Chemical Synthesis and Engineering of Cytokines of Potential Therapeutic Interest

Dr Paolo Botti

Dottorato in Scienze Biotecnologiche – XXI ciclo
Indirizzo Biotecnologie Molecolari
Università di Napoli Federico II

Dottorato in Scienze Biotecnologiche – XXI ciclo
Indirizzo Biotecnologie Molecolari
Università di Napoli Federico II

Chemical Synthesis and Engineering of Cytokines of Potential Therapeutic Interest

Dr Paolo Botti

Dottorando:	Dr. Paolo Botti
Relatore:	Prof. Ettore Benedetti
Correlatore:	Dr. Menotti Ruvo
Coordinatore:	Prof. Giovanni Sannia

Ad Elena e a tutta la mia famiglia

INDEX

Riassunto

Sintesi Chimica ed Ingegnerizzazione di Citochine a Potenziale Interesse Terapeutico

La recenti innovazioni nella sintesi chimica dei peptidi e delle tecniche di bioconiugazione hanno rilanciato nei primi anni novanta la sintesi chimica di polipetidi con catene superiori a 100 residui. Nel 1994, il gruppo del Prof. S. Kent ha descritto per la prima volta sulla rivista *Science* la ligazione di due peptidi non protetti in acqua a pH neutro, reazione che ha generato un solo polipeptide avente un legame ammidico nativo al sito di ligazione, tecnica chiamata Native Chemical Ligation (NCL). Native Chemical Ligation ha permesso negli anni seguenti la sintesisi chimica di diverse decine di polypeptidi e proteine di piccola e media taglia. In particolare nel 2003 e' stata presentata sulla rivista *Science* la sintesi chimica totale di SEP (Synthetic Eritropoietin Protein), molecola poi entrata in fase clinica I.
In collaborazione con il Prof. C. Dinarello, Univerista' del Colorado, Denver, si e' deciso di intraprendere la sintesi chimica di due citokine di rilenvante interesse biologico e terapeutico: Interleukina 18 e Interleukina 32.

H-IL-18 consta di 157 residui ed e' una citokina proinfiammatoria che appartiene alla famiglia dei ligandi IL-1, inizialmente clonata come fattore di induzione di IFN-γ, ha un ruolo chiave in diverse patologie infiammatorie, allergiche ed autoimmunitarie. In particolare, Il-18 inducendo la produzione di IFN-γ da parte delle cellule T, e' stata utilizzata con successo nella stimolazione della difesa immunitaria contro tumori in topi e dopo studi preclinici di tossicita' su scimmie e'al momento in Clinical Trials per il trattamento di patologie tumorali. H-IL-18 ricombinante e' sensibile all'ossidazione atmosferica ed e' prona a formare dei prodotti di oligomerizzazione attraverso ponti disolfuro inter-molecolari che rendono la molecola inattiva. Data la rilevanza delle applicazioni mediche di h-IL-18, (in Clinical Trials per il trattamento di patologie tumorali) esiste la necessità di trovare una forma stabile della proteina che possa garantirne un uso clinico per le diverse applicazioni.

L'Interleukina IL-32 e' una nuova Interleukina recentemente scoperta dal gruppo del Prof. C.Dinarello presso l'Universita' del Colorado. Il-32 esiste in 4 differenti varianti. La sintesi dell'H-Il-32 e' divenuta prioritaria nello sviluppo del lavoro di tesi a causa della rilevanza scientifica che ha la proteina nei processi infiammatori e automminitari. La sintesi chimica a questo stadio della ricerca riveste un ruolo vitale nella conferma dei dati scientifici ottenuti precedentemente con materiale ricombinante a causa della difficolta' nell'espressione della proteina, sia dal punto di vista della quantita' del materiale prodotto sia dal punto di vista della purezza in particolare riguardo alla presenza di tossine o di altro materiale biologico potenzialmente bioattivo. Qualora infatti il polipeptide di sintesi riproducesse l'attivita' biologica della proteina ricombinante confermerebbe l'esistenza di nuovi pathways infiammatori ed autoimmunitari. H-Il-32 consiste di 164 residui. In collaborazione con il Prof. Dinarello, abbiamo sintetizzato la porzione N-terminale della proteina γ (gamma) isoform che a priori dovrebbe possedere attivita' biologica analoga a quella umana. La parte N-terminale consiste di 103 residui con 4 cisteine e due ponti disolfuro. In base ai risultati biologici ottenuti, in un secondo momento eventualmente si pensera' alla sintesi della proteina umana nella sua completezza.

Obiettivi.

1) H-IL-18

Sintesi Chimica della proteina wild-type per compararla con quella ricombinante ed eventualmente modificazione della proteina per renderla più stabile e/o per migliorare le sue proprieta' biofarmaceutiche.

YFGK LESKLSVIRN LNDQVLFIDQ GNRPLFEDMT DSDCRDNAPR TIFIISMYKD SQPRGMAVTI SV**KC**EKISTL SCENKIISFK **EM**NPPDNIKD TKSDIIFFQR SVPGHDNKMQ FESSSYEGYF L**AC**EKERDLF KLILKKEDEL GDRSIMFTVQ NED
(I siti di ligazione sono sottolineati in grassetto)

2) H-IL-32 : Sintesi Chimica della parte N-terminale γ (gamma) isoform (103 residui) per compararla con quella ricombinante.

M**C**FPKVLSDD MKKLKARMVM LLPTSAQGLG AWVS**AC**DTKD TVGHPGPWRD KDPAL**WC**Q**LC** LSSQHQAIER FYDKMQNAES GRGQVMSSLA ELEDDFKEGY LET
(I siti di ligazione sono sottolineati in grassetto)

Risultati conseguiti.

1) H-IL-18

La sequenza della proteina e' stata divisa in 4 frammenti per poter essere sintetizzata utilizzando Native Chemical Ligation. Tutti i frammenti (**I**, **II**, **III** e **IV**) sono stati sintetizzati e purificati. Particolarmente difficili sono state le sintesi dei frammenti I e III che sono stati sintetizzati piu' volte per ottimizzare le rese.

2) H-IL-32

Come strategia di sintesi, Il polipeptide e' stato diviso in quattro frammenti. La sintesi piu' difficile e' stato il segmento C-terminale che e' stato risintetizzato 2 volte. Tutti gli altri frammenti non hanno avuto alcun problema. Il polipeptide e' stato effettivamente sintetizzato nella sua interezza via Native Chemical Ligation e nei test effettuati dal Prof. Dinarello presso L'Universita' del Colorado a Denver ha mostrato un'attivita' biologica (in presenza di costimuli quali MDP o LPS) comparabile con la proteina intera confermando quindi.la validita' dei dati ottenuti con la proteina ricombinante. Infatti il polipeptide in presenza di MDP e' stato capace, in differenti modelli di indurre rispettivamente produzione di IL-6, IL-1β and TNFα. Inoltre sono stati testati anche I tre frammenti a maggior numero di residui utilizzati per la sintesi,

IL-32: Peptidi testati per attivita' biologica.

1) Il-32 gamma isoform 1-103
2) MCFPKVLSDD MKKLKARMVM LLPTSAQGLG AWVSA (Fragment **V**)
3) CDTKD TVGHPGPWRD KDPALW (Fragment **VI**)
4) C LSSQHQAIER FYDKMQNAES GRGQVMSSLA ELEDDFKEGY LET (Fragment **VIII**)

I Frammenti **V**, **VI** e **VIII** non hanno pero' evidenziato alcuna attivita' biologica rilevante anche in presenza di costimuli.

Summary

Recent advances in the Solid Phase Peptide Synthesis (SPPS) have led to the synthesis of polypeptides in a robust and reproducible manner. Nowadays peptide fragments up to 60-70 residues can be routinely synthesized with high degree of homogeneity. Nevertheless, to chemically synthesize a polypeptide or a small protein of 100 residues or more a way to chemically join two peptide fragments is a necessity. Several chemoselective ligations have been developed to join unprotected peptide fragments in aqueous solution, however most of them do not form a native amide bond with a native structure. The most robust and widely used amide forming ligation is referred to as Native Chemical Ligation (NCL), which during the last decade has extensively demonstrated its usefulness for the rapid preparation of small-medium size proteins with high level of homogeneity. Such methodology exploits a chemoselective reaction between two unprotected fragments, a C-terminal thioester and a N-terminal cysteine, in aqueous solution at neutral pH thus generating a native amide bond at the ligation site.
In spite of the advances of the recombinant techniques to express, isolate and purify the desired protein, the chemical protein synthesis can still overcome some unresolved issues related to the use of recombinant material, such as refolding following removal of high levels of urea required for extraction from E.coli. In addition, the sensitivity of mammalian tissue to endotoxins and other products of microbial origin interfere and obscure the biological properties of recombinant cytokines.
In this study we discuss and tackle the chemical synthesis of two protein therapeutics, IL-18 and IL-32, which would greatly benefit from the chemical synthesis of their modified analogues.

IL-18

HIL-18 is a proinflammatory cytokine which belongs to the family of IL-1 ligands. Initially cloned as an IFN-γ inducing factor, plays a key role in many inflammatory diseases including allergy and autoimmune diseases. Recently, after preclinical studies on monkeys, therapeutic approaches using recombinant IL-18 have been examined for treatment of cancers including a clinical trial in humans indicating the necessity for a more stable form that will allow reproducible and reliable therapy for clinical uses. Recombinant h-IL-18 has been found to be susceptible of loss of biological activity upon storage and handling. Some authors indicated the misfolding of the molecule as cause of the loss of biological activity, while others challenged the hypothesis of the misfolding indicating the polymerization byproducts as cause of the inactivation. A possible explanation is the relative instability of intra-molecular disulfide bridges to atmospheric oxidation, which render the molecule prone to inter-molecular disulfide bond formation.
Due to the relevance of the medical implications and the potential impact on the treatment of some diseases (the molecule is currently under Clinical Trials to treat different types of tumors) there is an unmet need to find a more stable form for the protein that could lead to a superior drug candidate for the envisaged medical applications.

IL-32

IL-32 is a new inflammatory cytokine. Although IL 32 does not share sequence homology with known cytokine families, IL 32 induces various cytokines, human TNFα and IL 8 in THP 1 monocytic cells as well as mouse TNFα and MIP 2 in RAW macrophage cells. IL-32 activates typical cytokine signal pathways of nuclear factor kappa B and p38 mitogen activated protein kinase.
Human IL 32 exists as four splice variants and IL 32 from other species were found as expressed sequence tag clones in the databank.
The major problem with IL-32 is first the issue of contamination from E. coli products. There is synergism between peptidoglycans from the bacterial wall with IL-32 and therefore is not possible to unambiguously evaluate IL-32 alone.
H-IL-32 γ (gamma) isoform consists of 164 residues. In collaboration with the group of Prof. C. Dinarello, we wanted to synthesize the N-terminal segment (103 residues) of the γ (gamma) isoform. IL-32 binds to proteinase-1 (PR3) with a high affinity. When recombinant IL-32 was exposed to PR3 for 5 minutes, there was an increase in biological activity suggesting that the cleavage by PR3 generated a product more active. Prolonged exposure to PR3 resulted in loss of IL-32 activity. Upon examining the potential cleavage sites for PR3 in IL-32, the N-terminal 103 amino acids would be generated upon a short exposure to PR3. Other cytokines such as IL-1β and IL-8 also increase their biological activity following a short incubation with PR3. Therefore, the N-terminal 103 amino acid fragment was thought to posses same biological activity compared to the whole protein, possibly owning easier handling properties due to reduced aggregation. Structurally, it possesses 103 residues with 4 cysteines and two potentials disulfide bridges.

Objectives:

H-IL-18: Chemical synthesis of the wild-type protein to compare with the recombinant version and upon obtained results Chemical protein engineering to stabilize and enhance its pharmaceutical properties.
H-IL-18 consists of 157 residues.

YFGK LESKLSVIRN LNDQVLFIDQ GNRPLFEDMT DSDCRDNAPR TIFIISMYKD SQPRGMAVTI SV**KC**EKISTL SCENKIISFK **EM**NPPDNIKD TKSDIIFFQR SVPGHDNKMQ FESSSYEGYF L**AC**EKERDLF KLILKKEDEL GDRSIMFTVQ NED
(Bold and underlined are ligation sites)

Results:

All four fragments according to our strategy were finally successfully synthesized and purified. At the time being, the ligation steps to assembly the protein are still in progress.

H-IL-32: Total chemical synthesis of the N-terminal part of the γ (gamma) isoform (103 residues) to compare with the recombinant version. As in the case of IL-18 since the length of the polypeptide in matter is still significantly above the capabilities

of the SPPS, we had to use a chemical ligation strategy, notably Native Chemical Ligation.

M**C**FPKVLSDD MKKLKARMVM LLPTSAQGLG AWVS**AC**DTKD TVGHPGPWRD KDPAL**WC**Q**L**LSSQHQAIER FYDKMQNAES GRGQVMSSLA ELEDDFKEGY LET
(Bold and underlined are ligation sites)

Results:
The full-length N-terminal part 1-103 of the γ (gamma) isoform was synthesized and sent along with the intermediate peptide fragments to Prof. Dinarello for biological activity tests.

IL-32: Peptides tested for biological activity

5) Il-32 gamma isoform 1-103
6) MCFPKVLSDD MKKLKARMVM LLPTSAQGLG AWVSA (Fragment **V**)
7) CDTKD TVGHPGPWRD KDPALW (Fragment **VI**)
8) C LSSQHQAIER FYDKMQNAES GRGQVMSSLA ELEDDFKEGY LET (Fragment **VIII**)

Il-32 gamma isoform 1-103 N-terminal amino acids was found to be active in the presence of the co-stimulus) MDP. Previous studies published in Proceedings of the National Academy of Sciences demonstrated that MDP synergizes with IL-32 gamma for the induction of IL-6. This synergism was shown to be dependent on caspase-1 via the induction of IL-1β. The importance of the observation is that MDP triggers the intracellular receptor known as NOD2 (Nuclear Oligomerization Domain). Patients with a forward shift mutation in NOD2 have Crohn's Disease. The finding that synthetic N-terminal IL-32 103 amino acids synergized with MDP was essential since the synthetic 103 peptide lacked any microbial products. Indeed, the polypeptide in presence of MDP was able, in different models, to induce respectively production of IL-6, IL-1β and TNFα.
Such results thus confirm and validate the original data obtained with the whole protein produced with recombinant techniques. Not unexpectedly, the single fragments (Fragment **V**, **VI** and **VIII**) utilized to build the protein as described above were each found not to be functional

1. Introduction

1.1 Relevance of Chemical Protein Synthesis

Protein therapeutics are continuously increasing attention and gaining wider market in the pharmaceutical business.[1] Recent advances in the bioconjugation techniques as well as in the production of biocompatible polymers (i.e. Pegylation) allowed enhanced protein therapeutics candidates with improved half-life and reduced side effects (immunogenicity)[2].
Nevertheless, the chemical protein synthesis can still overcome some unresolved issues related to the use of recombinant material.[3,15]
In this study we will discuss the chemical synthesis of two protein therapeutics, IL-18 and IL-32, which would benefit from the chemical synthesis of their modified analogues.

1.2 Overview of Chemical Protein Synthesis

Peptide synthesis began in 1901 by Emil Fisher with the synthesis of glycyl-glycine, the first synthesis of a free peptide.

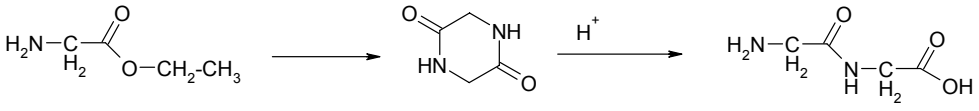


Fig 1.1 Emil Fisher and his synthesis of glycyl-glycine,1901, from M. Goodman[4]

After the synthesis of the gly-gly dipeptide many improvements in the technique occurred along the years, however until the discovery of the Solid Phase Peptide Synthesis (SPPS)[5] by Bruce Merrifield (Nobel Laureate) only small peptides could be prepared. Using SPPS, polypeptides of various sizes could be synthesized and upon significant advances in the field[6] small proteins[7] and large polypeptides can be now obtained. All methods developed thereafter along the years to bridge the gap from peptide to chemical protein synthesis are intrinsically derived from the Merrifield's invention. However, still nowadays synthetic peptides above 60-70 residues cannot be routinely prepared in good yield with a high degree of homogeneity. Above such limit, poor yields and high impurity often characterize the product made by SPPS.

Incomplete Sequences[8]

These are formed during the assembly of a sequence :with a 100 amino acids synthesis the final yield should be, maximally :

Coupling yield	final yield
99.99	99.00
99.90	90.47
99.60	66.97
99.30	49.53
99.00	36.60
98.00	13.26

Starting with 250 mg of a resin with a substitution of 0.4 mmoles/g, assuming a coupling yield of 99.0 % the theoretical final total yield is 1.1 g, of which 396 mg are the correct peptide, and 704 mg are impurities, material in which one or more amino acids are missing. These numbers consider only chain assembly, and *do not take* into account side reactions during cleavage and handling.
RP-HPLC cannot resolve the impurities and separate the pure peptide since the change in hydrophobicity due to a single amino acid over 100 a.a. is too small to be resolved with current methods.
Therefore to chemically approach proteins of small and medium size (e.g. polypeptide chains greater than 100 a.a. residues) there are two possible solutions:

1) Developments of new methods for purification of intermediates and final products to selectively isolate the desired compound in the complex mixture after SPPS.

2) The necessity to introduce of novel technology to join together 2 or more peptide segments with ease and high efficiency.

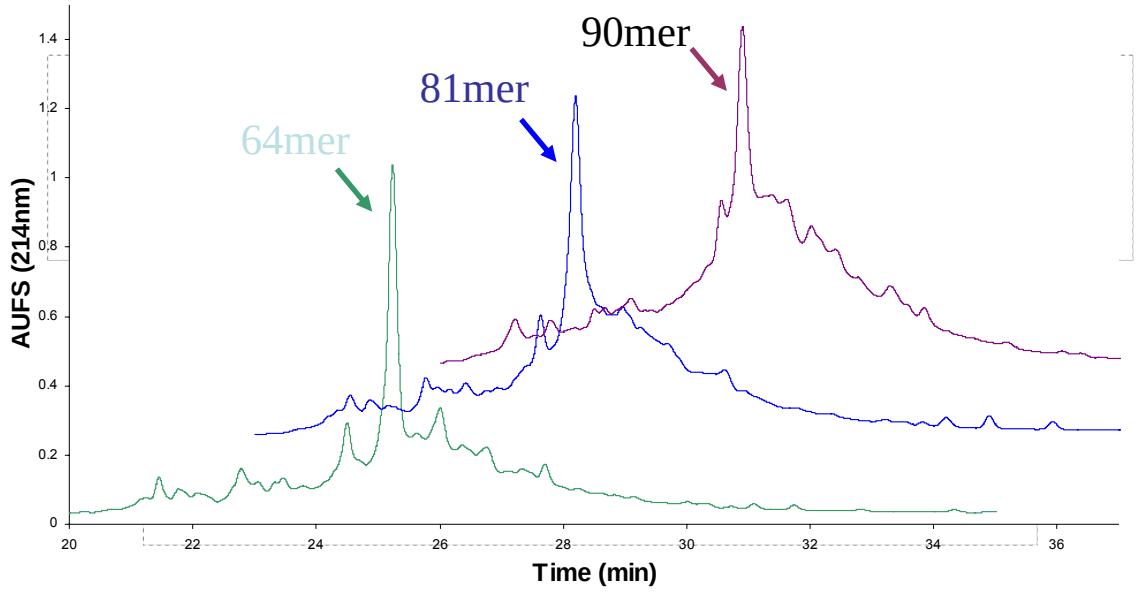


Fig 1.2 Synthesis of peptide fragments: HPLC profile of crude peptides of different size of the same molecule[9].

1. Development of new purification methods relays on the attachment of a chemical probe at the end of the synthesis with a unique chemical moiety. Thus, only the full length material functionalized with the unique probe, is finally amenable to further react in a chemoselective way with a specific chemical group[10]. Finally, the chemoselective reaction allows the isolation of the full length material. In other cases, an hydrophobic and reversible chemical probe (i.e.Fmoc)[11] or Tbfmoc[12] were added

at the end of the synthesis, allowing easier purification. The impossibility to remove full length material that underwent to chemical modifications (side reactions, incomplete protecting group cleavage) still greatly limits such approach.

2. Possible approaches for fragment condensation:

A) Fragment condensation with protected peptides
B) Fragment condensation via enzyme catalysis (peptide ligase)
C) Chemical Ligation with unprotected peptide fragments

A) Fragment condensation with protected peptides:

The fragment condensation approach is the oldest strategy to synthesize large polypeptides and proteins. Such approach can be further divided in fragment condensation with (i) fully protected peptides and (ii) minimal protecting group strategy. The first approach with fully protected peptides was pioneered by Sakakibara,[13] and can count on many successes amongst which the total synthesis of GFP[14]. Nevertheless, the strategy presents obvious difficulties: purification the fully protected peptide fragments, solubilization of protected peptide fragments at suitable concentration, and finally possible racemization when the activated residue is not a glycine.

The second approach with partially unprotected peptides is based on a low enthalpic activation, which allows hydroxyl moieties to be unprotected thus giving advantages in aqueous solubility. This approach was further improved by Ramage[15] by using a transfer active ester condensation technique and although some interesting molecules were synthesized, it has not found many followers.

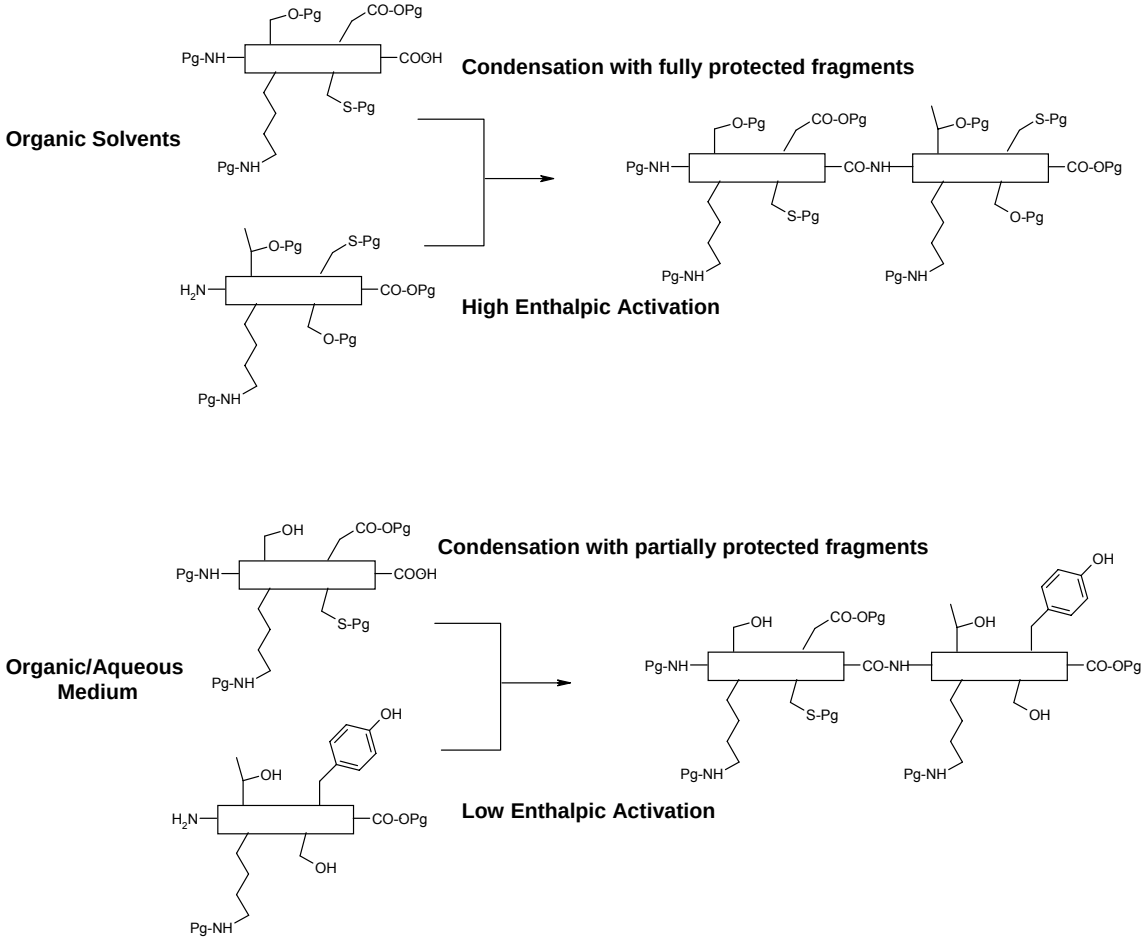


Fig 1.3 Fragment condensation with protected peptides

B) Fragment condensation via enzyme catalysis (peptide ligase):
The approach was pioneered by Wells[16] and co-workers at Genentech. It is based on the reversed engineering of enzymes (subtiligase), which under appropriate conditions allows peptide bond formation (ligation) rather than cleavage (proteolysis). The approach had a significant impact related to the publication on Proceedings of the National Academy of Sciences, however along the years the technique did not find many followers possibly due to limitation on the sequence alignment suitable for the site of ligation.
C) Chemical Ligation with unprotected peptide fragments:
Chemical Ligation with unprotected peptide fragments can be further divided into the following categories:
i) Chemoselective Coupling (Chemical Ligation) of Unprotected Peptides
Formation of an « Analog Structure »
ii) Native Chemical Ligation
Formation of a Native amide Bond with a Native Structure

i) Chemoselective Coupling of Unprotected Peptides in aqueous medium with formation of a non native structure « Analog Structure »

The chemoselective coupling, can be further divided into two classes:
Chemoselective Ligation via Thiol Chemistry and Chemoselective Ligation via Weak Base-Carbonyl Chemistry

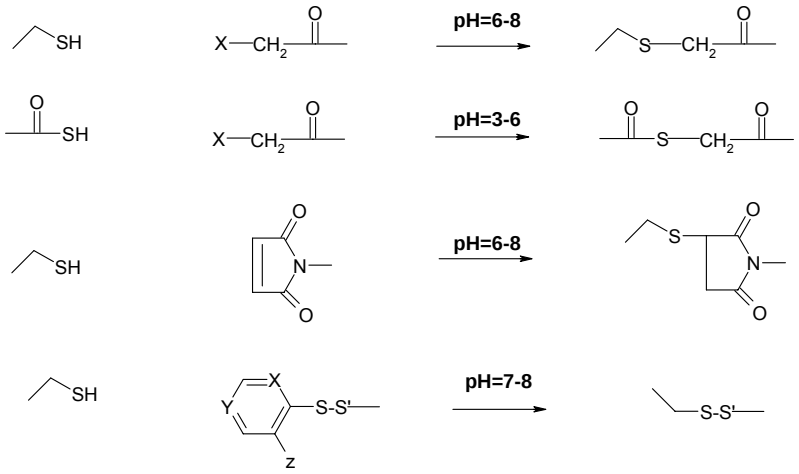


Fig 1.4 Chemoselective Ligation via Thiol Chemistry

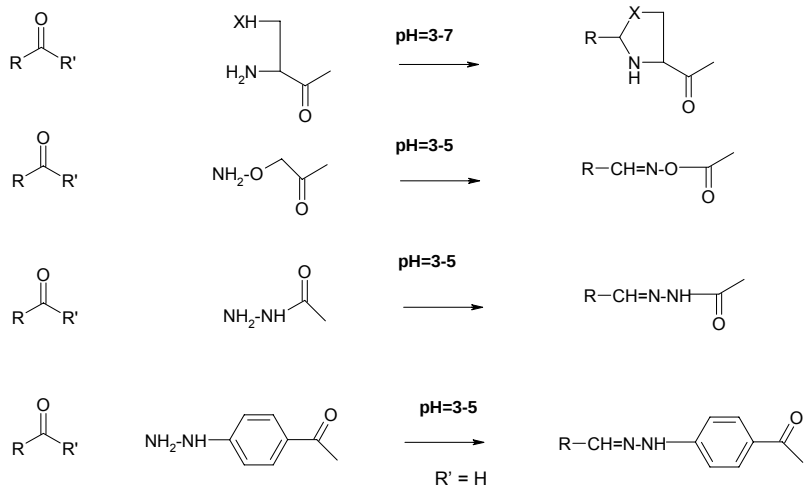


Fig 1.4 Chemoselective Ligation via weak-base carbonyl chemistry[17]

The most relevant example of thiol based ligation is the total Synthesis of HIV-1 Protease via thioester forming chemical ligation by the group of Kent[18]. In the early 90s, many groups (especially *via* recombinant techniques) unsuccessfully tried to produce and isolate the HIV-1 protease. The work of Kent's group allowed the isolation of the fully active protein and based upon the information obtained from crystallography[19] medicinal chemist could design small molecules and peptides that are now at the basis of the anti-HIV treatment.

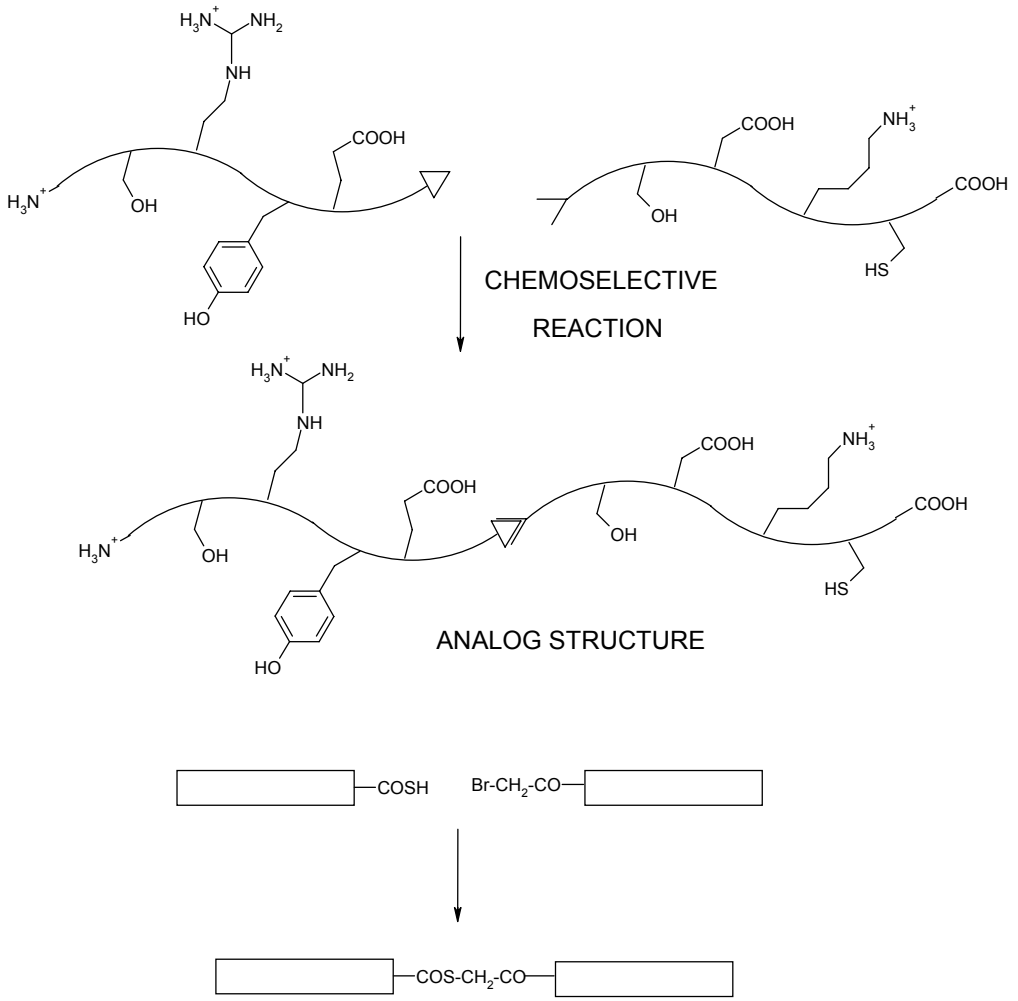


Fig 1.5 Chemoselective Ligation via thioester formation

Concerning Ligation via weak-base carbonyl chemistry, the most efficient and widely used is the oxime formation developed by K.Rose[20]. In another approach, the synthesis of a GCSF analogue[21] was achieved using hydrazone formation.
The ligation scheme proposed by Tam[22] is of particular interest since generates a native amide bond with a non native structure. Amongst the carbonyl based ligations is the strategy that best approaches a native architecture by structure similarity: in fact replacement of a native proline with hydroxymethyl thioproline as final product of the ligation proposed by Tam did not affect biological activity of a HIV-1 protease fragment.[23]

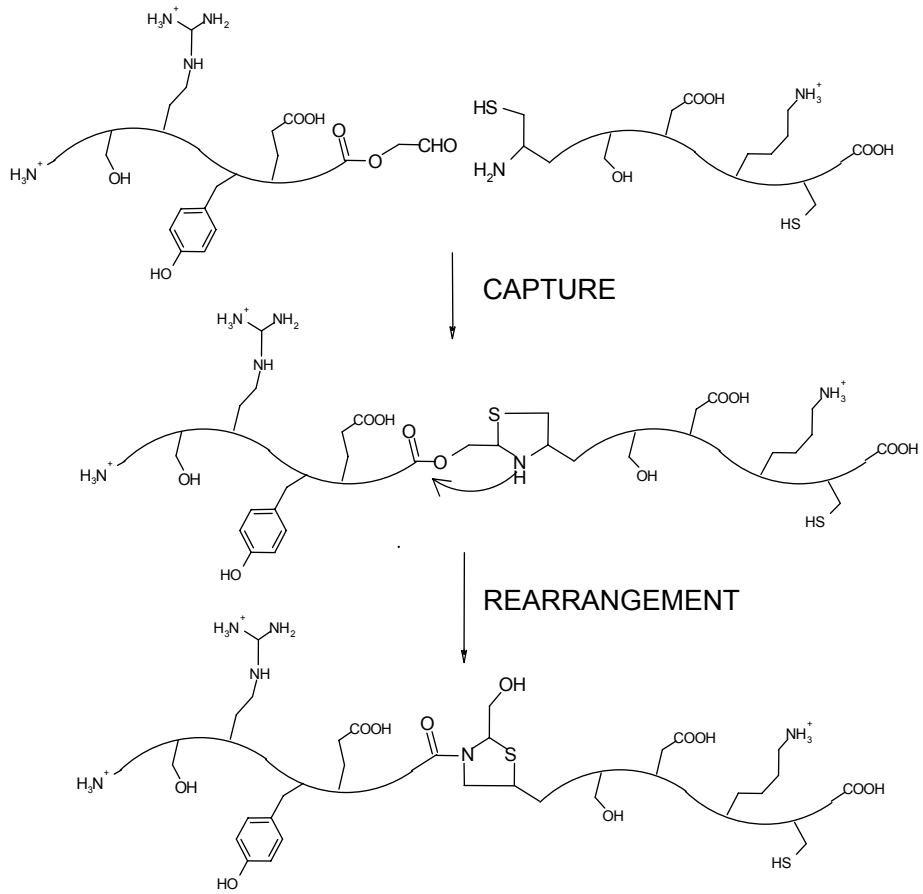


Fig 1.6 Chemoselective Ligation carbonyl based via etherocyclic ring formation: formation of a native amide bond with a non-native structure

ii) Native Chemical Ligation (NCL): generation of a native amide bond with a native structure

Native Chemical Ligation (NCL)[24] has been extensively demonstrated to be the key to the rapid preparation of small and medium size proteins with high level of homogeneity. In the NCL scheme a C-terminal thioester fragment reacts chemoselectively in aqueous solution at neutral pH with an N-terminal cysteine peptide. In the first step of the reaction transthioesterification takes place by thiol exchange between the free thiol of the N-terminal cysteine and the thioester moiety on the other molecule. The newly generated thioester then undergoes an S to N acyl shift due to the proximity of the amino group to the thioester functionality, thus generating a native amide bond at the ligation site. Native Chemical Ligation (NCL) allows the chemical synthesis of proteins from database sequence in a timely and cost-effective manner. NCL has become the method of choice for the rapid preparation of small and medium cysteine rich proteins.[25]

The requirement for cysteine at the site of peptide ligation is an intrinsic restriction of the standard NCL strategy. In proteins, cysteine occurs with a

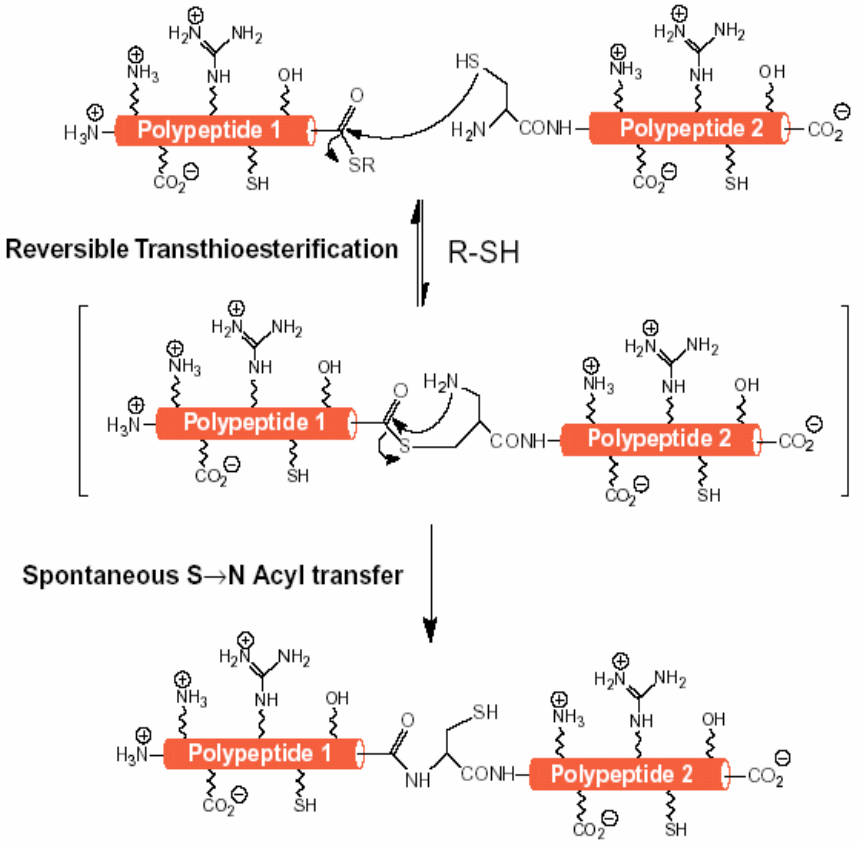


Fig 1.7 Native Chemical Ligation (NCL)

relatively low frequency; many protein sequences do not have suitably disposed cysteines for the NCL strategy and some lack cysteine residues entirely.
Thus, the continuously increasing demand of large polypeptides and proteins has led during the last few years to the development of removable auxiliary groups to extend the applicability of NCL to ligation and cyclization of unprotected peptides in aqueous solution at an X-X site (where X= any amino acid)
The first extension of the NCL was presented by Canne et al.[26] who exploited two acid stable N^{α}-thiol containing auxiliaries, designed to mimic the side chain of a native cysteine. The reductively cleavable Nα-(oxyethanethiol) 1 that has a 1,3-aminothiol structure, showed a relatively slow amide forming rate while the Nα–(2-mercaptoethyl) 2, which was not designed to be cleavable, showed superior rearrangement rates (Scheme 1).
The interesting propensity of the latter to a faster ligation prompted to the design of a novel type of auxiliary group that could take advantage of the superior ligation rates of 2 and allow its facile and complete removal after ligation.

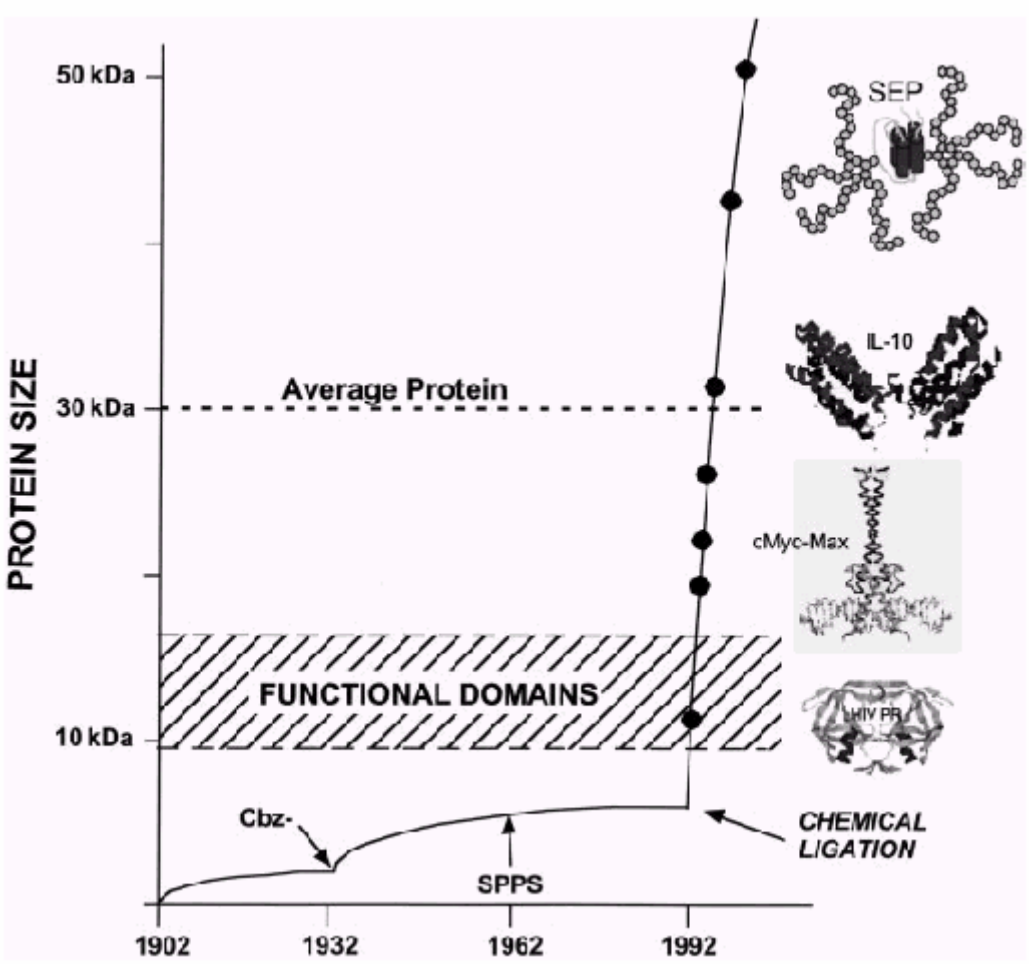


Fig1.8 From SBH Kent[27]

A case in point is the Nα-(1-phenyl-2-mercaptoethyl) auxiliary[28], which combines the faster ligation rate of the Nα–(2-mercaptoethyl) group through a 5 membered ring intermediate, enabling at the same time its complete removal after ligation. Indeed, the 1-phenyl substitution on the Nα–(2-mercaptoethyl) template generates a benzylamine derivative that is completely stable in the strong acidic condition used to cleave the peptide from the resin (HF, TFA) making the auxiliary compatible with both Boc and Fmoc like chemistries. Furthermore, because of the S to N acyl shift intrinsic mechanism of NCL, benzylamine to benzylamide conversion occurs during the reaction making the auxiliary readily cleavable under similar or milder conditions originally required for the cleavage of the peptide.

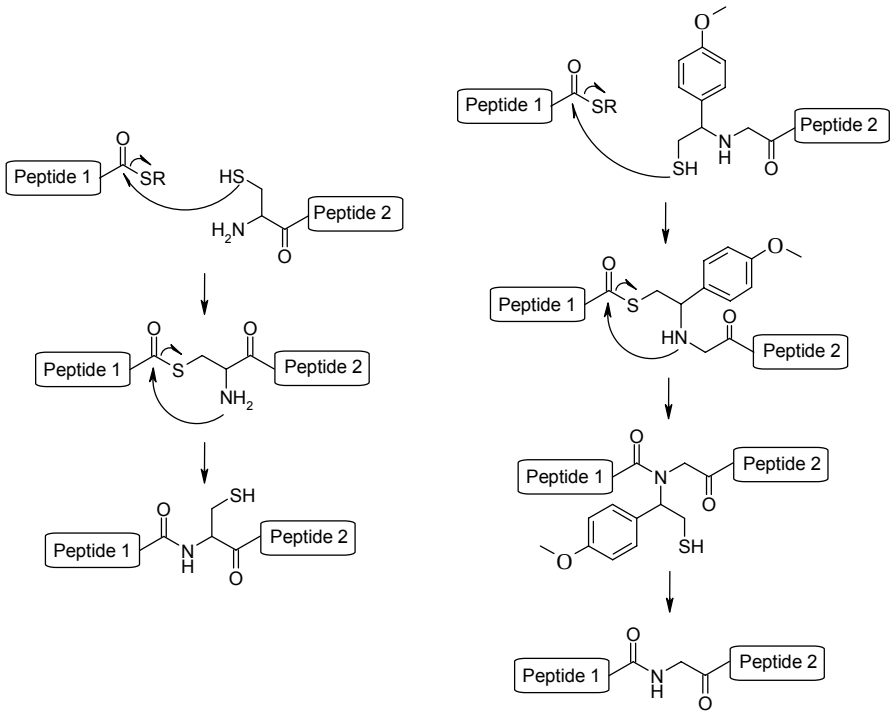


Fig 1.9 Native Chemical Ligation and Extended Chemical Ligation

Furthermore such auxiliary group has been successfully employed for the synthesis of cyclic peptides with potential therapeutic interest.[29]
However, for all different types of auxiliary groups[30] that have been developed to extend the application of NCL to other ligation sites, the slower ligation rates especially with large fragments and the additional step required to cleave the auxiliary post-ligation have reduced their practical utility and use.
Recently, a novel method to synthesize proteins through a chemical ligation using unprotected peptide segments has been presented[31]. Such scheme does not make use of auxiliary groups, instead originally exploits the features of some side chain removable functionalities. Ligation rates are high, comparable to NCL and the residues available for ligation are more frequent than cysteine. Furthermore the whole process is “one pot” and at the end a native polypeptide is obtained directly in the ligation mixture.
The new strategy takes advantage of the benzylic moiety of the aromatic residues to release a mercaptan under specific conditions. We use phenylcysteine to exploit ligation with thioester fragments and the alkylation post-ligation is designed to produce a benzylic thioether suitable for cleavage. Benzyl groups are commonly employed in peptide synthesis to protect side chain functionalities via ether (Ser and Thr) and thioether (Cys) bonds. Thus, after ligation, alkylation of the free mercaptan generates a benzyl thioether suitable for cleavage.

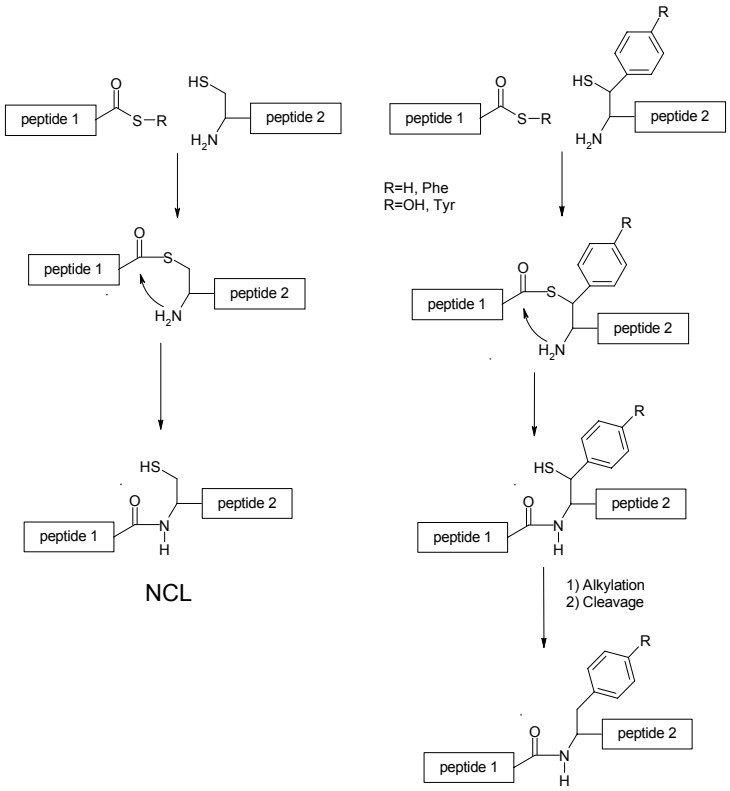


Fig 1.10 Native Chemical Ligation at Aromatic Residues (Phe, Tyr and Trp).

1.3 Problems related to the clinical use of therapeutic cytokines

Although cytokine biology has advanced rapidly with the ability to produce recombinant molecules in E.coli, the field still suffers from problems related to low biological activities and the problems of endotoxin contamination as well as peptidoglycan. This latter issue is of paramount importance in testing new cytokines such as interleukin-IL-32. However, it is particularly a problem for some cytokines such as IL-1b and IL-18. In the case of IL-1b, the natural molecule is produced as an inactive precursor and the intracellular enzyme caspase-1 cleaves the precursor into an active cytokine. However, the N-terminal peptide in E. coli exhibits variable biological activities (variable specific activities in terms of units/milligram). The variability appears to be due to more than one reason. For example, deletion of the terminal methionine from the E- coli product enhances specific activity and arginine in position number 4 appears to be critical. Most of the problems with human recombinant IL-1b are now resolved but it remains an excellent example of the problem of assessing biological activities of new molecules. Naturally, endotoxin content always remains a significant confounding issue. In the case of IL-1b, the ability to study its pyrogenic properties requires strict removal of endotoxins. IL-18 is a recent example. Recombinant mouse IL-18 produced in E.coli is biologically active in the picomolar range, similar to several proinflammatory cytokines. However, human IL-18 remains problematic in that one needs nanomolar concentrations.

1.4. IL-18

H-IL-18[32] is a cytoplasmic protein produced as biologically inactive 24 kDa precursor that requires to be cleaved by caspase-1 into the biologically active mature 18 kDa molecule. IL-18 is a proinflammatory cytokine which belongs to the family of IL-1 ligands, initially cloned as an IFN-γ inducing factor, plays a key role in many inflammatory diseases including allergy and autoimmune diseases[33]. Notably, IL-18 by inducing production of IFN-γ from T cells, has been successfully used to boost immune-defense in mice to treat tumors[34]. Recently, after preclinical studies on monkeys[35], therapeutic approaches using recombinant IL-18 have been examined for treatment of cancers including a clinical trial in humans indicating the need for a more stable form that will allow reproducible and reliable therapy for clinical uses.

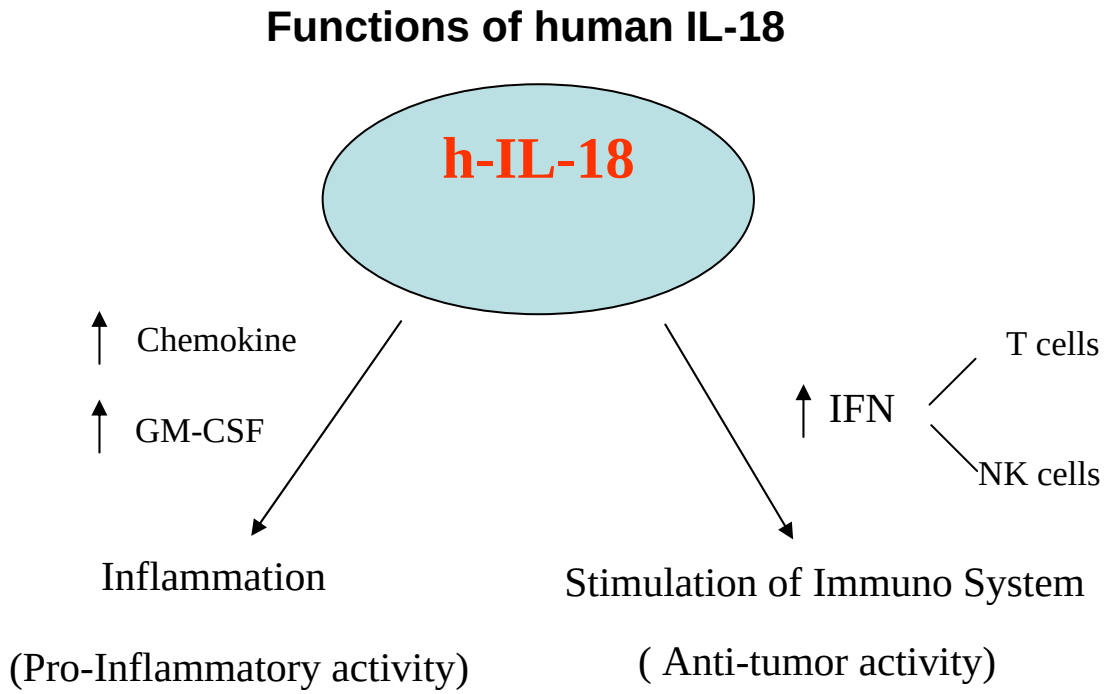


Fig 1.11 Functions of IL-18

Recombinant h-IL-18 has been found to be susceptible of loss of biological activity upon storage and handling. IL 18 induces Interferon gamma but depending on the folding of the molecule and the state of disulfide bonds (S-S) reduction, the induction of IFNγ (gamma) can be different by a factor of 10 or more. Naturally this is a problem in human trials. Depending on the specific activity (induction of IFNγ units/ milligram of IL-18), this can vary if the protein is not stable or folded correctly. These issues of low activities of IL-18 were discussed by the group of C. Dinarello.[36] Some authors[37] indicated the misfolding of the molecule as cause of the loss of biological activity.
Recently though, another paper[38] challenged the hypothesis of misfolding indicating the polymerization byproducts as cause of the inactivation. A possible explanation is

the relative instability of intra-molecular disulfide bridges to atmospheric oxidation, which render the molecule prone to inter-molecular disulfide bond formation.
Due to the relevance of the medical implications and the potential impact on the treatment of some diseases (the molecule is currently under Clinical Trials to treat different types of tumors) there is an unmet necessity to find a more stable form for the protein that could lead to a superior drug candidate for the envisaged medical applications.

1.5 IL-32

IL-32 is a new inflammatory cytokine.
14 years ago a cytokine like molecule described as natural killer cell transcript 4 (NK4)[39] was discovered, but its function remained unknown until recently.
Increased gene expression of NK4 has been reported in PBMC from patients receiving high dose IL 2 therapy for malignant melanoma[40], but the function of NK4 has remained unknown until the recent study by the group of Dinarello[41].
NK4 was then found to be a novel inflammatory cytokine and re-named asIL-32, and its biological function, regulation of production, genomic structure and signal transduction has been described[33]. Although IL 32 does not share sequence homology with known cytokine families, IL 32 induces various cytokines, human TNFα and IL 8 in THP 1 monocytic cells as well as mouse TNFα and MIP 2 in RAW macrophage cells. IL-32 activates typical cytokine signal pathways of nuclear factor kappa B and p38 mitogen activated protein kinase.
Human IL 32 exists as four splice variants and IL 32 from other species were found as expressed sequence tag clones in the databank.

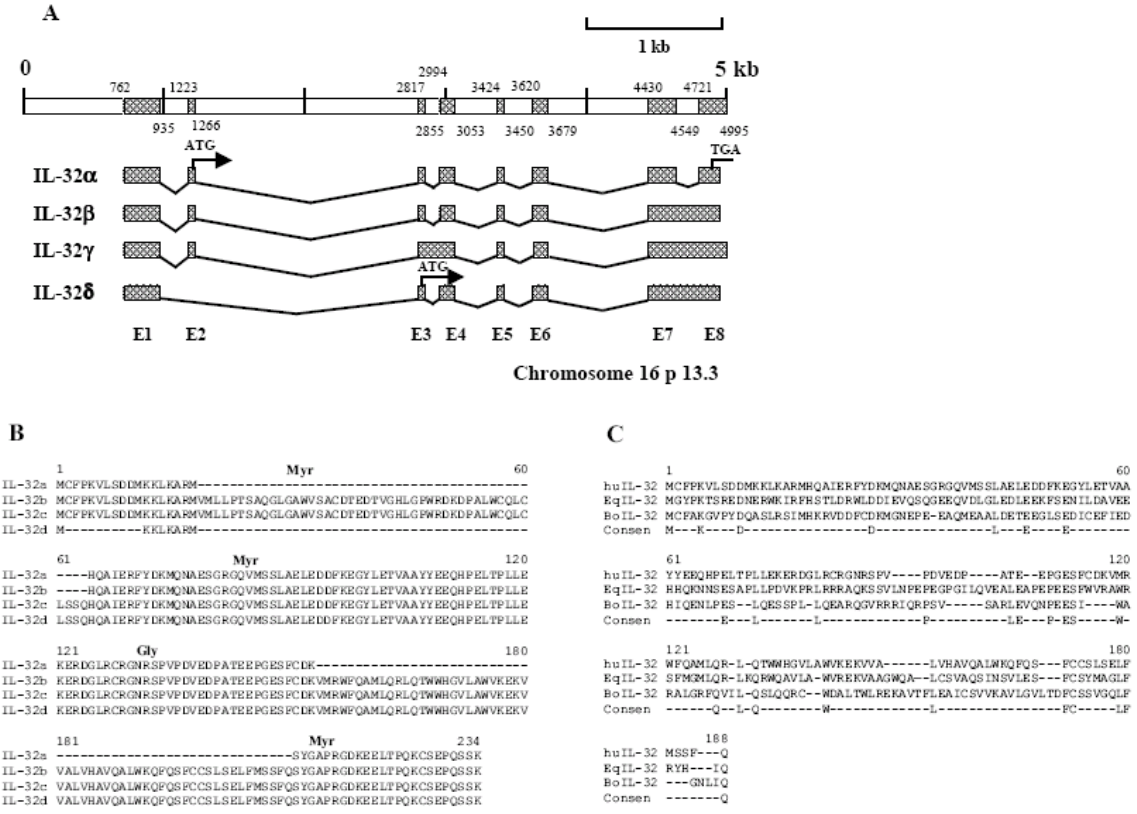


Fig 1.12 The genomic organization of four IL-32 splice variants and other species of IL-32 by amino acid sequence alignment, from Soo-Hyun Kim et others[39].

The major problem with IL-32 is first the issue of contamination from E. coli products. There is synergism between peptidoglycans from the bacterial wall with IL-32 and therefore, so is not possible to evaluate IL-32 alone.
The next problem is aggregation of the gamma isoform:
The alpha isoform is produced in high amounts from E. coli but the activity is extremely low compared to the gamma isoform.
Finally the gamma isoform production is also low. [42]
h-IL-32 γ (gamma) isoform consists of 164 residues. In collaboration with the group of Prof. C. Dinarello, we wanted to synthesize the N-terminal fragment (103 residues) of the γ (gamma) isoform[50]. The N-terminal fragment was thought to posses same biological activity compared to the whole protein, possibly owning easier handling properties due to reduced aggregation[33]. Structurally, it possesses 103 residues with 4 cysteines and two potentials disulfide bridges.
There is also a need for IL-32 reagents for making neutralizing antibodies. First, the synthetic peptides and antibodies to these may help in the development of better ELISA detection methods. Indeed, having synthetic active peptides for IL-32 (such as 103 IL 32 γ and even smaller peptides), biologists can make monoclonal antibodies and look for epitopes for neutralization. These may then become therapeutic agent in treating Myelodysplastic Syndrome[43], COPD[44], rheumatoid arthritis[45], M. tuberculosis[46] and Crohn's Disease[55] .
There is also a therapeutic role for IL 32 in HIV-1[47]

2. Aim of the thesis

Objectives:
1) H-IL-18

Chemical synthesis of the wild-type protein to compare with the recombinant version and upon obtained results Chemical protein engineering to stabilize and enhance its pharmaceutical properties.
H-IL-18 consists of 157 residues. [48]

Sequence of h-IL-18

YFGK LESKLSVIRN LNDQVLFIDQ GNRPLFEDMT DSDCRDNAPR TIFIISMYKD SQPRGMAVTI SV**K**C**C**EKISTL SCENKIISFK **E**M**M**NPPDNIKD TKSDIIFFQR SVPGHDNKMQ FESSSYEGYF L**A**C**C**EKERDLF KLILKKEDEL GDRSIMFTVQ NED
(Bold and underlined are ligation sites)

Fragment **I**:
YFGK LESKLSVIRN LNDQVLFIDQ GNRPLFEDMT DSDCRDNAPR TIFIISMYKD SQPRGMAVTI SVK- C^{α}SR

Fragment **II**:
C(thiaz)EKISTL SCENKIISFK E(Fm)[49] - C^{α}SR

Fragment **III**:
hCys[50](thiaz)NPPDNIKD TKSDIIFFQR SVPGHDNKMQ FESSSYEGYF LA- C^{α}SR
Fragment **IV**
CEKERDLF KLILKKEDEL GDRSIMFTVQ NED

2) H-IL-32 :

Total chemical synthesis of the N-terminal part of the γ (gamma) isoform (103 residues) to compare with the recombinat version. Again, as in the case of IL-18 since with the length of the polypeptide in matter is significantly above the capabilities of the SPPS, we had to use a chemical ligation strategy.

M**C**FPKVLSDD MKKLKARMVM LLPTSAQGLG AWVS**AC**DTKD TVGHPGPWRD KDPAL**WC**Q**LC** LSSQHQAIER FYDKMQNAES GRGQVMSSLA ELEDDFKEGY LET (Bold and underlined are ligation sites)

MCFPKVLSDD MKKLKARMVM LLPTSAQGLG AWVSA- C^{α}SR (Fragment **V**)
C(thiaz)DTKD TVGHPGPWRD KDPALW C^{α}SR (Fragment **VI**)
C(thiaz)QL- C^{α}SR (Fragment **VII**)
C LSSQHQAIER FYDKMQNAES GRGQVMSSLA ELEDDFKEGY LET (Fragment **VIII**)

N-terminal fragment:1-35 Cα-COS-R
Intermediate fragment: 36-56: nterminal cys(thiazolidine) Cα-COS-R
C-terminal fragment 57-103: Ligation 57-59 with 60-103 to overcome synthetic and solubility issue of the fragment

3. Results and discussion

Peptides were synthesized using Boc chemistry In Situ neutralization chemistry as pioneered by Steve Kent[5]. Cα-thioester peptides were prepared according to the protocol of Hackeng et al.[51]. Chemical ligation of peptide fragment was carried out from C termini to N termini direction, N terminal 1-2 amino thiol (i.e. cysteine) was protected through thiazolidine formation.[52]. Chemical ligations were carried out as described in the material and methods section.

3.1 IL-18

Results Obtained

Originally, the synthesis was designed in 3 fragments ligation: 1-67, 68-126 and the C-terminal fragment 127-157. Particularly difficult was the synthesis of the N terminal fragment both for the size and solubility problems. The N-terminal fragment 1-67-C-terminal thioester gave initially solubility problems leading to poor yields after purification. Finally the 1-67 was successfully synthesized adding 3 lysines residues on the thioester part to overcome solubility problems due to its polycationic feature as described in Fig **3.1**.

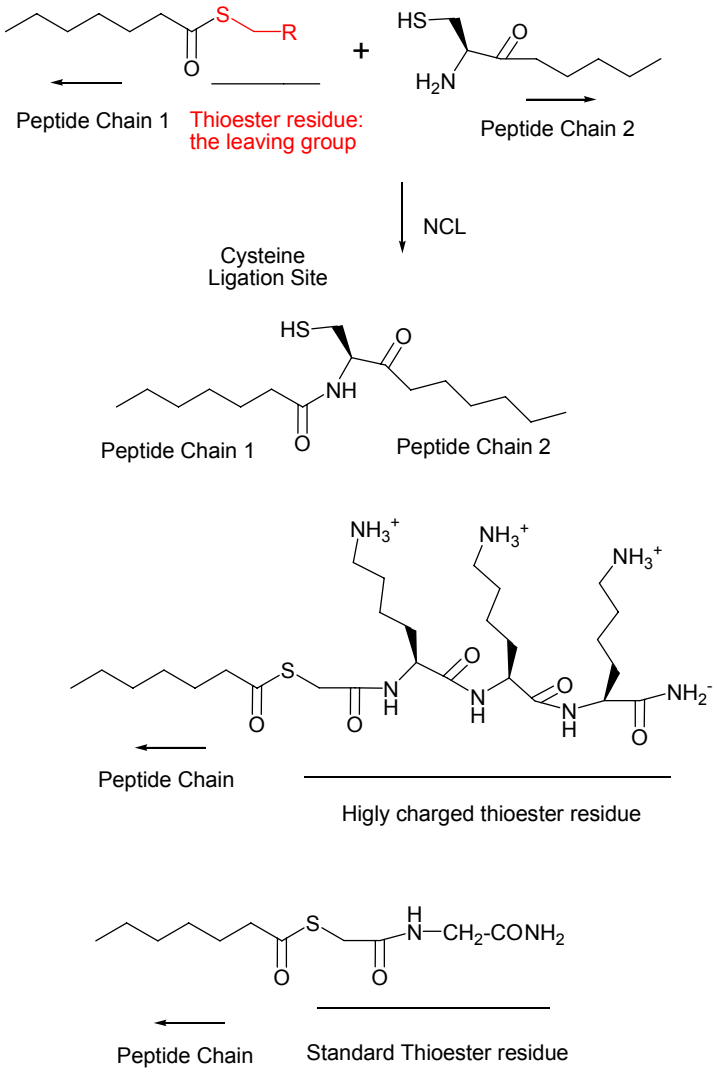


Fig. **3.1** Thioester comparison

IL18 Synthetic Strategy

Due to its high number of residues, the N terminal fragment was synthesized from residue 1 untill threonine 34 utilizing standard In Situ neutralization[5] chemistry at 0.2 mmol scale resin substitution; Then 50 % of the resin was removed, thus the synthesis was finally carried out and completed at 0.1 mmol scale with capping at the end of each cycle using acetilglycine.

The C terminal portion 127-157 was prepared in straightforward way without any special modification, however the middle fragment 68-126 was impossible to be synthesized even upon several unsuccessful attempts. Thus we changed strategy and in order to shorter the size of the fragment we decided to use the homocysteine ligation[41] developed by the group of Tam, who exploited thioester ligation with a N-terminal homocysteine residue resulting in a native methionine residue generation upon post ligation S-methylation. Thus the original middle fragment (68-126) was further divided in two segments, 68(CysThiaz)-85Glu(OFM)-thioester and hcys(Thiaz)86-126-thioester. The Glutamic residue position 85 C^{α} thioester needed to be protected during ligation to avoid formation of undesired gamma glutamic peptide backbone[40]. The homocysteine with 1,3 amino thiol masked through heterocyclic ring (like thiazolidine for the cysteine) is not commercially available. Thus, the protection of the 1,3 aminothiol through heterocyclic ring formation occurred after HF cleavage in TFA buffer as described by Villain et al.[53]

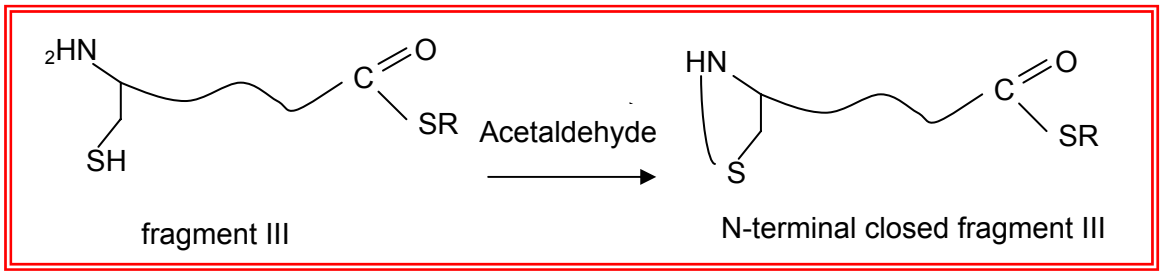


Fig.**3.2** IL-18 Thiazolidine formation

All four fragments according to our new strategy were finally successfully synthesized and purified. At the time being, the ligation steps to assembly the protein are still in progress. N-terminal protected fragment III was then ligated with fragment IV forming fragment III+IV. Trial ligation occurred almost quantitavely by HPLC measurement, however a des-Asn (Asparagine deletion), impurity carried over from chain assembly of fragment III is still present after ligation (Fig **3.13**).

Characterization of IL-18 Fragments

Fragment **I**, IL 18 N-terminal 1-67-CαSR, expected M+H 8157.4, found ES MS 8157, Maldi MS 8160.5
Fragment **II**, IL 18 Cys(Thiaz)68-85-E(OFm)- CαSR, expected M+H 2448, found ES MS 2449, Maldi MS 2450
Fragment **III**, IL 18 hcys(PG)86-126-CαSR, expected M+H 5118, found ES MS 5119
Fragment **IV**, IL 18 C terminal127-157, expected M+H 3743, found ES MS 3743, Maldi MS 3742

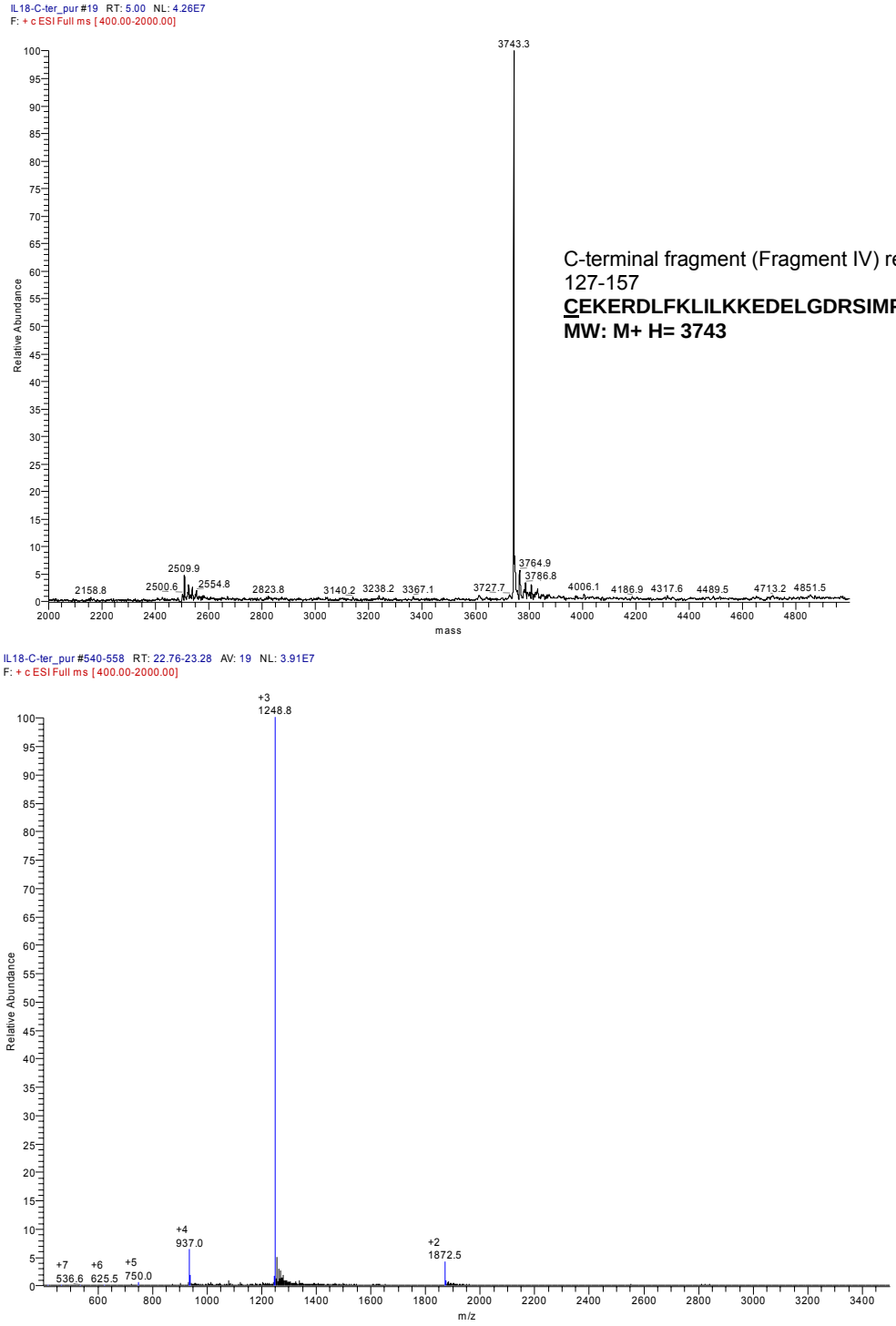


Fig.**3.3** IL-18 C Terminal Frangment (**IV**): MS analysis of purified fragment

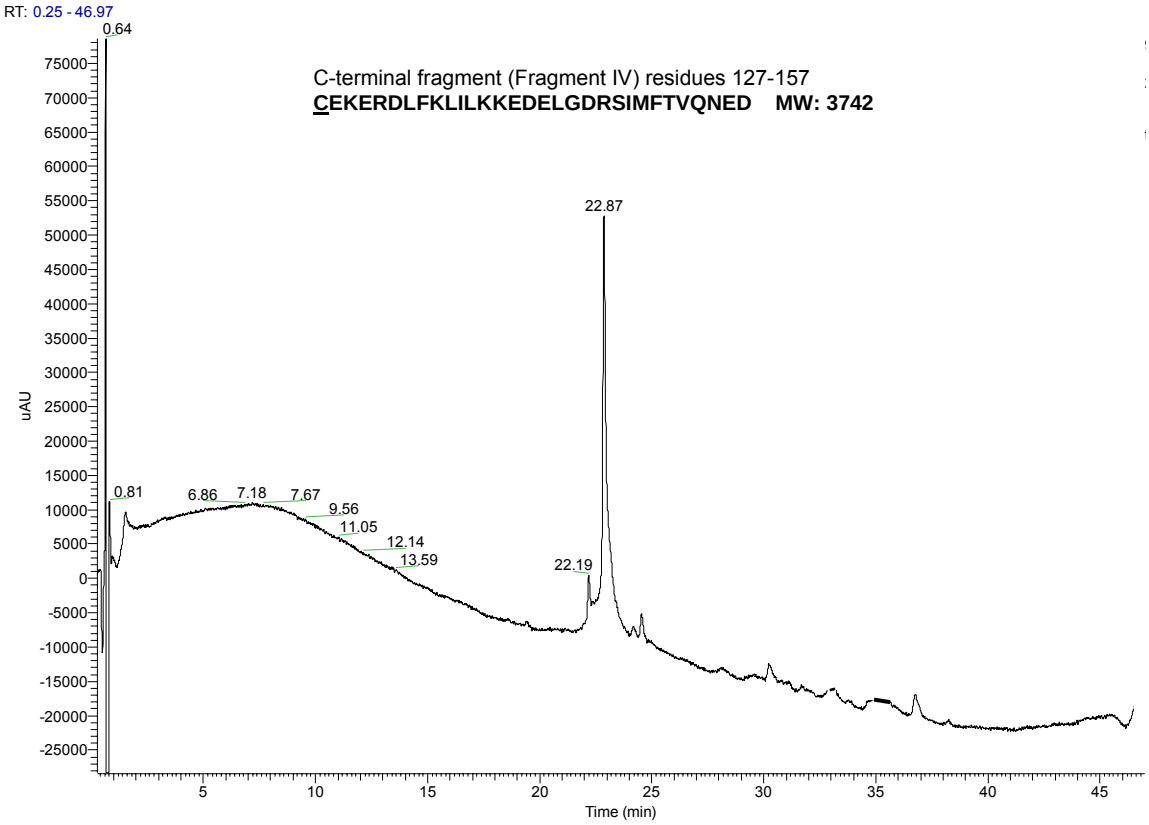


Fig.**3.4** IL-18 C Terminal fragment: HPLC analysis

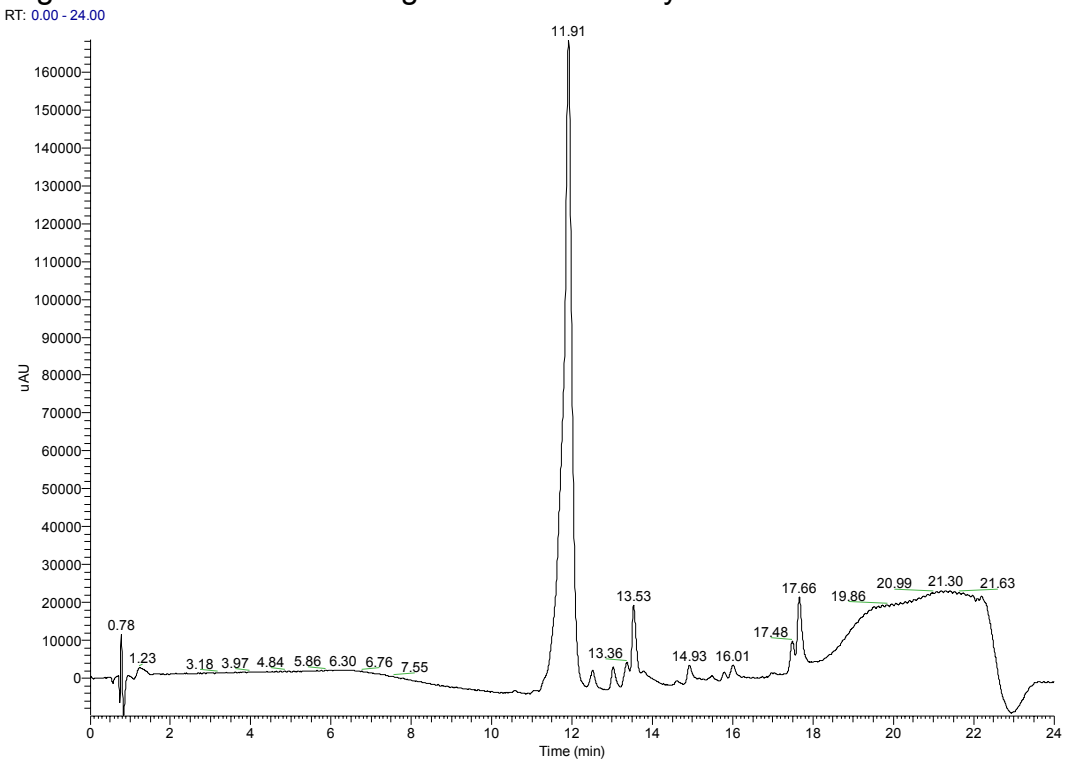


Fig.**3.5** HPLC analysis of IL-18 Fragment **III**

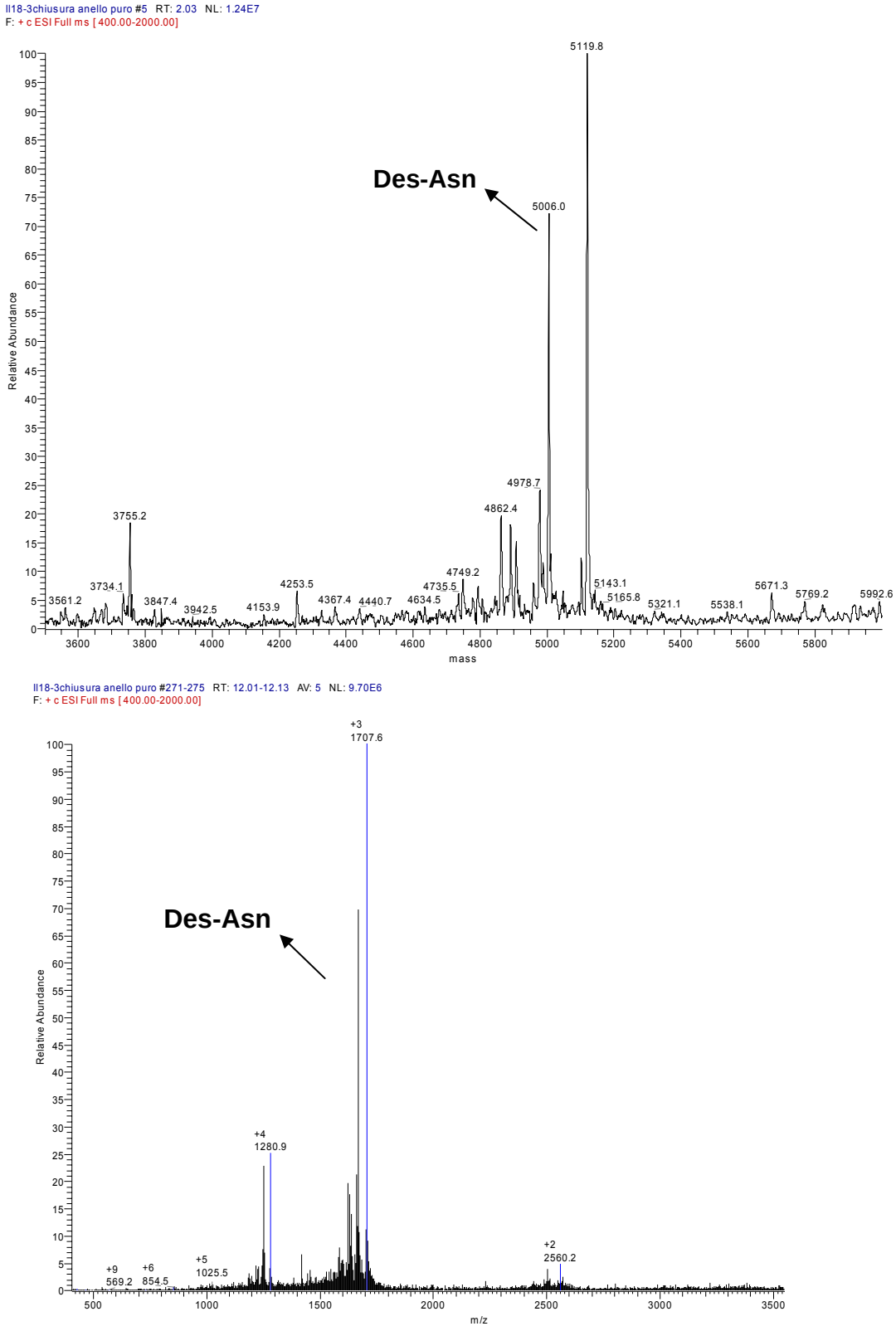


Fig.**3.6** IL-18 Fragment **III** MS Analysis. Note impurity present due to uncompleted coupling of asparagine (Asn).

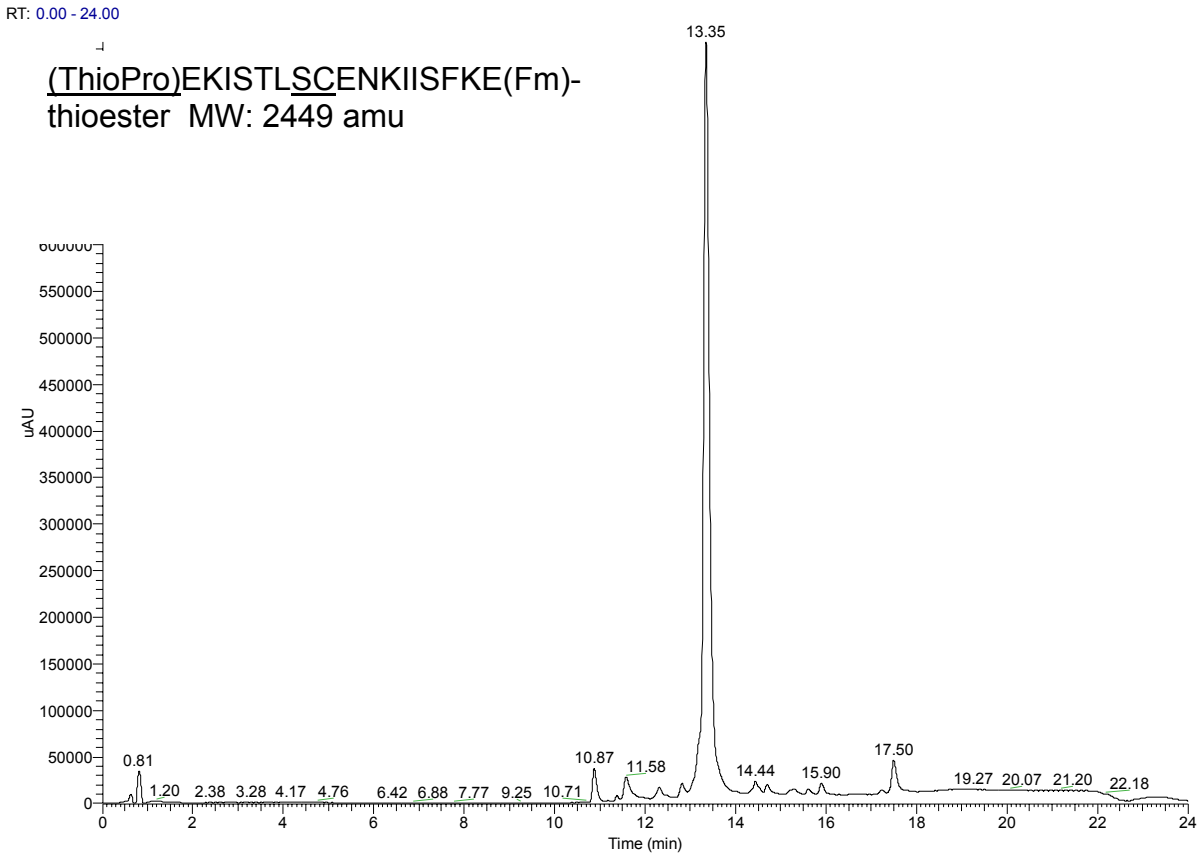


Fig.**3.7** IL-18 Fragment **II**: HPLC analysis

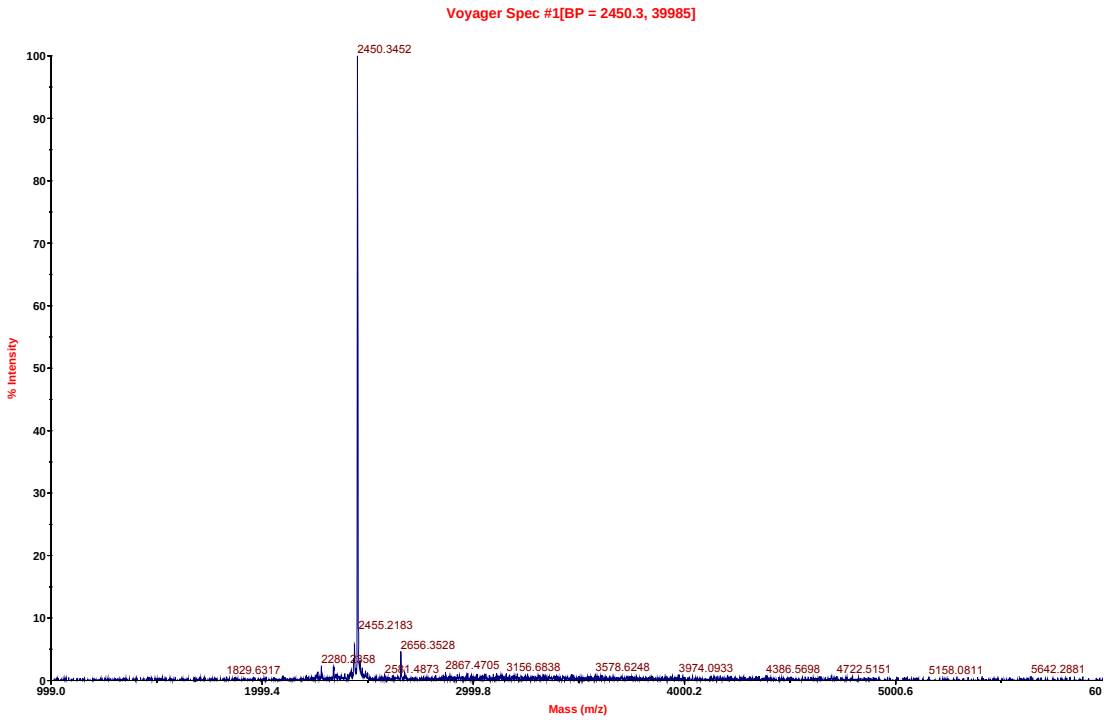


Fig.**3.8** IL-18 Fragment **II**:Maldi MS analysis

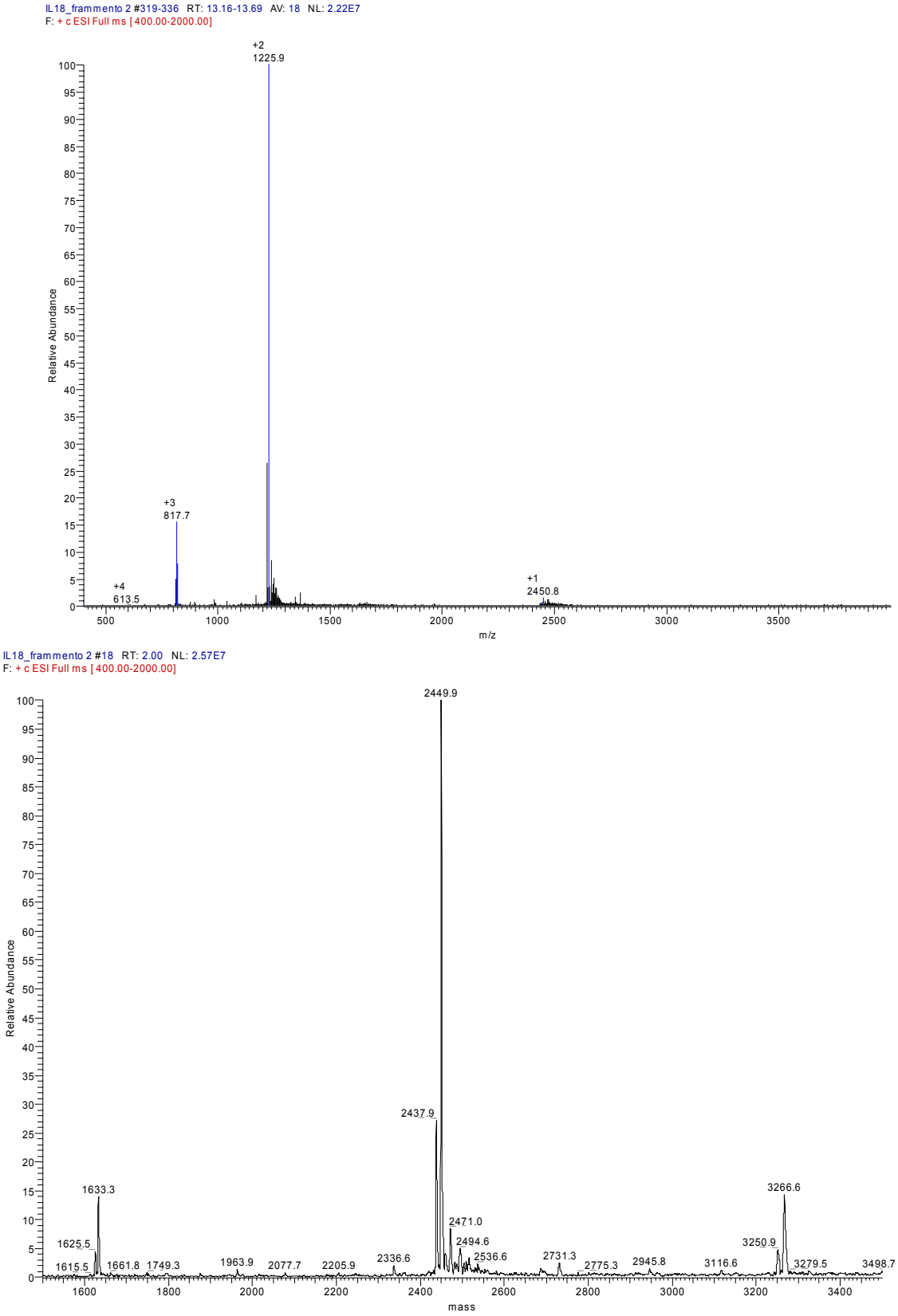


Fig.**3.9** IL-18 Fragment **II**: MS analysis

YFGK LESKLSVIRN LNDQVLFIDQ GNRPLFEDMT DSDCRDNAPR TIFIISMYKD SQPRGMAVTI SVK- CαSR **MW: 8157 amu**

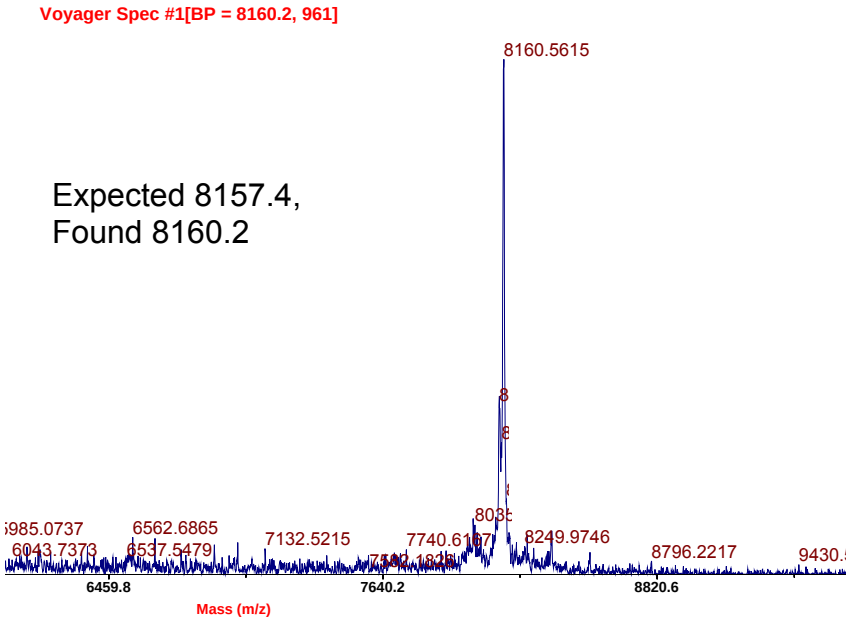


Fig.**3.10** IL-18 N Terminal (Fragment **I**) 1-67. Maldi MS analysis

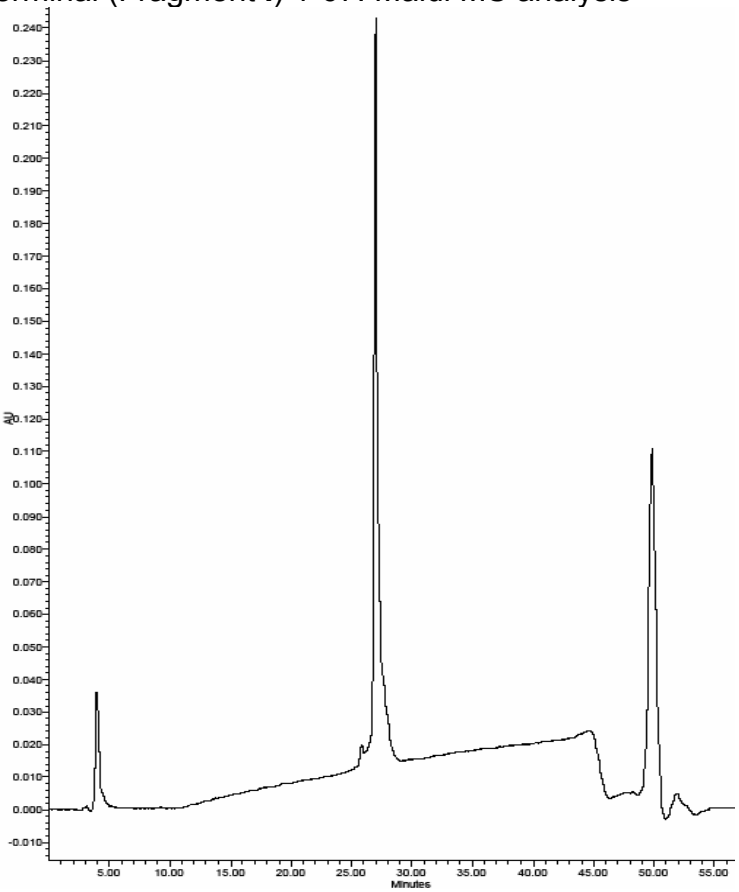


Fig.**3.11**IL-18 N Terminal (Fragment **I**) 1-67. HPLC analyis,

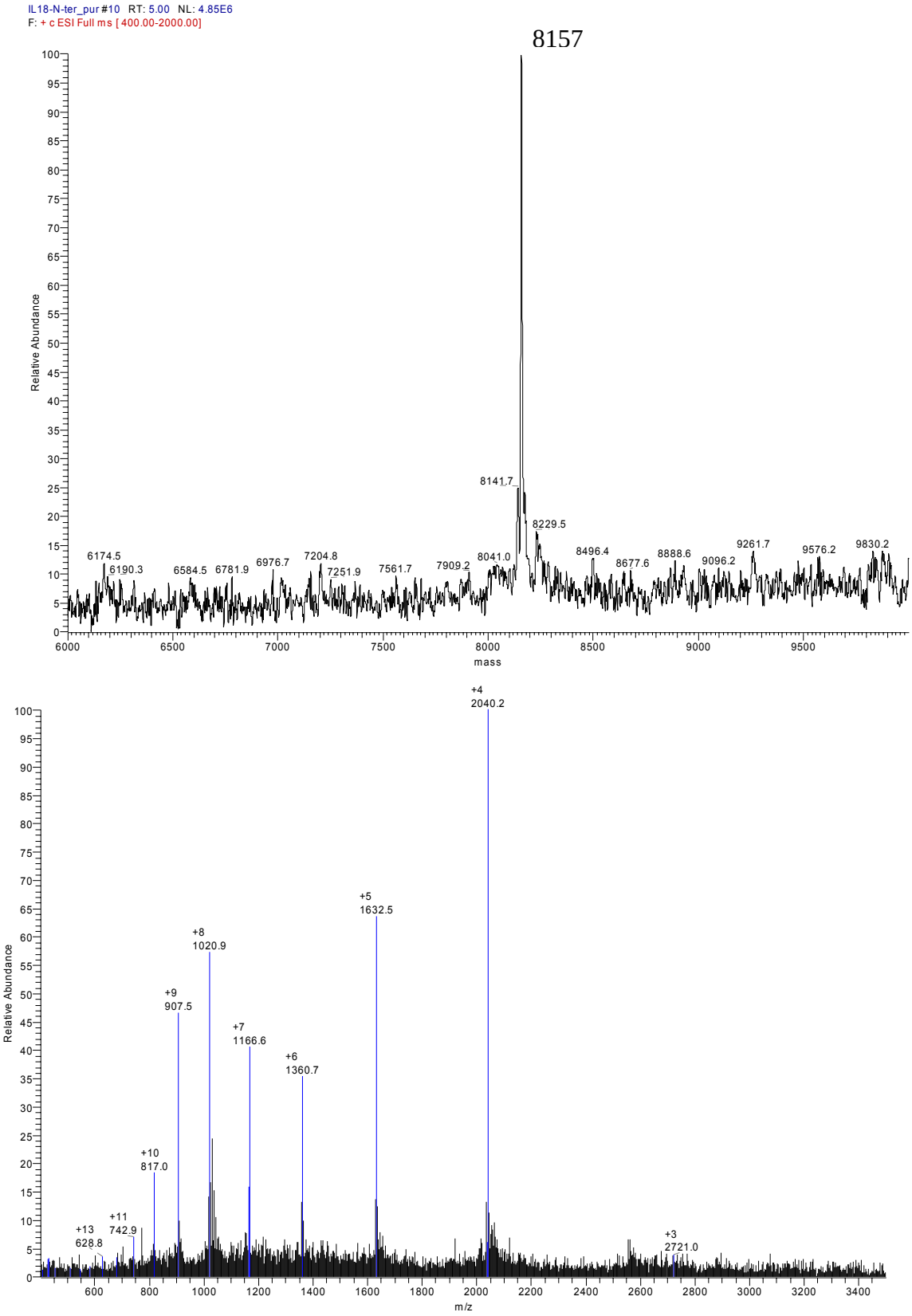


Fig.**3.12** IL-18 N Terminal (Fragment **I**) 1-67. MS analysis

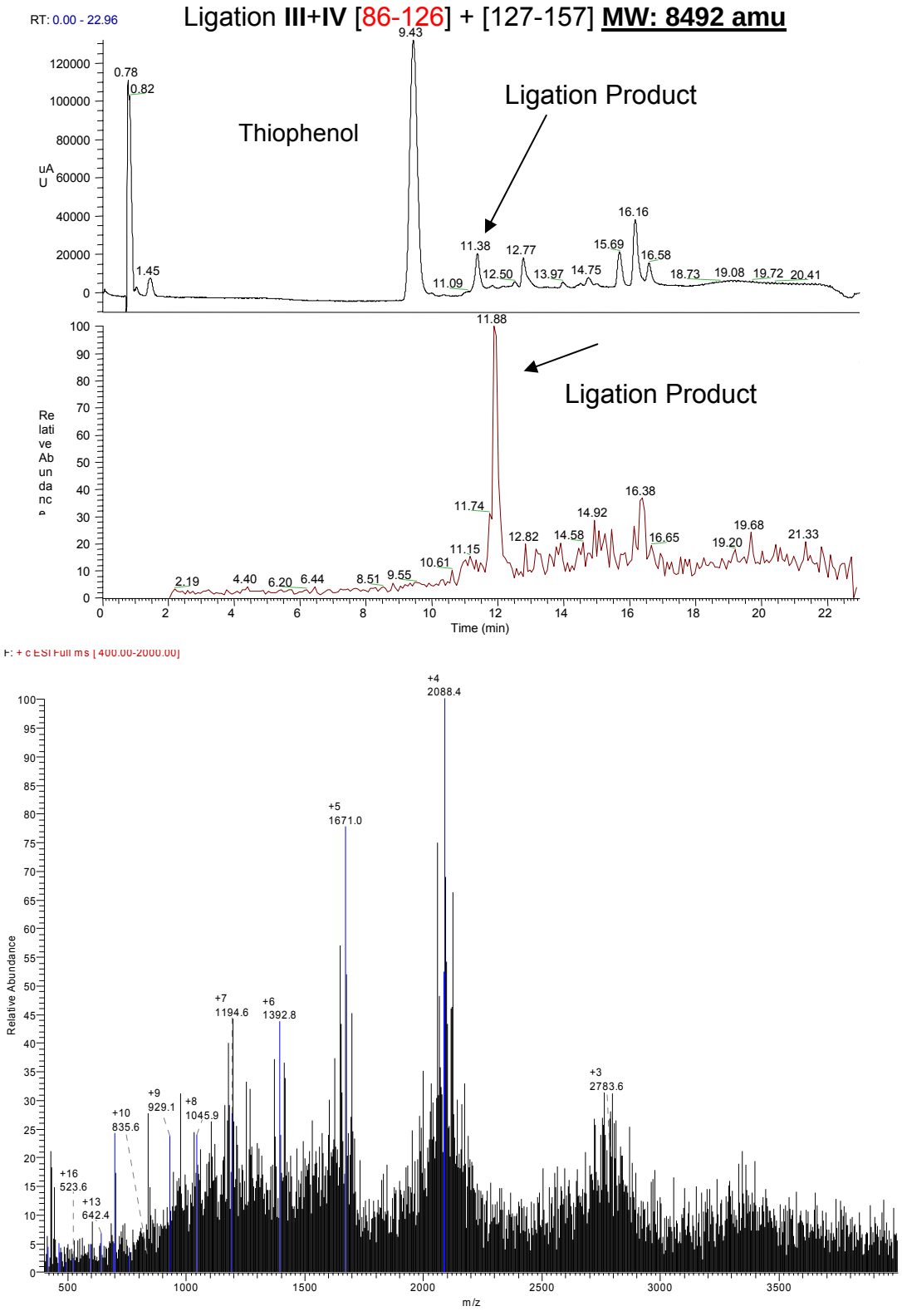


Fig.**3.13** IL-18 Ms Product Ligation Fragment **III and IV**

Ligation Strategy to assembly IL-18

Ligation product from fragments **III** and **IV** will be first protected at the cys 127 (ligation site) with Acm group under strong acidic conditions as described by the work of Albericio[54]. Similarly, Fragment **II** was protected at the cys 76 with Acm (Fig. 3.13 below). Then after selective deprotection of the 1,3 amino thiol of the N terminal hcys 86-157, the N-terminal free hcys fragment can undergo to ligation with Cα-thioester Fragment **II** cys(Acm) 76 protected**.**

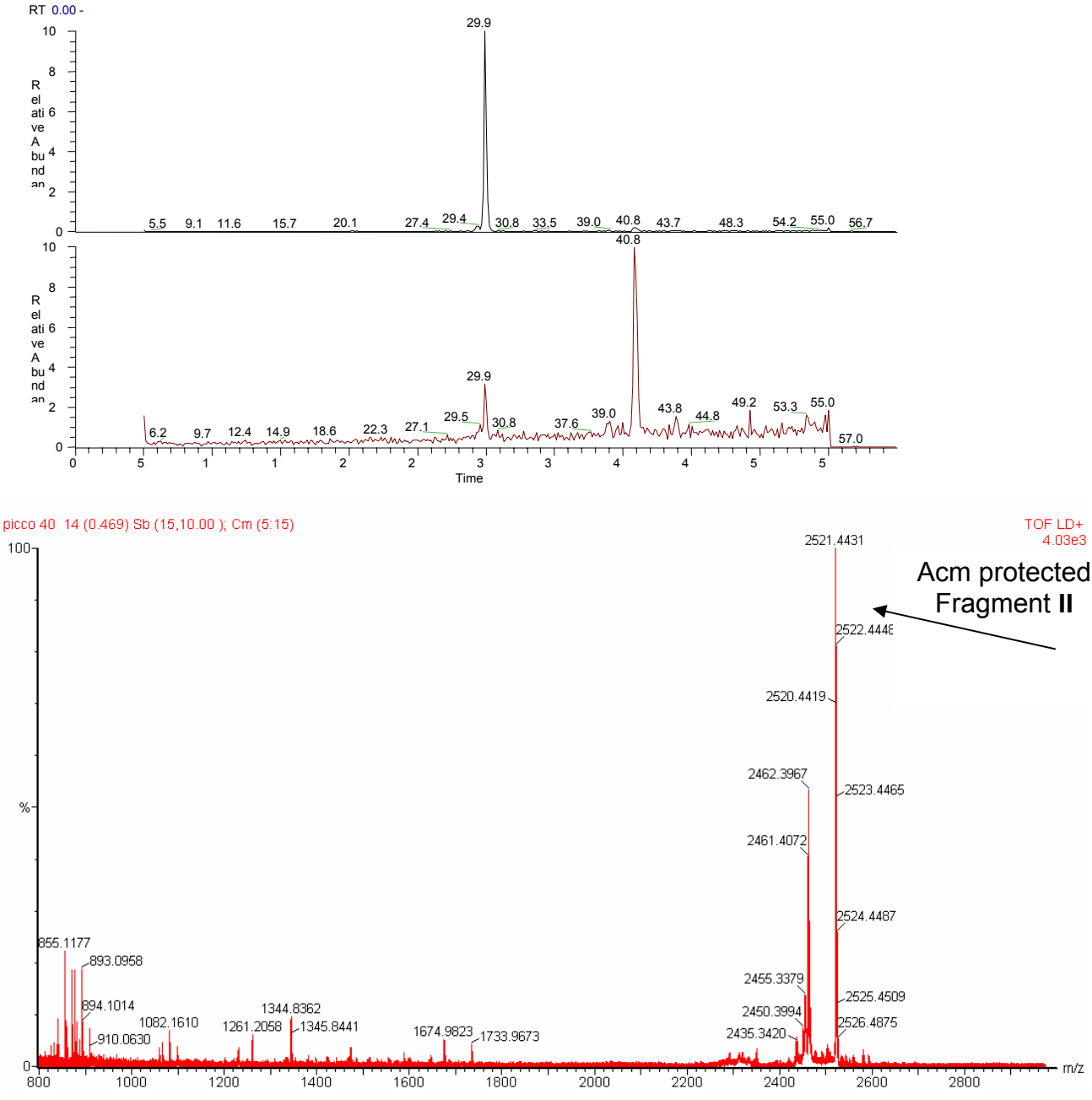
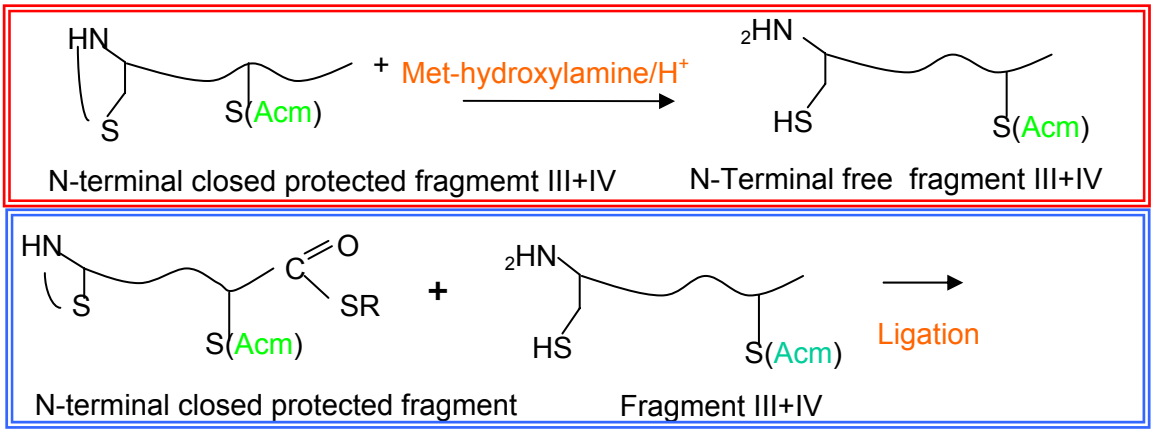
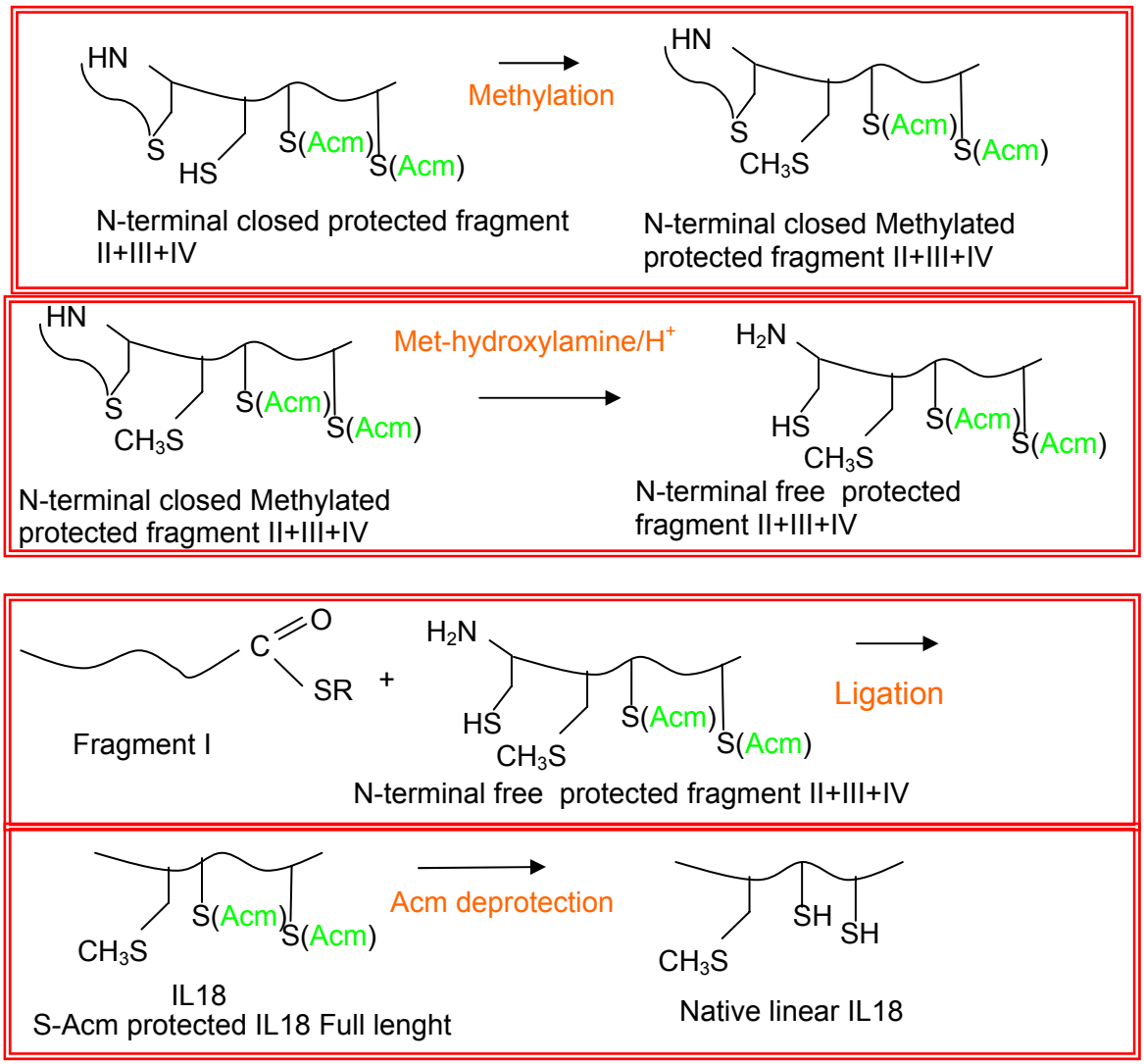


Fig.**3.14** IL-18 Fragment **II** Acm protection



Under this strategy, the ligation product of segments **II** +**III** + **IV** will own a linear polypeptide with only one free thiol moiety, which upon methylation results in a native fragment with methionine 86 at the ligation site. Finally, upon deprotection of the 1,2 amino thiol of the cys 68, the N-terminal free cys polypeptide 68-157 is amenable to react with the thioester fragment **I** (1-67) to yield, after cys Acm deprotection, the linear full length native chemically assembled IL-18 1-157.



Cartoons describing the IL-18 synthesis development

3.2 IL-32

At certain point during the development of our study the chemical synthesis of h-IL-32 became a priority due to its relevance in inflammatory and autoimmune diseases. Although several of immune regulatory molecules have been discovered during the past decade, many phenomena remain unexplained in understanding immune regulation. The functional portfolio of the cytokine IL 32 can now help unravel the complexity of cytokine biology. In spite of many advances in the recombinant protein expression, the difficulties encountered in the expression of the protein variants concerning both the quantity and purity of the obtained material (absence of toxins and other biological contamination such as for example peptido-glycans) rendered the chemical synthesis of the polypeptide truly necessary in order to unambiguously confirm the scientific data so far obtained by C. Dinarello and co-workers. Paolo, you can insert the references here). Indeed, should the polypeptide chemically synthesized (and thus devoid of toxins and other biological contamination) confirming the previous results obtained with the recombinant material, that would finally prove the existence of new inflammatory and immune pathways. We began our work by synthesizing the N-terminal γ (gamma) isoform consisting of 103 residues. Indeed, studies from C.Dinarello and co-workers indicated it as the shortest polypeptide that should own biological activity comparable to the entire functional protein. This was based on the cleavage site of PR3[55]. Assuming the polypeptide in matter could then show the desired or at least significant biological activity, in a second phase of the project the synthesis of the whole protein could be tackled. However, having a synthetic peptide without any contamination will be of great value in finding a specific receptor. The receptor can be a surface receptor on an intracellular receptor.
Our initial strategy was to synthesize the polypeptide by Native Chemical Ligation assembly of only three peptide segments. However the middle fragment, when synthesized the first time, showed unexpected solubility problems. Thus it was further divided in two segments: a short fragment of 3 residues (Fragment **VII**) 57-59 and fragment **VI,** 36-56 thus making a total of four fragments. The C terminal fragment **VIII,** 60-103 was synthesized two times. The other 3 fragments were synthesized without any changes from standard Boc chemistry with In situ neutralization. Each fragment was purified by HPLC semi-preparative and each ligation reaction was first tried on analytical scale (0.5-1.0 mg scale) and then ligation was scaled up to multi-mgs scale. Finally the ligation mixture was purified by semi-preparative HPLC.
The full-length N-terminal part 1-103 of the γ (gamma) isoform was synthesized and sent with the intermediate fragments to Prof. Dinarello for biological activity tests.

Characterization of IL-32 Fragments and ligation product

MCFPKVLSDD MKKLKARMVM LLPTSAQGLG AWVSACDTKD TVGHPGPWRD
KDPALWCQLC LSSQHQAIER FYDKMQNAES GRGQVMSSLA ELEDDFKEGY LET
Theoretical pI/Mw: 5.26 / 11570.25

IL-32 C-terminal 60-103 Fragment **VIII** =Expected 5042.5, found 5043
IL-32 57-59 thioester, Fragment **VII** MW_{th}=561
IL-32 36-56 thioester, Fragment **VI** MW_{th}= 2760
IL-32 1-35 thioester, Fragment **V** MW_{th}= 4011

Ligation Products :

IL 32 57-103 (Ligation fragments **VII** +**VIII**) expected, found 5398; (expected after thiazolidine ring opening 5386., found 5385);
IL 32 36-103 (Ligation fragments **VI** + C-terminal 57-103), , expected 7775, found 7775; (7763.5 expected after thiazolidine ring opening, found 7763.8);
IL-32 1-103, Ligation fragment **V** with 36-103, IL-32 gamma 1-103, expected, 11570.2, found, 11570; After folding, expected 11566, found: 11568

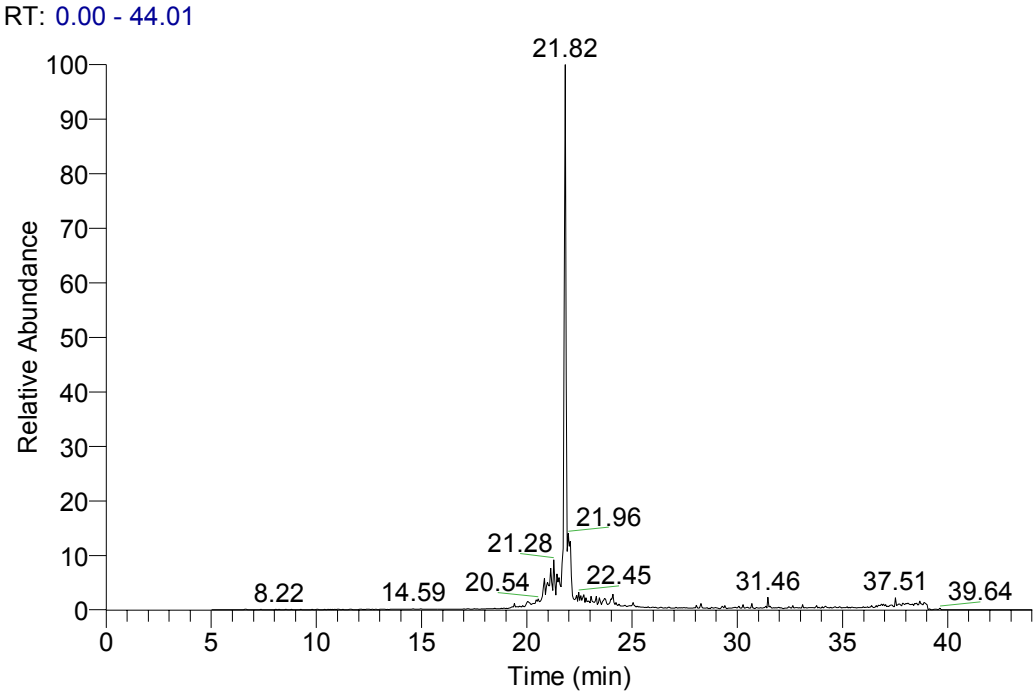
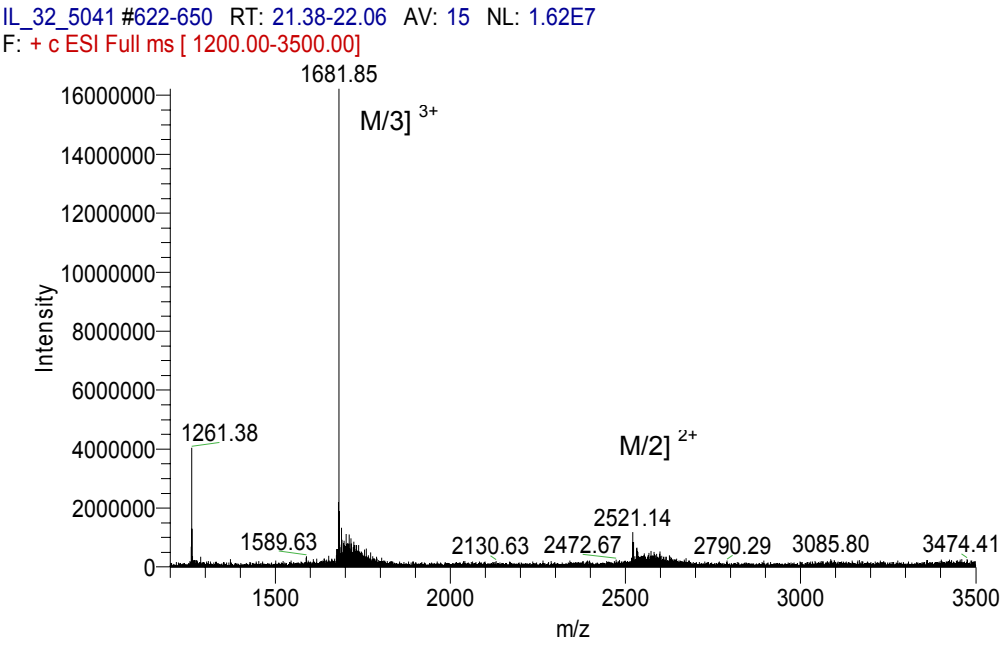


Fig.**3.15** IL-32 Fragment **VIII,** HPLC analysis**.**



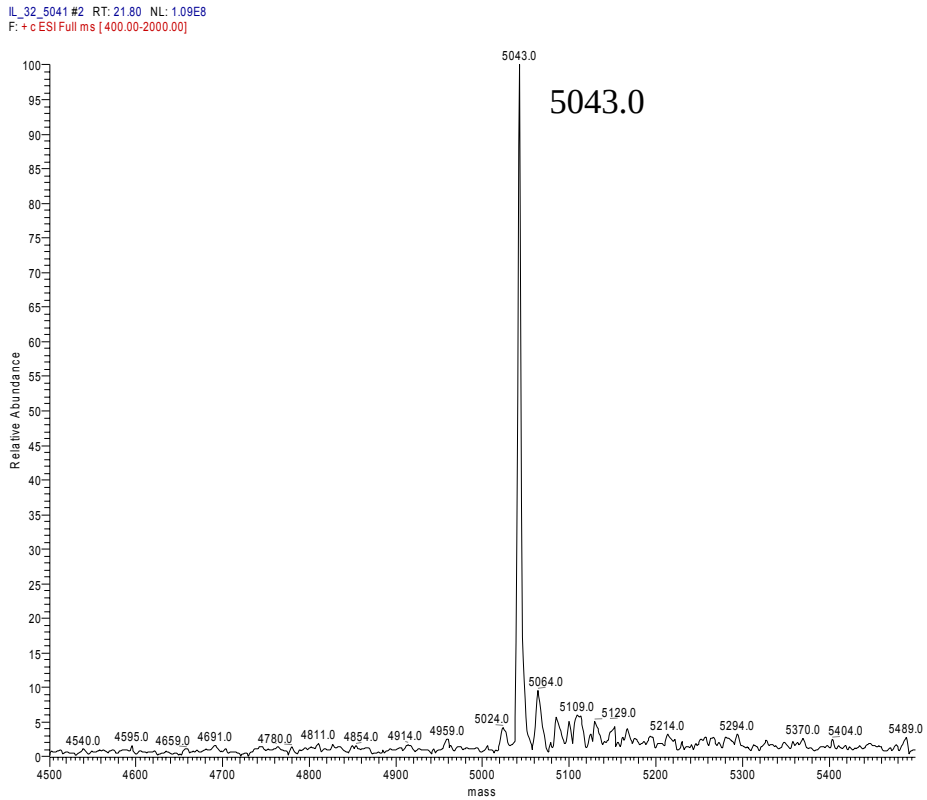
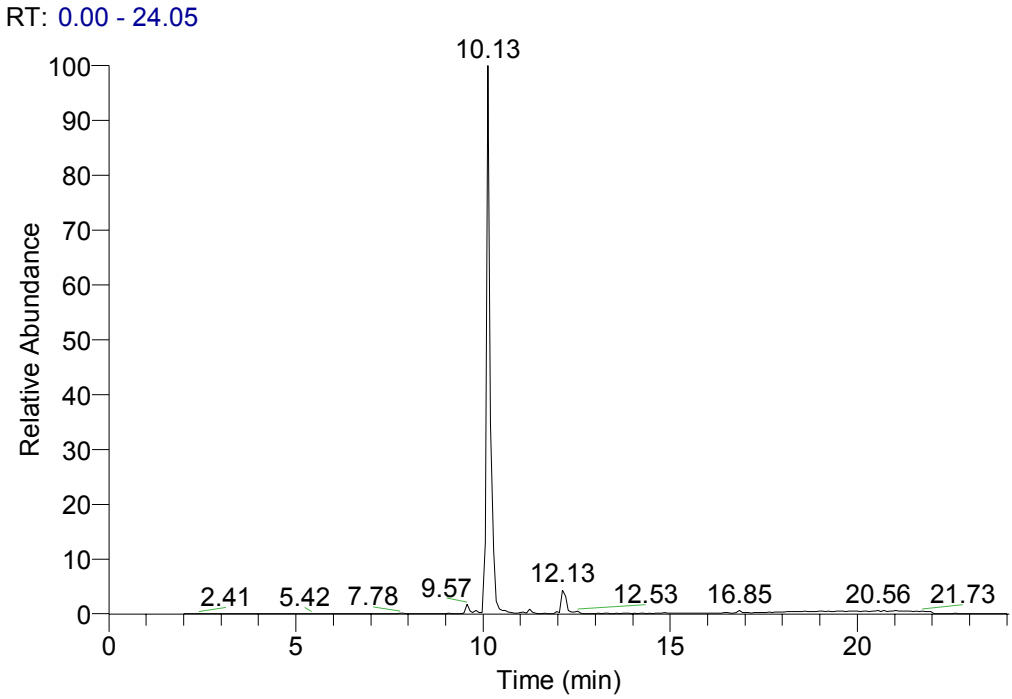


Fig.**3.16** IL-32 Fragment **VIII,** MS analysis**.**



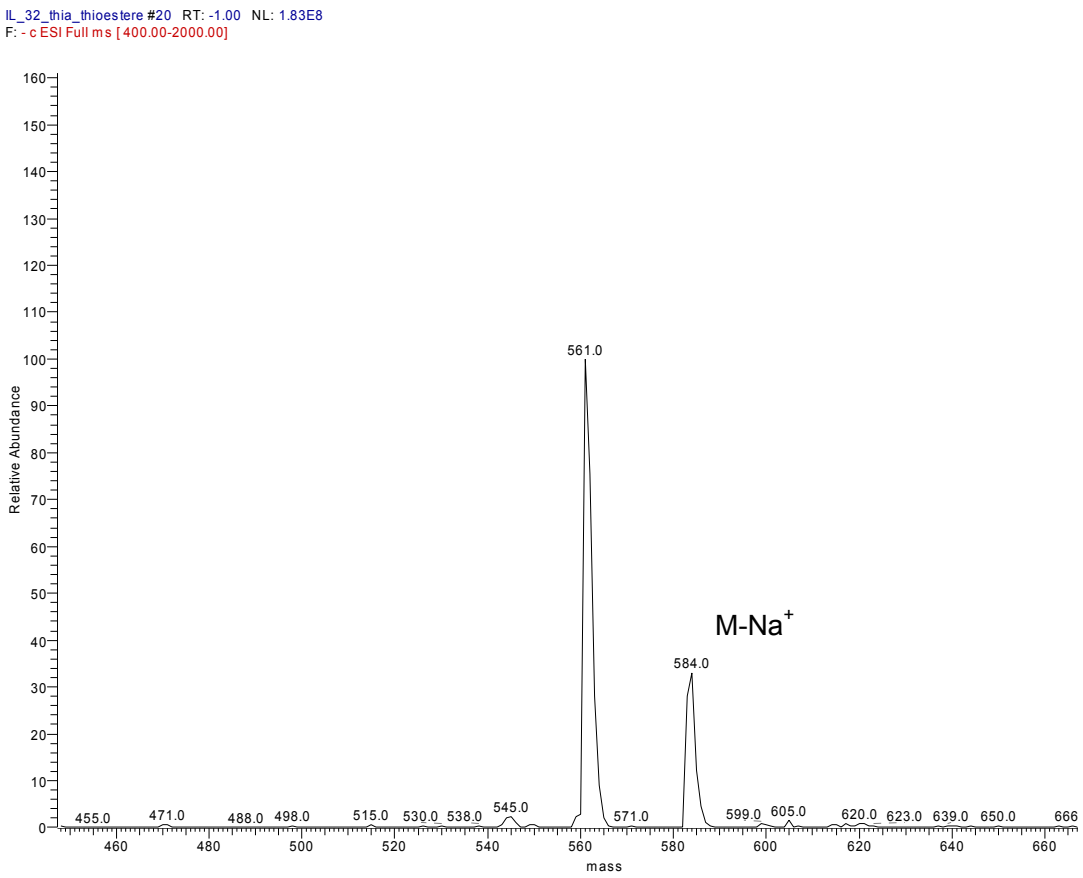


Fig.**3.17** IL-32 Fragment **VII,** HPLC and MS analyses**.**

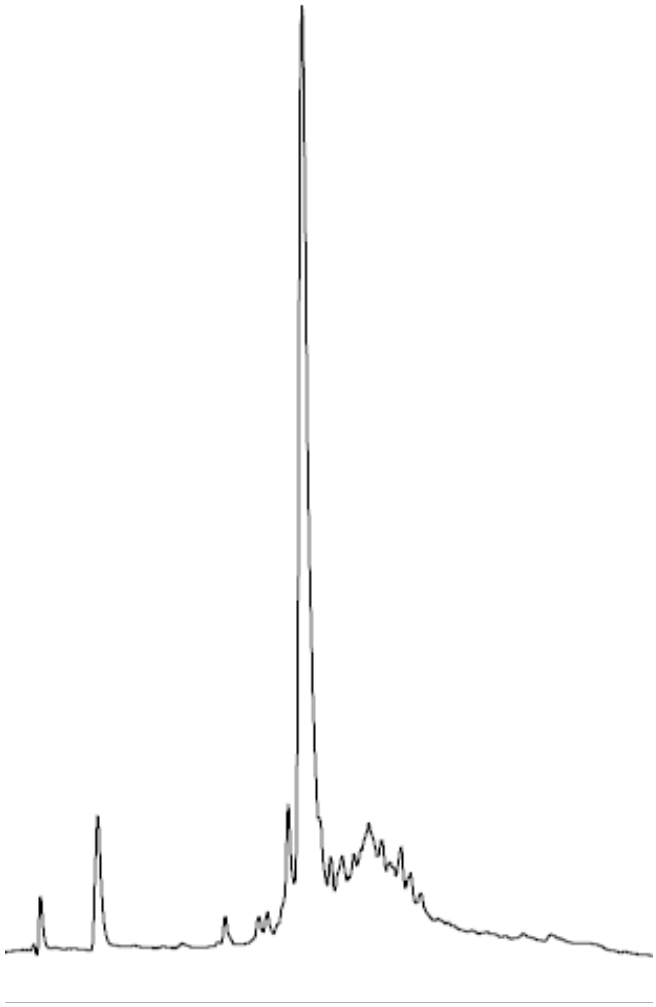
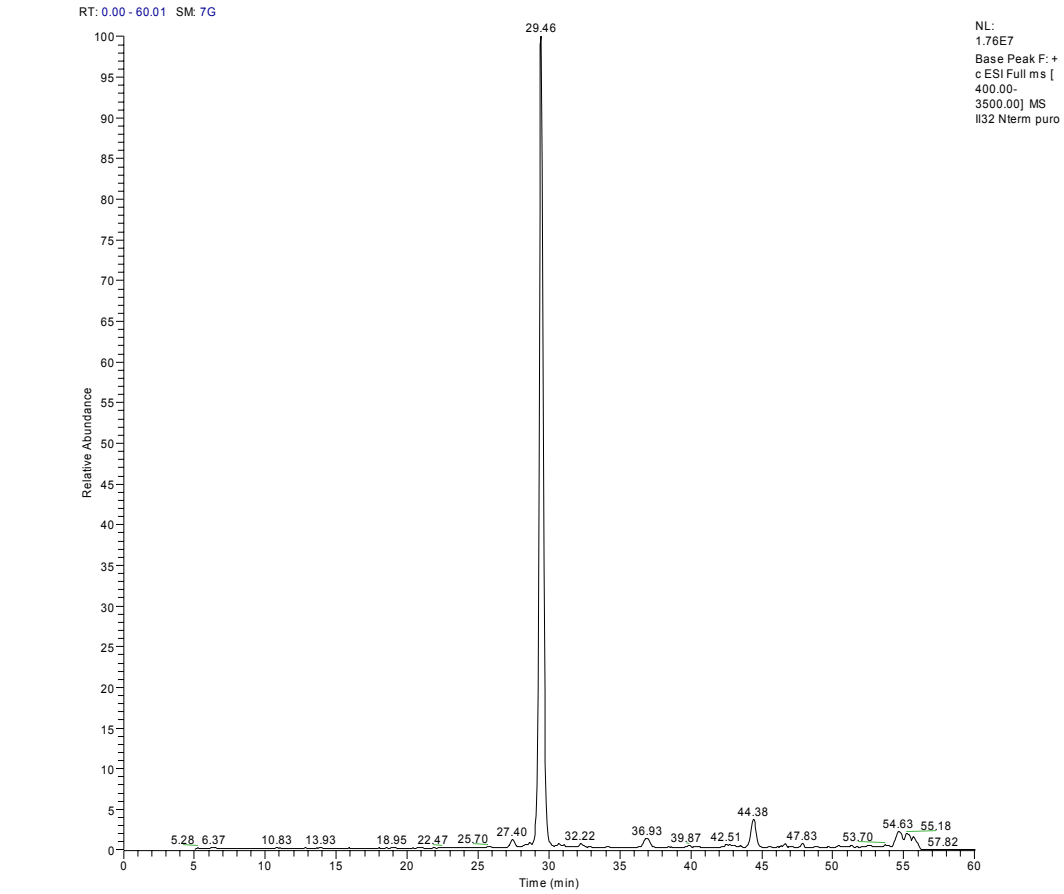
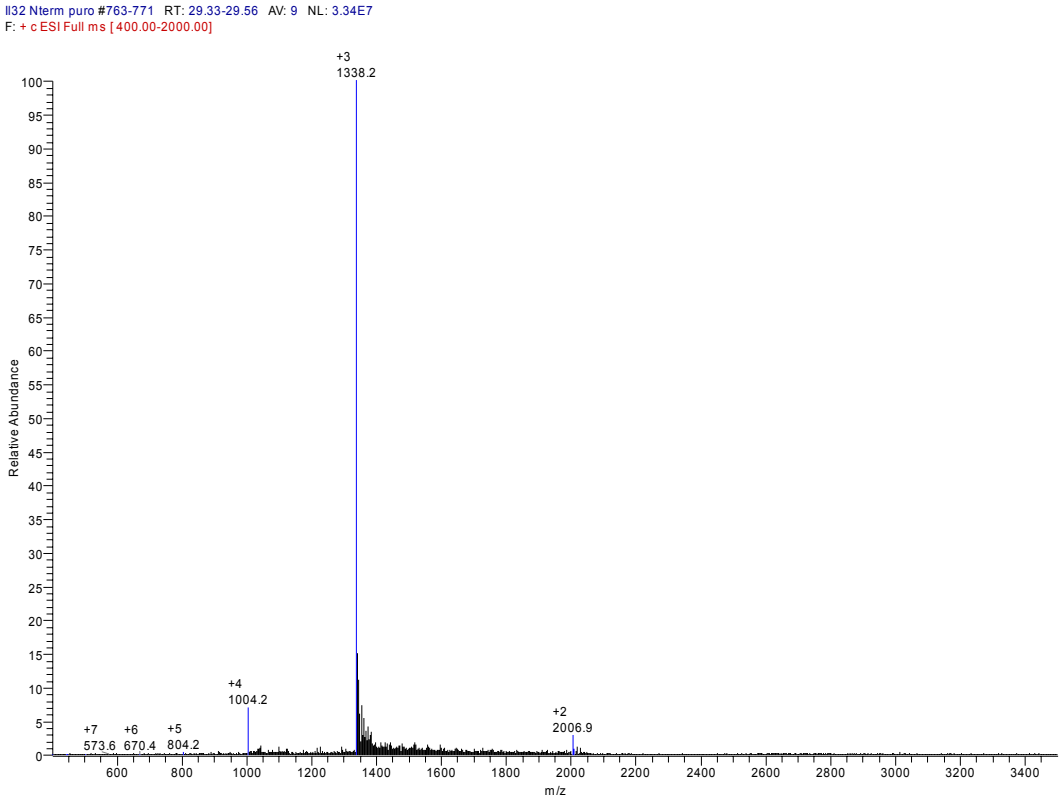


Fig.**3.18** IL-32 Fragment **VI,** HPLC of Crude


RT: 0.00 - 60.01 SM: 7G
NL:
1.76E7
Base Peak F: +
c ESI Full ms [
400.00-
3500.00] MS
II32 Nterm puro
Relative Abundance
Time (min)
29.46
44.38
5.28
6.37
10.83
13.93
18.95
22.47
25.70
27.40
32.22
36.93
39.87
42.51
47.83
53.70
54.63
55.18
57.82


II32 Nterm puro #763-771 RT: 29.33-29.56 AV: 9 NL: 3.34E7
F: + c ESI Full ms [400.00-2000.00]
Relative Abundance
m/z
+3
1338.2
+4
1004.2
+2
2006.9
+7
573.6
+6
670.4
+5
804.2

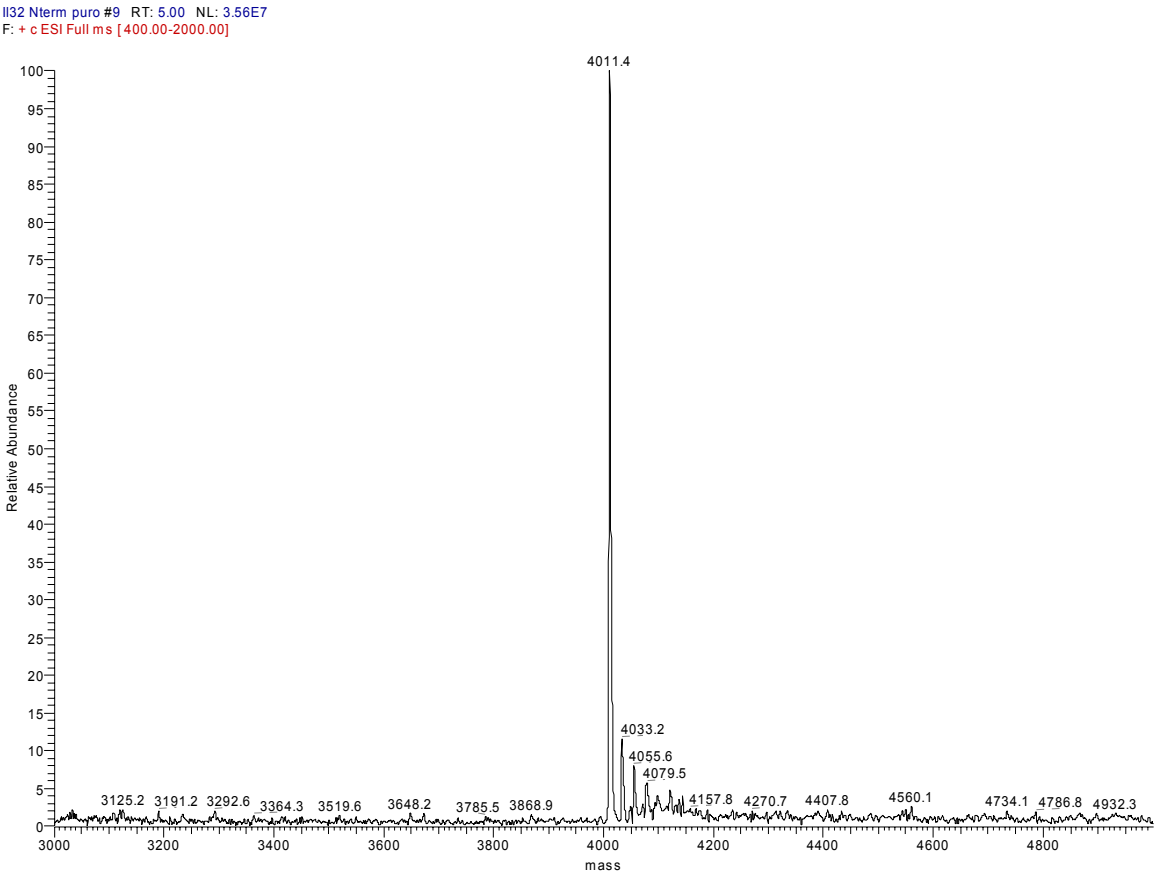


Fig.**3.19** IL-32 N-Terminal Fragment **V**: HPLC and MS:

IL-32 Ligation Strategy

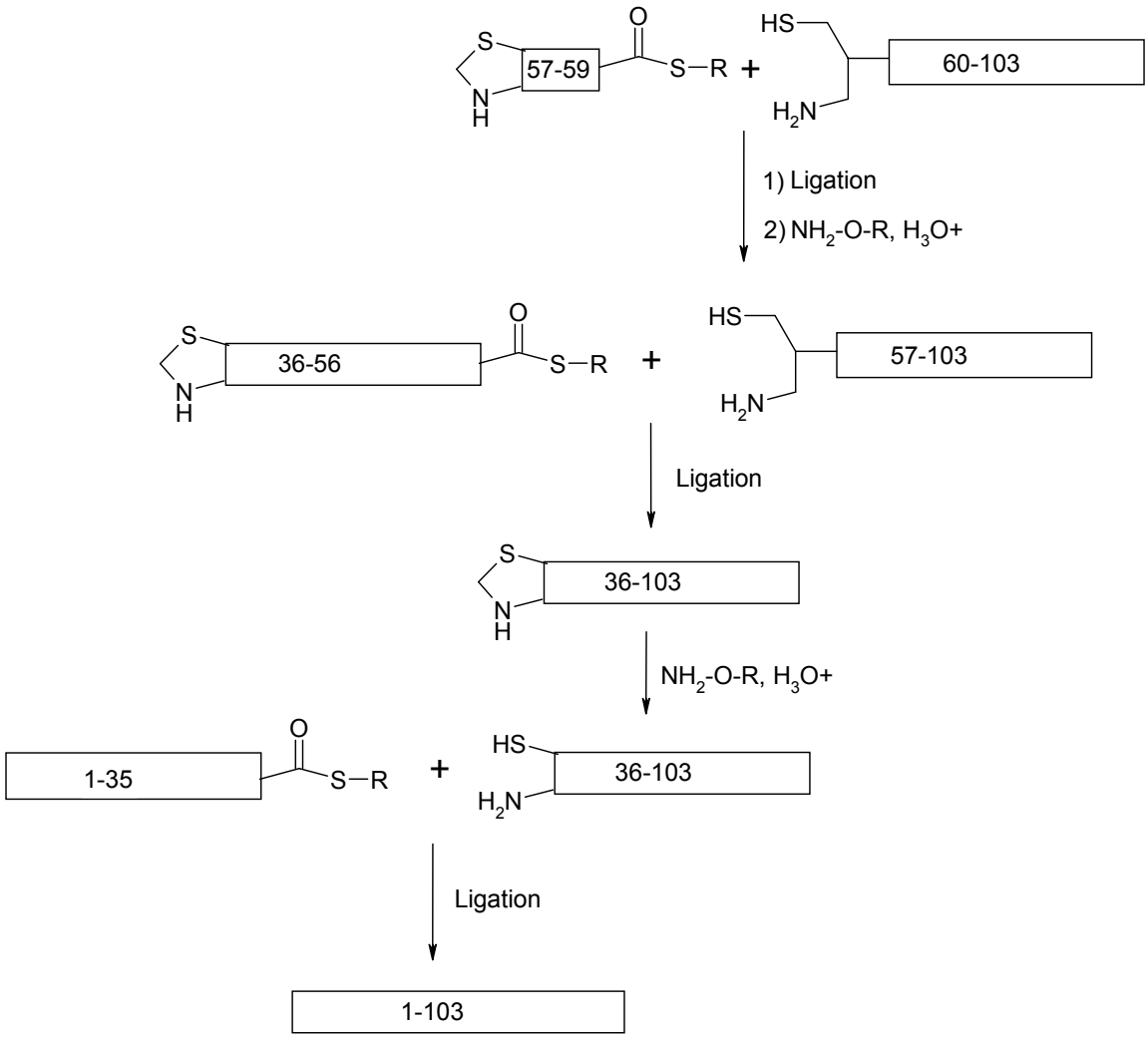


Fig.**3-20**. IL-32 Synthetic Strategy

IL-32 synthesis through ligation reactions

Ligation IL-32 57-59+60-103, MW_{th}=5398 amu

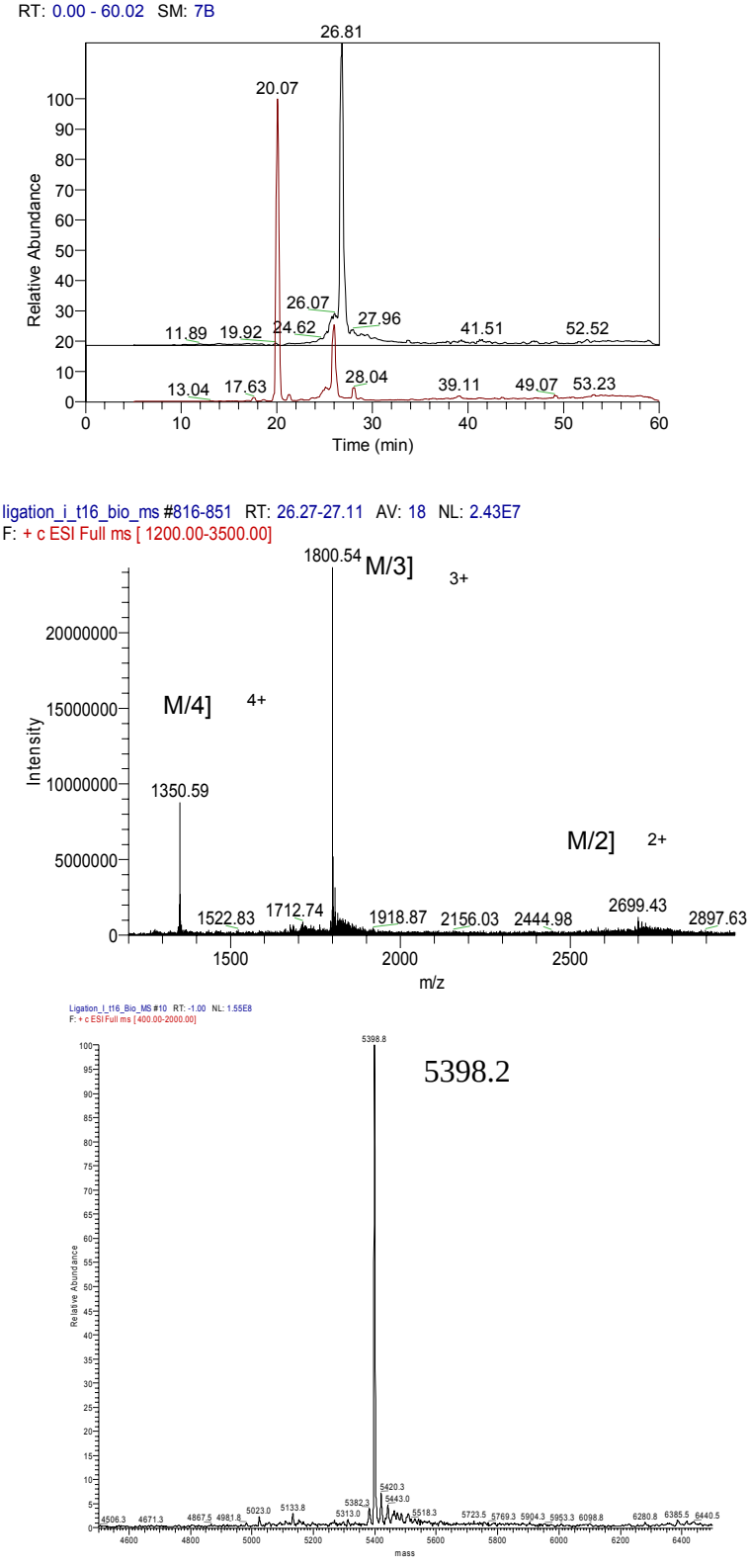


Fig.**3.21** IL-32 Ligation Fragments **VII** and **VIII**

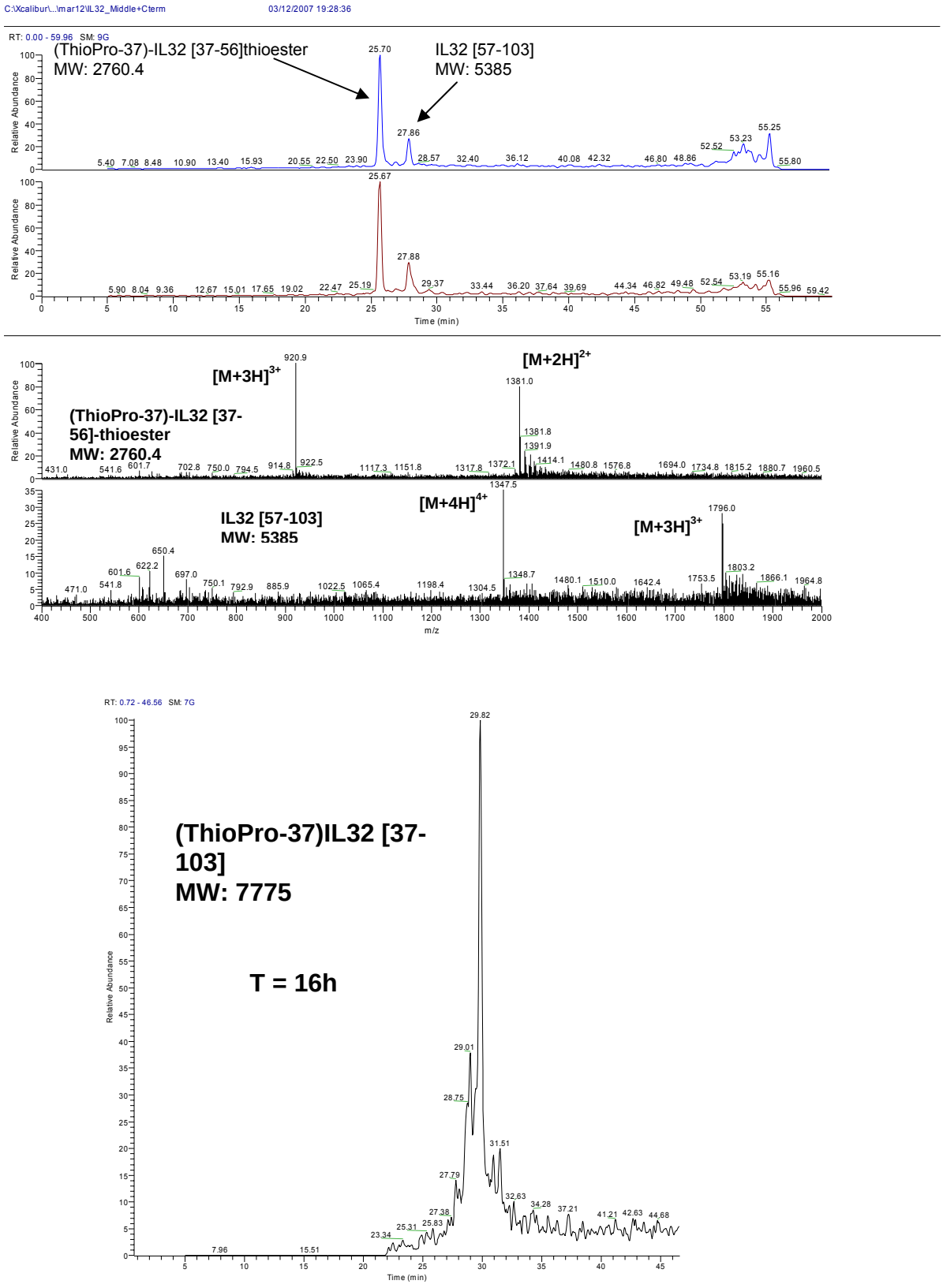
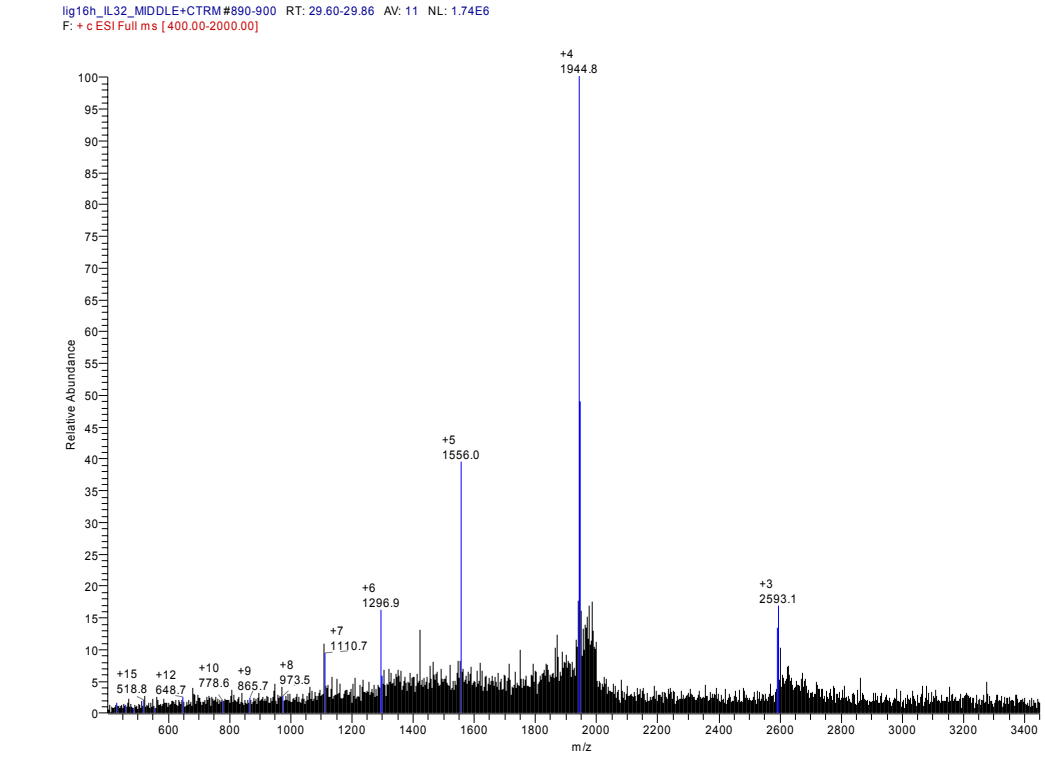


Fig. **3.22**. Ligation Middle Fragment **VI**+C-term Fragment 57-103



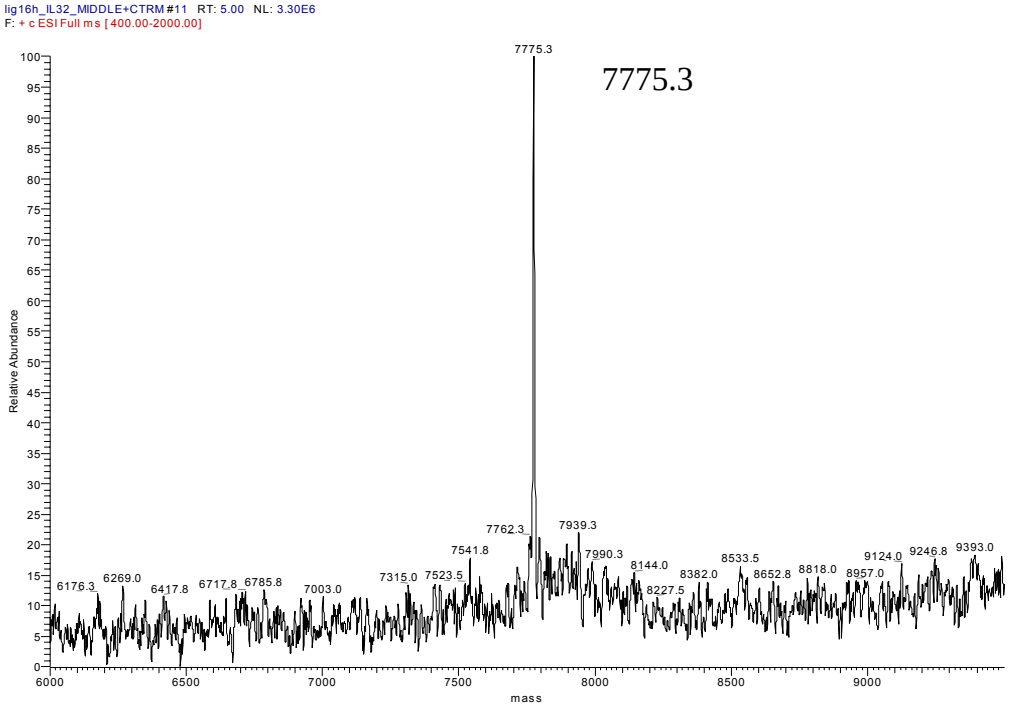
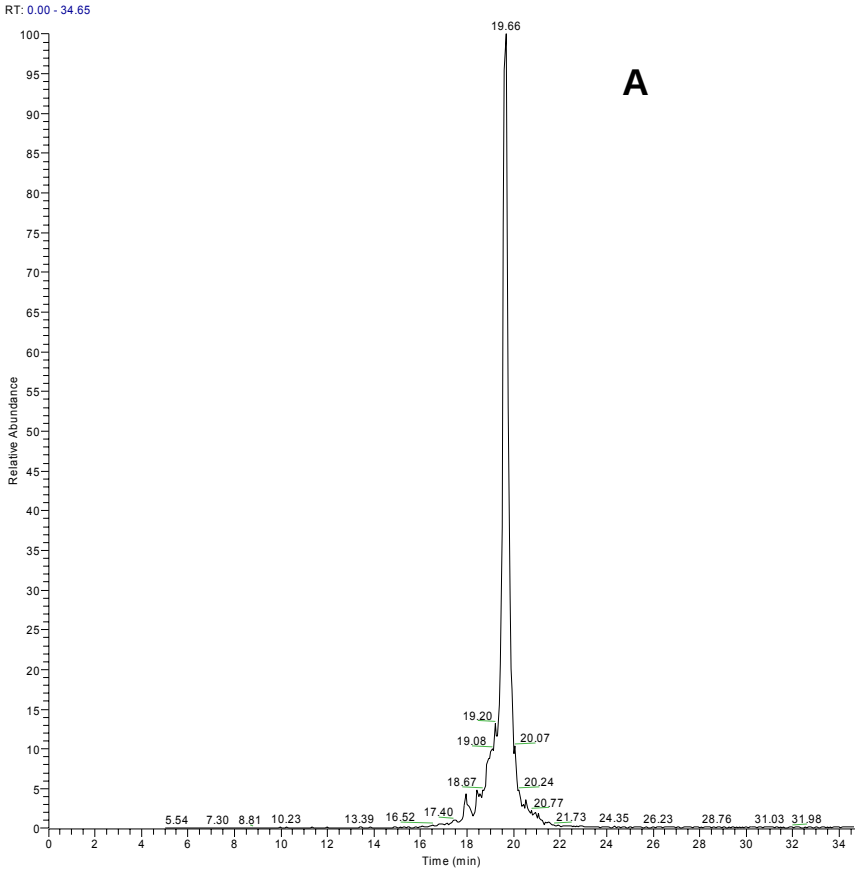


Fig.**3.23** MS IL-32 Ligation Fragments **VII** and C-term Fragment 57-103

Ligation to form IL-32 36-103. Il-32 C terminal 57-103, 10mg was reacted Il 32 Fragment **VI** (middle segment) 5.7 mg, at pH 7.00 in presence of thiophenol as catalyst. Ligation product was purified on HPLC semi-preparative on a C4 column using a gradient 10% b to 60% B in 60 minutes. After purification we recovered 9.36 mg, corresponding to a yield of 65%.
The ligation product was then treated with O-methyl hydroxylamine (0.5M) in acid condition to open thiazolidine ring to free the n-terminal 1,2 amino thiol.



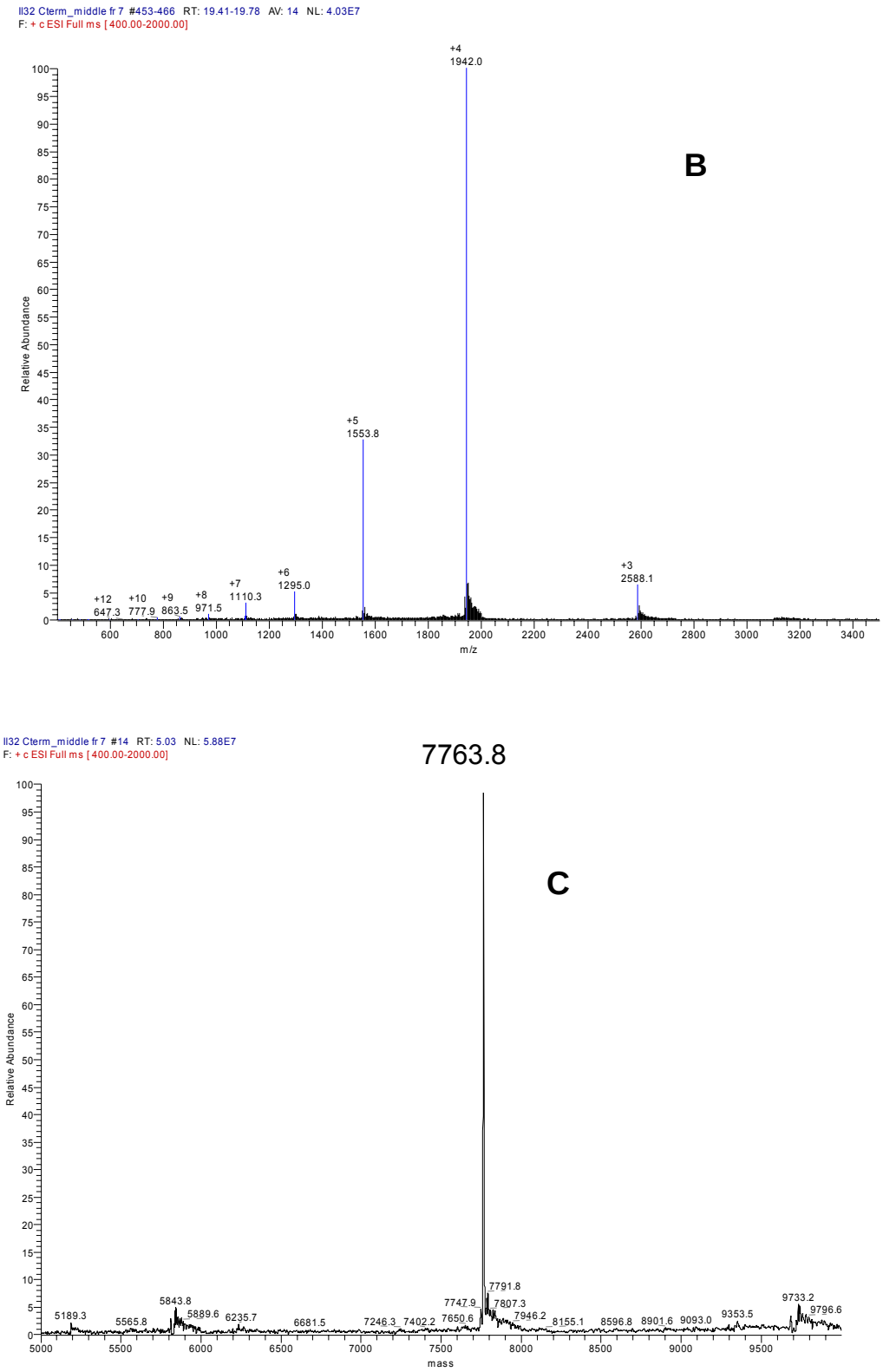


Fig.**3.24** IL-32 ligation product 36-103 after thiazolide ring opening : **A**) HPLC, **B**) Ms raw charged state product, **C**) deconvoluted MS product

Ligation to form IL-32 1-103: IL-32 Fragment **V** (1-35), 3 mg and 5 mg of purified N-terminal free cysteine 36-103 were ligated at neutral pH with thiophenol as catalyst under standard condition already reported. Ligation product was purified on HPLC

semi-preparative on a C4 column. Recovered 1.4 mg after purification, with yield of 20%

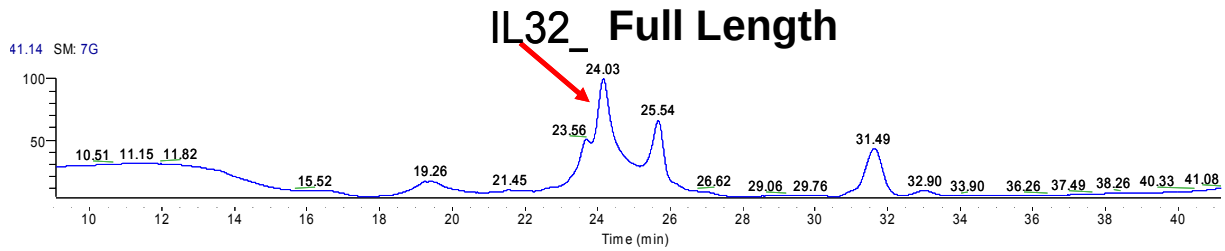
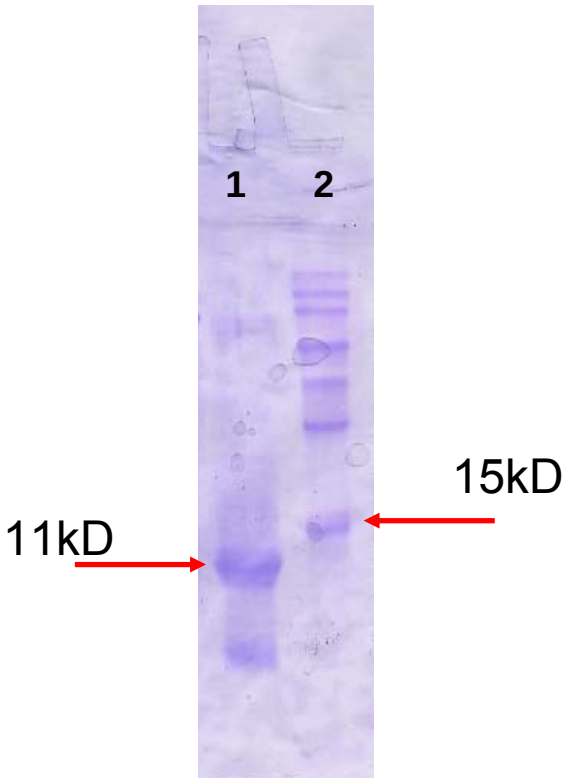


Fig.**3.25** IL-32 HPLC Ligation Fragments **V** and 36-103



1 – IL32_full_length

2 – MW marker

Fig.**3.26** IL-32 Gel of linear full length IL- 32 1-103

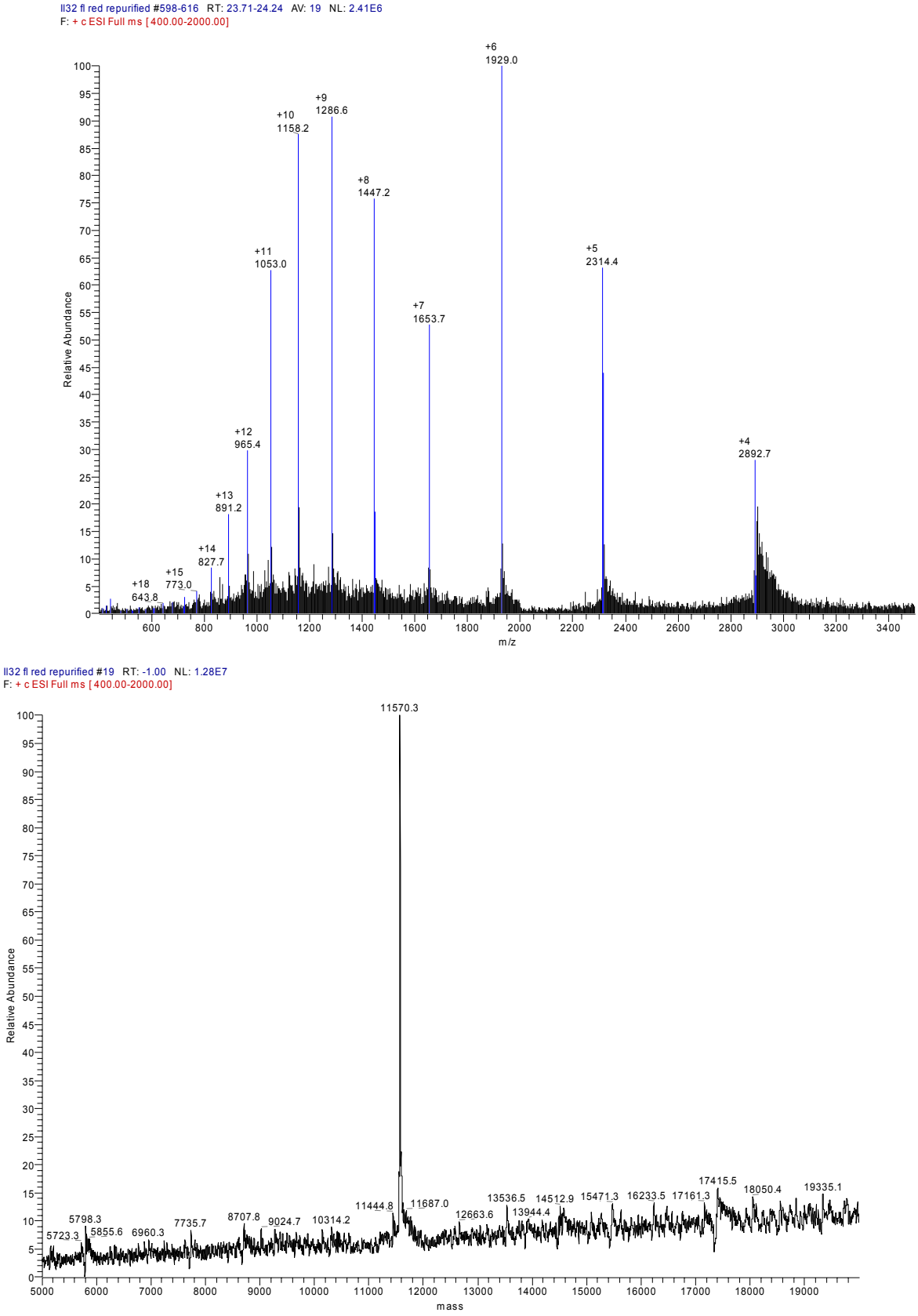


Fig.**3.27** IL-32 MS of linear IL- 32 1-103

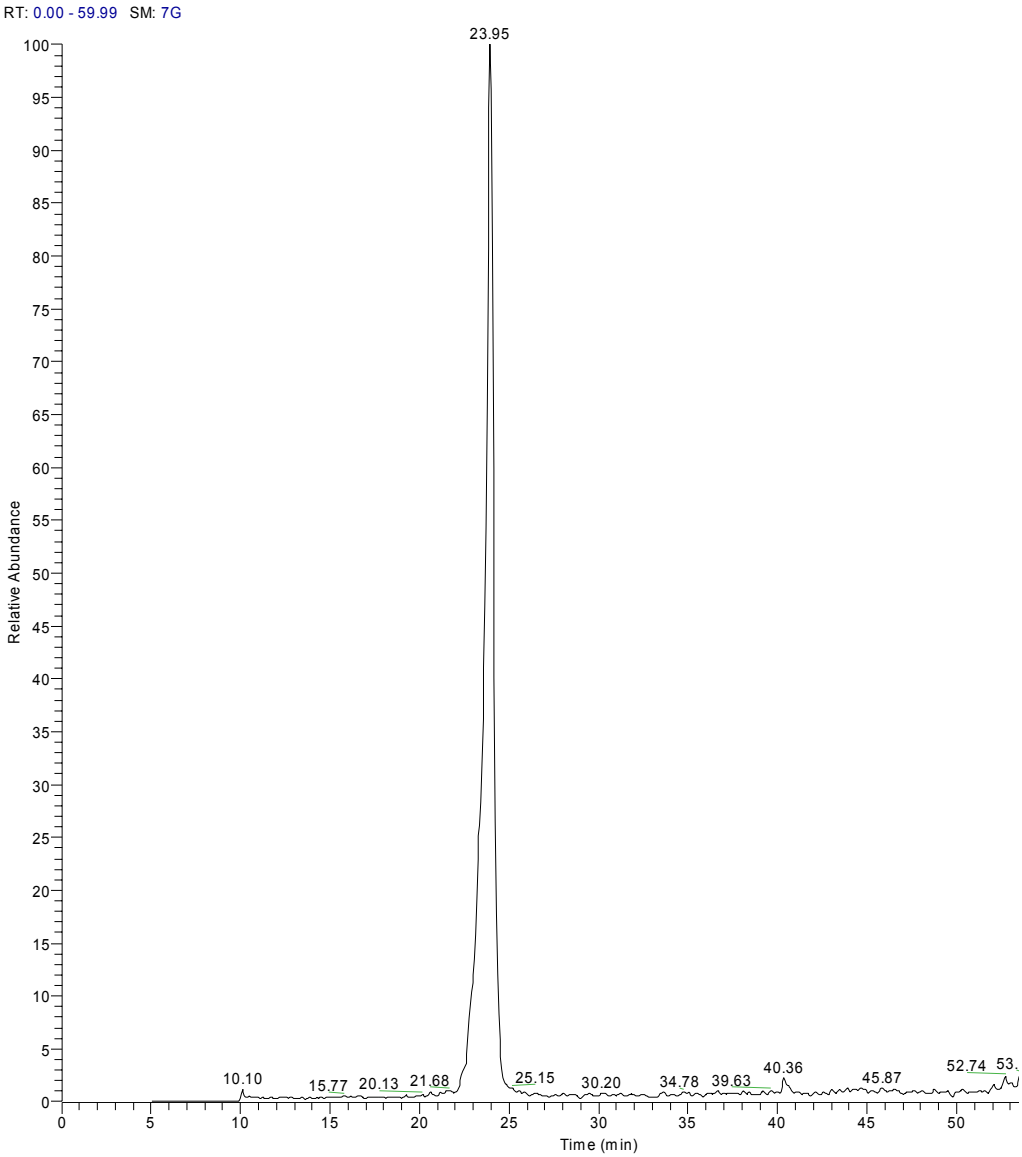


Fig.**3.28** IL-32 HPLC of linear IL- 32 1-103

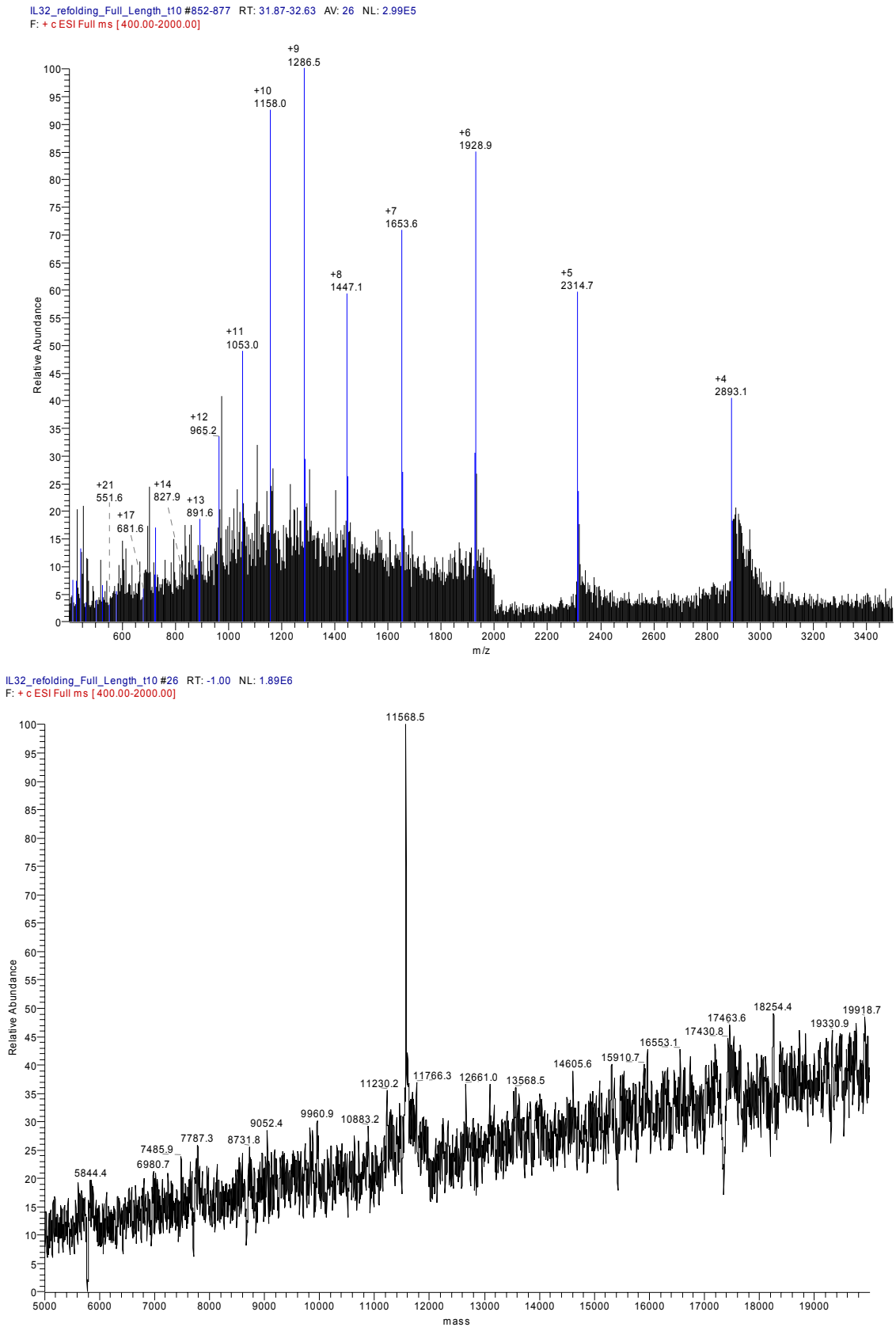


Fig.**3.29** IL-32 MS refolding IL- 32 1-103

3.3 IL-32: Peptides tested for biological activity

Il-32 peptides tested:

9) Il-32 gamma isoform 1-103
10) MCFPKVLSDD MKKLKARMVM LLPTSAQGLG AWVSA (Fragment **V**)
11) CDTKD TVGHPGPWRD KDPALW (Fragment **VI**)
12) C LSSQHQAIER FYDKMQNAES GRGQVMSSLA ELEDDFKEGY LET (Fragment **VIII**)

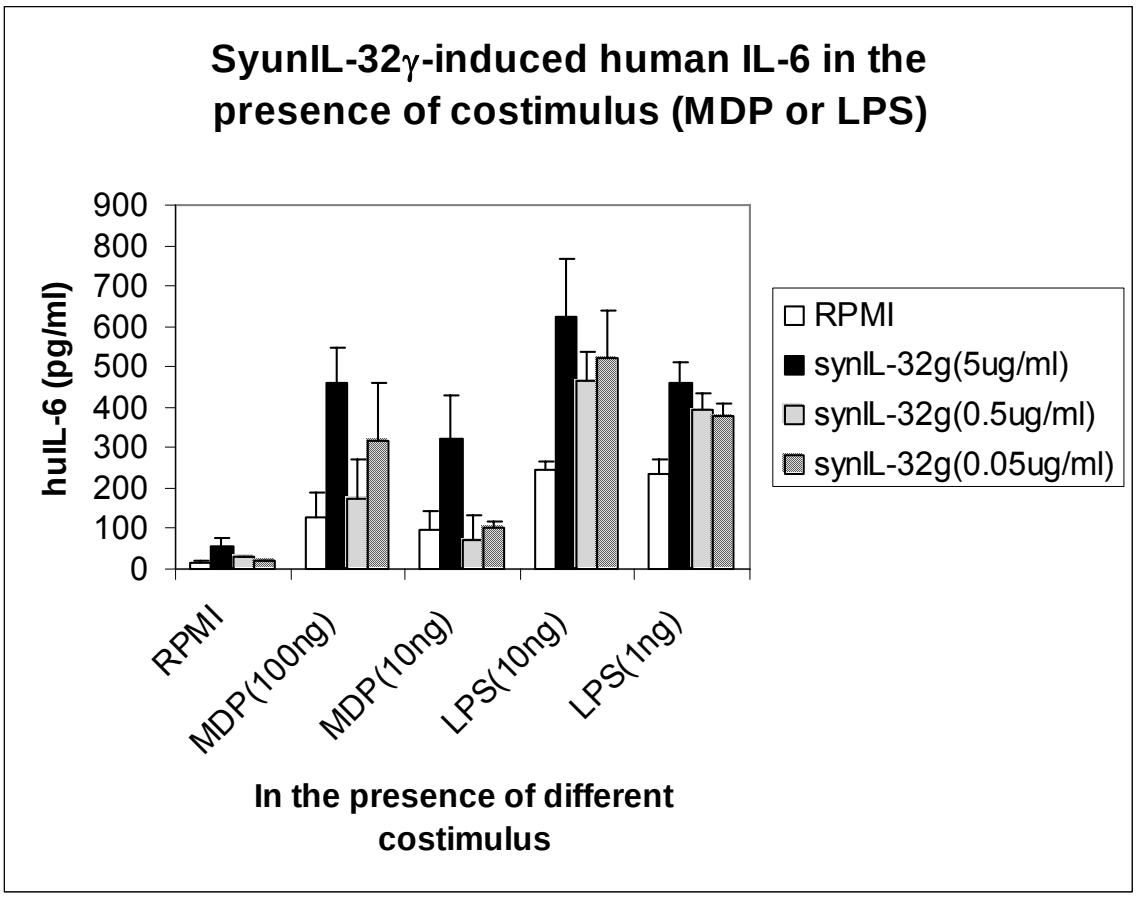
Results of Biological Test

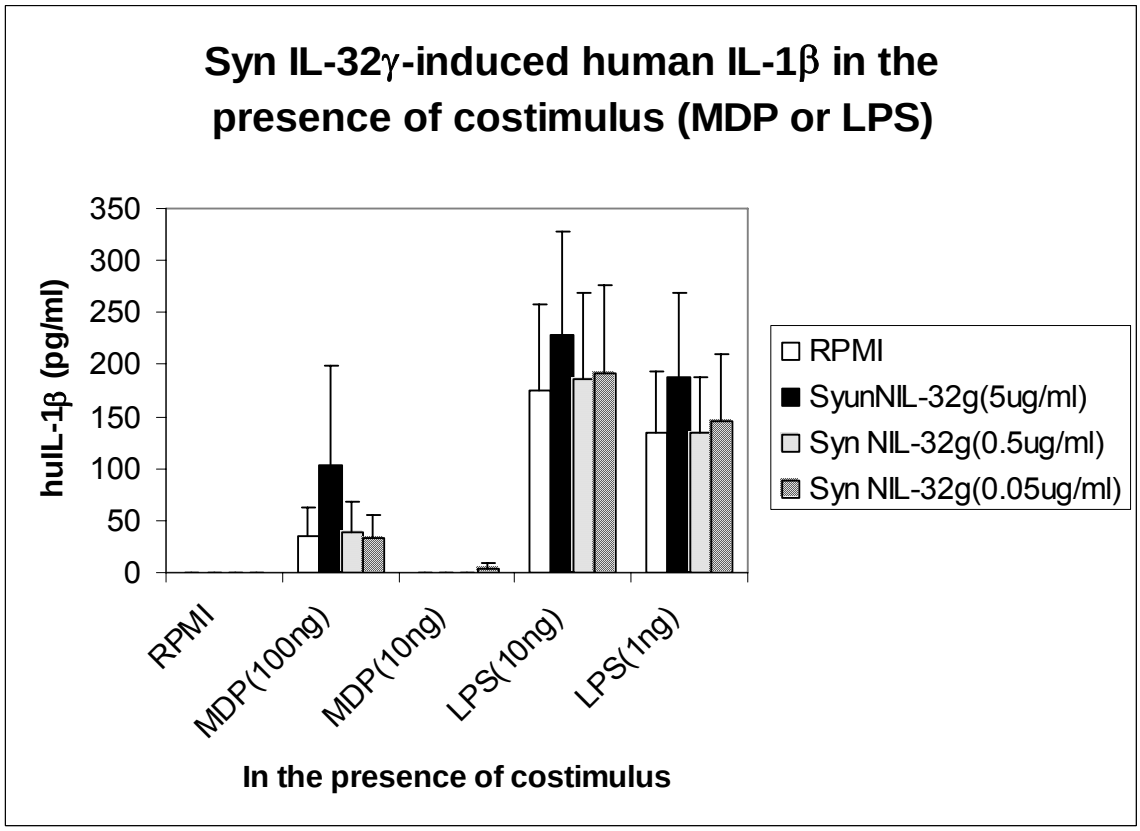
1) Il-32 gamma isoform 1-103 was found to be active in presence of the costimulus MDP. Indeed, the polypeptide in presence of MDP was able, in different models, to induce respectively production of IL-6, IL-1β and TNFα[56]. The synthetic IL-32γ 1-103 was also active in combination with LPS. In addition, IL-32 1-103 was active in differentiation of human monocytes into macrophages, see Netea (Proceedings of the National Academy of Sciences).[57]

Such results thus confirm and validate the original data obtained with the whole protein produced with recombinant techniques.

Unfortunately the single fragments (Fragment **V**, **VI** and **VIII**) utilized to build the protein as described above were all found not to be functional.

Synthetic N-terminal IL-32g induced cytokines in human PBMC, courtesy of Prof.C. Dinarello





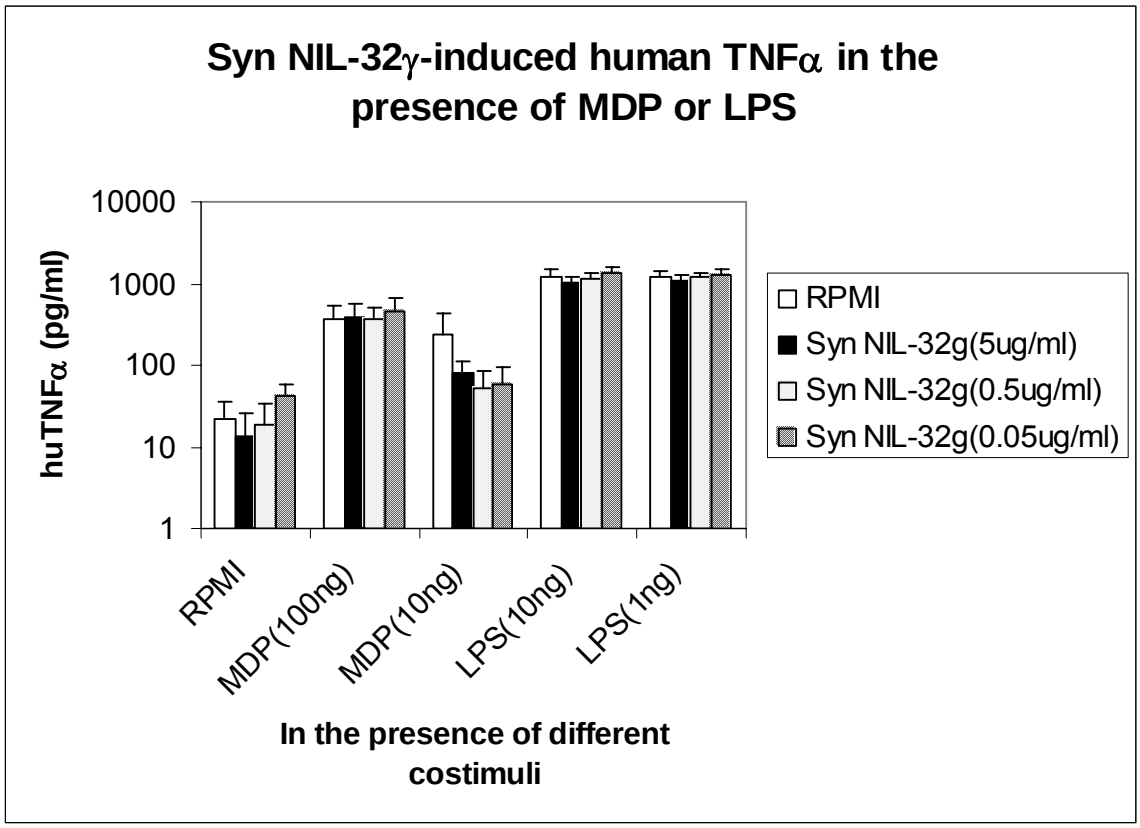


Fig. 3.30 IL 32 Biological assays

4. Conclusions

During the last decade, many advances both on peptide synthesis and ligation chemistry rendered accessible the total chemical synthesis and engineering of medium size proteins up to 180 residues.
We successfully synthesized all peptide fragments of Il-18 suitable for protein synthesis using Native Chemical Ligation and we successfully assembled and tested the Il-32 γ (gamma) isoform.
While the assembly of the IL-18 protein is still in progress, the biological tests of IL-32 1-103 gamma isoform we synthesized confirmed and validated the original data obtained with the whole protein produced with recombinant techniques and the synthetic material will be soon tested in many in vivo models by the group of Prof. C. Dinarello. The chemically synthesized IL-32 γ 1-103 N-terminal amino acids was found to be active in the presence of the co-stimulus MDP. Previous studies published in Proceedings of the National Academy of Sciences demonstrated that MDP synergizes with IL-32 gamma for the induction of IL-6. This synergism was shown to be dependent on caspase-1 via the induction of IL-1β. The importance of the observation is that MDP triggers the intracellular receptor known as NOD2 (Nuclear Oligomerization Domain). Patients with a forward shift mutation in NOD2 have Crohn's Disease. The finding that synthetic N-terminal IL-32 103 amino acids synergized with MDP was essential since the synthetic 103 peptide lacked any microbial products. Indeed, the polypeptide in presence of MDP was able, in different models, to induce respectively production of IL-6, IL-1β and TNFα.
There is also a need for IL-32 reagents for making neutralizing antibodies. The synthetic peptides and antibodies to these may help in the development of better ELISA detection methods. Furthermore the synthetic material will be useful to make monoclonal antibodies and look for epitopes for neutralization to cure Myelodysplastic Syndrome, COPD, rheumatoid arthritis, M. tuberculosis and Crohn's Disease.
Chemically synthesized IL-18 will unambiguously elucidate the mechanism of its inactivation testing the theory about disulfide bridging and its enhanced version may constitute an effective therapeutic agent for some form of cancer.

5. Experimental Section

Materials

Amino acids used for Boc chemistry peptide synthesis were, Tyr(2BrZ), Asp(OChx), Thr(Bzl), Ser(Bzl). Boc amino acids were obtained from Peptides International, USA. HBTU was from Luxembourg Science. DCM, DMF, TFA were purchased from Fluka. Pam resin was purchased from Novabiochem.
Protected Na-Fmoc-amino acid derivatives, coupling reagents and Rink amide MBHA resin were purchased from Calbiochem-Novabiochem
(Laufelfingen, Switzerland). S-trityl-mercaptopropionic acid was obtained from Peptides International (Louisville, Kentucky 40224 USA), DIEA was from Applied Biosystem (Foster City, CA). All other chemicals were commercially available by Sigma-Aldrich or Fluka (Bucks, Switzerland) or LabScan (Stillorgan, Dublin, Ireland) and were used as received unless otherwise stated. D2O was purchased by Sigma Aldrich (99.96% D).

Solid Phase Synthesis

All peptides were prepared by SPPS using machine-assisted protocols on a custom-modified Applied Biosystems model 430A and 4333A peptide synthesizers. Peptides were synthesized on the appropriate Boc-aminoacyl-OCH2-Pam preloaded resin. Thioester resins are prepared according to the procedure of Hackeng et al. starting from the MPAG resin (b-mercapto-propionic acid-Glycine) following a published procedure using PyBop as activator. SPPS was performed using in situ neutralization with the HBTU activation procedure for Boc chemistry as described by Kent. Bromoacetic acid was coupled manually to the N-terminal free amine peptide resin after standard neutralization cycles. After chain assembly was completed, the desired peptide was simultaneously deprotected and removed from the resin by treatment with anhydrous HF containing 5% p-cresol, and lyophilized. Peptides were characterized by analytical RP-HPLC Waters 2690 HPLC module with 214 nm UV detection, using a Symmetry 300 C8 column (5 mm, 0.46x 15 cm) with a linear gradient of buffer B in buffer A typically over 30 minutes at 1.0 mL/mi or by on a Shimadzu 10A-LC using a Phenomenex.C18 column (Torrance, CA), 4.6 x 250 mm 5 mm, eluted with an H2O/0.1%TFA (A) and CH3CN/0.1% TFA (B).
Peptide fragments were purified by preparative RP-HPLC on a Waters 600 HPLC module, using a C18 Waters DeltaPak preparative column (5 ´ 25 cm; 10 mm). at 12 mL/min.
Peptide identity was confirmed by MALDI/TOF STR DE (PE Biosystems, Foster City, CA) and ESI-MS that was performed with a Bruker Esquire 3000 Ion Trap (Bruker Daltonics, Bremen, DE).
.

Ligation reaction

The C-terminal unprotected peptide (1.5 equivalent) and the N–terminal peptide (1 equivalent) were solubilized in freshly degassed 6M guanidinium hydrohloride, 0.2 M sodium phosphate, pH 7.5 buffer at a concentration of 2.5 mM each. After addiction of 1% thiophenol the pH was adjusted to 7.0. Aliquots of this solution were treated with equal volumes of BME for 5 minutes and then of 10% TCEP for 10 minutes to completely hydrolyse any thiol abduct before HPLC analysis. Ligation reactions

analysis was conducted using a gradient starting from either 10 % of B with a gradient slope of 1.66 % of B for minute. Base line resolution of the different isomerization product was achieved using a gradient starting at 20 % of B and with an increase of 0.33 % B for minutes. Peptide identity was confirmed by using MALDI-MS and/or HPLC and ESI-MS until reaction completion.
Conversion of Thz-to Cys-peptide was performed by 20% BME for 30 minutes,
adding then 3% acetic acid and 0,5 M methoxy-amine×HCl, causing the pH of the reaction mixture to drop to »4. The reaction was vortexed for 3 hours. Then, this solution was treated with 10% TCEP for 10 minutes to completely hydrolyse any thiol adduct before HPLC analysis. The reaction was monitored by using MALDI-MS and/or HPLC and electro-spray ionization-MS until completion.

Post cleavage S-Acm peptide protection

The IL-18 Fragment **II**, 2.5 mg, ~ 1 micromol, was dissolved in 400 uL of water, then 800 uL of 1:3 TFMSA:TFA were added. The solution was stirred while an argon atmosphere was created. 30 eq. of Acm-OH dissolved in 100 uL of 1:3 TFMSA:TFA were added and the mixture was left under stirring over night. The solution was monitored by HPLC before and after reaction. A MALDI-MS spectrum was also recorded on a solution aliquot to assess the MW increase (expected ΔM = 72 amu).
Results:
Upon HPLC analysis, the main peak of the N-terminus protected thioester, shifted about 10 minutes while still retaining an Fmoc spectrum due to the C-terminal Fm-protected Glutamic acid. Importantly, MALDI-MS analysis showed the expected increase of 72 amu, confirming the incorporation of the Acm group.
C(Thiaz)EKISTLSC(Acm)ENKIISFKE(Fmoc)-Thioester Expected 2521, found 2521

HPLC

Analytical RP-HPLC (reverse-phase high performance liquid chromatography) was performed using a column 250 × 4 mm i.d. packed with Nucleosil 300-A, 5µm C8 particles. The flow rate was 0.7 mL/min, and effluent was monitored at 214 nm. Peptides were purified on a C8 semi-preparative column (250 × 10 mm i.d. Nucleosil 300-A, 5 µm particle size) at a flow rate of 4 mL/min monitored at 214 nm, or purified on a Vydac C8 preparative column (250 × 22 mm i.d., 10 µm particle size) at 15 mL/min monitored at 214nm. Solvents used in RP-HPLC were as follows: A, 0.1% TFA (1 g TFA in 1 L H_2O); B, 0.1% TFA in 90% acetonitrile (1 g of TFA mixed with 100 mL of H_2O and then brought to 1 L with acetonitrile). Generally, the conditions used in analytical work were 2 min isocratic at 20% B followed by a linear gradient to 65% B over 45 min. In preparative work a linear gradient (usually of the same slope) was used after 2min at initial conditions. All peptide components were collected manually at the detector exit, evaporated at room temperature under reduced pressure, frozen, and then recovered by lyophilization.

Mass spectrometry

Electrospray ionization mass spectrometry (ESI-MS) was performed in positive ion mode on a Platform-II instrument (Micromass, Manchester, England). Samples were introduced at 10 µL/min in solvent acetonitrile/water/formic acid (49.9:49.9:0.2). MALDI-TOF mass spectrometry was performed in linear mode using sinapinic acid as matrix on a Voyager Elite machine (Applied Biosystems Foster City, CA) equipped with delayed extraction. External calibration was performed on the electrospray

machines using a solution of horse heart apomyoglobin and on the MALDI-TOF machine using the mixture of peptides supplied by the manufacturer.

6. References

[1] Global Protein Therapeutics Market Analysis, (2007) 3B-Bharat, Book Bureau editions

[2] (a) Cancer Treat Rev. (2002) Apr;28 Suppl A:13-6; (b) Drug Discov Today. (2005) Nov 1;10(21):1451-8.; (3) Expert Opin Drug Deliv. 2008 Apr;5(4):371-83

[3] (a) TW Muir, PE Dawson, SBH Kent, (1997) Methods in Enzymology 289, 266-298; (b) GG Kochendoerfer, et al. (2003) *Science.* 7;299 (5608):884-7;

[4] M.Goodman, W.C and N.D.Smith (2003), Journal of Peptide Science, 9, 594-603

[5] RB. Merrifield, (1963)J.A.C.S,85: 2149-2154,

[6] (a) LA Carpino et al. (1972) J.Org.Chem. 37, 3404-3405; (b) G. Barany and RB Merrifield, (1980) The Peptides, Analysis, Synthesis, Biology. Vol 2, Academic press, N.Y., 1-284; (c) M Bodansky, (1993) Principles of peptide Synthesis. 2nd Ed. Springer-Verlag, Berlin.

[7] A. Wlodawer et al. (1989) Science, 254, 216-221; M. Schnolzer, P. Alewood,. A. Jones, D. Alewood,.S. B. H Kent, (1992), Int. J. Peptide Protein Res.*40*, 180-193

[8] Unpublished data From Matteo Villain and Keith Rose

[9] Unpublished data from M. Villain and H. Gaertener. Synthesis made using Boc chemistry with in situ-neutralization, ref. 5.

[10] (a) T.J. Lobl et al. (1988), Anal. Biochem, 170, 502, ,;(b) Villain M, Vizzavona J, Rose K,. (2001) Chem Biol. Jul; 8(7):673-9.

[11] H.Ball, Mascagni P, (1992) Int. J. Peptide Protein Res., *40*, 370;

[12] Ramage R, Raphy G, Tet.Lett. 1992, 33, 385

[13] Sakakibara S. Biopolymers. (1999); 51(4):279-96.

[14] Y Nishiuchi, T Inui, H Nishio, J Bódi, T Kimura, FI Tsuji, S Sakakibara. (1998) Proc Natl Acad Sci U S A. Nov 10;95(23):13549-54.

[15] P.Wang et al. (1999) J.Pept.Res., 53, 673-677

[16] TK Chang, DY.Jackson, J.P. Burnier, and J.A. Wells, (1994) Proc Natl Acad Sci U S A. 91, 12544-12548

[17] JP Tam, YA Lu, (2000), Fmoc Solid Phase peptide synthesis, PAS Oxford University Press, 243-263

[18] M.Schnolzer and S.B.Kent, (1992) Science 256, 221

[19] S.B.H. Kent,.G.R. Marshall, A..Wlodawer, R.M.Perlmutter,.(2000); Science, 288, 1590

[20] K.Rose, (1994) JACS 116, 30-33

[21] H Gaertner., R Offord., R Cotton., D Timms., R Camble., and K Rose.(1994) J.Biol.Chem 269, 7224

[22] C.Liu and J.P.Tam, (1994) JACS, 116, 4149-4153,

[23] C.Liu , C.Rao,and J.P.Tam,JACS, (1995),117, 3893-3899,

[24] P.E.Dawson, T.W. Muir,. I. Clark-Lewis, S.B.H. Kent (1994).Science *266*, 776-779.

[25] J. Wilken, S. B. H Kent., (1998) Cur. Opin.Biotechnol*.*, *9*, 412-426; P.E. Dawson S. B. H Kent., (2000) Annu.Rev.Biochem., *69*, 923-960

[26] L.E. Canne,.S.J...Bark, S.B.H.Kent, (1996) J.Am.Chem.Soc*.,* , *118*, 5891-5896

[27] S.B.H. Kent, Journal of Peptide Science, (2003), 9, 574-593

[28] P.Botti, M.R., Carrasco, S.B.H Kent,.(2001),Tetrahedron Lett*. 42*, 1831-1833.; D.W.Low, , M.G.Hill, M.R Carrasco,., S.B.H. Kent,, P.Botti, (2001), Proc. Natl. Acad. Sci. USA, *98*, 6554-6559

[29] V.M.F. Cardona,. O.Hartley, P.Botti, (2003) J.Pept.Res*, 61,* 152-157.
[30] J.Offer, C.N.C.Boddy and P.E. Dawson, (2002) J.Am.Chem.Soc.,124(17), 4642; Paolo Botti, and Sylvie Tchertchian. (2005) Protein and Peptides Letters, *12*(8), 729-735
[31] S. Tchertchian,et al. (2005) Understanding Biology Using Peptides, Proceedings of Nineteen American Peptide Symposium, San Diego, CA, USA,. Pages 61-63
[32] The Cytokines Facts Books, (2001) second edition, Academic Press,
[33] (a) H. Okamura et al.. (1995) Nature, 378, 88-91; (b) S.Uschio et al. (1996) J..Immunol.,156, 4274-4279; (c) K.Nakanishi et al. (2001) Cytokine Grow Factor Rev. , 12, 53-72
[34] (a) T.Osaki et.al. J.Immunol., (1998), 160, 1742-1749; (b) H. Fukumoto et al. Japan J. Cancer Research, (1997), 88, 501-505
[35] (a) H.Nagai, et al. (2000), Cancer Investigation, 18, 206-213; (b) D.J. Herzyk et al. (2003), Toxicol.Pathol. 31, 554-561
[36] Lee et al. (2004), Proc. Natl. Acad. Sci. USA, 101, 8815-8820.
[37] S. Kikkawa et al. Biochemical and Biophysical Research Communication, (2001), *281*, 461-467
[38]Y.Yamamoto et al. (2004), Biochemical and Biophysical Research Communication, 317, 181-186
[39] CA Dahl et al., (1992) J.Immunology 148, 597-603
[40] Panelli et al., (2002) Genome Biol.Jun 25;3(7):RESEARCH0035. Epub 2002 Jun 25.
[41] Soo-Hyun Kim et al. (2005), Immunity, 22, 131-142
[42] Unpublished data from C. Dinarello
[43] A.M Marcondes,.et al. (2008). Dysregulation of IL-32 in myelodysplastic syndrome and chronic myelomonocytic leukemia modulates apoptosis and impairs NK function. Proc Natl Acad Sci U S A 105:2865-2870.
[44] F Calabrese,.et al., (2008). IL-32, a novel proinflammatory cytokine in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 178:894-901.
[45] (a) L.A Joosten,.et al., (2006). IL-32, a proinflammatory cytokine in rheumatoid arthritis. Proc Natl Acad Sci U S A 103:3298-3303.
(b) N., F Cagnard,et al..(.2005). Interleukin-32, CCL2, PF4F1 and GFD10 are the only cytokine/chemokine genes differentially expressed by in vitro cultured rheumatoid and osteoarthritis fibroblast-like synoviocytes. Eur Cytokine Netw 16:289-292.; (c) H.Shoda, et.al, (2006). Interactions between IL-32 and tumor necrosis factor alpha contribute to the exacerbation of immune-inflammatory diseases. Arthritis Res Ther 8:R166.
[46] M.G Netea,.et al.. (2006). Mycobacterium tuberculosis induces interleukin-32 production through a caspase- 1/IL-18/interferon-gamma-dependent mechanism. PLoS Med 3:e277.
[47] M.F Nold, et al..,(2008). Endogenous IL-32 controls cytokine and HIV-1 production. J Immunol 181:557-565.
[48] Bold and underlined are the ligation sites.
[49] M.Villain, H.Gaertner and P. Botti (2003) Eur. J. Org. Chem., 3267-3272
[50] J.P. Tam and Q. Yu, (1998), Biopolymers, 46, 319
[51] T.M Hackeng,.et al.. (2001) Proc. Natl. Acad. Sci. USA, *96*, 10068-10073
[52] (a) M Villain,. J.Vizzavona,; H Gaertner,. (2001), Peptides: the Wave of the Future. Proceedings of Seventeenth American Peptide Symposium, San Diego, Ca, USA; June 9–14, 107–108. (b) M Villain,.; J.Vizzavona,; H Gaertner,.(2002) Collected

Papers of the Seventh Symposium of Innovation and Perspectives in Solid Phase Synthesis, Southampton, England, UK; 39–42.
[53] M Villain,;H Gaertner,.; P Botti**,**. (2003) in Peptides: Peptide Revolution: Genomics, Proteomics & Therapeutics. Proceedings of Eighteen American Peptide Symposium, Boston, MA, USA; July19–23, , 71–72
[54] E. Marcucci et al.,(2008) J. Comb. Chem. 10, 69-78
[55] D Novick,et.al.. (2006). Natl Acad Sci U S A in press:
[56] M.G Netea,.et. al., (2005).. IL-32 synergizes with nucleotide oligomerization domain (NOD) 1 and NOD2 ligands for IL-1b and IL-6 production through a caspase 1-dependent mechanism. Proc Natl Acad Sci U S A 102:16309-16314
[57] M.G.,Netea et.al,. (2008). Interleukin-32 induces the differentiation of monocytes into macrophage-like cells. Proc Natl Acad Sci U S A 105:3515-3520..

Abbreviations

The utilized abbreviations for natural amino acids and for peptides are in agreement with the IUPAC-IUB Commission on Biochemical Nomenclature (IUPAC-IUB, 1984; 1989).

- **ACN** : Acetonitrile
- **BME** : 2 mercapto-ethanol
- **t-Boc** tert-butyloxycarbonyl
- **t-Bu**: ter-butyl
- **CD:** Circular Dichroism
- **COSY:** COrrelation SpectroscopY
- **1D, 2D, 3D:** mono-, bi-, three-dimensional
- **DCM:** Dichloromethane
- **DIEA**: Diisopropylethylamine
- **DMF:** Dimethylformamide
- **DMSO**: Dimethyl sulfoxide
- **DMAP:** Dimethylamino-pyridine
- **Dnp:** 2,4-dinitrophenyl
- **DTT:** 1,4-dithiothreitol
- **ESI:** ElectroSpray Ionization
- **Fmoc**: 9-fluorenylmethyloxycarbonyl
- **HATU:**O-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
- **HBTU:** 2-(1H-Benzotriazole-1-yl)-1,1,3,3,tetramethyluronium hexafluorophosphate
- **HF:** hydrogen fluoride
- **HOBt**: 1-hydroxybenzotriazole
- **MALDI-TOF:** Matrix Assisted Laser Desorption Ionization -Time of Flight
- **NMR:** Nuclear Magnetic Resonance
- **NMM: *N*-methylmorpholine;**
- **NCL:** native chemical ligation
- **NOE:** Nuclear Overhauser Enhancement
- **NOESY:** Nuclear Overhauser Enhancement and exchange SpectroscopY

- **OdiMeOPac: (**2´´-4´´-dimethoxy)-phenacyl group;
- **OFm:** 9-fluorenylmethyl ester;
- **OMop:** 1-methyl-2-oxo-2-phenylethyl ester;
- **OPac:** phenacyl esters**;**
- **OPse:** phenylsulfonyl)ethyl ester;
- **OTce:** 2,2,2-trichloroethyl esters;
- **Pam:** 4-hydroxymethylphenylacetic acid;
- **PyBop:** benzotriazolyl N-oxytrispyrrolidinophosphonium hexafluorofluorophosphate
- **Rink Amide MBHA Resin**: 4–2', 4'- Dimethoxyphenyl – Fmoc - aminomethyl)-phenoxyacetamidonorleucyl-MBHA resin
- **RP-HPLC:** Reverse Phase-High Performance Liquid Cromatography
- **SPPS**: Solid Phase Peptide Synthesis
- **TFA:** Trifluoroacetic acid
- **Thz:** Thiazolidine
- **TMS:** trimethylsilane
- **TOCSY:** TOtal Correlation SpectroscopY
- **Tris:** Tris(hydroxymethyl)aminomethane
- **Trt**: trityl
- **UV-vis**: *Ultra* Violet- Visible

Papers

1. Semisynthetic Analogues of PSC-Rantes, a potent Anti-HIV Protein. Hubert Gaertner, Robin Offord, Paolo Botti, Gabriel Kuenzi and Oliver Hartley, Bioconjugate Chemistry,, 2008 Feb;19(2):480-9.
2. The Chemical synthesis of the Gstl protein by NCL on a X-Met site. Angela saporito, Daniela Marasco, Angela Chambery, Paolo Botti, Simona Monti, Carlo Pedone and Menotti Ruvo, Biopolymers, 2006, Dec 5;83(5):508-18.
3. Chemical synthesis of proteins and circular peptides using Nα-(1-phenyl-2-mercaptoethyl) auxiliaries. Paolo Botti, and Sylvie Tchertchian. Protein and Peptides Letters, 2005, 12(8), 729-735
4. Native Chemical Ligation at Aromatic Residues. S. Tchertchian, F. Oplinger, M. Paolini, S. Manganiello, S. Raimondi, B. Depresle, N. Dafflon, H. Gaertner, and P. Botti. Understanding Biology Using Peptides, Proceedings of Nineteen American Peptide Symposium, San Diego, CA, USA, 2005. Pages 61-63

5. Congress Comunications

Selected for oral presentation at International Symposia:

1) Chemistry and Biology of Peptides 2007**,** 6th July at the Rhodes House, Oxford, UK.

2) 19th American Peptide Symposium, San Diego, USA, June 2005,

Article

Semisynthetic Analogues of PSC-RANTES, a Potent Anti-HIV Protein

Hubert Gaertner, Robin Offord, Paolo Botti, Gabriel Kuenzi, and Oliver Hartley

Bioconjugate Chem., **2008**, 19 (2), 480-489 • DOI: 10.1021/bc7003044 • Publication Date (Web): 08 January 2008

Semisynthetic Analogues of PSC-RANTES, a Potent Anti-HIV Protein

Hubert Gaertner, Robin Offord,† Paolo Botti,‡ Gabriel Kuenzi, and Oliver Hartley*

Department of Structural Biology and Bioinformatics, Centre Médical Universitaire, 1 rue Michel Servet, 1211 Geneva 4, Switzerland. Received August 13, 2007; Revised Manuscript Received October 8, 2007

New HIV prevention methods are needed, and among those currently being explored are "microbicides", substances applied topically to prevent HIV acquisition during sexual intercourse. The chemokine analogue PSC-RANTES (N^{α}(*n*-nonanoyl)-*des*-Ser^{1}-[L-thioprolyl^{2}, L-cyclohexylglycyl^{3}]-RANTES(4–68)) is a highly potent HIV entry inhibitor which has shown promising efficacy in its initial evaluation as a candidate microbicide. However, a way must be found to produce the molecule by cheaper means than total chemical synthesis. Since the only noncoded structures are located at the N-terminus, a possible solution would be to produce a protein fragment representing all but the N-terminal region using low-cost recombinant production methods and then to attach, site specifically, a short synthetic fragment containing the noncoded N-terminal structures. Here, we describe the evaluation of a range of different conjugation chemistries in order to identify those with potential for development as economical routes to production of a PSC-RANTES analogue with antiviral activity as close as possible to that of the parent protein. The strategies tested involved linkage through oxime, hydrazone/hydrazide, and Ψ[CH_2−NH] bonds, as well as through a peptide bond obtained either by a thiazolidine rearrangement or by direct α-amino acylation of a protein fragment in which 4 of the 5 lysine residues of the native sequence were replaced by arginine (the fifth lysine is essential for activity). Where conjugation involved replacement of one or more residues with a linker moiety, the point in the main chain at which the linker was introduced was varied. The resulting panel of 22 PSC-RANTES analogues was evaluated for anti-HIV activity in an entry inhibition assay. The [Arg^{25,45,56,57}] PSC-RANTES analogue has comparable potency to PSC-RANTES, and one of the oxime linked analogues, 4L-57, has potency only 5-fold lower, with scope for improvement. Both represent promising leads for development as microbicide compounds that could be produced at low cost via semisynthesis.

INTRODUCTION

New prevention methods are urgently required to combat the HIV/AIDS pandemic, which is currently running at a rate of approximately 5 million new infections per year (*1*). One of the strategies currently being developed relates to the use of "microbicides"—antiviral products that can be applied topically to genital mucosa prior to sexual intercourse, thereby preventing HIV transmission from person to person (*2*).

PSC-RANTES (*3*) is a chemokine analogue that has shown promising efficacy in preclinical studies related to its use as a microbicide (*4*). In its totally synthetic form, however, PSC-RANTES is likely to be impossibly expensive for worldwide use (*5*). Since the non-natural structures in PSC-RANTES are all located within an N-terminal pharmacophore, semisynthetic strategies could be envisaged in which a protein fragment representing all but the N-terminus of the sequence is produced by low-cost microbial fermentation, followed by site-specific conjugation of a short synthetic peptide containing the N-terminal pharmacophore.

One strategy to produce PSC-RANTES semisynthetically would involve the use of native chemical ligation (*6*) between a protein fragment beginning at Cys^{10} and an N-terminal fragment with a C-terminal thioester that contains the noncoded pharmacophore. We chose not to consider this route, however, since it would involve the production of a long and relatively expensive thioester peptide, as well as a folding step (formation of disulfide bridges) after the ligation. Instead, we chose to focus on approaches involving conjugation conditions that are compatible with the presence of the native disulfide bonds in the protein fragment and which could be adapted to the use of shorter, less expensive synthetic N-terminal fragments. Hence, the first set of strategies we envisaged involved the formation of oxime, hydrazone/hydrazide, and Ψ[CH_2−NH] bonds. These require N-terminal Ser or Thr residues to be engineered into the protein fragment (see linker-based strategy in Figure 1). With N-terminal Ser or Thr, periodate oxidation of the amino alcohol to a glyoxylic acid residue, O=CH−CO-, allows the specific conjugation to an aminooxy- or hydrazide-functionalized peptide through the formation of an oxime or hydrazone bond. In an another linker strategy, reductive amination under slightly acidic conditions was used to conjugate the short synthetic N-terminal peptide fragment functionalized with an aminoacetaldehyde moiety to the protein fragment. This approach, though less specific because it is only based on the difference of reactivity between α- and ε-amine groups toward aldehydes, has the advantage of generating the shortest linker—involving the substitution of only one amino acid residue—but the disadvantage of introducing a positively charged group into the protein backbone (Scheme 1). All of the linker-based strategies lead to modification to both the backbone and the side chains of the

* Corresponding author. Tel. +41 22 379 54 75. Fax +41 22 379 55 02.

† Present address Mintaka Foundation for Medical Research, 14 chemin des Aulx, 1228 Plan-les-Ouates, Geneva, Switzerland.

‡ Present address ArisGen, 14 chemin des Aulx, 1228 Plan-les-Ouates, Geneva, Switzerland.

[1]Abbreviations: 4L-5,7H, etc., see description of linker terminology in the Discussion section; AOP-RANTES, the aminooxypentane oxime of [glyoxylyl^{1}]RANTES (2-68); Apm, aspartimide; Chg, cyclohexylglycine; EDC, *N*-(3-dimethylaminopropyl)-*N*′-ethylcarbodiimide; HM-Thiaz, 2-hydroxymethylthazolididine; MES, morpholinoethanesulfonic acid; NHS, *N*-hydroxysuccinimide; NNY-RANTES, N^{α}(*n*-nonanoyl)-RANTES(2-68); PrS, thioproline; Pse, 2-(phenylsulphonyl)ethyl; PSC-RANTES, (N^{α}(*n*-nonanoyl)-*des*-Ser^{1}-[L-thioprolyl^{2}, L-cyclohexylglycyl^{3}]-RANTES(4-68)); RANTES, the traditional name for the chemokine CCL5 (Swissprot: CCL5_HUMAN).

10.1021/bc7003044 CCC: $40.75
Published on Web 01/08/2008

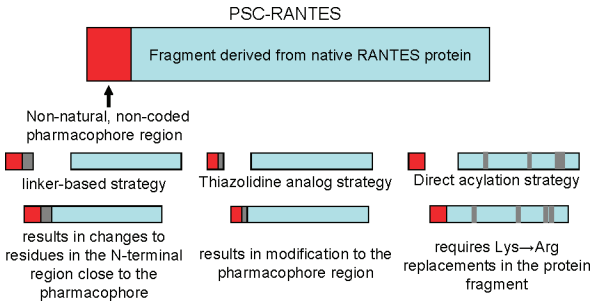


Figure 1. Schematic representation of the semisynthetic strategies used in this study. PSC-RANTES (upper panel) can be considered to consist of a protein fragment derived from native RANTES (cyan), and which could be produced cheaply by biosynthesis, to which a short, synthetic N-terminal pharmacophore region containing non-natural, noncoded structures (red) is joined. All three of the semisynthesis strategies we evaluated in this study necessitate the introduction of modifications (indicated in gray) to the PSC-RANTES protein.

protein at the linkage site, and some of them also change the register of the backbone by adding atoms (Scheme 1). Hence, we opted to create panels of analogues for each linkage type, varying the site at which the linkage was introduced (Table 1).

Yet another possible strategy requires the protein fragment to have N-terminal cysteine. The resulting 1-amino, 2-thiol structure at the N-terminus can be specifically conjugated to a glycolaldehyde-functionalized ester of the short synthetic N-terminal fragment through formation of a thiazolidine, which can be further rearranged to a 2-hydroxymethyl-thioproline with the generation of an authentic peptide bond (Scheme 2). The drawback here is that the structure of the pharmacophore unit of PSC-RANTES cannot be precisely reproduced (see thiazolidine analogue strategy in Figure 1).

Our last strategy involved direct acylation of the amino terminus of protein fragments in which either four or five of the five lysine residues present in native RANTES had been replaced by arginine (see direct acylation strategy in Figure 1). This approach has the advantage of generating a natural peptide bond at the linker site but the disadvantage of changing structures elsewhere in the protein that may be important for activity.

The PSC-RANTES analogues from the different groups were evaluated for their anti-HIV activity. We discuss the results obtained with regard to structure–activity relationships at the anti-HIV pharmacophore of PSC-RANTES and suggest that two of the above strategies have potential for development as economic routes to production of semisynthetic PSC-RANTES analogues that could be effective microbicides.

EXPERIMENTAL PROCEDURES

Reagents and Chemicals. Unless otherwise specified, all solvents and reagents were obtained from Fluka (Buchs, Switzerland), were of analytical or higher grade, and were used without further purification. All amino acids were purchased from Orpegen (Heidelberg, Germany). The resins were from NeoMPS (Strasbourg, France) or Bachem (Bubendorf, Switzerland). Water was repurified using a Milli-Q system (Millipore, Inc., Bedford, MA).

HPLC. Analytical RP-HPLC (reverse-phase high-performance liquid chromatography) was performed using a column 250 × 4 mm i.d. packed with Nucleosil 300-A, 5 μm C8 particles. The flow rate was 0.7 mL/min, and effluent was monitored at 214 nm. Peptides were purified on a C8 semi-preparative column (250 × 10 mm i.d. Nucleosil 300-A, 5 μm particle size) at a flow rate of 4 mL/min monitored at 214 nm, or purified on a Vydac C8 preparative column (250 × 22 mm i.d., 10 μm particle size) at 15 mL/min monitored at 214 nm. Solvents used in RP-HPLC were as follows: A, 0.1% TFA (1 g TFA in 1 L H_2O); B, 0.1% TFA in 90% acetonitrile (1 g of TFA mixed with 100 mL of H_2O and then brought to 1 L with acetonitrile). Generally, the conditions used in analytical work were 2 min isocratic at 20% B followed by a linear gradient to 65% B over 45 min. In preparative work, a linear gradient (usually of the same slope) was used after 2 min at initial conditions. All peptide components were collected manually at the detector exit, evaporated at room temperature under reduced pressure, frozen, and then recovered by lyophilization.

Mass Spectrometry. Electrospray ionization mass spectrometry (ESI-MS) was performed in positive ion mode on a Platform-II instrument (Micromass, Manchester, England). Samples were introduced at 10 μL/min in solvent acetonitrile/water/formic acid (49.9:49.9:0.2). MALDI-TOF mass spectrometry was performed in linear mode using sinapinic acid as the matrix on a Voyager Elite machine (Applied Biosystems, Foster City, CA) equipped with delayed extraction. External calibration was performed on the electrospray machines using a solution of horse heart apomyoglobin and on the MALDI-TOF machine using the mixture of peptides supplied by the manufacturer.

Synthesis of RANTES Fragments. Although the route to bulk manufacture of the larger RANTES fragments would ultimately be biosynthesis, for exploratory experimental purposes it was more rapid and convenient to prepare them by total chemical synthesis. The two proteins and ten protein fragments described in Table S1 of the Supporting Information were prepared by polymer-supported organic synthesis of the two segments that would have been formed if they had been cleaved between residues 33 and 34. *t*-Butoxycarbonyl chemistry was used, as described in ref (*7*). The segments were coupled in a native chemical ligation (*6*). Folding of the resulting molecule with concomitant disulfide formation was carried out in the presence of a Cys-SH/$(Cys\text{-}S)_2$ redox couple except for [Cys^2]-RANTES(2-68), where conjugation to the N-terminal peptide fragment was performed before the folding step. Purity and integrity of the resulting analogues were routinely verified by HPLC and mass spectrometry.

Synthesis of *n*-Nonanoyl Peptides. The short N-terminal synthetic fragments that contain the pharmacophore all possess an N-terminal *n*-nonanoyl group as in PSC-RANTES (Scheme 1). These *n*-nonanoyl peptides were synthesized as described in ref (*8*). Crude material obtained after ether precipitation (following HF resin cleavage) was used without HPLC purification for further C-terminal functionalization. For synthesis of *n*-nonanoyl-PrS-Chg-Ser-Ser-Asp^6, Boc-Asp(OPse)-OH was used to maintain side-chain protection during further C-terminal modification of the peptide. Because of difficulties with the deprotection, this peptide was also synthesized using Fmoc chemistry, using 2-chlorotrityl chloride resin and cleaving the fully protected peptide with acetic acid/trifluoroethanol/dichloromethane (2:2:6) for 2 h.

Synthesis of Heterobifunctional Spacer Reagents Containing Aminooxy Groups (Scheme 3). *I. 2-(Boc-aminooxy)ethylamine (III). N-[2-(Tert-butyl-N-hydroxycarbamoyl)ethyl]phthalimide (I).* Conditions of synthesis are similar to those described in ref (*9*). *tert*-Butyl *N*-hydroxycarbamate (1.32 g, 10 mmol) was dissolved in anhydrous DMF (7 mL), NaH (380 mg, 60% dispersion in mineral oil, 9.5 mmol) was added at room temperature, and the mixture stirred for 1 h, then cooled to 0 °C, treated with 2-bromoethylphthalimide (2.03 g, 8 mmol) for 2 h at 0 °C, and then left at room temperature overnight. The material was precipitated in cold water (100 mL) and the precipitate redissolved in EtOAc. The organic layer was

Scheme 1. Backbone Structure of PSC-RANTES Analogues Resulting from the Use of Different Linker-Based Strategies

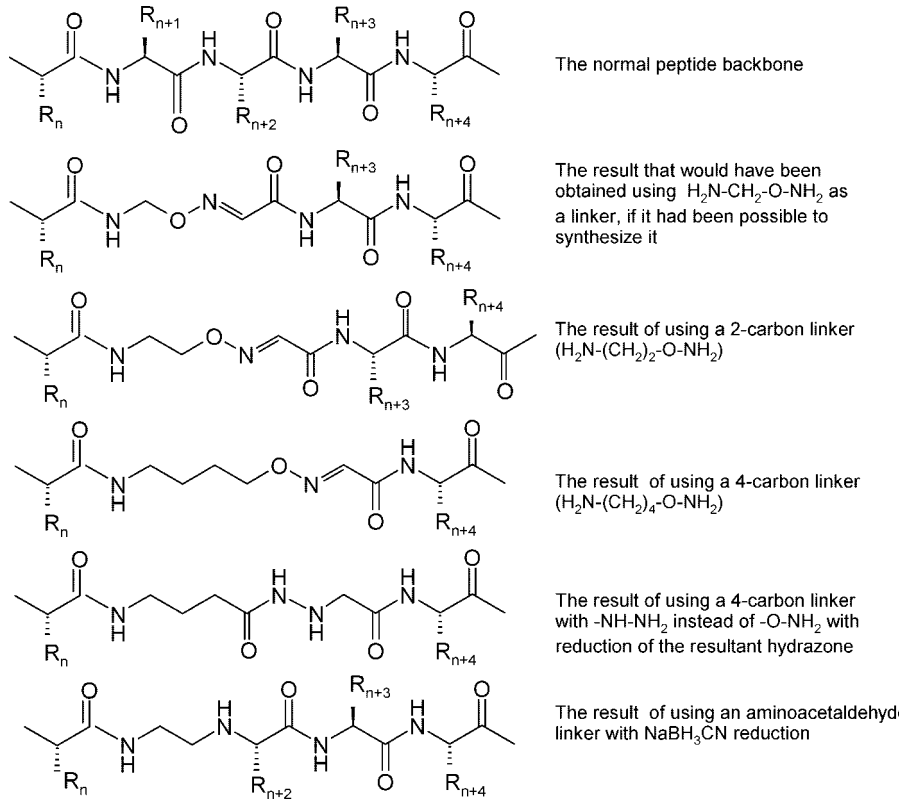


Table 1. N-Terminal Structures of PSC-RANTES Analogues Produced Using Linker-Based Strategies[a]

Compound	N-terminal structure	Linker used
2L-45	*n*-nonanoyl-PrS-Chg-^^^^^^^-Asp-Thr-Thr-Pro	2-carbon oxime
2L-56	*n*-nonanoyl-PrS-Chg-Ser-^^^^^^^-Thr-Thr-Pro	2-carbon oxime
2L-67	*n*-nonanoyl-PrS-Chg-Ser-Ser-^^^^^^^-Thr-Pro	2-carbon oxime
2L-78	*n*-nonanoyl-PrS-Chg-Ser-Ser-Asp-^^^^^^^-Pro	2-carbon oxime
4L-46	*n*-nonanoyl-PrS-Chg-^^^^^^^^^^-Thr-Thr-Pro	4-carbon oxime
4L-57	*n*-nonanoyl-PrS-Chg-Ser-^^^^^^^^^^-Thr-Pro	4-carbon oxime
4L-68	*n*-nonanoyl-PrS-Chg-Ser-Ser-^^^^^^^^^^-Pro	4-carbon oxime
4L-46H	*n*-nonanoyl-PrS-Chg-^^^^^^^^^^-Thr-Thr-Pro	4-carbon hydrazone
4L-57H	*n*-nonanoyl-PrS-Chg-Ser-^^^^^^^^^^-Thr-Pro	4-carbon hydrazone
4L-68H	*n*-nonanoyl-PrS-Chg-Ser-Ser-^^^^^^^^^^-Pro	4-carbon hydrazone
4L-46HR	*n*-nonanoyl-PrS-Chg-^^^^^^^^^^-Thr-Thr-Pro	4-carbon methyl-hydrazine
4L-57HR	*n*-nonanoyl-PrS-Chg-Ser-^^^^^^^^^^-Thr-Pro	4-carbon methyl-hydrazine
4L-68HR	*n*-nonanoyl-PrS-Chg-Ser-Ser-^^^^^^^^^^-Pro	4-carbon methyl-hydrazine
1L-4	*n*-nonanoyl-PrS-Chg-^^^-Ser-Asp-Thr-Thr-Pro	1-carbon ψ[CH_2-NH]
1L-5	*n*-nonanoyl-PrS-Chg-Ser-^^^-Asp-Thr-Thr-Pro	1-carbon ψ[CH_2-NH]
1L-6	*n*-nonanoyl-PrS-Chg-Ser-Ser-^^^-Thr-Thr-Pro	1-carbon ψ[CH_2-NH]
1L-7	*n*-nonanoyl-PrS-Chg-Ser-Ser-Asp-^^^-Thr-Pro	1-carbon ψ[CH_2-NH]
1L-7 *des*	*n*-nonanoyl-PrS-Chg-Ser-Ser-Apm-Thr-^^^-Pro	1-carbon ψ[CH_2-NH]

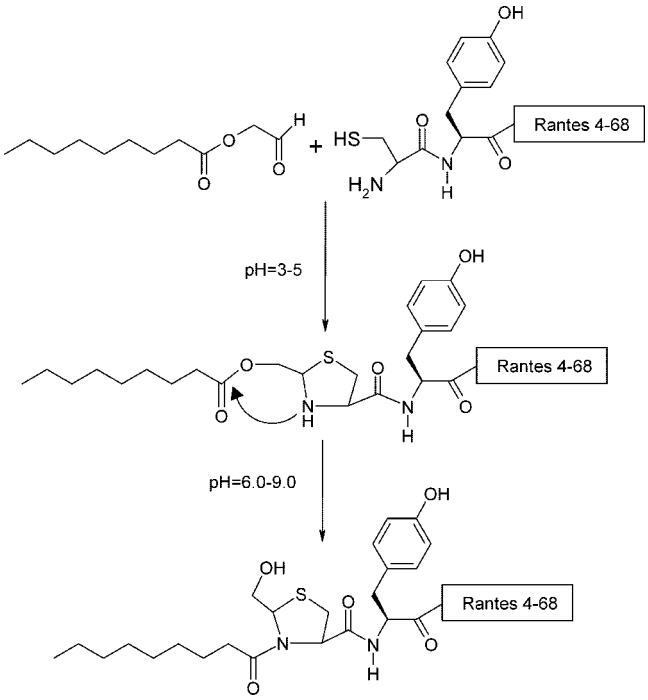
[a] The residues substituted by the linkers are indicated using ∧∧∧; the structures of the different linkers can be found in Scheme 1. See the appropriate footnote for the abbreviations for noncoded amino acids.

successively washed with a saturated NaCl solution and then water. The crude product (**I**) was dried over $MgSO_4$ and solvent removed. Yield 1.3 g (53%).

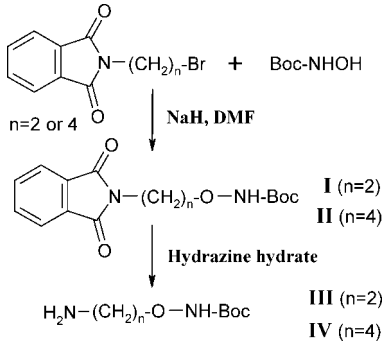
N-[2-(tert-butyl-N-hydroxycarbamoyl)ethyl]amine (III). A mixture of 1.1 g of *N*-[2-(*tert*-butyl-*N*-hydroxycarbamoyl)ethyl]phthalimide (**I**) and 1.4 mL hydrazine hydrate in 20 mL ethanol was refluxed overnight at 90 °C. The solution was cooled, and the precipitated phthalyl hydrazide was separated by centrifugation. The supernatant was diluted with water and acidified with 10% TFA. A further precipitate was separated by centrifugation and the deprotected material isolated from the supernatant by preparative HPLC (yield 215 mg, 0.75 mmol, 21%), showing the expected mass (ESI-MS M + H found 177.2, theor. M + H 177.0).

II. 4-(Boc-aminooxy)butylamine (IV). The same protocol was used starting from *tert*-butyl *N*-hydroxycarbamate (400 mg, 3 mmol), NaH (114 mg, 2.85 mmol), and 4-bromobutylphthalimide (675 mg, 2.4 mmol). The recovered *N*-[4-(*tert*-butyl-*N*-hydroxycarbamoyl)butyl]phthalimide (**II**) (630 mg, 1.7 mmol, 70%) was treated with a 2-fold molar excess of hydrazine

Scheme 2. Ligation through Thiazolidine Ring Formation



Scheme 3. Synthesis of Heterobifunctional Spacer Reagents Containing Aminooxy Groups



hydrate to afford, after isolation and HPLC purification as above, an oily product (**IV**) (105 mg, 0.33 mmol, 19%) with the expected mass (ESI-MS M + H found 204.8, theor. M + H 205.0).

Synthesis of Aminooxy-Functionalized n-Nonanoyl Peptides. Coupling of the 2- and 4-Carbon Linkers to the C-Termini of the n-Nonanoyl Peptides. *n*-Nonanoyl peptide (25 μmol) was dissolved in 50 μL DMF and mixed with equimolar amounts of solid NHS and diisopropylcarbodiimide. A 3-fold molar excess of Boc-aminooxy-linker (**III** or **IV**) was dissolved in 50 μL DMF and added to the *n*-nonanoyl peptide solution, and diisopropylethylamine, equivalent to a 4-fold molar excess over *n*-nonanoyl peptide, was added to the reaction mixture which was then incubated overnight at room temperature. The mixture was then acidified with 10 μL acetic acid and the conjugation product isolated by preparative HPLC using a gradient of 30−70% B over 40 min. In the case of *n*-nonanoyl-PrS-Chg-Ser-Ser-Asp(OPse)-OH, the (OPse) protecting group was removed afterward by 2 h treatment at 37 °C in 3 mL of 0.1 M $NaHCO_3$, 10% α-mercaptoethanol, 40%CH_3CN, pH 9.0 (*10*). The lyophilized material was deprotected by a 5 min treatment in 1 mL neat TFA. The TFA was evaporated under N_2, the material redissolved in CH_3CN/water and freeze-dried.

Synthesis of Hydrazide-Functionalized *n*-Nonanoyl Peptides. These were made manually at the 0.1 mmol scale according to standard techniques of Fmoc peptide chemistry (*11*) starting with 2-methoxy-4-alkoxybenzyloxycarbonyl hydrazide resin (Bachem, Switzerland). The *n*-nonanoyl peptides were deprotected and cleaved from the resin by vortexing with 1.5 mL of a mixture consisting of TFA (95%), 2.5% H_2O, and 2.5% triisopropylsilane (v/v) for 1 h at room temperature and filtered to remove resin. The solution was evaporated under N_2, and the oily material taken up in CH_3CN−water and lyophilized. The resulting *n*-nonanoyl peptides could be used in the conjugation reaction without any further purification.

Synthesis of Aldehyde-Functionalized Reagents and *n*-Nonanoyl Peptides. *n*-Nonanoyl-2,2′-dimethoxyethyl ester was prepared by reaction of *n*-nonanoic acid cesium salt with bromoacetaldehyde dimethylacetal in dimethylformamide (DMF) at 60 °C for 48 h (*12*). *n*-Nonanoyl peptide (25 μmol) dissolved in 200 μL DMF was mixed with a 5-fold excess of aminoacetaldehyde diethylacetal and an equimolar amount of (benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyBOP) reagent. The solution was incubated for 3 h at room temperature, acidified, and the conjugation product isolated by preparative HPLC using a gradient of 30–70% B over 40 min. The acetal group was cleaved by a 5 min treatment in 1 mL neat TFA. The TFA was evaporated under N_2, the *n*-nonanoyl glycoaldehyde ester was redissolved in CH_3CN/ water, freeze-dried, and used immediately for the conjugation step. In the case of the aldehyde derivative of the fully protected peptide, *n*-nonanoyl-PrS-Chg-Ser(tBu)-Ser(tBu)-Asp(OtBu), the protecting groups were removed all at once by a 1 h treatment in neat TFA.

Conjugation through Oxime and Hydrazone Bonds. The truncated proteins with N-terminal serine residues were dissolved in water at 2 mg/mL, mixed with an equal volume of 2% (w/v) NH_4HCO_3 buffer, containing 40 mM methionine, pH 8.3, and oxidized in the presence of a 20-fold excess of sodium metaperiodate. After 10 min incubation in the dark, the reaction was quenched by the addition of a 10 000-fold excess of ethylene glycol. The oxidized protein was retained on a Sep-Pack cartridge, eluted with 80% CH_3CN in 0.1% TFA, and freeze-dried. The protein was redissolved (4 mg/mL) in a 0.1 M acetate (Na) buffer containing 20 mM methionine, 40% CH_3CN, apparent pH 4.6, and either a 2-fold excess of aminooxy-functionalized *n*-nonanoyl peptide or a 6-fold excess of hydrazide-functionalized *n*-nonanoyl peptide was added, dissolved in the same buffer at a concentration of 4 mg/mL. After overnight incubation, the conjugation product was isolated by semipreparative HPLC, using a 1 × 25 cm Macherey-Nagel column at 4 mL/min. In the case of hydrazone conjugates, the hydrazone bond could be reduced by incubation of the recovered and lyophilized material in 400 μL of 0.4 M $NaBH_3CN$ in the same buffer as used previously for 4 days in the dark (*13*), before HPLC purification.

Conjugation through a Thiazolidine Analogue. Unfolded [Cys^2]RANTES(2-68) was dissolved in 0.1 M sodium phosphate, 6 M GnCl, pH 6.0, to a concentration of approximately 5 mg/mL, a 50 molar excess of freshly TFA-deprotected *n*-nonanoyl glycolaldehyde ester was added and the pH readjusted to 5.0. After a 1 h incubation, the conjugation product was isolated by preparative HPLC and freeze-dried. The thiazolidine conjugate was then redissolved in 0.1 M sodium phosphate, 6 M GnCl, pH 7.5, at 4 mg/mL and the solution incubated for 4 days at 37 °C, with daily readjustment of the pH to 7.5. The solution was acidified to pH 3.0 and further incubated with 5 mg tris(2-carboxyethyl)phosphine for 1 h, before isolation of the rearranged thiazolidine analogue by

semipreparative HPLC. The material was then refolded under the same conditions as given above.

Conjugation through a Methylene Amino, Ψ[CH_2−NH] Bond. Truncated RANTES (2 mg/mL in water) was reacted at a molar ratio of 1:10 with 5 mg/mL solution freshly deprotected *n*-nonanoyl aminoacetaldehyde peptide in 80% acetonitrile and mixed with an equal volume of 0.2 M acetate(Na) buffer containing 50% acetonitrile, pH 5.2. The reaction took place at room temperature for 15 h. Sodium cyanoborohydride (1 M) dissolved in the same buffer was added to the reaction to a final cyanoborohydride concentration of 0.3 M. The reaction medium was further incubated at room temperature for at least 6 h before isolation of the conjugation product by semipreparative HPLC.

Conjugation via N-Terminal Acylation of Protein Fragments in Which Lysine Residues Had Been Replaced by Arginines. The complementary short N-terminal peptide, *n*-nonanoyl-PrS-Chg, was activated with NHS by a 1 h incubation with equimolar amounts of HO-Su and EDC in DMSO. The modified protein fragment, [$Arg^{25,45,56,57}$] RANTES-(4–68) (see Table S1 in the Supporting Information), was dissolved in 0.1 M MES buffer containing 50% DMSO, apparent pH 5.5, to a concentration of approximately 4 mg/mL, a 20 molar excess of freshly activated peptide was added, and the solution stirred over 15 h. The solution was then acidified with acetic acid, diluted 10-fold with water, and the conjugation product isolated by semipreparative HPLC using a 1 × 25 cm Macherey-Nagel column at 4 mL/min.

HIV Entry Inhibition Assay. Envelope-mediated cell fusion assays were carried out as described in ref (*14*) with HeLa-P5L (*15*) and HeLa-Env-ADA (*16*) cell lines. Experiments were performed in triplicate, and dose-inhibition curves were fitted by using *Prism* software (GraphPad, CA).

RESULTS

Production of Oxime Analogues. Construction of the oxime analogues was driven by the idea of synthesizing the shortest possible stable linker, functionalized with an amine and an aminooxy group. Our initial strategy involved using a one-carbon linker, which would have entailed the replacement of only two residues and would have kept the backbone of the semisynthetic protein in register with that of PSC-RANTES (Scheme 1). However, these compounds proved inaccessible due to their high instability (data not shown), and it was therefore necessary to investigate linkers with two or more carbon atoms. As shown in Scheme 1, incorporation of a two- or four-carbon linker results in the replacement of two or three residues, respectively. While the two-carbon linker has the advantage of substituting one less residue that the four-carbon linker, it has the disadvantage of introducing a one-atom shift, relative to PSC-RANTES, of the register of the backbone of the N-terminal region of the semisynthetic protein (Scheme 1). Since it is likely that linker position will influence the biological activity of the resulting protein, we opted to generate a panel of analogues for each linker, moving the site of the linker within the N-terminal region of the protein. The nomenclature used for this series of analogues was 2L-*x*,*y* for the two-carbon linker and 4L-*x*,*y* for the four-carbon linker, with *x* and *y* denoting the positions of the first and last substituted residue (Table 1).

We chose to produce the aminooxy-functionalized *n*-nonanoyl peptides through reaction of N-(bromoalkyl)phthalimides with *tert*-butyl *N*-hydroxycarbamates. A 2-fold excess of aminooxy-functionalized peptide was sufficient to drive the conjugation reaction to completion. All intermediates in the construction of the different conjugates were characterized by mass spectrometry and described in Table S1 in the Supporting Information.

Production of Hydrazone and Hydrazide Analogues. We used 2-methoxy-4-alkoxybenzyloxycarbonyl-hydrazide resin to produce a range of *n*-nonanoyl peptides hydrazide-functionalized with a two-carbon linker. These could be reacted with N-terminal Ser or Thr protein fragments that had been oxidized with periodate to generate a range of hydrazone-linked PSC-RANTES analogues. Reduction of the hydrazone bond with $NaBH_3CN$ gave us easy access to the corresponding methyl hydrazine linked analogues, which allow more backbone flexibility at the linkage site. The nomenclature used for this group of compounds is 4L-*x*,*y*H for the hydrazone analogues and 4L-*x*,*y*HR for the reduced methyl hydrazine analogues, with *x* and *y* denoting the positions of the first and last substituted residue (Table 1).

Production of Ψ[CH_2−NH] Analogues. Reductive amination, in which an aldehyde reacts with the N-terminal α-amino group of a target protein in the presence of sodium cyanoborohydride as a reducing agent, has been recently explored as a route to selectively modify proteins at the α-amino group without protecting the ϵ-amino groups on lysine residues (*17–19*). This approach is less restrictive than other established semisynthetic strategies, since it does not require the presence of an N-terminal residue such as Ser, Thr, or Cys. Under only slightly acidic conditions, the α-amino group (pK_a in the region of 8) is deprotonated to a significantly greater extent than the ϵ-amino groups of lysine residues (pK_a in the region of 10). Reaction of a peptide aldehyde with the α-amino group results in the replacement of one residue in the peptide chain by an aminoethylene moiety (Scheme 2).

We opted to test this conjugation approach using the same group of protein fragments that was used to generate the oxime and hydrazone analogues, together with a corresponding group of aminoacetaldehyde-conjugated *n*-nonanoyl peptides. The nomenclature used for this series of analogues was 1L-*x*, with *x* and *y* denoting the positions of the substituted residues (Table 1). It was necessary to change the strategy for the functionalization of one of the *n*-nonanoyl peptides, namely, *n*-nonanoyl-PrS-Chg-Ser-Ser-Asp(OPse)-OH, since the Asp deprotection step carried out under basic conditions resulted in an almost quantitative dehydration of the compound with the formation of an aspartimide. While we opted to retain the *n*-nonanoyl peptide carrying the aspartimide modification and use it to produce an extra analogue (1L-7 *des*, Table 1), it was possible to obtain the desired Asp analogue by using the orthogonal (OtBu) protecting group for the Asp side chain, with the fully protected peptide being cleaved from the 2-chlorotrityl chloride resin with acetic acid.

Not unexpectedly, given the limited selectivity of the approach, each conjugation reaction gave rise to detectable overreaction products resulting from the conjugation of the aminoacetaldehyde-functionalized *n*-nonanoyl peptides with the ϵ-amino groups of the protein (Figure 2).

Production of Thiazolidine Analogues. Proteins carrying an N-terminal cysteine residue can be site-specifically modified at the N-terminus with an aldehyde under acidic conditions (*20*), and the use of a glycolaldehyde ester opens the way to a rearrangement that regenerates a true peptide bond (*21*), although both the condensation and rearranged products give a mixture of two diastereoisomers due to the creation of a new asymmetric carbon at position 2 of the thiazolidine ring. We chose to use this approach to generate two PSC-RANTES analogues. First, conjugation of an *n*-nonanoyl glycolaldehyde ester to [Cys^2]-RANTES(2-68) generates as the final product N^{α}*n*-nonanoyl-[2-hydroxymethythiazolidine^2]RANTES(2-68), an analogue very similar in structure to N^{α}(*n*-nonanoyl)-RANTES(2-68), the molecule known as NNY-RANTES. This was a lead compound, already quite potent in itself, identified during the development of PSC-RANTES (*3*). Second, since the technology developed by the group of Schultz (*22*) enables a single noncoded amino

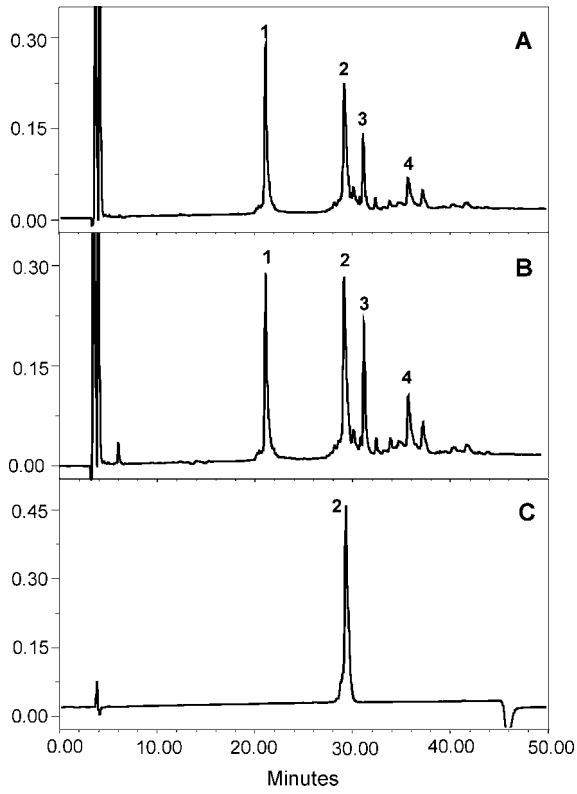


Figure 2. Analytical HPLC of the reductive amination reaction of *n*-nonanoyl-PrS-Chg-Ser^4-NH-CH_2-CHO with [Asp^6]RANTES(6-68). (A) Reaction mixture 2 h after addition of $NaBH_3CN$ following an overnight incubation. (B) Reaction mixture 20 h after addition of $NaBH_3CN$. (C) *n*-Nonanoyl-PrS-Chg-Ser^4-NH-CH_2-CH_2-RANTES(6-68), after semipreparative HPLC purification. Peak 1, [Asp^6]RANTES(6-68); Peak 2, the formed conjugate, *n*-nonanoyl-PrS-Chg-Ser^4-NH-CH_2-CH_2-RANTES(6-68); Peak 3, excess *n*-nonanoyl-PrS-Chg-Ser^4-NH-CH_2-CHO; Peak 4, the overreaction bisubstituted conjugate.

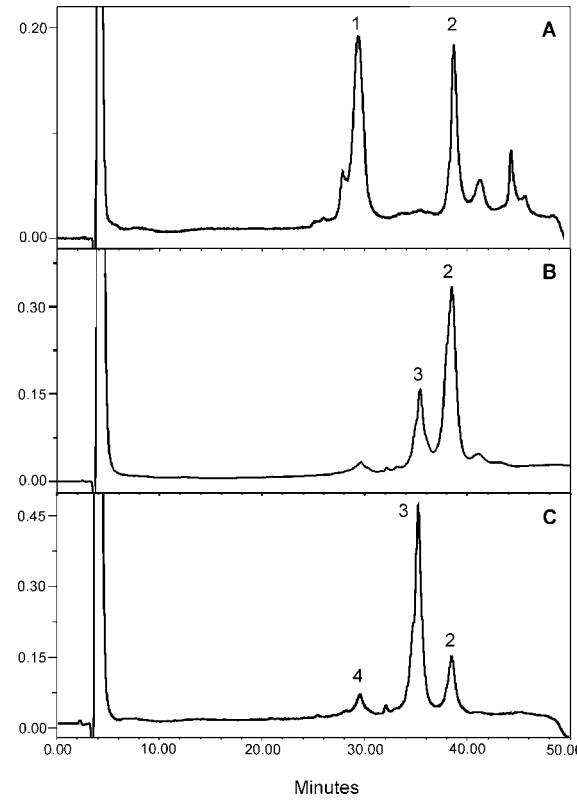


Figure 3. Progress in the construction of *n*-nonanoyl-HMThiaz^2-Chg^3-RANTES(4-68). (A) Reaction of [Cys^2, Chg^3]RANTES(2-68) with *n*-nonanoyl glycolaldehyde ester after 45 min incubation at pH.5.0. (B,C) The O to N acyl transfer at pH 7.5 after 1 and 4 days incubation at 37 °C, respectively. Peak 1, [Cys^2, Chg^3]RANTES(2-68); Peak 2, the thiazolidine conjugate; Peak 3, the O,N acyl rearranged conjugate; Peak 4, the hydrolyzed condensation product [HMThiaz^2, Chg^3]RANTES(2-68).

acid to be site-specifically incorporated into recombinant proteins and works particularly well with aromatic or cyclic residues, we opted to perform the same reaction on [Cys^2, Chg^3]RANTES(2-68), in order to produce the analogue N^{α}(*n*-nonanoyl)-[2-hydroxymethylthazolididine^2, Chg^3]RANTES(2-68), an analogue for which the only structural difference relative to PSC-RANTES is a hydroxymethyl side chain (present in two diastereisomeric forms) on the carbon at position 2 of the thiazolidine ring. If the latter analogue were found to have useful biological properties, the methods of Schultz and co-workers could presumably be adapted to the production of the [Cys^2, Chg^3]RANTES(2-68) as a totally recombinant molecule, onto which, as we have just seen, the *n*-nonanoyl 2-hydroxymethythazoline moiety could be introduced with total site specificity.

For the purposes of this part of the study, the folding of the protein with concomitant formation of disulfide bridges was performed after conjugation of the peptide to enable us to avoid the need to develop the new folding conditions for totally synthetic protein fragments carrying an N-terminal cysteine residue. For both conjugation reactions, incubation with a 50-fold excess of the *n*-nonanoyl glycolaldehyde ester resulted in an HPLC shift of the protein of more than 8 min, as shown in Figure 3A for the [Cys^2, Chg^3]RANTES(2-68). The isolated thiazolidine ester derivatives were resolubilized in a sodium phosphate buffer pH 7.5 to allow O to N acyl rearrangement, as previously described (*21*). This reaction was stopped after 4 days of incubation (Figure 3C), since no improvement could be observed over a longer period of time (data not shown). The incomplete reaction may be related to the presence of contaminating *n*-nonanoyl glycolaldehyde ester, which is difficult to detect and totally eliminate after the first reaction step due to its very low optical absorbance, and which would be highly reactive with the protein's ϵ-amino groups under the more basic conditions of the second reaction. Starting with 15 mg unfolded protein, 5 mg of condensation product were isolated, which finally gave 1 mg of the rearranged compound.

Anti-HIV Activity of PSC-RANTES Analogues. The anti-HIV activity of the full set of PSC-RANTES analogues, together with reference compounds AOP-RANTES, NNY-RANTES, and PSC-RANTES, was determined in a series of cell fusion assays. Each compound was assayed in at least four independent experiments, with PSC-RANTES used as an internal standard in each experiment. Results are shown in Figure 5. The measured potency of PSC-RANTES across the total of 41 independent assays (mean IC_{50} value of 24 pM) is in agreement with previously published values (*3*), as are those for AOP-RANTES and NNY-RANTES (mean IC_{50} values of 0.96 nM and 190 pM, respectively).

Oxime-Based Analogues. In general, the oxime-based analogues with two-carbon linkers were less potent than the corresponding analogues with four-carbon linkers (both 2L-45 and 2L-56 are less potent than 4L-46; both 2L-56 and 2L-67 are less potent than 4L-57; both 2L-67 and 2L-78 are less potent than 4L-68). While a number of the analogues in this series show anti-HIV activity that compares well to the earlier lead molecules, AOP-RANTES and NNY-RANTES, the best analogue in this series, 4L-57 (IC_{50} value of 130 pM), is still approximately 5-fold less potent than PSC-RANTES.

Hydrazone/Methyl Hydrazine-Based Analogues. Each of the hydrazone analogues showed comparable (4L-46H) or lower

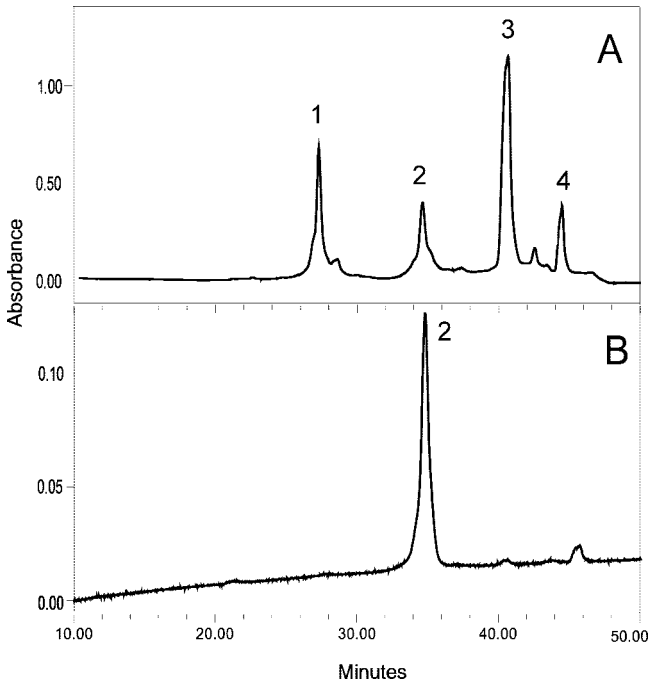


Figure 4. Production of PSC-RANTES[$Arg^{25,45,56,57}$] via direct acylation of [$Arg^{25,45,56,57}$]RANTES(4-68) with *n*-nonanoyl-PrS-Chg. Analytical HPLC of the reaction medium after 15 h incubation at room temperature (A) and after isolation of the conjugation product (B). Peak 1, unreacted [$Arg^{25,45,56,57}$]RANTES(4-68); Peak 2, PSC-RANTES[$Arg^{25,45,56,57}$]; Peak 3, excess *n*-nonanoyl-PrS-Chg. Peak 4 is presumed to be the product of over-reaction, with acylation of both the α- and ϵ-amino groups of the protein fragment.

activity (4L-57H, 4L-68H) to the corresponding four-carbon linker oxime analogues. Reduction to methyl hydrazide did not affect the activity of two of the analogues (4L-57HR, 4L68HR), but strongly reduced the activity of a third (4L-46HR).

Ψ[CH_2−NH] Analogues. The use of the reductive amination approach gave us access to a one-carbon linker strategy, which has the advantages of retaining backbone register while requiring replacement of only one residue. Interestingly, this strategy provided little advantage over the four-carbon linker oxime approach, which also retained backbone register but necessitated the replacement of three residues. While 1L-4 shows slightly higher potency than 4L-46, 1L-5 was less potent than any of the 4L oxime compounds, 1L-6 showed lower potency than 4L-57, and 1L-7 showed lower potency than both 4L-57 and 4L-68.

Thiazolidine Analogues. The protein N^{α}(*n*-nonanoyl)-[HMThiaz^2]RANTES(2-68) shows considerable anti-HIV activity (IC_{50} value of 120 pM), a value comparable to that of our earlier lead compound, NNY-RANTES, to which it has a strong structural resemblance (see Discussion section). When in addition to the substitution of 2-hydroxymethyl thiazolidine for Pro^2 we replaced Tyr^3 by cyclohexylglycine, the potency of the new analogue was less (IC_{50} value of 0.98 nM) by an approximately 40-fold reduction compared to the lead compound PSC-RANTES in spite of the close structural resemblance of the two proteins.

"All-Arg" Analogues. If the protein fragment were to contain no lysine residues, acylation by the short N-terminal fragment would be totally site-specific onto the α-amino group of the protein fragment. This could be achieved by replacing all the lysine residues in the native sequence of the protein fragment (positions 25, 33, 45, 56, and 57) with arginine. Had the properties of the resulting protein [$Arg^{25,33,45,56,57}$]PSC-RANTES been as attractive as those of PSC-RANTES itself, it might well have led to a highly economic production process, in spite of the noncoded residues. However, upon synthesis and testing of this analogue, we found that, despite its having an N-terminal region identical to that of PSC-RANTES, it has an approximately 15-fold reduction in potency relative to PSC-RANTES (Figure 5, IC_{50} value of 380 pM). This prompted us to speculate that one or more of the lysine residues in the protein, although distant in the linear sequence of the protein from its putative pharmacophore, may play an important role in its anti-HIV activity. On the basis of ongoing studies of a different group of PSC-RANTES analogues (*23*), as well as studies in which Lys residues of native RANTES were substituted (*24*), we postulated that Lys^{33} might be a key component of the PSC-RANTES pharmacophore. We found, in agreement with this hypothesis, that [$Arg^{25,45,56,57}$]PSC-RANTES showed potency indistinguishable from that of PSC-RANTES (IC_{50} value of 35 pM). It was therefore desirable to see if, despite the presence of one ϵ-amino group in the relevant protein fragment as well as the α-amino group, acylation by the short N-terminal peptide

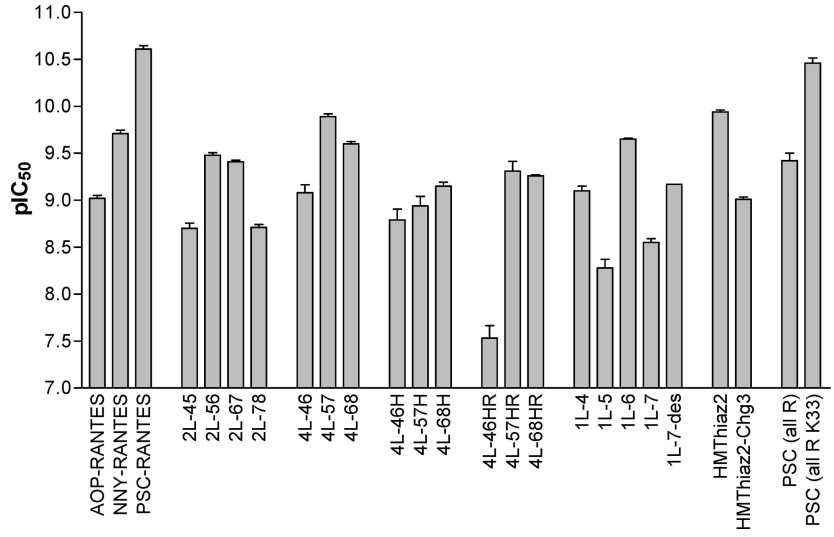


Figure 5. Anti-HIV potency of semisynthetic PSC-RANTES analogues and reference compounds. At least four independent dose–response experiments were performed for each compound using a cell fusion assay, Bars represent the mean potency from the group of experiments performed (pIC_{50}; −log(molar IC_{50} value)), with error bars representing the SEM. See text for the definition of the linker nomenclature. HMThiaz2 = N^{α}(*n*-nonanoyl)-[HMThiaz^2]RANTES(2-68), HMThiaz2-Chg3 = N^{α}(*n*-nonanoyl)-[HMThiaz^2, Chg^3]RANTES(2-68).

could be directed to the α-amino group alone by an appropriate choice of reaction conditions (see below).

Production of Analogues by N-Terminal Acylation of Protein Fragments Having No, Or Few, Lysines. The satisfactory potency of [$Arg^{25,45,56,57}$]PSC-RANTES, mentioned above, led us to consider a pH-controlled acylation of the α-amino group of [$Arg^{25,45,56,57}$]RANTES(4-68) with *n*-nonanoyl-PrS-Chg to produce the full-length construct. The reaction had to be performed in a semiorganic medium to solubilize the highly hydrophobic N-terminal peptide. DMSO turned out to be the most appropriate cosolvent for this reaction, leading to the formation of the wanted product, eluting at approximately 35 min (Figure 4A). The other solvents investigated (DMF, THF) resulted in the formation of a double peak in this region (data not shown). However, even in DMSO, while the higher reactivity of the α-amino group was clearly responsible for a high yield of the wanted product, under the reaction conditions investigated there was some formation of overreaction products, eluting at approximately 44 min (Figure 4A).

DISCUSSION

In this study, we set out to identify strategies with potential for development as economical routes to production of a PSC-RANTES analogue with antiviral activity as close as possible to that of the parent protein. The bulk of the protein would be produced at low cost by biosynthesis prior to attachment to a short N-terminal pharmacophore region that contains the non-natural, noncoded structures and which would be synthesized separately. High-yield production of recombinant RANTES fragments has been demonstrated using *E. coli* with expression methods directed toward inclusion body formation (*24*), and we have successfully used this approach to produce samples of N-terminally truncated derivatives of RANTES (O.H. et al., unpublished results). While expression of proteins via inclusion body formation often leads to incomplete cleavage of the initiating methioine (e.g., ref *25*), this problem can be overcome by the use of N-terminal peptide tags featuring specific enzymatic cleavage sites (e.g., ref *26*). The criteria we chose to adopt when exploring the suitability of the different ligation strategies were as follows: ease and cost of production of the N-terminal fragment, ease and yield of the conjugation reaction, and anti-HIV activity of the resulting analogue. These considerations led us to decide not to explore semisynthesis of PSC-RANTES via native chemical ligation; while it would certainly be a feasible approach, it would necessitate the production of a long and relatively expensive thioester peptide, as well as a folding step after the ligation.

We therefore considered several semisynthesis approaches, some of which necessitate the introduction of a nonpeptide linker region into the protein. Since incorporation of the linker leads to the replacement of backbone and side chain atoms of the protein, and the likelihood of inactivation of a protein increases with the number of substitutions introduced into key regions, we attempted to keep the linkers as short as possible. We varied the site of introduction of the linker, and hence the site at which atoms of the original protein would be changed.

Taken together, study of the three linker-based strategies that we used shows that certain parts of the N-terminal region would appear to be more amenable to the introduction of linker moieties than others. We found overall that substitution at positions 4 and 8 led to a greater reduction in biological activity than substitution at positions 5, 6, and 7. Finally, we did not find that use of a shorter linker necessarily leads to better retention of activity. Indeed, the most potent protein that we constructed with a linker, the oxime-linked analogue 4L-57, has the longest linker region tested, one which requires three residues to be replaced. This four-carbon oxime linker performed generally better than the shorter two-carbon linker that we tested, perhaps because the two-carbon linker modifies the backbone register of the semisynthetic protein with respect to PSC-RANTES.

It was not our purpose in this exploratory study to completely optimize the processes evaluated. In the case of the four-carbon linkers, a useful next step might be to use molecular diversity (library) methods to change the pharmacophore: it is possible that changes there could offset the effect of the linker on activity.

It was possible to directly compare the four-carbon linker oxime linkage strategy with the corresponding one using hydrazones. The activity of the hydrazone analogues was generally lower than that of the corresponding oxime ones, indicating that the activity of the protein is influenced by either the composition or flexibility of the linkers, or both. Reducing the hydrazone linker to methyl hydrazide would be expected to increase linker flexibility, and at certain positions in the protein, this led to a slight increase in activity. In the case of 4L-46HR, however, reduction to a hydrazide led to a more than 10-fold loss in potency. This loss might be related to the fact that in this case the secondary -NH- group in the reduced linker could, when ionized, place a positive charge in proximity to the negative charge of the Asp^6 side chain in the native protein and a strong ionic interaction here might distort the configuration of the whole pharmacophore.

Further evidence that introduction of a positive charge into the backbone can affect activity is provided by results obtained with Ψ[CH−NH] analogues. While these analogues have the shortest linker region that we tested, one which requires the replacement of only one residue, they introduce a full positive charge into the backbone, since reduction of the Schiff base results in a secondary amine with a pK_a in the 10.0–11.0 range. Once again, 1L-5, the analogue in the Ψ[CH_2−NH] series in which the backbone charge is introduced adjacent to Asp^6, shows the greatest reduction in activity. The highest activity in this series of conjugates was obtained for 1L-6, which means that replacement of Asp^6 by an aminoethylene moiety with concomitant alkylation of the nitrogen of Thr^7 is better tolerated. Overall, however, the performance of the Ψ[CH_2−NH] analogues suggests that use of a cumbersome but uncharged linker region does less harm to activity than introduction of a shorter linker carrying a positive charge.

Ligation According to Tam's Thiazolidine Strategy. This method (*21*) yields a native peptide bond in the backbone structure with, however, the introduction of a hydroxymethyl thiazolidine residue immediately C-terminal to the junction site. This residue can be thought of as an analogue of proline (present at position 2 of RANTES and of the earlier lead protein NNY-RANTES) and of thioproline (present at this position in the current lead PSC-RANTES). We therefore used the strategy to incorporate 2-hydroxymethyl thiazolidine at position 2 of NNY-RANTES, and it was encouraging to find that it retains the potency (IC_{50} value of 120 pM) of the parent molecule. We also produced a 2-hydroxymethyl thiazolidine^2 analogue of our best protein lead PSC-RANTES using the same chemistry. This time the protein fragment was [Cys^2-Chg^3]RANTES(2-68), which should be possible to obtain in recombinant form using the Schultz technology (*22*). The conjugation was successful, but the resulting molecule exhibited a significant loss of activity (IC_{50} value of 0.98 nM) compared to the parent PSC-RANTES (IC_{50} value of 24 pM). This apparently paradoxical result recalls previous observations that certain modifications at position 3 of RANTES analogues that are normally beneficial lead to a significant loss in potency when made in the presence of normally beneficial or neutral changes at position 2 (*3*). Thus, PSC-RANTES itself is the result of simultaneously replacing Pro^2 of NNY-RANTES ((*N*^{α}(*n*-nonanoyl)-*des*-Ser^1-RANTES(2-

68)) with thioproline and Tyr^3 by cyclohexylglycine. Substitution of position 2 alone in NNY-RANTES had given only a slight improvement in potency, substitution of position 3 alone had given a 3.4-fold improvement, while in PSC-RANTES substitution of both together gave a 5-fold improvement. However, if position 2 was substituted by 4-hydroxyproline, which also gave no significant change in potency if introduced alone into NNY-RANTES, the simultaneous introduction of cyclohexylglycine in position 3 led neither to an improvement nor to the maintenance of potency, but to an approximately 100-fold decrease (*3*).

All-Arg Analogues. Had [$Arg^{25,33,45,56,57}$]PSC-RANTES (in which all the lysines in the sequence are replaced by arginines) shown comparable activity to PSC-RANTES itself, it would have been possible to use the short N-terminal pharmacophore peptide to acylate with complete specificity the sole amino group of the corresponding protein fragment, namely, the α-amino group. Unfortunately, it showed a significant loss in activity relative to PSC-RANTES. However, [$Arg^{25,45,56,57}$]PSC-RANTES showed activity comparable with that of PSC-RANTES, confirming earlier indications as to the importance of Lys^{33} for the anti-HIV activity of PSC-RANTES (*23*). [$Arg^{25,45,56,57}$]PSC-RANTES therefore represents a potent candidate molecule that could be produced by semisynthesis given sufficient selectivity of acylation between the protein fragment's one α-amino group and its one ϵ-amino group. In an initial practical test of this strategy, we obtained a sufficient degree of selectivity to offer the hope that optimization might lead to a practically useful process. We would not expect the replacement of the four lysines by arginine to increase the immunogenicity of the protein, but this point would need to be investigated experimentally.

We conclude that, in particular, both the four-carbon oxime linker and the direct acylation of [$Arg^{25,45,56,57}$]RANTES (4-68) represent routes to powerful semisynthetic anti-HIV compounds that show potential for optimization.

ACKNOWLEDGMENT

This work was supported by the Mintaka Foundation for Medical Research, Geneva, and the Foundation Dormeur Vaduz. We thank Dr. Reto Stöcklin of Atheris SA (Geneva) for giving us access to the Micromass Platform instrument.

Supporting Information Available: Supplementary Table S1. This material is available free of charge via the Internet at http://pubs.acs.org.

LITERATURE CITED

(1) Who, U. A. (2006) *AIDS Epidemic Update*, UN AIDS/World Health Organisation, Geneva, Switzerland.

(2) Lederman, M. M., Offord, R. E., and Hartley, O. (2006) Microbicides and other topical strategies to prevent vaginal transmission of HIV. *Nat. Rev. Immunol. 6*, 371–382.

(3) Hartley, O., Gaertner, H., Wilken, J., Thompson, D., Fish, R., Ramos, A., Pastore, C., Dufour, B., Cerini, F., Melotti, A., Heveker, N., Picard, L., Alizon, M., Mosier, D., Kent, S., and Offord, R. (2004) Medicinal chemistry applied to a synthetic protein: development of highly potent HIV entry inhibitors. *Proc. Natl. Acad. Sci. U.S.A. 101*, 16460–5.

(4) Lederman, M. M., Veazey, R. S., Offord, R., Mosier, D. E., Dufour, J., Mefford, M., Piatak, M., Jr., Lifson, J. D., Salkowitz, J. R., Rodriguez, B., Blauvelt, A., and Hartley, O. (2004) Prevention of vaginal SHIV transmission in rhesus macaques through inhibition of CCR5. *Science 306*, 485–7.

(5) Moore, J. P. (2005) Topical microbicides become topical. *N. Engl. J. Med. 352*, 298–300.

(6) Dawson, P. E., Muir, T. W., Clark-Lewis, I., and Kent, S. B. (1994) Synthesis of proteins by native chemical ligation. *Science 266*, 776–9.

(7) Wilken, J., and Kent, S. B. (1998) Chemical protein synthesis. *Curr. Opin. Biotechnol. 9*, 412–26.

(8) Wilken, J., Hoover, D., Thompson, D. A., Barlow, P. N., McSparron, H., Picard, L., Wlodawer, A., Lubkowski, J., and Kent, S. B. (1999) Total chemical synthesis and high-resolution crystal structure of the potent anti-HIV protein AOP-RANTES. *Chem. Biol. 6*, 43–51.

(9) Carrasco, M. R., Brown, R. T., Serafimova, I. M., and Silva, O. (2003) Synthesis of N-Fmoc-O- (N′-Boc-N′-methyl)-aminohomoserine, an amino acid for the facile preparation of neoglycopeptides. *J. Org. Chem. 68*, 195–7.

(10) Villain, M., Gaertner, H., and Botti, P. (2003) Native chemical ligation with aspartic and glutamic acids as C-terminal residues: scope and limitations. *Eur. J. Org. Chem. 2003*, 3267–3272.

(11) McNamara, J. F., Lombardo, H., Pillai, S. K., Jensen, I., Albericio, F., and Kates, S. A. (2000) An efficient solid-phase strategy for the construction of chemokines. *J. Pept. Sci. 6*, 512–8.

(12) Wang, S. S., Gisin, B. F., Winter, D. P., Makofske, R., Kulesha, I. D., Tzougraki, C., and Meienhofer, J. (1977) Facile synthesis of amino acid and peptide esters under mild conditions via cesium salts. *J. Org. Chem. 42*, 1286–90.

(13) Gaertner, H. F., Offord, R. E., Cotton, R., Timms, D., Camble, R., and Rose, K. (1994) Chemo-enzymic backbone engineering of proteins. Site-specific incorporation of synthetic peptides that mimic the 64–74 disulfide loop of granulocyte colony-stimulating factor. *J. Biol. Chem. 269*, 7224–30.

(14) Sabbe, R., Picchio, G. R., Pastore, C., Chaloin, O., Hartley, O., Offord, R., and Mosier, D. E. (2001) Donor- and ligand-dependent differences in C-C chemokine receptor 5 reexpression. *J. Virol. 75*, 661–71.

(15) Simmons, G., Clapham, P. R., Picard, L., Offord, R. E., Rosenkilde, M. M., Schwartz, T. W., Buser, R., Wells, T. N. C., and Proudfoot, A. E. (1997) Potent inhibition of HIV-1 infectivity in macrophages and lymphocytes by a novel CCR5 antagonist. *Science 276*, 276–9.

(16) Pleskoff, O., Treboute, C., Brelot, A., Heveker, N., Seman, M., and Alizon, M. (1997) Identification of a chemokine receptor encoded by human cytomegalovirus as a cofactor for HIV-1 entry. *Science 276*, 1874–8.

(17) Basu, A., Yang, K., Wang, M., Liu, S., Chintala, R., Palm, T., Zhao, H., Peng, P., Wu, D., Zhang, Z., Hua, J., Hsieh, M. C., Zhou, J., Petti, G., Li, X., Janjua, A., Mendez, M., Liu, J., Longley, C., Zhang, Z., Mehlig, M., Borowski, V., Viswanathan, M., and Filpula, D. (2006) Structure-function engineering of interferon-beta-1b for improving stability, solubility, potency, immunogenicity, and pharmacokinetic properties by site-selective mono-PEGylation. *Bioconjugate Chem. 17*, 618–30.

(18) Lee, H., Jang, I. H., Ryu, S. H., and Park, T. G. (2003) N-terminal site-specific mono-PEGylation of epidermal growth factor. *Pharm. Res. 20*, 818–25.

(19) Nie, Y., Zhang, X., Wang, X., and Chen, J. (2006) Preparation and stability of N-terminal mono-PEGylated recombinant human endostatin. *Bioconjugate Chem. 17*, 995–9.

(20) Zhang, L., and Tam, J. P. (1996) Thiazolidine formation as a general and site-specific conjugation method for synthetic peptides and proteins. *Anal. Biochem. 233*, 87–93.

(21) Liu, C. F., and Tam, J. P. (1994) Peptide segment ligation strategy without use of protecting groups. *Proc. Natl. Acad. Sci. U.S.A. 91*, 6584–8.

(22) Wang, L., Xie, J., and Schultz, P. G. (2006) Expanding the genetic code. *Annu. Rev. Biophys. Biomol. Struct. 35*, 225–49.

(23) Gaertner, H., Lebeau, O., Borlat, I., Cerini, F., Kuenzi, G., Melotti, M., Fish, R., Offord, R., Springael, J.-Y., Marc Parmentier, M., and Hartley, O. (2008) *Protein Eng. Des. Sel.*, in press.

(24) Martin, L., Blanpain, C., Garnier, P., Wittamer, V., Parmentier, M., and Vita, C. (2001) Structural and functional analysis of the RANTES-glycosaminoglycans interactions. *Biochemistry 40*, 6303–18.

(25) Proudfoot, A. E., Power, C. A., Hoogewerf, A. J., Montjovent, M. O., Borlat, F., Offord, R. E., and Wells, T. N. (1996) Extension of recombinant human RANTES by the retention of the initiating methionine produces a potent antagonist. *J. Biol. Chem. 271*, 2599–603.

(26) Shin, N. K., Kim, D. Y., Shin, C. S., Hong, M. S., Lee, J., and Shin, H. C. (1998) High-level production of human growth hormone in *Escherichia coli* by a simple recombinant process. *J. Biotechnol. 62*, 14351.

BC7003044

Angela Saporito[1]
Daniela Marasco[1]
Angela Chambery[2]
Paolo Botti[3]
Simona M. Monti[1]
Carlo Pedone[1,4]
Menotti Ruvo[1]

[1] *Istituto di Biostrutture e Bioimmagini del CNR, Sezione Biostrutture, NAPOLI, via Mezzocannone 16, 80134 Napoli, Italy*

[2] *Dipartimento di Scienze della Vita, Seconda Università di Napoli, via Vivaldi 43, I-81100, Caserta, Italy*

[3] *Department of Molecular Medicine, University of Geneva 1, rue Michel Servet, 1211, Geneva, Switzerland*

[4] *Dipartimento delle Scienze Biologiche, Università Federico II, via Mezzocannone, 16, 80134 Napoli, Italy*

The Chemical Synthesis of the GstI Protein by NCL on a X-Met Site

Received 12 May 2006; revised 10 July 2006; accepted 1 August 2006
Published online 3 August 2006 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bip.20582

Abstract: *The small GstI protein (63 amino acids) of* Rhizobium leguminosarum *is the endogenous inhibitor of the* glnII *(glutamine synthetase II) gene expression. It has been suggested that GstI has a predominantly* β*-structure and mediates the block of translation and stabilization of* glnII *mRNA through direct binding to its 5′ untranslated region. Because of the unavailability of adequate amounts of purified recombinant protein, the mechanism as well as the protein tridimensional structure remain very poorly understood. To obtain the full-length protein, we have undertaken the chemical synthesis of the protein by different approaches. In a first attempt, the stepwise synthesis was unsuccessful, with strong aggregation experienced on the N-terminal side, after residue 44 from the C-terminus. In a second approach, we set up the conditions to carry out a native chemical ligation (NCL). Albeit the protein contains two Cysteine residues, located at positions 40 and 47, to minimize the size of the N-terminal segment to be synthesized, we have devised an alternative strategy of ligation on Met32, utilizing homoCys as the ligating moiety and then alkylating*

Correspondence to: M. Ruvo; e-mail: menotti.ruvo@unina.it
A. Saporito and D. Marasco contributed equally to this work.
This article includes supplementary material available at http://www.interscience.wiley.com/jpages/0006-3525/suppmat

Biopolymers, Vol. 83, 508–518 (2006)

the resulting polypeptide with methyl iodide. New conditions to quantitatively methylate thiol groups in complex polypeptides have been conceived, obtaining the protein in very good yields and purity. A CD spectroscopy investigation has revealed that the protein does not adopt canonical secondary structures but is very rich in β-structure (~ 60%), in agreement with a previous study carried out on samples obtained by recombinant methods.

This article was originally published online as an accepted preprint. The "Published Online" date corresponds to the preprint version. You can request a copy of the preprint by emailing the Biopolymers editorial office at biopolymers@wiley.com

Keywords: *native chemical ligation; homocysteine; peptide synthesis; circular dichroism*

INTRODUCTION

The small protein Glutamine synthetase translational Inhibitor (GstI, 63 amino acids) of *Rhizobium leguminosarum* inhibits the expression of the *glnII* (glutamine synthetase II) gene, providing one of the control mechanisms of bacterial ability to assimilate nitrogen.[1] A primary control is determined by the nitrogen transcriptional regulator, NtrC,[2] that directly activates *glnII* and represses *gsti*, which is found upstream of *glnII* and is transcribed divergently with respect to *glnII*. It has been suggested that GstI mediates a block of translation and stabilization of *glnII* mRNA through direct binding to its 5′ untranslated region, but this mechanism remains to be elucidated.[1,2] Remarkably, when expressed in the heterologous background of *Escherichia coli*, the protein appeared particularly sensitive to proteolytic degradation in the very N-terminal part, suggesting a high chain flexibility and thus a poorly defined structure in this protein region. A mutational analysis (alanine-scanning) carried out on the entire protein sequence has also indicated the residues essential for *glnII* expression inhibition, which are T3, T11, Y13, Q14, F15, V17, K18, R21, N22, M32, E33, E35, Y49, H50, R55, and S63.[3] Recombinant protein preparations from *E. coli* have been obtained by a combination of diafiltration and gel filtration techniques in the presence of a cocktail of protease inhibitors and a preliminary analysis has shown that GstI has a monomeric, prevailingly β-structure, as described for other single-stranded RNA binding motifs.[3]

The reported difficulties in obtaining the full-length protein by recombinant methods and the need for adequate amounts to carry out new investigations on protein structure and activity have led us to undertake the chemical synthesis of the protein by different approaches. In a first attempt, the stepwise synthesis were unsuccessful, with strong aggregation experienced after residue 44 from the C-terminus. In a second approach, we set up the conditions to carry out a native chemical ligation (NCL).[4] NCL is a very elegant and straightforward approach that has enormously broadened the benefits of solid phase synthesis,[5,6] opening the route to the preparation of small to medium size proteins by chemical methods.[6] Albeit the protein contains two Cysteine residues, they are located in position 40 and 47; therefore, to minimize the size of the N-terminal segment that had to be synthesized, we devised an alternative strategy of ligation on Met32 utilizing homoCys as the ligating moiety and then methylating the resulting polypeptide with MeI.[7,8] Thus far, homocysteine has been rarely utilized in ligation reactions,[7] thus kinetic and thermodynamic details are poorly understood. However, the presence of an additional methylene group would probably affect reaction rates and yields, as the six-membered ring of the intermediate homocysteinyl thioester[4,6] could less favorably undergo the proximity-driven S- to N-acyl migration to form a homocysteinyl amide bond. We have prepared the protein that was then characterized by analytical techniques to assess the formation of the desired amino acid and by CD spectroscopy to investigate the secondary structure of the native protein.

RESULTS AND DISCUSSION

Protein Synthesis and Characterization

A strategy for the chemical synthesis of the small, 63-amino-acid-long GstI protein (Figure 1) has been designed to overcome the difficulties associated with the heterologous expression in bacterial hosts. Though the synthesis of a 63-residue polypeptide falls well within the actual limits of the stepwise solid phase method, we observed an extensive aggregation after assembling the first 44 amino acids starting from the C-terminus (see Supplementary Material S1). Aggregation persisted even when coupling and deprotection time were extended and when a variety of common coupling reagents and solvents were used (not shown).

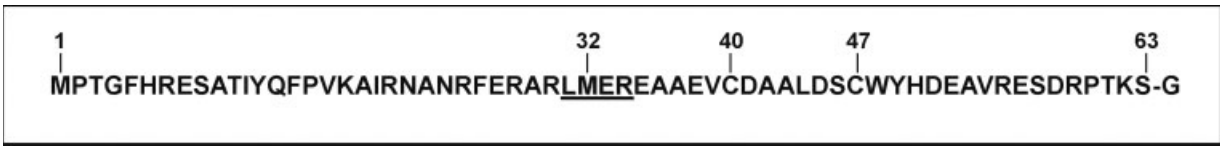
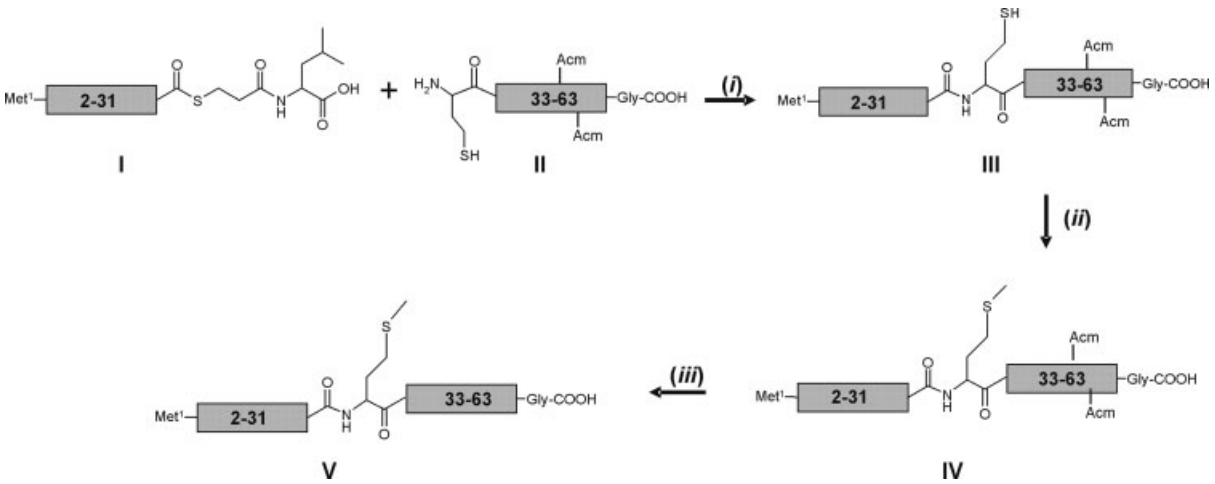


FIGURE 1 Primary sequence of the 63 amino acid protein GstI. The methionine 32 lies exactly in the middle of the sequence. The peptide 31–34, obtainable by trypsin digestion and containing the internal methionine is underlined. Cysteines 40 and 47, protected with Acm until the last synthetic step, are also shown. A C-terminal glycine has been added for synthetic convenience.

The new strategy, depicted in Scheme 1, contemplates the chemical synthesis of two fragments of similar size, which are then ligated using the NCL method on the Leu31-Met32 site. The ligation reaction is achieved by means of a N-terminal homoCys residue that is subsequently converted into a native methionine by an alkylation reaction with methyl iodide.[7] Cysteines on positions 40 and 47 have been opportunely protected with acetamidomethyl (Acm) until alkylation on position 32 is achieved. homoCys32-GstI [32–63]Gly-Cys(Acm40,47) (product II in Scheme 1) has been obtained by coupling fluorenylmetoxycarbonyl (Fmoc)-L-homoCys(Trt)-OH to GstI[33–63]Gly-Cys(Acm40,47). A glycine residue has been added to the C-terminal end of the polypeptide to reduce the risks of racemization by coupling the first amino acid to resin. After Fmoc removal and cleavage from the resin, the polypeptide has been purified to homogeneity and characterized by LC–MS before its use in the subsequent reactions. Fragment 1–31 has been successfully prepared as a C-terminal thioester by solid phase peptide synthesis following several alternative approaches. Direct synthesis on a TAMPAL resin[9–11] by Boc chemistry, followed by RP–HPLC purification, provided a highly pure product with the expected MW (Exp./Theor.: 4048.2 ± 0.2 /4047.9 amu). Ethanethiol or Fmoc-Cys-NH_2 C-terminal thioesters of the same fragment were also successfully prepared using an Fmoc chemistry-based approach. In this case, the fully protected 1–31 peptide was produced using a chlorotrityl-chloride resin and quantitatively converted in solution to the corresponding thioesters, in about 12 h, by using PyBOP/diisopropylethylamine (DIEA) in DCM and a large excess of thiols.[12] Conversion of carboxylic acids into the correponding thioesters using these conditions reportedly occurs with low levels of racemization.[12,13] These thioesters showed reactivities toward the homocysteine on the N-terminus of the 32–63 fragment that are comparable to those of the thioester I (not shown). However, given the availability of larger amounts, this latter derivative was used to produce the ligated protein. GstI[1–31]-COS$(CH2)_2$-CO-Leu-COOH (I) and homoCys32-GstI[1–63]Gly-Cys(Acm40,47) (II) reacted in about 16 h (see Figure 2) to give the expected 64-residue-long polypeptide, homoCys32-GstI[1–63]Gly-Cys(Acm40,47) (III), with an average final 60% conversion yield based on the initial limiting thioester reagent. A reaction side product was detected in this



SCHEME 1 Synthetic steps required for the assembling of GstI[1-63]-Gly. (I) 200 m*M* phosphate buffer, 6 *M* guanidinium hydrochloride, 1% thiophenol, pH 7, 12 h, RT. (II) 100 m*M* bicarbonate buffer, 10% DMF, 5 m*M* TCEP, 0.1 *M* MeI; 1 min, RT, then 1 *M* DTT added. (III) AgOTf/HCl/DTT.

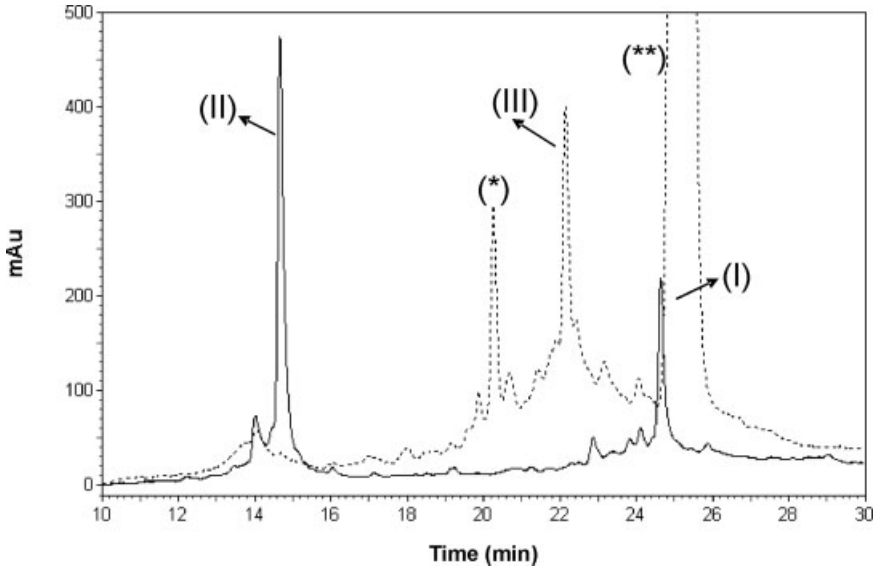


FIGURE 2 HPLC analysis of the ligation reaction between the thioester (**I**) and the homoCys derivative (**II**) at $t = 0$ (solid line) and after 16 h (dashed line). The peak indicated by (*) is due to the transthioesterification of (**I**) with β-Me. The peak indicated as (**) is due to the thiophenol.

reaction, whose MW corresponds to unreacted GstI[1–31] β-mercaptoethanol (β-Me) thioester that forms upon addition of the reducing agent. The presence after 16 h of unreacted GstI[1–31] suggests that the ligation reaction with the homocysteine is less efficient than with cysteine. Indeed, the same reaction carried out using Cys32-GstI[32–63]-Gly went to completion within the same time interval (see Supplementary Material S2). Starting from about 2 μmol of thioester, about 7 mg of crude ligated polypeptide were recovered, which, after purification, were further reduced to about 4 mg. The purified polypeptide was subsequently alkylated to obtain the native Met32 protein with the two Acm protections still attached. Previous reports have suggested the use of reagents such as methyl *p*-nitrobenzene sulphonate or the same methyl iodide utilized in this work. Nevertheless, following these approaches[7,8] to our polypeptide, we have observed a large set of side products derived by unwanted hypermethylation on other reactive residues. In our hands, neither the full-length polypeptide nor other less complex model peptides underwent a satisfactory methylation under the suggested conditions.[7,8] Indeed, the starting compounds were recovered mostly hypermethylated after the treatment. Under the newly devised conditions, we made use of the highly reactive methyl iodide in very large excess over the thiol group, which was maintained in the fully reduced state by the presence of the thiol-free Tris(2-carboxyethyl) phosphine (TCEP) reagent. In 1 min, the reaction was complete and the polypeptide appeared not to be contaminated by side products. By monitoring the reaction for the subsequent 30 min, we observed that, up to 5 min, no side reactions occurred. Thereafter, other +14 amu and +15 amu derivatives were readily produced, corresponding to undesired methylations. The +14 amu side products possibly derived from methylation of lysines and appeared more hydrophobic upon RP–HPLC analysis, whereas, the +15 derivatives, likely due to oxidation of methionines, were eluted slightly earlier. In our conditions, that is, 100 m*M* carbonate buffer, pH 8.0, containing 5 m*M* TCEP, the methylation with MeI proceeded smoothly in about 1 min with quantitative conversion to the desired product (see Figure 3A–C). A very clean product, with the expected MW was detected by LC–MS analysis (Exp./Theor.: 7530.8 $\pm$ 0.5/7530.47 amu). The highly reactive reagent was then readily quenched by the addition of a large excess of DTT. Attempts to carry out the reaction with a smaller excess of MeI demonstrated that the conversion required extended reaction times leading, by consequence, to many side products. The polypeptide was purified to homogeneity by RP–HPLC (3 mg recovered) and submitted to an extensive characterization by trypsin digestion and chemical degradation by CNBr. The addition of trypsin should provide a short tetrapeptide, corresponding to GstI[31–34] (sequence: LMER, see Figure 1) containing the rebuilt methionine. Conversely, CNBr treatment leads to specific peptide bond cleavage on Met-X sites, with the concurrent production of a C-terminal homoserinelattone moiety in place of the original methionine. After trypsin treatment, the

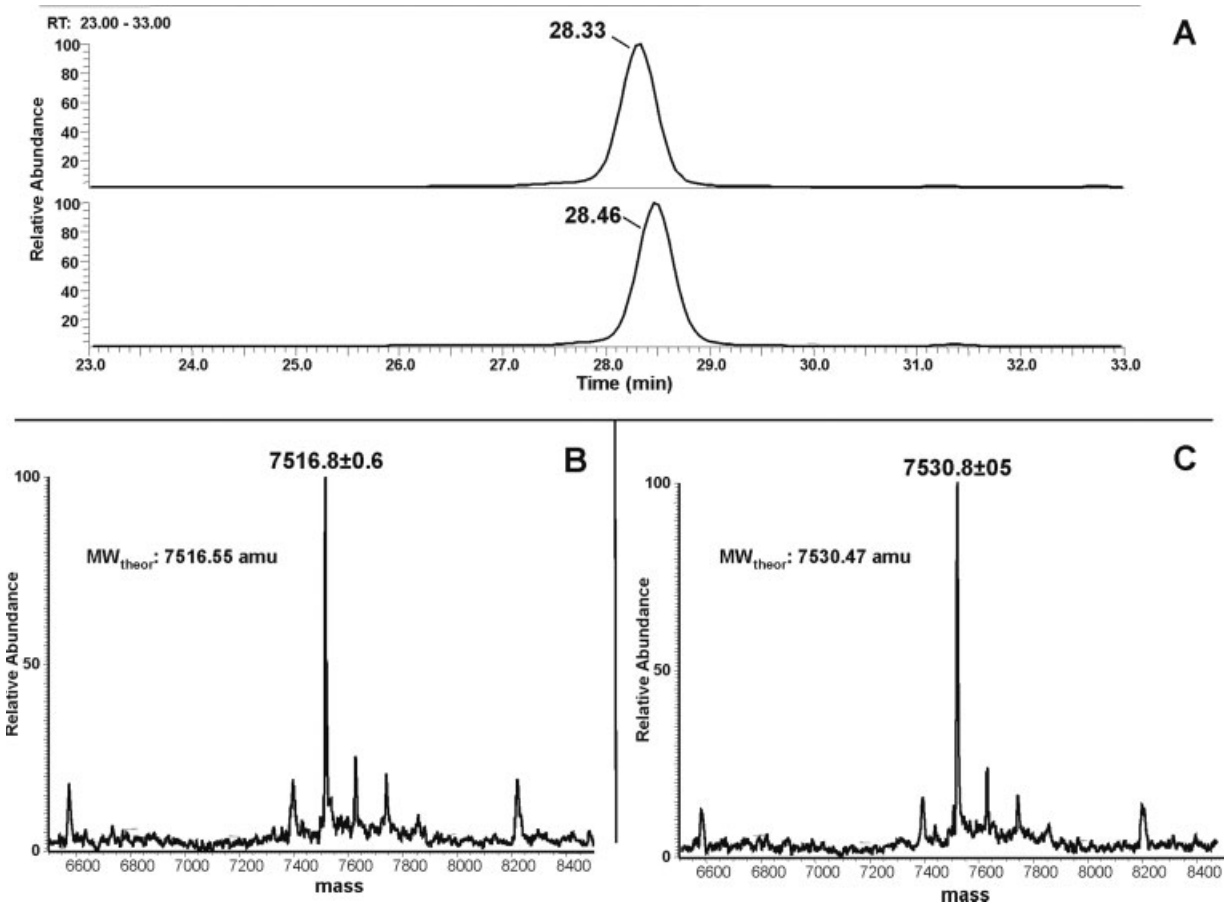


FIGURE 3 (A) LC–MS monitoring of the methylation reaction on polypeptide (**III**). The reaction is complete after 1 min. (B and C) The deconvoluted mass spectra of peaks at Rt 28.33 and 28.46 min are reported.

mixture was analyzed by LC–MS shown in the chromatogram reported in Figure 4A, where, as indicated, the expected fragments were detected. For the tetrapeptide LMER we determined an experimental MW of 547.42 amu, in very good agreement with the expected 547.28 amu value. The contextual MS/MS fragmentation provided the sequence of the expected tetrapeptide (see Supplementary Material S3). The chemical cleavage by CNBr provided two main products with the expected MW (Figure 4B). The presence of the tryptic tetrapeptide and the expected fragmentation by CNBr strongly supported the fact that methylation occurred only on the desired side chain thiol of the homocysteine in position 32.

To produce the fully deprotected protein, treatment with silver trifluoromethanesulfonate (AgOTf)/HCl/DTT was carried out, bringing some changes to the reported method of Tamamura et al.[14]

Following this procedure, we could not recover any product; indeed, after the HCl/dimethylsulfoxide (DMSO) treatment, we believe that most product was still complexed to Ag^+ ions, remaining trapped within the AgCl precipitate. In this instance, DTT is particularly useful, as it can also revert the oxidation of methionines[15] that typically occurs in this reaction. Under the new conditions, the polypeptide was almost quantitatively recovered in the DTT-containing fraction, though in the reduced form. After protein purification, about 1.5 mg of pure final product was recovered and characterized by LC–MS, observing the expected 144 amu loss due to the Acm protections removal. The final product (V) had a shorter retention time compared with the starting product (IV) (Figure 5A) and a purity level higher than 90% by RP–HPLC. The experimental MW (Figure 5B) was in very good agreement with the theoretical value of 7388.31 amu (average MW) calculated assuming the cysteine in the reduced state.

CD Spectroscopy

A CD analysis of the fully reduced polypeptide in phosphate buffer was carried out to determine the

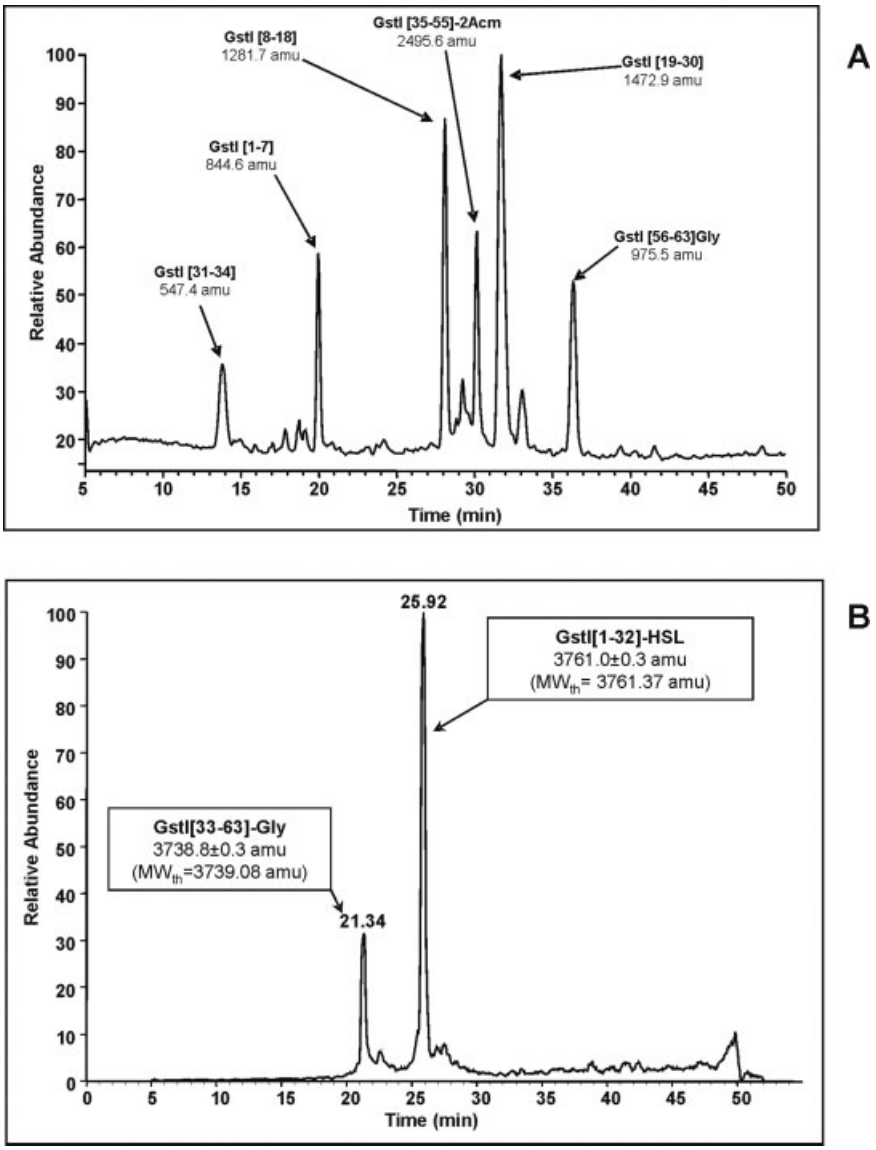


FIGURE 4 (A) LC–MS analysis of the peptide mixture obtained following trypsin treatment of the synthetic GstI[1–63]Gly-Cys(Acm40,47). Peptide fragments and MW are indicated. MW have been calculated all as monoisotopic. MS/MS has also been carried out on the eluted peptides (see Supplementary Material S3). (B) HPLC analysis of the mixture after CNBr treatment. Two main products are detected with the MW expected by a splitting on the Met^{32}-Glu^{33} bond and formation of a homoserine lactone (HSL) moiety on the carboxyl side of the methionine.

global protein fold under these conditions and to perform a comparison with previous results obtained on the recombinant protein. In Figure 6, a comparison of spectra obtained on the full-length protein and on the two precursor peptides is reported. As shown, the two fragments displayed spectra characteristic of random coils, with single minima at about 195 nm, while the protein spectrum, though not fitting any canonical structure, exhibited a strong minimum at about 205 nm and a second weak band at around 222 nm, indicative of some higher conformational preferences of the full-length polypeptide. Remarkably, within the 190–240 nm spectral region, the CD curve of synthetic GstI[1–63]-Gly is very similar to that reported by Napolitani et al.[3] for the recombinant polypeptide, acquired under the same buffer conditions, suggesting that they adopt very similar, prevailingly β-structure conformations. Indeed, upon deconvolution of the CD curve, we found the following composition of secondary structure: $\alpha = 15.3\%$, $\beta = 60.7\%$, random coil $= 24.0\%$, that, unless for a modest α-helix contribution, is comparable to that reported for the recombinant protein ($\alpha = 0\%$;

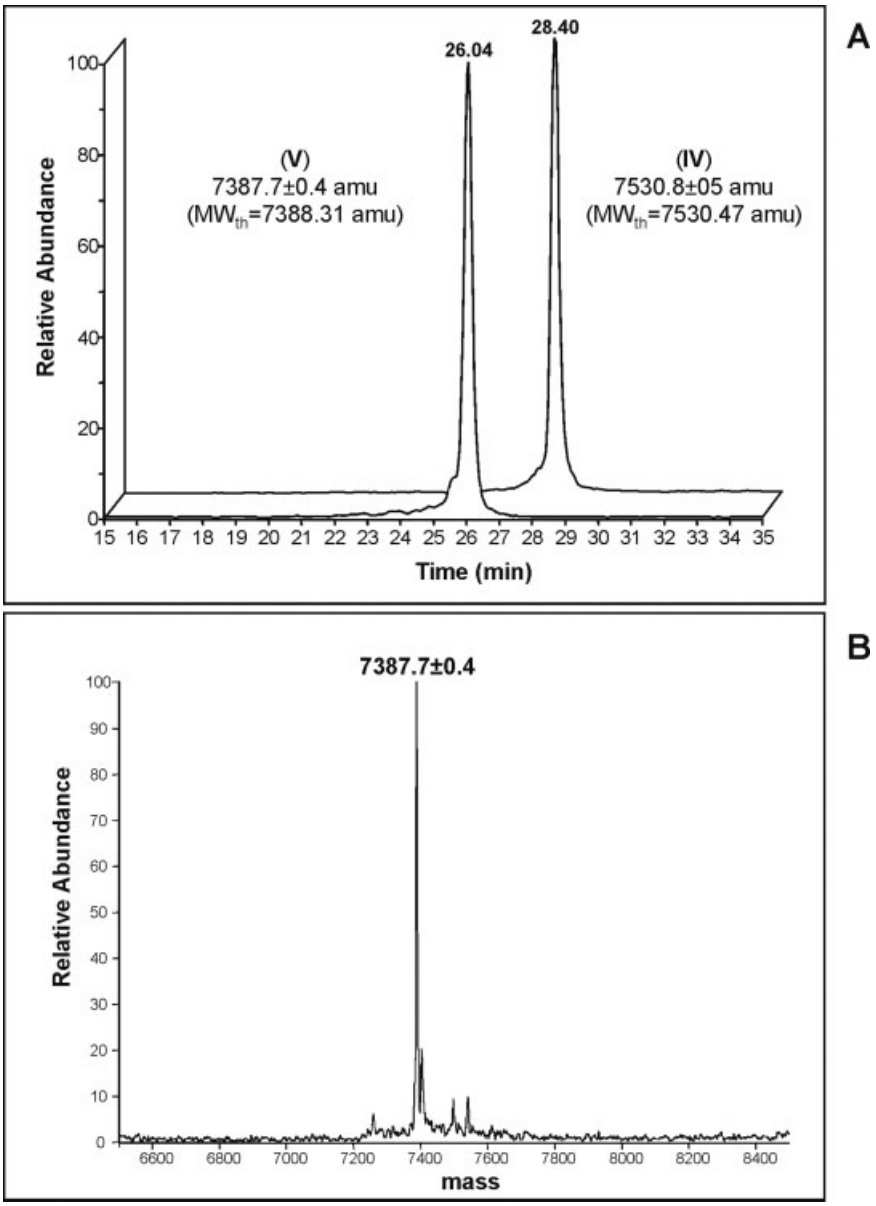


FIGURE 5 Acm removal reaction on product (**IV**) to obtain the final protein (**V**). (A) An overlay of chromatograms before and after treatment. (B) Deconvoluted mass spectrum relative to the main peak. The polypeptide is in the reduced form.

$\beta = 58.7\%$; coil $= 41.3\%$, see reference[3]). The entire analysis has been carried on the reduced protein, as, though cysteines oxidation status is currently not known in vivo, they reportedly do not have any involvement in GSII inhibition. Indeed, mutation of Cys40 into alanines does not alter the protein activity and mutation of Cys47 only slightly affects the GstI inhibitory activity.[3] Furthermore, while Cys47 is conserved between *Rhizobium leguminosarum* (Rl)-GstI and *Sinorhizobium meliloti* (Sm)-GstI and Cys40 is shifted in position, no cysteine residues are present in *Agrobacterium tumefaciens* (At)-GstI, suggesting that, if these proteins fulfill the same function,[3] cysteines should be reduced in (Rl)-GstI.

CONCLUSION

The chemical synthesis of the small protein GstI by an NCL approach on a X-Met site is described in this report. The synthesis has been successful, observing a 10% overall yield (0.2 μmoles over the 2 μmoles of the starting thioester). Though the synthetic strategy has already been described, the key steps of thiol

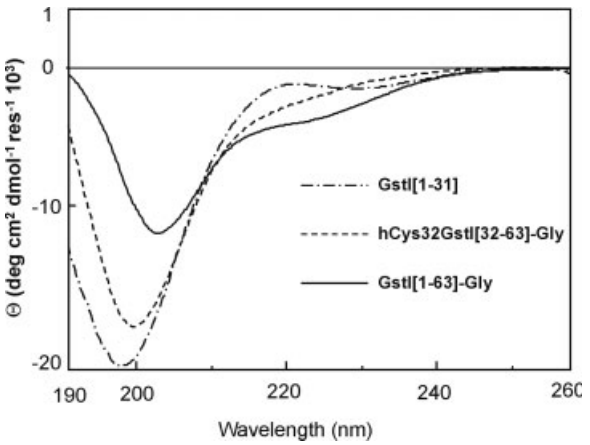


FIGURE 6 CD characterization of synthetic GstI[1–63]-Gly. Spectra relative to the starting 1–31 and 32–63 polypeptides are also reported. Spectra have been recorded between 260 and 190 nm.

methylation and Acm removal have been optimized. The final protein, in the reduced form, has also been preliminarly characterized by CD spectroscopy to assess the presence of secondary structures. The molecule is fully soluble and does not adopt canonical structure in aqueous buffers at neutral pH, but deconvolution of the CD curve has shown that the protein adopts a prevailingly β-structure conformation, according to a previous report.[3] These results suggest that the synthetic protein, as obtained through this process, has structural properties very similar to those described for the recombinant protein.

EXPERIMENTAL

Materials

Fmoc-Gly-4-hydroxymethylphenoxyacetic (HMP)-derivatized polystyrene resin (PS) and chlorotrytilchloride (Cl-TrtCl) resins were purchased from NovaBiochem (Laufelfingen, Switzerland). 9-Fluorenylmethoxycarbonyl (Fmoc)- or t-Butyloxycarbonyl (t-Boc)–derivatized amino acids (purity > 99%) and activating agents such as 1-*H*-Benzotriazolium, 1-[bis(dimethylamino)methylene]-hexafluorophosphate(1-), 3-oxide (HBTU) and *N*-hydroxybenzotriazole (HOBt) were from Inbios (Pozzuoli, Italy) and Novabiochem (Laufelfingen, Switzerland). Fmoc-L-homoCys(Trt)-OH was from Anaspec (San Jose, CA, USA). Synthesis-grade dichloromethane (DCM), dimethylformamide (DMF), methanol, trifluoroacetic acid (TFA), diethylether, and HPLC-grade water and acetonitrile were from LabScan (Dublin, Ireland). PyBOP was from Applied Biosystem (Foster City, CA, USA). Reagents such as tri-isopropylsilane, ethyl sulphide, thiophenol, methyl iodide, piperidine and DIEA were all from Sigma–Aldrich (Milan, Italy). HPLC columns were from Phenomenex (Torrance, CA, USA) and from ThermoElectron (Milan, Italy). The LC–MS systems equipped with ESI sources were from ThermoElectron (Milan, Italy).

METHODS

Synthesis of GstI[1–63]-Gly by NCL. The C-terminal fragment was prepared by automatic solid-phase synthesis using a 433A Applied Biosystems automatic synthesizer following the Fmoc methodology on a 100 μmoles scale. 20% Piperidine in DMF was used as Fmoc deprotecting agent in all steps. Activation was performed using a 0.9 HBTU–HOBt and 2.0 DIEA molar excess over amino acid (1.0 mmol). The C-terminal fragment was assembled from Glu33 to Ser63 on an Fmoc–Gly–HMP resin (substitution 0.53 mmol/g) to avoid racemization of the C-terminal amino acid, thus the resulting fragment GstI(33–63) contained an additional C-terminal glycine. On the resulting resin Fmoc-L-homoCys(Trt)-OH was coupled (1 h, RT) using a fivefold excess and PyBOP/DIEA (50% DCM/DMF v/v) as activating agents. After Fmoc removal, the resin was dried under vacuum. The fragment, homoCys32Met-GstI[32–63]-Gly (polypeptide **II**) was cleaved using a 94:3:3 v/v/v TFA–H_2O–tri-isopropylsilane (TIS) mixture (2.0 mL/100 mg resin, 1.5 h reaction time at RT), precipitated in cold ethyl ether (Et_2O), and lyophilized from a 50% H_2O/CH_3CN solution. 20 mg of crude polypeptide, were again dissolved in 25% H_2O/CH_3CN containing 0.1% TFA and 50 m*M* DTT, incubated at 37°C for 2 h, and then purified by RP–HPLC on a COMBI C18 5 × 2.1 cm ID column (10 μm, Phenomenex) using a gradient of CH_3CN, 0.1% TFA (Solvent B; Solvent A was H_2O, 0.1% TFA) from 20 to 45% over 15 min (flow 20 mL/min) and monitoring at 210 nm. Preparative purifications were carried out on a LC8 Shimadzu HPLC system (Shimadzu Corporation, Milan, Italy). Fractions were characterized by LC–MS analysis (0.2 μg/fraction) to assess purity and molecular weight. A LCQ DCA XP Ion Trap mass spectrometer (ThermoElectron, Milan, Italy) equipped with an OPTON ESI source, operating at 4.8 kV needle voltage and 320°C and with a complete Surveyor HPLC system, comprising an MS pump, an autosampler, and a photo diode array (PDA) was used. A narrow bore 50 × 2 mm C18 Biobasic column, 300 Å, 3 μm (ThermoElectron) was used for these analyses, applying a gradient of CH_3CN, 0.05% TFA (Solvent B) from 5 to 55% over a period of 40 min. The sample tray was kept at 20°C, while the column was maintained at 25°C. Solvent A was H_2O, 0.08% TFA. Mass spectra were continuously recorded in a mass interval of 400–2000 amu (normal mode) and 1200–3500 amu (high mode), in positive modality (hereafter, LC–MS conditions **A**). Multicharge spectra were deconvoluted using the BioMass program implemented in the Bioworks 3.1 package, provided by the spectrometer manufacturer. Mass calibration was performed automatically by means of selected multiple charged ions in the presence of a calibrant (UltraMark, ThermoElectron). All masses are reported as averaged values. Pure fractions were pooled and lyophilized.

Synthesis of Fully Protected GstI[1–31]-COOH. The N-terminal fragment GstI[1–31]-COOH was prepared by automatic Fmoc chemistry, using a 433A Applied Biosystems automatic synthesizer. The synthesis was carried out on a chlorotritylchloride resin onto which the protected residue Fmoc-Leu-OH was loaded using standard procedures.[16] Standard side chains protections compatible with Fmoc chemistry were used throughout the synthesis for all residues, except Cys40 and Cys47, which were introduced as TFA-stable Acm derivatives. The N-terminal methionine was introduced as Boc derivative. 20% Piperidine in DMF was used as the Fmoc deprotecting agent in all steps. Activation was performed using a 0.9 HBTU–HOBt and 2.0 DIEA molar excess over amino acid (1.0 mmol). The resulting protected polypeptide (630 mg) was removed from the solid support by 1%TFA in DCM (v/v) treatments (5 treatments of 10 mL, 2 min, RT); filtrating in a 5% DIEA in DCM solution (v/v), the solvent was evaporated and the residual oil dried under vacuum at RT overnight. To characterize the polypeptide, about 1 mg of product was treated with 500 μL of a 94:3:3 v/v/v TFA–H_2O–TIS mixture for 1 h at RT, precipitated by adding cold Et_2O and characterized by LC–MS under the reported conditions **A**.

Preparation of the Fmoc-L-Cys-CONH$_2$. To try to use a odorless thiol instead of Et-SH for the preparation of C-terminal thioesters, we prepared the cysteine derivative Fmoc-L-Cys-$CONH_2$. This reagent, which is particularly useful also by virtue of the characteristic UV-absorption spectrum, was prepared by solid-phase synthesis on 100 μmol of RINK amide resin, subst. 0.53 mmol/g. To this aim, the resin was swollen in DCM–DMF 1:1 for 1 h, was extensively treated with 20% piperidine in DMF to remove the Fmoc protection, and then was washed thoroughly with plenty of DMF. 500 μmol of Fmoc-L-Cys(Trt)-OH was dissolved in 1.0 mL of DMF and poured onto the resin, then 1.0 mL of 0.5 *M* PyBOP in DCM and 1.0 mL of 2 *M* DIEA in NMP were added and the suspension was mixed for 2 h at RT. The amino acid was cleaved from the resin by treatment with 3.0 mL of a 94:3:3 TFA–H_2O–TIS v/v/v mixture and purified by solid phase extraction on a manually packed C18 cartridge. The amino acid was characterized by LC–MS using a complete Surveyor HPLC system and an MSQ single quadrupole mass spectrometer. The analysis was performed on a Proteo C8 150 × 4.6 mm ID column (Phenomenex) using a gradient from 20 to 80% CH_3CN, 0.1% TFA over 35 min (flow rate was 0.8 mL/min). Probe and cone of the mass spectrometer were set at 3.0 kV and 35 V, respectively, while the temperature was set at 330°C.

Preparation of the GstI[1–31] C-terminal Thioesters. Fully protected GstI[1–31]-COOH was thioesterified with EtSH (6-fold molar excess over the peptide[12] or the odorless Fmoc-L-Cys-$CONH_2$ (6-fold molar excess). Reactions were carried out in neat DCM at RT (5–16 h) using PyBOP (6-fold molar excess over the peptide) and DIEA (12-fold molar excess) as activating agents.[12] The reactions were monitored by removing small aliquots of solution, evaporating the solvent, and performing the complete deprotection with a 94:3:3 v/v/v TFA–H_2O–TIS mixture (500 μL/0.2 mL DCM solution, 1.5 h at RT). The resulting crude material was then analyzed by LC–MS under the reported LC–MS conditions **A**. When reactions were not completed, further reagents were added until more than 90% conversion was observed. The reaction mixtures were evaporated and treated with a 94:3:3 v/v/v TFA–H_2O–TIS mixture (3 mL/100 mg oil, 1.5 h at RT), and the crude products precipitated by adding cold Et_2O and purified on a COMBI C18 5 × 2.1 cm ID column (10 μm, Phenomenex) applying a gradient of Solvent B from 20 to 45% over 15 min (flow 20 mL/min; 20 mg per run), monitoring at 210 nm. Purest fractions were characterized by LC–MS analysis (0.2 μg/fraction, LC–MS condition **A**) identifying the products by molecular mass determination.

Synthesis of GstI[1–31]-CO-S-(CH$_2$)$_2$-CO-Leu-COOH. This thioester was directly prepared on a TAMPAL (trityl-associated mercaptopropionic acid-leucine) resin.[9–11] The resin was prepared by coupling *S*-trityl mercaptopropionic acid to Boc-Leu-Pam (HBTU/HOBt/DIEA) and was used after removal of the trityl protecting group with two 1-min treatments with 2.5% tri-isopropylsilane and 2.5% H_2O in TFA. The thioester bond was formed with the C-terminal Leu31 by using standard in situ–neutralization peptide-coupling protocols for 1 h. Boc chemistry was accomplished using in situ neutralization/HBTU activation procedure for all amino acids.[17] Commercially available Boc amino acids with classical protections were used throughout the synthesis. Each synthetic cycle consisted of N^{α}-Boc removal by a 1- to 2-min treatment with neat TFA, a 1-min DMF flow wash, a 10- to 20-min coupling time with 1.0 mmol of preactivated Boc-amino acid in the presence of excess DIEA, and a second DMF flow wash. N^{α}-Boc-amino acids (1.1 mmol) were preactivated for 3 min with 1.0 mmol of HBTU (0.50 *M* in DMF) in the presence of excess DIEA (3 mmol). After coupling of Gln residues, a DCM flow wash was used before and after deprotection by using TFA to prevent possible high-temperature (TFA/DMF)–catalyzed pyrrolidone formation. After chain assembly was completed, the peptide was deprotected and cleaved from the resin by treatment with anhydrous HF for 1 h at 0°C with 5% *p*-cresol as scavenger. The peptide was precipitated with Et_2O, dissolved in aqueous acetonitrile, and lyophilized. The final thioester was purified by RP–HPLC on a COMBI C18 5 × 2.1 cm ID column (10 μm, Phenomenex) applying a gradient of Solvent B from 20 to 45% over 15 min (flow 20 mL/min, 20 mg per run), monitoring at 210 nm. Fractions were characterized by LC–MS analysis (0.2 μg/fraction, LC-MS conditions **A**), assessing purity and molecular weight.

Preparation of homoCys32-GstI[1–63]Gly-Cys(Acm)40,47 (III) by NCL

GstI[1–31]-CO-SEt or GstI[1–31]-CO-S(Fmoc-Cys-NH_2) or GstI[1–31]-CO-S-(CH2)$_2$-CO-Leu-COOH (**I**) were interchangeably used in ligation reactions. In a typical experiment, 2 μmoles of a thioester were dissolved in

2.0 mL sodium phosphate buffer 200 m*M*, guanidine hydrochloride 6 *M*, pH 7.8. The solution was mixed with a phosphate buffered solution (200 m*M*, guanidine hydrochloride 6 *M*, pH 7.8) of **II** (0.9 equiv of a 1.0 m*M* solution); 1% thiophenol (v/v) was directly added and then the pH was checked to be $\sim$ 7.[4,6] The reactions were monitored by RP–HPLC analysis for 16 h using a Jupiter 250 $\times$ 4.6 mm ID RP-18 column (Phenomenex) operating at 1.0 mL/min. Aliquots of this solution were treated with equal volumes of β-Me for 30 min and then with 10% TCEP for 10 min to completely hydrolyze any thiol adduct before RP–HPLC analysis. Peaks were collected and identified by LC-MS analysis (conditions **A**). The ligated products were purified using a Jupiter 25 $\times$ 1 cm ID RP18 column operating at 5 mL/min. Effluents were monitored at 210 nm. Fractions corresponding to the ligated product **III** were pooled, characterized by LC–MS (conditions **A**), and lyophilized. The ligation between (**I**) and Cys32-GstI[32–63]-Gly was also carried out under the same conditions to compare the reactivity of homocysteine with cysteine.

Alkylation of homoCys-containing Polypeptides

***Method A: Methylation with Methyl* p-*Nitro-benzene-sulphonate (MpNBS)*[7].** 0.7 mg of Product (**III**) were dissolved in 200 m*M* phosphate buffer, pH 8.6. To this solution 1.1 mg of MpNBS dissolved in CH_3CN was added, corresponding to a 50-fold excess over the thiol group. The reaction was allowed to proceed for 1 h at RT and afterward it was blocked by adding 25% TFA in water. The reagent excess was removed by three subsequent extractions with Et_2O and the resulting product was characterized by LC–MS under the reported conditions **A**.

Method B: Methylation with Methyl Iodide. The alkylation reaction with MeI first was optimized on Cys-containing model peptides to avoid hyperalkylation. Reactions were carried out by changing alkylating agent excess and reaction time. The effects of reducing agents, like TCEP, were also evaluated. Most experiments were performed using the peptide LC1 that contained two cysteines (sequence: CQSLLNSGMRC-NH2). Peptide aliquots of 0.25 μmol (0.5 μmol total thiols) were dissolved in 400 μL of 100 m*M* bicarbonate buffer, pH 8.3, and separately treated with increasing molar excesses (10, 50, 100, 1000) of MeI, always dissolved in 100 μL DMF (5 m*M* peptide final concentration). Reactions were monitored by LC–MS analysis following retention time and MW changes after 0, 1, 5, 15, and 30 minutes. To this aim, aliquots were removed at the indicated times and the reactions quenched by 1 *M* DTT addition up to 250 m*M* final concentration and then TFA by bringing the pH to about 2. Polypeptide **III** was alkylated with MeI to produce Met32-GstI[1–63]Gly-Cys(Acm40,47) (**IV**) under the following conditions: 3.0 mg of polypeptide (about 0.15 μmol) were dissolved in 2.0 mL of 100 m*M* bicarbonate buffer, pH 8.3, 5 m*M* TCEP. The solution was left at RT for 10 min to allow reduction of any disulphide, thereafter 100 μL of 1.0 *M* MeI in DMF was added and the solution was left to stand for 1 min. 120 μL of 1.0 *M* DTT was finally added and the pH was brought to about 2 with pure TFA. The product **IV** was purified by RP–HPLC using a Jupiter 25 $\times$ 1 cm ID RP-18 column (Phenomenex) equilibrated with 5% solvent B at 5 mL/min, monitoring at 210 nm. A gradient from 5 to 60% solvent B in 50 min was applied after injection and elution of the low MW components. The purest fractions were collected and analyzed by LC–MS under the reported conditions **A**.

Proteolytic Digestions and Chemical Fragmentation

The polypeptide **IV** (40 μg) was dissolved in 30 μL of 0.1 *M* Tris–Cl, pH 8.4, containing 20 m*M* $CaCl_2$ and 10% acetonitrile. 1-Chloro-3-tosylamido-4-phenyl-2-butanone–treated trypsin was added at a final enzyme-to-substrate ratio of 1:100 (w/w) and the mixture left at 37°C for 5 h. CNBr fragmentation was performed in 70% formic acid according to Gross.[18,19]

Tryptic and CNBr peptides were separated by means of a modular CapLC system directly connected to the Z-spray source of a Q-TOF Micro (Waters Corporation, Milford, MA, USA). The sample (about 3 pmol), diluted at a concentration of 1 pmol/μL, was loaded onto a C18 precolumn (5 mm length $\times$ 300 μm ID) at a flow rate of 20 μL/min and desalted for 5 min with a solution of 0.1% formic acid. The sample was then directed onto a Symmetry-C18 analytical column (10 cm $\times$ 300 μm ID) using 5% CH_3CN, containing 0.1% formic acid at a flow rate of 5 μL/min, and eluted increasing the CH_3CN concentration. Electrospray MS and MS/MS data were acquired on a Q-TOF Micro mass spectrometer fitted with a Z-spray nanoflow electrospray ion source. The mass spectrometer was operated in the positive ion mode with a source temperature of 80°C and with a potential of 3000 V applied to the capillary probe. LC–MS data of tryptic peptides were acquired with the instrument operating in data directed analysis (DDA) MS/MS mode. The MS/MS data were processed automatically, and de novo sequencing was obtained by means of ProteinLynx Global Server 2.0.

Acm Removal

Acm was removed from the polypeptide by successive treatments with AgOTf[14] and 1 *M* DTT/aqueous 1 *M* HCl. Briefly, the polypeptide (0.25 μmol, ca. 2 mg) was dissolved in a solution (200 μL) of 1.5% (v/v) anisole in TFA. A 40-fold molar excess of AgOTf then was added and the reaction was left to proceed for 1.5 h at 4°C. Dry Et_2O was added to the reaction to afford powder. After being washed three times with Et_2O the product was left overnight at 4°C. Afterward, it was treated with 50% 1.0 *M* DTT/1 *M* HCl at RT for 4 h and, after removal of AgCl precipitate by filtration, the filtrates were diluted in H_2O and then characterized by LC–MS using the described method. Final products were finally purified by RP–HPLC under the conditions described in the previous section.

CD Spectroscopy

CD experiments were carried out on a J-715 Jasco spectropolarimeter using a 0.1 cm path length quartz cuvette. A phosphate (10 m*M*, pH 7.0) buffered solution of protein (1×10^{-5} *M*) was used to record spectra between 260 and 190 nm. The instrument was calibrated with an aqueous solution of D-10-(+)-camphorsulphonic acid[20] at 290 nm. Three spectra for each sample were recorded, averaged, and transformed in molar ellipticity/residue, [θ]. CD spectra were recorded with a time constant of 4 s, a 2 nm bandwidth, and a scan rate of 5 nm/min, were signal averaged over at least five scans, and were baseline corrected by subtracting a buffer spectrum. The spectrum of GstI[1–63]-Gly was deconvoluted using the CD Spectra Deconvolution software, version 2.1.[21]

This work was supported by the Ministero dell'Istruzione e dell'Università e della Ricerca (MIUR) FIRB 2001 RBNE01ARR4 and FIRB 2003 RBNE03PX83. We also gratefully acknowledge the National Research Council (CNR) and the technical support of Dr. Giuseppe Perretta.

REFERENCES

1. Spinosa, M.; Riccio, A.; Mandrich, L.; Manco, G.; Lamberti, A.; Iaccarino, M.; Merrick, M.; Patriarca, E. J. Mol Microbiol 2000, 37, 443–452.
2. Tatè, R.; Mandrich, L.; Spinosa, M. R.; Riccio, A.; Lamberti, A.; Iaccarino, M.; Patriarca, E. J. Mol Plant Microbe In 2001, 14, 823–831.
3. Napolitani, C.; Mandrich, L.; Riccio, A.; Lamberti, A.; Manco, G.; Patriarca, E. J. FEBS Lett 2004, 558, 45–51.
4. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. Science 1994, 266, 776–779.
5. Fields, G. B.; Noble, R. L. Int J Pept Protein Res 1990, 35, 161–214.
6. Dawson, P. E.; Kent, S .B. H. Annu Rev Biochem 2000, 69, 923–960.
7. Tam, J. P.; Yu, Q. Biopolymers 1998, 46, 319–327.
8. Pachamuthu, K.; Schmidt, R. R. Synlett 2003, 5, 659–662.
9. Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. Proc Natl Acad Sci USA 1999, 96, 10068–10073.
10. Aimoto, S. Biopolymers 1999, 51, 247–265.
11. Camarero, J. A.; Cotton, G. J.; Adeva, A.; Muir, T. W. J Pept Res 1998, 51, 303–316.
12. Von Eggelkraut-Gottanka, R.; Klose, A.; Beck-Sickinger, A. G.; Beyermann, M. Tetrahedron Lett 2003, 44, 3551–3554.
13. Kajihara, Y.; Yoshihara, A.; Hirano, K.; Yamamoto, N. Carbohydr Res 2006, 341, 1333–1340.
14. Tamamura, H.; Otaka, A.; Nakamura, J.; Okubo, K.; Koide, T.; Ikeda, K.; Ibuka, T.; Fujii, N. Int J Pept Protein Res 1995, 45, 312–319.
15. Houghten, R. A.; Li, C. H. In Peptides, Proceedings of the American Peptide Symposium; Goodman, M.; Meienhofer, J., Eds.; Wiley; New York, NY, 1977; 5th ed. pp. 458–460.
16. Bernhardt, A.; Drewello, M.; Schutkowski, M. J Pept Res 1997, 50, 143–152.
17. Schnolzer, M.; Alewood Jones, P. A.; Alewood, D.; Kent, S. B. Int J Pept Protein Res 1992, 40, 180–193.
18. Gross, E. Methods Enzymol 1967, 11, 238–255.
19. Skopp, R. N.; Lane, L. C. Appl Theor Electrophoresis 1988, 1, 61–64.
20. Johnson, W. C. Jr. Proteins 1990, 7, 205–214.
21. Boehm, G. CD Spectra Deconvolution; Institut fuer Biotechnologie, Martin Luther Universitaet, Hall Wittenberg, Germany, 1996. Available at: http://bioinformatik.biochemtech.uni-halle.de.

Reviewing Editor: Stephen Blacklow

Chemical Synthesis of Proteins and Circular Peptides Using N^{α}-(1-Phenyl-2-Mercaptoethyl) Auxiliaries

Paolo Botti[†,‡]* and Sylvie Tchertchian[†,‡]

[†]*Department of Structural Biology and Bioinformatics, University of Geneva, 1 rue Michel Servet, 1211 Geneva 4, Switzerland and Geneprot Inc., 2 Pré-de-la-Fontaine, 1217 Meyrin/GE Switzerland and* [‡]*Geneprot Inc., 2 Pré-de-la-Fontaine, 1217 Meyrin/GE Switzerland.*

Abstract: An overview of the applications of N^{α}-(1-phenyl-2-mercaptoethyl) auxiliary is presented. We describe the on resin preparation (C^{α}-carboxy and thioester) of N^{α}-auxiliary derivatives of glycine and the synthesis and incorporation of preformed N^{α}-auxiliary derivatives of glycine and alanine with the protection schemes, including the thiazolidine strategy for SPPS. Such approaches allowed the synthesis of the protein cytochrome b562 as well as native circular peptides after successful removal of the auxiliary.

Keywords: native chemical ligation, peptide cyclization, auxiliary group, protecting group.

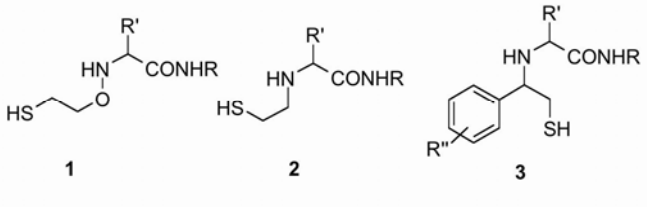
INTRODUCTION

During the last decade, total chemical synthesis has been established as the most rapid and efficient way to chemically gain access to small and proteins 1]. Up to the present, many ligation chemistries have been developed [2]. However, Native Chemical Ligation (NCL) [3] has been demonstrated to be uniquely effective. In the NCL scheme, a C-terminal thioester fragment reacts chemoselectively with an N-terminal cysteine peptide in aqueous solution at neutral pH. In the first step of the reaction, transthioesterification takes place by thiol exchange between the free thiol of the N-terminal cysteine and the thioester moiety on the other peptide. The newly generated thioester then undergoes an S to N acyl shift due to the proximity of the amino group to the thioester functionality, thus generating a native amide bond at the ligation site. NCL has also been used to generate cyclic peptides via an intramolecular reaction [4]. The requirement for cysteine at the site of peptide ligation is an intrinsic restriction of the standard NCL strategy. In proteins, Cys occurs with relatively low frequency; many proteins do not have suitably-disposed cysteines for the NCL strategy and some proteins lack Cys residues entirely. Thus, continuing demand for large polypeptides and proteins has to the development of removable auxiliary groups to extend the applicability of NCL to ligation and cyclization of unprotected peptides in aqueous solution at an X-X site (where X= any amino acid) [5]. In this paper, an exhaustive description of N^{α}-(1-phenyl-2-mercaptoethyl)-based [6] auxiliaries will be given, including synthetic strategies (on resin derivatization and synthesis of the preformed amino acids ready to be used in SPPS) and applications in both intermolecular (ligation) and intramolecular (cyclization) reactions.

*Address correspondence to this author at the Department of Structural Biology and Bioinformatics, University of Geneva, 1 rue Michel Servet, 1211 Geneva 4, Switzerland and Geneprot Inc., 2 Pré-de-la-Fontaine, 1217 Meyrin/GE Switzerland; Tel: (+41) 22 379 55 23, Fax: (+41) 22 379 55 02; E-mail: Paolo.Botti@medecine.unige.ch

DEVELOPMENT OF N^{α}-(1-PHENYL-2-MERCAPTO-ETHYL) AUXILIARIES.

The first extension of NCL was presented by Canne et al [5a], who exploited two acid-stable N^{α}-thiol containing auxiliaries, designed to mimic the side-chain of a native cysteine. The reductively cleavable N^{α}-(oxyethanethiol) **1** that has a 1,3-aminothiol structure, showed a relatively slow amide formation rate while the N^{α}-(2-mercaptoethyl) **2**, which was not designed to be cleavable, showed superior rearrangement rates (Scheme **1**).

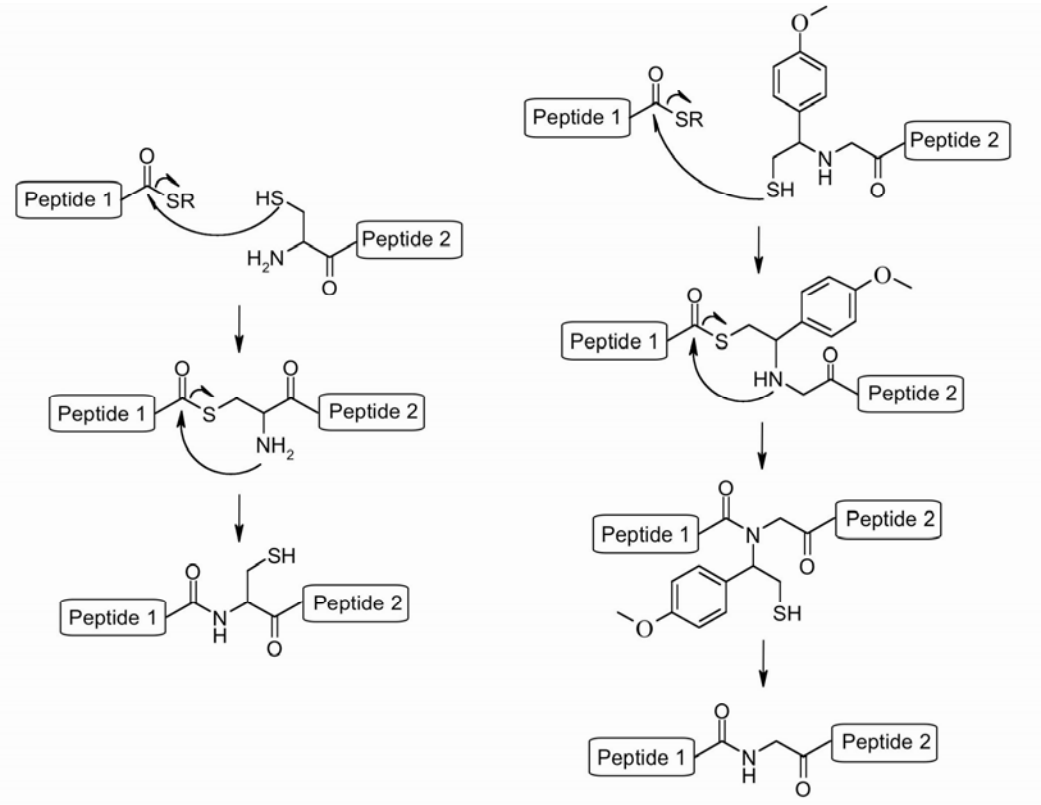


Scheme 1. Different auxiliary groups.

The interesting propensity of the latter toward a faster ligation prompted us to design novel auxiliary groups that could take advantage of the superior ligation rates of **2** and allow its facile and complete removal after ligation.

A case in point is the N^{α}-(1-phenyl-2-mercaptoethyl) auxiliary [6a-c] **3**, which combines the faster ligation rate of the N^{α}–(2-mercaptoethyl) group through a five-membered ring intermediate, enabling at the same time its complete removal after ligation. Indeed, the introduction of a 1-phenyl group on the N^{α}-(2-mercaptoethyl) template generates a benzylamine derivative that is completely stable under the strong acidic conditions used in SPPS to cleave peptide-resins (HF or TFA) making the auxiliary group compatible with both Boc- and Fmoc-based chemistries. Furthermore, because of the S to N acyl shift intrinsic to the mechanism of NCL, benzylamine to benzylamide conversion occurs during the reaction making the auxiliary readily cleavable under similar or milder conditions originally required for the

0929-8665/05 $50.00+.00



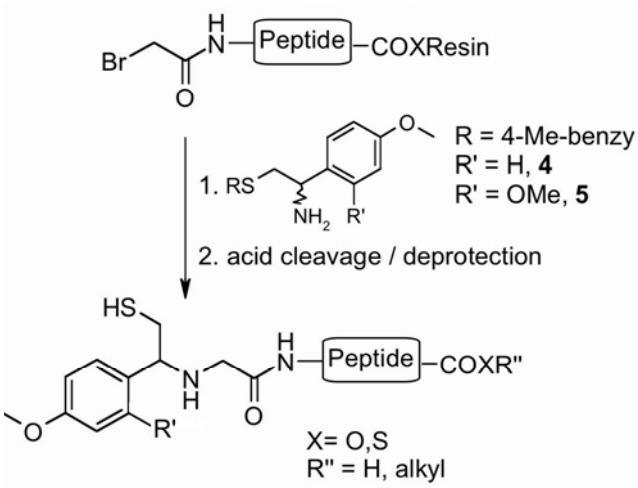
Scheme 2. Native Chemical Ligation and extended Chemical Ligation.

cleavage of the peptide (Scheme **2**). In order to unambiguously monitor reaction progress (proof of the amide bond formation) and to distinguish between the rearranged and unrearranged products that posses same mass, a sample of the ligation mixture must be treated with an excess of 2-mercaptoethanol. Indeed, the newly added mercaptan is necessary to intercept any unrearranged thiolactone, thus generating, in case of unrearranged material, a new specie with characteristic mass and HPLC retention time. Finally, since the auxiliary can be cleaved only when is linked to an amide group (benzylamines are stable to strong acidic conditions), its removal proves that the product is the desired material.

PREPARATION OF N-TERMINAL GLYCINE N^{α}-(1-PHENYL-2-MERCAPTOETHYL) PEPTIDES *VIA* ON-RESIN BROMIDE DISPLACEMENT

Our original scheme relied on the reaction of the appropriate benzylamine with an N-terminal bromo-acetylated peptide on a solid support (Scheme **3**). Two different benzylamine derivatives were synthesized [6b]: 1-(4-methoxyphenyl)-2-(4-methylbenzylthio)-ethylamine **4** and 1-(2,4-dimethoxyphenyl)-2-(4-methylbenzylthio)-ethyl-amine **5**.

Both amine precursors, **4** and **5,** were easily synthesized in three steps from the bromoacetophenone derivative of choice. Finally, reaction of such amines with a resin-bound bromoacetyl-peptide generated, after cleavage/deprotection steps, the desired N^{α}-(1-phenyl-2-mercaptoethyl) peptides. N-terminal glycine model peptides of identical sequence bearing either a mono or dimethoxy phenyl substituted N^{α}-(1-phenyl-2-mercaptoethy) auxiliary were reacted with the same peptide thioester fragment to evaluate both the ligation and deprotection steps. While the removal of the auxiliary from both products was accomplished in quantitative way, the ligation involving N^{α}-1-(4-methoxyphenyl)-2-mercaptoethyl auxiliary showed superior yield [6b] thus directing our focus on the mono-methoxy phenyl-substituted analogue for further development. After the preliminary work on model peptides which focused on the optimization of ligation conditions and auxiliary cleavage, we attempted the total chemical synthesis of cytochrome b562, a functional protein with no cysteine residues [6c] . In such a



Scheme 3. On-resin auxiliary formation.

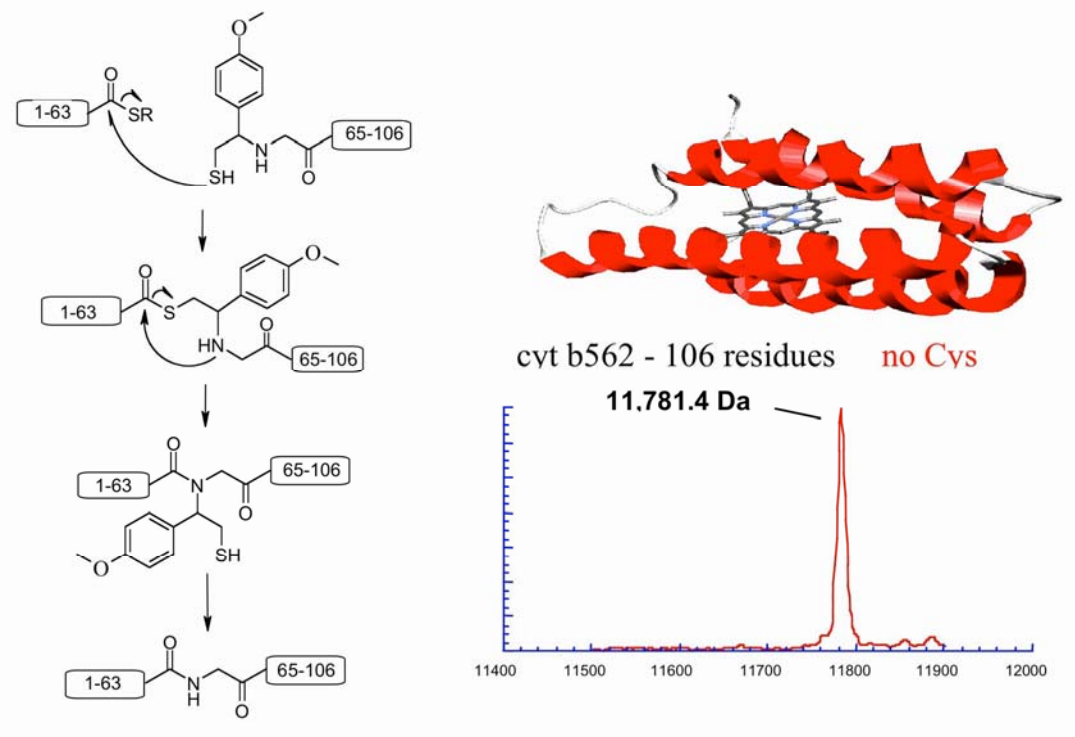


Figure 1. Synthesis of Cytochrome b562.

synthesis (Figure **1**), the C-terminal portion of cytochrome b652 (64-106) was prepared using 1-(4-methoxyphenyl)-2-(4'-methylbenzylthio)-ethylamine **4** *via* S_N2 displacement of the highly activated bromide from the N-terminal bromoacetyl peptide on a solid support.

Two variants of the N-terminal portion were synthesized. These include cytochrome b562 (1-63) with either Met or SeMet at position 7, both obtained as thioester fragments ready for use in ligation reactions. Both of the ligations were completed after overnight incubation. Upon auxiliary removal *via* HF treatment, both proteins were properly folded in the presence of the corresponding heme group.

CYCLIZATION REACTION

The on-resin incorporation of the auxiliary moiety *via* S_N2 displacement of the bromide from a N-terminal bromoacetyl peptide has also been applied to the synthesis of circular peptides *via* an intramolecular reaction of linear precursors having both reactive species, the thioester and the 1,2 amino-thiol moieties, on the same molecule [6d]. Since thioesters are known to be susceptible to aminolysis, a possible side reaction could be the loss of peptide due to a large excess of the incoming primary amine. The reaction has to be conducted using a minimal excess (1.1 equiv.) of amine **4** in the presence of 1.5 equiv. of base (DIEA).

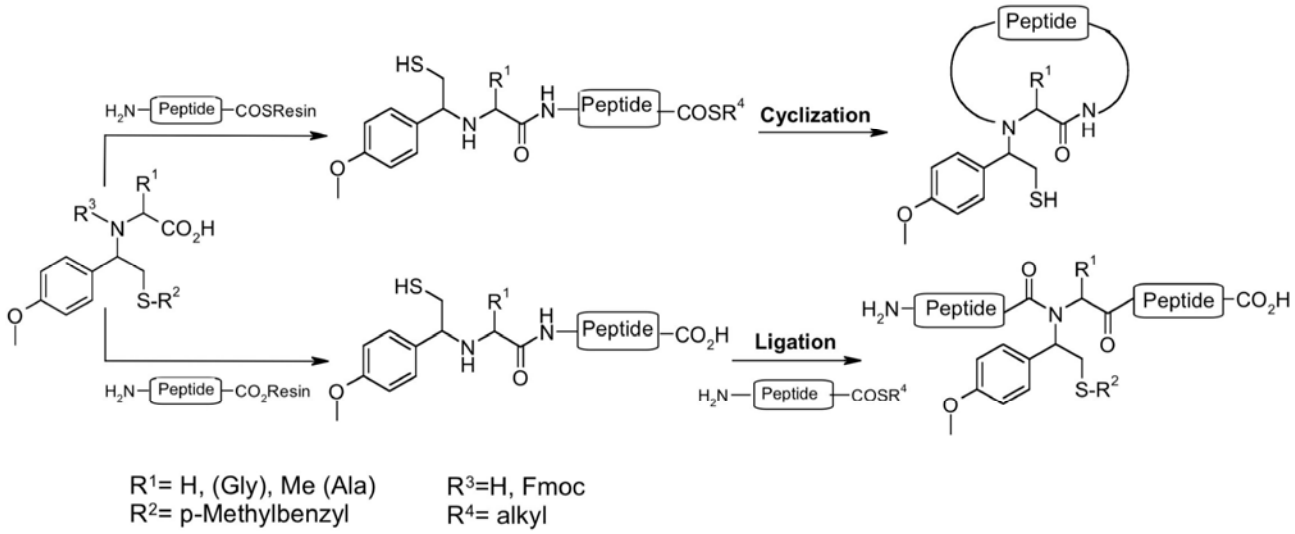
Under such controlled conditions, no detectable loss of resin weight was observed after overnight incubation at room temperature. After the cleavage/deprotection step, the desired cyclic peptide was obtained using thiophenol as catalyst. Several experiments were carried out to optimize cyclization conditions, and the most critical parameter was pH precisely set to 7.4. Finally, the subsequent facile removal of the auxiliary group yielded the desired cyclic peptide with native structure.

PREPARATION OF N-TERMINAL N^{α}-(1-PHENYL-2-MERCAPTOETHYL) PEPTIDES USING PREFORMED N^{α}-(1-PHENYL-2-MERCAPTOETHYL) AMINO ACIDS

The former strategy, although successfully applied in both inter- and intramolecular reactions at the glycine site, presents a few limitations. Indeed, the on-resin preparation of peptide segments featuring a substituted residue at the N-terminus may generate a side product on the 2-bromoacetyl residue precursor through an elimination mechanism [5a]. Such phenomenon can be further complicated either by the size of the desired peptide (especially with large fragments due to incomplete acyclation reactions) or by the type of resin (C^{α}-thioester resin), which can be susceptible to aminolysis/hydrolysis depending on the conditions. We realized that incorporation of the auxiliary group into a peptide chain (especially on long peptides) through on-resin bromine displacement could limit the real usefulness of the auxiliary group. Therefore we set out to synthesize amino acid derivatives incorporating the auxiliary moiety ready to be coupled in SPPS (Scheme **4**).

Based on such premises, we designed a simple synthetic strategy to obtain N^{α}-(1-phenyl-2-mercaptoethyl) amino acids ready to be incorporated into any C^{α}-functionalized peptide resin. Thus, we prepared an auxiliary derivatized glycine and alanine. Alanine was selected for its high abundance in proteins and small side-chain.

Importantly, our synthetic strategy [6f], which is based on the reductive amination step (Scheme **5**), allows the direct



Scheme 4. Application of preformed N^a-(1-phenyl-2-mercaptoethyl) amino acid.

conjugation of an amino acid with the auxiliary template. It is interesting to remark that such a reaction generates a new stereogenic center at the former carbonyl of the auxiliary skeleton when optically pure amino acids (e.g., alanine) are used. The two resulting diastereomers can be separated and individually evaluated for the ligation or cyclization reactions. Consequently, using such a synthetic scheme, we have been able to synthesize the N^{α}-unprotected derivative **5**, the diastereomers **6a** and **6b**, the acid-labile protected Boc-N^{α}(Aux)-Gly-OH **7** and Boc-N^{α}(Aux)-Ala-OH **8b** and the base-labile Fmoc-N^{α}(Aux)-Gly-OH **9** and Fmoc-N^{α}(Aux)-Ala-OH **10b** (Scheme **5**).

COUPLING USING OF PROTECTED OR UNPROTECTED N^{α}- (1-PHENYL-2-MERCAPTO-ETHYL) AMINO ACIDS

The coupling of N^{α}-unprotected derivatives **5**, **6a** and **6b** requires the use of 1.1 equiv. of activated amino acid to avoid the generation of oligomerized side products. Under such controlled conditions, all unprotected derivatives **5**, **6a** and **6b** were successfully coupled to a model peptide C^{α}-thioester.

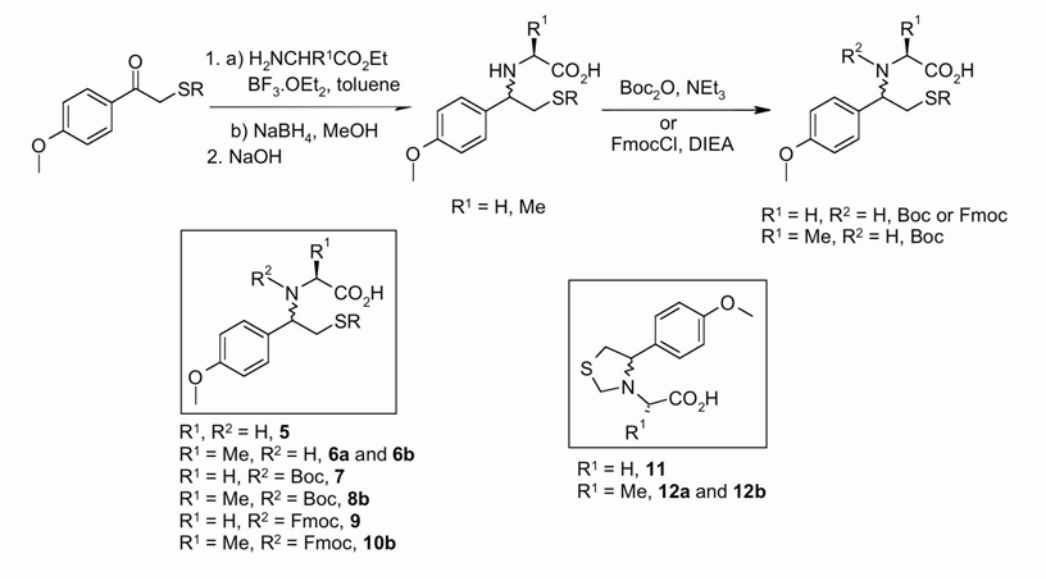
Boc-protected derivatives were also readily coupled to the peptide resin. However, they cannot be used due to the TFA mediated cleavage of the auxiliary group, which probably takes place prior or concomitantly to the Boc group deprotection. Indeed, stability tests carried out in solution on Boc-N^{α}(Aux)-Gly-OH **7** with different percentages of TFA in DCM showed that the Boc-protected amino acid is extremely sensitive to acidic conditions. Even at TFA concentrations as low as 5%, both the auxiliary and the Boc groups were completely cleaved in 30 minutes. Consequently, we turned our attention to a base-labile protection strategy. To be compatible with both ligation and cyclization strategies, a suitable protecting group must be safely removed from a C^{α}-thioester peptide. Recently Clippingdale et *al*. [7] reported the use of DBU for the synthesis of peptide C^{α}-thioesters *via* Fmoc-based chemistry. Although such methodology is applicable only to the preparation of small peptides, its use in a single step presents practically no side reactions.

Thus, the Fmoc-N^{α}(Aux)-Ala-OH compound 10b was successfully coupled to a peptide resin C^{α}-thioester and the Fmoc group was removed using 1% DBU in anhydrous DMF. No appreciable amount of resin was lost during the Fmoc deprotection step. The desired peptide was obtained after standard HF cleavage.

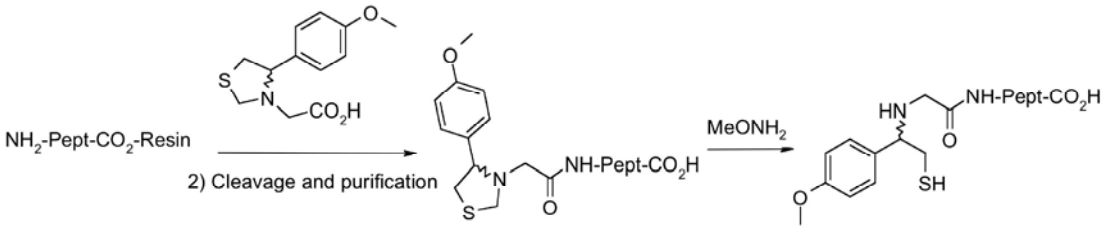
THIAZOLIDINE PROTECTION SCHEME

Thiazolidine derivatives have been successfully used to protect the N-terminal α-aminothiol moiety of cysteine during syntheses of peptides [8]. Furthermore, such structures have been exploited as N-terminal protection for cysteine in the multi-fragment synthesis of different natively folded proteins [9,10]. Accordingly, in ligation strategies involving three or more peptide fragments, the 1,2-aminothiol of the N-terminal (Aux)-middle segment must be protected to avoid cyclization [6d-11] or polymerization. To this end, the thiazolidine structure provides a convenient temporary protection for the N-terminal Cys residue for peptide segments containing a C-terminal thioester. During the ligation reaction the thiazolidine group effectively blocks the 1,2 aminothiol moiety of the N-terminal cysteine thus preventing cyclization or polymerization.

Our aim was to take advantage of the secondary amine of the 1,2 aminothiol moiety of our auxiliary group to prepare SPPS-ready N^{α}(Aux)-amino, which could be also suitable for the multi-fragment ligation strategy. A case in point was the thiazolidine formation with the secondary amine and the thiol groups of the N^{α}(Aux)-derivatized amino acid that generates a convenient temporary tertiary amine, suitable for SPPS. We thus prepared the thiazolidine derivatives **11** and diastereomers **12a** and **12b** (Scheme **5**), which were conveniently obtained from their precursors **5**, **6a** and **6b** in a two-step procedure [6f]. The thiazolidine **11** was coupled to a model peptide made by Fmoc chemistry. Analysis of the



Scheme 5. N^{α}-(1-phenyl-2-mercaptoethyl) amino acids.



Scheme 6. Thiazolidine protection scheme.

crude material after TFA cleavage showed no traces of ring opening (Scheme **6**).

Finally, the free 1,2-aminothiol group of the N-terminal auxiliary-derived glycinyl peptide was produced by treatment with O-methylhydroxylamine under acidic conditions (identical to those reported for N-terminal cysteine derivatives) in the post cleavage mixture [8-10]. Future work will present a more exhaustive study on thiazolidine derivatives including the use of alanine derivatives in SPPS and the exploitation of such protection scheme in Boc chemistry.

CYCLIZATION STUDIES USING N^{α}(AUX)-AA-OH DERIVATIVES.

The coupling of Gly derivative **5** to a model peptide of sequence GSYRFG on a thioester resin yielded fragment G(Aux)-GSYRFG-C^{α}OSR, which quantitatively cyclized in 0.5 h under standard conditions. Subsequent removal of the auxiliary group by HF treatment provided the native peptide as a single peak by HPLC. The two Ala-based diastereomers, **6a** and **6b**, were also coupled to the same model peptide GSYRFG-C^{α}OS-resin. After HF cleavage, each peptide was purified by HPLC. Both linear purified peptide diastereomers, A(Aux)-GSYRFG-C^{α}OSR derived from **6a** and **6b**, were cyclized as before affording each one a mixture of two cyclic compounds with the same mass (with superior conversion for peptide diastereomer derived from **6a**). These results suggest epimerization at the chiral center of the alanine residue.

HPLC analysis of the crude reactions after TFA removal of the auxiliary group in both cases still showed the presence of two peaks with the same mass corresponding to the cyclized peptide without the auxiliary group. Evaluation of the epimerization ratios were done by comparison to the *D*- and *L*-Ala cyclic peptides (AGSYRFG) synthesized on an oxime resin [12]. The results showed 15% and 30% epimerization of the Ala residue for the peptides derived from **6a** and **6b,** respectively. Such high extent of racemization (especially for the diastereomer **6b**) could not come from hydrolysis of the ester due to the different ^{1}H-NMR spectrum and HPLC retention time of the two diastereomers [6f]. Since reductive amination has widely been used in peptide chemistry for the preparation of N-alkykated amino acids [13] and is known to give no or little epimerization [14], we set out to determine whether the major source of epimerization resulted from the coupling or from the cyclization step.

Determination of the epimerization step was accomplished by detailed HPLC analysis of both the linear precursors and the cyclization products. Indeed, whereas both linear precursors showed a single symmetrical peak on a C-8 column, a non-symmetrical shape was obtained on a C-18 column. Thus, when the A(aux)-GSYRFG diastereomer derived from **6a** was repurified on a C-18

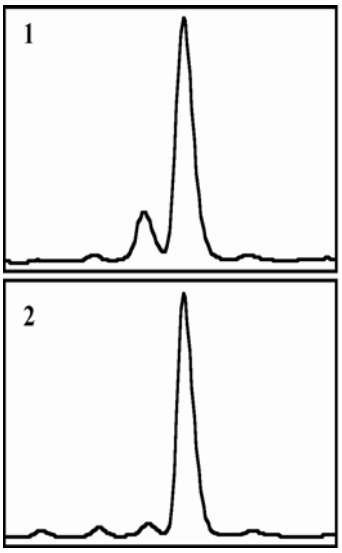


Figure. 2. HPLC analysis of cyclization products: A(Aux)GSYRFG cyclic peptides derived from **6a** obtained before (1) and after (2) a second purification of the linear starting material.

column, the cyclization reaction produced only one peak with the expected mass for the cyclized material (Figure **2**). Finally, our synthetic efforts to synthesize and separate amino acids **6a** and **6b** were validated by the different reactivities of these Ala derivatives during the coupling (epimerization ratios) steps. Unfortunately, our preliminary efforts to reduce the extent of epimerization by using carbodiimide mediated coupling (DCC) with alternatively HOBt or HOOBt (3-hydroxy-3,4-dihydro-4-oxo-1,2,3-benzotriazine) as activating agent failed. Similar epimerization results were also obtained when the Fmoc-Ala derivative **9b** was employed.

WORK DONE BY OTHERS

Other groups have also contributed to the exploitation of the N^{α}-(1-phenyl-2-mercaptoethyl) auxiliary groups. The work presented by Marinzi et *al*. [15] described the scaffold-mediated amide bond formation, but it did not address how to remove the auxiliary group after the ligation step. An interesting photo-removable version of this auxiliary group has been also presented by Kawakami et *al*. [16] and later by Marinzi et *al*. [17]. More recently, Clive et *al*. [18] has described the synthesis of different N^{α}-auxiliary amino acid templates although they were not used in SPPS.

CONCLUSION

Our work has broadened the application of Native Chemical Ligation (NCL) by extending the number of available ligation sites. Using the Extended Chemical Ligation (ECL) strategy at X-Gly sites, we have readily synthesized and folded cytochrome b562, a functional protein devoid of cysteine residues. This method diminished the sequence depending restrictions imposed by NCL. The same methodology has also been applied to an intramolecular reaction, thus facilitating the chemical synthesis of cyclic peptides from linear unprotected precursors. We have also made available the preformed auxiliary derivatized alanine and glycine residues ready for SPPS. This makes our technology suitable for both ligation and cyclization of unprotected peptide segments and even more sequence independent. Future work will focus on trying to find different coupling reagents and/or conditions that eliminate or reduce epimerization of the N^{α}-(Aux)-Ala derivatives during SPPS.

REFERENCES

[1] (a) Wilken, J.; Kent S.B.H. (**1998**) *Cur. Opin. Biotechnol.*, *9*, 412-426. (b) Kochendoerfer, G.G.; Kent S.B.H. (**1999**) *Cur. Opin.Chem Bio*., *3*, 665-671. (c) Dawson P.E.; Kent S.B.H. (**2000**) *Annu. Rev. Biochem*., *69*, 923-960.

[2] (a) Schnolzer M.; Kent, S.B.H. (**1992**) *Science*, *256*, 221-225. (b) Rose, K. (**1994**) *J. Am. Chem. Soc*., *116*, 30-33. (c) Gaertner H., Offord R., Cotton R., Timms D., Camble R., Rose K. (**1994**) *J. Biol. Chem*.**,** *269*, 7224-7230. (d) Liu, C.F.; Tam, J.P. (**1994**) *Proc. Natl. Acad. Sci. USA*, *91*, 6584-6588. (e) Shao, J.; Tam, J.P. (**1995**) *J. Am. Chem. Soc*.**,** *117*, 3894-3899. (f) Tam, J. P.;.Yu, Q. (**1998**) *Biopolymers*, *46*, 319-327. (g) Tam, J. P.; Miao, Z.; (**1999**) *J. Am. Chem. Soc*., *121*, 9013-9022 (h) Saxon, E.; Bertozzi, C.R. (**2000**) *Science*, *287*, 2007-2010. (i) Saxon, E.; Armstrong, J.I.; Bertozzi, C.R. (**2000**) *Org. Lett*., *2*, 2141-2143. (l) Nilsson, B.L.; Kiessling L.L.; Raines, R.T. (**2000**) *Org. Lett*., *2*, 1939-1941. (m) Nilsson, B.L.; Kiessling L.L.; Raines, R.T. (**2001**) *Org. Lett*., *3*, 9-12. (n) Nilsson, B.L.; Hondal, R.J.; Soellner, M.B.; Raines, R.T. (**2003**) *J. Am. Chem. Soc*.**,** *125*, 5268-5269.

[3] (a) Dawson; P.E.; Muir, T.W.; Clark-Lewis, I.; Kent, S.B.H. (**1994**) *Science*, *266*, 776-779. (b) Tam, J. P.; Lu, Y.A.; Liu, C.F.; Shao, J. (**1995**) *Proc. Natl. Acad. Sci. USA*, *92*, 12485-12489. (c) Canne, L.E.; Botti, P.; Simon, R.J.; Chen, Y.; Dennis, E.A.; Kent, S.B.H. (**1998**) *J. Am. Chem. Soc*., *121*, 8720-8727. (d) Kochendoerfer, G.G.; Chen, S.Y.; Mao, F.; Cressman, S.; Traviglia, S.; Shao, H.; Hunter, C.L.; Low, D.W.; Cagle, E.N.; Carnevali, M,; Gueriguian, V.; Keogh, P.J.; Porter, H.; Stratton, S.M.; Wiedeke, M.C.; Wilken, J.; Tang, J.; Levy, J.J.; Miranda, L.P.; Crnogorac, M.M.; Kalbag, S.; Botti, P.; Schindler-Horvat, J.; Savatski, L.; Adamson, J.W.; Kung, A.; Kent ,S.B. H.; Bradburne, J.A. *Science*. (**2003**) Feb 7;299(5608):884-7. (e) Rose, K.; Bougueleret, L.; Baussant, T.; Boehm, G.; Botti, P.; Colinge, J.; Cusin, I.; Gaertner, H.; Gleizes, A.; Heller, M.; Jimenez, S.; Johnson, A.; Kussmann, M.; Menin, L.; Menzel, C.; Ranno, F.; Rodruiguez-Tomé, P.; Rogers, J.; Saudrais, C.; Villain, M.; Wetmore, D.; Bairoch, A.; Hochstrasser, D. *Proteomics* (**2004**), *4*, 2125-2150

[4] (a) Zhang, L.; Tam, J.P (**1997**) *J. Am. Chem. Soc*., *119*, 2363-2370; (b) Camarero, J.A.; Pavel, J.; Muir, T.W. (**1998**) *Angew. Chem. Int. Ed*., *37*, 347-349. (c) Camarero, J.A.; Cotton, G.J.; Adeva, A.; Muir, T.W. (**1998**) *J.Peptide Res. 51*, 303-316.

[5] (a) Canne, L.E.; Bark, S.J.; Kent, S.B.H. (**1996**) *J. Am. Chem. Soc*., *118*, 5891-5896. (b) Shao, Y.; Lu, W.; Kent, S.B.H. (**1998**) *Tetrahedron Lett.*, *39*, 3911-3914. (c) Offer, J.; Dawson, P.E. (**2000**) *Org. Lett*., *2*, 23-26. (d) Kawakami, T.; Akaji, K.; Aimoto, S. (**2001**) *Org. Lett.*, *3*, 1403-1405, (e) Vizzavona, J.; Dick, F.; Vorherr,T. (**2002**) *Bioorg. & Med. Chem. Lett.*,12,1963-1965. (f) Offer, J.; Boddy, C.N.C.; Dawson, P.E. (**2002**) *J. Am. Chem. Soc*., *124*, 4642-4647.

[6] (a) Botti, P; Carrasco, M.R.; Low, D.W *Proceedings of the 26th European Peptide Symposium*, 10-15 September 2000, Montpellier, France. (**2001),** 385-386**.** Editions EDK, Paris. (b) Botti, P.; Carrasco, M.R.; Kent, S.B.H. (**2001**) *Tetrahedron Lett., 42*, 1831-1833. (c) Low, D.W.; Hill, M.G.M; Carrasco, M.R.; Kent, S.B.H.; Botti, P. (**2001**) *Proc. Natl. Acad. Sci. USA*, *98*, 6554-6559. (d) Cardona,V.M.F.; Hartley O.; Botti, P. (**2003**) *J. Pept. Res*.**,** *61*, 152-157. (f) Tchertchian, S.; Hartley, O.; Botti, P. (**2004**) *J. Org. Chem*, *69*, 9208-9214.

[7] Clippingdale, A.B.; Barrow, C.J.; Wade, J.D. (**2000**) *J. Peptide Science*, *6*, 225-234.

[8] (a) King, F.E.; Clark-Lewis, J.W.; Smith, G.R.; Wade, R. (**1959**) *J. Chem. Soc*., 2264. (b) Ratner, S.; Clarke, H.T. (**1937**) *J. Am. Chem. Soc*., *59*, 200. (c) Sheehan, J.C.; Yang, D.-D.H. (**1958**) *J. Am. Chem. Soc.*, *80*, 1158. (d) Hiskey, R.G.; Tucker, W.P. (**1962**) *J.*

Am. Chem. Soc., *84*, 4789. (e) Kemp, D.S.; Carey, R.I. (**1989**) *J. Org. Chem.*, *54*, 3640. (f) Wöhr, T.; Rohwedder, B.; Wahl, F.; Mutter, M. (**1994**) *J. Am. Chem. Soc*., *118*, 9218-9224.

[9] (a) Villain, M.; Vizzavona, J.; Gaertner, H. (**2001**) *Peptides: the Wave of the Future. Proceedings of Seventeenth American Peptide Symposium*, San Diego, Ca, USA; June 9–14, 107–108. (b) Villain, M.; Vizzavona, J.; Gaertner, H. (**2002**) *Collected Papers of the Seventh Symposium of Innovation and Perspectives in Solid Phase Synthesis*, Southampton, England, UK, 39–42.

[10] Bang, D.; Kent, S.B.H (**2004**) *Angew. Chem. Int. Ed., 43*, 2534-2538.

[11] Villain, M.; Gaertner, H.; Botti, P. (**2003**) in *Peptides: Peptide Revolution: Genomics, Proteomics & Therapeutics. Proceedings of Eighteen American Peptide Symposium*, Boston, MA, USA; July19–23, 71–72.

[12] Osapay, G.; Profit, A.; Taylor, J.W. (**1990**) *Tetrahedron Lett*., *31*, 6121-6124.

[13] (a) Johnson, T.; Quimbell, M.; Owen, D.; Sheppard, R.C. (**1993**) *J.Chem.Soc., Chem.Commun*.,*4*, 369-372. (b) Aurelio, L.; Brownlee, R.T.C.; Hughes, A.B.; (**2004**) *Chem. Rev*., *104*, 5823-5846. (c) Taranov, V.I.; Kadyrov, R.; Riermeier, T.H.; Fisher, C.; Boerner, A. (**2004**) *Adv. Synth. Catal*., *346*, 561-565.

[14] Najera,C. (**2002**) *Synlett*., *9*, 1388-1403.

[15] Marinzi, C.; Bark, S. J.; Offer, J.; Dawson, P. (**2001**) *Bioorg. Med. Chem*., *9*, 2323-2328.

[16] Kawakami, T.; Aimoto, S.; *Tetrahedron Lett.* (**2003**), *44*, 6059-6061.

[17] Marinzi, C.; Offer, J.; Longhi, R.; Dawson, P. (**2004**) *Bioorg. Med. Chem*., *12*, 2749-2757.

[18] Clive, D.L.J.; Hisaindee, S.; Coltart, D.M. (**2003**) *J. Org. Chem*., *68*, 9247-9254.

Understanding Biology Using Peptides
Sylvie E. Blondelle (Editor)
American Peptide Society, 2005

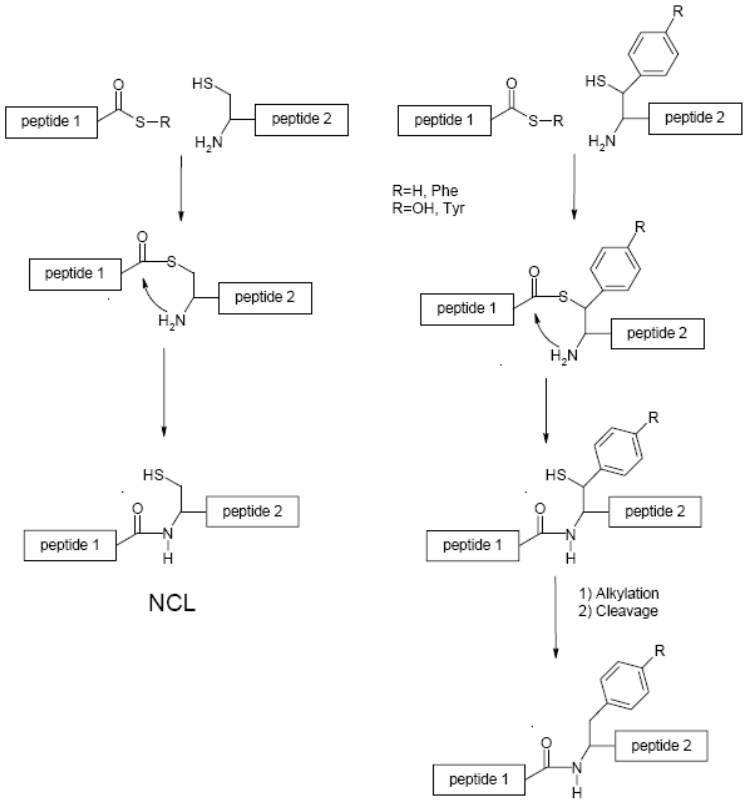
Native Chemical Ligation at Aromatic Residues

Sylvie Tchertchian[1], Fred Opligger[1], Marianne Paolini[1], Sonia Manganiello[1], Sylvain Raimondi[1], Benoît Depresle[1], Nicolas Dafflon[1], Hubert Gaertner[1] and Paolo Botti[1,2]

[1]Geneprot Inc. 2, Pré-de-la-Fontaine 1217 Meyrin/Geneva Switzerland;
[2]DBSB Department, University of Geneva, 1, rue Michel Servet, 1211 Geneva, Switzerland.

Introduction

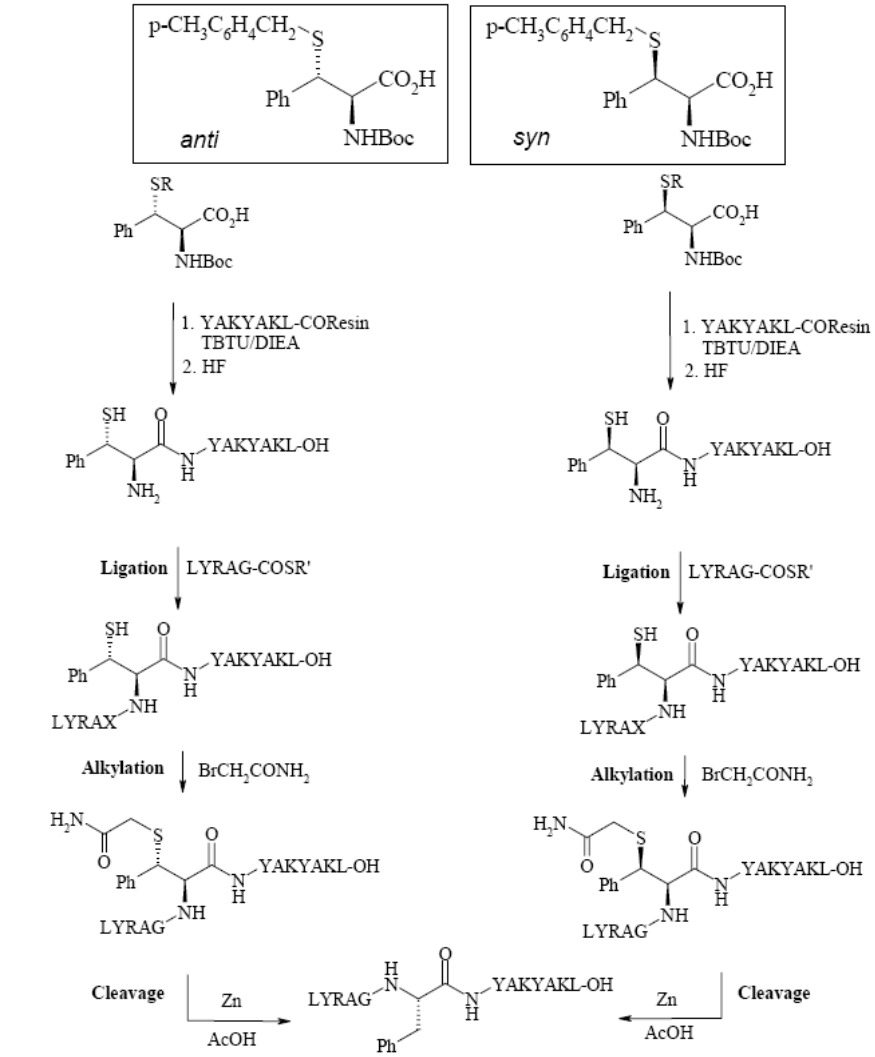
The last decade has provided extensive demonstration of the key role played by Native Chemical Ligation (NCL) for the preparation of small and medium size proteins [1]. Yet the requirement for cysteine at the site of ligation in standard NCL has limited its flexibility. Recently, different types of auxiliary groups [2-4] have been developed to extend the application of NCL to other ligation sites. However, the slower ligation rates especially with large fragments and the additional step required to cleave the auxiliary post-ligation have reduced their utility.



Scheme 1. NCL at aromatic residues (Phe, Tyr and Trp).

Results and Discussion

Our strategy does not make use of N-alkyl auxiliary groups [2,4], instead originally exploits the feature of particular side chain functionalities to generate a removable 2-mercapto amino group to mimic NCL. Our scheme takes advantage of the benzylic site of the aromatic residues to release a mercaptan under specific conditions. We use phenylcysteine [5] to exploit ligation with thioester fragments and the alkylation post-ligation is designed to produce a benzylic thioether suitable for cleavage. Benzyl groups are commonly employed in peptide synthesis to protect side chain functionalities *via* ether (Ser and Thr) and thioether (Cys) bonds. Thus, after ligation, alkylation of the free mercaptan generates a benzyl thioether suitable for cleavage.



Scheme 2. NCL using phenylcysteine derivatives.

Both N and S protected *anti* and *syn* phenylcysteine derivatives are coupled to a model peptide resin of sequence YAKYAKL. After HF cleavage and purification each peptide is exploited in ligation with thioester model peptides of sequence LYRAX-CO-SR, where X is alternatively Gly, Ala or Ile. As control, we also ligate CYAKYAKL standard with thioester fragments respectively with X = Ala and Ile.

Table 1.

Thioester Fragment	C-Terminal Fragment		
	Ligation with Z-YAKYAKL after 1h[a]		
LYRAX-CO-SR	^{HS}F-YAKYAKL *anti*	^{HS}F-YAKYAKL *syn*	C-YAKYAKL
LYRAG-CO-SR	> 95%	> 95%	-
LYRAA-CO-SR	90%	> 95%	> 95%
LYRAI-CO-SR	20%	~ 50%	40%

[a]*Ligation yield based on HPLC measured at λ 214nm.*

Ligation crudes of both reactions using N-terminal *anti* and *syn* ^{HS}F-YAKYAKL peptides with LYRAG-CO-SR fragment are treated with excess of bromoacetamide at ~ pH 8 followed by ether extractions. Then approximately 20%v/v of acetic acid is added to each mixture with subsequent final addition of 30-fold excess of Zn dust. In both cases the reduction is completed in 1hr at r.t., generating a major identical product, which has mass and HPLC retention time identical to the native standard reference peptide (LYRAGFYAKYAKL) synthesized *via* SPPS.

In conclusion, we developed a new method to ligate peptide fragments at non-Cys residue. Ligation rates are high, and in the case of the *syn* phenylcysteine fully comparable with the ligation at cysteine. The overall process is "one pot" and the reductive cleavage post alkylation is easy and efficient. No racemization is detected by comparing both final products with reference standard peptides (LYRAGFYAKYAKL) assembled *via* SPPS with either L or D-Phe.

References

1. Dawson, P. E., Muir, T. W., Clark-Lewis, I. and Kent, S. B., *Science* **266**, 776-779 (1994).
2. Canne, L. E., Bark S. J. and Kent S. B. *J. Am. Chem. Soc*. **118** 5891-5896 (1996).
3. Offer, J. and Dawson, P. E. *Org. Lett*., **2**, 23-26 (2000).
4. Botti, P., Carrasco, M. R. and Kent, S. B. H. *Tetrahedron Lett*. **42**, 1831-1833 (2001).
5. Xiong, C., Wang W. and Hruby, V. J. *J. Org. Chem*. **67**, 3514-3517 (2002).