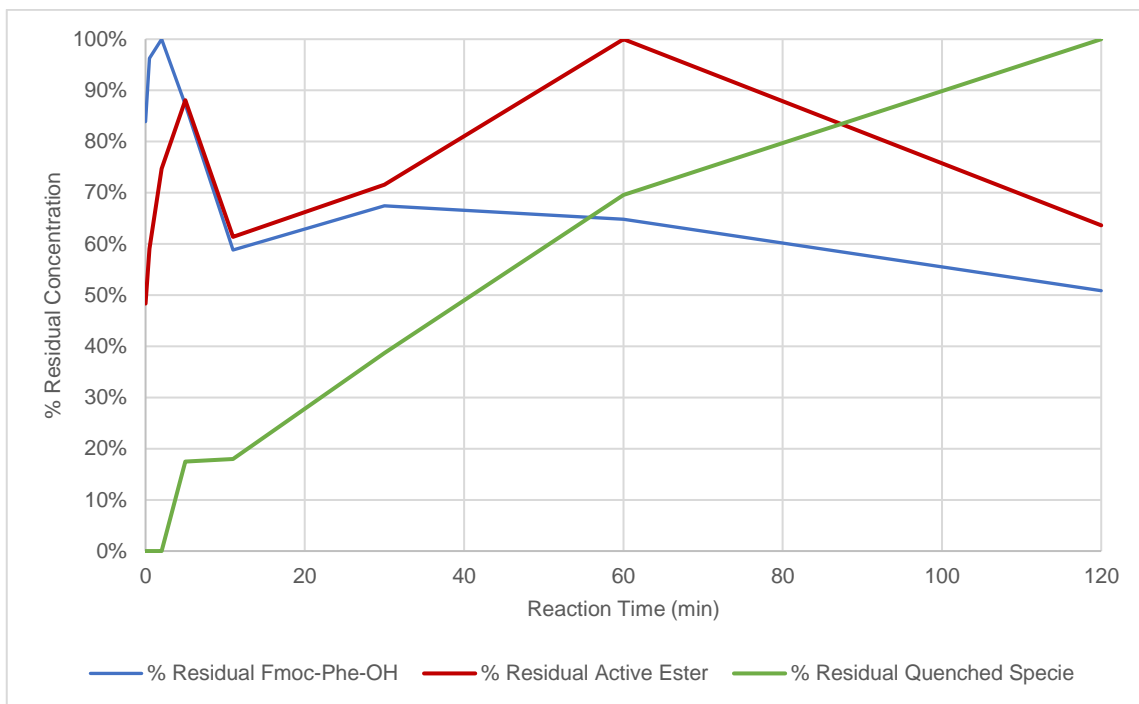
Figure 4.8 - Graph of the evolution of the concentrations of the species (2 equiv. Fmoc-Phe-OH and 2.1 equiv. aniline) during quenching performance.
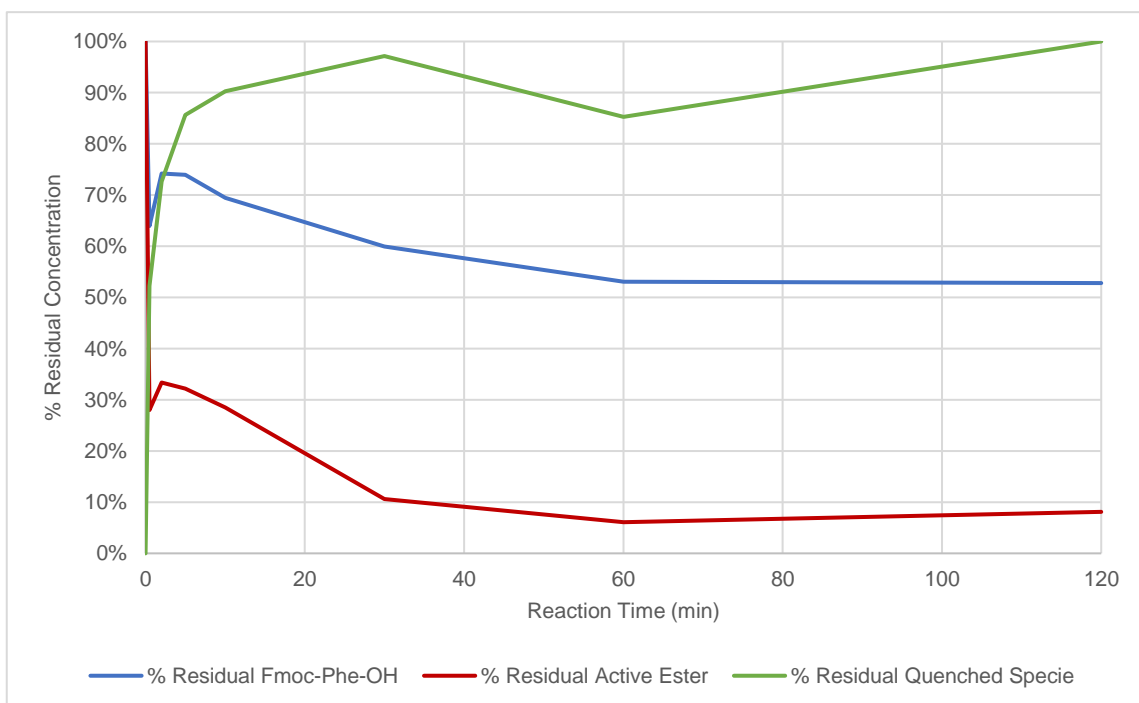
Figure 4.9 - Graph of the evolution of the concentrations of the species (3 equiv. Fmoc-Phe-OH and 2.1 equiv. thiomalic acid) during quenching performance.
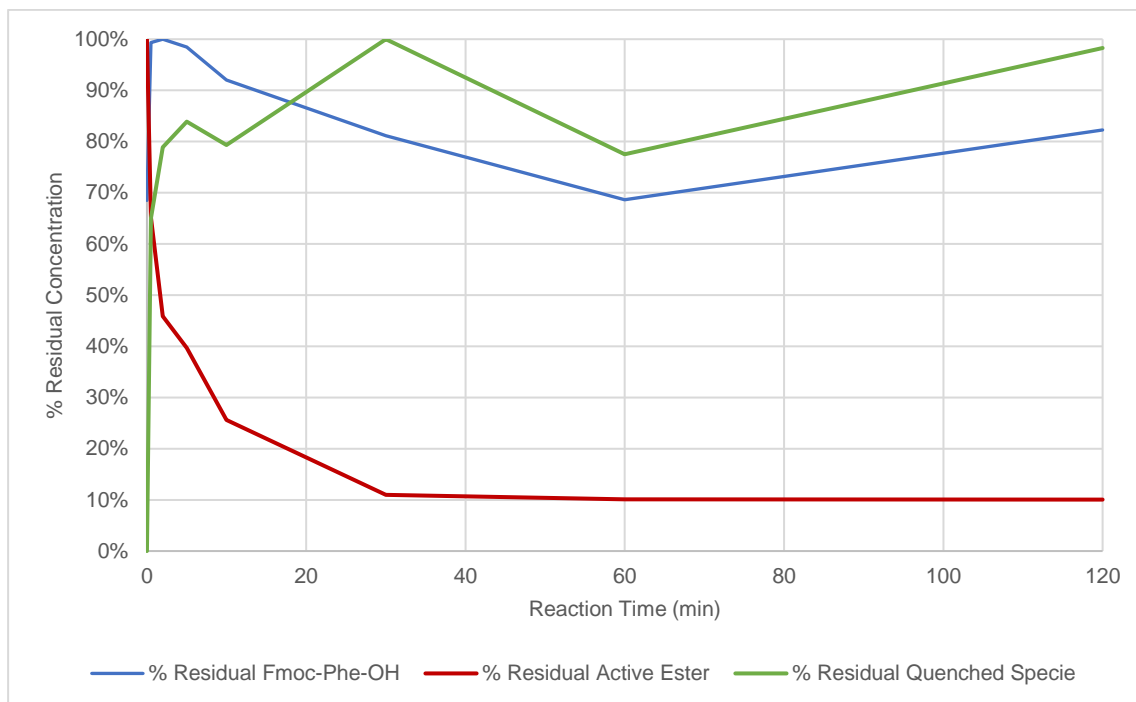
Figure 4.10 - Graph of the evolution of the concentrations of the species (2 equiv. Fmoc-Phe-OH and 2.1 equiv. thiomalic acid) during quenching performance.

In the 2 equivalents of Fmoc-Phe-OH, the active ester reached 50% of their initial concentration in 18 seconds with piperidine and in 100 seconds with thiomalic acid. A smaller time difference was founded in the 3 equivalents of Fmoc-Phe-OH, with a $t_{1/2}$ of 17.45 seconds for piperidine and a $t_{1/2}$ of 21 seconds for thiomalic acid. In both amino acid concentrations, piperidine is faster to complete the 50% of the quenching reactions, being that piperidine show almost the same time in the two performances. Therefore, it is expected that piperidine presents the better quenching rate constant.

Overall, the species concentrations profiles are not faultless. This can occur due to undesired side reactions, contaminated samples, or experimental errors (namely in measurements). Also, some limitations were taked into account as detailed further in section 4.4.3. After the amino acid quenching performance was realized for the three quenchers reagents, the kinetic models were determined.

4.3. Kinetic modelling

4.3.1. Kinetic model of quenched specie formation

The kinetic model and the rate constant, k_q, of quenched specie formation for the different quenchers was deduced by the models described in Chapter 3, section 3.3. Figure 4.11 and Figure 4.12 shows the results for 3 amino acid equivalents and 2 amino acid equivalents, respectively, and Table 4.2 presents the quenching rate constants.
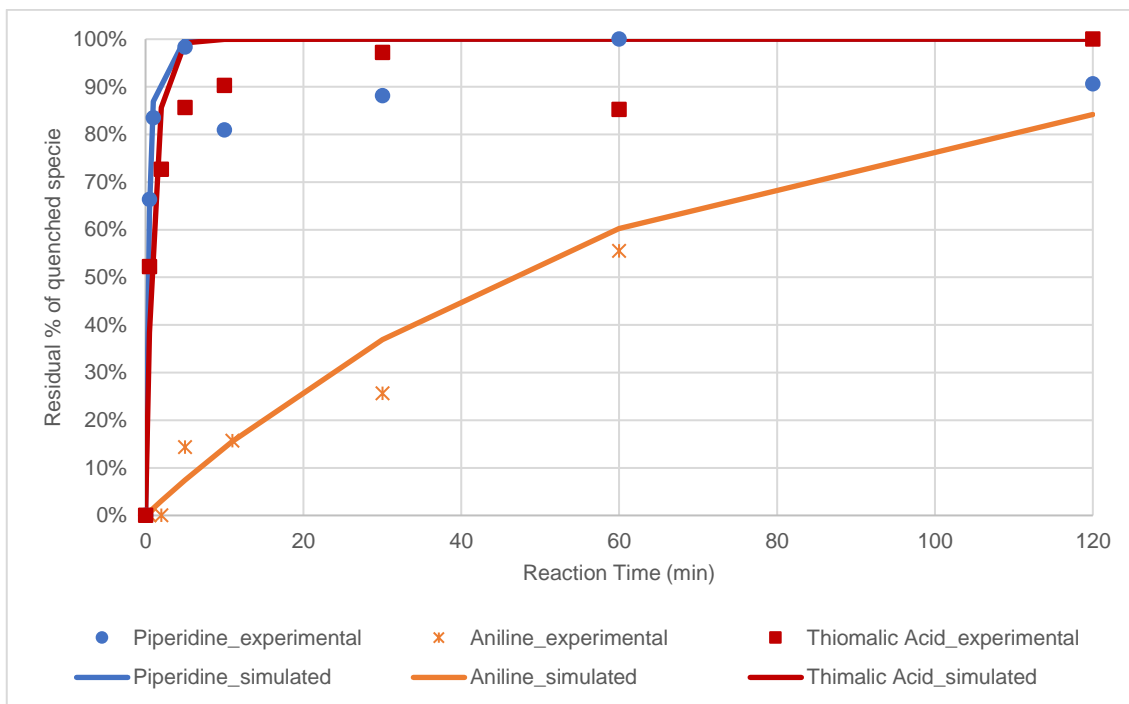
Figure 4.11 - A kinetic model of quenched specie formation during the 3 equiv. Fmoc-Phe-OH quenching reaction.
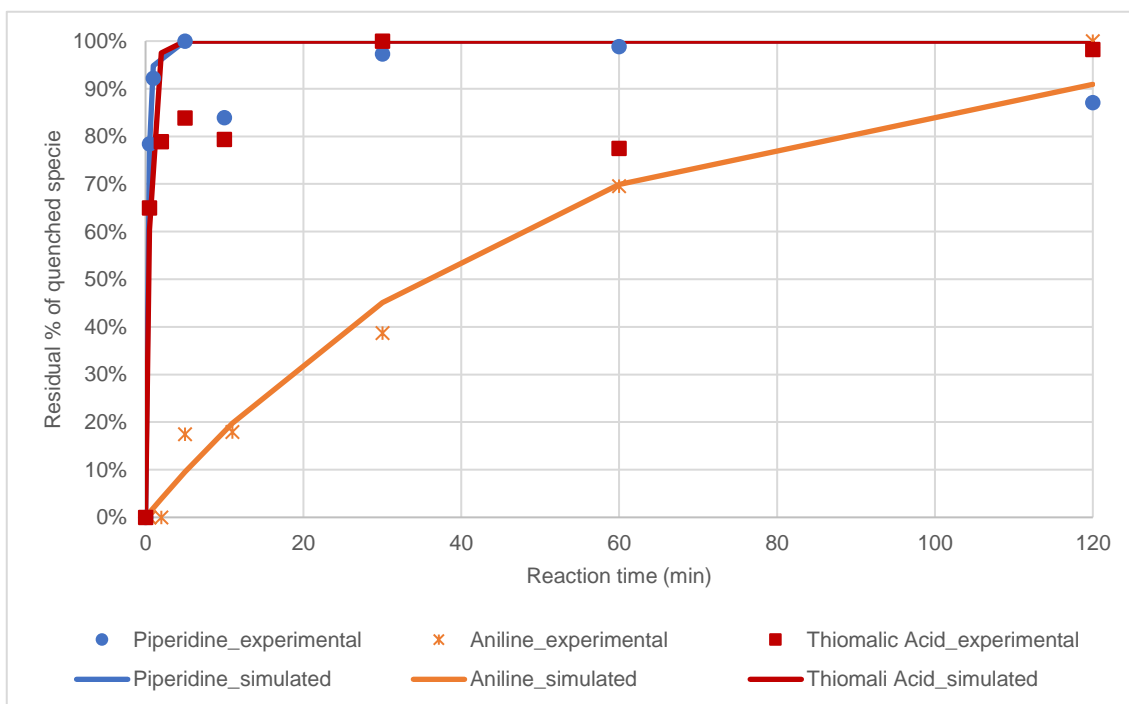
Figure 4.12 - A kinetic model of quenched specie formation during the 2 equiv. Fmoc-Phe-OH quenching reaction.

Table 4.2 - Quenching rate constant (k_q).

Quencher	K_q (min^{-1}) (3 Equiv.)	K_q (min^{-1}) (2 Equiv.)
Piperidine	2.0340	2.9615
Aniline	0.0154	0.0200
Thiomalic Acid	0.9696	1.8486

In both cases, piperidine shows the best constant rate, followed by thiomalic acid and, for last, by aniline. The piperidine completed the reaction within 5 minutes in the two scenarios, same as thiomalic acid with 2 amino acids equivalents scenario. In the 3 amino acid equivalents, the thiomalic acid completed the reaction in 10 minutes, not far from the first ones. Of note, aniline did not finish the reaction within 120 minutes, showing itself as the slower quencher.

4.3.2. Kinetic model of amino acid reduction

The kinetic model and the rate constant, k_a, of amino acid reduction for the different quenchers was deduced by the models described in Chapter 3, section 3.3. Figure 4.13 and Figure 4.14 show the results for 3 amino acid equivalents and 2 amino acid equivalents, respectively, and Table 4.3 presents the amino acid reduction rate constants.
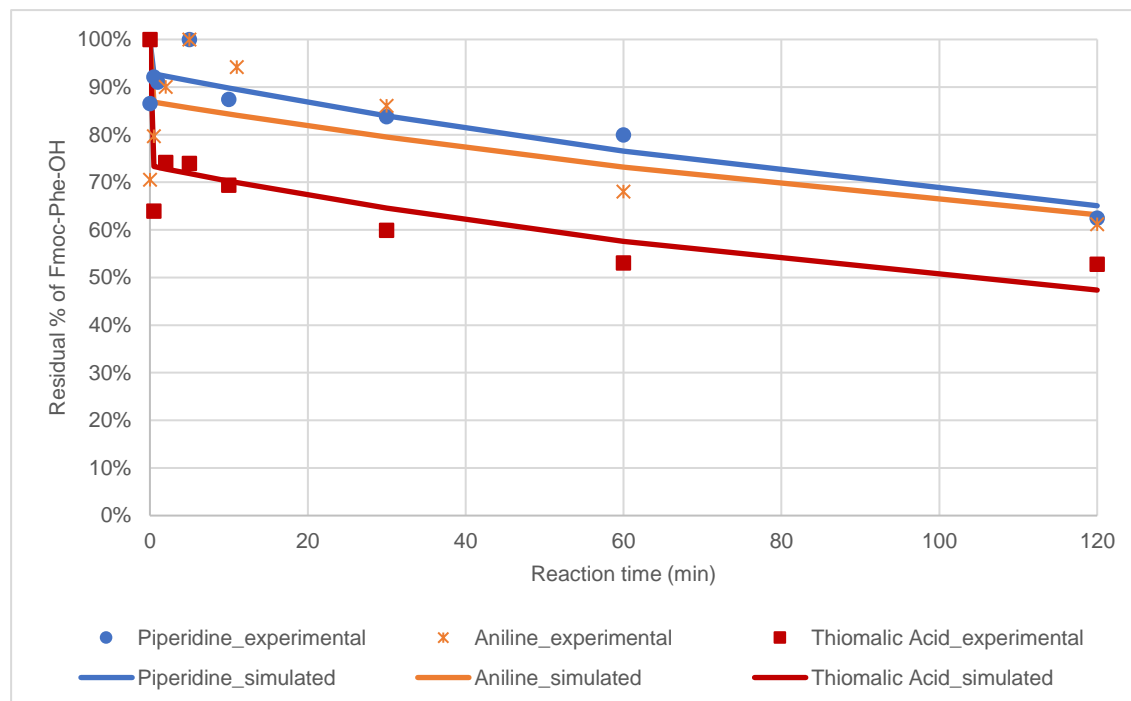
Figure 4.13 - A kinetic model of Fmoc-Phe-OH reduction during the 3 equiv. Fmoc-Phe-OH quenching reaction.
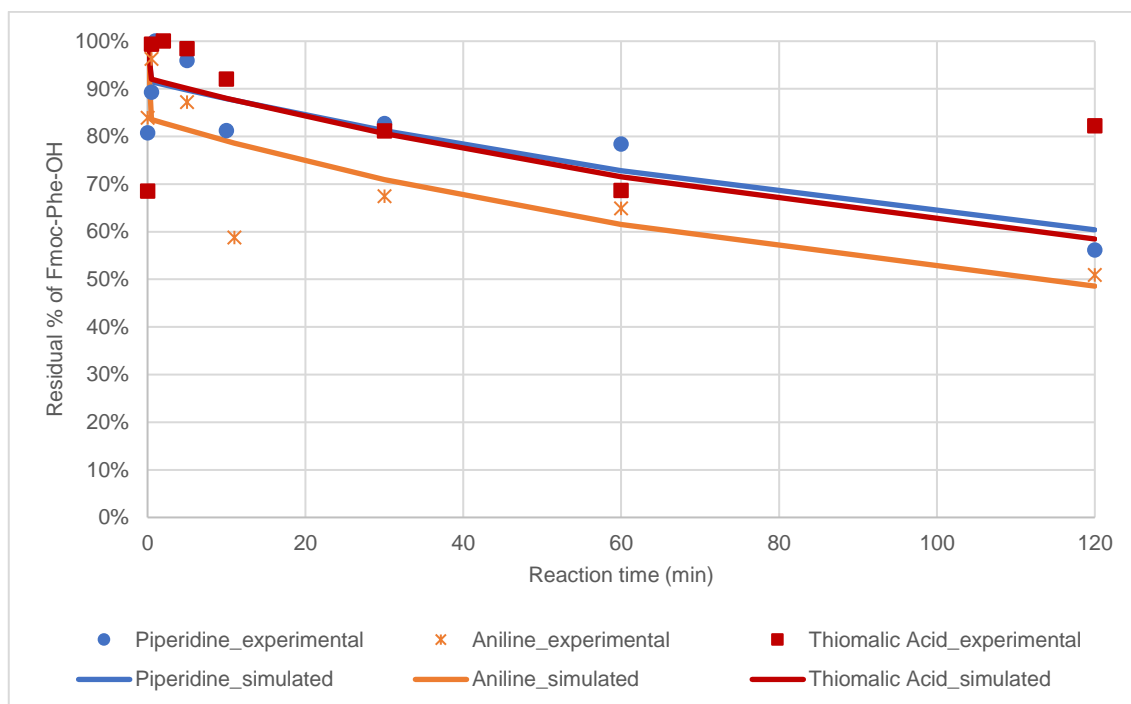
Figure 4.14 - A kinetic model of Fmoc-Phe-OH reduction during the 2 equiv. Fmoc-Phe-OH quenching reaction

Table 4.3 – Amino acid reduction rate constant (k_a).

Quencher	K_a (mmol.min^{-1}) (3 Equiv.)	K_a (mmol.min^{-1}) (2 Equiv.)
Piperidine	0.0136	0.0250
Aniline	0.0128	0.0383
Thiomalic Acid	0.0221	0.0277

Analysing Figure 4.13 (3 equivalents), at the end of the reaction the piperidine scenario shows the biggest concentration of amino acid not converted, followed by aniline and thiomalic acid respectively. This can be explained since the conversion of active ester in quenched specie is faster than the amino acid reduction. The same behaviour can be observed in the 2 equivalents scenario, Figure 4.14.

According, the thiomalic acid follow the piperidine in Figure 4.14. However, in Figure 4.13, more amino acid has converted in the thiomalic acid scenario, that may be due to side reactions occurred.

4.4. Factors affecting the quenching rate constant

Factors as polarity (TPSA), hydrophobicity (LogP), pH (pKa), time of reaction, and molecular weight of the quenching reagent can influence the quenching reaction, therefore affects the quenching constant rate (k_q). It would be interesting to understand if exist a correlation between each factor, illustrated in Table 4.4, and the k_q. However, due to it only was possible done few experiments, there is not enough data to prove these correlations.

Table 4.4 – Quenching reagents properties.

	pKa	LopP	TPSA	Molecular Weight
Piperidine	11.12	0.84	12 Å²	85.15
Aniline	4.60	0.9	26 Å²	93.13
Thiomalic Acid	4.68	-0.46	75.6 Å²	150.16

The following graphs, Figure 4.15 - Figure 4.18, were designed with the k_q of the reaction of 2 equivalents of Fmoc-Phe-Oh and 2.1 equivalents of quenching reagent.

As observed in the graphs, Figure 4.15 - Figure 4.17, no correlation between LogP, TPSA or molecular weight and quenching rate constant was found, once all R^2 are nearly zero, Table 4.5. Nevertheless, it is impossible to conclude that correlations do not exist with the limited data accessed.
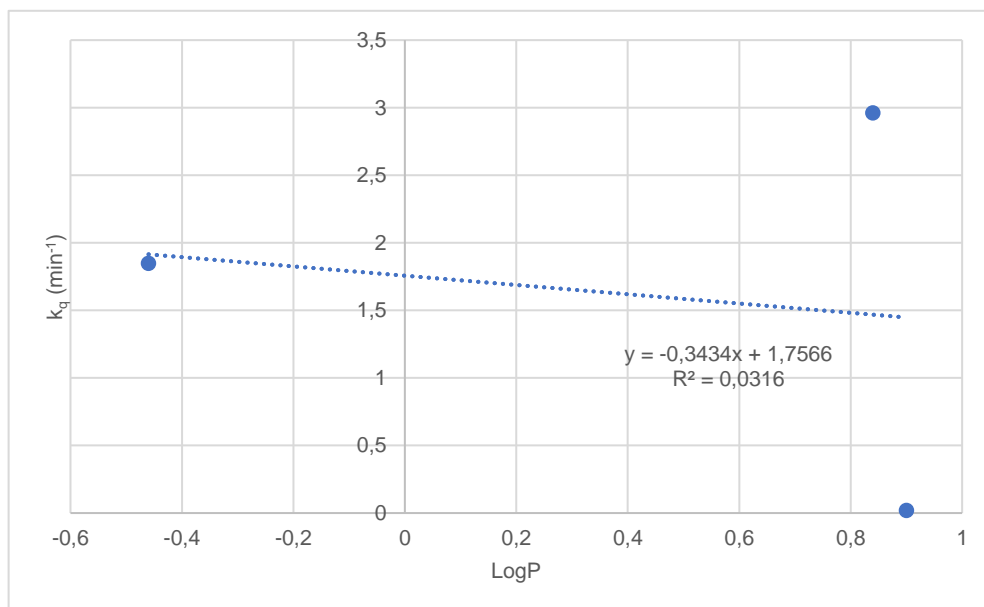
Figure 4.15 – Graph of k_q as function of quenching reagents LogP.
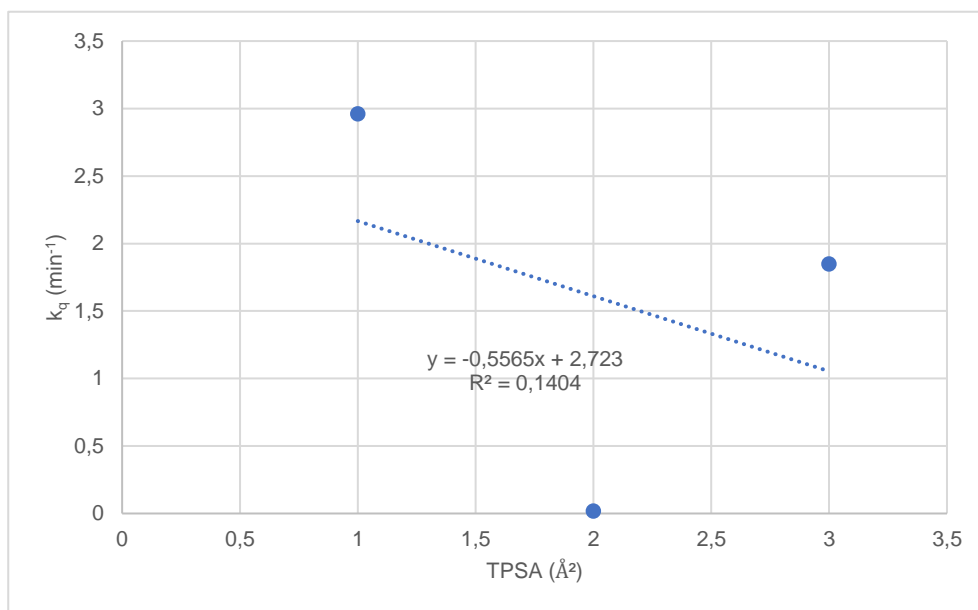
Figure 4.16 - Graph of k_q as function of quenching reagents TPSA.
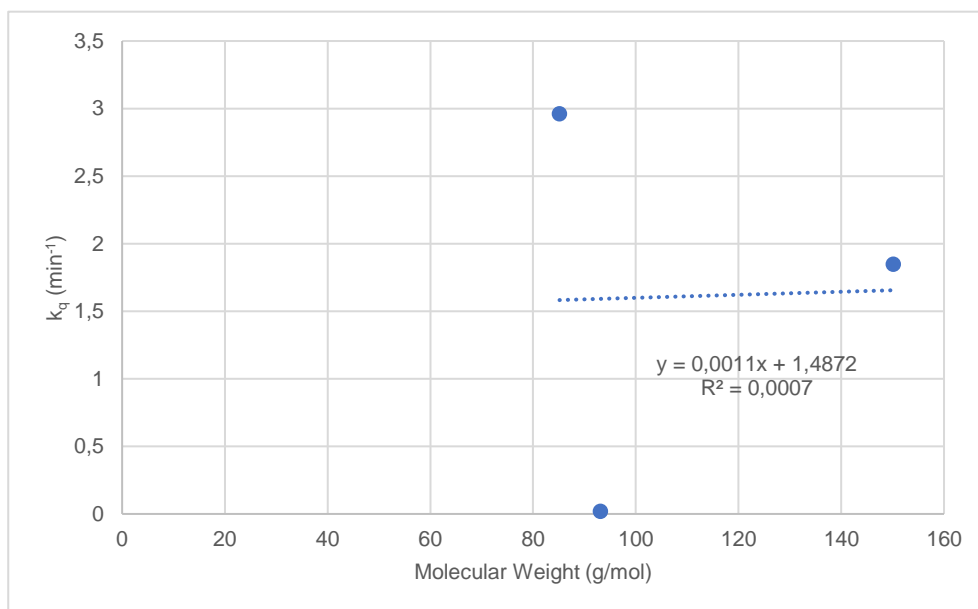
Figure 4.17 - Graph of k_q as function of quenching reagents molecular weight.
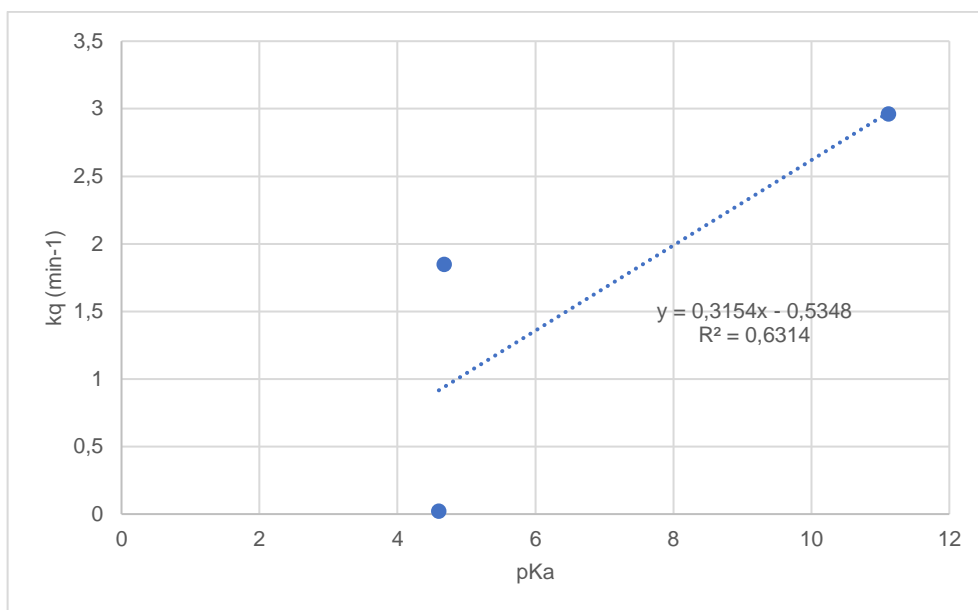
Figure 4.18 - Graph of k_q as function of quenching reagents pKa.

Table 4.5 - Linear regressions of the different correlations.

	Linear Regression	R^2
LogP vs. kq	-0.3434x + 1.7566	0.0316
TPSA vs. kq	-0.5565x + 2.723	0.1404
Molecular Weight vs. kq	0.0011x + 1.4872	0.0007
pKa vs. kq	0.3154x - 0.5348	0.6314

Figure 4.18 shows a possible linear relation between pKa and k_q, that can be observed by an increase of the quenching rate constant when the reagent pKa was increased, meaning that the quenching reaction benefits in a more basic reagent. The data collected cannot confirm this correlation and it would be necessary to do more trials. In the future, more experiments should be done to conclude if and how these factors affect the quenching performance. Also, for a more complete analysis, the quenching reaction should be made in different amino acids (in this case, it was only made in Fmoc-Phe-OH) to detect how the amino acid and the correspondent active esters properties can also affect the quenching rate.

4.4.1. Quenching reagents solubility

The solubility of the quenching reagent in the solvent 65 NMP: 35 THF, was predicted in the *COSMOtherm* software, with the methodology described in Chapter 3, section 3.4, to

understand if a relation between the reagent solubility and the quenching reaction efficiency. The solubility results are presented in Table 4.6.

Table 4.6 - Quenching reagents solubility.

Reagent	Solvent	S (mol/l)
Aniline	65 NMP : 35 THF	3.793
Thiomalic Acid	65 NMP : 35 THF	3.756
Piperidine	65 NMP : 35 THF	3.054

All reagents presented an excellent solubility value in the reaction solvent, what is essential because if the reagent is not soluble the quenching is not possible. Even aniline, that is determined the slower quencher, presents a great solubility. Whereby, it seems that with those three quenchers, there is no relation between the solubility and the k_q.

4.4.2. Amino acid Concentration

Since the quenching reaction was performed in two different Fmoc-Phe-OH concentrations, it was compared to the quenching rate constant k_q obtained in each one, Table 4.7.

When compared the results in Table 4.7, it can be detected that a lower Fmoc-Phe-OH concentration (2 equivalents) obtained better quenching rate constant results for all reagents. One of the reasons for this happens is that a lower amino acid concentration allows a more controlled reaction and the side reactions were suppressed. The performance with an excess of amino acid translates into a performance with a bigger amount of active ester, whereby is more favourable to occur side reactions, that will decrease the quenching reaction efficiency. Although there are only two concentrations to be compared, it ascertained that bigger amounts of the quenching reagent are beneficial for the reaction.

Table 4.7 - Quenching rate constants (k_q) for different Fmoc-Phe-OH concentrations.

Quencher	K_q (min^{-1}) (3 Equiv.)	K_q (min^{-1}) (2 Equiv.)
Piperidine	2.0340	2.9615
Aniline	0.0154	0.0200
Thiomalic Acid	0.9696	1.8486

4.4.3. Limitations

This analysis presented a few limitations that need to be mentioned:

1 – The data collected were few. It allows choosing the best reagent between the three quenching reagents proposed. However, makes it impossible to ascertain the relation between the reagent properties and the quenching performance.

2 – It was assumed that the quenching reaction would be completed in two hours and the maximum of the quenched specie would be achieved in those two hours. However, the quenched specie could be continued to be formed after the two hours. The clearest case where this probably happened was in the aniline quenching reaction. Therefore, the product concentration at two ours is not corresponded to the 100% of formation.

3 – As referred in Chapter 3, the maximum of the quenched specie concentration would be, ideally, at two hours but it does not correspond with some results (specifically, in the piperidine analysis). The product concentration decreases probably due to uncontrolled deprotections reactions. However, it was assumed that the concentration of the quenched specie was only resultant of the quenching reaction.

4.5. Conclusion

To conclude this chapter, the study of quenching condition was carried out successfully and the reactions conducted with 2 and 3 equivalents of Fmoc-Phe-OH and 2.1 equivalents of reagent proved to be the best option conditions. From there, the quenching rate constant (k_q) and the amino acid reduction rate constant (k_a) was determined for both conditions with the three reagents: piperidine, aniline, and thiomalic acid. The values obtained for k_a were 0.025, 0.028, and 0.038 min^{-1} and for k_q were 2.962, 1.849, and 0.020 min^{-1}, corresponding to piperidine, thiomalic acid and aniline, respectively. Surely, it can be affirmed that piperidine (that also presented the shorter half time reactions) showed to be the best quencher for Fmoc-Phe-Oh, followed by thiomalic acid and aniline, respectively. Also, it was carried out a study about the quenching reagent properties and their effect in the k_q. However, due to the lack of data, the only correlation that was possible to identify, but still needs to be confirmed posteriorly, was with the pKa. The amino acid concentration exhibited to be also an important factor, whereby it would be interesting to study different amino acid and comprehend the effect of their properties. Therefore, all the proposed objectives in section 4.1 were achieved.

5. Results and discussion: Peptide solubility Estimations

5.1. Introduction

During the peptide synthesis process, one of the major factors to take in consideration during all the process is the solubility. Both the solubility of the reagents as the solubility of the products need to be considered once that they can have a big influence in the process yield. This because, as mentioned, the presence of solid particles will increase the mass resistance and, consequently, decrease the efficiency of the process. In this project, the focus will be the solubility of the peptides. Usually, the solubility of new peptides is experimental tested in different solvents. However, this methodology requires an experimental effort, is very time consuming and produce a lot of waste, making the global cost of peptide synthesis increasing. Solubility can be predicted by solubility parameters, as Hildebrand and Hansen Parameters, by the premiss that ''*like dissolves like*'' or by COSMO-RS calculations, through the software's enounced in Chapter 3.

Therefore, the solubility parameters of 4 different free amino acids (Fmoc-Gly-OH, Fmoc-Ser(tBu)-OH, Fmoc-Tyr(tBU), and Fmoc-Arg(Pbf)-OH) and their respective peptides (dimer, 5 mer, 10 mer, and 20 mer) will be predicted and compared to the solubility parameters of several solvents to choose the best and worst solvents. The amino acids were selected to represent different types of amino acids, as following: Gly – apolar; Ser(tBu) – uncharged polar; Arg(Pbf) – charged polar; Tyr(tBu) – uncharged polar.

Also, the solubility of the same amino acids and peptides will be predicted by the COSMO-RS approach. The best and worst solvents for each one will be selected, and some will be tested experimentally to observe the accuracy of the projections made. The same analysis will be made with the amino acids attached to the hub by a wang linker and a rink amide linker.

To summarize, the main objectives of this chapter are:
1. Estimate the Hildebrand and Hansen solubility parameters to the free peptides and to the peptides attached to the hub (wang nanostar and rink amide nanostar).
2. Estimate the solubility of the free peptides and to the peptides attached to the hub (wang nanostar and rink amide nanostar) by COSMO-RS approach.
3. Predict the best and worst solvents for the peptide by both methods.
4. Test the solvents experimentally.
5. Determine the accuracy of the estimations.
6. Identify the effect of the hub (by different linkers) and the Fmoc affects the peptide solubility.

5.2. Hildebrand and Hansen solubility parameters estimation

This study started with the estimation of the solubility parameters. Hildebrand and Hansen Parameters were estimated by different software's, as stated in Chapter 3, section 3.4. To start, it was also necessary to estimate the solubility parameters, Hildebrand and Hansen, of the solvents, present in Table 5.1.

Table 5.1 - Hildebrand and Hansen Parameters values of the solvents.

	δ_D (MPa$^{1/2}$)	δ_P (MPa$^{1/2}$)	δ_H (MPa$^{1/2}$)	δ (MPa$^{1/2}$)
THF	17.056	6.075	6.910	18.800
MeTHF	16.992	6.008	5.325	18.793
NMP	18.238	12.553	7.742	23.455
DMSO	17.802	16.557	10.367	20.900
GVL	18.487	14.173	7.926	24.607
PC	19.246	17.021	5.220	26.218
NFM	18.189	13.556	10.619	23.900
NBP	16.619	4.581	4.632	17.850
Butanol	15.902	5.815	15.620	18.000

5.2.1. Amino acids and free peptides estimations

The range of amino acids tested were: glycine (Fmoc-Gly-OH); protected serine (Fmoc-Ser(tBu)-OH); protected tyrosine (Fmoc-Tyr(tBu)-OH); and Protected Arginine (Fmoc-Arg(Pbf)-OH). Also, the respective peptides of each amino acid were tested was dimer, 5 mer, 10 mer and 20 mer. The results obtained are presented in the following Table 5.2. For some peptides, namely, serine (10 and 20 mer) and tyrosine (10 and 20 mer) it was not possible to estimate the Hansen Parameter value due to a software error. For glycine, the machine learning *Polymer Genome* could not recognize the amino acid, so the Hildebrand Parameter was calculated by Eq. 2.4. Both problems happen to arginine, whereby arginine is not present in this estimation. The amino acids and peptides structures used in the estimations are presented in Appendix II. The obtained results were then applied in the Hildebrand and Hansen Model. The Hildebrand Model affirms that a solvent and a peptide with a Hildebrand parameter difference value within 2 MPa$^{1/2}$ is a good solvent and the Hansen Model supports that a peptide and a solvent with a R_a within 8 MPa$^{1/2}$ is a good solvent, otherwise the solvent is considered a non-solvent for that peptide.

$$\left| \delta_{pept} - \delta_{solv} \right| < 2 \qquad (Eq.\,5.1)$$

$$Ra = \sqrt{4(\delta_{D1} - \delta_{D2})^2 + (\delta_{P1} - \delta_{P2})^2 + (\delta_{H1} - \delta_{H2})^2} < 8 \qquad (Eq.\,5.2)$$

Table 5.2 - Hildebrand and Hansen Parameters values of the (protected) free peptides.

	δ_D (MPa$^{1/2}$)	δ_P (MPa$^{1/2}$)	δ_H (MPa$^{1/2}$)	δ (MPa$^{1/2}$)
Fmoc-Gly-OH	19.29	7.70	13.98	25.03
Fmoc-Gly-OH (dimer)	18.75	8.34	20.34	28.89
Fmoc-Gly-OH (5 mer)	18.69	8.37	20.56	29.02
Fmoc-Gly-OH (10 mer)	18.14	4.70	24.39	30.75
Fmoc-Gly-OH (20 mer)	18.12	4.61	25.03	31.24
Fmoc-Ser(tBu)-OH	18.75	8.70	11.72	20.60
Fmoc-Ser(tBu)-OH (dimer)	17.83	8.63	20.02	20.70
Fmoc-Ser(tBu)-OH (5 mer)	17.61	8.01	13.02	21.60
Fmoc-Ser(tBu)-OH (10 mer)	-	-	-	22.00
Fmoc-Ser(tBu)-OH (20 mer)	-	-	-	21.60
Fmoc-Tyr(tBu)-OH	18.89	8.02	13.76	21.20
Fmoc-Tyr(tBu)-OH (dimer)	18.38	7.54	11.46	21.50
Fmoc-Tyr(tBu)-OH (5 mer)	18.08	8.08	14.37	22.10
Fmoc-Tyr(tBu)-OH (10 mer)	-	-	-	22.10
Fmoc-Tyr(tBu)-OH (20 mer)	-	-	-	18.50

The calculations made for both models are presented in Appendix III. From the solvents within the model's criteria, the best two were selected as best solvents and, from the solvents out of criteria, the worst two were selected as the worst solvents. The solvents selected for the amino acids and corresponding peptides by the Hildebrand and Hansen models are announced in Table 5.3 and Table 5.4, respectively.

Table 5.3 and Table 5.4 revealed a clear difference in the solvent's choice for each peptide. The cells that do not present any solvent means that none of the solvents is within the criteria. The best solvents obtained by both models present solid differences. While by the Hildebrand Model, the DMSO shows a strong predominance as the best solvent, followed by THF and NMP, for serine and tyrosine, by the Hansen Model, DMSO is not recommended as the best solvent, being their strong predominance substituted by the NFM. In both models, the glycine only presented suitable solvents for the amino acid, and none solvent within the criteria for the peptides (dimer, 5 mer, 10 mer, 20 mer). It is notorious the similarity in the serine and tyrosine results, and it is explained since both present the same polarity (uncharged polar) and, hence, similar σ-profiles.

The worst solvents in both models, although there is a slight difference, are almost the same ones, except in the glycine peptides. In this case, by Hildebrand Model, the butanol solvent

showed to be one of the worst, which was replaced by the PC solvent in the Hansen Model approach. PC and NBP showed a strong domain in the worst solvents in both models.

Table 5.3 - Best and worst solvents for the (protected) free peptides by Hildebrand Model.

	Best Solvents		Worst Solvents	
Fmoc-Gly-OH	GVL	NFM	Butanol	NBP
Fmoc-Gly-OH (dimer)	-	-	Butanol	NBP
Fmoc-Gly-OH (5 mer)	-	-	Butanol	NBP
Fmoc-Gly-OH (10 mer)	-	-	Butanol	NBP
Fmoc-Gly-OH (20 mer)	-	-	Butanol	NBP
Fmoc-Ser(tBu)-OH	DMSO	THF	GVL	PC
Fmoc-Ser(tBu)-OH (dimer)	DMSO	THF	GVL	PC
Fmoc-Ser(tBu)-OH (5 mer)	DMSO	NMP	NBP	PC
Fmoc-Ser(tBu)-OH (10 mer)	DMSO	NMP	NBP	PC
Fmoc-Ser(tBu)-OH (20 mer)	DMSO	NMP	NBP	PC
Fmoc-Tyr(tBu)-OH	DMSO	-	NBP	PC
Fmoc-Tyr(tBu)-OH (dimer)	DMSO	-	NBP	PC
Fmoc-Tyr(tBu)-OH (5 mer)	DMSO	NMP	NBP	PC
Fmoc-Tyr(tBu)-OH (10 mer)	DMSO	NMP	NBP	PC
Fmoc-Tyr(tBu)-OH (20 mer)	THF	MeTHF	GVL	PC

Table 5.4 - Best and worst solvents for the (protected) free peptides by Hansen Model.

	Best Solvents		Worst Solvents	
Fmoc-Gly-OH	NFM	-	PC	NBP
Fmoc-Gly-OH (dimer)	-	-	PC	NBP
Fmoc-Gly-OH (5 mer)	-	-	PC	NBP
Fmoc-Gly-OH (10 mer)	-	-	PC	NBP
Fmoc-Gly-OH (20 mer)	-	-	PC	NBP
Fmoc-Ser(tBu)-OH	NFM	NMP	NBP	PC
Fmoc-Ser(tBu)-OH (dimer)	Butanol	-	NBP	PC
Fmoc-Ser(tBu)-OH (5 mer)	NFM	Butanol	NBP	PC
Fmoc-Tyr(tBu)-OH	NMP	NFM	NBP	PC
Fmoc-Tyr(tBu)-OH (dimer)	THF	NFM	PC	-
Fmoc-Tyr(tBu)-OH (5 mer)	Butanol	NFM	NBP	PC

The growth of the peptide chain can impact in the peptide solubility and modify the solvents choice. By the Hildebrand Model, this can be observed from glycine amino acid to peptide dimer, from serine peptide dimer to 5 mer, and from tyrosine peptide dimer to 5 mer and 10 mer to 20 mer. By the Hansen Model, with the increasing of peptide, there are constantly an alteration in one of the best solvents. This factor should be considerate in the final solvent choice and a good solvent for all products preferred.

5.2.2. Peptides attached to the Hub – Wang nanostar

The estimations process was repeated to the peptides attached to a wang linker nanostar hub. However, as mentioned above, due to a system error, the Hansen solubility parameter was not possible to estimate. Therefore, this analysis was made only with the Hildebrand solubility parameter by *Polymer Genome*. An example of a peptide (protected tyrosine 5 mer) attached to wang nanostar hub is presented in Appendix II, Figure II.5.

Table 5.5 - Hildebrand Parameter value for the peptides attached to wang nanostar.

	δ (MPa$^{1/2}$)
Fmoc-Ser(tBu)-OH	19.9
Fmoc-Ser(tBu)-OH (dimer)	20.2
Fmoc-Ser(tBu)-OH (5 mer)	20.8
Fmoc-Ser(tBu)-OH (10 mer)	20.7
Fmoc-Ser(tBu)-OH (20 mer)	17.1
Fmoc-Tyr(tBu)-OH	19.9
Fmoc-Tyr(tBu)-OH (dimer)	20.3
Fmoc-Tyr(tBu)-OH (5 mer)	21.0
Fmoc-Tyr(tBu)-OH (10 mer)	20.0

Table 5.5 shows the values obtained to the Hildebrand solubility parameter for serine and tyrosine amino acids and peptides. The glycine and arginine peptides were not included in this analysis once the *Polymer Genome* did not could recognize thus as input. The Hildebrand Model was again applied and the best and worst solvents for the peptides attached to the wang nanostar hub were discovered. The calculations of the Hildebrand model are, as for the free peptides, in Appendix III and the best and worst solvents results are illustrated in Table 5.6. The results were very consistent since they presented the same best solvents and worst solvents for all peptides studied, except for the serine 20 mer peptide.

Table 5.6 - Best and worst solvents for the peptides attached to wang nanostar by Hildebrand Model

	Best Solvents		Worst Solvents	
Fmoc-Ser(tBu)-OH	DMSO	THF	PC	GVL
Fmoc-Ser(tBu)-OH (dimer)	DMSO	THF	PC	GVL
Fmoc-Ser(tBu)-OH (5 mer)	DMSO	THF	PC	GVL
Fmoc-Ser(tBu)-OH (10 mer)	DMSO	THF	PC	GVL
Fmoc-Ser(tBu)-OH (20 mer)	NBP	THF	PC	GVL
Fmoc-Tyr(tBu)-OH	DMSO	THF	PC	GVL
Fmoc-Tyr(tBu)-OH (dimer)	DMSO	THF	PC	GVL
Fmoc-Tyr(tBu)-OH (5 mer)	DMSO	THF	PC	GVL
Fmoc-Tyr(tBu)-OH (10 mer)	DMSO	THF	PC	GVL

5.2.3. Peptides attached to the Hub – Rink amide nanostar

For the peptides attached to the rink amide nanostar hub, the same estimation process to the peptides attached to the wang nanostar hub was made. The results obtained are illustrated in Table 5.7.

Table 5.7 - Hildebrand Parameter value for the peptides attached to rink amide nanostar

	δ (MPa$^{1/2}$)
Fmoc-Gly-OH	21.7
Fmoc-Gly-OH (dimer)	22.4
Fmoc-Gly-OH (5 mer)	24.0
Fmoc-Gly-OH (10 mer)	25.6
Fmoc-Gly-OH (20 mer)	25.9
Fmoc-Ser(tBu)-OH	21.2
Fmoc-Ser(tBu)-OH (dimer)	21.3
Fmoc-Ser(tBu)-OH (5 mer)	21.3
Fmoc-Ser(tBu)-OH (10 mer)	20.7
Fmoc-Ser(tBu)-OH (20 mer)	16.3
Fmoc-Arg(Pbf)-OH	21.5
Fmoc-Arg(Pbf)-OH (dimer)	21.6
Fmoc-Arg(Pbf)-OH (5 mer)	21.8
Fmoc-Arg(Pbf)-OH (10 mer)	21.7
Fmoc-Tyr(tBu)-OH	21.2
Fmoc-Tyr(tBu)-OH (dimer)	21.3
Fmoc-Tyr(tBu)-OH (5 mer)	21.4
Fmoc-Tyr(tBu)-OH (10 mer)	20.7

Then, the Hildebrand model was applied and the peptide – solvent parameter difference was calculated (Appendix III). The results of the solvents that showed to be the best and worst ones are illustrated in Table 5.8. For the worst solvents, PC, NBP and GVL showed to be the solvents with the poorest results for serine, arginine, and tyrosine. For glycine, butanol was determined as one of the worst solvents. The DMSO and NMP showed a strong predominance in the best solvents, namely for serine, arginine, tyrosine and glycine amino acid and dimer. An example of a peptide (protected arginine 5 mer) attached to rink amide nanostar hub is presented in Appendix II, Figure II.6.

Table 5.8 - Best and worst solvents for the peptides attached to rink amide nanostar by Hildebrand Model.

	Best Solvents		Worst Solvents	
Fmoc-Gly-OH	DMSO	NMP	NBP	PC
Fmoc-Gly-OH (dimer)	NMP	DMSO	NBP	Butanol
Fmoc-Gly-OH (5 mer)	NFM	NMP	NBP	Butanol
Fmoc-Gly-OH (10 mer)	PC	GVL	NBP	Butanol
Fmoc-Gly-OH (20 mer)	PC	GVL	NBP	Butanol
Fmoc-Ser(tBu)-OH	DMSO	THF	PC	GVL
Fmoc-Ser(tBu)-OH (dimer)	DMSO	NMP	PC	NBP
Fmoc-Ser(tBu)-OH (5 mer)	DMSO	NMP	PC	NBP
Fmoc-Ser(tBu)-OH (10 mer)	DMSO	THF	PC	GVL
Fmoc-Ser(tBu)-OH (20 mer)	NBP	Butanol	PC	GVL
Fmoc-Arg(Pbf)-OH	DMSO	NMP	PC	NBP
Fmoc-Arg(Pbf)-OH (dimer)	DMSO	NMP	PC	NBP
Fmoc-Arg(Pbf)-OH (5 mer)	DMSO	NMP	PC	NBP
Fmoc-Arg(Pbf)-OH (10 mer)	DMSO	NMP	PC	NBP
Fmoc-Tyr(tBu)-OH	DMSO	NMP	PC	GVL
Fmoc-Tyr(tBu)-OH (dimer)	DMSO	NMP	PC	NBP
Fmoc-Tyr(tBu)-OH (5 mer)	DMSO	NMP	PC	NBP
Fmoc-Tyr(tBu)-OH (10 mer)	DMSO	THF	PC	GVL

After all estimations were finished, a study was carried out to determine the effect of the hubs (wand and rink amide nanostars) in the solubility of the peptide studied.

5.3. COSMO-RS solubility estimations

Since that for some amino acids and peptides were not possible to predict the best and worst solvents, a COSMO-RS approach was then applied, using the software COSMO*therm,* to

overcome those estimations faults. In this way, this approach was also applied to all amino acids and respective peptides to compare the predications of both approaches – the solubility parameters and COSMO-RS approach. Therefore, the logarithmic molar solubility and solubility (mol/L) of all amino acids and their respective peptides was calculated for each solvent (DMSO, THF, MeTHF, NMP, NBP, NFM, GVL, PC, Butanol), being the results presented in Appendix III.

In Appendix III, when obtained a logarithmic solubility value of 0 or larger than 0, only indicates that the two compounds are miscible, therefore, the predictive solubility value isn't real (unless that experimental data supports it) whereby those solvents would not be chosen as best solvents. This happens due to inappropriate behaviour of the model that misestimates the free energy of fusion, ΔG_{fus}, that need to be considerate during the solubility estimation (explained in 3.4). However, to provide the maximum inside of the solvents set, positive values for the logarithmic solubility are allowed here as best solvents.

The results of the best and worst solvents for the amino acids are illustrated in the following Table 5.9.

Table 5.9 - Best and worst solvents for the (protected) free peptide by COSMO-RS.

	Best Solvents		Worst Solvents	
Fmoc-Gly-OH	NFM	MeTHF	PC	Butanol
Fmoc-Gly-OH (dimer)	GVL	MeTHF	PC	Butanol
Fmoc-Gly-OH (5 mer)	NMP	NFM	PC	Butanol
Fmoc-Gly-OH (10 mer)	NMP	NFM	PC	Butanol
Fmoc-Gly-OH (20 mer)	DMSO	-	PC	MeTHF
Fmoc-Ser(tBu)-OH	GVL	-	PC	Butanol
Fmoc-Ser(tBu)-OH (dimer)	GVL	NFM	PC	Butanol
Fmoc-Ser(tBu)-OH (5 mer)	NMP	THF	PC	Butanol
Fmoc-Ser(tBu)-OH (10 mer)	NMP	THF	PC	GVL
Fmoc-Ser(tBu)-OH (20 mer)	-	-	PC	GVL
Fmoc-Arg(Pbf)-OH	THF	NBP	PC	Butanol
Fmoc-Arg(Pbf)-OH (dimer)	THF	NMP	PC	Butanol
Fmoc-Arg(Pbf)-OH (5 mer)	THF	NMP	PC	Butanol
Fmoc-Arg(Pbf)-OH (10 mer)	-	-	PC	Butanol
Fmoc-Tyr(tBu)-OH	NFM	GVL	PC	Butanol
Fmoc-Tyr(tBu)-OH (dimer)	NFM	MeTHF	PC	Butanol
Fmoc-Tyr(tBu)-OH (5 mer)	THF	DMSO	PC	Butanol
Fmoc-Tyr(tBu)-OH (10 mer)	THF	NMP	PC	Butanol

Table 5.9 shows that the solvent PC was chosen as one of the worst solvents for all amino acids and respective peptides since PC presented the lower solubility values. The glycine 20 mer peptide presented MeTHF as the second worst solvents and serine 10 and 20 mer peptides presented GVL, unlike the rest that presented butanol as the second solvent with lower solubility.

Concerning the best solvents, NMP and NFM presented a strong predominance among all amino acids and respective peptides. It must be noted that, with the growth of peptide chain, the peptide best solvent choice changed. The biggest change is noted from peptide dimer to 5 mer at glycine, serine, and arginine. For tyrosine, it only changed from amino acid to peptide dimer. The cells that do not present any solvent means that the solubility presented by the remain (or all in serine 20 mer and arginine 10 mer) solvents is very low, whereby they cannot be chosen as best solvent. As referred, besides the logarithmic solubility meaning they are miscible and therefore they are good solvents, those solvents were not selected as the best ones since they cannot be compared with experimental data.

This approach was also applied to the peptides attached to wang and rink amide nanostar hub, however the results obtained showed low solubility for all solvents, being that none proved to be a suitable solvent. The low solubility was expected, however the values obtained were very low, and one of the possibilities of this happened, it could be the peptides were considered big structures that can corrupt the *cosmo* files.

5.4. Hub and Fmoc effect in solubility parameters estimations

As can be observed in the next sections, the suitable and not suitable solvents had changed among the analysis. Therefore, to understand the effect of adding a hub to amino acids or changing the hub will affect the solubility of the peptides, it was made a comparison between the solvents and parameters obtained in which analysis. For the free peptides, the best and worst solvents were also obtained without the protecting group, Fmoc. In this way, it will be verified if the Fmoc will make a significant change on the solubility of the peptides.

5.4.1. Fmoc effect on the solubility of the peptides

The study the effect of the Fmoc in the solubility started with an analysis between the predicted absolute solubility for the amino acids by COSMO-RS method with and without Fmoc, Figure 5.1.

As can be observed, the amino acids of glycine, serine and arginine presented a bigger predictive absolute solubility in the same solvents when they have the protecting group Fmoc, even without including the samples that obtained a predictive logarithmic solubility of zero (highlighted with red diamonds, ♦).
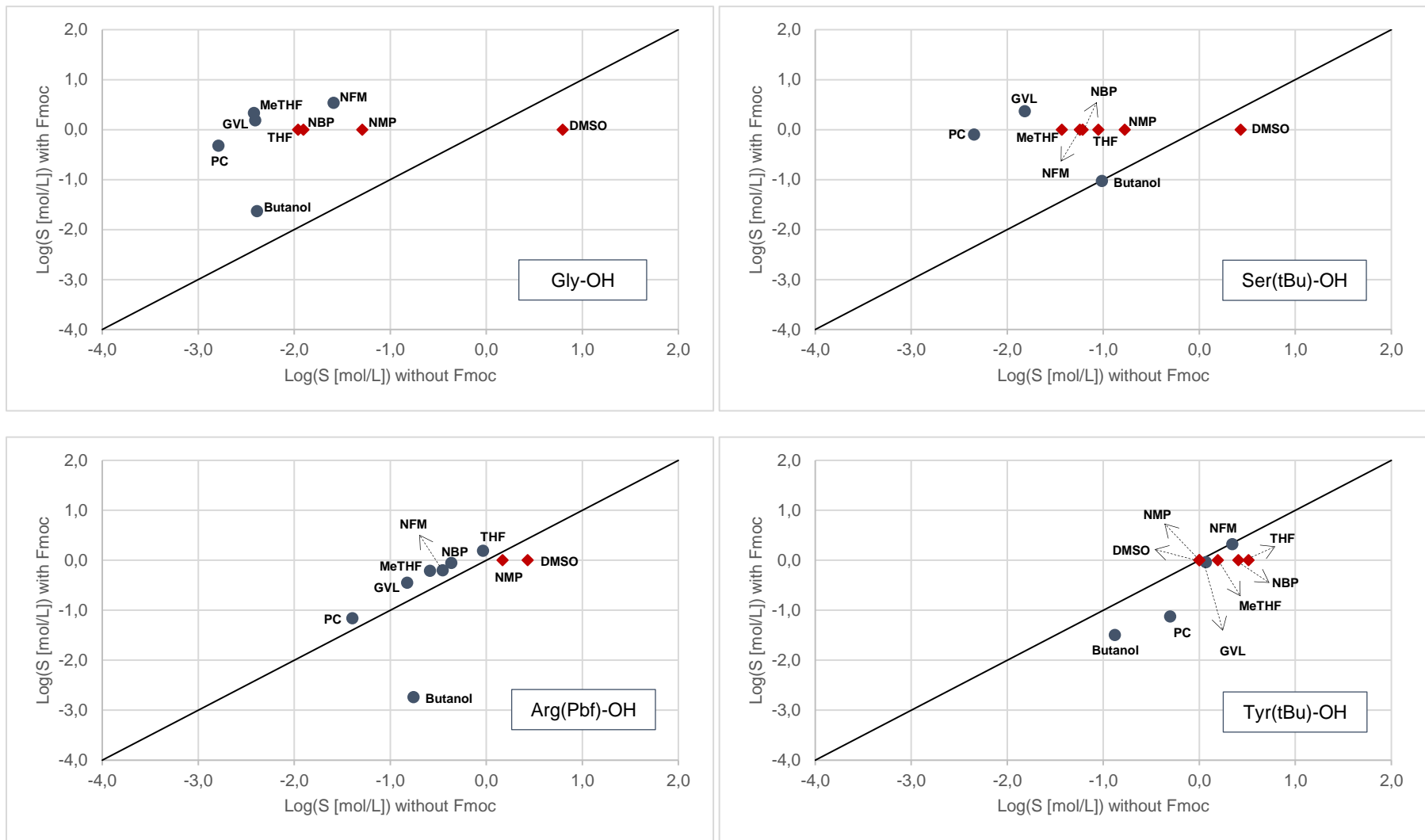
Figure 5.1 - Predicted absolute solubility without the Fmoc versus predicted absolute solubility with Fmoc of the amino acids.

The protecting group Fmoc is very soluble in organic solvents, so would be expecting that do not negatively influence the peptide solubility. A hydrogen bond on the free peptide has to be broken to link the peptide to Fmoc, contributing to an easier dissolution of the peptide and increasing the rate of solubility. The dipole moment is also a factor that can contribute to the solubility variance since it can be modified when the Fmoc is added. Glycine and serine exhibited the higher growths on solubility in the overall solvent set, whereby is expected that the best solvent choice will be affected. Contrary to the others, tyrosine presented better global solubility without the Fmoc.

To consolidate the results obtained by COSMO-RS, the Hildebrand Model was applied to the amino acids and respective peptides with and without the Fmoc for the same solvents set, Figure 5.2.

Figure 5.2 shows that for the majority of the peptides the Hildebrand Parameter did not had a big change, what not will influence in the final solvent choice. The amino acid and peptides that presented the most significant variation (blue triangles, ▲) in Hildebrand Parameter value were glycine amino acid and respective peptides (5, 10 and 20 mer) and tyrosine peptide 20 mer. Those amino acid and peptides presented the behaviour expect according the COSMO-RS method with the decrease of the Hildebrand Parameter, meaning a general increasing in rate of solubility.
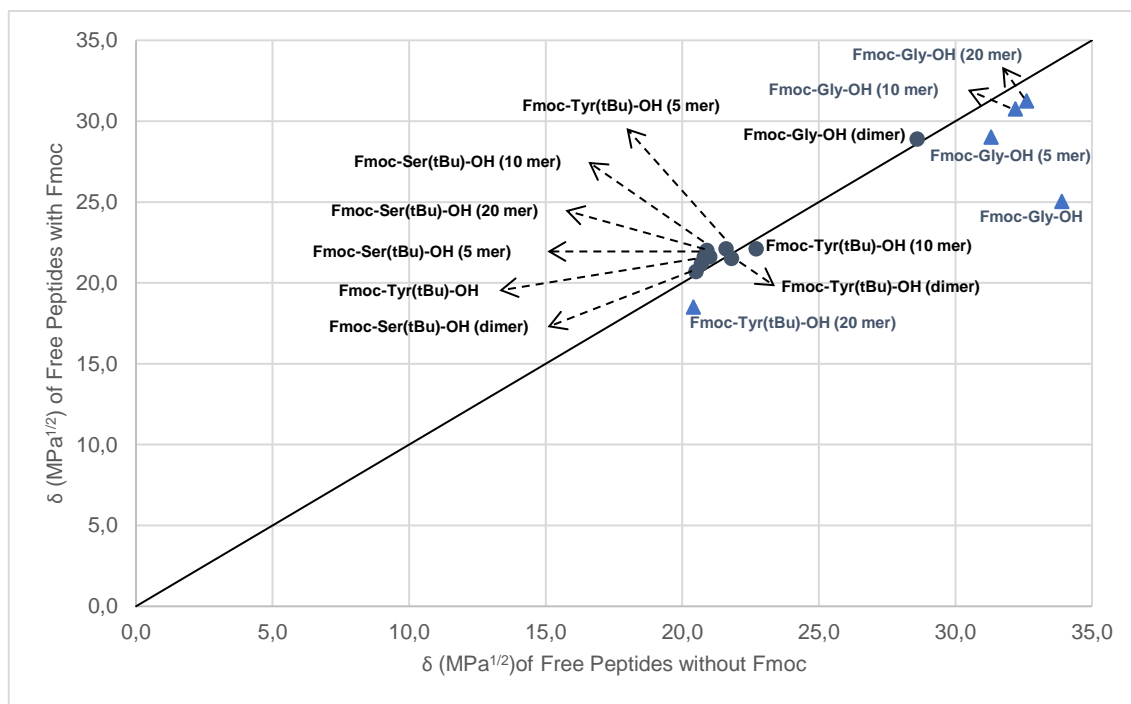
Figure 5.2 - Predicted Hildebrand Parameter without the Fmoc versus predicted Hildebrand Parameter with Fmoc of the amino acids and free peptides.

For a more detailed analysis in the behaviour of the solubility parameter with and without the Fmoc, the analysis was extended for the Hansen Model, with special attention for those that presented bigger variations on the Hildebrand Parameter, being those highlighted in the graphs with triangles symbols (▲), Figure 5.3 - Figure 5.5.

The majority of the peptides presented an increase of the Dispersion Parameter (δ_D), a decrease of the Hydrogen Bonding Parameter (δ_H) and maintain of the Polar Parameter (δ_P) when added the protector group Fmoc. The differences presented between the Dispersion and Hydrogen Bonding Parameter equilibrates the total Hildebrand Parameter, whereby the most of peptides do not showed big changes in this parameter, as observed in Figure 5.2.

What stands out about the peptides with the greatest differences in the Hildebrand Parameter from the rest of the peptides is the decrease in the Polar Parameter (δ_P) when added the Fmoc. The value of Polar Parameter derives from the dipole moment, whereby for those amino acid and peptides (glycine amino acid, 5, 10 and 20 mer), the dipole moment decreased when the Fmoc was incorporated, facilitating their dissolution.

Overall, both approaches, COSMO-RS and Solubility Parameters, presented a similar result, whereupon the Fmoc does not have a negative effect in the peptide solubility and even is beneficial in some cases (for example, glycine).
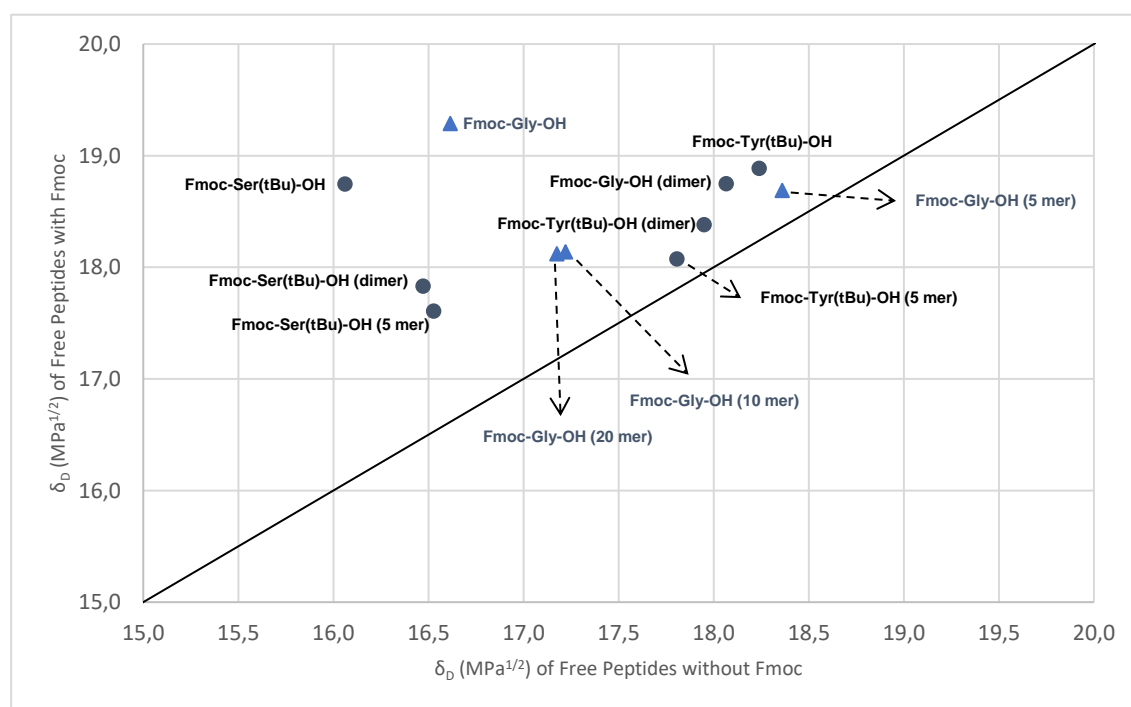
Figure 5.3 - Predicted Dispersion Parameter without the Fmoc versus predicted Dispersion Parameter with Fmoc of the amino and free peptides.
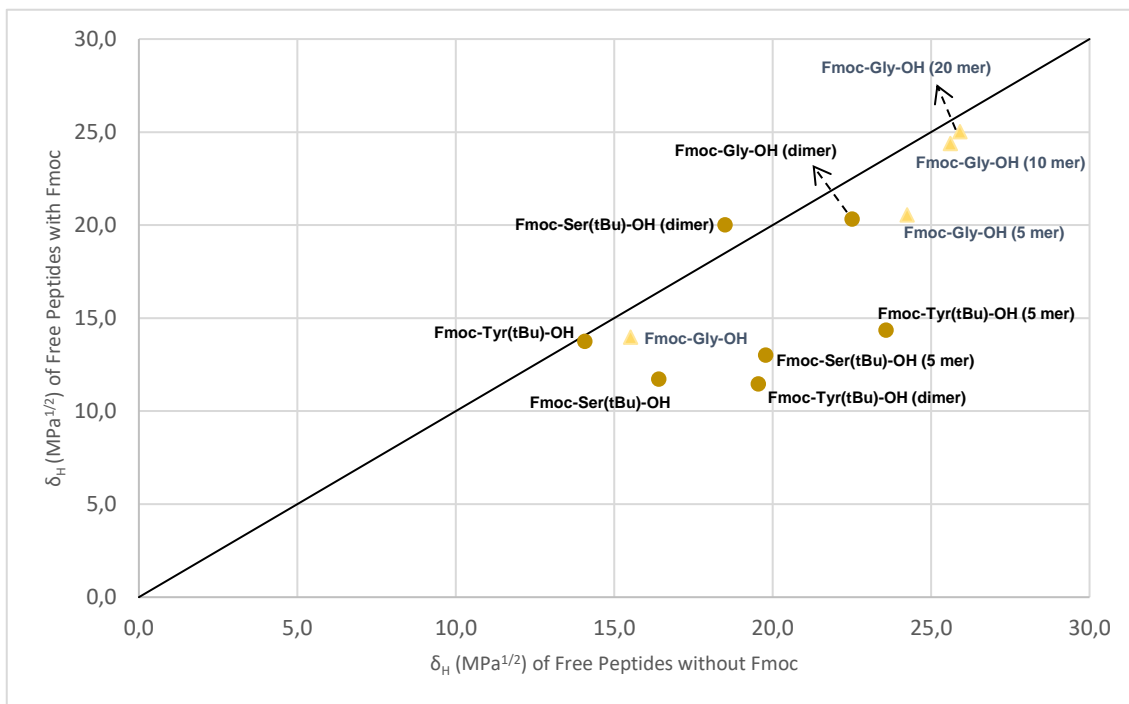
Figure 5.4 - Predicted Hydrogen Bonding Parameter without the Fmoc versus predicted Hydrogen Bonding Parameter with Fmoc of the amino and free peptides.
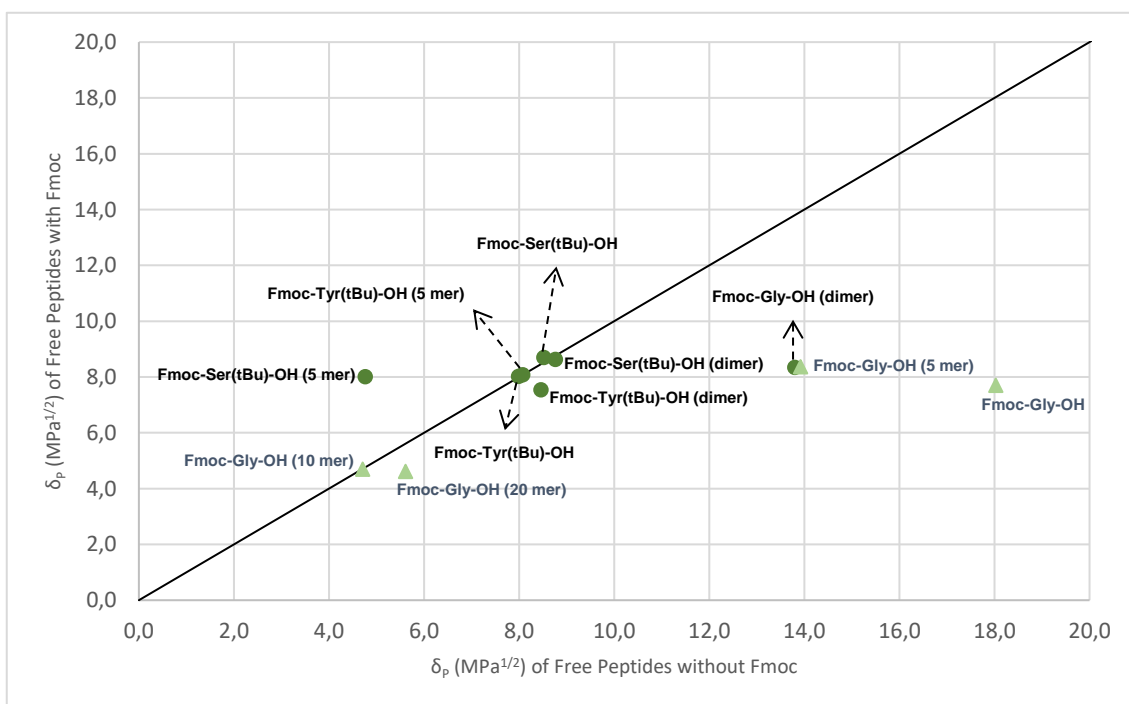
Figure 5.5 - Predicted Polar Parameter without the Fmoc versus predicted Polar Parameter with Fmoc of the amino and free peptides.

In relation to best solvent choice, concerning the solubility parameters, there were no extreme variations. Although glycine was the most affected by adding the Fmoc in their solubility parameters, none solvent remains within the criteria of the models (mainly for the peptides). Regarding the COSMO-RS results, in the protected amino acids, more solvents obtained logarithmic solubility of zero (even the solvents that presented to be the best for the unprotected amino acids), being the difference of solvents choice more flagrant, Table 5.10.

Table 5.10 - Best solvents choice for the protected and unprotected amino acids by COSMO-RS.

	Without Fmoc		With Fmoc	
Glycine	DMSO	NMP	NFM	MeTHF
Serine	DMSO	NMP	GVL	-
Arginine	DMSO	NMP	THF	NBP
Tyrosine	THF	NBP	NFM	GVL

5.4.2. Hub effect on the solubility of the peptides

The hub effect on solubility study was made with the Hildebrand Model. Therefore, the Hildebrand Parameter for the free amino acids and corresponding peptides was plotted versus the Hildebrand Parameter for the amino acids and corresponding peptides attached to wang nanostar and to rink amide nanostar, Figure 5.6 and Figure 5.7, respectively.

Overall, the Hildebrand Parameter of the peptides attached to both hubs decreased, being the most affected the serine 20 mer and tyrosine 10 mer in wang nanostar analyse and glycine amino acid, their respective peptides and serine peptide 20 mer in rink amide nanostar analyse. The decrease of the Hildebrand Parameter means a general increase in rate of solubility and applying the Hildebrand Model, more solvents were within the criteria. However, the peptides when attached to hub presented a predicted absolute solubility decrease, by COSMO-RS, compared to the free peptides, which causes discord between the two models.

Indeed, the peptides attached to the hub are bigger structures than the free peptides and with a larger volume, whereby it would be expected that their dissolution was more difficult. Therefore, it can be deduced that the Hildebrand Model cannot be applied to the peptides attached to hub and the Hildebrand Parameter failed to predict the solvent-peptide performance.
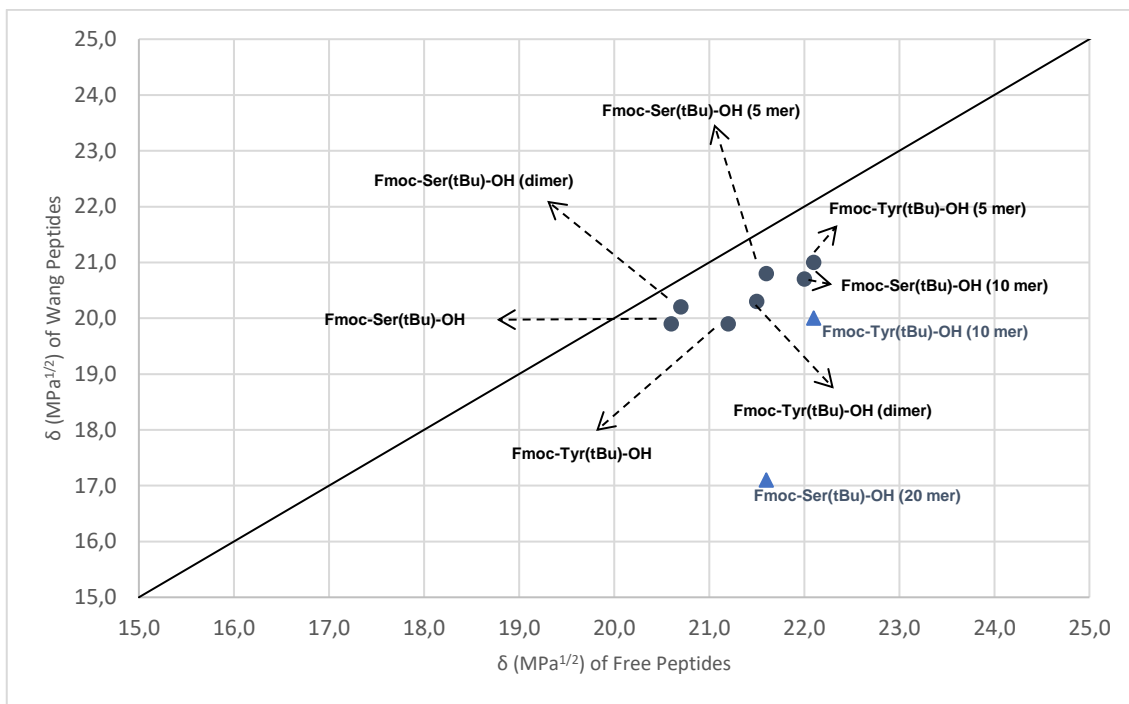
Figure 5.6 - Predicted Hildebrand Parameter of Free Peptides versus predicted Hildebrand Parameter of Peptides attached to Wang Hub.
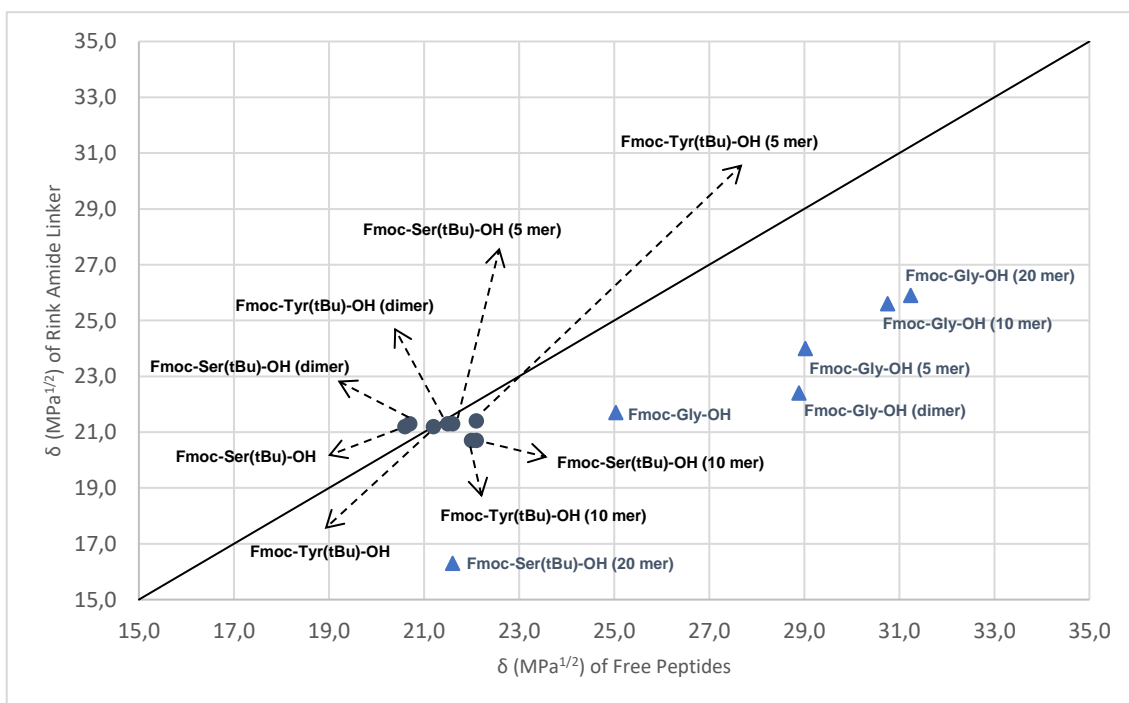
Figure 5.7 - Predicted Hildebrand Parameter of Free Peptides versus predicted Hildebrand Parameter of Peptides attached to Rink Amide Hub.

5.5. Solubility Parameters Methods *versus* COSMO-RS approach

Both approaches, the solubility parameters methods and COSMO-RS were used to predict the most and less suitable solvents for the peptides, without the need for experimental trials. The COSMO-RS is based on the combination of quantum chemical calculations with statistical thermodynamics and the Solubility Parameters methods is based on the principle '*like dissolves like*''. However, because the two approaches were applied to the same peptides with the same solvents, it would be expected similar results.

Therefore, it would be interesting to verify if the different approach presented similar or diverse results. For that, the results obtained were presented in the following Table 5.11, that included the best solvents obtained by COSMO-RS approach, Hildebrand Model and Hansen Model. It must be noted that the arginine is not present in this comparison, since it was not possible to predict the solubility parameters, whereby it was not possible to make conclusions.

Table 5.11 - Best solvents for the (protected) free peptides by different approaches.

	COSMO-RS		Hildebrand		Hansen	
Fmoc-Gly-OH	NFM	MeTHF	GVL	NFM	NFM	-
Fmoc-Gly-OH (dimer)	GVL	MeTHF	-	-	-	-
Fmoc-Gly-OH (5 mer)	NMP	NFM	-	-	-	-
Fmoc-Gly-OH (10 mer)	NMP	NFM	-	-	-	-
Fmoc-Gly-OH (20 mer)	DMSO	-	-	-	-	-
Fmoc-Ser(tBu)-OH	GVL	-	DMSO	THF	NFM	NMP
Fmoc-Ser(tBu)-OH (dimer)	GVL	NFM	DMSO	THF	Butanol	-
Fmoc-Ser(tBu)-OH (5 mer)	NMP	THF	DMSO	NMP	NFM	Butanol
Fmoc-Ser(tBu)-OH (10 mer)	NMP	THF	DMSO	NMP		
Fmoc-Ser(tBu)-OH (20 mer)	-	-	DMSO	NMP		
Fmoc-Tyr(tBu)-OH	NFM	GVL	DMSO	-	NMP	NFM
Fmoc-Tyr(tBu)-OH (dimer)	NFM	MeTHF	DMSO	-	THF	NFM
Fmoc-Tyr(tBu)-OH (5 mer)	THF	DMSO	DMSO	NMP	Butanol	NFM
Fmoc-Tyr(tBu)-OH (10 mer)	THF	NMP	DMSO	NMP		

Analysing Table 5.11, it can be verified that glycine showed similar results between the Hildebrand and Hansen models but, even though it share NFM as best solvent for amino acid with COSMO-RS, by this method the peptides predictions for the peptides are very distinct of the solubility models. Serine, besides some exceptions as NMP for 5 and 10 mer from COSMO-RS method and Hildebrand Model, presented different solvent choices for the respective amino acid or the same peptide among the three models. Tyrosine exhibited a similar behaviour to serine

and the best solvent chosen between the three models presents some variations, except the amino acid and dimer with NFM from COSMO-RS and the Hansen Model and to 5 mer with DMSO from COSMO-RS and the Hildebrand Model. Overall and considering the differences showed, it is expected that one of these approaches be more accurate than others.

To complete the analysis, the worst solvents obtained by the different approaches were compared, Table 5.12. The analysis of the worst solvents shows more similarities than the best solvents analysis.

Table 5.12 – Worst solvents for the (protected) free peptides by different approaches.

	COSMO-RS		Hildebrand		Hansen	
Fmoc-Gly-OH	PC	Butanol	Butanol	NBP	PC	NBP
Fmoc-Gly-OH (dimer)	PC	Butanol	Butanol	NBP	PC	NBP
Fmoc-Gly-OH (5 mer)	PC	Butanol	Butanol	NBP	PC	NBP
Fmoc-Gly-OH (10 mer)	PC	Butanol	Butanol	NBP	PC	NBP
Fmoc-Gly-OH (20 mer)	PC	MeTHF	Butanol	NBP	PC	NBP
Fmoc-Ser(tBu)-OH	PC	Butanol	GVL	PC	NBP	PC
Fmoc-Ser(tBu)-OH (dimer)	PC	Butanol	GVL	PC	NBP	PC
Fmoc-Ser(tBu)-OH (5 mer)	PC	Butanol	NBP	PC	NBP	PC
Fmoc-Ser(tBu)-OH (10 mer)	PC	GVL	NBP	PC		
Fmoc-Ser(tBu)-OH (20 mer)	PC	GVL	NBP	PC		
Fmoc-Tyr(tBu)-OH	PC	Butanol	NBP	PC	NBP	PC
Fmoc-Tyr(tBu)-OH (dimer)	PC	Butanol	NBP	PC	NBP	PC
Fmoc-Tyr(tBu)-OH (5 mer)	PC	Butanol	NBP	PC	NBP	PC
Fmoc-Tyr(tBu)-OH (10 mer)	PC	Butanol	NBP	PC		

For glycine, the results of COSMO-RS approach present similarities or with the Hildebrand Model (with Butanol) and with Hansen Model (with PC). Both Solubility Models also share NBP as the worst solvent. For serine and tyrosine, all models share PC as the worst solvent for all peptides.

Serine presented a different second worst solvent different among all methods, and tyrosine presented a second worst solvent (NBP) shared only between the Solubility Parameters Models. Due to the variations found in the determination of the less suitable solvents, it evidences once again that one approach should be more accurate than the others.

5.6. Predictions accuracy

To validate the methods applied in the solubility predictions, the experimental solubility of some amino acids in respective solvents was experimental determined, by the procedure detailed in Chapter 3, section 3.5. Table 5.13 presents the amino acids and solvents verified, the respective experimental solubility for a 100^2-fold diluted saturated solution and predicted absolute solubility.

Table 5.13 - Experimental and Predicted Absolute Solubilities of the amino acids in Butanol or THF, for a 100^2-fold saturated solution.

	Experimental		Predicted		
	S (mg/mL)	log S(mg/mL)	S (mg/mL)	log S(mg/mL)	Error
Fmoc-Gly-OH in Butanol	4.5	0.650	6.9	0.840	0.19
Fmoc-Gly-OH in THF	271.7	2.434	297.3	2.473	0.04
Fmoc-Ser(tBu)-OH in Butanol	191.7	2.283	36.1	1.557	-0.73
Fmoc-Arg(Pbf)-OH in Butanol	216.0	2.334	1.2	0.071	-2.26
Fmoc-Arg(Pbf)-OH in THF	231.7	2.365	1007.4	3.003	0.64

The COSMO-RS model overpredicted the solubility of glycine and arginine(in THF), contrary the solubility of the remain protected amino acids that were underestimated, Figure 5.8. The RMSE is 1.10 log units, however there is a significant outlier in the data set: Fmoc-Arg(Pbf)-OH in Butanol (highlighted with a red diamond, ♦).

The solubility of Fmoc-Arg(Pbf)-OH in butanol is severely underestimated, less 2.26 log units, by the model. None explanation could be found for this error, since the model obtained reasonable predictions for Fmoc-Arg(Pbf)-OH in THF and for other protected amino acids in butanol.

With the outlier removed, the overall RMSE reduces to 0.49 log units, which arises mainly from the estimated ΔG_{fus}. By using experimental data for T_{melt} and ΔH_{fus} to calculate the ΔG_{fus}, it would be expected a lower error.

Glycine, the apolar amino acid, presented the better accuracy between the model predictions and the real solubility, with a RMSE of 0.14 log units. Chapter 2, section 2.4, stated that also the Hildebrand Model has better accuracy for apolar amino acids. However, more experiments have to be done to determine a relation between the methods accuracy and the polarity of the amino acids.

Some experiments were repeated for a 500-fold diluted saturated solution, Table 5.14, to move away from the y-intercept. In the a 100^2-fold diluted saturated solution, the results presented a substantial variation when the intercept varied – defined based on calibration curve or zero.
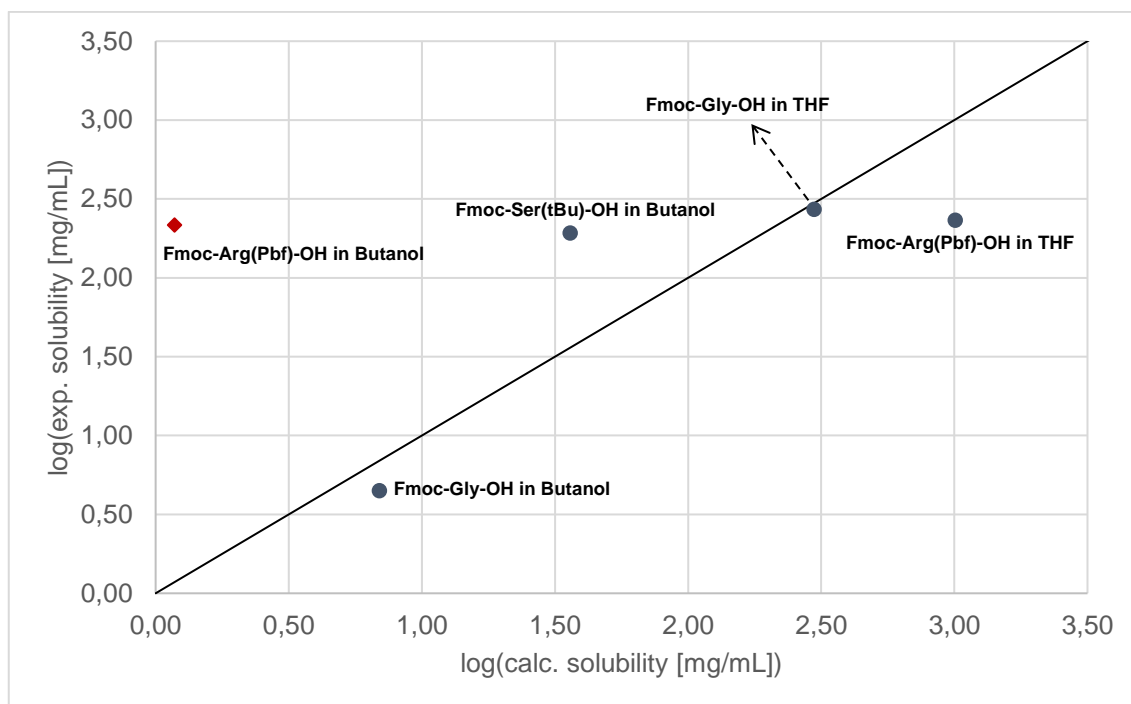
Figure 5.8 - Predicted absolute solubility versus experimental data for a 100^2-fold diluted saturated solution.

With a less diluted saturated solution, the interpolation will be less sensitive to change in intercept value. All the results relative to this section - the peak areas of calibration curve and saturated solution, calibration curve slope and intercept variations - are exposed in Appendix IV.

Analysing Table 5.14, the COSMO-RS model overpredicted the solubility of arginine in THF but underestimate the solubility of serine and glycine in butanol. The overall RMSE is 0.74 log units, Figure 5.9, that once again occurs principally from the estimated ΔG_{fus} without experimental data.

Table 5.14 - Experimental and Predicted Absolute Solubilities of the amino acids in Butanol or THF, for a 500-fold saturated solution..

	Experimental		Predicted		
	S (mg/mL)	log S(mg/mL)	S (mg/mL)	log S(mg/mL)	Error
Fmoc-Arg(Pbf)-OH in THF	181.108	2.26	1007.448	3.00	0.75
Fmoc-Ser(tBu)-OH in Butanol	102.094	2.01	36.051	1.56	-0.45
Fmoc-Gly-OH in Butanol	19.216	1.28	6.925	0.84	-0.44

Must be noted that two free peptides were tested – Ser(tBu) 5 mer and Gly 10 mer – but the solubility was very low to be possible construct a calibration curve. Overall, the COSMO-RS presents a helpful tool to predict the amino acids best solubility and determine which solvents must be experimental tested in the solvent choice process.
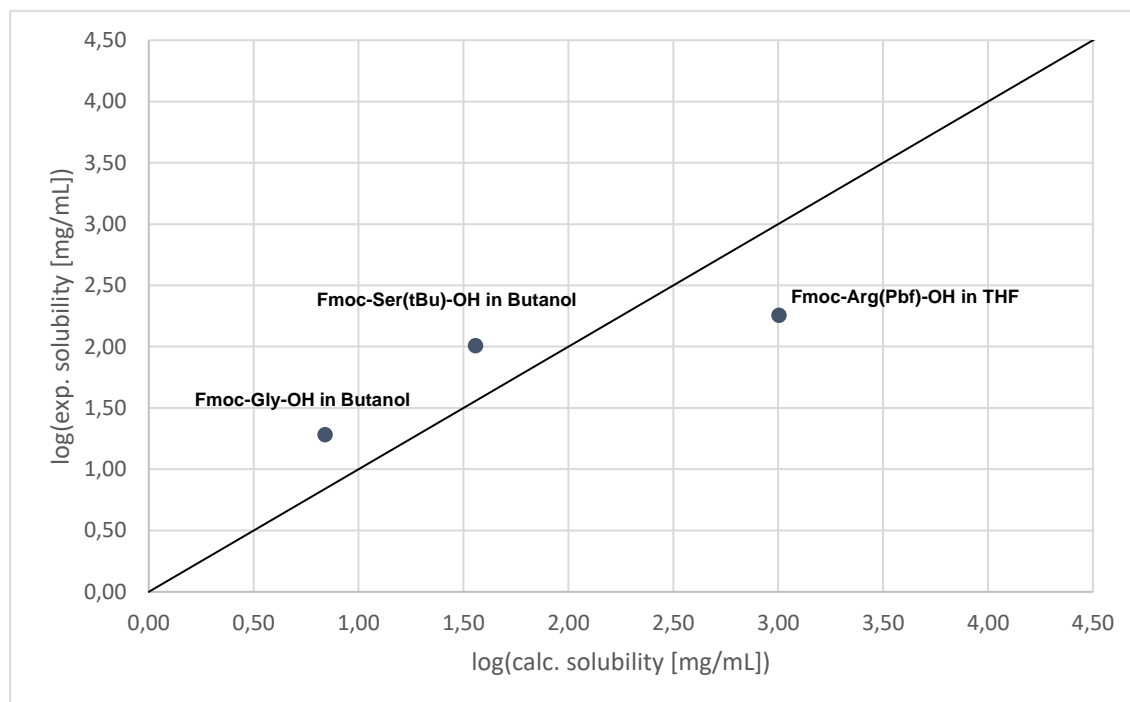
Figure 5.9 - Predicted absolute solubility versus experimental data for a 500-fold diluted saturated solution.

In relation to Hildebrand and Hansen Parameter Models, those models only predict which are suitable solvents and which are not, whereby it would be necessary testing all solvents used in the analysis to observe if the predictions corresponded with the real solubility, however, that was not possible.

5.7. Conclusion

To conclude this chapter, the estimations of the free peptides were carried out successfully by the three models proposed – Hildebrand Parameter Model, Hansen Parameter Model and COSMO-RS approach – and the best and worst solvents for the peptides by each model determined. The Hildebrand Model and COSMO-RS were also applied to the peptides attached to the wang nanostar hub and rink amide nanostar hub, being that the two models presented a significant results variation, studied posteriorly. The hub effect study, allowed to verify that the Hildebrand Model failed to predict the solubility of the peptides attached to hub, since it is known that

the solubility is very low and the COSMO-RS results are according, incompatible with the Hildebrand model results. Also, a study of the effect of protecting group, Fmoc, on the solubility of the amino acids and peptides was carried out, where was found that the Fmoc does not presented a negative effect in the solubility and a significant variation in the solvents choices are not verified. The solvents choices of the three methods for the free amino acids and peptides was compared, where was notorious differences and, therefore, it is expected that one method is more accurate than the others. Finally, the experimental solubility of some protected amino acids was discovered and compared to the predicted, presenting satisfactory results. Excepting the verification of the Hildebrand and Hansen Models results, all proposed objectives in section 5.1 were achieved.

6. Conclusion

The present work was focused on liquid-phase peptide synthesis optimisation by via quenching reagent and by peptides solubility estimation. Piperidine was found as best quenching reagent and COSMO-RS validated as support method to implement in peptide solubility determination, whereby the main goal of this project was accomplished.

For the process optimisation, a peptide synthesis review was performed. It was confirmed that the standard process to determine the peptides (products) solubility via experimental process, besides require experimental effort, is time-consuming and expensive. Also, the quenching reagent usually used, piperidine, is a hazard compound and propitious to side reactions. Past studies were revised where other promising quenching reagents were found – aniline and thiomalic acid – and also solubility estimation approaches were found – Solubility Parameters approach (includes Hildebrand Parameter Method and Hansen Parameter Method) and COSMO-RS approaches.

In the quenching optimisation, after a equivalents study starting with 5 amino acids equivalent (Fmoc-Phe-OH and Fmoc-Ser(tBu)-OH) and 2.4 piperidine equivalents, it was determined that the best conditions for a more controlled reaction without the formation of by-products were 2 or 3 amino acid equivalents and 2.1 quenching reagent equivalents. After the quenching performance with all reagents, it was verified that piperidine was the most effective, since it presented the best k_q, followed by thiomalic acid and, lastly, by aniline. It was also verified that lower concentrations generate best performances, with the k_q of 2.96 min^{-1} for piperidine, 1.85 min^{-1} for thiomalic acid and 0.02 min^{-1} for aniline. However, it was not possible to prove if k_q is influenced by factors as polarity, hydrophobicity, or molecular weight of the quenching reagent, due to the lack of data. Therefore, piperidine should be adopted in LPPS since provides good efficiency, essential to increase the yield and purity of the peptides synthesised.

Relating to peptide solubility estimation, the Solubility Parameters Models and the COSMO-RS approach were applied to free peptides (glycine, protected serine, protected tyrosine and protected arginine) and to the peptides attached to hub by wang linkers and by rink amide linkers. The best and worst solvents for each amino acid and peptide from each method were determined. DMSO was nominated as the overall best solvation solvent, followed by NFM, that was also determined as the best in the green solvent set. In relation to peptides attached to the hubs (wang and rink amide nanostars), the results showed disagreements between the Hildebrand and COSMO-RS models, that posteriorly was proved that Hildebrand failed in the solubility predictions of those peptides. A study of the effect of the Fmoc in the solubility of peptides demonstrated that the protecting group does not have a negative influence in the solubility, reflected in

minor changes in the choice of solvents. Additionally, a comparison between the best solvents selected by each approach was realized and, although some similarities were found, it was also founded differing results, whereby is expected that one of the approaches be more accurate than the others.

Lastly, a validation of the COSMO-RS was performed and the errors between the solubility predicted and real determined. For a 100^2-fold diluted saturated solution, a RMSE for the five experiments was of 1.10 log units, increased by the outlier of Fmoc-Arg(Pbf)-OH in butanol. By removing it, the RMSE decreased to 0.49 log units. The experiment was repeated for a less diluted (500-fold) saturated solution and a RMSE of 0.74 log units obtained. Considering all the experiments, COSMO-RS presented better accuracy for the apolar amino acid (glycine) than for the polar amino acids (serine and arginine). Unfortunately, the validation of the Hildebrand and Hansen Models was not possible. Since the overall RMSE do not exceed 1 log unit, the results were considered satisfactory for the amino acids, encouraging the implementation of COSMO-RS as a support method for solubility determination. However, to use the model for directly choose the best solvent of the amino acids or to the peptides, a better validation needs to be performed.

6.1 Future directions

The aim of the projected was achieved, however, for a more complete study and improved results, still exists some points that can be amended. Therefore, the suggestion of the following steps on the quenching reagent optimisation are:
- Expand the quenching performance to more different reagents. It will allow understanding how the reagents properties as pH, molecular weight and hydrophobicity can affect the quenching reaction. Also, it suggested performing in five different amino acids to be possible to establish a correlation between the quenching rate constant and the amino acids properties.

- Enlarge the set of quenching reagent candidates for more green reagents. Since the excess reagent will be removed and disposed, it would be favourable to evaluate the reagent toxicity. Also, the reagent cost could be included in the ponderation of the final choice.

To conclude, the last recommendations of the next steps are regardless the solubility of the peptide estimations:
- Increase the number of experiments to determine each method accuracy according to the polarity of the amino acids. In the present work, apolar amino acids presented a better accuracy than polar amino acids. However, only a few experiments were realized.

- Included solvent mixtures in the set of candidates. The COSMO-RS approach and the Solubility Parameters approach allow to estimate the solubility and solubility parameters, respectively, for solvents mixtures, whereby the set of solvents candidates can be improved.

- Both approaches, COSMO-RS and Solubility Parameters, should be validated separately with experimental data for more peptides to determine the correct accuracy of each method. Therefore, the Hildebrand and Hansen Solubility Parameters should be determined quantitatively by using an inverse gas chromatography or a UV spectroscopy and the COSMO-RS validation extended for more amino acids and peptides.

References

[1]   K. Fosgerau and T. Hoffmann, "Peptide therapeutics: Current status and future directions," *Drug Discov. Today*, vol. 20, pp. 122–128, 2015, doi: 10.1016/j.drudis.2014.10.003.

[2]   S. J. Tzotzos, "Peptide Drugs of the Decade," 2020. doi: https://www.eurpepsoc.com/peptide-drugs-decade/.

[3]   "Peptide Therapeutics Market: Key Insights," *Future Market Insights*, 2020. https://www.futuremarketinsights.com/reports/peptide-therapeutics-market.

[4]   "Peptide Therapeutics Market - Growth, Trends, and Forecasts (2020 - 2025)," *Research and Markets*, 2020. https://www.researchandmarkets.com/reports/4896193/peptide-therapeutics-market-growth-trends-and?utm_source=dynamic&utm_medium=GNOM&utm_code=ddshw2&utm_campaign=1403928+-+Global+Peptide+Therapeutics+Market+(2020+to+2025)+-+Growth%2C+Trends%2C+and+Forecast.

[5]   F. Albericio and A. El-Faham, "Choosing the Right Coupling Reagent for Peptides: A Twenty-Five-Year Journey," *Org. Process Res. Dev.*, vol. 22, pp. 760–772, 2018, doi: 10.1021/acs.oprd.8b00159.

[6]   R. K. Henderson *et al.*, "Expanding GSK's solvent selection guide – embedding sustainability into solvent selection starting at medicinal chemistry," *Green Chem.*, vol. 13, pp. 854–862, 2011, doi: 10.1039/c0gc00918k.

[7]   F. Paquin, J. Rivnay, A. Salleo, N. Stingelin, and C. Silva, "Multi-phase semicrystalline microstructures drive exciton dissociation in neat plastic semiconductors," *J. Mater. Chem. C*, vol. 3, pp. 10715–10722, 2015, doi: 10.1039/b000000x.

[8]   S. B. Lawrenson, R. Arav, and M. North, "The greening of peptide synthesis," *Green Chem.*, vol. 19, pp. 1685–1691, 2017, doi: 10.1039/c7gc00247e.

[9]   O. Al Musaimi, B. G. De La Torre, and F. Albericio, "Greening Fmoc/: T Bu solid-phase peptide synthesis," *Green Chem.*, vol. 22, pp. 996–1018, 2020, doi: 10.1039/c9gc03982a.

[10]   C. M. Alder *et al.*, "Updating and further expanding GSK's solvent sustainability guide," *Green Chem.*, vol. 18, pp. 3879–3890, 2016, doi: 10.1039/c6gc00611f.

[11]   A. Isidro-Llobet *et al.*, "Sustainability Challenges in Peptide Synthesis and Purification: From R&D to Production," *J. Org. Chem.*, vol. 84, pp. 4615–4628, 2019, doi: 10.1021/acs.joc.8b03001.

[12]   N. L. Benoiton, *Chemistry of Peptide Synthesis,* pp. 1-21, 25-29, 1st ed. Taylor & Francis Group, LLC, 2006. ISBN: 9781574444544.

[13]   J. Klose, A. El-Faham, P. Henklein, L. A. Carpino, and M. Bienert, "Addition of HOAt dramatically improves the effectiveness of pentafluorophenyl-based coupling reagents," *Tetrahedron Lett.*, vol. 40, pp. 2045–2048, 1999, doi: 10.1016/S0040-4039(99)00089-1.

[14]   B. Ucar, T. Acar, P. Pelit Arayici, M. Sen, S. Derman, and Z. Mustafaeva, "Synthesis and Applications of Synthetic Peptides," *IntechOpen*, vol. 1, pp. 1–17, 2019, doi: 10.5772/intechopen.85486.

[15]   O. F. Luna, J. Gomez, C. Cárdenas, F. Albericio, S. H. Marshall, and F. Guzmán, "Deprotection reagents in Fmoc solid phase peptide synthesis: Moving away from piperidine?," *Molecules*, vol. 21, pp. 1–12, 2016, doi: 10.3390/molecules21111542.

[16]   W. Chen, "Membrane Enhanced Peptide Synthesis (MEPS) – Process Development and Application," Imperial College London, 2015.

[17] J. Howl, *Peptide Synthesis and Applications*, pp. 1-25, 1st ed. Humana Press, 2005. ISBN: 1588293173.

[18] B. R. B. Merrifield, "Solid-Phase Peptide Synthesis," in *Advances in Enzymology and Related Areas of Molecular Biology*, vol. 32, 1969, pp. 221–289.

[19] S. W. J. So, "The Novel Membrane Enhanced Peptide Synthesis Process," Imperial College London, 2009.

[20] S. So, L. G. Peeva, E. W. Tate, R. J. Leatherbarrow, and A. G. Livingston, "Organic solvent nanofiltration: A new paradigm in peptide synthesis," *Org. Process Res. Dev.*, vol. 14, pp. 1313–1325, 2010, doi: 10.1021/op1001403.

[21] E. Bayer, M. Mutter, R. Uhmann, J. Polster, and H. Mauser, "Kinetic Studies of the Liquid Phase Peptide Synthesis," *J. Am. Chem. Soc.*, vol. 96, pp. 7333–7336, 1974, doi: 10.1021/ja00830a027.

[22] J. Yeo, L. Peeva, P. Gaffney, F. Albericio, and A. G. Livingston, "TIDES Europe: Oligonucleotides and Peptide Therapeutics," in *Liquid Phase Peptide Synthesis via Nanostar-Sieving*, 2019.

[23] M. Belmares *et al.*, "Hildebrand and hansen solubility parameters from molecular dynamics with applications to electronic nose polymer sensors," *J. Comput. Chem.*, vol. 25, pp. 1814–1826, 2004, doi: 10.1002/jcc.20098.

[24] C. M. Hansen, "50 Years with solubility parameters - Past and future," *Prog. Org. Coatings*, vol. 51, pp. 77–84, 2004, doi: 10.1016/j.porgcoat.2004.05.004.

[25] John Burke, "Solubility Parameters: Theory and Application," *B. Pap. Gr.*, vol. 3, 1984, doi: https://cool.culturalheritage.org/coolaic/sg/bpg/annual/v03/bp03-04.html.

[26] S. Venkatram, C. Kim, A. Chandrasekaran, and R. Ramprasad, "Critical Assessment of the Hildebrand and Hansen Solubility Parameters for Polymers," *J. Chem. Inf. Model.*, 2019, doi: 10.1021/acs.jcim.9b00656.

[27] C. M. Hansen, *Hansen Solubility Parameters: A User's Handbook*, pp. 1-24, 95-110, 203-227. 2nd ed. CRC Press, 2007. ISBN: 9788578110796.

[28] B. Hossin, K. Rizi, and S. Murdan, "Application of Hansen Solubility Parameters to predict drug-nail interactions, which can assist the design of nail medicines," *Eur. J. Pharm. Biopharm.*, vol. 102, pp. 32–40, 2016, doi: 10.1016/j.ejpb.2016.02.009.

[29] B. Sanchez-Lengeling, L. M. Roch, J. D. Perea, S. Langner, C. J. Brabec, and A. Aspuru-Guzik, "A Bayesian Approach to Predict Solubility Parameters," *Adv. Theory Simulations*, vol. 2, pp. 1–10, 2018, doi: 10.1002/adts.201800069.

[30] A. Benazzouz, L. Moity, C. Pierlot, V. Molinier, and J. M. Aubry, "Hansen approach versus COSMO-RS for predicting the solubility of an organic UV filter in cosmetic solvents," *Colloids Surfaces A Physicochem. Eng. Asp.*, vol. 458, pp. 101–109, 2014, doi: 10.1016/j.colsurfa.2014.03.065.

[31] B. Delley, "The conductor-like screening model for polymers and surfaces," *Mol. Simul.*, vol. 32, pp. 117–123, 2006, doi: 10.1080/08927020600589684.

[32] A. Klamt, "The COSMO and COSMO-RS solvation models," *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, vol. 8, pp. 1–11, 2017, doi: 10.1002/wcms.1338.

[33] S. Abbott, *Solubility Science : Principles and Practice*. pp. 1-15, 60-90, 92-101. 3rd ed. Steven Abbott TCNF Ltd, 2017. ISBN: 9781605954844.

[34] A. Klamt, "Conductor-like screening model for real solvents: A new approach to the quantitative calculation of solvation phenomena," *J. Phys. Chem.*, vol. 99, pp. 2224–2235, 1995, doi: 10.1021/j100007a062.

[35] COSMOlogic, "COSMOthermX: A Graphical User Interface to the COSMOtherm Program." pp. 37–42, 2011.

[36]   A. del P. Sánchez-Camargo, M. Bueno, F. Parada-Alfonso, A. Cifuentes, and E. Ibáñez, "Hansen solubility parameters for selection of green extraction solvents," *TrAC - Trends Anal. Chem.*, vol. 118, pp. 227–237, 2019, doi: 10.1016/j.trac.2019.05.046.

[37]   S. Yaqub, B. Lal, B. Partoon, and N. B. Mellon, "Investigation of the task oriented dual function inhibitors in gas hydrate inhibition: A review," *Fluid Phase Equilib.*, vol. 477, pp. 40–57, 2018, doi: 10.1016/j.fluid.2018.08.015.

[38]   B. Schröder, L. M. N. B. F. Santos, I. M. Marrucho, and J. A. P. Coutinho, "Prediction of aqueous solubilities of solid carboxylic acids with COSMO-RS," *Fluid Phase Equilib.*, vol. 289, pp. 140–147, 2010, doi: 10.1016/j.fluid.2009.11.018.

[39]   A. Klamt, F. Eckert, and W. Arlt, "COSMO-RS: An alternative to simulation for calculating thermodynamic properties of liquid mixtures," *Annu. Rev. Chem. Biomol. Eng.*, vol. 1, pp. 101–122, 2010, doi: 10.1146/annurev-chembioeng-073009-100903.

[40]   D. Takahashi, T. Inomata, and T. Fukui, "AJIPHASE®: A Highly Efficient Synthetic Method for One-Pot Peptide Elongation in the Solution Phase by an Fmoc Strategy," *Angew. Chemie - Int. Ed.*, vol. 56, pp. 7803–7807, 2017, doi: 10.1002/anie.201702931.

[41]   J. Lopez, S. Pletscher, A. Aemissegger, C. Bucher, and F. Gallou, "N -Butylpyrrolidinone as Alternative Solvent for Solid-Phase Peptide Synthesis," *Org. Process Res. Dev.*, vol. 22, pp. 494–503, 2018, doi: 10.1021/acs.oprd.7b00389.

[42]   S. Chung, "Further Developments and Optimisation in Nanostar Sieving Technology for Peptide Synthesis– In Search for Optimum Solvent Composition," Imperial College London, 2019.

[43]   A. Kumar, Y. E. Jad, A. El-Faham, B. G. de la Torre, and F. Albericio, "Green solid-phase peptide synthesis 4. γ-Valerolactone and N-formylmorpholine as green solvents for solid phase peptide synthesis," *Tetrahedron Lett.*, vol. 58, no. 30, pp. 2986–2988, 2017, doi: 10.1016/j.tetlet.2017.06.058.

[44]   H. Moussout, H. Ahlafi, M. Aazza, and H. Maghat, "Critical of linear and nonlinear equations of pseudo-first order and pseudo-second order kinetic models," *Karbala Int. J. Mod. Sci.*, vol. 4, pp. 1–11, 2018, doi: 10.1016/j.kijoms.2018.04.001.

[45]   C. Kim, A. Chandrasekaran, T. D. Huan, D. Das, and R. Ramprasad, "Polymer Genome: A Data-Powered Polymer Informatics Platform for Property Predictions," *J. Phys. Chem. C*, vol. 122, pp. 17575–17585, 2018, doi: 10.1021/acs.jpcc.8b02913.

[46]   A. Niederquell, N. Wyttenbach, and M. Kuentz, "New prediction methods for solubility parameters based on molecular sigma profiles using pharmaceutical materials," *Int. J. Pharm.*, vol. 546, pp. 137–144, 2018, doi: 10.1016/j.ijpharm.2018.05.033.

[47]   G. Y. Han and L. A. Carpino, "The 9- Fluorenylmethoxycarbonyl Amino-Protecting Group," *J. Org. Chem.*, vol. 37, pp. 3404–3409, 1972.

Appendix I

The areas under the peak, Table I.1 - Table I.5, used for Figure 4.1 - Figure 4.10, in Chapter 4 - Equivalents study (page 39).

Table I.1 - Area under the peaks for 5 equiv. of amino acid and 2.4 equiv. of piperidine.

Time(min)	Fmoc-Phe-OH				Fmoc-Ser(tBu)-OH			
	Amino Acid	Active Ester	Quenched Specie	*By-product*	Amino Acid	Active Ester	Quenched Specie	*By-Product*
0	2966.97	803.2	413.3	0	1913.69	2699.4	617.7	0
1	8057	136.4	3523.3	29.8	4865.41	348.7	6484.5	74.1
3	8288.85	175.4	4887	43.9	4618.97	416.3	7063.2	96.7
5	7721.48	228.4	5862.1	57.2	3570.6	513.2	7279.3	127.2
11	5611.94	294.8	7459.8	104.2	3505.3	831.9	7771.5	210.1
30	4989.43	546.2	8323.6	203.3	2956.07	1428.6	7462.2	383.7
60	4706.85	1035.8	9856.3	413.6	2516.9	2358.1	7374.5	667.2
120	3834.02	1391.4	8555.4	834	1809.87	3979.1	6167.4	834

Table I.2 - Area under the peaks for 5 equiv. of amino acid and 6 equiv. of piperidine.

	Fmoc-Phe-OH			Fmoc-Ser(tBu)-OH		
Time(min)	Amino Acid	Active Ester	Quenched Specie	Amino Acid	Active Ester	Quenched Specie
0	6571.6	2114.4	210.3	2318.2	5.5	0.0
1	7059.5	1069.9	4239.3	2521.5	352.1	5117.8
2	6814.4	1308.5	4460.7	1978.6	453.6	4813.7
5	6763.7	1817.1	4765.8	2076.4	570.8	5403.2
10	6987.8	2998.4	5137.2	1907.0	1086.9	5441.4
30	5038.9	5776.3	3963.6	1764.5	2443.4	5311.9
60	4104.4	9681.0	3077.3	1251.0	3989.4	4143.6
120	3125.7	10990.6	1297.4	691.1	5436.4	2297.7

Table I.3 - Area under the peaks for 3 and 2 equiv. of Fmoc-Phe-OH and 2.1 equiv. of piperidine.

	3 Equiv. Fmoc-Phe-OH			2 Equiv. Fmoc-Phe-OH		
Time (min)	Amino Acid	Active Ester	Quenched Specie	Amino Acid	Active Ester	Quenched Specie
0	765.4	733.7	0.0	576.3	555.3	0.0
0.5	814.8	103.0	2332.1	637.1	93.9	1547.9
1	805.2	103.5	2934.9	713.5	86.8	1820.4
5	883.9	103.9	3458.5	684.3	76.5	1973.5
10	772.9	101.6	2847.0	579.4	78.8	1657.1
30	740.7	99.5	3099.8	590.1	48.8	1921.0
60	706.8	93.9	3516.0	559.3	23.9	1951.0
120	552.3	31.6	3186.4	400.6	36.8	1719.3

Table I.4 - Area under the peaks for 3 and 2 equiv. of Fmoc-Phe-OH and 2.1 equiv. of thiomalic acid.

Time (min)	3 Equiv. Fmoc-Phe-OH			2 Equiv. Fmoc-Phe-OH		
	Amino Acid	Active Ester	Quenched Specie	Amino Acid	Active Ester	Quenched Specie
0	3323.2	1286.5	0.0	1284.8	375.8	0.0
0.5	2125.3	359.8	568.2	1862.9	245.0	428.7
2	2465.1	429.0	790.0	1875.4	172.3	520.0
5	2458.7	413.7	931.1	1846.1	149.2	553.0
10	2307.9	366.5	981.6	1725.5	96.1	523.2
30	1991.7	136.4	1056.5	1521.3	41.3	659.4
60	1764.4	78.3	927.0	1287.3	38.0	511.1
120	1754.8	104.4	1087.1	1542.3	37.8	647.8

Table I.5 - Area under the peaks for 3 and 2 equiv. of Fmoc-Phe-OH and 2.1 equiv. of aniline.

Time (min)	3 Equiv. Fmoc-Phe-OH			2 Equiv. Fmoc-Phe-OH		
	Amino Acid	Active Ester	Quenched Specie	Amino Acid	Active Ester	Quenched Specie
0	2057.5	759.0	0.0	1862.1	615.3	0.0
0.5	2322.7	1092.3	0.0	2136.5	751.9	0.0
2	2624.0	1299.9	0.0	2219.0	951.1	0.0
5	2914.5	1658.7	130.7	1934.5	1121.1	108.7
11	2745.9	1585.8	142.7	1304.6	781.5	111.9
30	2509.9	1736.5	233.1	1496.2	911.0	241.0
60	1984.2	1635.6	505.1	1439.0	1272.8	433.3
120	1784.3	1645.6	909.9	1128.7	809.7	623.0

The amino acids and peptides structures, Figure II.1 - Figure II.4, used for in Chapter 5 (page 55). An example of a peptide attached to wang nanostar hub is presented in Figure II.5 and an example of a peptide attached to rink amide nanostar hub is presented in Figure II.6.
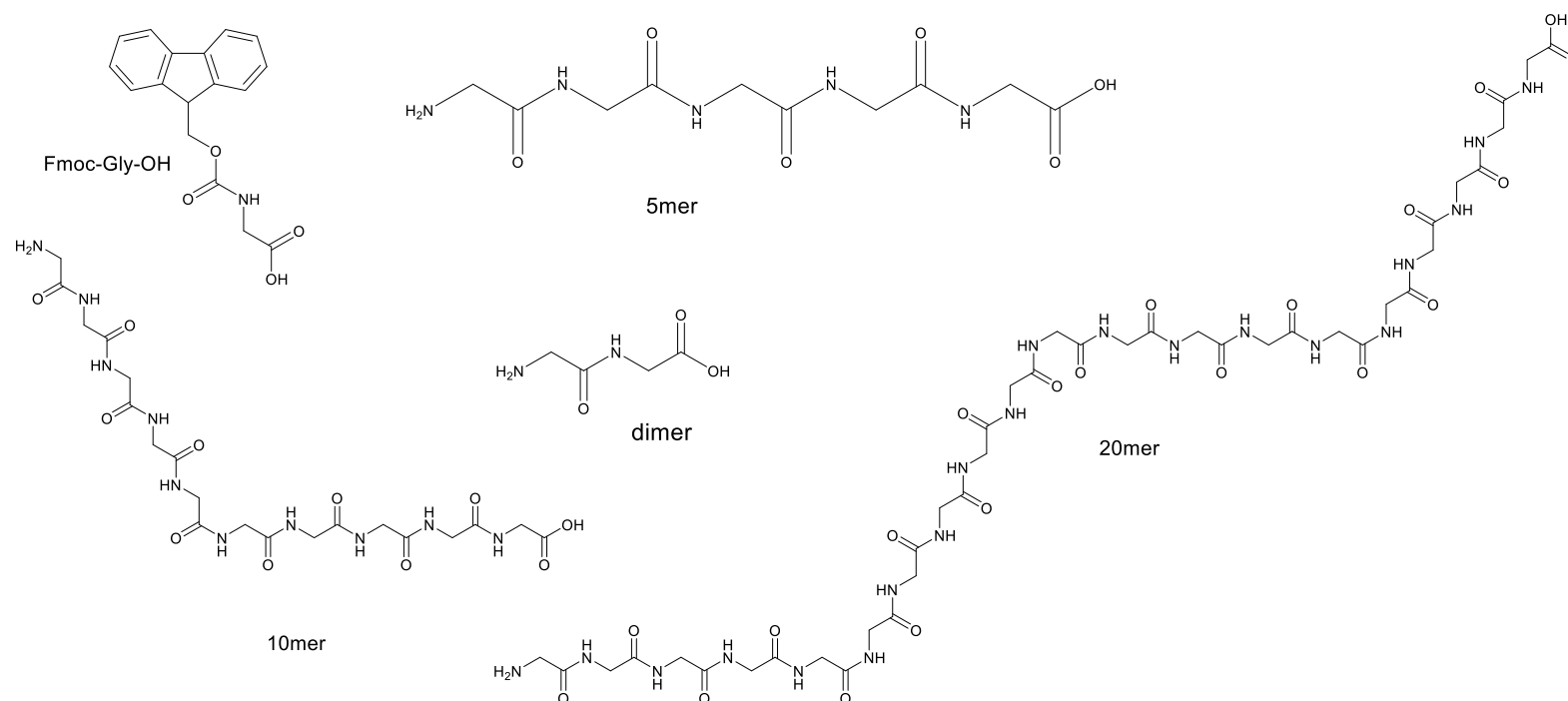
Figure II.1 - Glycine (protected - Fmoc) amino acid and respective peptides (dimer, 5, 10 and 20 mer) structure.
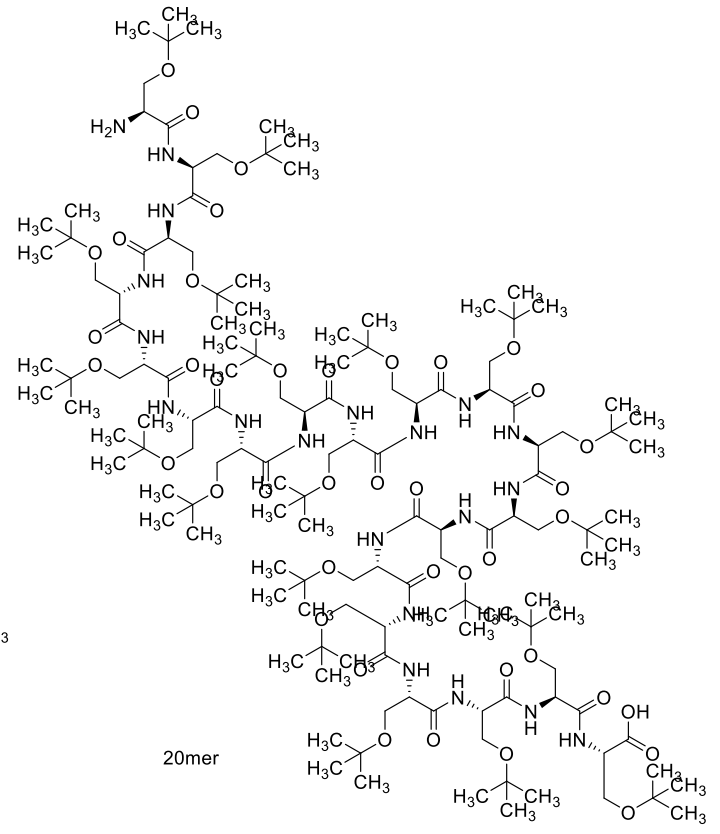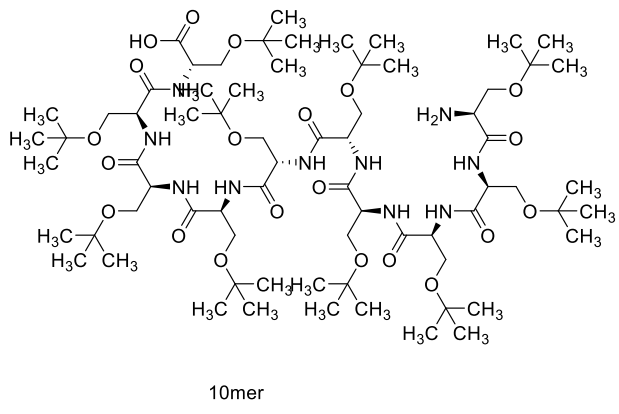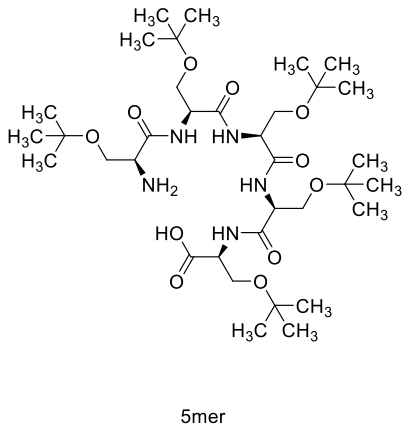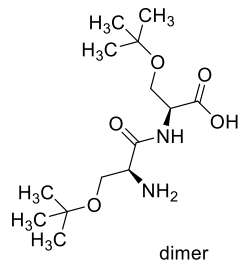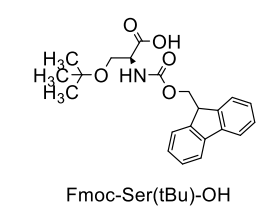
Figure II.2 - Protected (tBu) serine (protected – Fmoc) amino acid and respective peptides (dime, 5, 10 and 20 mer) structure.

Fmoc-Tyr(tBu)-OH

dimer

5mer

10 mer
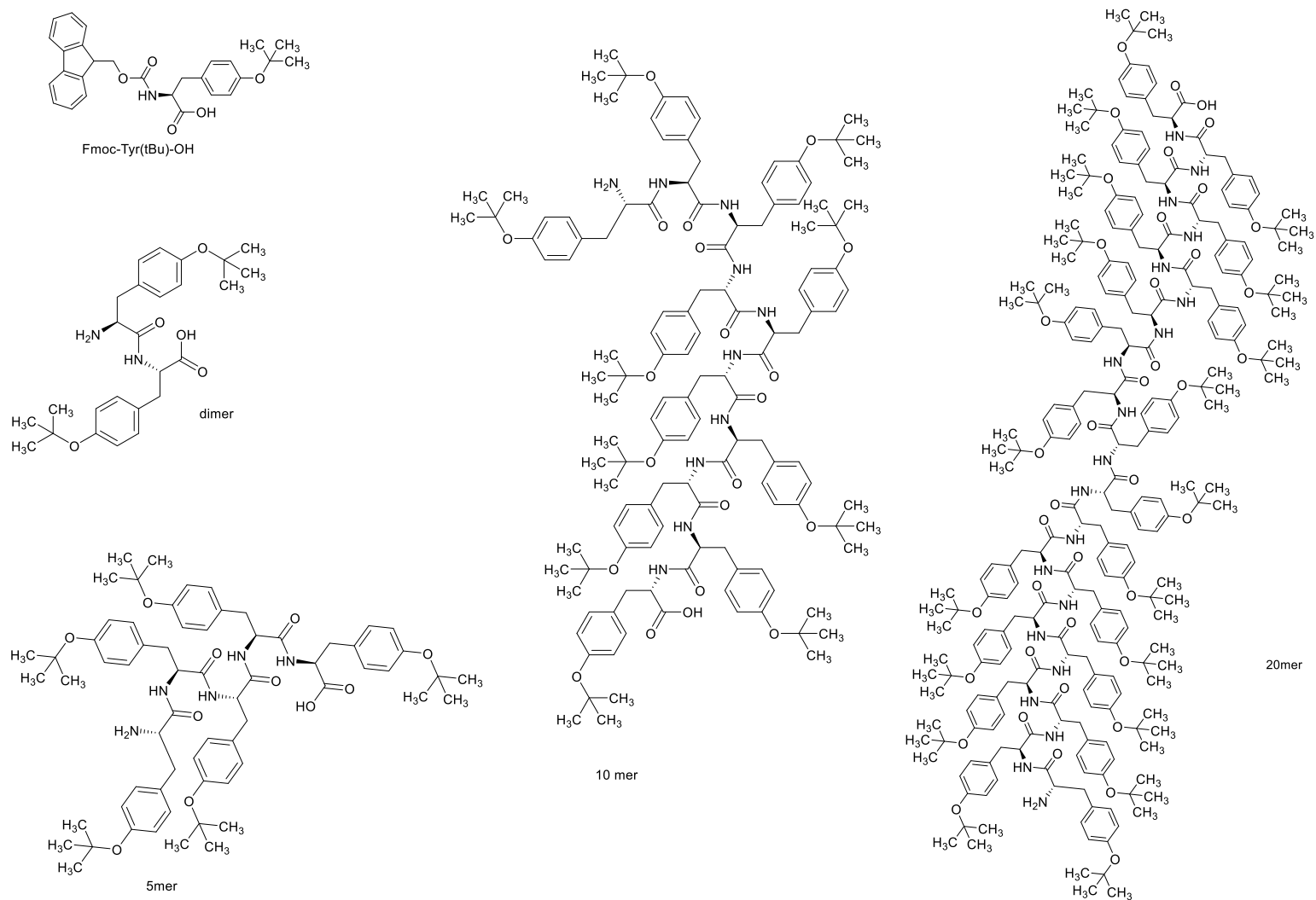
20mer

Figure II.3 - Protected (tBu) tyrosine (protected – Fmoc) amino acid and respective peptides (dimer, 5, 10 and 20 mer) structure.

90

Fmoc-Arg(Pbf)-OH

dimer

5mer

10mer
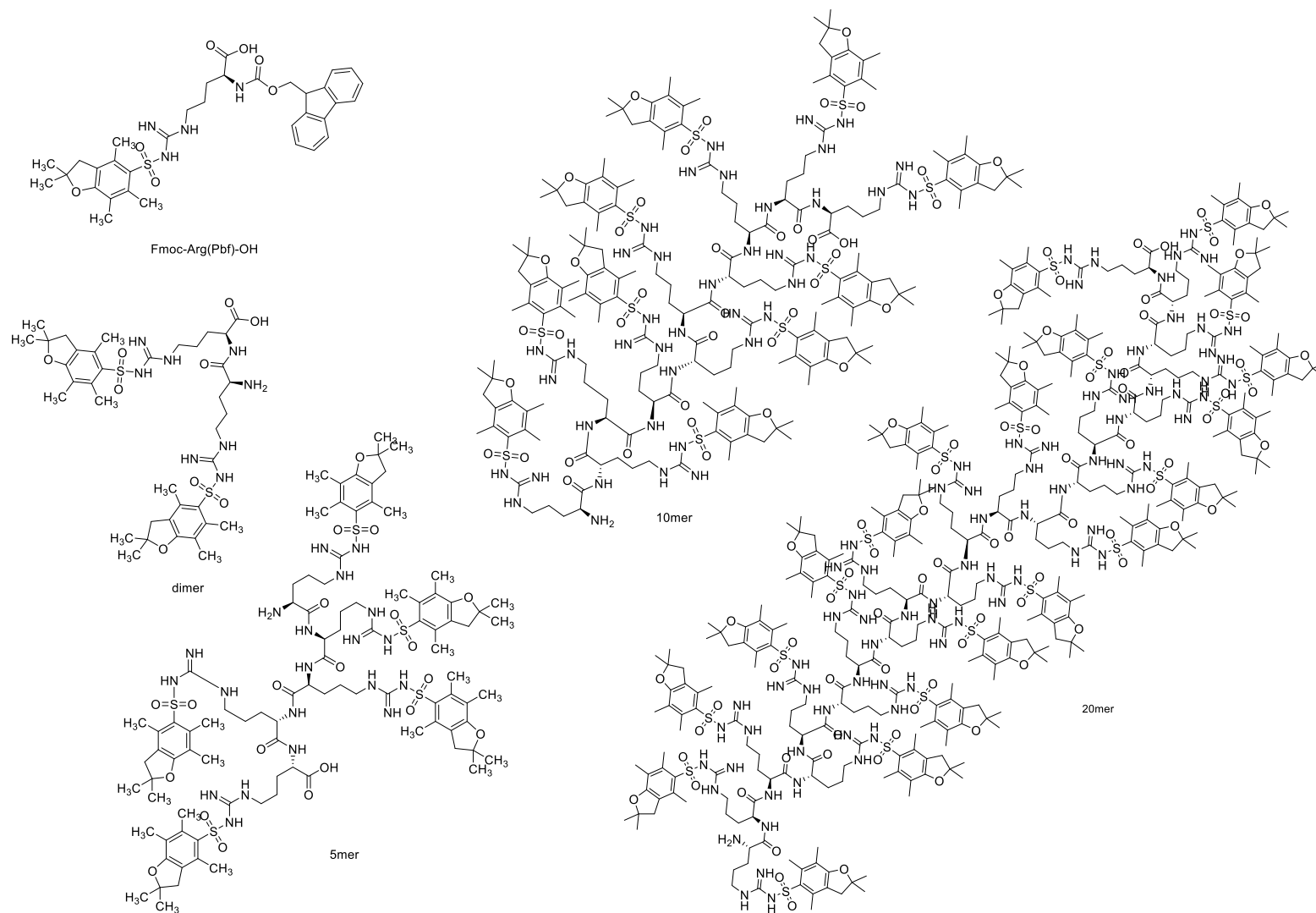
20mer

Figure II.4 - Protected (Pbf) arginine (protected – Fmoc) amino acid and respective peptides (dimer, 5, 10 and 20 mer) structure.
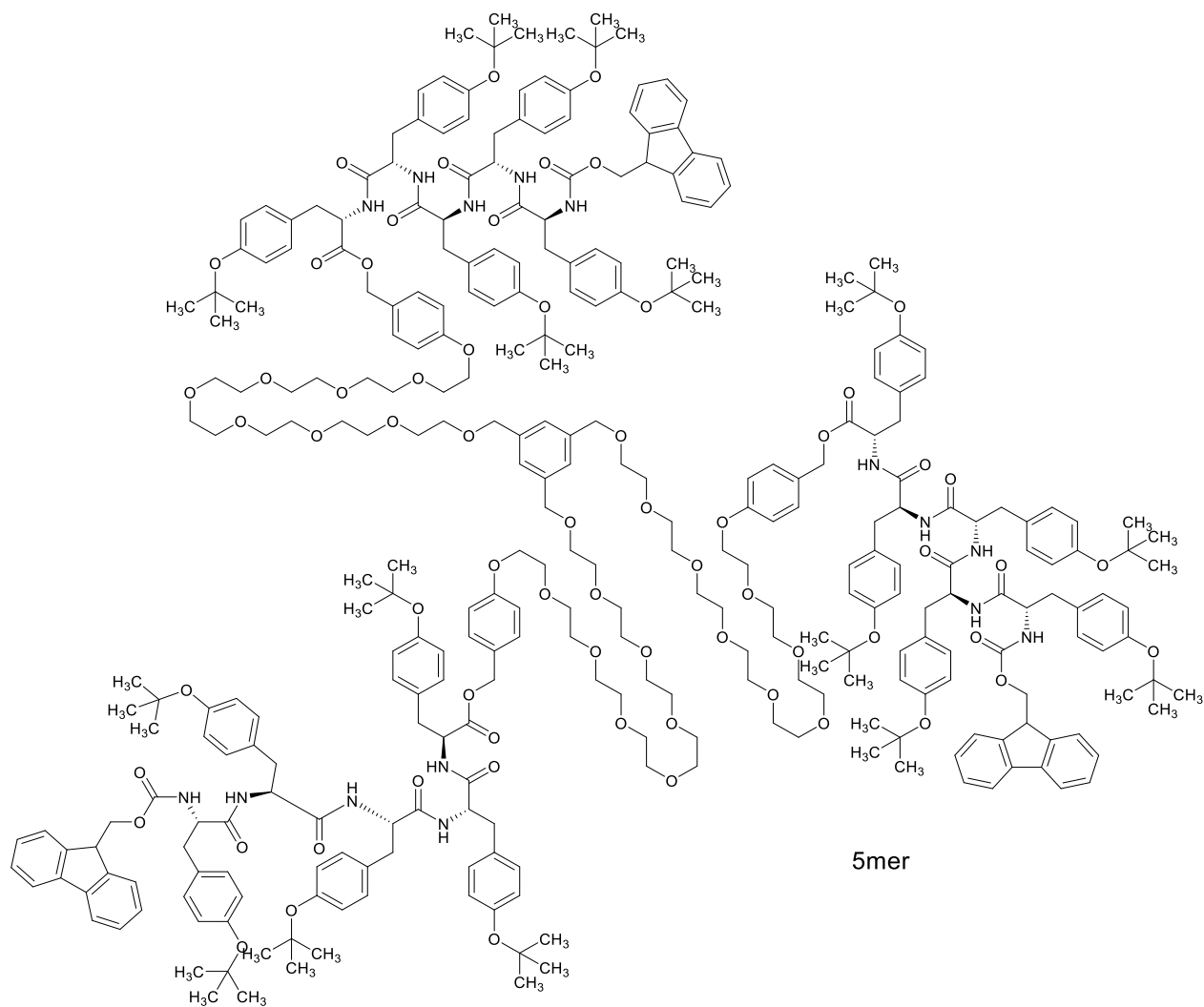
Figure II.5 - Protected (tBu) tyrosine 5 mer to wang nanostar hub.
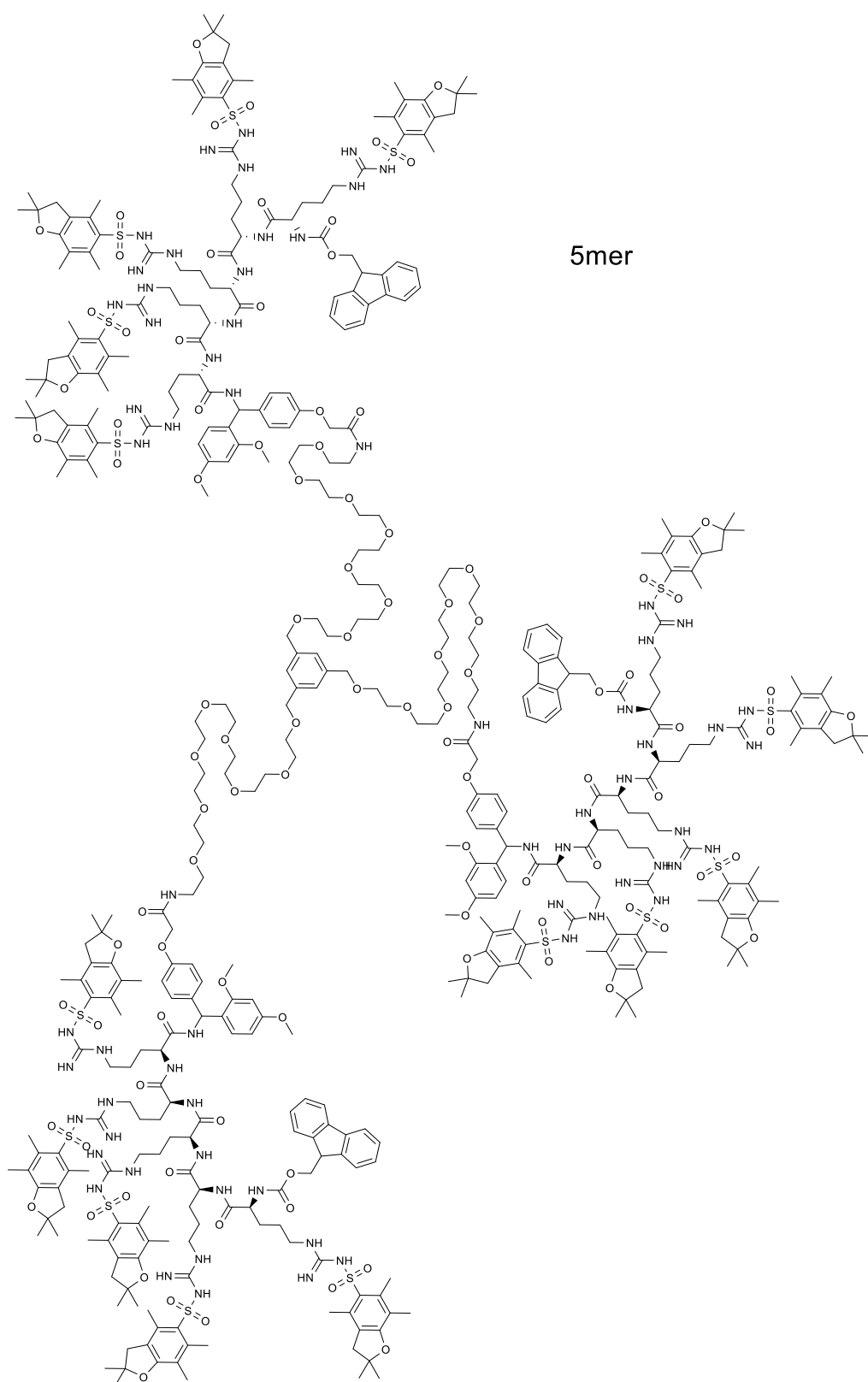
5mer

5mer

Figure II.6 - Protected (Pbf) arginine 5 mer to rink amide nanostar hub.

Appendix III

The results of: Hildebrand and Hansen Model, Table III.1 - Table III.7; COSMO-RS approach for the protected (Fmoc) amino acids and peptides, Table III.8 - Table III.11; COSMO-RS approach for the unprotected amino acids and peptides, Table III.12 - Table III.15, used in Chapter 5 (page 55).

Table III.1 - Hildebrand and Hansen models results for the (protected) glycine.

	Fmoc-Gly-OH		Fmoc-Gly-OH (dímer)		Fmoc-Gly-OH (5mer)		Fmoc-Gly-OH (10mer)		Fmoc-Gly-OH (20mer)	
	$\|\delta_{pept}-\delta_{solv}\|$	Ra	$\|\delta_{pept}-\delta_{solv}\|$	Ra	$\|\delta_{pept}-\delta_{solv}\|$	Ra	$\|\delta_{pept}-\delta_{solv}\|$	Ra	$\|\delta_{pept}-\delta_{solv}\|$	Ra
THF	6.2	8.5	10.1	14.0	10.2	14.2	12.0	17.7	12.4	18.3
MeTHF	6.2	9.9	10.1	15.6	10.2	15.8	12.0	19.2	12.4	19.9
NMP	1.6	8.2	5.4	13.3	5.6	13.5	7.3	18.4	7.8	19.0
DMSO	4.1	10.0	8.0	13.1	8.1	13.2	9.9	18.4	10.3	18.9
GVL	0.4	9.0	4.3	13.7	4.4	13.9	6.1	19.0	6.6	19.6
PC	1.2	12.8	2.7	17.5	2.8	17.6	4.5	22.9	5.0	23.5
NFM	1.1	7.1	5.0	11.1	5.1	11.3	6.9	16.4	7.3	17.0
NBP	7.2	11.2	11.0	16.7	11.2	16.9	12.9	20.0	13.4	20.6
Butanol	7.0	8.0	10.9	8.1	11.0	8.2	12.8	9.9	13.2	10.5

Table III.2 - Hildebrand and Hansen models results for the (protected) serine and tyrosine.

	Fmoc-Ser(tBu)-OH		Fmoc-Ser(tBu)-OH (dimer)		Fmoc-Ser(tBu)-OH (5mer)		Fmoc-Ser(tBu)-OH (10mer)	Fmoc-Ser(tBu)-OH (20mer)	Fmoc-Tyr(tBu)-OH		Fmoc-Tyr(tBu)-OH (dimer)		Fmoc-Tyr(tBu)-OH (5mer)		Fmoc-Tyr(tBu)-OH (10mer)	Fmoc-Tyr(tBu)-OH (20mer)
	$|\delta_{pept}-\delta_{solv}|$	Ra	$|\delta_{pept}-\delta_{solv}|$	Ra	$|\delta_{pept}-\delta_{solv}|$	Ra	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	Ra	$|\delta_{pept}-\delta_{solv}|$	Ra	$|\delta_{pept}-\delta_{solv}|$	Ra	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$
THF	1.8	6.4	1.9	13.4	2.8	6.5	3.2	2.8	2.4	8.0	2.7	5.5	3.3	8.0	3.3	0.3
MeTHF	1.8	7.8	1.9	15.0	2.8	8.0	3.2	2.8	2.4	9.5	2.7	6.9	3.3	9.5	3.3	0.3
NMP	2.9	5.6	2.8	12.9	1.9	7.1	1.5	1.9	2.3	7.6	2.0	6.3	1.4	8.0	1.4	5.0
DMSO	0.3	8.2	0.2	12.5	0.7	9.0	1.1	0.7	0.3	9.4	0.6	9.2	1.2	9.4	1.2	2.4
GVL	4.0	6.7	3.9	13.4	3.0	8.2	2.6	3.0	3.4	8.5	3.1	7.5	2.5	8.9	2.5	6.1
PC	5.6	10.6	5.5	17.2	4.6	12.4	4.2	4.6	5.0	12.4	4.7	11.5	4.1	13.0	4.1	7.7
NFM	3.3	5.1	3.2	10.6	2.3	6.2	1.9	2.3	2.7	6.5	2.4	6.1	1.8	6.6	1.8	5.4
NBP	2.8	9.2	2.9	16.1	3.8	9.3	4.2	3.8	3.4	10.8	3.7	8.2	4.3	10.7	4.3	0.7
Butanol	2.6	75	2.7	6.5	3.6	4.8	4.0	3.6	3.2	6.6	3.5	6.7	4.1	5.1	4.1	0.5

Table III.3 - Hildebrand and Hansen models results for the (unprotected) glycine and arginine.

	NH₂-Gly-OH		NH₂-Gly-OH (dímer)		NH₂-Gly-OH (5mer)		NH₂-Gly-OH (10mer)		NH₂-Gly-OH (20mer)		NH₂-Arg(Pbf)-OH		NH₂-Arg(Pbf)-OH (dímer)	
	$|\delta_{pept}-\delta_{solv}|$	Ra	$|\delta_{pept}-\delta_{solv}|$	Ra	$|\delta_{pept}-\delta_{solv}|$	Ra	$|\delta_{pept}-\delta_{solv}|$	Ra	$|\delta_{pept}-\delta_{solv}|$	Ra	$|\delta_{pept}-\delta_{solv}|$	Ra	$|\delta_{pept}-\delta_{solv}|$	Ra
THF	15.1	14.8	9.8	17.5	12.5	19.2	13.4	18.8	13.8	19.0	8.5	11.5	12.4	17.5
MeTHF	15.1	15.8	9.8	19.0	12.5	20.7	13.4	20.3	13.8	20.6	8.5	13.0	12.4	19.1
NMP	10.4	10.0	5.1	14.8	7.8	16.6	8.7	19.6	9.1	19.6	3.8	9.1	7.7	17.1
DMSO	13.0	5.9	7.7	12.5	10.4	14.2	11.3	19.3	11.7	19.1	6.4	8.0	10.3	16.2
GVL	9.3	9.3	4.0	14.6	6.7	16.3	7.6	20.2	8.0	20.1	2.7	9.2	6.6	17.4
PC	7.7	11.6	2.4	17.7	5.1	19.4	6.0	24.2	6.4	24.0	1.1	12.9	5.0	21.2
NFM	10.0	7.3	4.7	11.9	7.4	13.6	8.3	17.5	8.7	17.4	3.4	6.4	7.3	14.7
NBP	16.1	17.3	10.8	20.3	13.5	22.0	14.4	21.0	14.8	21.3	9.4	14.4	13.3	20.1
Butanol	15.9	12.3	10.6	11.4	13.3	12.8	14.2	10.4	14.6	10.6	9.3	7.4	13.2	9.7

Table III.4 - Hildebrand and Hansen models results for the (unprotected) serine and tyrosine.

	NH$_2$-Ser(tBu)-OH		NH$_2$-Ser(tBu)-OH (dímer)		NH$_2$-Ser(tBu)-OH (5mer)		NH$_2$-Ser(tBu)-OH (10mer)	NH$_2$-Ser(tBu)-OH (20mer)	NH$_2$-Tyr(tBu)-OH		NH$_2$-Tyr(tBu)-OH (dimer)		NH$_2$-Tyr(tBu)-OH (5mer)		NH$_2$-Tyr(tBu)-OH (10mer)	NH$_2$-Tyr(tBu)-OH (20mer)
	$|\delta_{pept}-\delta_{solv}|$	Ra	$|\delta_{pept}-\delta_{solv}|$	Ra	$|\delta_{pept}-\delta_{solv}|$	Ra	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	Ra	$|\delta_{pept}-\delta_{solv}|$	Ra	$|\delta_{pept}-\delta_{solv}|$	Ra	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$
THF	0.2	10.0	1.7	11.9	2.0	13.0	2.1	2.2	1.9	7.8	3.0	13.0	2.8	16.9	3.9	1.6
MeTHF	0.2	11.5	1.7	13.5	2.0	14.5	2.1	2.2	1.9	9.3	3.0	14.6	2.8	18.4	3.9	1.6
NMP	4.5	10.5	3.0	11.9	2.7	14.7	2.6	2.5	2.8	7.8	1.7	12.5	1.9	16.5	0.8	3.1
DMSO	1.9	10.6	0.4	11.6	0.1	15.3	0.0	0.1	0.2	9.4	0.9	12.2	0.7	15.7	1.8	0.5
GVL	5.6	11.3	4.1	12.5	3.8	15.6	3.7	3.6	3.9	8.7	2.8	13.0	3.0	16.9	1.9	4.2
PC	7.2	15.4	5.7	16.6	5.4	19.8	5.3	5.2	5.5	12.8	4.4	16.9	4.6	20.6	3.5	5.8
NFM	4.9	8.8	3.4	9.8	3.1	13.1	3.0	2.9	3.2	6.5	2.1	10.3	2.3	14.1	1.2	3.5
NBP	1.2	12.5	2.7	14.5	3.0	15.2	3.1	3.2	2.9	10.5	4.0	15.6	3.8	19.4	4.9	2.6
Butanol	1.0	2.8	2.5	4.3	2.8	4.5	2.9	3.0	2.7	5.4	3.8	6.3	3.6	9.1	4.7	2.4

Table III.5 - Hildebrand model results for the serine and tyrosine attached by wang linker.

	Fmoc-Ser(tBu)-OH	Fmoc-Ser(tBu)-OH (dimer)	Fmoc-Ser(tBu)-OH (5mer)	Fmoc-Ser(tBu)-OH (10mer)	Fmoc-Ser(tBu)-OH (20mer)	Fmoc-Tyr(tBu)-OH	Fmoc-Tyr(tBu)-OH (dimer)	Fmoc-Tyr(tBu)-OH (5mer)	Fmoc-Tyr(tBu)-OH (10mer)
	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$
THF	1.1	1.4	2.0	1.9	1.7	1.1	1.5	2.2	1.2
MeTHF	1.1	1.4	2.0	1.9	1.7	1.1	1.5	2.2	1.2
NMP	3.6	3.3	2.7	2.8	6.4	3.6	3.2	2.5	3.5
DMSO	1.0	0.7	0.1	0.2	3.8	1.0	0.6	0.1	0.9
GVL	4.7	4.4	3.8	3.9	7.5	4.7	4.3	3.6	4.6
PC	6.3	6.0	5.4	5.5	9.1	6.3	5.9	5.2	6.2
NFM	4.0	3.7	3.1	3.2	6.8	4.0	3.6	2.9	3.9
NBP	2.1	2.4	3.0	2.9	0.7	2.1	2.5	3.2	2.2
Butanol	1.9	2.2	2.8	2.7	0.9	1.9	2.3	3.0	2.0

Table III.6 - Hildebrand model results for the glycine and serine attached by rink amide linker.

	Fmoc-Gly-OH	Fmoc-Gly-OH (dímer)	Fmoc-Gly-OH (5mer)	Fmoc-Gly-OH (10mer)	Fmoc-Gly-OH (20mer)	Fmoc-Ser(tBu)OH	Fmoc-Ser(tBu)OH (dimer)	Fmoc-Ser(tBu)OH (5mer)	Fmoc-Ser(tBu)OH (10mer)	Fmoc-Ser(tBu)OH (20mer)
	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$
THF	2.9	3.6	5.2	6.8	7.1	2.4	2.5	2.5	1.9	2.5
MeTHF	2.9	3.6	5.2	6.8	7.1	2.4	2.5	2.5	1.9	2.5
NMP	1.8	1.1	0.5	2.1	2.4	2.3	2.2	2.2	2.8	7.2
DMSO	0.8	1.5	3.1	4.7	5.0	0.3	0.4	0.4	0.2	4.6
GVL	2.9	2.2	0.6	1.0	1.3	3.4	3.3	3.3	3.9	8.3
PC	4.5	3.8	2.2	0.6	0.3	5.0	4.9	4.9	5.5	9.9
NFM	2.2	1.5	0.1	1.7	2.0	2.7	2.6	2.6	3.2	7.6
NBP	3.9	4.6	6.2	7.8	8.1	3.4	3.5	3.5	2.9	1.6
Butanol	3.7	4.4	6.0	7.6	7.9	3.2	3.3	3.3	2.7	1.7

Table III.7 - Hildebrand model results for the arginine and tyrosine attached by rink amide linker.

	Fmoc-Arg(Pbf)OH	Fmoc-Arg(Pbf)-OH (dimer)	Fmoc-Arg(Pbf)-OH (5mer)	Fmoc-Arg(Pbf)-OH (10mer)	Fmoc-Tyr(tBu)OH	Fmoc-Tyr(tBu)-OH (dimer)	Fmoc-Tyr(tBu)-OH (5mer)	Fmoc-Tyr(tBu)-OH (10mer)
	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$	$|\delta_{pept}-\delta_{solv}|$
THF	2.7	2.8	3.0	2.9	2.4	2.5	2.6	1.9
MeTHF	2.7	2.8	3.0	2.9	2.4	2.5	2.6	1.9
NMP	2.0	1.9	1.7	1.8	2.3	2.2	2.1	2.8
DMSO	0.6	0.7	0.9	0.8	0.3	0.4	0.5	0.2
GVL	3.1	3.0	2.8	2.9	3.4	3.3	3.2	3.9
PC	4.7	4.6	4.4	4.5	5.0	4.9	4.8	5.5
NFM	2.4	2.3	2.1	2.2	2.7	2.6	2.5	3.2
NBP	3.7	3.8	4.0	3.9	3.4	3.5	3.6	2.9
Butanol	3.5	3.6	3.8	3.7	3.2	3.3	3.4	2.7

Table III.8 - COSMO-RS solubility results for the (protected) glycine.

Solvent	Fmoc-Gly-OH			Fmoc-Gly-OH (dímer)			Fmoc-Gly-OH (5mer)			Fmoc-Gly-OH (10mer)			Fmoc-Gly-OH (20mer)		
	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)
NMP	0.000	0.000	1.000	0.000	0.000	1.000	-0.270	0.276	1.887	-0.503	0.019	1.045	-10.181	-9.172	6.73E-10
MeTHF	-0.514	0.334	2.159	-0.397	0.367	2.329	-2.728	-1.724	0.019	-3.861	-2.855	1.40E-03	-18.199	-17.192	6.42E-18
Butanol	-2.664	-1.633	0.023	-3.185	-2.153	0.007	-5.455	-4.423	3.78E-05	-7.195	-6.163	6.87E-07	-24.209	-23.177	6.65E-24
DMSO	0.000	0.000	1.000	0.000	0.000	1.000	0.000	0.000	1.000	0.000	0.000	1.000	-2.039	-0.951	0.112
GVL	-0.740	0.186	1.533	-0.550	0.287	1.936	-2.374	-1.344	0.045	-3.710	-2.674	2.12E-03	-15.974	-14.937	1.16E-15
NBP	0.000	0.000	1.000	0.000	0.000	1.000	-1.502	-0.701	0.199	-2.252	-1.433	0.037	-14.182	-13.354	4.42E-14
NFM	-0.182	0.535	3.428	0.000	0.000	1.000	-1.249	-0.326	0.473	-1.955	-0.984	0.104	-12.501	-11.501	3.16E-12
PC	-1.370	-0.320	0.478	-1.300	-0.266	0.543	-3.128	-2.045	9.02E-03	-4.952	-3.869	1.35E-04	-17.480	-16.396	4.01E-17
THF	0.000	0.000	1.000	0.000	0.000	1.000	-1.624	-0.575	0.266	-2.314	-1.237	0.058	-14.748	-13.654	2.22E-14

Table III.9 - COSMO-RS solubility results for the (protected) serine.

Solvent	Fmoc-Ser(tBu)-OH			Fmoc-Ser(tBu)-OH (dímer)			Fmoc-Ser(tBu)-OH (5mer)			Fmoc-Ser(tBu)-OH (10mer)			Fmoc-Ser(tBu)-OH (20mer)		
	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)
NMP	0.000	0.000	1.000	0.000	0.000	1.000	-0.016	0.002	1.006	-1.731	-0.846	0.142	-18.670	-17.661	2.18E-18
MeTHF	0.000	0.000	1.000	0.000	0.000	1.000	-1.193	-0.386	0.412	-3.833	-2.828	1.49E-03	-21.189	-20.182	6.57E-21
Butanol	-2.049	-1.027	0.094	-2.157	-1.138	0.073	-3.398	-2.368	4.29E-03	-4.384	-3.352	4.45E-04	-15.363	-14.331	4.67E-15
DMSO	0.000	0.000	1.000	0.000	0.000	1.000	0.000	0.000	1.000	0.000	0.000	1.000	-16.627	-15.460	3.47E-16
GVL	-0.301	0.369	2.337	-0.874	-0.034	0.925	-2.340	-1.322	0.048	-5.309	-4.272	5.34E-05	-22.289	-21.252	5.60E-22
NBP	0.000	0.000	1.000	0.000	0.000	1.000	-0.913	-0.313	0.486	-3.336	-2.510	3.09E-03	-21.161	-20.333	4.64E-21
NFM	0.000	0.000	1.000	-0.334	0.221	1.662	-1.640	-0.720	0.191	-4.232	-3.231	5.87E-04	-21.679	-20.679	2.10E-21
PC	-1.087	-0.101	0.793	-2.038	-0.974	0.106	-3.805	-2.722	1.90E-03	-7.544	-6.460	3.47E-07	-24.617	-23.533	2.93E-24
THF	0.000	0.000	1.000	0.000	0.000	1.000	-0.196	-0.016	0.965	-2.085	-1.063	8.65E-02	-18.564	-17.471	3.38E-18

Table III.10 - COSMO-RS solubility results for the (protected) tyrosine.

Solvent	Fmoc-Tyr(tBu)-OH			Fmoc-Tyr(tBu)-OH (dimer)			Fmoc-Tyr(tBu)-OH (5mer)			Fmoc-Tyr(tBu)-OH (10mer)		
	log10(x_solub)	log10(S)	S (mol/l)	log10(x_solub)	log10(S)	S (mol/l)	log10(x_solub)	log10(S)	S (mol/l)	log10(x_solub)	log10(S)	S (mol/l)
NMP	0.000	0.000	1.000	0.000	0.000	1.000	-2.643	-1.647	2.25E-02	-4.357	-3.349	4.47E-04
MeTHF	0.000	0.000	1.000	-0.250	0.162	1.451	-3.743	-2.737	1.83E-03	-5.712	-4.705	1.97E-05
Butanol	-2.521	-1.493	3.21E-02	-3.347	-2.315	4.84E-03	-6.760	-5.728	1.87E-06	-10.158	-9.126	7.48E-10
DMSO	0.000	0.000	1.000	0.000	0.000	1.000	-1.965	-0.880	1.32E-01	-4.927	-3.760	1.74E-04
GVL	-0.922	-0.036	0.921	-1.098	-0.223	0.598	-5.413	-4.376	4.21E-05	-7.398	-6.361	4.35E-07
NBP	0.000	0.000	1.000	0.000	0.000	1.000	-3.738	-2.911	1.23E-03	-6.138	-5.310	4.90E-06
NFM	-0.214	0.324	2.110	-0.692	-0.002	0.996	-4.762	-3.761	1.73E-04	-7.316	-6.316	4.83E-07
PC	-2.200	-1.127	0.075	-2.311	-1.241	5.74E-02	-7.356	-6.272	5.34E-07	-9.610	-8.526	2.98E-09
THF	0.000	0.000	1.000	0.000	0.000	1.000	-2.556	-1.482	3.30E-02	-3.826	-2.734	1.84E-03

Table III.11 - COSMO-RS solubility results for the (protected) arginine.

Solvent	Fmoc-Arg(Pbf)-OH			Fmoc-Arg(Pbf)-OH (dímer)			Fmoc-Arg(Pbf)-OH (5mer)			Fmoc-Arg(Pbf)-OH (10mer)		
	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)
NMP	0.000	0.000	1.000	-0.915	-0.225	0.595	-1.004	-0.513	3.07E-01	-18.124	-17.115	7.67E-18
MeTHF	-1.076	-0.212	0.614	-2.767	-1.767	1.71E-02	-4.119	-3.114	7.70E-04	-23.641	-22.634	2.32E-23
Butanol	-3.773	-2.741	1.82E-03	-5.680	-4.647	2.25E-05	-10.215	-9.183	6.56E-10	-25.774	-24.742	1.81E-25
DMSO	0.000	0.000	1.000	0.000	0.000	1.000	0.000	0.000	1.00E+00	-12.881	-11.714	1.93E-12
GVL	-1.410	-0.451	0.354	-3.895	-2.859	1.39E-03	-5.379	-4.342	4.55E-05	-25.259	-24.222	6.00E-25
NBP	-0.690	-0.055	0.882	-2.163	-1.351	0.045	-3.420	-2.594	2.55E-03	-22.547	-21.719	1.91E-22
NFM	-1.056	-0.201	0.629	-2.835	-1.840	1.45E-02	-4.061	-3.061	8.68E-04	-23.068	-22.067	8.57E-23
PC	-2.228	-1.159	6.94E-02	-5.558	-4.474	3.36E-05	-7.729	-6.645	2.26E-07	-28.982	-27.898	1.26E-28
THF	-0.325	0.191	1.553	-1.512	-0.546	0.285	-1.775	-0.851	1.41E-01	-19.380	-18.287	5.16E-19

Table III.12 - COSMO-RS solubility results for the (unprotected) glycine.

Solvent	Gly-OH			Gly-OH (dimer)			Gly-OH (5 mer)			Gly-OH (10 mer)			Gly-OH (20 mer)		
	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)
NMP	-2.301	-1.292	0.051	-1.750	-0.742	0.181	-3.015	-2.007	0.010	-3.500	-2.493	3.22E-03	-15.119	-14.110	7.76E-15
MeTHF	-3.426	-2.420	3.80E-03	-2.999	-1.992	1.02E-02	-5.135	-4.128	7.45E-05	-7.837	-6.831	1.48E-07	-23.427	-22.420	3.80E-23
Butanol	-3.421	-2.388	4.09E-03	-3.251	-2.218	6.05E-03	-4.824	-3.792	1.62E-04	-9.946	-8.914	1.22E-09	-25.717	-24.684	2.07E-25
DMSO	-0.402	0.795	6.235	0.000	0.000	1.000	-0.647	0.309	2.039	0.000	0.000	1.000	-6.230	-5.063	8.66E-06
GVL	-3.445	-2.408	0.004	-3.020	-1.983	0.010	-4.517	-3.480	3.31E-04	-6.271	-5.234	5.83E-06	-20.098	-19.061	8.68E-20
NBP	-2.734	-1.906	0.012	-2.263	-1.435	0.037	-3.978	-3.150	7.08E-04	-5.545	-4.717	1.92E-05	-19.117	-18.289	5.14E-19
NFM	-2.591	-1.590	0.026	-2.155	-1.155	0.070	-3.489	-2.488	3.25E-03	-4.367	-3.366	4.30E-04	-16.637	-15.637	2.31E-16
PC	-3.876	-2.792	1.61E-03	-3.495	-2.411	3.88E-03	-4.875	-3.791	1.62E-04	-6.776	-5.692	2.03E-06	-20.828	-19.744	1.80E-20
THF	-3.054	-1.961	0.011	-2.539	-1.446	0.036	-4.296	-3.203	6.27E-04	-6.099	-5.005	9.88E-06	-20.060	-18.967	1.08E-19

Table III.13 - COSMO-RS solubility results for the (unprotected) serine.

Solvent	Ser(tBu)-OH			Ser(tBu)-OH (dimer)			Ser(tBu)-OH (5 mer)			Ser(tBu)-OH (10 mer)			Ser(tBu)-OH (20 mer)		
	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)
NMP	-1.782	-0.778	0.167	-1.477	-0.496	0.319	-0.663	-0.050	0.890	-5.639	-4.630	2.34E-05	-16.393	-15.384	4.13E-16
MeTHF	-2.436	-1.431	0.037	-2.565	-1.560	0.028	-1.593	-0.656	0.221	-7.614	-6.607	2.47E-07	-19.792	-18.785	1.64E-19
Butanol	-2.045	-1.015	0.097	-2.500	-1.471	0.034	-1.233	-0.355	0.442	-1.942	-0.984	0.104	-13.892	-12.860	1.38E-13
DMSO	-0.623	0.431	2.700	0.000	0.000	1.000	-0.708	-0.021	0.953	-3.787	-2.622	2.39E-03	-12.768	-11.601	2.51E-12
GVL	-2.853	-1.817	0.015	-3.108	-2.071	0.008	-1.503	-0.556	0.278	-7.869	-6.832	1.47E-07	-21.648	-20.611	2.45E-21
NBP	-2.044	-1.216	0.061	-1.983	-1.160	0.069	-1.332	-0.581	0.262	-7.048	-6.220	6.03E-07	-19.064	-18.236	5.81E-19
NFM	-2.243	-1.244	0.057	-2.239	-1.243	0.057	-1.447	-0.539	0.289	-7.199	-6.198	6.34E-07	-20.181	-19.181	6.60E-20
PC	-3.429	-2.345	0.005	-3.972	-2.888	0.001	-1.949	-0.904	0.125	-9.078	-7.994	1.01E-08	-24.985	-23.901	1.26E-24
THF	-2.143	-1.052	0.089	-2.021	-0.938	0.115	-0.846	-0.099	0.796	-6.095	-5.001	9.97E-06	-16.833	-15.739	1.82E-16

Table III.14 - COSMO-RS solubility results for the (unprotected) tyrosine.

	Tyr(tBu)-OH			Tyr(tBu)-OH (dimer)			Tyr(tBu)-OH (5 mer)			Tyr(tBu)-OH (10 mer)		
Solvent	log10(x_solub)	log10(S)	S (mol/l)	log10(x_solub)	log10(S)	S (mol/l)	log10(x_solub)	log10(S)	S (mol/l)	log10(x_solub)	log10(S)	S (mol/l)
NMP	0.000	0.000	1.000	-0.691	0.093	1.238	-2.544	-1.548	0.028	-3.800	-2.793	0.002
MeTHF	-0.724	0.193	1.560	-1.781	-0.798	0.159	-3.636	-2.631	2.34E-03	-5.684	-4.678	2.10E-05
Butanol	-1.906	-0.881	0.132	-3.257	-2.226	0.006	-4.586	-3.554	2.79E-04	-9.166	-8.134	7.34E-09
DMSO	0.000	0.000	1.000	0.000	0.000	1.000	-2.097	-0.981	0.105	-2.390	-1.281	0.052
GVL	-0.899	0.069	1.171	-2.516	-1.484	0.033	-4.685	-3.648	2.25E-04	-7.950	-6.913	1.22E-07
NBP	-0.338	0.406	2.549	-1.301	-0.512	0.308	-3.492	-2.665	2.16E-03	-5.695	-4.868	1.36E-05
NFM	-0.528	0.343	2.204	-1.752	-0.776	0.168	-4.250	-3.250	5.63E-04	-7.010	-6.009	9.80E-07
PC	-1.359	-0.305	0.496	-3.578	-2.495	0.003	-6.132	-5.048	8.95E-06	-10.742	-9.658	2.20E-10
THF	-0.325	0.513	3.260	-1.126	-0.153	0.704	-2.570	-1.492	0.032	-3.682	-2.591	0.003

Table III.15 - COSMO-RS solubility results for the (unprotected) arginine.

	Arg(Pbf)-OH			Arg(Pbf)-OH (dimer)			Arg(Pbf)-OH (5 mer)			Arg(Pbf)-OH (10 mer)		
Solvent	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)	log10(x_solub)	log10(S)	S(mol/l)
NMP	-0.624	0.169	1.477	-3.556	-2.548	2.83E-03	-4.624	-3.615	2.43E-04	-24.909	-23.901	1.26E-24
MeTHF	-1.563	-0.587	0.259	-6.004	-4.998	1.01E-05	-8.973	-7.967	1.08E-08	-33.085	-32.079	8.34E-33
Butanol	-1.771	-0.760	0.174	-5.879	-4.847	1.42E-05	-11.444	-10.412	3.87E-11	-30.336	-29.303	4.97E-30
DMSO	-0.055	0.429	2.688	-0.824	-0.059	0.873	-0.359	-0.373	0.423	-15.682	-14.515	3.05E-15
GVL	-1.843	-0.824	0.150	-6.285	-5.248	5.65E-06	-9.569	-8.532	2.94E-09	-32.941	-31.904	1.25E-32
NBP	-1.150	-0.365	0.432	-4.896	-4.068	8.55E-05	-7.442	-6.614	2.43E-07	-30.280	-29.452	3.53E-30
NFM	-1.413	-0.455	0.351	-4.979	-3.978	1.05E-04	-7.588	-6.587	2.59E-07	-29.419	-28.418	3.82E-29
PC	-2.474	-1.395	0.040	-7.558	-6.474	3.35E-07	-11.878	-10.794	1.61E-11	-36.464	-35.380	4.17E-36
THF	-0.998	-0.036	0.921	-4.710	-3.617	2.42E-04	-6.292	-5.198	6.34E-06	-28.116	-27.023	9.49E-28

The peak areas of the calibration curves and saturated solutions, Table IV.1 and Table IV.2, used in Chapter 5, section 5.6 (page 72).

Table IV.1 - Areas under the peaks of the calibration curves and 100^2-fold saturated solution and respective slope and intercept.

Calibration Curve	Fmoc-Gly-OH in THF	Fmoc-Gly-OH in Butanol		Fmoc-Ser(tBu))-OH in Butanol		Fmoc-Arg(Pbf)-OH in THF		Fmoc-Arg(Pbf)-OH in Butanol	
Peak Area (0 mg/mL)	0	0		0		0		0	
Peak Area (0.01 mg/mL)	142	138		147		16		21	
Peak Area (0.1 mg/mL)	1628	1277		1826		1529		527	
Peak Area (0.5 mg/mL)	10834	6186		12464		10502		1806	
Slope	21901	12347		25225		21349		3573	
Intercept	-189	17	0	-238	0	-244	0	44	0
Peak Area (Saturated Solution)	406	23		246		251		121	
Experimental Solub. (mg/mL)	271.68	4.46	18.63	191.70	97.48	231.67	117.43	215.23	338.37
Predicted Solub. (mg/mL)	297.32	6.92		36.05		1007.45		1.18	

Table IV.2 - Areas under the peaks of the calibration curves and 500-fold saturated solution and respective slope and intercept.

Calibration Curve	Fmoc-Gly-OH in Butanol		Fmoc-Ser(tBu)-OH in Butanol		Fmoc-Arg(Pbf)-OH in THF	
Peak Area (0 mg/mL)	0		0		0	
Peak Area (0.01 mg/mL)	138		147		16	
Peak Area (0.1 mg/mL)	1277		1826		1529	
Peak Area (0.5 mg/mL)	6186		12464		10502	
Slope	12347		25225		21349	
Intercept	17	0	-238	0	-244	0
Peak Area (Saturated Solution)	492		4913		7489	
Experimental Solub. (mg/mL)	19.22	19.92	102.09	97.38	181.11	175.40
Predicted Solub. (mg/mL)	6.92		36.05		1007.45	