# The Implication of Side-reactions in Iterative Biopolymer Synthesis: The Case of Membrane Enhanced Peptide Synthesis (MEPS)

Wenqian Chen, Mahdi Sharifzadeh<sup>1</sup>, Nilay Shah, Andrew G. Livingston

Department of Chemical Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom

# ABSTRACT

Membrane enhanced peptide synthesis (MEPS) improves conventional liquid phase synthesis by purifying intermediate products via filtration. A challenging aspect of MEPS is the propagation of unreacted materials and by-products throughout the iterative synthesis. In this study, we first develop and validate a model of MEPS. The model is then applied to investigate the implications of side-reactions (i.e. formation of error sequences) due to incomplete reaction and insufficient removal of amino acids after coupling. Sensitivity analysis shows that increasing the reaction extent for all couplings from 90% to 100% reduces the yield of truncated sequences from 29% to 0%. The formation of repeating sequences is found to be negligible in all case studies due to the large diavolume of post-deprotection diafiltration. This study provides useful insights into the operation of MEPS with particular emphasis on the control of error sequence formation. These insights are transferable to other sequence-controlled biopolymer syntheses.

# **KEYWORDS**

Iterative biopolymer synthesis, Membrane enhanced peptide synthesis (MEPS), Error sequences, Modeling and validation, Sensitivity analysis

<sup>&</sup>lt;sup>1</sup> Corresponding Author: Dr Mahdi Sharifzadeh; Room C603, Roderic Hill Building, South Kensington Campus, Imperial College London, London SW7 2AZ, UK. E-mail: <u>mahdi@imperial.ac.uk</u> ; Tel: +44 (0)7909450107

# 1. INTRODUCTION

Peptides are biopolymers that control, direct and coordinate inter- and intracellular communications and cellular functions in many living systems.<sup>1,2</sup> As the traditional field of small-molecule medicines is relatively mature, researchers in the pharmaceutical industry have turned their attention to biologics such as peptides.<sup>1,3–8</sup>



Figure 1. Comparison of the LPPS and MEPS routes for peptide synthesis.

Peptides can be synthesized chemically via step-wise addition of amino acid (AA) monomers.<sup>9–11</sup> The conventional chemical synthesis route consists of solid phase peptide synthesis (SPPS)<sup>12,13</sup> and liquid phase peptide synthesis (LPPS) (Figure 1).<sup>14</sup> Membrane enhanced peptide synthesis (MEPS) improves the LPPS route by replacing extraction, precipitation, microfiltration and drying with *diafiltration*<sup>15–18</sup>, and has the advantage of significantly shorter processing time.<sup>19,20</sup> Previous studies have shown that this iterative approach can achieve high yield and purity for peptide, oligonucleotide and mono-disperse polyethylene glycol (PEG).<sup>21–23</sup>

Wenqian Chen, Mahdi Sharifzadeh<sup>\*</sup>, Nilay Shah, and Andrew G. Livingston. Implication of Side Reactions in Iterative Biopolymer Synthesis: The Case of Membrane Enhanced Peptide Synthesis. Ind. Eng. Chem. Res., 2017, 56 (23), 6796–6804.

In MEPS, diafiltration is an operational procedure for removing impurities such as by-products, excess reagents (i.e. amino acids and piperidine) from a liquid mixture. Continuous addition of fresh solvent into the feed tank compensates the volumetric decrease due to liquid permeating through the membrane and leaving the system as waste.<sup>24</sup>

The operation of MEPS (Figure 1) involves repeated cycles of reaction and diafiltration:

- 1. Coupling amino acid to the anchored peptide via conventional Fmoc-chemistry.<sup>25</sup>
- 2. Removal of excess amino acid, other coupling reagents and by-products via constant volume diafiltration.
- 3. N-terminus deprotection of the Fmoc-protected anchored peptide by piperidine.
- 4. Removal of piperidine, other deprotection reagents and by-products via constant volume diafiltration.
- 5. Repeating steps 1 4 until the target peptide sequence is achieved.
- 6. Cleavage and global deprotection to obtain the target peptide.

The soluble anchor<sup>23,26</sup> plays a vital role in MEPS (Figure 1). In general, peptides can easily pass through membrane during filtration due to their linear structures. However, after they are attached to an anchor which has significantly higher molecular weight and normally a branched structure, the anchored peptide is retained by the nanofiltration membrane during constant volume diafiltration. At the same time, the excess reagents and by-products are completely removed through the membrane. This is critical for minimizing side-reactions, which can occur in the presence of residual (i.e. unreacted excess) amino acids and piperidine.

However, there is a trade-off between the purity and overall yield: as the impurities permeate through the membrane, so does a small fraction of the valuable product.<sup>15</sup> Since the branched structure of anchor significantly improves the rejection of the anchor over its linear counterpart according to previous studies,<sup>21,27</sup> it is preferable to select an anchor with multiple arms attached to the core.<sup>23</sup>

Sequence-controlled polymerization is of paramount importance, especially for biopolymers such as peptides and oligonucleotides, whose sequences determine their biological functions. A considerable number of biological and chemical approaches have been developed for preventing or minimizing the formation of error sequences.<sup>28–34</sup>

From the process operation perspective, two key decision variables in MEPS are the reaction extent for each coupling and the extent of removal of amino acid and piperidine before a new coupling. These variables have significant implications for the formation of general error sequences (i.e. exclusive of

sequence-specific side reactions<sup>35</sup>). The resulting error sequences can be classified into three categories<sup>36,37</sup> (using a target peptide AA(3)-AA(2)-AA(1) as an example):

- 1. Truncated sequence (e.g. AA(2)-AA(1)).
- 2. Deletion sequence (e.g. AA(3)-AA(1)).
- 3. Repeating sequence (e.g. AA(2)-AA(2)-AA(1)).

The truncated sequence refers to AA(2)-AA(1) that misses the coupling with AA(3) due to an incomplete reaction (i.e. reaction extent < 100%). Similarly, the deletion sequence is formed due to the incomplete reaction between AA(1) and AA(2). On the other hand, the repeating sequence AA(2)-AA(2)-AA(1) is formed when AA(2)-AA(1) reacts with the residual AA(2) instead of AA(3) in the third coupling

Using MEPS as a case study, the present research conducts a systematic analysis on the error sequence formation in step-growth polymerization. The article is organized into three parts. In the first part, a generic model of MEPS is developed in order to study the formation of target biopolymers and associated error sequences. The key observations are applied for model reduction and complexity management. In the second part, the model is applied to the synthesis of a hexapeptide and then validated with the experimental data. Finally, the validated model is used to demonstrate the implications of the two key variables (reaction extent for each coupling and the extent of removal of amino acid before the next coupling) for the formation of error sequence.

# 2. A GENERIC MODEL FOR MEMBRANE ENHANCED PEPTIDE SYNTHESIS (MEPS)

In this section, we develop a generic model of MEPS for studying the effects of the reaction extent and the diavolumes used in the diafiltrations on the propagation of error sequences in this iterative synthesis. The aim is to draw general conclusions which apply to peptide synthesis regardless of the sequence.

As discussed earlier, MEPS starts with the attachment of the first amino acid onto the anchor, which aids the subsequent anchored peptide to be retained by the membrane during diafiltrations. The addition of subsequent amino acids is achieved with cycles of coupling, post-coupling diafiltration, deprotection and post-deprotection diafiltration as shown in Figure 2.

Wenqian Chen, Mahdi Sharifzadeh\*, Nilay Shah, and Andrew G. Livingston. Implication of Side Reactions in Iterative Biopolymer Synthesis: The Case of Membrane Enhanced Peptide Synthesis. Ind. Eng. Chem. Res., 2017, 56 (23), 6796–6804.



Figure 2. Steps for synthesizing a peptide with n amino acids before cleavage and global deprotection in MEPS.

In each coupling, excess amino acid is activated and then reacts with the amino-terminus (H<sub>2</sub>N-) of the anchored peptide. Once the coupling is complete, post-coupling diafiltration is performed to remove the excess activated amino acid, coupling reagents and the by-products of coupling. The purified anchored peptide proceeds to deprotection, where the temporary amino-terminus protecting group (Fmoc-) is removed by piperidine. Once the deprotection is complete, post-deprotection diafiltration is performed to remove the piperidine, deprotection reagent and by-products. The purified deprotected anchored peptide undergoes another cycle of reactions and diafiltrations for elongation. Detailed experimental procedures can be found in Section 3.

# 2.1 REACTION NETWORK AND IMPLICATION OF SIDE-REACTIONS

The formation of error sequences is a general problem in peptide synthesis. Even when all deprotections are 100% complete, truncated and deletion sequences can be formed if couplings fail to proceed to completion.

For instance, in the first and second couplings (n=1 & n=2 in Figure 3), the anchor can miss coupling with AA(1) and then react with AA(2) to form the deletion sequence Fmoc-AA(2)-Anchor. The anchor can also miss coupling with AA(1) and then reacts with AA(1) again in the second coupling, forming the truncated sequence  $H_2N$ -AA(1)-Anchor at the second coupling (i.e. n=2).

Wenqian Chen, Mahdi Sharifzadeh\*, Nilay Shah, and Andrew G. Livingston. Implication of Side Reactions in Iterative Biopolymer Synthesis: The Case of Membrane Enhanced Peptide Synthesis. Ind. Eng. Chem. Res., 2017, 56 (23), 6796–6804.



n	Number of error sequences	n	Number of error sequences	n	Number of error sequences	n	Number of error sequences
1	1	6	1956	11	1.09E+08	16	5.69E+13
2	4	7	1.37E+04	12	1.30E+09	17	9.67E+14
3	15	8	1.10E+05	13	1.69E+10	18	1.74E+16
4	64	9	9.86E+05	14	2.37E+11	19	3.31E+17
5	325	10	9.86E+06	15	3.55E+12	20	6.61E+18

Figure 3. Combinations of possible reactions in each coupling, where n is the coupling number.

Repeating sequences can also be formed in the presence of residual amino acid from the previous coupling due to incomplete removal. Using the first and second couplings (Figure 3) again as example, after the anchor reacts with AA(1) in the first coupling, the anchored peptide can react with the residual AA(1) in the second coupling to form the repeating sequence  $H_2N-AA(1)-AA(1)$ -Anchor.

Overall, the number of reaction pathways (Figure 3) increases exponentially with the coupling number (n) according to Equation 1. For example, there are more than 5,000 reaction pathways in the last coupling of a 6-mer (i.e. n=6).

Number of reaction pathways = 
$$(n + 1)!$$
 (1)

However, since different reaction pathways can produce the same sequence, the number of unique sequences in each coupling is smaller than the number of reaction pathways (Equation 1). For example, when n=2 (Figure 3),  $H_2N$ -AA(1)-Anchor can remain unreacted, whereas  $H_2N$ -Anchor from the previous coupling (n=1) can react with  $H_2N$ -AA(1) to form the same product,  $H_2N$ -AA(1)-Anchor.

 Table 1. Unique sequences in each coupling.

Unique sequences (coupling number = n)	Number of sequences
Anchor	$1 = \frac{n!}{(n-0)!}$
AA(n)-Anchor,, AA(1)-Anchor	$n = \frac{n!}{(n-1)!}$
AA(n)-AA(1)-Anchor,, AA(1)-AA(1)-Anchor AA(n)-AA(2)-Anchor,, AA(1)-AA(2)-Anchor  AA(n)-AA(n-1)-Anchor,, AA(1)-AA(n-1)-Anchor	$n \times (n-1) = \frac{n!}{(n-2)!}$
AA(n)-AA(1)-AA(1)-Anchor,, AA(1)-AA(1)-AA(1)-Anchor AA(n)-AA(2)-AA(1)-Anchor,, AA(1)-AA(2)-AA(1)-Anchor  AA(n)-AA(n-1)-AA(n-2)-Anchor,, AA(1)-AA(n-1)-AA(n-2)-Anchor	$n \times (n-1) \times (n-2) = \frac{n!}{(n-3)!}$
	··· ···
AA(n)-AA(1)AA(1)-AA(1)-Anchor,, AA(1)-AA(1)AA(1)-AA(1)-Anchor AA(n)-AA(2)AA(1)-AA(1)-Anchor,, AA(1)-AA(2)AA(1)-AA(1)-Anchor  AA(n)-AA(n-1)AA(2)-AA(1)-Anchor,, AA(1)-AA(n-1)AA(2)-AA(1)-Anchor	$n! = \frac{n!}{(n-n)!}$
Target sequence = <u>AA(n)-AA(n-1)AA(2)-AA(1)-Anchor</u>	$Sum = 1 + n + n \times (n - 1) + \dots + n! + n!$

A closer examination of the reaction network reveals an interesting pattern for the formation of unique sequences. In order to quantify the total number of unique sequences, sequences are grouped according to their length (i.e. number of amino acids) in Table 1. For example, the unreacted anchor is in the second row, whereas one amino acid attached to the anchor (e.g. AA(1)-Anchor) is in the third row. It is observed that the number of sequences with  $\alpha$  amino acids, where  $\alpha \leq n$ , is given by Equation 2.

Number of unique sequences with 
$$\alpha$$
 amino acids =  $\frac{n!}{(n-\alpha)!}$  (2)

As a result, the total number of unique sequences can be calculated by Equation 3, which is the summation of all formulations in the second column of Table 1. Interestingly, when n approaches infinity, the well-known Euler's number appears (Equation 4).<sup>38</sup> Since there is only one target sequence in each coupling, the number of error sequences can then be calculated by Equation 5, showing that the number of unique error sequences increase exponentially with the coupling number (n).

Number of unique sequences 
$$= \frac{n!}{n!} + \frac{n!}{(n-1)!} + \frac{n!}{(n-2)!} + \dots + \frac{n!}{1!} + \frac{n!}{0!}$$
 (3a)

Number of unique sequences = 
$$n! \times \left[\sum_{i=0}^{n} \left(\frac{1}{i!}\right)\right]$$
 (3b)

$$e = \sum_{i=0}^{\infty} \left(\frac{1}{i!}\right) \tag{4}$$

Number of error sequences = 
$$\left[n! \times \sum_{i=0}^{n} \left(\frac{1}{i!}\right)\right] - 1$$
 (5)

For instance, the synthesis of a 5-mer can produce more than 300 error sequences, whereas more than six trillion  $(6.61 \times 10^{-18})$  error sequences can be formed in the synthesis of a 20-mer (Figure 3). The exponential growth of error sequences makes the process simulation extremely complicated. Therefore, it is necessary to apply justifiable assumptions that suppress the computational complexities.

Although the *diversity* of error sequences increases exponentially with the number of coupling (Equation 5), the *quantity* of error sequences that originate from a mis-coupling should be conserved. For instance, the unreacted anchor from the first coupling (H<sub>2</sub>N-Anchor) has three possible reaction pathways in the second coupling (Figure 4a): 1) remains unreacted; 2) reacts with residual H<sub>2</sub>N-AA(1) from the first coupling; 3) reacts with Fmoc-AA(2). At the end of the second coupling, the total number of moles of these error sequences must be equal to the total number of moles of the unreacted anchor at the beginning.

This observation means the reaction pathway can be greatly simplified by only including new error sequences immediately upon formation, which represent all the corresponding error sequences in the subsequent reactions. For example, the unreacted anchor in the first coupling represents all three corresponding error sequences in the second coupling (Figure 4b). In general, the error sequence representatives have 100% conservation of quantity during reactions, but can be removed through the membrane during diafiltrations.



Figure 4. Combinations of possible reactions in each coupling, where n is the coupling number.

Another important experimental insight about MEPS is the fact that the presence of residual amino acid from two couplings or more before the current coupling can be ignored due to the extensive washing by the two diafiltrations per addition of amino acid. For instance, in the third coupling (n=3), residual AA(1) can be ignored and only AA(2) and AA(3) are considered for the formation of the error sequences.

Wenqian Chen, Mahdi Sharifzadeh\*, Nilay Shah, and Andrew G. Livingston. Implication of Side Reactions in Iterative Biopolymer Synthesis: The Case of Membrane Enhanced Peptide Synthesis. Ind. Eng. Chem. Res., 2017, 56 (23), 6796–6804.

Considering these observations, only two relevant error sequences (truncated and repeating sequences) are formed in each coupling (Figure 5). The truncated sequence is formed due to incomplete coupling, while the repeating sequence is the result of residual amino acid from the previous coupling. As a result, the simplified model needs to consider only a manageable number of error sequences, while preserving the model fidelity.



n	Number of error sequences	n	Number of error sequences	n	Number of error sequences	n	Number of error sequences
1	1	6	11	11	21	16	31
2	3	7	13	12	23	17	33
3	5	8	15	13	25	18	35
4	7	9	17	14	27	19	37
5	9	10	19	15	29	20	39

Figure 5. Combinations of possible reaction pathways in the tailored model, where n is the coupling number.

For example, in the sixth coupling (n=6), only 11 error sequences are considered in the simplified model, which is only 0.6% of all possible error sequences. The effect of simplification is even more evident for the synthesis of a 20-mer, as the number of error sequences necessary for modeling the last coupling will be only 39 instead of 6.61 trillion.

# 2.2 MASS BALANCE IN REACTIONS AND DIAFILTRATIONS

The process model calculates the mass balance (*mole-based*) of the anchored peptide, error sequences, amino acids and piperidine in couplings, deprotections and diafiltrations. The process consists of steps, which are assigned specific step numbers (i). For example, the first coupling between anchor and Fmoc-AA(1) is step 1 and the following diafiltration is step 2. The quantities of the anchored peptide, error sequences, amino acids and piperidine are calculated at the beginning and the end of each step.

A steady state model was developed and the consumption of amino acid and formation of anchored peptide was determined by the extent of the reactions in a specific coupling ( $\xi_{coupling}$  in Equation 6 and

7). In deprotections, piperidine remained unconsumed (Equation 8), whereas the formation of deprotected anchored peptide is determined by the reaction extent ( $\xi_{deprotection}$  in Equation 9).

Any unreacted peptide in a coupling becomes the truncated sequence in the next coupling. As explained earlier, the quantity of these truncated sequences is conserved in the coupling (Equation 10). The formation of repeating sequence in a specific coupling is determined by the reaction extent and quantity of residual amino acid from the previous coupling (Equation 11).

$$M_{amino-acid-f} \quad (i,j) = M_{amino-acid-initia} (i,j) \times \left[1 - \varepsilon_{coupling}(i)\right] \tag{6}$$

$$M_{product-final}(i,k) = M_{product-initia} (i,k-1) \times \varepsilon_{coupling}(i)$$
(7)

$$M_{piperidine-fi}$$
  $(i) = M_{piperidine-initial}(i)$  (8)

$$M_{product-final}(i,k) = M_{product-initi} \quad (i,k-1) \times \varepsilon_{deprotection}(i) \tag{9}$$

$$M_{product-final}(i,l) = M_{product-initi} \quad (i,l)$$
(10)

$$M_{error-repeat-}$$
  $(i,m) = M_{amino-acid-in}$   $(i,j) \times \varepsilon_{coupling}(i)$  (11)

where *i* is the step number (1, 5, 9, ... for coupling; 3, 7, 11, ... for deprotection). j (1, 2, 3, ...) is the species number (dimensionless) for amino acid. *k* (1, 2, 3, ...) and *l* (1, 2, 3, ..., k-1) are the species numbers (dimensionless) for anchored peptide. *m* (=j) is the species number (dimensionless) for repeating sequence. M<sub>XXX-initial</sub> and M<sub>XXX-final</sub> are quantities of a compound (XXX = amino acid, piperidine or anchored peptide) (mol) at the beginning and the end of each coupling.

For all diafiltrations, the quantities of components at the end are calculated by the equation for singlestage membrane system (Equation 12)<sup>24</sup>, which only requires the inputs of diavolume (W) and rejection of each component (R). Diavolume (W) refers to the ratio between cumulative permeate volume and the system liquid volume.<sup>39</sup>

$$M_{XXX-final}(i) = M_{XXX-initia}(i) \times e^{-W(i) \times (1 - R_{XXX})}$$
(12)

where *W* is the diavolume (dimensionless) and  $R_{XXX}$  is rejection of a compound (XXX = amino acid, piperidine or anchored peptide) (dimensionless).

In order to quantify the target and error sequences relative to the starting material (i.e. anchor), the overall yields of these sequences are calculated at the end of each step (i.e. reaction or diafiltration). The overall yield (*mole-based*) is defined as the quantity of the particular sequence relative to the quantity of the anchor at the beginning of the first coupling (Equation 13).

$$Overall \ yield = \frac{M_{product-final}(i,k)}{M_{anchor-init} \quad (1)} \times 100\%$$
(13)

where *i* is the step number (1, 5, 9, ... for coupling; 3, 7, 11, ... for deprotection) and *k* is the species numbers (dimensionless) for anchored peptide (1, 2, 3, ...).

In summary, the key decision variables in the current model are the reaction extent for each coupling and the solvent diavolumes for post-coupling and post-deprotection diafiltrations. The impact of these variables on the formation of error sequences and the overall yield is discussed in Section 4.

### 3. EXPERIMENTAL STUDY OF MEMBRANE ENHANCED PEPTIDE SYNTHESIS (MEPS)

The development of the generic model was based on the valuable insights about the overall reaction network. As a case study, the generic model was applied to the synthesis of a model hexapeptide (sequence and anchor not disclosed due to collaborator confidentiality agreement). Experiment was performed in order to validate the model. All the materials and procedures involved are reported in the following.

# 3.1 MATERIALS AND SINGLE-STAGE MEMBRANE SYSTEM

All chemicals were used as received from the commercial suppliers unless stated otherwise. All amino acids and 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were purchased from Merck Millipore, whereas diisopropylethylamine (DIEA), 1-hydroxybenzotriazole hydrate (HOBt) and piperidine were purchased from Sigma-Aldrich. Tetrahydrofuran (THF) was purchased from VWR. The organic-solvent-stable ceramic membrane (Material: TiO<sub>2</sub>/Al<sub>3</sub>O<sub>3</sub>, 19-channel, length = 25 cm, channel diameter = 3.5 mm, pore size = 0.9 nm, molecular weight cut-off (MWCO) =  $450 \text{ g} \cdot \text{mol}^{-1}$ ) was purchased from Inopor GmbH.

The single-stage cross-flow membrane system (Figure 6) for MEPS consisted of a feed tank, a membrane unit, a heat exchanger and several tanks for solvent, permeate and drainage. The feed tank had a working volume between 0.4 L and 2.5 L. The system was pressurized by compressed nitrogen from the top of the feed tank with a maximum working pressure of 25 bar. Pump 1 pumped fresh solvent into the system against the compressed nitrogen during constant volume diafiltration. Pump 2 was a circulation pump (flow rate  $\approx 1.7 \text{ L} \cdot \text{min}^{-1}$ ) which ensured good mixing in the system and controlled the flow direction of the liquid.

Wenqian Chen, Mahdi Sharifzadeh<sup>\*</sup>, Nilay Shah, and Andrew G. Livingston. Implication of Side Reactions in Iterative Biopolymer Synthesis: The Case of Membrane Enhanced Peptide Synthesis. Ind. Eng. Chem. Res., 2017, 56 (23), 6796–6804.



Figure 6. Single-stage membrane system used for MEPS.

# **3.2 EXPERIMENTAL PROCEDURE**

MEPS consisted of three major steps before cleavage and global deprotection. The coupling of amino acid with anchored peptide and N-terminus deprotection was repeated for the elongation of the peptide. Once a reaction was finished, constant volume diafiltration was performed: fresh solvent was pumped constantly into the system to maintain a constant liquid volume in the system, as the excess reagents and by-products permeated through the membrane. Each post-coupling and post-deprotection diafiltration required 4 and 14 diavolume of fresh THF respectively. All the reactions and diafiltrations took place in the single-stage membrane system (Figure 6).

# (i) LOADING OF FIRST AMINO ACID ONTO ANCHOR

Fmoc-protected amino acid (1.05 eq) was dissolved in THF, followed by HOBt (1.05 eq) and DIEA (2.00 eq). HBTU (1.05 eq) was dissolved separately in DMF and the solution was added to the reaction solution. Anchor (1.00 eq) was added to the reaction solution, which was stirred at room temperature for 1 hour. The reaction was monitored by high performance liquid chromatography (HPLC). After the reaction was complete, diafiltration was performed with 4 diavolumes of fresh THF (400 mL per diavolume).

# (ii) N-TERMINUS DEPROTECTION

Piperidine and HOBt was added into the purified anchored peptide/THF solution (concentrations of piperidine and HOBt were 5 weight% and 0.1M respectively). The reaction solution was stirred at room

temperature for 2 hours. The reaction was monitored by HPLC. After the reaction was complete, the diafiltration was performed with 14 diavolumes of fresh THF (400 mL per diavolume).

## (iii) COUPLING OF AMINO ACIDS WITH ANCHORED PEPTIDE

Fmoc-protected amino acid (1.05 eq with respect to the starting anchor quantity), HOBt (1.05 eq) and DIEA (2.00 eq) were added into the purified deprotected anchored peptide/THF solution. HBTU (1.05 eq) was dissolved separately in DMF and the solution was added to the reaction solution. The reaction solution was stirred at room temperature for 1 hour. The reaction was monitored by HPLC. After the reaction was complete, diafiltration was performed with 4 diavolumes of fresh THF (400 mL per diavolume).

# 4. PROCESS MODEL VALIDATION WITH EXPERIMENTAL DATA

The experimental data of MEPS of the model hexapeptide before cleavage and global deprotection is presented in Table 2. The anchored peptide and error sequences share the same high rejection values (99.7%) with the anchor, whereas amino acids and piperidine have relatively low rejection (33.0%). All deprotections achieved completion and all couplings had close to 100% reaction extent. Each post-coupling and post-deprotection diafiltration had 4 and 14 diavolumes respectively.

<b>P</b> ainstion (aming agids)	33.0%
Rejection (annuo acius)	33.070
<b>Rejection (piperidine)</b>	33.0%
Rejection (anchor)	99.7%
Rejection (anchored peptide)	99.7%
Rejection (deletion sequence)	99.7%
Rejection (repeating sequence)	99.7%
Reaction extent (all deprotections)	100%
Reaction extent (coupling 1 - 5)	99%
Reaction extent (coupling 6)	95%
Diavolume (post-coupling)	4
Diavolume (post-deprotection)	14

Table 2. Experimental data of MEPS for model validation.

These data were used in the process model of MEPS for calculating the mass balance of the anchored peptide product and error sequences. The purity and yields of anchored peptide product and error sequences from experiment and simulation are presented in Table 3. The simulation data have good agreement with the experimental ones, showing that the model has a good representation of the actual process.

 Table 3. Experimental and modelling results of single-stage MEPS.

	Experimental	Simulation
	1	

Wenqian Chen, Mahdi Sharifzadeh\*, Nilay Shah, and Andrew G. Livingston. Implication of Side Reactions in Iterative Biopolymer Synthesis: The Case of Membrane Enhanced Peptide Synthesis. Ind. Eng. Chem. Res., 2017, 56 (23), 6796–6804.

Purity (%)	88	88
Yield* (anchored peptide product) (%)	71	68
Yield* (error sequences) (%)	10	7

\*The overall yield was before cleavage and global deprotection (i.e. the peptide was still bound to the anchor).

### 4.2 SENSITIVITY ANALYSIS

Since MEPS is a combination of reactions and diafiltrations, the operating conditions of reactions and diafiltrations have significant influences on the overall performance (i.e. overall yield of anchored peptide product vs overall yield of error sequences).

For couplings and deprotections, a close to 100% reaction extent means more desired product and less truncated sequence passing down to the subsequent reactions. Incomplete couplings can be caused by side-reactions<sup>35</sup> or hindrance for mass transfer. Side-reactions can be avoided by adopting a suitable chemical route, whereas hindrance for mass transfer can be overcome by increasing the concentration of reactants or the reaction time.<sup>22</sup>

For post-coupling and post-deprotection diafiltrations respectively, the removal of amino acid and piperidine prevents the side-reaction in the following reactions. For example, piperidine was reduced to less than the detection threshold, in order to prevent the deprotection of peptide chain in the following coupling. Amino acid must be removed before the next coupling in order to prevent the formation of repeating sequences. The removal of amino acids and piperidine in a diafiltration is proportional to the diavolumes used.

After validating the process model with experimental data, sensitivity analysis was performed on the two key decision variables of MEPS, namely the reaction extent for all couplings as well as the diavolume of post-coupling diafiltrations. Based on the experimental data (Table 2), the reaction extent for all deprotections was fixed at 100% and the diavolume of all post-deprotection diafiltrations was fixed at 14.

# 4.2.1 EFFECT OF POST-COUPLING DIAFILTRATION ON THE OVERALL YIELD

In this sensitivity analysis, the implications of excess amino acids were studied. All reactions were assumed 100% completion, in order to decouple their effects. Here, the diavolume of post-coupling diafiltrations was varied in order to investigate the effect of post-coupling diafiltration on the overall yield of the final anchored peptide product.

In each post-coupling diafiltration, amino acid is removed through the membrane. As shown in Figure 7, the quantity of residual amino acid in the system decreases exponentially as the diavolumes increase.

For example, when 3.00 equivalent amino acid is used in a coupling, 2.00 equivalent is unreacted. After 8 diavolumes, only 0.01 equivalent amino acid remains in the system.



Figure 7. Removal of amino acid in post-coupling diafiltration.

The overlapping lines in Figure 8 show that the post-coupling diafiltration has negative impact on the overall yield of the final anchored peptide product, regardless how many equivalent of amino acid is used in the coupling. This means the decrease of the overall yield is mainly attributed to the loss of anchored product through membrane during diafiltration rather than the formation of repeating sequences.





Figure 9 shows that the formation of repeating sequences is negligible even in the worst case scenario where 3.00 equivalent of amino acid was used in each coupling without any post-coupling diafiltration. This finding is surprising as the presence of unreacted amino acid definitely leads to the formation of repeating sequences in the next stage.



Figure 9. Effect of diavolume of each post-coupling diafiltration on the overall yields of repeating sequences.

The modelling data in Table 4 and 5 reveal that significant removal of amino acid occurs in the postdeprotection diafiltration, where the quantity of piperidine is reduced from 12.49 equivalent to 0.01 equivalent (i.e. 99.9% removal). As a result of this demanding requirement of residual piperidine level, 10.6 diavolumes are needed per post-deprotection diafiltration. With the same diavolume, the residual amino acid present in the system is also reduced by 99.9%. For example, when 3.00 equivalent amino acid is used per coupling, only 0.002 equivalent amino acid is left after the diafiltration with 10.6 diavolume (Table 5). As a result, the formation of repeating sequences is not an issue even when no post-coupling diafiltration is conducted.

 Table 4. Piperidine removal in post-deprotection diafiltration.

Diavolume (post-deprotection)	
Quantity of piperidine at the beginning of post-deprotection diafiltration	
Quantity of residual piperidine at the end of post-deprotection diafiltration	

Diavolume	10.6			
Quantity of amino acid at the beginning of coupling	3.00 eq			
Quantity of amino acid at the beginning of post-coupling diafiltration	2.00 eq			

Table 5. Amino acid removal in post-deprotection diafiltration.

	Quantity of amino acid at the end of post-coupling diafiltration	0.002 eq	1
--	--	----------	---

### 4.2.2 EFFECT OF COUPLING COMPLETION ON THE OVERALL YIELD

In this sensitivity analysis, the reaction extent was varied from 100% to 90% for all couplings. Figure 10 shows that this has greater impact on the overall yield of anchored peptide product and error sequences than the diavolume of each post-coupling diafiltration. The overall yield of anchored peptide product decreases considerably from 80% to 41% as the reaction extent for all couplings decreases from 100% to 90%.



Figure 10. Effect of reaction extent on the overall yields of target and error sequences.

The formation of the truncated sequences becomes prominent as the reaction extent for all couplings decreases (Figure 10). For instance, when the reaction extent is 90%, the truncated sequences account for 29% of the final yield, whereas the overall yield of the target sequence is only 41%. Therefore, the formation of truncated error sequences can be prevented by allowing all reactions to proceed to completion.

On the other hand, the formation of repeating sequences is negligible ( $\sim 0.004\%$  overall yield) and insensitive to the change in reaction extent for all couplings due to the extensive removal of amino acid in post-deprotection diafiltrations.

#### 4. CONCLUSIONS

The reaction network for peptide synthesis is highly complex as the number of possible side-products (error sequences) grows exponentially with the number of couplings. Nonetheless, many pathways can lead to the same error sequences. This study quantifies the unique error sequences, demonstrating that

Wenqian Chen, Mahdi Sharifzadeh<sup>\*</sup>, Nilay Shah, and Andrew G. Livingston. Implication of Side Reactions in Iterative Biopolymer Synthesis: The Case of Membrane Enhanced Peptide Synthesis. Ind. Eng. Chem. Res., 2017, 56 (23), 6796–6804.

it is in the order of n factorial (n!) where n is the number of coupling. Considering all unique error sequences in the peptide synthesis process poses a formidable computational challenge. Based on the mechanistic insights of the peptide chemistry and membrane filtration, a process model of MEPS was developed to account for the major error sequences. The process model was validated with the experimental data for the MEPS of a 6-mer, showing good agreement between the experimental and modelling results. Sensitivity analysis was then conducted on two key decisions variables in MEPS, namely the extent of the coupling reaction and the diavolume of each post-coupling diafiltration. Modelling results demonstrated that the reaction extent for all couplings has significant impact on the overall yield of the final anchored peptide product. Increasing the reaction extent for all couplings from 90% to 100% improves the overall yield of the final anchored product from 41% to 80%, while reducing the overall yield of truncated sequences from 29% to 0%. On the other hand, the post-coupling diafiltrations are found to be unnecessary, as 99.9% of unreacted amino acid is removed together with piperidine in the post-deprotection diafiltration. The sensitivity analysis also shows that the formation of truncated sequences is more prominent than the formation of repeating sequences in all case studies, reaffirming the assumption made in model development that the residual amino acids from two couplings or earlier are negligible.

Wenqian Chen, Mahdi Sharifzadeh<sup>\*</sup>, Nilay Shah, and Andrew G. Livingston. Implication of Side Reactions in Iterative Biopolymer Synthesis: The Case of Membrane Enhanced Peptide Synthesis. Ind. Eng. Chem. Res., 2017, 56 (23), 6796–6804.

### ORCID

Wenqian Chen: 0000-0001-8867-3012 Mahdi Sharifzadeh: 0000-0002-7895-5646 Nilay Shah: 0000-0002-8906-6844 Andrew G. Livingston: 0000-0003-0074-1426

# Notes

The authors declare no competing financial interest.

### ACKNOWLEDGEMENT

Part of the current work was supported by the European Community's Seventh Framework Programme (FP7/2007-2013) [grant number 238291].

### REFERENCES

- (1) Korhonen, H.; Pihlanto, A. Bioactive Peptides: Production and Functionality. *Int. Dairy J.* **2006**, *16* (9), 945.
- (2) Danquah, M.; Agyei, D. Pharmaceutical Applications of Bioactive Peptides. *OA Biotechnol. [E]* **2012**, *1*, 1.
- (3) Kelley, W. S. Therapeutic Peptides: The Devil Is in the Details. *Nat Biotech* 1996, 14 (1), 28.
- (4) Leader, B.; Baca, Q. J.; Golan, D. E. Protein Therapeutics: A Summary and Pharmacological Classification. *Nat Rev Drug Discov* **2008**, *7* (1), 21.
- Trusheim, M. R.; Aitken, M. L.; Berndt, E. R. Characterizing Markets for Biopharmaceutical Innovations: Do Biologics Differ from Small Molecules? *Front. Heal. POLICY Res.* 2010, 13 (1), 4.
- (6) Thayer, A. M. Improving Peptides. *Chem. Eng. News* **2011**, *89* (22), 13.
- (7) da Silva, F. P.; Machado, M. C. C. Antimicrobial Peptides: Clinical Relevance and Therapeutic Implications. *Peptides* **2012**, *36* (2), 308.
- (8) Craik, D. J.; Fairlie, D. P.; Liras, S.; Price, D. The Future of Peptide-Based Drugs. *Chem. Biol. Drug Des.* **2013**, *81* (1), 136.
- (9) Lloyd-Williams, P.; Albericio, F.; Giralt, E. *Chemical Approaches to the Synthesis of Peptides and Proteins*; CRC Press, 1997; Vol. 10.
- (10) BODANSZKY, M. In Search of New Methods in Peptide Synthesis. A Review of the Last Three Decades. *Int. J. Pept. Protein Res.* **2009**, *25* (5), 449.
- (11) Isidro-Llobet, A.; Álvarez, M.; Albericio, F. Amino Acid-Protecting Groups. Chem. Rev. 2009, 109 (6), 2455.
- (12) Albericio, F. Solid-Phase Synthesis: A Practical Guide; CRC Press, 2000.
- (13) Benoiton, N. L. Solid-Phase Synthesis. Chem. Pept. Synth. Taylor Fr. London 2005, 125.
- (14) Okada, Y. Synthesis of Peptides by Solution Methods. Curr. Org. Chem. 5 (1).

Wenqian Chen, Mahdi Sharifzadeh\*, Nilay Shah, and Andrew G. Livingston. Implication of Side Reactions in Iterative Biopolymer Synthesis: The Case of Membrane Enhanced Peptide Synthesis. Ind. Eng. Chem. Res., 2017, 56 (23), 6796–6804.

- (15) van Reis, R.; Saksena, S. Optimization Diagram for Membrane Separations. J. Memb. Sci. 1997, 129 (1), 19.
- (16) Foley, G. Ultrafiltration with Variable Volume Diafiltration: A Novel Approach to Water Saving in Diafiltration Processes. *Desalination* **2006**, *199* (1), 220.
- (17) Wang, L.; Yang, G.; Xing, W.; Xu, N. Mathematic Model of the Yield for Diafiltration Processes. *Sep. Purif. Technol.* **2008**, *59* (2), 206.
- (18) Takači, A.; Žikić-Došenović, T.; Zavargó, Z. Mathematical Model of Variable Volume Diafiltration with Time Dependent Water Adding. *Eng. Comput.* **2009**, *26* (7), 857.
- (19) So, S.; Peeva, L. G.; Tate, E. W.; Leatherbarrow, R. J.; Livingston, A. G. Membrane Enhanced Peptide Synthesis. *Chem. Commun.* **2010**, *46* (16), 2808.
- (20) So, S.; Peeva, L. G.; Tate, E. W.; Leatherbarrow, R. J.; Livingston, A. G. Organic Solvent Nanofiltration: A New Paradigm in Peptide Synthesis. Org. Process Res. Dev. 2010, 14 (6), 1313.
- (21) Kim, J. F.; Gaffney, P. R. J.; Valtcheva, I. B.; Williams, G.; Buswell, A. M.; Anson, M. S.; Livingston, A. G. Organic Solvent Nanofiltration (OSN): A New Technology Platform for Liquid-Phase Oligonucleotide Synthesis (LPOS). *Org. Process Res. Dev.* 2016, 20 (8), 1439.
- (22) Székely, G.; Schaepertoens, M.; Gaffney, P. R. J.; Livingston, A. G. Beyond PEG2000: Synthesis and Functionalisation of Monodisperse Pegylated Homostars and Clickable Bivalent Polyethyleneglycols. *Chem. - A Eur. J.* 2014, 20 (32), 10038.
- (23) Castro, V.; Noti, C.; Chen, W.; Cristau, M.; Livignston, A.; Rodriguez, H.; Albericio, F. Novel Globular Polymeric Supports for Membrane-Enhanced Peptide Synthesis. *Macromolecules* 2017, 50 (4), 1626.
- (24) Kim, J. F.; Freitas da Silva, A. M.; Valtcheva, I. B.; Livingston, A. G. When the Membrane Is Not Enough: A Simplified Membrane Cascade Using Organic Solvent Nanofiltration (OSN). *Sep. Purif. Technol.* 2013, *116*, 277.
- (25) FIELDS, G. B.; NOBLE, R. L. Solid Phase Peptide Synthesis Utilizing 9-Fluorenylmethoxycarbonyl Amino Acids. *Int. J. Pept. Protein Res.* **2009**, *35* (3), 161.
- (26) Gravert, D. J.; Janda, K. D. Organic Synthesis on Soluble Polymer Supports: Liquid-Phase Methodologies. *Chem. Rev.* **1997**, *97* (2), 489.
- (27) Zheng, F.; Li, C.; Yuan, Q.; Vriesekoop, F. Influence of Molecular Shape on the Retention of Small Molecules by Solvent Resistant Nanofiltration (SRNF) Membranes: A Suitable Molecular Size Parameter. J. Memb. Sci. 2008, 318 (1–2), 114.
- (28) Lutz, J.-F.; Ouchi, M.; Liu, D. R.; Sawamoto, M. Sequence-Controlled Polymers. *Science* (80-.). 2013, 341 (6146), 1238149.
- (29) Lewandowski, B.; Bo, G. De; Ward, J. W.; Papmeyer, M.; Kuschel, S.; Aldegunde, M. J.; Gramlich, P. M. E.; Heckmann, D.; Goldup, S. M.; D'Souza, D. M.; et al. Sequence-Specific Peptide Synthesis by an Artificial Small-Molecule Machine. *Science (80-. ).* **2013**, *339* (6116), 189.
- (30) Fukukawa, K.; Ueda, M. Sequence Control in One-Step Polycondensation. In *Polymer Science: A Comprehensive Reference, 10 Volume Set*; 2012; Vol. 5, pp 71–93.
- (31) Espeel, P.; Carrette, L. L. G.; Bury, K.; Capenberghs, S.; Martins, J. C.; Duprez, F. E.; Madder, A. Multifunctionalized Sequence-Defined Oligomers from a Single Building Block. *Angew. Chemie - Int. Ed.* 2013, *52* (50), 13261.

Wenqian Chen, Mahdi Sharifzadeh\*, Nilay Shah, and Andrew G. Livingston. Implication of Side Reactions in Iterative Biopolymer Synthesis: The Case of Membrane Enhanced Peptide Synthesis. Ind. Eng. Chem. Res., 2017, 56 (23), 6796–6804.

- (32) Vavourakis, N.; Leondiadis, L.; Ferderigos, N. A New Tag Reagent for Efficient Capping and Easy Separation of Deletion Peptides. *Tetrahedron Lett.* **2002**, *43* (46), 8343.
- (33) Eggen, I. F.; Bakelaar, F. T.; Petersen, A.; Ten Kortenaar, P. B. W.; Ankone, N. H. S.; Bijsterveld, H. E. J. M.; Bours, G. H. L.; El Bellaj, F.; Hartsuiker, M. J.; Kuiper, G. J.; et al. A Novel Method for Repetitive Peptide Synthesis in Solution without Isolation of Intermediates. *J. Pept. Sci.* 2005, 11 (10), 633.
- (34) Leibfarth, F. A.; Mattson, K. M.; Fors, B. P.; Collins, H. A.; Hawker, C. J. External Regulation of Controlled Polymerizations. *Angewandte Chemie - International Edition*. 2013, pp 199– 210.
- (35) Yang, Y. Side Reactions in Peptide Synthesis; Academic Press, 2015.
- (36) Hancock, W. S.; Prescott, D. J.; Vagelos, P. R.; Marshall, G. R. Solvation of the Polymer Matrix. Source of Truncated and Failure Sequences in Solid Phase Synthesis. J. Org. Chem. 1973, 38 (4), 774.
- (37) Azev, V. N.; Mustaeva, L. G.; Gorbunova, E. Y.; Molchanov, M. V; Rodionov, I. L. Formation of Truncated Peptide by-Products via Sequence-Specific Formyl Group Transfer from Trp (For) Residues to Nα in the Course of Boc-SPPS. J. Pept. Sci. 2013, 19 (10), 607.
- (38) Maor, E.; Sopka, K. R. The Story of E: E The Story of a Number; 1995; Vol. 33.
- (39) Székely, G.; Bandarra, J.; Heggie, W.; Sellergren, B.; Ferreira, F. C. Organic Solvent Nanofiltration: A Platform for Removal of Genotoxins from Active Pharmaceutical Ingredients. *J. Memb. Sci.* 2011, *381* (1), 21.

Wenqian Chen, Mahdi Sharifzadeh<sup>\*</sup>, Nilay Shah, and Andrew G. Livingston. Implication of Side Reactions in Iterative Biopolymer Synthesis: The Case of Membrane Enhanced Peptide Synthesis. Ind. Eng. Chem. Res., 2017, 56 (23), 6796–6804.

Table of Contents Graphic

