Polymers for proteomics: Development of polyglycidol-based soluble ester-linked aldehydes and their peptide binding characteristics

by

Patrick Edmund Beaudette

B.Sc., University of British Columbia, 2004

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Biochemistry and Molecular Biology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2009

© Patrick Edmund Beaudette, 2009
ABSTRACT

High molecular weight hyperbranched polyglycerol (HPG) was selected for development as a soluble polymer support for the targeted selection and release of primary amine containing peptides from a complex mixture. HPG has been functionalized with an ester-linked aldehyde group, that can be cleaved off under basic conditions, for the binding of primary amine containing peptides via a reductive alkylation reaction. Once bound, the high molecular weight of the polymer facilitates separation from a complex peptide mixture by employing either a 30-kDa molecular weight cut-off membrane or precipitation in acetonitrile. Following separation and washing of the polymer to remove unbound peptides and reagents, hydrolysis of the ester linker releases the bound peptide into solution for analysis by mass spectrometry. Released peptides retain the polymer linker moiety and are therefore characteristically mass-shifted. Four cleavable aldehyde polymers (PG1, PG2, PG3, and PG5) and a cleavable succinimidyl succinate polymer (PG4) have been prepared and characterized, each demonstrating the ability to selectively bind primary amine peptides from a complex mixture containing blocked (dimethylated tertiary amine) and unblocked (primary amine) peptides. The polymers have low non-specific binding properties and exhibit a range of reactivity and binding capacity. The specificity towards primary amines was designed to capture the cleavage products generated during proteolysis of a peptide library for the purpose of protease specificity profiling. PG2 was selected as the most promising candidate for having the best binding capacity of the tested polymers and no detectable non-specific binding effects.
# TABLE OF CONTENTS

Abstract ............................................................................................................................. ii  
Table of Contents .............................................................................................................. iii  
List of Tables ..................................................................................................................... iv  
List of Figures .................................................................................................................. v  
List of Abbreviations ....................................................................................................... vi  
Acknowledgements ......................................................................................................... vii  
Co-authorship Statement ............................................................................................... viii  

**CHAPTER 1. INTRODUCTION** ......................................................................................... 1  
References ......................................................................................................................... 16  

**CHAPTER 2. POLYMERS FOR PROTEOMICS: DEVELOPMENT OF POLYGLYCIDOL-BASED SOLUBLE ESTER-LINKED ALDEHYDES AND THEIR PEPTIDE BINDING CHARACTERISTICS** .......................................................... 21  
Introduction ..................................................................................................................... 21  
Materials and Methods ................................................................................................. 24  
Results ............................................................................................................................ 43  
Discussion ...................................................................................................................... 65  
References ......................................................................................................................... 71  

**CHAPTER 3. CONCLUDING DISCUSSION** ..................................................................... 74  
References ......................................................................................................................... 83
**LIST OF TABLES**

Table 2.1. PG1 developmental polymer batches .............................................. 29
Table 2.2. Synthetic test peptides ................................................................. 37
Table 2.3. Oxidation of PG1 polymers with periodic acid .............................. 45
Table 2.4. Peptide B species in hydrolysate ................................................... 49
LIST OF FIGURES

Figure 1.1. Schematic illustration of a soluble-polymer based procedure to isolate primary amine containing peptides from a complex mixture........15

Figure 2.1. Chemical syntheses of ester-linked aldehyde polymers PG1 and PG2.................................................................27

Figure 2.2. Chemical syntheses of cleavable polymers PG3, PG4 and PG5........32

Figure 2.3. \(^1\)H NMR spectrum of HPG, PG1-acid and PG1-diol.........................44

Figure 2.4. PG1 test binding and release of peptide B........................................48

Figure 2.5. Effect of aldehyde density on dialkylation ........................................50

Figure 2.6. Effect of reaction time on dialkylation................................................50

Figure 2.7. Effect of increasing PG1:peptide ratio.............................................53

Figure 2.8. Reactivity of peptide-binding polymers and binding capacity........55

Figure 2.9. Dehydration of peptide N-terminus with polymer linker covalently attached.................................................................57

Figure 2.10. Assessment of non-specific binding to polymer..............................60

Figure 2.11. Purification of peptide A from a complex mixture..........................62

Figure 2.12. PICS: PG2 pullout of MMP2 cleavage products..........................64
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACN</td>
<td>acetonitrile</td>
</tr>
<tr>
<td>ADP</td>
<td>1-amino-3,3-diethoxypropane</td>
</tr>
<tr>
<td>APD</td>
<td>3-amino-1,2-propanediol</td>
</tr>
<tr>
<td>CHCA</td>
<td>α-cyanohydroxycinnamic acid</td>
</tr>
<tr>
<td>CID</td>
<td>collision-induced dissociation</td>
</tr>
<tr>
<td>CHO</td>
<td>aldehyde</td>
</tr>
<tr>
<td>COFRADIC</td>
<td>combined fractional diagonal chromatography</td>
</tr>
<tr>
<td>DMAP</td>
<td>dimethylaminopyridine</td>
</tr>
<tr>
<td>DMF</td>
<td>dimethylformamide</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethylsulfoxide</td>
</tr>
<tr>
<td>DNP</td>
<td>2,4-dinitrophenyldihydrazine</td>
</tr>
<tr>
<td>ESI</td>
<td>electrospray ionization</td>
</tr>
<tr>
<td>GPC</td>
<td>gel permeation chromatography</td>
</tr>
<tr>
<td>$^1$H NMR</td>
<td>proton nuclear magnetic resonance</td>
</tr>
<tr>
<td>HPG</td>
<td>hyperbranched polyglycerol</td>
</tr>
<tr>
<td>ICAT</td>
<td>isotope-coded affinity tag</td>
</tr>
<tr>
<td>LC</td>
<td>liquid chromatography</td>
</tr>
<tr>
<td>MALDI-TOF</td>
<td>matrix-assisted laser desorption/ionization time-of-flight</td>
</tr>
<tr>
<td>MMP</td>
<td>matrix metalloproteinase</td>
</tr>
<tr>
<td>MS</td>
<td>mass spectrometry</td>
</tr>
<tr>
<td>MS/MS</td>
<td>tandem mass spectrometry</td>
</tr>
<tr>
<td>MudPIT</td>
<td>multidimensional protein identification technology</td>
</tr>
<tr>
<td>MWCO</td>
<td>molecular weight cut-off</td>
</tr>
<tr>
<td>NHS</td>
<td>N-hydroxysuccinimide</td>
</tr>
<tr>
<td>PAMAM</td>
<td>polyamidoamine dendrimer</td>
</tr>
<tr>
<td>PADP</td>
<td>3,3-diethoxy propanoate</td>
</tr>
<tr>
<td>PICS</td>
<td>proteomic identification of cleavage sites</td>
</tr>
<tr>
<td>RP</td>
<td>reversed-phase</td>
</tr>
<tr>
<td>SCX</td>
<td>strong cation exchange</td>
</tr>
<tr>
<td>SoPIL</td>
<td>soluble polymer-based isotopic labeling</td>
</tr>
<tr>
<td>TAILS</td>
<td>terminal amino isotopic labeling of substrates</td>
</tr>
<tr>
<td>TFA</td>
<td>trifluoroacetic acid</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

I would like to thank my supervisors, Dr. Chris Overall and Dr. Jay Kizhakkedathu, for their ideas, patience and encouragement over these last couple of years. I would also like to acknowledge my supervisory committee, comprised of Dr. Leonard Foster and Dr. Clive Roberts, for their time and helpful advice. I would like to acknowledge the contributions made by Dr. Nick Rossi, who prepared additional materials, performed the $^1$H NMR analysis and provided valuable input in my synthetic chemistry endeavours. As for the entire Overall Lab, I could not have imagined a better and more nurturing working environment, and I feel privileged to have spent my time in the company of so many wonderful scientists. Finally, I would like to thank my family for their unabating support and confidence.
CO-AUTHORSHIP STATEMENT

In collaboration with my supervisors, Dr. Overall and Dr. Kizhakkedathu, I designed and implemented the research goals outlined herein. With the exception of $^1$H NMR analyses and the preparation of polymers PG3, PG4 and PG5 by Dr. Nick Rossi, all work was performed by the author. The manuscript presented in Chapter Two of this thesis was prepared by the author with advice and feedback from my supervisors. Sections in Material and Methods relating to the preparation of PG3, PG4 and PG5 were adapted from contributions by Dr. Nick Rossi.
Chapter 1: INTRODUCTION

Sample Complexity in Proteomics

Proteomics routinely deals with a level of sample complexity that would confound other disciplines. At the protein level, genomics predicts a complement of approximately 30,000-40,000 proteins based on DNA sequence, without taking into account the numerous isoforms arising from splice variations and post-translational modifications. Complicating this further is the “shotgun proteomics” approach whereby whole proteins are digested into peptides by an endoprotease cleaving at a canonical site, usually trypsin which cleaves C-terminal to lysine and arginine. In plasma, for example, the 30,000 proteins present would yield upwards of 900,000 tryptic peptides. In addition to having a discrete mass that can be measured with accuracy and precision, peptides have more favorable solubility profiles and ionize more easily compared to proteins, making them more amenable to analysis by electrospray mass spectrometry (ESI-MS). MS/MS fragmentation of peptides yields amino acid sequence information that is used to identify the parent proteins. While MS instrumentation continues to make significant gains in sensitivity, resolution and mass accuracy, the duty cycle of the mass spectrometer is such that so-called “under-sampling” of proteomic samples hinders peptide identification. With a million peptides to analyze and an upper limit of 7,000 MS/MS spectra acquired per hour, the great majority of peptides are never analyzed. Worse yet, less than 20% of the spectra can be assigned to a peptide with confidence.

One of the problems relating to under-sampling is that multiple peptides entering the ion source simultaneously experience ion suppression, that is the preferential ionization of
certain peptides at the expense of others\textsuperscript{3}. Peptides that are not ionized do not enter the mass analyzer. Fortunately, a great number of separation techniques are available to sort peptides according to their physicochemical properties, each discrete property referred to as a ‘dimension’. Most mass spectrometers are equipped with an in-line chromatography source; a sample is injected and passes through a reversed-phase column before reaching the electrospray needle. A gradient of increasing organic solvent elutes peptide in the sample off the column and into the ion source over a period of time, with a longer gradient correlating to a reduced incidence of peptide co-elution. Reversed-phase chromatography is the most common single dimension separation and is based on the differing relative hydrophobicity of the peptides.

### Multidimensional Peptide Chromatography

Additional dimensions include separation according to molecular weight\textsuperscript{9}, electrostatic charge\textsuperscript{10} and isoelectric point\textsuperscript{11}. A highly successful in-line separation called ‘multidimensional protein identification technology’ (MudPIT) uses a strong cation exchange (SCX) column as a second dimension, arranged in series with the reversed phase column\textsuperscript{12}. Peptide samples loaded onto the SCX column are fractionated by step-wise elution by increasing the ionic strength of the buffer to displace charged peptides from the stationary phase. Each SCX elution step is transferred to the reversed-phase column and subjected to a gradient elution into the MS. Therefore if ten SCX fractions are generated, there are ten gradient elutions required. While this method does indeed reduce the sample complexity and increase the coverage of peptides by MS, it also significantly increases the amount of LC-MS instrument time required and therefore reduces how many samples can be analyzed in a given period of time. Should the separation be augmented by a third dimension, such as separation by isoelectric point
or molecular weight, the required MS time would increase by as many fold as there are additional fractions. Despite this extra burden, multidimensional chromatography is the method of choice to analyze as many peptides in a sample as possible, as is typical for protein identification	extsuperscript{13} and quantitation	extsuperscript{14} types of experiments.

**Targeted Selection of Peptides**

By utilizing targeted selection of a subset of peptides, certain fields of proteomics can add an additional dimension of sample sorting without changing the number of samples to be analyzed. Positive selection specifically isolates the peptides of interest from a complex mixture, while negative selection enriches for these peptides by depleting others from solution. The study of protein post-translational modifications frequently uses positive selection methods prior to traditional single or multi-dimensional LC-MS analysis. Immobilized antibodies	extsuperscript{15} or metal cations	extsuperscript{16} pull out phosphorylated peptides and proteins, and glycosylated peptides can be enriched by means of lectin chromatography	extsuperscript{17}. In each instance, the target analyte represents a fraction of the available peptide pool, allowing the researcher to effectively ignore the background ‘noise’ peptides. Some quantitative proteomic approaches target specific amino acids present in peptides. Isotope-coded affinity tagging (ICAT) uses a sulfhydryl specific iodoacetamide group to biotinylate the relatively rare cysteine-containing peptides that are later captured by avidin chromatography	extsuperscript{18}. Serum proteomics makes use of negative selection to deal with the ten orders of magnitude separating the most and least abundant serum proteins. Immunodepletion of the most abundant proteins, especially serum albumin, prior to tryptic digestion enhances the analysis of the lower abundance species	extsuperscript{19,20}. Targeted selection of analytes is more efficient as more MS time is dedicated to studying only those peptides that have the desired information.
Polymers as Solid Supports for Targeted Selection Strategies

Non-biological polymers are composed of linear or branched chains of a repeating chemical functionality. When serving as the physical support for other chemical processes, the ideal polymeric backbone must be able to endure the reaction conditions, with simple hydrocarbons and alkyl ether structures common for their low reactivity and high chemical and mechanical stability\(^21\). The polymer monomer functionality is modified to allow for the covalent attachment of other groups. In organic chemistry, hydrophobic polystyrene resins, stabilized by divinylbenzene cross-links are the most common and serve as platforms for synthetic reactions in a variety of solvents\(^22\). Upon these base resins a large variety of linker groups have been attached to support and facilitate multi-step synthetic processes. Following each reaction step, the insoluble product-polymer conjugate is easily separated and residual unreacted chemical reagents washed away. Release of the final product is dependent on the linker, with a large number of cleavage types available to custom tailor the process, from simple acidic or basic hydrolysis to nucleophilic, metal- and photo-catalyzed mechanisms\(^22\). Easy manipulation of the polystyrene beads lends itself to automated combinatorial approaches, including the preparation of large, complex peptide libraries\(^23\).

In proteomics, a ligand attached to solid-phase polymer beads or resins is used to select for a target analyte in an aqueous solution. Hydrophilic polymers comprised of cross-linked polysaccharide chains are the most common. Agarose is a common resin, its surface hydroxyl groups activated by cyanogen bromide for the covalent attachment of a variety of ligands. Separation from unbound contaminants can be achieved in several ways. Beads can be packed into a column as a stationary phase through which
sample solutions, washing and elution buffers can be flowed. Alternatively, the beads can be separated by simple centrifugation or filtration. The nature of the interaction between the polymer-linked ligand and the target analyte will dictate the stringency of the washing protocol. Analytes captured by non-covalent means must be treated gently and washed with mild buffers so as not to disrupt or denature structures involved in the binding. Covalently captured analytes, however, will only be released by a chemical cleavage and therefore can withstand the harsher washing conditions required to effectively remove non-specific contaminants.  

**Solid-phase Polymers in Targeted Selection**

Antiphosphotyrosine and streptavidin are examples of ligands frequently coupled to the surface of agarose resins to capture their target peptides by non-covalent affinity. As the immobilized ligands are proteins, certain precautions must be taken to maintain their structural integrity lest the basis for the affinity interaction is destroyed. Indeed, elution from such resins is possible by applying harsh denaturing solvents to unfold the protein. Elution by competitive displacement of the analyte by flooding the resin with high concentrations of an excess ligand (e.g. biotin elution from streptavidin) is also effective and helps to preserve the resin for future use. Both the polymer itself and the immobilized proteins provide extensive surfaces capable of non-specifically binding a range of contaminants. There are also agarose resins functionalized with reactive groups designed for the solid-phase covalent capture of peptides. Hydrazide-functionalized resins have been used to enrich for peptides containing tryptophan residues and glycosylated peptides, in each case the peptide is specifically modified to incorporate an aldehyde group that reacts with hyrazide. Cysteine-peptides, frequently targeted by their highly nucleophilic sulfhydryl side-chain, have been linked to
resins functionalized with haloaceto\textsuperscript{30} or maleimido\textsuperscript{31} groups. The purification of azide-containing peptides on an alkyne resin by copper (I) catalyzed click chemistry (3+2 cycloaddition) has also been described\textsuperscript{32}. In each of the above methods, peptides purified by solid-phase covalent capture are released from the resin by means of a chemical or enzymatic cleavage\textsuperscript{29}. This follows rigorous washing to remove contaminants, involving solvents such as methanol, acetonitrile, dichloromethane and 8 M urea to disrupt non-specific interactions responsible for contaminant retention.

The use of insoluble resins for the affinity or covalent capture of peptides in targeted selection techniques has been widespread. Although very successful, solid-phase chemistry is not without its downsides. Largely due to the heterogeneous conditions involved, reaction kinetics tend to be non-linear and solution-phase chemical reagents have unequal access to the immobilized functionalities\textsuperscript{33}. In addition the macromolecular supports provide many surfaces for nonspecific binding that can result in contamination and sample loss\textsuperscript{24}. Only relatively recently have soluble polymers become popular as physical supports for peptide enrichment procedures. Soluble polymers permit the restoration of the homogeneous liquid-phase conditions and the associated kinetic enhancements while maintaining the benefit of facile product purification by means of filtration or precipitation\textsuperscript{21,34}. An ideal soluble polymer support should meet the following requirements. It should be commercially available or easily prepared in the laboratory, chemical and mechanical stability of the backbone is essential, it should be chemically functionalized to allow for easy attachment of organic moieties, and it should have good solubilizing power so it remains soluble even when hydrophobic entities are attached. Finally, as polymer properties can vary with chain
length, the range of molecular weights (i.e. the polydispersity) should be as narrow as possible\textsuperscript{21}.

**Targeted Selection and Proteolysis**

Proteolysis is a post-translational modification of great significance as it alters both protein structure and function irreversibly\textsuperscript{35}. Moreover, every protein within a proteome experiences proteolysis. For example, N-terminal methionine\textsuperscript{36} and signal peptide removal\textsuperscript{37}, propeptide cleavage to activate a zymogen in protein maturation\textsuperscript{38} as well as general degradation. In addition, proteolysis can facilitate the careful regulation of entire signaling cascades by precise proteolytic activation/inactivation of a number of important proteins, for example the protein factors involved in blood coagulation or the processing of chemokines in inflammation\textsuperscript{39}. Proteases form the second largest enzyme family, with 569 members in man\textsuperscript{40}. Inappropriate or uncontrolled proteolysis has been linked to a number of pathologies, including cancer\textsuperscript{41} and chronic inflammatory diseases\textsuperscript{39}, as well as viral\textsuperscript{42} and parasitic\textsuperscript{43} diseases. For this reason, proteases have been the focus of drug discovery programs in the pharmaceutical industry, with 5-10\% of human proteases under investigation as potential therapeutic targets\textsuperscript{44}. There are extensive efforts to map out the degradome\textsuperscript{45}, the complete repertoire of proteases that are produced by an organism, and proteomics techniques have proven to be powerful tools for this effort. Understanding a protease involves the study of two critical facets: elucidating the endogenous substrate repertoire and hence function of the protease, and the characterization of the protease active site (cleavage site specificity). Targeted selection methods are therefore needed to study those peptides and proteins undergoing proteolysis.
Substrate discovery screens are vital to understanding the biological role of a protease and its function in a disease process, especially with regard to its utility as a therapeutic target. The protease web\(^{41}\) refers to the interconnections and associations of various protease cascades and circuits; perturbation of one process (i.e. by protease inhibition) can have unpredictable and deleterious effects on an organism if it is not done with precision. A key to understanding proteases is to identify all of the natural substrates a protease will act on in vivo, since proteins are embedded in complex proteomes and linked to other proteases in the protease web. Proteomics techniques are utilized to identify protease cleavage sites, branded by the formation of a new primary amine created by hydrolysis of a peptide bond in a native proteome\(^{46}\). The challenge is to identify these neo-N-termini and distinguish these from the natural N-termini present in a proteome. While complex enough, this task is rendered more problematic by the presence of the N-termini of the tryptic peptides that are invariably generated in most proteomics workflows.

The combined fractionation diagonal chromatography technique (COFRADIC)\(^{47}\) blocks the primary amines of natural and neo-N-termini by acetylation after protease treatment but before tryptic digestion. Following trypsinization, LC-fractionated internal tryptic peptides bearing a primary amine react with 2,4,6-trinitrobenzenesulfonic acid and become markedly more hydrophobic, resulting in a retention time shift during a subsequent LC analysis. This allows natural and protease-generated N-termini to be sorted for a more focused analysis. Inherent in this technique is the necessity for multiple chromatographic runs, as typically 96 fractions are generated for MS/MS analysis, so consuming significant instrument time. This problem has been addressed in our laboratory by employing a targeted selection method that adds another dimension to
sample simplification without creating more fractions and without necessitating more MS
time. In this method, termed TAILS (terminal amino isotope labeling of substrates)\textsuperscript{46},
isotopically heavy or light formaldehyde blocks the natural and neo N-terminal primary
amines by dimethylation. These peptides are separated from tryptic peptides, which are
irreversibly bound to a novel soluble aldehyde polymer that facilitates their removal from
solution. Isotopic labeling provides a quantitative aspect to compare samples, subtract
background proteolysis and determine which peptides are natural N-termini and which
are protease cleavage sites. The polymer used is a highly-branched polyglycidol that is
functionalized with terminal aldehydes. Its impressive binding capacity and low non-
specific binding properties make it an attractive reagent for proteomics.

Cleavage site specificity involves the recognition of an amino acid sequence by a
protease. The conventional nomenclature designates residues C-terminal to the actual
cleavage site as ‘prime’ (P') and those N-terminal as ‘non-prime’ (P)\textsuperscript{48}. These residues
form favorable interactions with complimentary subsites (S' and S, respectively) in the
protease active site. By compiling large numbers of protease cleavage sites, a
consensus sequence is established. Such information is invaluable to the design of
protease inhibitor drugs as well as synthetic peptide analogues used in cleavage
assays\textsuperscript{44,49}.

Peptide libraries are useful tools for cleavage site specificity studies as they provide a
wealth of sequence variety that the test protease is presented with. Peptide libraries
prepared by synthetic combinatorial chemistry incorporate a fluorescent quencher and
fluorophore, such that cleavage is detected by fluorescence emission of the free
fluorophore\textsuperscript{50,51}. Edman sequencing of pools of cleaved peptides, as opposed to
discrete peptides, determines the consensus prime side and non-prime residues. The process is very time-consuming, requiring two consecutive rounds of Edman sequencing, and prevents the analysis of subsite cooperativity\textsuperscript{27,52}. In a method called proteomic identification of cleavage sites (PICS)\textsuperscript{27}, developed in our laboratory, natural peptide libraries are prepared by tryptic digestion of a proteome, streamlining the process by making use of databases to identify the sequence of the prime and non-prime residues simultaneously. All primary amines in the library are blocked by dimethylation\textsuperscript{53,54} prior to treatment with the test protease. Proteolysis generates a new primary amine on the prime-side cleavage products, and this moiety is targeted for enrichment using a positive selection method that conjugates it to a cleavable NHS-biotin reagent. Streptavidin chromatography extracts these cleavage products from solution for LC-MS/MS analysis, identifies the sequence and the source peptide, and provides the sequence of the non-prime side from the database. PICS can identify hundreds of cleaved sequences in a single analysis, making it a powerful technique for determination of cleavage site specificity and a general amino acid consensus motif.

The use of streptavidin Sepharose beads for the solid-phase pull-out of cleavage products in PICS, while powerful, has drawbacks. Compared to the TAILS method of targeted selection, the binding capacity in PICS is lower and the non-covalent nature of the biotin-streptavidin interaction precludes rigorous washing. This fact coupled to the heterogeneous reaction conditions results in non-specific retention and loss of peptides. Therefore it was hypothesized that the TAILS polymer could be adapted for PICS by incorporation of a cleavable linker between the polyglycidol backbone and the peptide-binding aldehyde, to provide a positive selection technique for PICS with enhanced capacity and reduced non-specific binding and peptide losses.
Soluble Polymers in Targeted Selection

In proteomics workflows, the class of soluble polymer that has been adapted for targeted peptide selection are known as dendrimers. Dendrimers, also known as ‘starburst polymers’, are a type of branching polymer with a well-defined architecture. Multiple annular shells of covalently linked branching units surround a core, giving a highly symmetrical spherical shape with maximized terminal density of functional groups on the surface. The branching is so precise that the number of surface groups can be predicted mathematically, making dendrimers essentially a unimolecular solution of a singular molecular weight species. Polyamidoamine (PAMAM) was the first dendrimer described in the literature, and it features a surface functionalized with the reactive primary amine group. In its unmodified form, PAMAM has been used to covalently bind phosphopeptides. Phosphate groups activated with carbodiimide and imidazole react with the polymer amines to form stable phosphoramidate bonds. A 5-kDa molecular weight cut-off membrane facilitates separation and treatment with acid hydrolyzes the bond and release the bound phosphopeptides into solution. Compared to the same reaction on an amine-functionalized solid-phase resin, the capture of phosphopeptides on the soluble PAMAM dendrimer returned higher yields due to improved homogeneous reaction kinetics and the higher reagent excess permitted by the densely functionalized polymer.

PAMAM has been modified for the positive selection of other peptides. In a method referred to as soluble polymer-based isotopic labeling (SoPIL) the surface amines have isotope-coded cleavable linkers attached with cysteine-reactive bromoaceto groups. Following acid-catalyzed release, identical Cys-peptides from two samples can be distinguished and quantified relative to each other. Once again, significant kinetic gains
were made over an analogous solid-phase peptide labeling system. In addition to covalent capture, dendrimer-immobilized metal ions take advantage of the strong and selective bidentate chelation of phosphate groups to zirconium oxide or titanium oxide to purify phosphopeptides by affinity. In each above instance, separation of the dendrimer-peptide conjugate has been afforded by size-selective filtration. A bifunctionalized PAMAM has also been created with a chemical handle that enables a solid-phase pullout of the dendrimer-peptide conjugate. Following binding of peptides, an alkyne group on the dendrimer reacts to an azide group on a resin. This copper-(I)-catalyzed click chemistry reaction is highly potent and proceeds to completion even at low reagent concentrations. Homogeneous solution-phase peptide capture is combined with solid-phase isolation of the dendrimer for washing and peptide release.

Dendrimers such as PAMAM largely fulfill the requirements of soluble polymers as physical supports in targeted selection techniques. They have inert backbones and are mechanically stable, a variety of organic structures can be attached and they have essentially a polydispersity of one. The preparation of dendrimers is not trivial, though, and complicated, multi-step syntheses require significant expertise. Not only does this place upper-limits on the molecular weight (generation 5 PAMAM is 28 kDa), but also makes them expensive and less accessible to research groups. Hyperbranched polymers, by contrast, offer many of the same advantages as a dendrimer. As the name suggests, they are highly branched structures, though this branching is not as uniform as for dendrimers. They have a spherical shape and a surface that is richly functionalized with modifiable groups, anchor points for other chemical groups. Importantly, as high molecular weight hyper-branched polymers can be prepared economically in a single high-yield step, they may be of more practical use than
dendrimers. For example, polyglycidol is a flexible and highly water-soluble hyper-branched polymer with an inert aliphatic polyether backbone and a surface with many modifiable 1,2-diol groups. It has been demonstrated to have more thermal and oxidative stability than the related polyethylene glycol, a linear polymer that is ubiquitous in many biochemical and pharmaceutical applications. Like polyethylene glycol, polyglycidol is biocompatible and has shown promise in such applications as drug delivery and antigen-masking in artificial blood products. As for proteomics, the aldehyde polymer in the TAILS procedure for negative selection of internal tryptic peptides was prepared by periodate oxidation of polyglycidol. Up to 3200 aldehydes could be incorporated on one polymer molecule with a molecular weight of 565 kDa, nearly twenty times the size of the generation 5 PAMAM and with more groups per unit mass. The high molecular weight allows for easy separation, and the size and extent of functionality are easily tuned by variation of synthesis and oxidation parameters. An impressive binding capacity of 2.5 mg of peptide per 1 mg of polymer and low non-specific binding make this an enticing base polymer as a soluble support for other applications. Such polymers can be easily prepared in two days starting from simple starting materials, making them accessible to a wider group of researchers.

**Project Aim: Preparation and Development of a New Class of Soluble Support**

The goal of this work was the preparation of a soluble polymeric reagent for the targeted selection of the prime-side peptidic cleavage products generated during proteolytic treatment of a peptide library. Such a reagent

Here we modified polyglycidol for use in PICS and other proteomics approaches by making it cleavable. Peptide primary amine groups react by reductive alkylation on the
immobilized aldehydes to form a covalent bond, and a base-labile ester group in the linker makes it possible to retrieve bound peptides from the polymer for MS analysis following isolation. In this way, the polymer can be used to add a dimension of sample simplification to proteomic workflows by enriching for primary amine containing peptides without generating more fractions. Figure 1.1 schematically illustrates the procedure used to enrich for primary amine containing peptides from a complex mixture.
Figure 1.1. Schematic illustration of a soluble-polymer based procedure to isolate primary amine-containing peptides from a complex mixture.
REFERENCES


14. Ralhan, R. et al. iTRAQ-multidimensional liquid chromatography and tandem mass spectrometry-based identification of potential biomarkers of oral epithelial


28. Foettinger, A., Leitner, A. & Lindner, W. Selective enrichment of tryptophan-containing peptides from protein digests employing a reversible derivatization with


CHAPTER 2. POLYMERS FOR PROTEOMICS: DEVELOPMENT OF SOLUBLE ESTER-LINKED ALDEHYDES AND THEIR PEPTIDE BINDING CHARACTERISTICS

Introduction

Liquid chromatography/tandem mass spectrometry (LC-MS/MS) has become the primary methodology for the rapid identification and measurement of proteins in biological samples\textsuperscript{1,2}. In shotgun proteomics approaches, intact proteins are typically first broken down into their component peptides by digestion with a protease of canonical specificity such as trypsin\textsuperscript{3}. Compared to proteins, tryptic peptide masses can be determined with much more accuracy and precision\textsuperscript{4}, and are also more amenable to analysis by MS due to improved ionization and solubility properties. However, with each protein yielding many tryptic peptides, proteolytic digestion of a whole proteome significantly increases the sample complexity, so reducing the possibility that every protein present will be identified by at least two or more unique peptides\textsuperscript{5}.

To overcome what has been termed ‘undersampling’\textsuperscript{6}, peptide fractionation can be performed before MS analysis to reduce sample complexity. Fewer peptides simultaneously entering the MS ion source minimizes deleterious ion suppression effects and results in higher-quality MS/MS spectra and an increased number of confident peptide identifications. Separation at the peptide level by in-line tandem strong cation exchange (SCX) / reversed phase (RP) chromatography sorts and separate peptides according to their electrostatic properties and relative hydrophobicity prior to injection into the mass spectrometer ion source. Alternatively, fractionation can be performed at the protein level, before trypsin digestion. Separation based on

\textsuperscript{1} A version of this chapter will be submitted for publication. Beaudette, P.; Rossi, N.; Overall, C.M. and Kizhakkedathu, J. Preparation of cleavable soluble polyglycidol-based reagents for targeted peptide selection in proteomics.
electrophoretic mobility, isoelectric point, size-exclusion and affinity chromatography are all useful methods to separate a complex mixture of proteins into simpler fractions. However, while effective at reducing sample complexity, each additional dimension of sample fractionation exponentially increases the total number of samples to analyze, significantly increasing the amount of MS time required and bioinformatics analysis.

Alternatively, for certain proteomics studies, for example those involving post-translational modifications or relative quantification between samples, complex peptide mixtures are simplified by enriching for a subset of target peptides. For example, His-containing peptides are purified by nickel cation affinity chromatography, and phosphorylated peptides are captured using immobilized metal oxides (titanium, zirconium) or antibodies directed towards phosphate groups. Primary amines and cysteine sulfhydryl groups can be specifically labeled with biotin and biotinylated peptides captured on a streptavidin column. In such solid-phase methods, the affinity ligands are physically supported on an insoluble matrix, often resins or beads comprised of highly cross-linked polymers. These macromolecular supports can be packed into columns or given magnetic properties to facilitate purification by physically separating them from unbound components in the sample mixture. However, the interaction between the solution-phase peptides and the solid-phase ligand is not ideal and suffers from non-linear kinetics due to unequal access between reactive groups. Additionally, the heterogeneous nature of the reaction can result in non-specific sample loss at the solid-liquid phase interface, particularly for low-abundance species.

As an alternative, replacing insoluble resins and beads with a soluble polymer as the support allows for the restoration of more favorable homogeneous reaction conditions.
and improved kinetics/efficiency of coupling\textsuperscript{18,19}. Soluble dendrimers such as polyamidoamine (PAMAM)\textsuperscript{20} have been modified to incorporate both affinity ligands for noncovalent capture of phosphopeptides\textsuperscript{21}, and chemically reactive groups for covalent capture of phosphopeptides\textsuperscript{22} and Cys-containing peptides\textsuperscript{15}. The large peptide-dendrimer complex is easily separated from low molecular weight species using 5-kDa molecular weight cutoff filtration, and the bound peptides eluted off or released by cleavage of an acid-labile linker. Covalent capture of target peptides has the added benefit of permitting a more stringent washing and removal of non-specific contaminants\textsuperscript{17}.

In the study of proteolysis as a post-translational modification, the primary amine group generated by hydrolysis of an amide bond can serve as a handle to direct enrichment efforts\textsuperscript{23}. The primary amine at the N-terminus of the peptide cleavage products can be modified to alter the peptide hydrophobicity and change its chromatographic behavior as performed in combined fractional diagonal chromatography (COFRADIC)\textsuperscript{24}, or affinity captured by resin-immobilized streptavidin following biotinylation of peptides using small-molecule reagents\textsuperscript{13} or enzymes\textsuperscript{25}. Our group has demonstrated the utility of a soluble aldehyde-functionalized polyglycerol polymer capable of irreversibly binding internal tryptic peptides to enrich for amine-blocked N-terminal peptides in solution\textsuperscript{23}. In this method, termed TAILS (terminal amino isotope labeling of substrate), the N-terminal peptides are analyzed and sequenced to simultaneously identify cleavage sites and the identity of the substrates. In another method aimed at profiling the cleavage specificity and active site of a protease termed PICS (proteomic identification of cleavage sites)\textsuperscript{26}, cleaved peptides derived from protease-treatment of an amine-blocked, proteome-derived peptide library are conjugated to biotin and purified by streptavidin Sepharose
resin. However, due to the non-covalent nature of this interaction and the heterogenous conditions, non-specific binding and peptide losses are common.

Here we present a novel, water-soluble polymer for proteomics. Highly-branched polyglcerol, a dendrimer-like polymer, has been functionalized with ester-linked, amine-reactive aldehyde groups attached to the polymer. Primary-amine containing peptides are selectively and covalently attached by reductive alkylation and later released back into solution for MS analysis by base hydrolysis, the mass shifted characteristically because the cleaved polymer linker remnant remains attached to the N-terminus. The procedure is illustrated schematically in Figure 1.1. This cleavable aldehyde polymer is shown to have no non-specific binding and possess significantly more reactive binding groups per milligram compared to the strepatavidin Sepharose resin. Peptide enrichment by negative (e.g. TAILS) and positive (e.g. PICS) selection processes is an effective way to reduce sample complexity by adding an extra dimension in the separation process but without increasing the total number of samples to analyze.

2.1. MATERIALS & METHODS.

All chemicals were purchased from Sigma-Aldrich Canada (Oakville, ON) unless otherwise noted. All solvents were purchased from Fisher Scientific or BDH Chemicals. Peptides were chemically synthesized by means of solid-phase methods and were purified by reversed-phase high-performance liquid chromatography (HPLC)\textsuperscript{27}.

**Gel-permeation chromatography.** Polymer molecular weights were determined by gel-permeation chromatography (GPC) on a Waters 2690 separation module equipped
with a detector from Viscotek Corporation that utilized refractive index and multi-angle light-scattering. An aqueous 0.1 N NaNO₃ solution was used as the mobile phase at a flow-rate of 0.8 mL/min, and the stationary phase was an Ultrahydrogel linear column from Waters with bead size 6-13 µm (elution range 10³ – 5 x 10⁶ Da).

**Proton nuclear magnetic resonance.** The polymer product (5 mg) was dissolved in 1 mL deuterated solvent (Cambridge Isotopes, Inc) and the ¹H NMR spectrum acquired on a 300 MHz Bruker Avance Instrument.

**Thermogravimetric analysis.** Polymer concentrations in aqueous solution were found by thermogravimetric analysis (TGA) using a TA instruments Q500. A tared platinum pan holding 50 µL of sample was heated from room temperature to 300 °C at a rate of 20 °C/min to determine the weight percent of polymer in solution.

**Matrix-assisted laser desorption ionization time-of-flight mass spectrometry.** All peptide samples were analyzed by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry (MS) on one of two instruments: an Applied Biosystems (ABI) 4700 MALDI TOF-TOF, or an ABI Voyager DE-STR MALDI-TOF. The instrument was set up in positive reflector mode. Peptide samples were in a 1:1 mixture of acetonitrile (ACN)/0.5% trifluoracetic acid (TFA) containing 5 mg/ml α-cyanohydroxycinnamic acid (CHCA) as the matrix.

**Liquid chromatography / tandem mass spectrometry (LC-MS/MS) and MS/MS data analysis.** Peptide samples derived from the proteomics identification of cleavage sites (PICS) assay were analyzed injected onto a LC Packings capillary LC system (Dionex)
coupled to a quadrupole time-of-flight mass spectrometer (QStar Pulsar, Applied Biosystems). Samples were injected onto a column loaded with PepMap C18 resin (Dionex) and eluted using a 5-40% gradient of organic phase (buffer B) over 90 min. Buffer A was 2% acetonitrile (ACN) and 0.1% formic acid, while buffer B was 85% acetonitrile and 0.1% formic acid. Peak lists of the original wiff QSTAR data were converted to the mzXML format and peptides identified using the IPI database by X! Tandem. The mass tolerance was 0.4 Da for parent ions and 0.2 Da for fragment ions with up to two missed cleavages. Chymotrypsin was defined to cleave C-terminal to tryptophan, tyrosine, leucine and phenylalanine. Fixed modifications included carboxyamidomethylation of cysteine thiols (+57.02 Da), dimethylation of lysine ε-amino groups (+28.03 Da) and linker modification of the N-terminus (+139.15 Da). The possibility of linker hydration was allowed by a variable modification of 18.01 Da.

**Synthesis of ester-linked aldehyde polymers.**

**Synthesis of PG1.**

i) **Reaction of polyglycerol with succinic anhydride (Quantitative conversion of HPG 1,2-diols to ester-linked acid).** Hyper-branched polyglycerol (HPG, 444 kDa, Figure 2.1A) was prepared according to Sunder et al. HPG (0.7 g, 10 mmoles of hydroxyls) was dissolved in anhydrous pyridine (10 mL) with succinic anhydride (1 g, 10 mmoles) and 4-dimethylaminopyridine (75 mg, DMAP) and stirred for 48 h at room temperature. Ice-cold acetone (75 mL) was used to precipitate the polymer from solution, before being centrifuged for 20 min at 7000 x g. The pellet was rinsed with acetone, dissolved in ddH$_2$O (12 mL, NANOpure) and dialyzed using a 10-kDa molecular weight cut-off (MWCO) regenerated cellulose membrane (Thermo Slide-a-lyzer) for 90 min with three changes of ddH$_2$O. Finally the dialyzed solution was
Figure 2.1. Chemical synthesis of ester-linked aldehyde polymers PG1 and PG2.

A. Structural representation of polyglycerol (HPG), a dendritic and water-soluble polymer. The hydroxyl groups have been modified with each of the five R groups shown on the right.

B. Schematic of the synthesis of the first-generation cleavable aldehyde polymer, PG1.

C. Schematic of the synthesis of the second-generation cleavable aldehyde polymer, PG2.
lyophilized overnight to afford 1.03 g (59% yield) of a brittle white solid (PG1-acid, Figure 2.1B).

\(^1\)H NMR (400 MHz, d-D\(_2\)O): \(\delta\) 2.3-3.0 (broad, [-O-CO-CH\(_2\)-CH\(_2\)-COOH]), 3.4-4.0 (broad, [CH\(_2\) and CH]), 4.1-4.5 (broad, [-CH\(_2\)-O-CO-]), 4.7-4.9 (sharp, [H\(_2\)O]), 5.0-5.3 (broad, [-CH-O-CO-]).

**ii) Formation of ester-linked 1,2-diols (Reaction of PG1-acid with 3-amino-1,2-propanediol).** PG1-acid (0.45 g, 2.6 mmoles of hydroxyls) was dissolved in N,N-dimethylformamide (DMF, 15 mL) with gentle heating. Diisopropylcarbodiimide (DIC, 0.67 mL, 4.3 mmoles) and 3-amino-1,2-propanediol (APD, 0.4 g, 4.4 mmoles) were added to the stirring solution in four equal parts over the course of 2.5 h. Stirring was continued for 1.5 h. The reaction was stopped via addition of 110 mL ice-cold 0.1 M phosphate buffer, pH 7. This solution was dialyzed using 6 to 8-kDa MWCO regenerated cellulose tubing (SpectraPor) against 4 L of ddH\(_2\)O for 24 h with five water changes and filtered through a 0.45 \(\mu\)m PVDF membrane (Millipore) to remove any particulates. Lyophilization yielded 0.48 g (69% yield) of a fluffy white hygroscopic solid (PG1-diol, Figure 2.1B).

\(^1\)H NMR (400 MHz, d-D\(_2\)O): \(\delta\) 1.1-1.3 (sharp, H\(_2\)O, adsorbed), 2.3-3.0 (broad, [-O-CO-CH\(_2\)-CH\(_2\)-COOH]), 2.9-3.1 (broad, [CH\(_2\)-CO-NH]), 3.1-4.0 (broad, CH\(_2\) and CH), 4.1-4.5 (broad, [CH-O-CO-]), 4.7-4.9 (sharp, [H\(_2\)O]), 5.0-5.3 (broad, [-CH\(_2\)-O-CO-]).

**iii) Formation of ester-linked aldehydes (Periodic acid oxidation of 1,2-diols).** The oxidative cleavage of 1,2-diols to aldehydes to generate PG1 (Figure 2.1B) was achieved by treatment with periodic acid (HIO\(_4\)). PG1-diol (25 mg, 0.6 nmoles, 94 \(\mu\)moles) was dissolved in 1.1 mL 0.5 M HEPES with HIO\(_4\) (40 mg, 175 \(\mu\)moles) and stirred for 3 h at room temperature, then dialyzed overnight against 2 L ddH\(_2\)O with five
water changes using a 10-kDa MWCO membrane. All aldehyde polymers were frozen in liquid nitrogen and stored at -80 °C.

Several batches of PG1 were prepared during the development stage of this project in order to refine the synthesis, in particular to maximize the coupling of PG1-acid carboxylic acid groups to APD and to develop a protocol to better understand the peptide binding properties of this novel ester-linked aldehyde polymer. These polymers are designated as PG1-I, II, III, IV, V and VI and are listed in Table 2.1.

### Table 2.1. PG1 developmental polymer batches.

<table>
<thead>
<tr>
<th>PG1 Polymers</th>
<th>MW (Da)</th>
<th>mols aldehydes per mol polymer</th>
</tr>
</thead>
<tbody>
<tr>
<td>PG1-I</td>
<td>200,000</td>
<td>476 ± 57</td>
</tr>
<tr>
<td>PG1-II</td>
<td>444,000</td>
<td>99 ± 15</td>
</tr>
<tr>
<td>PG1-III</td>
<td>444,000</td>
<td>495 ± 40</td>
</tr>
<tr>
<td>PG1-IV</td>
<td>444,000</td>
<td>204 ± 11</td>
</tr>
<tr>
<td>PG1-V</td>
<td>444,000</td>
<td>115 ± 15</td>
</tr>
<tr>
<td>PG1-VI</td>
<td>444,000</td>
<td>191 ± 26</td>
</tr>
</tbody>
</table>

### Synthesis of PG2.

i) **Reaction of polyglycerol with succinic anhydride (Limited conversion of HPG hydroxyls to ester-linked acid).** Rather than the quantitative conversion sought in the first step of PG1 synthesis, for PG2 only a fraction of the available 1,2-diol groups present on HPG (444 kDa) were to be converted to the ester-linked acid form. HPG (0.25 g, 3.4 mmoles hydroxyls) and succinic anhydride (45 mg, 0.5 mmoles, twenty-
times lower molar ratio than for PG1-acid) were dissolved in anhydrous pyridine (5 mL) with DMAP (25 mg) and stirred for 24 h at room temperature. The polymer was precipitated from solution with ice-cold acetone (50 mL) and centrifuged for 20 min at 7000 x g. The pellet was rinsed with acetone, dissolved in ddH$_2$O (6 mL) and dialyzed for 2 h against 2 L ddH$_2$O with 3 water changes. The dialyzed polymer was lyophilized to afford 130 mg (50% yield) of a white, waxy solid (PG2-acid, Figure 2.1C).

$^1$H NMR (400 MHz, d-$D_2$O): $\delta$ = 2.4-2.6 (broad, [-O-CO-CH$_2$-CH$_2$-CO-COOH], 3.3-4.1 (broad, CH and CH$_2$).

**ii) Coupling of HPG-acid-2 to protected aldehyde groups (Reaction of PG2-acid with N-hydroxysuccinimide and 1-amino-3,3-diethoxypropane).** PG2-acid (120 mg, 1.6 mmoles hydroxyls) was dissolved in dimethylsulfoxide (DMSO, 2 mL) and combined with N-hydroxysuccinimide (NHS, 20 mg, 0.17 mmoles) and DIC (27 $\mu$L, 0.17 mmoles) and allowed to react overnight at room temperature to form the PG2-NHS ester intermediate. The next day 1-amino-3,3-diethoxypropane (ADP, 28 $\mu$L, 0.17 mmoles) was dissolved in D$_2$O (2 mL, Cambridge Isotopes, 99.9% D) and added drop-wise over 10 min to the stirring DMSO solution. Stirring was continued for 30 min. The reaction was stopped by addition of ice-cold 0.1 M phosphate buffer, (30 mL), pH 7. This solution was dialyzed using 6 to 8-kDa MWCO tubing (SpectraPor) against ddH$_2$O (4 L) for 24 h with 5 water changes, filtered through a 0.45 $\mu$m PVDF membrane and lyophilized to afford 128 mg (70% yield) of the product (PG2-acetal, Figure 2.1C).

$^1$H NMR (400 MHz, d$_6$-DMSO): $\delta$ 1.1-1.3 (t, -O-CH$_2$-CH$_3$), 1.6-1.7 ([-CH$_2$-CHOO-]), 2.2-2.3 [O-CO-CH$_2$-CH$_2$-COOH], 2.5 (sharp, DMSO residual), 3.0-4.0 (broad, CH and CH$_2$), 4.3-5.0 (broad, OH).

**iii) Deprotection of PG2-acetal (removal of acetal protecting groups to give ester-linked aldehyde).** PG2-acetal (50 mg) was dissolved in aqueous 0.4% trifluoroacetic
acid (TFA, 1 mL), pH 2 and stirred for 30 min at room temperature, after which 1 M HEPES (100 μL) was added to neutralize the pH. The solution was dialyzed for 24 h using a 10-kDa MWCO membrane against 2 L ddH₂O with five water changes. Acetal deprotection yields the aldehyde polymer PG2 (Figure 2.1C).

Synthesis of PG3.

i) Synthesis of propanoic acid 3,3-diethoxy propanoate (PADP). 3,3-diethoxy-1-propanol (0.300 g, 2.02 mmol) and succinic anhydride (0.304 g, 3.03 mmol) were dissolved in dichloromethane (15 mL). DMAP (0.370 g, 3.03 mmol) was added and the solution was left to stir overnight. Solvent was evaporated and the product was purified via silica gel column chromatography using a gradient from 100:0 to 80:20 dichloromethane:methanol. ¹H NMR (400MHz, CDCl₃, ppm) d: 1.14 (-CH₃); 1.94 (-CH₂CH₂CH-); 3.36 – 3.59 (COCH₂CH₂CO); 3.50 – 3.78 (CH₃CH₂O-); 4.14 (CH₂CO₂); 4.8 (CH).

ii) Synthesis of acetal containing HPG (PG3-acetal). HPG (Mₙ 444 kDa, 139 mg, 0.313 mmol) was dried under reduced pressure for 2 days before it was dissolved in dry DMF (2 mL) in a small round bottom flask containing a stirrer bar. Propanoic acid 3,3-diethoxy propanoate (PADP, 54.5 mg, 0.220 mmol) and DMAP (27.0 mg, 0.221 mmol) were added to the solution which was stirred for 20 min. DIC (23 mg, 0.183 mmol) dissolved in DMF (0.25 mL) was added slowly to the solution, which was left to stir overnight. The product was precipitated into acetone, dissolved in water, and dialysed against water with frequent changes for 24 h. The solution was then lyophilized and
Figure 2.2. Chemical synthesis of cleavable polymers PG3, PG4 and PG5.
A, B, and C: Schematic of the synthesis of PG3, PG4 and PG5, respectively.
PG4 uses an NHS-ester to bind peptides to the polymer. The synthesis of PG4 is identical to the intermediate PG2-NHS en route to PG2.
analyzed using $^1$H NMR; integration data showed that ~ 3% of available OH groups had been converted into ester-linked acetal protected aldehyde groups (Figure 2.2A). $^1$H NMR (400MHz, d-D$_2$O, ppm) d: 1.20 (CH$_3$CH$_2$–); 1.95 (CH$_3$CH$_2$–); 2.70 (–CO$_2$CH$_2$CH$_2$CO$_2$–); 3.5 – 4.05 (m, HPG: –CH$_2$O–, –CHO–).

iii) Deprotection of PG3-acetal.

PG3-acetal (50 mg) was dissolved in aqueous 0.4% TFA (1 mL), pH 2 and stirred for 30 min at room temperature, after which 1 M HEPES (100 µL) was added to neutralize the pH. The solution was dialyzed for 24 h using a 10-kDa MWCO membrane against 2 L ddH$_2$O with five water changes. Acetal deprotection yields the ester-linked aldehyde polymer PG3 (Figure 2.2A).


i) Synthesis of ester-linked acid polyglycerol. In a typical reaction, hyperbranched polyglycerol, HPG M$_n$ 30 kDa (1.5 g, 0.188 mmol) and 5 molar equivalents of succinic anhydride (0.0938 g, 0.938 mmol) was dissolved in DMF (10 mL). DMAP (0.115 g, 0.938 mmol) was slowly added to the solution, which was cooled in an ice bath and stirred overnight. The polymer was precipitated in acetone (60 mL) and centrifuged at 11,000 x g for 12 min to maximize yield. The supernatant was decanted and the product was re-dissolved in methanol (6 mL) and dropped into acetone (60mL) and centrifuged at 11,000 x g for 12 min. The supernatant was decanted, and the product dried vacuum at room temperature. Yield = 90%.

$^1$H NMR (400MHz, d-D$_2$O, ppm) 2.5 – 2.7 (–CO$_2$CH$_2$CH$_2$CO$_2$–); 3.5 – 4.05 (m, HPG: –CH$_2$O–, –CHO–).

ii) Synthesis of succinimidyl succinate-functionalized polyglycerol (PG4). As an example, the addition of N-hydroxysuccinimide to acid functionalized HPG (Mn 30 kDa,
10 acid groups per molecule) is described. Acid functionalized HPG (0.150 g, 0.0188 mmol) and N-hydroxysuccinimide (0.0130 g, 0.113 mmol) was dissolved in DMF (1.5 mL). DIC (0.0142 g, 0.113 mmol) was added to the solution which was left to stir overnight at room temperature. The solution was precipitated in acetone (20 mL), immediately re-dissolved in methanol (1.5 mL) and re-precipitated in acetone (20 mL). Separation of the product PG4 (Figure 2.2B) from the solvent layer was aided by centrifugation – the solution was spun at 11,000 x g for 12 min after precipitation into acetone. Acetone was decanted, and the polymer was dried for 5 min under reduced pressure at room temperature to remove as much residual acetone or methanol as possible.

\[ ^1H \text{ NMR (400MHz, d-DMSO, ppm)} \]
\[
2.72 (-CH_2CO_2C); 2.80 (succinimide: – COCH_2CH_2CO–); 2.94 (-CH_2CO_2N–); 3.25 – 3.75 (m, HPG: –CH_2O–, –CHO–), 4.35 – 4.75 (m, OH). 
\]

**Synthesis of PG5.**

**i) Synthesis of pentynoic acid 3,3-diethoxy propanoate (PADP).** 3,3-diethoxy-1-propanol (0.300 g, 2.02 mmol) and pentynoic acid (0.298 g, 3.03 mmol) were dissolved in dichloromethane (15 mL). 4-(Dimethylamino)pyridine (DMAP, 0.370 g, 3.03 mmol) was added and the solution was left to stir for 10 min. N,N'-Diisopropylcarbodiimide (DIC, 0.382 g, 3.03 mmol) in dichloromethane (5 mL) was then added and the solution left to stir overnight. Solvent was evaporated and the product was purified via silica gel column chromatography using a gradient from 90:10 to 50:50 hexane:diethyl acetate.

\[ ^1H \text{ NMR (400MHz, d-CDCl}_3, \text{ ppm)} \]
\[
d: 1.25 (-CH_3); 2.02 (-CH_2CH_2CH–; -CCH); 2.56 (-CCH_2CH_2CO_2–); 3.55 & 3.70 (CH_3CH_2O–); 4.24 (CH_2CO_2); 4.68 (CH). 
\]
ii) Synthesis of 5-azido valeric acid. 5-bromovaleric acid (18.0 g, 0.0994 mol) and thionyl chloride (30.0 mL, 0.412 mol) was added to a dry, 250 mL, one-neck rb flask equipped with a dropping funnel and a drying tube. The mixture was stirred for 3 h at room temperature, after which the remaining thionyl chloride was removed under reduced pressure. The product liquid (5-bromovaleryl chloride) was added drop-wise over a 1 h period to a solution of sodium methoxide (5.6 g) and methanol (150 mL) in an ice bath. The mixture was allowed to warm to room temperature for a further 2 h. The product (methyl 5-bromopentanoate) was collected by extraction with dichloromethane. After removal of solvent under reduced pressure, methyl 5-bromopentanoate was stirred in DMF with sodium azide. Subsequently, methyl 5-azido pentanoate was hydrolyzed using an aqueous solution of sodium hydroxide in a 1,4-dioxane/tetrahydrofuran mixture. \(^1\)H NMR (300MHz, d-CDCl\(_3\), ppm) \(\delta: 1.62-1.72 (-CH\(_2\)CH\(_2\)); 2.37-2.42 (-CH\(_2\)CO); 3.28-3.32 (N\(_3\)C\(_2\)H\(_2\)).

iii) Synthesis of azido functionalized HPG (HPG-N\(_3\)). HPG (M\(_n\) 444 kDa, 0.500 g, 1.12 mmol) was dried under high vacuum for several days and then dissolved in DMF (20 mL) in a one-necked rb flask. 5-azido valeric acid (0.095 g, 0.66 mmol), DIC (1.0 g, 7.9 mmol), and DMAP (0.10 g, 0.82 mmol) were added to the solution which was stirred for 48 h at room temperature. The solution was then dialyzed against water for 72 h with frequent changes in water. \(^1\)H NMR (300MHz, d-D\(_2\)O, ppm) \(\delta: 1.64 (-CH\(_2\)CO); 2.44 (-CH\(_2\)CO); 3.35 – 3.95 (m, HPG: –CH\(_2\)O–, –CHO–); 4.36 (-CH\(_2\)O\(_2\)C).

iv) Synthesis of acetal-containing HPG via Click chemistry (PG5-acetal). Azide-containing HPG (HPG-N\(_3\) 1%, M\(_n\) 444 kDa, 231 mg, 0.520 mmol) and PADP (11.9 mg, 52.0 mmol) were dissolved in a methanol:water mixture (80:20, 3.0 mL). CuSO\(_4\) (2.70 mg, 10.8 mmol) and ascorbic acid (4.28 mg, 22.9 mmol) was dissolved in the solution and left to stir for 24 h at room temperature. The solution was then precipitated into
acetone, added to an EDTA buffer solution (to remove Cu) and dialyzed against water (24 h) to isolate the product PG5 (Figure 2.2C) The solution was then lyophilized and analyzed using $^1$H NMR. $^1$H NMR (400MHz, d-D$_2$O, ppm) d: 1.25 (-CH$_3$); 1.45 – 1.75 (-CH$_2$CH$_2$CH$_2$CH$_2$-; CCH$_2$CH$_2$); 1.85 (-CH$_2$CH$_2$CH-; -CCH); 2.45 (-CCH$_2$CO$_2$-); 3.3 – 4.15 (m, CH$_2$-N-; CH$_3$CH$_2$O-, -CH$_2$O-, -CH-O-); 4.8 (CH).

v) Deprotection of PG-acetal groups to aldehyde groups (PG5) PG5-acetal (50 mg) was dissolved in aqueous 0.4% TFA (1 mL), pH 2 and stirred for 30 min at room temperature, after which 1 M HEPES (100 µL) was added to neutralize the pH. The solution was dialyzed for 24 h using a 10-kDa MWCO membrane against 2 L ddH$_2$O with five water changes. Acetal deprotection yields the ester-linked aldehyde polymer PG5 (Figure 2.2C).

Aldehyde titration. The aldehyde content of the polymer was assessed by an aldehyde titration method$^{29}$. Aldehyde-derivatized polymers (100-1500 µg) were combined in a solution with 50 mM 2,4-dinitrophenylhydrazine (DNP, Aldrich), 850 mM sulfuric acid (H$_2$SO$_4$, BDH Chemicals) and 4.5% ethanol for 30 min at room temperature. The polymer-DNP complex precipitated out of solution and was separated by centrifugation for 10 min at 16,250 x g. The pellet was resolubilized in DMSO (400 µL) and its optical absorbance measured at 365 nm (NanoDrop) and compared to a standard curve (0.01–1.0 mM) prepared with formaldehyde.

Solid content measurement. Two methods have been used for determining the concentration of aqueous solutions of polymer:
i) **Oven drying.** For PG1-I, II, III, IV and V polymer concentration was determined by drying a 200-μL aliquot of solution in an 80 °C oven for 24 h and measuring the residual weight.

ii) **Thermogravimetric analysis.** For other polymers, a thermogravimetric analysis (TGA) was performed.

**Peptide binding to aldehyde polymer** All peptides used in the development and characterizations of ester-linked aldehyde polymers are listed in Table 2.2.

### Table 2.2. Synthetic test peptides.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>[M+H]+</th>
<th>Sequence</th>
<th>Primary amine</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1058</td>
<td>VWESATPLR</td>
<td>yes</td>
</tr>
<tr>
<td>B</td>
<td>1159</td>
<td>INSTTCCYR</td>
<td>yes</td>
</tr>
<tr>
<td>C</td>
<td>1631</td>
<td>FYQLPSNQYQSTR</td>
<td>yes</td>
</tr>
<tr>
<td>D</td>
<td>1523</td>
<td>pyroQPVGINSTTCCYR</td>
<td>no</td>
</tr>
</tbody>
</table>

The peptide (1 nmole, from a 1 mM ddH₂O stock) was added into a tube containing aldehyde polymer (20-1000 μg), 50 mM HEPES (pH 6.5) and 20 mM sodium cyanoborohydride (NaCNBH₃, prepared fresh as a 1 M stock solution in ddH₂O). The solution was incubated overnight (16 h) with gentle agitation at room temperature. Afterwards, glycine (Fisher) was added to a final concentration of 100 mM and allowed to react for 1 h to quench any unreacted aldehyde groups. Sample was applied to a 30-kDa MWCO regenerated cellulose membrane (Amicon Microcon) and spin-filtered at 14,000 x g for 10 min to separate the peptide-polymer complex from unbound peptides. In the case of PG4, a 10-kDa MWCO membrane was used. 50 mM HEPES (200μL)
was applied to the membrane and again centrifuged for 10 min to wash the complex. This wash step was repeated four times. In initial experiments, unblocked peptide B and blocked peptide D (1 nmole each) were incubated with PG1-I (20 µg) overnight.

**Hydrolysis.** The peptide-polymer complex was reconstituted in 50 mM HEPES (100 µL) and recovered from the MWCO membrane via a 10 s inverted spin on a bench top microfuge. Ammonium hydroxide (NH₄OH, BDH Chemicals) was added to a final concentration of 100 mM and the sample heated for 3 h at 65 ºC with gentle agitation to hydrolyze the polymer ester group. After hydrolysis, the sample was briefly cooled on ice and applied to a second Microcon device and spin-filtered at 14,000 x g for 10 min, followed by 50 µL of 50 mM HEPES to rinse any residual peptide into the flow-through.

**Peptide desalting and MALDI analysis.** Solutions of peptides were acidified to pH 3 with TFA and subjected to desalting on an OMIX C18 pipet tip (Varian). After equilibration, binding and washing, peptides were eluted into 5 mg/mL α-cyanohydroxycinnamic acid (CHCA, 20 µL, Sigma) matrix in 1:1 ACN/0.5% TFA-H₂O and 0.7 µL was spotted onto a stainless steel target and MALDI-TOF mass spectra collected.

**Strategies for reducing peptide dialkylation.**

**a) Effect of aldehyde density on the polymer.** Peptides A, B, C and D (1 nmole each) were incubated with PG1-II or PG1-III (75 µg) overnight. To achieve the different aldehyde densities, different ratios of HIO₄ to polymer were used. PG1-diol (25 mg, 50 nmoles, 94 µmoles monomer) was reacted with 25 mg (110 µmoles) or 5 mg (22 µmoles) of HIO₄ to generate PG1-II and PG1-III, respectively.
b) Effect of reaction time. In two separate reactions, peptides A, B, C and D (1 nmole each) were incubated with PG1-IV (75 μg) overnight. Reaction 1 proceeded as per normal binding protocol. For reaction 2, NaCNBH₃ was introduced only for the last 30 min of the reaction period, prior to glycine quenching.

Increasing ratio of PG1:peptide. To further optimize peptide binding and investigate the relationship between peptide dialkylation and aldehyde concentration, peptides A, B, C and D (1 nmole each) were incubated with 50, 100, 200, 400 and 1000 μg of PG1-V (115 aldehydes/molecule) overnight with 20 mM NaCNBH₃.

Polymer reactivity and binding capacity. Four of the aldehyde polymers and the NHS polymer (PG4) were assessed with respect to their reactivity towards a standard peptide. For the non-cleavable aldehyde polymer HPG-CHO, a fixed amount of peptide A (VWESATPLR, 10 μg) was incubated with 25-1500 μg of polymer in a 100 μL volume with 50 mM HEPES and 20 mM NaCNBH₃ for 16 h as per binding protocol. Following binding, the peptide-polymer complex was spin-filtered with a 30-kDa MWCO membrane to separate it from unbound peptides. Peptide concentration in the flow-through was measured by absorbance at 280 nm (NanoDrop) and correlated to a calibration curve of 0-500 μg/mL peptide A. Curves were fitted using Prizm software (GraphPad). The same approach was taken for the cleavable polymers, PG1, PG2, PG3 and PG4.

Assessment of non-specific binding to polymer.

i) Bovine serum albumin tryptic digest and dimethylation of primary amines. Bovine serum albumin (BSA, Sigma) was reduced by treatment with 5 mM dithiothreitol
(DTT, Fisher) at 65 °C for 1 h, and cysteine sulfhydryl groups were alkylated by reaction with 15 mM iodoacetamide (Sigma) for 2 h at room temperature in the dark. Iodoacetamide was quenched with 30 mM DTT for 30 min and the denatured protein was precipitated by mixing with 8:1 acetone/methanol (Fisher) at -80 °C for 1 h. After the protein was resuspended in 1 M guanidine hydrochloride (Fisher) and 50 mM HEPES, pH 7.5, sequencing grade trypsin (Promega) was added at a ratio of 1:60 (w/w) and the digestion proceeded for 16 h at 37 °C. The protease was heat inactivated at 95 °C for 5 min, the sample split in two, and primary amines (N-terminal, lysine) dimethylated with 40 mM heavy (\(^{13}\text{CD}_2\text{O}\), Cambridge Isotopes) or light formaldehyde and 20 mM NaCNBH\(_3\) for 16 h at 37 °C. Dimethylated peptides were desalted by reverse-phase chromatography (Waters C18 Sep-Pak), concentrated to dryness on a SpeedVac and re-suspended in ddH\(_2\)O.

ii) Peptide incubation with polymer and washing. Light labeled peptides (20 µg) were incubated with 1000 µg of the indicated polymer (HPG, PG1 or PG2) in 100 µL of 50 mM HEPES (pH 6.5) and 20 mM NaCNBH\(_3\) for 16 h at room temperature. Heavy-labeled peptides were incubated in an identical solution, but in the absence of any polymer. The samples were spin-filtered with the 30-kDa MWCO Microcon and the flow-through collected. The regenerated cellulose membrane with the polymer on top was washed three times with 100-µL volumes of 50 mM HEPES, three times with 100-µL volumes of 2 M NaCl/50 mM HEPES and two times with 100-µL volumes of 20% ACN/50 mM HEPES. Each wash was collected in a separate tube. Cleavable polymers (PG1 and PG2) were recovered, subjected to the normal hydrolysis conditions, and filtered with a 30-kDa Microcon to remove polymer from the solution. Equivalent volumes of the initial flow-through, washes 1 to 8, and the hydrolysate of PG1 and PG2
were pooled and subjected to desalting by reversed-phase chromatography (Varian OMIX C18) before being analyzed by MALDI-TOF MS.

**Purification of primary-amine containing peptide from a complex mixture.** Peptide A (2 µg) and light-labeled BSA tryptic digest (20 µg) were incubated with PG2 (1000 µg) in 100 µL of 50 mM HEPES, pH 6.5, and 20 mM NaCNBH₃ for 16 h. Afterwards, the polymer-peptide complex was precipitated with ACN (500 µL) and centrifuged for 10 min at 14,000 x g. The supernatant was decanted and evaporated (SpeedVac) and reconstituted in 100 µL 50 mM HEPES. The pellet was re-suspended in 50 mM HEPES and spin-filtered using a 30-kDa MWCO Microcon device, then washed three times with 100-µL volumes of 50 mM HEPES, three times with 100-µL volumes of 2 M NaCl/50 mM HEPES and two times with 100-µL volumes of 20% ACN/50mM HEPES. PG2-peptide complex was reconstituted and hydrolyzed in 100 mM NH₄OH for 3 h at 65 °C. The supernatant, the flow-through from each wash step and the hydrolysate were desalted with OMIX C18 tips and analyzed by MALDI-TOF MS.

**Proteomic identification of cleavage sites (PICS) with PG2 purification.** A peptide library was prepared from HEK cell lysate²⁶. In brief, disulfide bonds were reduced by treatment with DTT and the free cysteine sulphydryl groups alkylated with iodoacetamide. The denatured proteins were cleaved into peptides by digestion with chymotrypsin, and the free 1°-amine groups (N-termini and lysine side chains) were blocked by reductive dimethylation with formaldehyde. Low molecular weight peptides (< 700 Da) were filtered out by size-exclusion chromatography, detergents removed by strong cation exchange and the peptides desalted by reverse-phase chromatography (AKTA, GE Healthcare).
Pro-human matrix metalloproteinase-2 (MMP2) was activated by incubation with 0.1 mM phenylmercuric chloride for 30 min at 37 °C and added to 200 μg of the blocked peptide library in cleavage assay buffer (100 mM HEPES, pH 7.5, 20 mM NaCl, 5 mM CaCl₂) at a ratio of 1:100 (w/w) and incubated for 16 h at 37 °C. The protease was heat-deactivated at 95 °C for 5 min, then the pH adjusted to 6.5 with TFA. PG2 (1,500 μg) was added to the protease-treated library and incubated at 22 °C for 20 h with 20 mM NaCNBH₃. Glycine (100 mM) was used to quench the reaction for 90 min, then the PG2-peptide complex precipitated with 500 μL ACN and centrifuged for 10 min at 16,250 x g. The pellet was resuspended in 50 mM HEPES, pH 7, spin-filtered with a 30-kDa MWCO Microcon and washed by spin-filtration with 3 x 200-μL 20% ACN/50 mM HEPES, 3 x 200-μL 2 M NaCl/50 mM HEPES and 3 x 200-μL 50 mM HEPES. The polymer was resuspended in 50 mM HEPES (200 μL) and recovered from the MWCO membrane by an inverted spin, NH₄OH was added to a concentration of 100 mM and the sample heated at 65 °C for 3 h. After hydrolysis of the ester linkages, the sample was briefly cooled on ice before being spin-filtered with a 30-kDa MWCO Microcon to separate the polymer from the free peptides, and the polymer rinsed with 200 μL of 50 mM HEPES. The flow-through and rinse were combined and brought to 1 mL with H₂O and the pH adjusted to pH 3 with TFA. The peptides were desalted by reverse-phase chromatography (Waters C18 Sep-Pak) and the eluent concentrated to near-dryness before resuspension in 0.3% formic acid and analysis by LC-MS/MS.
2.3. RESULTS.

Synthesis of PG1.

The first-generation ester-linked aldehyde polymer, PG1, was synthesized by the quantitative conversion of hydroxyl groups by first converting to i) ester-linked carboxylic acid groups (PG1-acid), then to ii) ester-linked 1,2-diols (PG1-diol). Oxidation of the ester-linked 1,2-diol groups gives the desired aldehyde form (PG1). The $^1$H NMR spectrum (Figure 2.3B) and aldehyde titration (Table 2.2) indicate almost quantitative conversion of 1,2-diols by succinic anhydride. The generation of ester-linked 1,2-diols from the PG1-acid precursor was by reaction with 3-amino-1,2-propanediol. The ester-linked aldehyde polymer was obtained by periodic acid oxidation. By changing the amount of periodic acid with respect to the 1,2-diols present, the amount of aldehydes generated could be varied (Table 2.3).

Periodate oxidation of 1,2-diols to aldehydes

Periodic acid treatment of the polymers oxidatively cleaves 1,2-diols into aldehydes: a polymer-linked aldehyde and free formaldehyde (Figure 2.1B). Reaction was performed in HEPES buffer (pH 7) to protect the pH sensitive ester linkage and maintain solubility of PG1-acid, which exhibited a tendency to precipitate out of solution at lower pH due to protonation of the carboxylic acid groups. Dialysis was required to eliminate all formaldehyde from the polymer solution as this would interfere with the aldehyde titration and would also result in undesired methylation of peptide amine groups during the binding reaction. Polymer supports with different number of aldehyde groups on the polymer were obtained by altering the ratio of HIO$_4$ to the number of 1,2-diols present (see Table 2.3).
Figure 2.3. $^1$H NMR spectrum of HPG, PG1-acid and PG1-diol

A. $^1$H NMR spectrum (D$_2$O) of the starting polymer material polyglycerol
B. $^1$H NMR spectrum (D$_2$O) of the intermediate PG1-acid
C. $^1$H NMR spectrum (D$_2$O) of the intermediate PG1-diol
Table 2.3. Oxidation of PG1 polymers with periodic acid.

<table>
<thead>
<tr>
<th>Polymer</th>
<th>MW (Da)</th>
<th>Polymer (mg)</th>
<th>HIO₄ (mg)</th>
<th>Mole Ratio (HIO₄:hydroxyls)</th>
<th>mols aldehydes/mol polymer</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPG-CHO</td>
<td>444,000</td>
<td>25</td>
<td>100</td>
<td>1.30</td>
<td>2447 ± 149</td>
</tr>
<tr>
<td>PG1-acid</td>
<td>444,000</td>
<td>25</td>
<td>40</td>
<td>1.22</td>
<td>79 ± 13</td>
</tr>
<tr>
<td>PG1-II*</td>
<td>444,000</td>
<td>25</td>
<td>25</td>
<td>1.16</td>
<td>495 ± 40</td>
</tr>
<tr>
<td>PG1-III*</td>
<td>444,000</td>
<td>25</td>
<td>5</td>
<td>0.23</td>
<td>99 ± 15</td>
</tr>
</tbody>
</table>

*PG1-II and PG1-III are derived from differential oxidation of the same batch of PG1-diol

Synthesis of PG2.

A lower aldehyde density was desired for the second generation of ester-linked aldehyde polymer and therefore fewer ester-linked carboxylic acid groups were required. A different strategy to incorporate ester-linked aldehydes onto the polymer support was followed (Figure 2.1C). We used an initial NHS activation of carboxylic acid followed by coupling 1-aminoo-3,3-diethoxypropane to generate an aldehyde precursor polymer, PG2-acetal. A ¹H NMR spectrum of the NHS intermediate has a signal at δ = 2.80 characteristic of the NHS moiety attached to the polymer via an ester linkage. The precursor acetal-protected aldehyde polymer (PG2-acetal) exhibits a triplet ¹H NMR signal in the region δ = 1.1-1.3 corresponding to the ethoxy groups. The desired aldehyde polymer (PG2, Figure 1C) was obtained by deprotection of PG2-acetal by trifluoroacetic acid. The aldehyde titration showed a content of 99 aldehyde groups per molecule.
**Synthesis of PG3.**

This ester-linked aldehyde polymer (Figure 2.2A) is nearly identical to PG2, with the exception that the protected aldehyde acetal functionality was attached via a second ester group, not an amide. $^1$H NMR chemical shifts are consistent with the structure. Deprotection to yield the aldehyde groups occurred by treatment with trifluoracetic acid, and aldehyde titration showed a content of 279 aldehyde groups per polymer.

**Synthesis of PG4.**

This ester-linked NHS polymer (Figure 2.2B) was produced to examine an alternative peptide-binding strategy in the context of polyglycerol. An NHS ester will react with only a single primary amine, eliminating the possibility of a dialkylation reaction (see section: ‘Strategies to reduce dialkylation’). The observed $^1$H NMR chemical shifts support the structure.

**Synthesis of PG5.**

This ester-linked aldehyde polymer (Figure 2.2C) was produced by a unique synthetic pathway, resulting in a unique linker moiety attaching the protected aldehyde to the polymer backbone. Deprotection of the aldehyde was afforded by treatment with trifluoroacetic acid. This polymer was designed to aid in understanding the dehydration of hydrolytically released peptides observed with the other polymers that share a common linker structure. The $^1$H NMR chemical shifts are consistent with the structure of PG5.
**PG1 test binding and release of Peptide B.**

The peptide binding and release properties of PG1 were evaluated by incubating the polymer with N-terminal primary-amine containing peptide B and a N-terminal amine-blocked peptide D (see Table 2.3). Following the coupling and separation, the polymer was hydrolyzed (Figure 2.4A) to release bound peptides and separated from the solution using a molecular weight cut-off membrane. A number of signals were observed in the MALDI-TOF mass spectrum of the hydrolysate (Figure 2.4B). A signal at 1302 m/z, corresponds to the anticipated mass shift of 143 Da for the linker modification (monoalkylation) of the peptide N-terminus (Figure 2.4A). Adjacent to this peak is another signal at 1284 m/z, representing a loss of 18 Da (dehydration) from the monoalkylated peak. A dialkylation of the peptide B N-terminus results in a mass shift of 286 Da and the expected signal would be at 1445 m/z. Indeed there is an additional cluster of signals for peptide B present at 1445, 1427 and 1409 m/z, corresponding to the dialkylated product and losses of 18 and 36 Da, respectively. Figure 2.4C illustrates the dialkylated peptide structures. A peak for blocked peptide D is present at 1523 m/z. Its unmodified mass indicates that this blocked peptide did not react with the polymer-linked aldehyde groups but has been carried over through the separation and washing steps that sought to eliminate it. The presence of multiple peaks for the same peptide was undesirable from the standpoint of peptide identification. In the case of peptide B above, the peptide recovered from the polymer was present at five m/z values (Table 2.4.)
Figure 2.4. PG1 test binding and release of peptide B

A. Peptide binds to the polymer through a reductive alkylation reaction between the peptide N-terminal and the polymer aldehyde. Treatment with 100 mM NH₄OH, pH 10, hydrolyzes the ester linker between the peptide and polymer, releasing the peptide. The peptide mass has increased due to the covalent attachment of the polymer linker group, forming species B1. Once hydrolyzed, this linker group can cyclize and dehydrate, forming species B1Δ. ('1' signifies monoalkylation; Δ signifies a loss of one H₂O molecule)

B. The structures of dialkyated peptide species with their corresponding mass shifts are shown. With two polymer linked aldehyde groups reacted with a single peptide N-terminal, two separate dehydration events can occur to form the three structures B2, B2Δ, and B2ΔΔ.

C. Unblocked peptide B (1 nmole) and blocked peptide D (1 nmole) were incubated with PG1-I (20 μg) overnight. Following separation and washing of the PG1-peptide complex, it was subjected to hydrolysis to release bound peptides, 30-kDa MWCO filtering to remove the polymer, and OMIX C18 desalting prior to MALDI-TOF MS analysis. Signals at 1302 and 1284 m/z are monoalkylated species B1 and B1Δ. Dialkylated peptide B species are present at higher abundance relative to the monoalkylated peptide B. Signals at 1445, 1427 and 1409 m/z represent B2, B2Δ and B2ΔΔ, respectively. Peptide D lacks a primary amine and cannot react with the polymer aldehydes. Therefore peptide D has not undergone the characteristic mass shift due to modification by the polymer linker and has been carried over into the hydrolysate sample through the washing steps.
Table 2.4. Peptide B species in hydrolysate.

<table>
<thead>
<tr>
<th>m/z</th>
<th>Sequence (identity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1284.5</td>
<td>INSTTCCYR, monoalkylated, -H₂O</td>
</tr>
<tr>
<td>1302.5</td>
<td>INSTTCCYR, monoalkylated</td>
</tr>
<tr>
<td>1409.6</td>
<td>INSTTCCYR, dialkylated, -2 H₂O</td>
</tr>
<tr>
<td>1427.6</td>
<td>INSTTCCYR, dialkylated, -H₂O</td>
</tr>
<tr>
<td>1445.6</td>
<td>INSTTCCYR, dialkylated</td>
</tr>
</tbody>
</table>

All signals represent the same amino acid sequence information. As illustrated in Table 2.4 for Peptide B, rather than one signal to select for sequencing by MS/MS, the peptide’s signal intensity has been diluted across five peaks. High signal intensity for MS/MS is important for confident identification of a peptide in a bioinformatics analysis. A peptide recovered in low amounts as five separate m/z values has a lower likelihood of being correctly sequenced due to this reduction of intensity. Therefore, further development of the polymer and other strategies were employed to eliminate the redundant peptide signals. As the proposed dehydration reaction was an inherent property of the polymer’s linker moiety, the dialkylation issue was given the foremost concern. Two strategies were designed to prevent this unwanted side-reaction.

**Strategies for reducing dialkylation: i) Effect of aldehyde density.**

The close proximity of amine-reactive aldehyde groups on the polymer surface was thought to play a role in promoting dialkylation. In separate reactions, three unblocked peptides (A, B and C) and one blocked peptide (D) were incubated with two polymers, PG1-II and PG1-III, having 99 and 495 aldehyde groups per molecule, respectively. MALDI-TOF mass spectra (Figure 2.5A) of the hydrolysate for PG1-III exhibit significant
Effect of aldehyde density on dialkylation

Unblocked peptides A, B and C (1 nmole each) and blocked peptide D (1 nmole) were incubated with either PG1-II (495 aldehydes/molecule) or PG1-III (99 aldehydes/molecule). PG1-II and PG1-III were derived from the same batch of precursor PG1-diol, and the different aldehyde densities achieved by variable oxidation with HIO₄. The table lists all of the peptide m/z values and their alkylation/hydration state. The elevated number of binding sites in PG1-II correlates to an increased abundance of the dialkylated peptides compared to PG1-III.

Effect of reaction time on dialkylation

A. Mechanism of the reductive alkylation reaction between a 1º amine and an aldehyde illustrating the formation of the Schiff base (imine) intermediate. NaCNBH₃ is the reducing agent responsible for driving the equilibrium to the right.

B. MALDI-TOF mass spectra of the hydrolysate of two separate reactions. Unblocked peptides A, B and C (1 nmole each) and blocked peptide D (1 nmole) were incubated with PG1-IV (75 aldehydes/molecule) overnight. In the reaction represented by the upper spectra, NaCNBH₃ was present at 20 mM for the duration of the incubation, 16 h. The lower panel shows the results for the shorter exposure to NaCNBH₃, 30 min after 15.5 h of equilibration between peptide amines and polymer aldehydes. The shorter reaction time correlates to a lower relative abundance of the dialkylated peptides.
levels of the dialkylated peptide products. Indeed, for peptides B and C the dialkylated product represents the major species in solution. PG1-II has 20% of the surface aldehyde-density as PG1-III, and in the MALDI-TOF mass spectrum of the hydrolysate the dialkylated peptide products are dramatically reduced in relative abundance to their monoalkylated counterparts. For all peptide species, both mono and dialkylated versions exhibit high relative levels of the characteristic dehydration peaks at -18 and -36 Da. All signals are tabulated and identified in Figure 2.5B.

**Strategies for reducing dialkylation: ii) Effect of shorter reaction time.**

The reductive alkylation reaction responsible for peptides binding covalently to the polymer aldehydes proceeds via an intermediate Schiff base, a structure having a double-bond between the peptide amine and the aldehyde carbon (Figure 2.6A). Subsequent reduction of the double bond to the single bond by NaCNBH$_3$ is an irreversible process that drives the reaction. Equilibrating the peptides with the polymer aldehydes in the absence of NaCNBH$_3$ means that this intermediate, a monoalkylated peptide-polymer conjugate, is as far as the reaction can advance. A short exposure to the reducing agent following equilibration completes the irreversible phase of the reaction and the limited time window would inhibit the desired secondary amine of the monoalkylated peptide from reacting further at an adjacent aldehyde site to become dialkylated.

For this study, PG1-IV (204 aldehydes/molecule) was used. Examining the MS spectra (Figure 2.6B) for the normal 16 h reaction period (NaCNBH$_3$ throughout present at 20 mM), one sees major dialkylation of peptide B (approximately 1:1 with monoalkylated
peptide B) and lower but still significant dialkylation of peptides A and C. The lower panel of Figure 2.6B represents a reaction in which the reducing agent was introduced only for the final 30 min, after the peptides and aldehydes had equilibrated to form Schiff base intermediates for the previous 15.5 h. For peptides A and C, the dialkylated species have been reduced to the level of noise, while for peptide B the dialkylated species that was so significant in the normal reaction has been dramatically reduced to less than 10% of its monoalkylated counterpart. For all of the polymer-conjugated peptide species MALDI-TOF spectra showed characteristic dehydration peaks at -18 and -36 Da. Both strategies for reducing dialkylation showed promise in limiting the generation of this undesired species.

Effect of PG1:peptide ratio.

To further optimize the peptide binding to the polymer support, the effect of different ratios of polymer to peptide in the binding reaction was studied. PG1-V (115 aldehydes/molecule) was used as this level of aldehyde content resulted in lower dialkylation of peptides. The mass spectra (Figure 2.7) illustrate that increasing the amount of PG1-V used resulted in differential recovery of primary amine peptides A, B and C. At the 50 µg level, monoalkylated peptide B is the dominant product, while peptides A and C are at a level below 10% relative to peptide B. As the amount of PG1 incubated with the peptides increased, so too did the levels of peptides A and C in the hydrolysate. However, the increasing concentration of aldehyde in the reaction did not lead to increased dialkylation of peptides, indicating that dialkylation is a result of intramolecular reactions. Peptide D is present in all spectra, indicating carry-over of this unbound peptide throughout the washing of the polymer.
Figure 2.7. Effect of increasing PG1:peptide ratio

Unblocked peptides A, B and C (1 nmole each) and blocked peptide D (1 nmole) were incubated with 50, 100, 200, 400 and 1000 µg of PG1-V (115 aldehydes/molecule). At 50 µg, monoalkylated peptide B species (B1, B1Δ) dominate the spectrum. As the amount of PG1-V in the reaction increases, so too do the relative abundances of the monoalkylated peptides A (A1, A1Δ) and C (C1, C1Δ) compared to peptide B (B1, B1Δ).
**Measurement of affinity and binding capacity of polymer-supported aldehydes or NHS esters.**

Four polymers were evaluated with respect to the activity of their aldehydes (HPG-CHO, PG1-VI, PG2 and PG3) or NHS esters (PG4) towards the primary amine of standard peptide A. Figure 2.8A plots the relationship between the number of nmol of peptide A bound to an increasing amount of polymer added to solution, with the amount of peptide A added fixed at 9.5 nmol. As the polymers used have a range of aldehydes per molecule, this was normalized to the number of nmol of aldehyde present per reaction. HPG-CHO, the non-cleavable aldehyde polymer, shows a high reactivity as indicated by the sharp increase in peptide binding with the increase in aldehyde. The HPG-CHO polymer was saturated at about 8 nmol of peptide A. Similar behaviour is observed for PG2, reaching a near-plateau at approximately 8.3 nmol of peptide bound and 334 nmol of aldehyde. The data points for PG1 have a much gentler slope and the relationship between nmol of peptide bound and nmol of aldehyde in solution is more linear rather than the hyperbolic shape observed for HPG-CHO and PG2. PG3 shows an even more linear relationship, and neither PG1 nor PG3 reach a saturation plateau at the maximum nmol of aldehydes, the solubility limit in this assay. HPG-NHS does not have an aldehyde but the reactivity of its NHS-ester primary amine binding site was studied in an analogous manner. It reaches a saturation of 4.5 nmol of peptide bound at 333 nmol aldehyde added, though there are not enough points to fully establish a plateau. Figure 2.8B compares the structures of the polymer linker (R) groups attached to the aldehyde and gives a binding capacity for each at both 150 and 300 nmol of aldehyde. The slope of the curves between 0 and 150 nmol of aldehyde were calculated as a means for comparing the binding affinity of peptide A towards the peptide reactive groups. From this data, HPG-CHO gave the highest binding reactivity.
Figure 2.8. Reactivity of peptide-binding polymers and binding capacity

A. 10 µg (9.5 nmoles) of peptide A was incubated with increasing amounts of the peptide binding polymers overnight: non-cleavable aldehyde polymer HPG-CHO, three cleavable aldehyde polymers PG1, PG2 and PG3, and the cleavable NHS ester polymer PG4. Plotted is the number of nanomoles of peptide A bound versus the number of binding sites (aldehyde or NHS ester) present.

B. Table contains the names and structures of the peptide-binding polymers. Reactivity towards peptide amines has been determined by the initial slope of the curves between 0 and 150 nmoles of binding site. Peptide A binding capacities are given at two different concentrations of peptide binding site, 150 and 300 nmoles of aldehyde or NHS ester.
and, among the cleavable polymers, PG2 had highest reactivity towards primary amine peptides.

**Dehydration mechanism.**

In the next step we investigated the possible dehydration mechanism observed during the MALDI-TOF analysis. The characteristic loss of 18 Da from monoalkylated peptides and 18 and 36 Da from dialkylated peptides illustrated in mass spectra of hydrolytically released peptides from PG1 and PG2 complexes corresponds to the loss of one or two molecules of H$_2$O (18 Da). The proposed mechanism for this dehydration is a cyclization of the modified N-terminus shown in Figure 2.9A. PG1, PG2, PG3 and HPG-NHS have in common the succinic anhydride-derived four-carbon unit as part of the linker moiety joining the peptide-binding site (aldehyde or NHS ester) to the polymer backbone via an ester group. When a conjugated peptide A is released by hydrolysis, the mass spectra of this hydrolysate for all of these polymers exhibits a signal at 18 Da less than the expected mass shift due to the linker attachment. PG3 possesses two base-hydrolyzable ester groups in the linker, (i) and (ii) (Figure 2.9D). Hydrolysis at group (i) generates a peptide species with this four-carbon unit with an adjacent dehydration peak at 18 Da less. Hydrolysis at (ii) removes this four-carbon unit from the peptide and no dehydration is observed, suggesting that presence of the carboxylic acid in the linker group is playing a role in the dehydration mechanism. In the case of PG3, it is predicted that dehydration would form a nine-membered ring between the amide and the terminal carboxylic acid. As this structure is less favored, the dehydration is less extensive and only accounts for about 10% of the monoalkylated peptide A signal.

To probe this dehydration mechanism we generated a fifth cleavable polymer, PG5 (Figure 2.9F), by a unique synthetic route (Figure 2.3C). PG5 lacks the structural
Figure 2.9. Dehydration of peptide N-terminus with polymer linker covalently attached

A. Proposed mechanism for the dehydration of the linker-modified peptide N-terminus by cyclization. B to F: Structures of the cleavable peptide-binding polymers with a peptide bound. PG1, PG2, PG3 and PG5 are aldehyde polymers while PG4 has an NHS-ester binding site. Note the common structural feature common to PG1, PG2, PG3 and PG4, the four carbon chain derived from reaction of polyglycerol with succinic anhydride. Arrows indicate the hydrolyzable ester groups. Both PG3 and PG5 possess two base-cleavable esters.

B. Structure of PG1 with peptide bound. MALDI-TOF mass spectra showing the cleaved peptide at the expected mass shift due to monoalkylation (PG1 linker modification) and another peak at 18 Da less (dehydration).

C. Second generation aldehyde polymer PG2. Spectra shows the monoalkylated peptide at the expected mass shift and its dehydrated version.

D. Cleavable polymer PG3. This polymer has two esters in its linker region and two possible hydrolysis products. Both products are observed in the spectra, but only hydrolysis at the left ester yields a peptide undergoing dehydration, likely forming a nine-membered ring between the terminal carbonyl and the amide (*).

E. Cleavable NHS polymer PG4 with peptide bound. Spectra shows the monoalkylated peptide at the expected mass shift and its dehydrated version.

F. A cleavable aldehyde polymer PG5 with peptide bound. This polymer's linker possesses two esters and two possible hydrolysis products. Both products are observed in the spectra, and neither undergoes a dehydration.
A. 

B. 

$\Delta m = 143.1 \text{ Da}$

PG1

C. 

$\Delta m = 157.2 \text{ Da}$

PG2

D. 

$\Delta m = 158.1 \text{ Da}$  

$\Delta m = 58.1 \text{ Da}$

PG3

E. 

$\Delta m = 100.1 \text{ Da}$

PG4

F. 

$\Delta m = 223.2 \text{ Da}$

$\Delta m = 58.1 \text{ Da}$

PG5

Mass-to-charge ratio (m/z)

Relative intensity %

1000 1050 1100 1150 1200 1250

58
feature common to the other polymers. PG5 possesses two base-labile ester groups in the linker. Peptide species corresponding to hydrolysis at both ester groups are observed, but in neither case are additional dehydration peaks associated, supporting the proposed mechanism.

Assessment of non-specific binding to polymers

Another important aspect of developing a soluble polymer is the low non-specific peptide binding, and this was studied using amine-blocked BSA tryptic peptides. The BSA tryptic peptides with dimethylated tertiary amines display no non-specific binding to the polymers HPG-CHO, PG1 and PG2. In the pooled 30-kDa MWCO flow-through, heavy-light labeled peptide pairs show an equal intensity in the MALDI-TOF MS spectra, the light peptides having previously been incubated in buffer with 1000 µg of polymer while the heavy peptides were incubated without polymer. Peptide pairs are separated by 6 Da or 12 Da in the case of peptides with a lysine residue. For HPG-CHO the peptide pairs were of equal (1:1) intensity in the flow-through (Figure 2.10A). For PG1, the peptide pairs (Figure 2.10B) are mixed between being of equal intensity and pairs where the light peptide has a higher abundance. The greater abundance of the light peptide relative to the heavy is more pronounced in the first wash fraction (100 µL 50 mM HEPES). The opposite trend is observed when heavy peptides are incubated with polymer (Figure 2.10B, bottom). Virtually all peptide signals have disappeared by the fourth wash (100 µL 2 M NaCl/50 mM HEPES), and hydrolysis of PG1 recovered from the incubation yields no further peptide signals. In the initial flow-through for PG2 (Figure 2.10C), the light peptide is more abundant in each peptide pair, on average about 30% greater, with the exception of pair C where the light peptide intensity fully doubles that of the heavy. As with PG1 this trend is again seen in the first wash fraction;
Light isotope formaldehyde labeled (dimethylated) BSA tryptic peptides (20 µg) were incubated overnight with one of the three aldehyde polymers (1000 µg) while heavy-isotope formaldehyde labeled BSA tryptic peptides were incubated with no polymer. Flow-through and wash fractions were pooled and analyzed by MALDI-TOF MS and the relative abundance of the heavy/light peptide pairs examined.

A. Non-cleavable polymer HPG-CHO. Peptide intensities are approximately 1:1 in the flow-through and >1 in Wash 1.

B. First-generation cleavable polymer PG1. Light-heavy peptide ratios are >1 in both the flow-through and in Wash 1.

C. Second-generation cleavable polymer PG2. Light-heavy peptide ratios are >1 in both the flow-through and in Wash 1.
the light-heavy peptide ratio is significantly greater than 1. Peptide signals have been reduced to baseline by the fourth wash, and hydrolysis of PG2 yielded no peptide signal. No significant differences in nonspecific binding were observed for the three tested polymers despite their different structures and degrees of modification.

**Purification of a single primary-amine peptide from a complex mixture**

The goal of this experiment was to demonstrate specific recovery of a primary amine containing peptide from a mixture of blocked BSA tryptic peptides. A modified binding protocol was used that incorporated an ACN precipitation of the PG2-peptide complex and a more vigorous washing regimen using 2 M NaCl/50 mM HEPES as well as 20% ACN/50 mM HEPES to wash the polymer to remove unbound peptides. Peptide A (m/z = 1058) is seen at high intensity in the initial experimental conditions, prior to addition of the polymer. Following incubation to bind the peptides and ACN precipitation, the supernatant was evaporated and reconstituted and subjected to MALDI-TOF MS analysis (Figure 2.11A). Peptide A has been nearly eliminated from the spectrum while the BSA tryptic peptides remain. As is evident from the spectrum of Wash 1 and Wash 5 these BSA peptide signals have disappeared. The hydrolysis of the washed PG2-peptide complex yields the expected signals for monoalkylated peptide A (Δm/z = 157; 1058 + 157 = 1215) and a prominent dehydration peak at -18 Da (m/z = 1197). Dialkylated peptide A is also detected, with peaks at 1354 and 1336 m/z. representing the loss of one and two molecules of H₂O, respectively. Notably, the dehydrated, monoalkylated peptide A is evident at low levels in the flow-through and washes, indicating premature hydrolysis of the polymer linker’s ester-moiety.
Figure 2.11. Purification of peptide A from a complex mixture

A. Unblocked peptide A (2 nmoles) and light-formaldehyde blocked BSA tryptic digest (20 μg) were incubated overnight with second-generation cleavable aldehyde polymer PG2 (1000 μg). MALDI-TOF mass spectra shows the initial conditions with no polymer added.

B. Following incubation, the 30-kDa MWCO flow-through contains only a near-baseline signal for peptide A, indicating near-complete depletion from the solution. Unbound blocked BSA tryptic peptides are also observed.

C, D. Wash 1 and 5 demonstrate removal of the BSA tryptic peptides from the PG2-peptide A complex.

E. Following washing of the PG2-peptide A conjugate and hydrolysis of the linker’s ester group, the MALDI-TOF spectrum of the hydrolysate shows recovery of the released peptide, present largely in its monoalkylated and dehydrated form.
**Proteomic identification of MMP2 cleavage sites**

This experiment was designed to demonstrate the utility of the cleavable aldehyde polymer, PG2, to capture and enrich for the cleavage products of MMP2 digestion of a peptide library. As the cleaved peptides would all possess a primary amine, they should covalently bind to the polymer. Figure 2.12A is a Pep3D map generated using the TPP proteomics data analysis engine, with the m/z values observed during the MS scan that were selected for MS/MS analysis plotted against the retention time. In the 400-700 m/z range, hundreds of signals are distributed across the span of the gradient, each one potentially a peptide cleavage product captured and released by PG2. However, database searching of the corresponding MS/MS spectra failed to return any confident peptide matches. Inspection of the MS/MS spectra revealed a common problem, Figure 2.12B illustrates a pattern seen in all acquired MS/MS spectra: A lack of ion signals in the range above approximately 500 m/z. A quadruply-charged peptide of 475.3 m/z was fragmented by collision-induced dissociation (CID). The peak at 140.1 m/z corresponds to a fragmentation of an N—C bond in the linker moiety (see inset of spectra). The peak at 225.2 may correspond to an a₁ fragment ion if the R group was a leucine or isoleucine side-chain. This is in agreement with a previous MMP2 PICS study that showed a strong preference for Leu or Ile in the P₁⁺ position. The fragmentation events dominate the spectra, severely limiting the appearance of ions at higher m/z values and limiting the possibility of a confident peptide identification due to lack of amino acid sequence information.
Figure 2.12. PG2 pullout of MMP2 cleavage products from a PICS screen

A. Pep3D representation of the LC-MS analysis, plotting the m/z values against the retention time. The intense signals are spread over the retention time, indicating a good chromatographic separation. Hundreds of peptides were selected during the MS scan for MS/MS analysis.

B. Sample MS/MS spectra illustrating the preferential fragmentation of the N—C indicated in the structure. Most fragments appear in the low (<500 Da) mass range, hindering a confident peptide identification due to insufficient amino acid sequence information.
2.4. DISCUSSION

We designed and synthesized a novel polymer useful for the enrichment of primary-amine containing peptides from a complex mixture of peptides. A series of water-soluble and base-cleavable ester-linked aldehyde polymers were prepared to covalently bind peptides with a primary amino group via an overnight reductive alkylation reaction. As amine-reactive groups, aldehydes offer longterm stability compared to NHS esters. Although NHS esters are common in amine-labeling approaches, they are less stable and deteriorate in aqueous systems due to hydrolysis. The high molecular weight of the polymer facilitates separation of the polymer with bound peptides from a complex peptide mixture using either 30-kDa molecular-weight cutoff filtration or precipitation in acetonitrile. To release the bound peptides after washing, base-catalyzed hydrolysis of the polymer-peptide conjugate cleaves the ester group, releasing the bound peptide into solution to be analyzed by MS. Attachment of the polymer linker moiety to the N-termini results in a characteristic mass shift that can distinguish the covalently bound peptides, which is useful to categorize bona fide bound peptides from any minor peptide contaminants present due to carry-over. Future versions of the polymer will incorporate isotopic labeling in the linker moiety for relative quantitation techniques. The predicted mass shift is also useful in polymer characterization to confirm the presence of the intended chemical functionality. Hence, this cleavable polymer is a useful reagent in proteomic workflows such as those aiming to identify the peptidic cleavage products generated by proteolytic digestion.

The polymer development process revealed an issue with significant dialkylation of peptides due to bidentate binding, resulting in multiple MS signals for the same peptide. An observed dehydration of hydrolytically released peptides was correlated to the linker
region by observing similar behaviour in a panel of cleavable aldehyde polymers. The same group of polymers was also assessed with respect to their reactivity toward a standard peptide and it was shown that the linker structure impacts this reactivity. The different aldehyde polymers unexpectedly show significant differences in reactivity towards the same standard peptide. PG2, with two methylene groups between the aldehyde and the linker amide, had the highest reactivity of the ester-linked aldehyde polymers and was selected as the candidate for PICS trials. PG4, a hyper-branched polyglycerol using NHS-ester chemistry to bind peptides was also tested. The polymers were shown to have no non-specific binding effects, as supported by incubation of HPG-CHO, PG1 and PG2 with and without isotopically-labeled, amine-blocked BSA tryptic peptides (Figure 2.10 A,B & C). Additionally, the recovery of Peptide A alone from a complex background of blocked BSA peptides (Figure 2.11) demonstrates the selectivity of the procedure. The undesirable carry-over of unbound peptides was reduced through a rigorous washing protocol incorporating both high-salt and ACN-containing buffers to disrupt any electrostatic and hydrophobic non-specific interactions.

Insoluble polymers are commonly employed in solid-phase chemistry procedures such as peptide synthesis and the preparation of large combinatorial libraries\textsuperscript{30,31}, with polystyrene beads forming the solid support. Highly cross-linked polysaccharide chains such as Sepharose form the resins and beads to which other functionalities have been covalently attached for the purposes of affinity chromatography. For example, antibodies attached to such solid supports are used in immunoprecipitation of proteins, and immobilized metal cations (e.g. Ni\textsuperscript{2+}) are routinely used to capture recombinant proteins bearing poly-histidine tags. At the peptide level, the conjugation of peptide N-termini\textsuperscript{13,26} or cysteine sulfhydryl groups\textsuperscript{14,32,33} to biotin followed by streptavidin affinity
chromatography has proven an effective way to enrich subsets of a population. In these techniques, the insoluble nature of the polymeric support facilitates purification, typically by packing the beads into a column and flowing samples, wash solvent and elution buffers over this stationary phase. While these methods have been successful, there are several drawbacks of solid-phase capture of molecules. The heterogeneous conditions can be subject to non-linear reaction kinetics\textsuperscript{15,16}, largely due to unequal access of the target molecules to the resin-linked chemical functionality. Additionally, problems with solvation and non-specific binding to the macromolecular structures can result in significant losses of peptides, particularly of non-abundant species\textsuperscript{17}.

To resolve these problems, soluble polymers afford a desirable remedy. A soluble polymer restores homogeneous reaction conditions, enhancing the reaction efficiency compared to biphasic systems while still taking advantage of the large size of the polymer to enable separation from a complex mixture. Dendrimers are a class of branched polymers with spherical structures and highly functionalized surfaces\textsuperscript{34}, ideal for attachment of organic moieties. Polyamidoamine (PAMAM) is an amine-functionalized dendrimer that has successfully been used to capture a variety of peptides from solution by covalent attachment. Phosphopeptides have been captured by the selective reaction of phosphate ester groups with the dendrimer’s amines\textsuperscript{22}. Compared to an amino-functionalized solid-phase support, capture of phosphopeptides on the soluble dendrimer exhibited improved kinetics and yield\textsuperscript{22}. Additionally, a bromoaceto-functionalized version of PAMAM has been used to bind Cys-containing peptides\textsuperscript{15}. In each case, the peptide-dendrimer conjugates are readily separated from a mixture and the peptides released from the polymer under acidic conditions, with size-filtration methods used to separate high and low molecular weight species at each step.
Dendrimers have interesting properties because of their extensive branching, uniform shape and size, and large number of surface-modifiable groups. However, their synthesis proceeds in a stepwise fashion and is tedious and very time-consuming (up to 30 days), requiring expertise in polymer synthesis methods. In contrast, hyper-branched polymers possess many of the same benefits as dendrimers and can be prepared more readily, often in a single overnight reaction in high yield and without specialized knowledge of polymer chemistry\textsuperscript{35}. Polyglycerol is a hyper-branched polymer that is water-soluble, stable, and richly functionalized with reactive hydroxyl groups. It has a more flexible structure than the rigid architecture of many dendrimers and this further enhances the reaction kinetics in solution. The surface hydroxyls are easily converted into other chemical functionalities, such as amine-reactive aldehyde groups.

The general peptide capture-and-release protocol designed for the ester-linked aldehyde polymers has proven effective. A pH of 6.5 was selected to promote the reaction through protonation of the aldehyde carbonyl oxygen. In Figure 2.11 we see that the 16 h binding has fully depleted Peptide A from the complex mixture, and post-washing hydrolysis yields only the desired hydrated and dehydrated signals for PG2 linker-modified Peptide A. The use of higher pH during the hydrolysis led to some peptide degradation, but no such degradation was observed for hydrolysis at pH 10. With respect to peptide binding and release behaviour, several unexpected hurdles presented during the development. The first was the MALDI-TOF MS observation of extensive peptide dialkylation in the hydrolysate (Figure 2.4C). The original polymer design called for complete polymer branch modification to maximize the number of peptide binding sites (Figure 2.1), resulting in close proximity of surface aldehyde groups. This promoted the second alkylation reaction as the singly bound peptide with a
reactive secondary amine nucleophile was constrained within close proximity to an adjacent aldehyde neighbour. Reduction of the number of aldehydes per molecule to increase the spacing proved to be a logical and effective solution to minimizing this undesired intramolecular reaction (Figure 2.5). Limiting reaction time was also effective but resulted in incomplete binding of the available peptides (Figure 2.6).

In addition to the dialkylation reaction, the dehydration of the released peptides further complicated MS spectra. As with dialkylation, the same peptide sequence information was split over multiple peaks rather than concentrating the ion intensity on only one signal (Figure 2.4C). For PG2, the cleavable polymer with the highest reactivity, the dehydration was very pronounced and indeed represented the abundant species observed in the MS spectrum. The suggested cyclization mechanism (Figure 2.9A), supported by the investigated panel of cleavable polymers, would not have been predicted without experimentation. In general, the abundance of the dehydrated peptide signal relative to the hydrated peptide signal was dependent on linker length. The linker of PG2 is one methylene group (CH$_2$) longer than for PG1, and the dehydrated peak dominates the MALDI-TOF spectrum (Figure 2.9B & C). The extra length will diminish steric hindrance effects and favour the cyclization into a five-membered ring. By contrast, PG4 has the shortest linker of all polymers and the spectrum is dominated by hydrated peptide due to the added steric hindrance of cyclization. This dehydration mechanism is analogous to the cyclization occurring in the preparation of maleimide, a similar 5-membered dicarbonyl amide ring$^{36}$. This feature of the polymer remains to be optimized, and alteration of the linker moiety is a critical factor.
Our group has previously demonstrated the utility of aldehyde-functionalized polyglycerol to simplify a complex peptide mixture in proteomics by removal of internal tryptic peptides\textsuperscript{23}. In this workflow, primary amines present at natural and protease-generated N-terminals are first blocked by dimethylation and the sample subjected to trypsin digestion. To enrich for the subset of amine-blocked N-terminal peptides by negative selection, the internal tryptic peptides are irreversibly bound to an aldehyde polymer and removed. However, in other approaches it is desirable to recover and analyze these bound peptides. For example, in PICS\textsuperscript{26}, a peptide library prepared by tryptic digestion has all primary amines blocked by dimethylation. Treatment with the test protease generates two peptides, representing amino acid sequences C-terminal and N-terminal to the cleaved bond. As reported, the N-terminal peptides are captured by conjugation with a cleavable biotin reagent followed by streptavidin affinity pull-out. A cleavable aldehyde polymer would offer the advantage of increased capacity, a simplified workflow, greater recovery and reduced cost. Here we have produced a working prototype polymer that can now be assessed in proteomic workflows designed to capture and release primary amine containing peptides with high efficacy, high recovery and low non-specific binding.
REFERENCES


A water-soluble polymer capable of binding primary-amine containing peptides has been developed for sample preparation in proteomics applications. In the course of this research, a hyper-branched polyglycidol polymer has been modified to generate amine-reactive aldehyde groups attached to the backbone via a cleavable ester linkage. Primary-amine containing peptides were selectively targeted, separated from a complex mixture, and subsequently released, enriching for these analytes and enhancing their analysis by MS. This was a method development project entailing the design and synthesis of the cleavable polymer as well as the implementation of it to bind primary amine containing peptides for the purpose of targeted selection from a complex peptide mixture. In shotgun proteomics, intact proteins are digested into smaller peptides with a protease of canonical specificity, often trypsin. Peptides are easier to analyze because of their improved solubility and ionization properties, and MS instrumentation can determine their masses with high accuracy and precision\(^1\). The generation of multiple peptides from every protein in a proteome creates an astounding level of sample complexity, and contributes to a problem termed MS undersampling. The MS is overwhelmed by the sheer number of analytes and only a small fraction of peptides, 10-20\%, will be assigned, leading to limited coverage as few proteins will be identified by two or more unique peptides\(^2\). Compounding this is the ion-suppression effects observed when multiple peptides simultaneously enter the MS ion source. Separation techniques designed to handle such complex peptide samples are an essential component of sample preparation.
Reversed-phase chromatography is near ubiquitous and often coupled directly to the MS to separate peptides based on their hydrophobic properties. Additional dimensions of sample preparation based on other physicochemical properties are also common. Proteins may be fractionated according to their molecular weight or isoelectric point prior to tryptic digestion, and tryptic peptides may be fractionated according to their electrostatic properties before reversed-phase chromatography. These multidimensional sample preparations do improve MS analysis and increase the number of peptides successfully identified, but the exponential increase in the number of total fractions from each dimension requires much more instrument time.

Targeted selection methods offer an additional dimension of sample simplification without generating more fractions to analyze by focusing on only those analytes that the researcher is interested in. This is particularly true in the study of post-translational modifications where only a small percentage of the total peptide pool possesses the features under investigation. Negative selection is applied in the substrate discovery technique TAILS\textsuperscript{3}. The aldehyde polymer is used to deplete the internal tryptic peptides from a complex peptide mixture, enriching for the natural and neo N-termini that represent the post-translational modification: a protease cleavage site. Positive selection is used in PICS, as the protease-cleaved peptides are specifically labeled with biotin and purified by streptavidin affinity chromatography. In each case, the targeted selection technique used eliminates those background analytes to improve the ionization and MS analysis of the peptides of interest.

The use of polymers as the physical supports in targeted selection techniques is not new. Highly cross-linked polysaccharides such as agarose form the solid support for a
variety of chemistries, including affinity capture of phosphopeptides with antiphosphotyrosine\(^4\) or immobilized metal ions\(^5\), or biotinylated peptides with immobilized streptavidin\(^6,7\). Peptides are also bound covalently, using chemistries specific for N-termini, cysteine\(^8,9\) and tryptophan\(^10\) residues, or for post-translationally added groups like phosphates and carbohydrates\(^11\). The solid support with bound peptides can be isolated by filtration or by using a column format and flowing through sample solution and wash solvents to remove non-specific contaminants. Covalent complexes can be washed with harsher organic and denaturing solvents to make this removal more effective, as the high salt concentration solvents and organic solvents disrupt non-specific electrostatic and hydrophobic interactions. The negative aspects of solid-phase chemistry stem from the heterogeneous nature of the system, manifested by non-linear kinetics caused by unequal distribution of reagents\(^12\), as well as peptide losses\(^13\). This has led to the rise in popularity of soluble polymers as physical supports for targeted selection techniques.

Soluble supports have previously gained popularity in organic combinatorial chemistry, where the homogeneous conditions recovered the kinetics and improved yields compared to solid-phase counterparts. In proteomics, however, the use of soluble polymers as physical supports is still in its infancy. An amine-functionalized dendrimer known as PAMAM has been used as is to bind phosphopeptides\(^14\). It has also been modified by the attachment of organic and inorganic moieties to capture phosphopeptides with immobilized metals\(^15\) or to bind and isotopically label cysteine-containing peptides in quantitative approaches\(^16\). A bifunctional PAMAM permitted liquid-phase binding of analytes followed by conjugation to a solid-support to facilitate separation, washing and elution\(^16\). Dendrimers such as PAMAM are expensive to buy
and very challenging and time-consuming to synthesize, and place restrictions on the upper limits of molecular weight. Soluble hyperbranched polyglycidol described here has similar properties to dendrimers. They are globular and have a high degree of functional groups for the attachment of organic moieties, and high molecular weight hyperbranched polyglycidol can be prepared quickly and cheaply, making them more accessible.

The synthetic component of the research described here involved standard organic chemistry reactions, though the polymeric context presented challenges to conventional characterization such as MS because high molecular-weight, polydisperse polymers do not have a single mass. While $^1$H NMR was employed, the wide range of proton microenvironments in solutions gives broad peaks and complex patterns. While helpful for confirmation of the existence and absence of certain groups and the success of a particular synthetic step, it was difficult to extract useful quantitative information from peak integrations.

The aldehyde-amine reductive alkylation chemistry that binds peptides covalently to the polymer has the potential for non-equivalent recovery of peptides due to amino acid-dependent binding kinetics. Figure 2.7 shows the relative abundance of Peptides A and C increasing with respect to peptide B as the amount of polymer in the reaction increases, suggesting that peptide A is bound preferentially when polymer is limiting. It would be worthwhile to further examine the binding behaviour of a range of synthetic peptides that represent a range of N-terminal residue types to determine if any amino acid-dependent binding behaviours exist.
The non-specific binding of peptides to the polymer would be detrimental to subsequent LC-MS/MS analysis, as unbound non-primary-amine containing peptides would be carried through the washing steps. It is these very peptides that this soluble polymer-based targeted selection approach was designed to eliminate, and their presence complicates the sample and hampers MS analysis. Using isotopically-labeled, amine-blocked BSA tryptic peptides, we have shown that non-specific binding of peptides is not a problem. The experiments revealed that non-specific binding to the regenerated cellulose MWCO membranes used in the separations does occur, and this would contribute to both peptide losses and carry-over of blocked peptides in polymer characterization studies. In the presence of polymer, the MWCO membranes retain more liquid volume after centrifugation, possibly due to blocking of membrane pores and formation of a hydrated polymer gel. This retained volume traps peptides that are later released, as supported by the relative increase in the polymer-incubated light peptide signal in the first wash fraction for HPG-CHO, PG1 and PG2. The rigorous washing protocol used in the non-specific binding assessment, employing both high-salt and acetonitrile-containing buffers, was also used in the targeted selection of Peptide A from a complex background of blocked BSA peptides. As Figure 2.11 illustrates, there was no carry-over of unbound peptides detectable in the hydrolysate, only the expected signals corresponding to recovery of bound Peptide A.

When PG2 was incorporated into the PICS workflow to enrich for peptides cleaved by MMP2, no peptides were identified. The Pep3D map (Figure 2.12A) indicates a large number of peptides observed by the MS. A preferential fragmentation within the linker moiety during CID-induced MS/MS prevented the generation of a complete fragment ion series of the peptide. Rather than the desired random fragmentation of inter-residue
amide bonds, the collision energy was directed to the labile N-terminal linker modification at the N—C bond that marks the interface between the peptide and the polymer. In Figure 2.12B, there are intense fragment ion peaks in the low-mass region of the MS/MS spectrum, while the complete ion series corresponding to fragmentation along the entire peptide backbone is absent. Database searching algorithms rely on matching ion peaks to acquire a confident peptide identification and the absence of the complete ion series prohibits this. Future LC-MS/MS analysis using CID will require the optimization of the collision energy imparted to selected parent ions to yield a uniform fragmentation pattern with the necessary high-mass ions required for peptide identification. An alternative MS/MS fragmentation approach called electron-capture dissociation (ECD) exists and is characterized as being gentler than CID on labile covalent modifications. ECD MS/MS therefore generates a more complete ion series for improved sequencing of peptides. Analysis of PICS samples on a MS equipped with this feature may result in more informative peptide fragmentation and yield confident peptide identifications.

Future development of the cleavable aldehyde polymer must explore how different linker structures may affect both reactivity and released peptide behaviour. As was demonstrated here, subtle variations in linker structure have significant consequences in this area, likely due to the variation in electron-withdrawing effects acting on the aldehyde. A decrease in electron density at the aldehyde position will enhance its electrophilic properties and make it more active towards the primary amine nucleophile. The preparation of more structural permutations may further the understanding of these effects and optimize primary amine reactivity and binding capacity. The incorporation of a disulfide bond in the linker may prove to be beneficial. First, it would offer an alternate
cleavage mechanism to base-catalyzed ester hydrolysis. Reduction with DTT would be both gentler and possibly faster than heating the polymer-peptide conjugate at 65 °C at pH 10. Second, the cleaved peptide N-terminus would have a free sulfhydryl, not a carboxylic acid, and would not be prone to cyclization and dehydration. In addition to making changes to the method of peptide release, the linker also gives an opportunity to use stable isotope labeling of peptides to facilitate a quantitative comparison of peptides from different sample conditions. In the context of PICS\textsuperscript{7} (proteolysis of a peptide library), for example, one could study kinetic differences between two sets of digestion conditions. Even the manner in which the polymer-peptide conjugate is separated from solution warrants further study. While convenient to use, regenerated cellulose MWCO membranes contribute to peptide loss by non-specific binding, as well as restrict washing solvents to those that are membrane-compatible. A bifunctional polymer with both peptide binding sites and a chemical handle for solid-phase pull-out would combine the efficient liquid-phase binding with effective solid-phase washing. Alternatively, the very large difference between the size of the polymer-peptide conjugate and unbound peptides could be exploited using size-exclusion chromatography, or the polymer could be immobilized within a column to create a system where sample, washing and hydrolysis solvents can be flowed through. There are many parameters involved in the optimization of a particular method. The synthesis of isotopically heavy peptide standards would enable a more quantitative study of the peptide binding behaviour and recovery while varying such parameters as the reagent ratios, time, temperature, and pH of binding and hydrolysis.

Hyperbranched polyglycidol has many possible applications as the physical support for targeted selection chemistries. In addition to aldehydes, other organic functional groups
could be attached. Cysteine-containing peptides could be bound through a maleimide or haloaceto moiety or phosphopeptides bound by carbodiimide-mediated formation of phosphoramidate bonds\textsuperscript{19} on an amine-functionalized polyglycidol. The same polymer used in the context of PICS would permit the non-prime side protease cleavage products to be selected for by similar formation of amide bonds. Essentially, immobilization of any amino acid-specific or posttranslational-modification specific tagging chemistry to polyglycidol via a cleavable linker would permit the simultaneous labeling and enrichment of the targeted peptides. NHS ester chemistry, widespread in peptide amine labeling approaches, was investigated here. It is an initially attractive choice as it is quick and the nature of the chemistry prevents dialkylation. However, NHS esters are susceptible to hydrolysis and would have short-term stability when attached to such a hydrophilic polymer as polyglycidol. Here they were prepared fresh immediately before use to circumvent these effects. An isotopically-coded linker could be incorporated into any of these strategies to enable labeling and quantitation of the captured and released target peptides.

A polymeric proteomics reagent need not be restricted to binding peptides. Small-molecule inhibitors could be conjugated to the polymer to facilitate the affinity capture of proteases in a complex mixture of native proteins, the Zn-chelating hydroxamate MMP inhibitor, marimastat, for example, to capture active MMPs. Enzymes themselves immobilized to the polymer could aid in the development of microfluidic reactors. Polymer-bound trypsin enables much higher enzyme concentrations and protease to substrate ratios in solution than can be achieved with free enzyme to significantly enhance digestion while preventing auto-proteolysis and contamination of samples with protease fragments\textsuperscript{20}.
In conclusion this thesis has demonstrated the utility of a polyglycidol-based reagent to enrich primary-amine containing peptides for the purpose of decreasing sample complexity in proteomic workflows. The reductive alkylation reaction is specific and there was no detectable interference from other peptide functional groups, while the soluble nature of the reagent enhances binding kinetics and minimizes sample loss. This work represents the initial steps in developing this novel reagent, and further development warrants additional studies aimed to better characterize the binding of peptides and the MS and MS/MS analysis of the captured species. I believe that with this refinement and optimization this reagent could be a useful tool to other proteomics researchers and could be readily adapted for a multitude of applications. I expect that similar applications of soluble polymers will continue to grow within the field of proteomics due to the noted advantages over solid-phase systems.
REFERENCES


