Università degli Studi di Firenze
Facoltà di Scienze Matematiche, Fisiche e Naturali
Dottorato di ricerca in Scienze Chimiche
XXII ciclo

Synthetic modified peptides to reproduce post-translational modifications and structures of pathologically relevant proteins

Stefano Carganico

Tutor
Prof. Anna Maria Papini

PhD Supervisor
Prof. Gianni Cardini

1 Introduction and Summary

The PhD work described in this thesis was developed in the context of a cooperation between the Laboratory of Peptide & Protein Chemistry & Biology of the University of Florence (Prof. Anna Maria Papini) and the Laboratory for Translational Research of Harvard Medical School (Prof. Michael Chorev).

The present thesis describes some applications of modified peptides to the study of diabetes and cancer disease forms with a special attention to the development of diagnostic/prognostic assays based on molecular tools. To this aim, we developed non proteinogenic modified amino acids, orthogonally protected for Solid Phase Peptide Synthesis peptide sequences bearing post-translational modifications and possible stabilized conformation to be used in various biochemical applications.

The goal of the **diabetic project** has been the development of a diagnostic/prognostic tool for type I and type II diabetes. A convergent approach of multiple different strategies proceeding in parallel has been adopted (section 1).

To this aim, a panel of new Fmoc-Lysine derivates bearing a glycation modification has been developed for SPPS of glycated peptides and proteins (section 1.1).

A systematic approach to the synthesis of a glycated hCD59(37-50) peptide antigen has been carried out (section 2.2.3).

The glycated antigen has been used to produce specific anti-glycated hCD59 antibodies that efficiently recognize the glycated protein (hCD59) *in vivo* (section 2.2.5).

A panel of glycated antigenic peptide probes has been generated and tested with un-competitive ELISA experiments against type I diabetic patients' sera (section 2.3.5). Preliminary results show the presence of specific

autoantibodies anti-glycated-hCD59 in a subfamily of diabetic patients (section 2.4.2).

The strategy for total synthesis by Native Chemical Ligation of glycated and un-glycated hCD59 protein has been set up (section 2.5.1).

Future developments will be the optimization of the biochemical assays base on the synthetic peptide probes and antibodies generated as well as the completion of the NCL synthesis of hCD59; it will also be undertaken an analytical study of the glycated metabolites in diabetic patients' sera; and finally, the synthetic peptide probes will be tested in uncompetitive ELISA assays against type II diabetic patients' sera (section 2.6.2).

The goal of the **cancer project** has been the development of new eIF4E Binding Protein (4E-BP) peptide inhibitors with stabilized conformation to be used as suppressors of Translation Initiation in a context of cancer therapy (section 3.3).

A collection of N^{α}-Fmoc protected unnatural amino acids bearing on the side chain azido or alkynyl functions was synthesized and introduced by SPPS in the fluoresceinated-4E-BP (621-636) peptide sequence to afford by Cu(I) catalyzed Huisgen reaction, a new collection of cyclopeptides containing the triazolyl moiety (section 3.5). The side-chain-to-side chain cyclization of linear peptides generated *via* click chemistry lead to cyclopeptides containing the triazolyl moiety linked to the α-carbon of the amino acids by alkyl chains of different lengths.

The collection of linear and cyclic peptides has been tested with Fluorescence Polarization Assay to measure the affinity for the eIF4E protein. The position and length of the triazole-bridge appears to play a critical role in enhancing and decreasing the affinity of the binding protein peptide analogs (4E-BP) for the target protein (eIF4E) (section 3.6.1).

A NMR/CD conformational study has been carried out on the cyclopeptides analogs containing triazolyl moiety. From preliminary data the triazole bridge

seems, as expected, to stabilize the α-helical structure of the peptides (section 3.7.1).

The completion of the synthesis and characterization of the 4E-BP cyclopeptides collection is in progress. The peptides presenting the best affinity with eIF4E will be selected for experiments of templated click cyclization free from metal catalyst and in presence of the protein (eIF4E) (section 3.5.2.4). The cell membrane permeability of the fluoresceinated peptides will also be assessed with a confocal microscope.

2 PART A: Glycation in diabetes; a post-translational modification useful for the development of biomarkers. A convergent approach aimed to diagnostic tools.

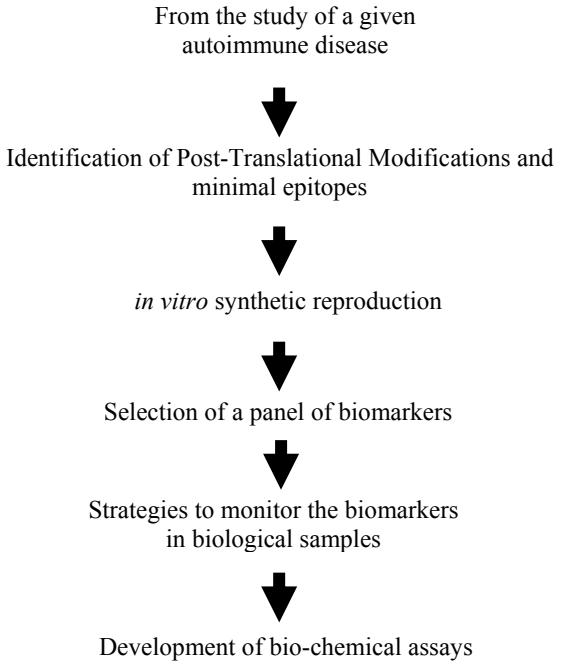
We have undertaken a convergent development of a diagnostic/prognostic tool for type I and type II diabetes. The designation of "convergent" points out that, different approaches have been forsaken in parallel in order to maximize the possibility of success (Scheme 2-1).

Our proof-of-concept is that aberrant post-translational modifications (PTMs) affecting specific proteins can trigger an autoimmune response. This process could be one of the factors contributing to autoimmune diseases development or being just a side effect of the autoimmune condition. In either case we concentrated our efforts in taking advantage of the biunivocal correlation between pathology and aberrant modifications in order to develop efficient diagnostic/prognostic tools. The second step is the selection of a panel of potential biomarkers of a given disease on the base of previously identified PTMs, synthetically reproduced.

The definition of a biological marker (biomarker) is the following:

a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention[1].

1 Biomarkers Definitions Working Group (2001) *Clinical Pharmacology and Therapeutics*, Bethesda, Md, 3.



Scheme 2-1 Convergent development of a diagnostic/prognostic tool

The third stage is to set up a biochemical experiment that allows identifying and quantifying the selected biomarker in a statistical significant number of biological samples (i.e., urines or sera). The most promising and successful of those biochemical assays will be optimized to deliver a prototype to be transferred for technical industrialization and becoming a fully reliable diagnostic/prognostic tools for final marketing.

We have focused our work on the subject of glycation in diabetes and we have selected four different types of biomarkers versus which we have directed our experiments as shown in Table 2–1.

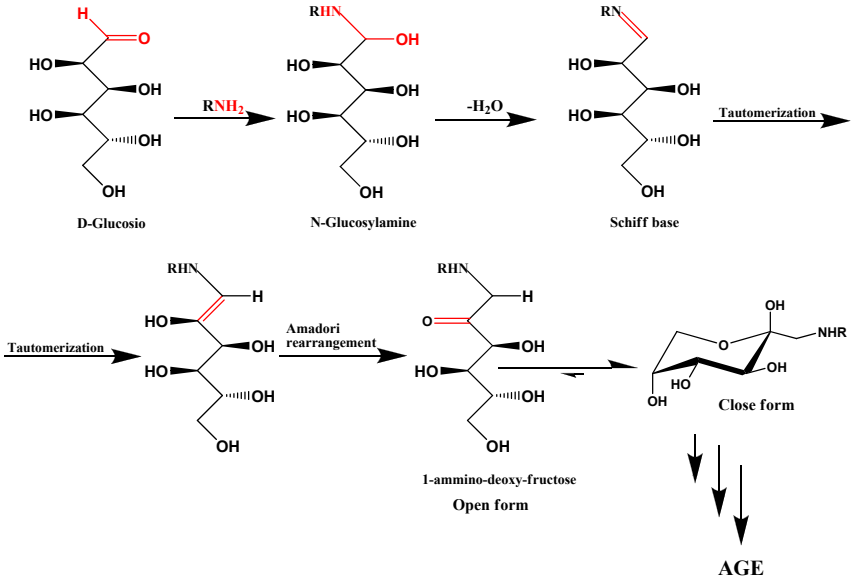
Strategy	Biomarker	Molecular tool	Application	Experiment
Direct approach (synthetic antibodies)	Glycated hCD59	Glycated hCD59(37-50) peptide antigen	Rise of mono and polyclonal anti glycated hCD59 antibodies	Identification of glycated CD59 rich tissues in diabetic patients biopsies
	Glycated hCD59	Glycated hCD59 protein (complete sequence, total synthesis by NCL)	Rise of mono and polyclonal anti glycated hCD59 antibodies with high specificity	Identification of glycated CD59 rich tissues in diabetic patients biopsies
Reverse approach (synthetic antigenic probes)	Anti-glycated hCD59 auto-antibodies	Glycated hCD59(37-50) peptide antigen	Recognition of anti-glycated hCD59 autoantibodies in diabetic patient sera	Uncompetitive ELISA screening of patient sera
	Unspecific anti-glycated auto-antibodies	Glycated antigenic probe (CSF114)	Recognition of anti-glycation autoantibodies in diabetic patient sera	Uncompetitive ELISA screening of patients sera
Amino acid analysis	Free glycated Lysine metabolites	Synthetic glycated Lysine	Synthetic glycated Lysine is used as a standard for metabolic free glycated Lysine in patient sera	UPLC-MS/HPLC analysis of patient sera
	Glycated Lysine residues (from hydrolysis of the proteins of patient sera)	Synthetic glycated Lysine	Synthetic glycated Lysine is used as a standard for metabolic free glycated Lysine in patient sera hydrolyzate	UPLC-MS/HPLC analysis of patient sera hydrolyzate

Table 2-1 Strategies for the development of a diagnostic-prognostic tool for diabetes based on the glycation process.

2.1 Glycation

Glycation of proteins through non-enzymatic reactions between glucose or other reducing sugars and reactive amino groups represents one of the most abundant processes involved in post-translational modification of proteins[2]. Spontaneous and reversible condensation of a reducing sugar and a free amino group of a protein forms an aldimine also known as the "Schiff base" that undergoes a rearrangement into the more stable ketoamine known also as the Amadori product[3] (Scheme 2.1-1). In the case of glucose, the initially formed Schiff base rearranges into the more stable 1-deoxyfructosyl moiety. Subsequent dehydration, condensation, fragmentation, oxidation, and cyclization reactions lead to the irreversible formation of Advanced Glycation End Products (AGEs). This process leads to inactivation of proteins and is involved in pathologies such as senile cataract[4], arteriosclerosis[5], vascular complications of diabetes[6], dysfunction of skin collagen[7], and neurodegenerative diseases such as Alzheimer's disease[8,9] and Parkinson disease[10].

2 Doyle, H.A. and Mamula, M.J. (2001) *Trends Immunol*., **22**, 443.
3 Ulrich, P. and Cerami, A. (2001) *Recent Prog. Horm. Res*., **56**, 1.
4 Lyons, T.J., Silvestri, G., Dunn, J.A., and Dyer, D.G. (1991) *Diabetes*, **40**, 1010.
5 Price, C.L. and Knight, S.C. (2007) *Curr. Pharm. Des.*, **13**, 3681.
6 Gugliucci, A.J. (2000) *Am. Osteopath. Assoc*., **100**, 621.
7 Avery, N.C. and Bailey, A.J. (2006) *Pathol. Biol. (Paris)*, **54**, 387.
8 Yan, S.D., Chen, X., Schmidt, A.M., Brett, J., Godman, G., Zou, Y.S., Scott, C.W., Caputo, C., Frappier, T., and Smith, M.A. (1994) *Proc.Natl. Acad. Sci. U.S.A.*, **91**, 7787.
9 Takeuchi, M. and Yamagishi, S. (2008) *Curr. Pharm. Des.*, **14**, 973.
10 Munch, G., Gerlach, M., Sian, J., Wong, A., and Riederer, P. (1998) *Ann. Neurol.*, **44**, 85.



Scheme 2.1-1 The glycation reaction.

The function of a glycated protein may be impaired if an amino group affected by glycation is in, or close to, its active site. For example, glycation of the b chains of hemoglobin gives rise to the glycated hemoglobins (HbA1), in which responsiveness to 2,3-diphosphoglycerate is decreased and oxygen affinity increased[11]. Also, glycation of the major thrombin inhibitor of the coagulation system, antithrombin III, decreases its affinity for heparin, possibly contributing to the hypercoagulable state associated with diabetes[12]. Even though proteins contain many surface amino groups, only a few are preferentially glycated. This intriguing observation was explained when the identification of glycated amino groups in proteins with known three-dimensional structure revealed that glycation preferably occurs at amino groups that are either close to an imidazole moiety or part of a Lysine doublet. Proximity ('5 Å) of an amino group to an imidazole moiety is the strongest

11 McDonald, M.J., Bleichman, M., and Bunn, H.F. (1979) *J. Biol. Chem.*, **254**, 702.
12 Ceriello, A., Giugliano, D., Quatraro, A., Stante, A., Consoli, G., Dello Russo, P., and D'Onofrio, F. (1987) *Diabete Metab.*, **13**, 16.

predictor of susceptibility to glycation[13]. This site specificity of protein glycation is the consequence of localized acid–base catalysis of the aldimineyketoamine tautomerization[14,15]. In conclusion the glycation phenomenon, thought non enzymatic, is sequence dependent

13 Fluckiger, R. and Strang, C.J. (1995) *Protein Sci.* **4**, 186.
14 Iberg, N. and Fluckiger, R. (1986) *J. Biol. Chem.* **261**, 13542.
15 Watkins, N.G., Neglia-Fisher, C.I., Dyer, D.G., Thorpe, S.R., and Baynes, J.W. (1987) *J. Biol. Chem.,* **262**, 7207.

2.1.1 *Diabetes*

Changes in human behavior and lifestyle over the last century have resulted in a dramatic increase in the incidence of diabetes worldwide. The number of adults with diabetes in the world will rise from 135 million in 1995 to 300 million in the year 2025[16] (Figure 2.1-1). Most cases will be of type II diabetes, which is strongly associated with a sedentary lifestyle and obesity[17]. This trend of increasing prevalence of diabetes and obesity has already imposed a huge burden on health-care systems and this will continue to increase in the future. Although type II diabetes is numerically more prevalent in the general population, type I diabetes is the most common chronic disease of children. But with the increasing prevalence of type II diabetes in children and adolescents, the order may be reversed within one to two decades[18].

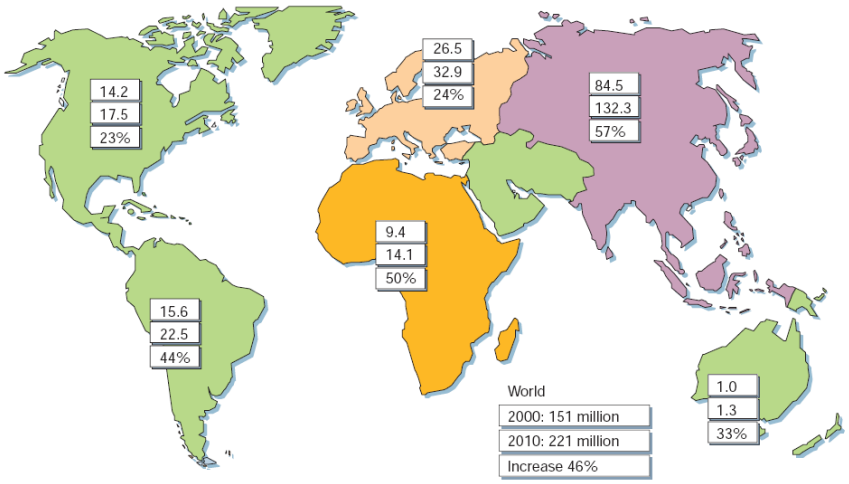


Figure 2.1-1 Number of people with diabetes (in millions) for 2000 and 2010 (top and middle values, respectively), and the percentage of increase.

16 King, H., Aubert, R.E., and Herman, W.H. (1998) *Diab. Care*., **21**,1414.
17 Kahn, B.B. and Flier, J.S. (2000) *J. Clin. Invest.*,**106**, 171.
18 Fagot-Gampagna, A. (2000) *J. Pediatr,*. **136**, 664.

Current estimates from different countries in Europe and the United States have shown that diabetes and its complications account for 8–16% of the total health costs for society and these will increase dramatically unless major efforts are made to prevent the ongoing epidemic[19].

19 Torben, H. (2002) *CURRENT SCIENCE*, **83**, 25.

2.1.1.1 Type I diabetes

Type I diabetes encompass several diabetic forms characterized by immunologically-mediated β-cell destruction, usually leading to absolute insulin deficiency and represent the 10% of all diabetic cases. Around the world there is a great variation in the type I diabetes incidence which is raising 3-5% per year due to environmental causes (Figure 2.1–2).

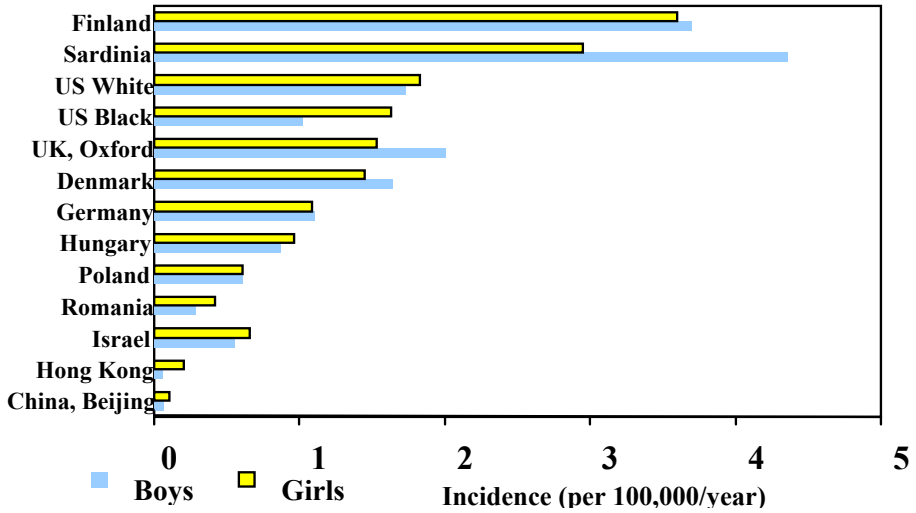


Figure 2.1-2 Type I diabetes incidence in caldron aged 0-14 other 100.000individuals.

The most common of this disease forms is known as Insulin-Dependent Diabetes Mellitus (IDDM) or juvenile-onset diabetes, results from autoimmune mediated destruction of the insulin-producing beta cells of the pancreas[20].The rate of destruction is quite variable, being rapid in some individuals and slow in others[21]. The rapidly progressive form is commonly

20 K.G., Alberti, and Zimmet, P.Z. (1998) *Diabetic Med.,* **15**, 539.
21 Zimmet, P.Z., Tuomi, T., Mackay, R., Rowley, M.J., Knowles, W., and Cohen, M. (1994) *Diabetic Med*., **11**, 299.

observed in children, but also may occur in adults[22]. The slowly progressive form generally occurs in adults and is sometimes referred to as Latent Autoimmune Diabetes in Adults (LADA). Some patients, particularly children and adolescents, may present with ketoacidosis as the first manifestation of the disease[23].Others have modest fasting hyperglycemia that can rapidly change to severe hyperglycemia and/or ketoacidosis in the presence of infection or other stress. Still others, particularly adults, may retain residual beta cell function, sufficient to prevent ketoacidosis, for many years[24].

Type I diabetes is characterized by severe complicances such as macro and microangiopathy, arteriosclerosis, retinopathy, nefropathy and neuropathy (Figure 2.1–3).

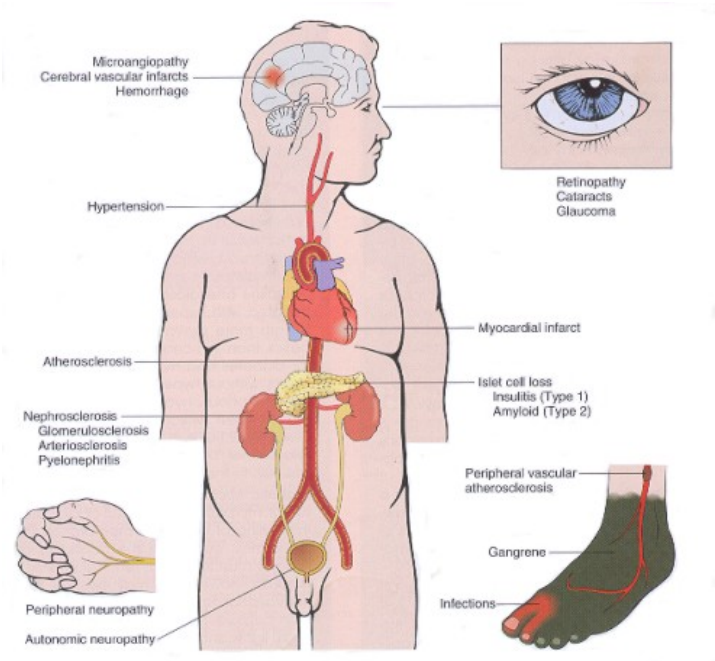


Figure 2.1-3 Type I diabetes complications.

22 Humphrey, A.R., McCarty, D.J., Mackay, I.R., Rowley, M.J., Dwyer, T., and Zimmet, P. (1998) *Diabetic Med.*, **15**, 113.
23 Japan and Pittsburgh Childhood Diabetes Research Groups (1985) *Diabetes*, **34**, 1241.
24 Zimmet, P.Z. (1995) *Diabetes Care*, **18**, 1050.

Markers of immune destruction, including islet cell autoantibodies, and/or autoantibodies to insulin, and autoantibodies to Glutamic Acid Decarboxylase (GAD65) are present in 85–90 % of individuals with type I diabetes mellitus when fasting diabetic hyperglycemia is initially detected[25]. The peak incidence of this form of type 1 diabetes occurs in childhood and adolescence, but the onset may occur at any age, ranging from childhood to the ninth decade of life[26]. There is a genetic predisposition to autoimmune destruction of beta cells, and it is also related to environmental factors that are still poorly defined. Although patients are rarely obese when they present with this type of diabetes, the presence of obesity is not incompatible with the diagnosis. These patients may also have other autoimmune disorders such as Graves' disease, Hashimoto's thyroiditis, and Addison's disease[27].

There are some forms of type 1 diabetes, called Idiopathic, which have no known etiology. Some of these patients have permanent insulinopenia and are prone to ketoacidosis, but have no evidence of autoimmunity[28]. This form of diabetes is more common among individuals of African and Asian origin. In another form found in Africans an absolute requirement for insulin replacement therapy in affected patients may come and go, and patients periodically develop ketoacidosis[29].

Studies measuring the expression of diabetes related autoantibodies in young children from birth suggest that the appearance of these markers is a major risk for the future development of type I diabetes[,30]. However, the role of

25 Verge, C.F., Gianani, R., Kawasaki, E., Yu, L., Pietropaolo, M., and Jackson, R.A., *Diabetes* (1996), **4**5, 926.
26 Mølbak, A.G., Christau, B., Marner, B., Borch-Johnsen, K., and Nerup, (1994) *J. Diabetic Med*, **11**, 650.
27 Betterle, C., Zanette, F., Pedini, B., Presotto, F., Rapp, L.B., and Monsciotti, C.M. (1983) *Diabetologia*, **26**, 431.
28 McLarty, D.G., Athaide, I., Bottazzo, G.F., Swai, A.B., and Alberti, K.G. (1990) *Diabetes,* Res Clin Pract, **9**, 219.
29 Ahre´n, B. and Corrigan, C.B. (1984) *Diabetic Med*, **2**, 262.
30 Yu, L., Robles, D.T., Abiru, N., Kaur, P., Rewers, M., and Kelemen, K., (2000) *Proc Natl Acad Sci USA*, **97**, 1701.

autoantibodies in the actual pathogenesis of type I diabetes has not been established in humans. In fact, a recent case report showed the development of type 1 diabetes in a patient with X linked agammaglobulinaemia, suggesting that autoantibodies are not needed for either the initiation or the progression of type I diabetes[31]. In general, type I diabetes is considered primarily a T cell mediated disease, and extensive evidence exists in both man and mouse to support this. Examination of islet tissue obtained from pancreatic biopsy from patients with recent onset type I diabetes confirms insulitis, with the presence of an infiltrate composed of CD4 and CD8 T lymphocytes, B lymphocytes, and macrophages, suggesting that these cells have a role in destruction of the _ cells[32]. Early studies in mice showed that anti-CD3 treatment prevented diabetes, and a trial using humanized anti-CD3 antibody in patients with new onset type I diabetes is under way[33]. Figure 2.1–4 describes a general model of β-cell destruction leading to type I diabetes. The initial interaction of genes and environmental factors seem to trigger an immune mediated response, with the appearance of autoantibodies as the first sign of β-cell destruction, followed eventually by the loss of the first phase insulin response. The progression to overt diabetes resulting in significant β-cell destruction is triggered by the development of a more aggressive T cell phenotype and a change in the Th1 to Th2 balance towards a more proinflammatory milieu. The expression of FasLigand on cytotoxic T cells also marks the progression to overt diabetes. Examination of islets during insulitis suggests that Fas mediated apoptosis occurs and therefore provides one possible mechanism of β-cell destruction[34].

---

31 Martin, S., Wolf-Eichbaum, D., Duinkerken, G., Scherbaum, W.A., Kolb, H., and Noordzij, J.G. (2001) *N. Engl. J. Med.*, **345**, 1036.
32 Imagawa, A., Hanafusa, T., Itoh, N., Waguri, M., Yamamoto, K. and Miyagawa, J. (1999) *Diabetologia*, **42**, 574.
33 Herold, K.C., Hagopian, W., Auger, J.A., Poumian-Ruiz, E., Taylor, L., and Donaldson, D. (2002) *N. Engl. J. Med.*, **346**, 1692.
34 Foulis, A.K., Liddle, C.N., Farquharson, M.A., Richmond, J.A., and Weir, R.S (1986) *Diabetologia*, **29**, 267.

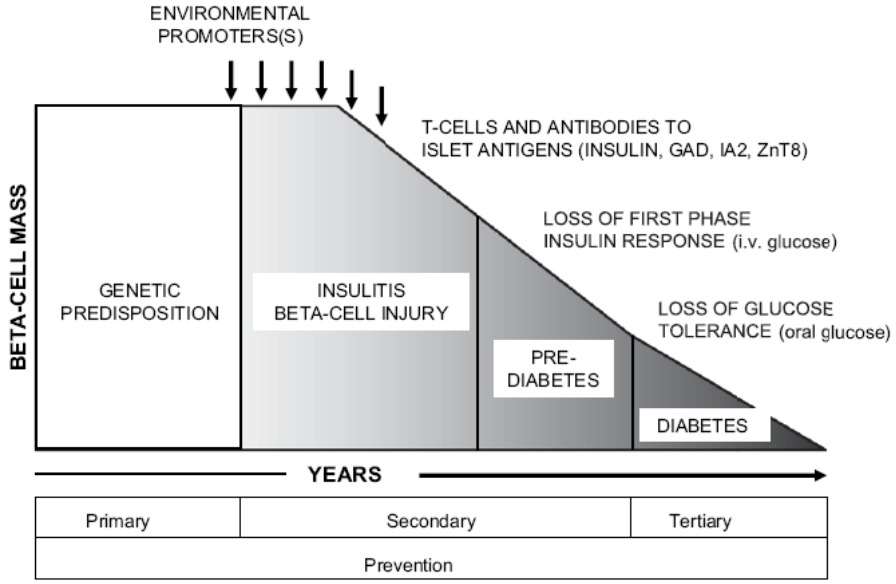


Figure 2.1-4 Model of phatogenesis and natural history of type I diabetes

Individuals with underlying islet autoimmunity who are at risk for type I diabetes can be identified years before symptomatic presentation by the presence of circulating autoantibodies to specific islet antigens: (pro)insulin, the molecular weight 65,000 isoform of glutamic acid decarboxylase (GAD65), tyrosine phosphatase-like insulinoma antigen 2 (IA2) and beta cell-specific zinc transporter 8. As can be seen in Figure 2.1-4 the preclinical factors mentioned above lay on the second half of the secondary prevention window. The environmental promoters (mostly un-known) lay upstream. Thus the decision to focus on glycation that can be considered as a factor characteristic of the early onset of the disease and that could be the core of a biochemical assay for type I diabetes early diagnosis.

2.1.1.2 Type II diabetes

Type II diabetes is not a single entity but a group of metabolic disorders characterized by the clinical feature of hyperglycemia which is due to insulin resistance or elative insulin deficiency, either of which may be present at the time that diabetes becomes clinically manifest[35] (Figure 2.1-5).

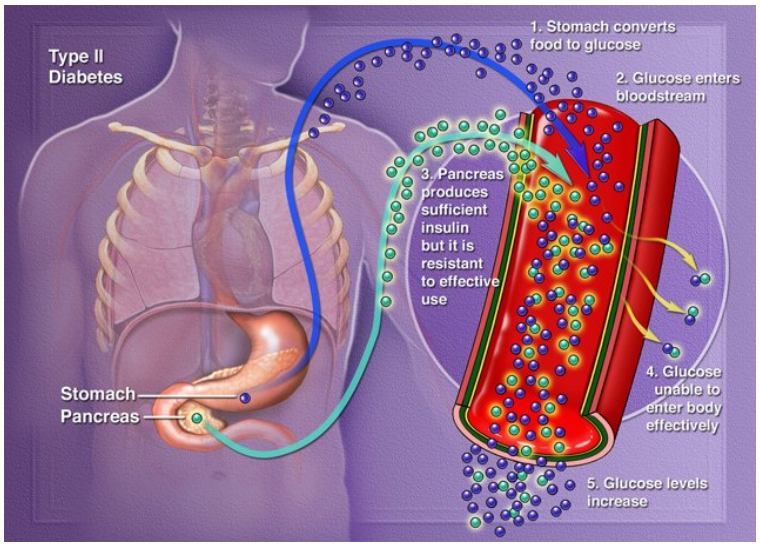


Figure 2.1-5 Pathway of the raising of glucose blood level in type II diabetes

There is continuing debate as to the primary aetiological factor for the syndrome. Genetic factors, visceral obesity, insulin resistance and endothelial dysfunction may all contribute either solely or jointly[36].

Type II diabetes complicances are very similar to those given by type I diabetes, namely the macro and microangiopathy, arteriosclerosis, retinopathy, nephropathy and neuropathy (Figure 2.1–4). This is due to the fact that such

35 Narayan, K.M. (2003) *JAMA*, **290**, 1884.
36 Reaven, G. (1988) *Diabetes,* **37**, 1595.

complications are primary generated by an excess of glucose[37] in the blood which is a common feature of all diabetic types. Glucose is responsible for glycation of the plasma soluble proteins and generation of the Advanced Glycation End Products, which play a key role in the complicances.

Type II diabetes is increasingly common, indeed epidemic, primarily because of increases in the prevalence of a sedentary lifestyle and obesity. Prevention of complications is a key issue because of the huge premature morbidity and mortality associated with the disease. The possibility of preventing the disease by interventions that affect the lifestyles of subjects at high risk for the disease is now the subject of numerous studies These have focused on people with Impaired Glucose Tolerance (IGT) commonly referred as "pre-diabetics"[38]. IGT is defined as hyperglycemia (with glucose values intermediate between normal and diabetes) following a glucose load[39], and affects at least 200 million people worldwide. It represents a key stage in the natural history of type II diabetes as these people are at much higher future risk than the general population for developing diabetes[40].

The type II diabetes diagnosis is established when it is observed a raise of the blood glucose level which is kept within a very narrow range of 7-12 mmol/L under normal physiological conditions[41] (Table 2.1-1).

37 Stitt, A.W., Jenkins, A.J., and Cooper, M.E. (2002) *Expert Opin. Investig. Drugs.*, **11**, 1205.
38 Eastman, R., Javitt, J., Herman, W., Dasbach, E. and Harris, M. (1996) In *Diabetes Mellitus: A Fundamental and Clinical Text,* eds Le-Roith, D., Taylor, S. and Olefsky, J. (Lippincott-Raven, New York) 621.
39 World Health Organization. Definition, Diagnosis and Classification of Diabetes mellitus and itsComplications, (1999) Department of Noncommunicable Disease Surveillance, Geneva,12.
40 Shaw, J., Hodge, A., de Courten, M., Chitson, P., and Zimmet, P. (1999) *Diabetologia,* **42**, 1050.
41 Zimmet, P., Alberti, K.G., and Shaw, J. (2001) *NATURE*, **414**, 782.

	Glucose concentration (mmol l^{-1})			
	Plasma		Whole blood	
	Venous	Capillary	Venous	Capillary
Diabetes mellitus				
Fasting	⩾7.0	⩾7.0	⩾6.1	⩾6.1
2-h post-glucose load	⩾11.1	⩾12.2	⩾10.0	⩾11.1
Impaired glucose tolerance				
Fasting concentration	<7.0	<7.0	<6.1	<6.1
2-h post-glucose load	7.8–11.0	8.9–12.1	6.7–9.9	7.8–11.0
Impaired fasting glucose				
Fasting	6.1–6.9	6.1–6.9	5.6–6.0	5.6–6.0
2-h post-glucose load	<7.8	<8.9	<6.7	<7.8

Table 2.1-1 Values for diagnosis of diabetes and other types of hyperglycemia

Interestingly the measurement of glycated hemoglobulin (HbA1c) is no more recommended as a diagnostic test for diabetes because of lack of standardized methodology resulting in varying nondiabetic reference ranges among laboratories[42].

The latter consideration adds value to our commitment of developing a new diagnostic/prognostic tool for diabetes based on the recognition of specific glycated sequences, amino acids and metabolites.

42 The Expert Commitee on the Diagnosis and Classification of Diabetes Mellitus (2003) *Biabetes Care,* **26**, 15.

2.1.2 *Correlation between glycation of hCD59 and complement-mediated diabetes complicances*

Humans are particularly prone to develop proliferative micro and macrovascular disease that mediating some of the most common long-term complicances of diabetes mellitus[43]. These vascular complicances of diabetes are caused by elevated blood glucose levels over long periods of time[44]. Understanding the molecular mechanisms that link hyperglycemia and the vascular proliferative disease in humans is essential for designing adequate animal models and therapeutic strategies for a condition that represents a leading cause of morbidity and mortality in the adult population. Glycation is considered a major pathophysiological mechanism causing tissue damage in diabetic subjects[45]. Reports of increased deposition of the membrane attack complex of complement (MAC) in blood vessels and kidneys of diabetic patients[46,47] suggest that there may be a link between complement activation and the development of chronic proliferative diabetic complications.

The complement system consists of a group of >12 soluble proteins that interact with one another in three distinct enzymatic activation cascades known as the classical, alternative, and lectin pathways. All activation pathways converge to form the membrane attack complex (MAC). The MAC is a circular polymer of 12–18 monomers of the C9 complement protein with the capacity to insert into cell membranes and form a transmembrane pore of an effective internal radius of 5–7 nm. Influx of salt and water through the MAC pore induces colloidosmotic swelling and lysis of MAC-targeted cells such as gram-negative bacteria or heterologous erythrocytes. We have

43 Duhault, J., and Koenig-Berard, E. (1997) *Therapie,* **52**, 375.
44 Nathan, D.M. (1996) *Ann. Intern. Med.* **124**, 86.
45 Brownlee, M., Vlassara, H., and Cerami, A. (1984) *Ann. Intern. Med.,* **101**, 527.
46 Weiss, J.S., Sang, D.N., and Albert, D.M. (1990) *Cornea,* **9**, 131.
47 Falk, R.J., Sisson, S.P., Dalmasso, A.P., Kim, Y., Michael, A.F., and Vernier, R.L. (1987) *Am. J. Kidney Dis.,* **9**, 121.

demonstrated that during the assembly of the MAC pore, there is a phase when the MAC pore is transient and reversible, permitting opening and closing/resealing of the plasma membrane[48,49,50]. These transient MAC pores can generate significant changes in the membrane permeability of autologous cells without compromising their viability and thereby mediate physiological and/or pathological responses[51].

Indeed, the MAC stimulates proliferation of fibroblasts and smooth muscle, mesangial, and other cells[52,53], in part by releasing growth factors such as basic fibroblast growth factor and platelet-derived growth factor from MAC-targeted endothelium[54]. The MAC also induces increased synthesis of extracellular matrix proteins by mesangial cells[55]. Thus, increased MAC deposition in diabetic tissues may induce the release of growth factors that would stimulate cell proliferation in the vascular wall and contribute to the development of vascular proliferative disease. In the kidneys, MAC-induced vascular proliferation and expansion of the extracellular matrix may contribute to the characteristic glomerulosclerosis of diabetic nephropathy[56]. Increased MAC deposition in diabetes is well documented but the underlying mechanism is poorly understood. Autologous MAC deposition is normally restricted because cells express complement regulatory membrane proteins such as DAF and CD59, which limit complement activation and MAC formation. In particular, CD59, a glycan phosphatidylinositol-linked membrane protein that is universally expressed in cells, restricts MAC assembly by interacting with the

48 Halperin, J.A., Nicholson-Weller, A., Brugnara, C., and Tosteson, D.C. (1988) *J. Clin. Invest.,* **82**, 594.
49 Halperin, J.A., Taratuska, A., Rynkiewicz, M., Nicholson-Weller, A. (1993) *Blood,* **81**, 200.
50 Acosta, J.A., Benzzaquen, L.R., Goldstein, D.J., Tosteson, M.T., and Halperin, J. (1996) *Mol. Med.,* **2**, 755.
51 Nicholson-Weller, A. and Halperin, J. (1993) *Immunol Res.,* **12**, 244.
52 Halperin, J.A., Taratuska, A. and Nicholson-Weller, A. (1993) *J. Clin. Invest.* **91**,1974.
53 Torzewski, J., Oldroyd, R., Lachmann, P., Fitzsimmons, C., Proudfoot, D., and Bowyer, D. (1996) *Arterioscler. Thromb. Vasc. Biol.* **16**, 673.
54 Benzaquen, L.R., Nicholson-Weller, A., and Halperin, J.A. (1994) *J. Exp. Med*., **179**, 985.
55 Wagner, C., Beer, M., Rother, K., and Hansch, G.M. (1994) *Exp.Nephrol.* **2**, 51.
56 Ziyadeh, F.N. (1993) *Am. J. Kidney Dis.,* **22**, 736.

terminal complement proteins C8 and C9, thus preventing C9 polymerization[57] (Figure 2.1-6).

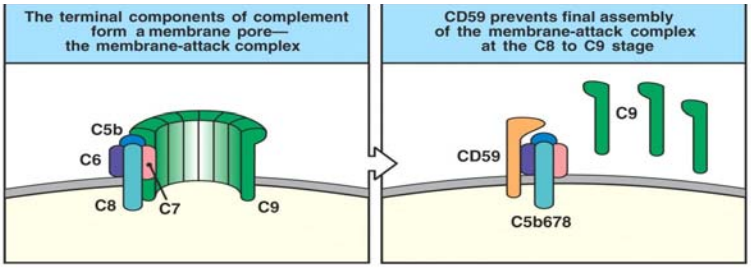


Figure 2.1-6 Complement/MAC regulation by CD59.

It has been demonstrated that hCD59 is prone to glycation at a specific Lysine site thanks to a glycation motif or consensus sequence[58]. The NMR structure of human CD59[59] reveals that Lysine-41 (K41) should be susceptible to glycation because of its critical proximity to the unique Histidine, Histidine-44 (H44), in the protein Figure 2.1-7. Furthermore, the fact that K41 is adjacent to tryptophan-40 (W40), a conserved amino acid that is essential for CD59 function[60,61] (18, 19), suggests that glycation of K41 may hinder the activity of CD59.

57 Ninomiya, H. and Sims, P.J. (1992) J. Biol. Chem. **267**, 13675.
58 Acosta, J., Hettinga, J., Fluckiger, R., Krumrei, N., Goldfine, A., Angarita, L., Halperin, J. (2000) *PNAS*, **97**, 5450.
59 Fletcher, C.M., Harrison, R.A., Lachmann, P.J., and Neuhaus, D. (1994) *Structure,* **2**, 185.
60 Bodian, D.L., Davis, J.S., Morgan, B.P., and Rushmere, N.K. (1997) *J. Exp. Med.,* **185**, 507.
61 Yu, J., Abagyan, R., Dong, S., Gilbert, A., Nussenzweig, V. and Tomlinson, S. (1997) *J. Exp. Med.*, **185**, 745.

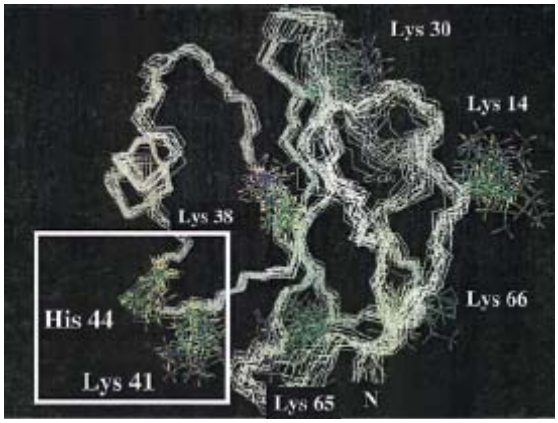


Figure 2.1-7 NMR structure of the protein backbone of human CD59. The figure shows the 20 lowest energy structures of human CD59 with all Lysine side chains and H44. The structures were superimposed for the backbone of the β-turn 41–44. The square highlights the K41–H44 glycation motif. The K41 is within 5.91 ± 1.44 Å of the D1 imidazolic nitrogen of H44.

Haplerin et al. extensively demonstrated that selective Lys 41 glycation deactivated hCD59 and destabilize MAC homeostasis:

(i) in vitro glycation of human CD59 inhibits its homologous restriction activity, (ii) replacement by site-directed mutagenesis of either K41 or H44 abolishes the sensitivity of human CD59 to glycation–inactivation, (iii) glycation of CD59 in human RBC (hRBC) increases their sensitivity to MAC-mediated lysis, (iv) glycation of human umbilical vein endothelial cells (HUVEC) renders them more sensitive to MAC-mediated growth factor release, and (v) glycated CD59 is present in human urine, indicating that CD59 is glycated in vivo

Thus glycation inhibiting the principal MAC inhibitor destabilize MAC regulation and promote diabetic complication. Glycation as a an aberrant-post-translational modification it is the key molecular factor that links complement to diabetic complications

Importantly, the H44 residue of human CD59 is not present in any other species in which CD59 has been sequenced (Table 2.1-2).

	Residue at position									
Species	38	39	40	41	42	43	44	45	46	47
Human	K	C	**W**	**K**	F	E	**H**	C	N	F
Mouse	R	C	**W**	**K**	Q	S	D	C	H	G
Rat	Q	C	**W**	**R**	F	S	D	C	N	A
Baboon	Q	C	**W**	**K**	F	A	N	C	N	F
African green monkey	Q	C	**W**	**K**	F	A	N	C	N	E
Owl monkey	R	C	**W**	**K**	F	E	D	C	T	F
Squirrel monkey	R	C	**W**	**K**	F	D	D	C	S	F
Pig	R	C	**W**	**R**	F	D	E	C	N	F
Marmoset	R	C	**W**	**K**	F	E	D	C	T	F

Table 2.1-2 Alignment of CD59 amino acid sequences from different species around residue W40

It has been propose that the presence of the glycation motif K41–H44 in human CD59 provides a possible molecular explanation for the propensity of humans to develop the combination of vascular complications that characterize human diabetes. Indeed most of the existing diabetic animal models do develop diabetes, for example iperglycemia, but don't show any or little trace of diabetic complication.

2.2 Building blocks for the synthesis of post-translationally modified glycated peptides and proteins.

Glycation of proteins through non-enzymatic reactions between glucose or other reducing sugars and reactive amino groups represent one of the more abundant processes involved in post-translational modification of proteins[2]. Spontaneous and reversible condensation of a reducing sugar and a free amino group of a protein forms an aldimine also known as the Schiff base that undergoes a rearrangement into the more stable ketoamine known also as the Amadori product[3] In the case of glucose, the initially formed Schiff base rearranges into the more stable 1-deoxyfructosyl moiety.

Growing evidence suggests that glycation occurs preferentially at specific glycation motifs characterized by acidic amino acids, mainly glutamate and Lysine residues that catalyze the glycation of nearby Lysines[62,63]. Proximity to histidine either in the primary or secondary structure was also suggested to promote glycation of adjacent Lysines[64,65]. Recent interest to fully characterize the glycation products and to use them as biomarkers and antigens for diagnosis and prognosis of disease, monitoring its progress and evaluation of the efficiency of therapy generated the need for glycated peptides representing the glycation motifs specifically modified by the 1-deoxyfructosyl. Today, syntheses of site-specific Amadori modified peptides are carried out on partially protected synthetic peptides in which only the lysyl residues designated for glycation are exposed while the rest are protected[66,67,68,69] . This

62 Johansen, M.B., Kiemer, L., and Brunak, S. (2006) *Glycobiology*, **16**, 844.
63Venkatraman, J., Aggarwal, K., and Balaram, P. (2001) *Chem. Biol*., **8**, 611.
64 Baynes, J.W., Watkins, N.G., Fisher, C.I., Hull, C.J., Patrick, J.S., Ahmed, M.U.; Dunn, and J.A.; Thorpe, S.R. (1989) *Clin. Biol. Res*., **304**, 43.
65 Acosta, J., Hettinga, J., Fluckiger, R., Krumrei, N., Goldfine, A., Angarita, L., and Halperin, J., (2000) *Proc. Natl. Acad. Sci. U.S.A*., **97**, 5450.
66 Frolov, A., Singer, D., and Hoffmann, R. (2006) *J. Pept. Sci*., **12**, 389.
67 Frolov, A., Singer, D., and Hoffmann, R. (2007) *J. Pept. Sci*., **13**, 862.

approach involves orthogonal protection strategies and suffers from low yields and elaborated purification schemes.

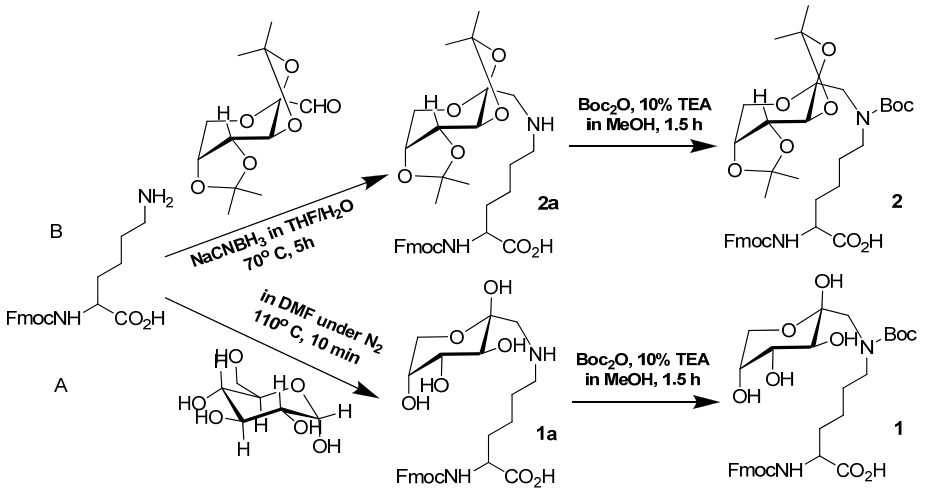
Stepwise assembly of site-specific Amadori modified peptides requires N^{α}-protected-N^{ε}-glycated-Lys s and represents a fully controlled and effective synthetic strategy. Herein, we report the synthesis, purification, and characterization of N^{α}-Fmoc, N^{ε}-Boc, N^{ε}-(1-deoxyfructosyl)Lysine s needed for Fmoc-based solid phase synthesis of Amadori modified peptides.

This study offers a controlled side-specific introduction of N^{ε}-Amadori-modified Lys residue into synthetic peptides during a stepwise assembly either in solution or solid phase methodologies. This strategy will overcome tow major problems associated with the modification of already assembled peptides: 1) lack of site-specificity in the introduction of the modification. 2) elaborate orthogonal protection scheme in an effort to achieve site-specificity. And 3) extremely low yields and complicated reaction mixtures due to side reactions following the direct thermal glycation. Adapting the conditions for generating Amadori peptides by direct thermal glycation in the presence of excess D-glucose[66] to the direct glycation of N^{α}-Fmoc-Lysine led to the synthesis of N^{α}-Fmoc-Lys[N^{ε}-1-deoxyfructosyl)]-OH (**1a**) in 67% yield (Scheme 2.2–1, pathway A). Preliminary attempt to use **1a** as a in stepwise assembly of peptides suggested that further protection of the ε-amino by the orthogonal Boc group may eliminate some of the observed side products. To this end Boc protection of **1a** yielded the pure N^{α}-Fmoc-Lys[N^{ε}-1-deoxyfructosyl, N^{ε}-Boc)]-OH (**1**) in 45% yield.

---

68 Stefanowicz, P., Kapczynska, K., Kluczyk, A., and Szewczuk, Z. (2007) *Tetrahedron Lett*., **48**, 967.

69 Jakas, A., Vinkovic, M., Smrecki, V., Sporec, M., and Horvat, S. (2008) *J. Pept. Sci*., **14**, 936.



Scheme 2.2-1. Synthesis of N$^{\alpha}$-Fmoc-Lys[N$^{\varepsilon}$-1-deoxyfructosyl, N$^{\varepsilon}$-Boc)]-OH (**1**) and N$^{\alpha}$-Fmoc-Lys[N$^{\varepsilon}$-(2,3;4,5-di-O-isopropylidene-1-deoxyfructosyl, N$^{\varepsilon}$-Boc)]-OH (**2**).

Anticipating that quantitative incorporation of **1** may require highly activated *N*$^{\varepsilon}$-Amadori modified building block and extended reaction times we sought the synthesis of an exhaustively protected product in which in addition to the primary and secondary amino function we also protected the hydroxyls on the carbohydrate moiety. To this end, reductive alkylation of the *N*$^{\alpha}$-Fmoc-Lys-OH by 2,3:4,5-di-*O*-isopropylidene-aldehydo-β-D-arabino-hexos-2-ulo-2,6-pyranose[70,71,72] in the presence of $NaCNBH_3$ led to the formation of **2a** in 22% (Scheme 2.2–1, pathway B). The pure fully protected *N*$^{\alpha}$-Fmoc-Lys[*N*$^{\varepsilon}$-(2,3;4,5-di-*O*-isopropylidene-1-deoxyfructosyl, *N*$^{\varepsilon}$-Boc)]-OH (**2**) was obtained in 67% yield. Figure 1 depicts the analytical RP-HPLC tracings obtained for the purified s **1** and **2**, and their precursors **1a** and **2a**.

In conclusion, we have developed N$^{\alpha}$-Amadori-containing N$^{\alpha}$-Fmoc-Lys-OH derivatives as new s for the synthesis of site specific Amadori-modified peptides. Contrary to previously published post-synthetically Amadori-

70 Brady, R.F. Jr. (1970) *Carbohydr. Res*., **15**, 35.
71 Cubero, I., and Lopez-Espinosa, M. (1990) *Carbohydr. Res*., **205**, 293.
72 Yuasa, Y., Ando, J., and Shibuya, S. (1996) *J. Chem. Soc., Perkin Trans. 1*, **343**, 793.

modified peptides, these building blocks will enable efficient and site specific incorporation of a major post-translational modification into bioactive and antigenic peptides by stepwise assembly.

2.2.1 *Peculiarities of the N^{ε}-1-deoxyfructosyl sugar moiety*

We have observed for those bulding block products carrying the unprotected sugar moiety, **1** and **1a**, a curious behavior on LC-MS. The mass chromatogram appears as a collection of peaks placed at precise intervals (Figure 2.2–1). We have realized that such peaks reproduce the fragmentation pattern of 1-deoxyfructosyl derivates reported in literature for ESI-MS-MS[73] (Scheme 2.2–2). Basely the 1-deoxyfructosyl residue looses water and formaldehyde units generating the oxonium, bis-dehydrated, pyrylium, furylium and immonium ions. However in our measures we have use the standard conditions of mild ionization of LC-ESI-MS which don't fragment most organic molecules.

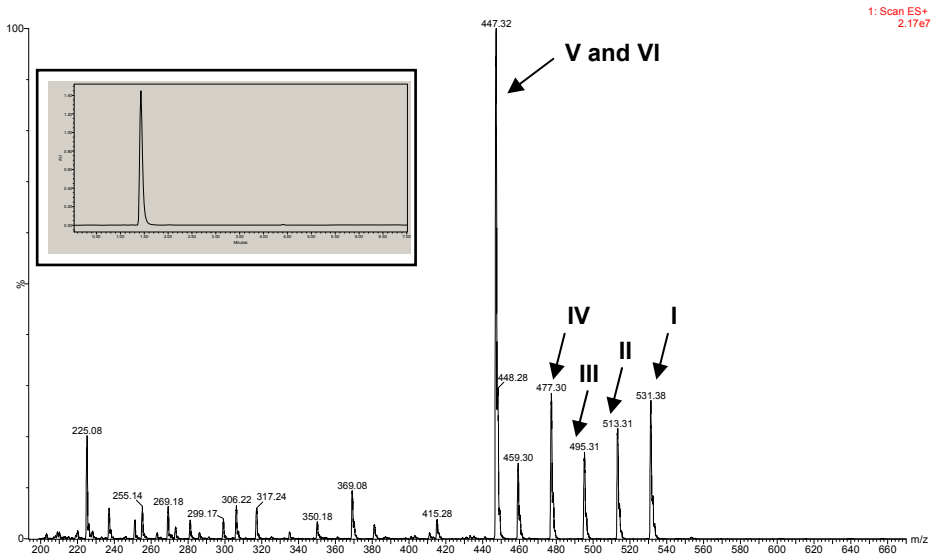
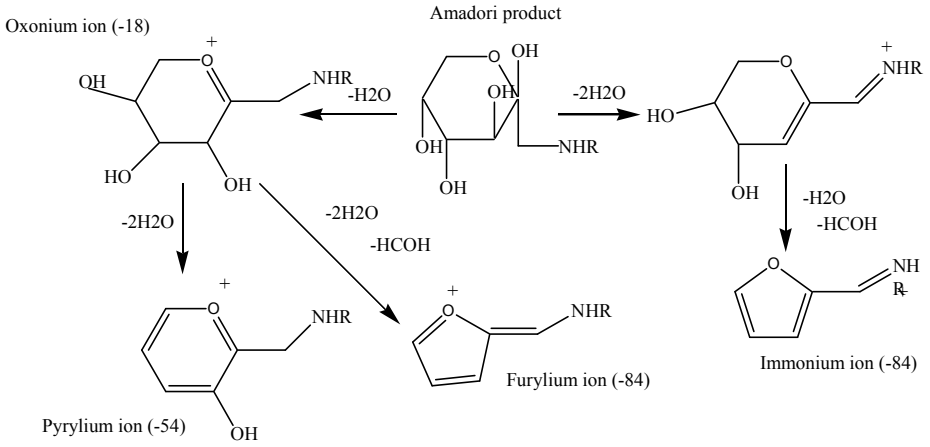


Figure 2.2-1 ESI-MS chromatogram of Nα-Fmoc-Lys(Nε-1-deoxyfructosyl)-OH, product **1a**, with the labeled peaks of the fragmentation ions [531=Amadori(I), 513= oxonium (II), 495=-2H2O (III), 477= pyrylium (IV), 447=furylium (V) and immonium (VI)]; in the small box is reported the analytical HPLC chromatogram of **1a**.

73 Frolov, A., Hoffmann, P., and Hoffmann, R. (2006) *J. Mass Spectrom.*, **41**, 1459.

From this consideration we speculate that the sugar moiety is particularly sensitive to ionization.



Scheme 2.2-2 Mass Fragmentation pattern of the Amadori forms(531=Amadori, 513= oxonium, 495=-2H_2O, 477= pyrylium, 447=furylium and immonium).

Another interesting feature has been displayed by glycated building block **1**, with the free glucose moiety and the Boc protection on the ε-amino. The LC chromatograms show a characteristic pattern of four peaks with exactly the same mass (Figure 2.2–2). Despite numerous attempts it has been impossible to isolate the different isomers because each fraction collected generated the same pattern regardless of the solvent and/or pH conditions applied. We speculate, in accordance from what was previously reported for short post-synthetically glycated peptides[74,75], that the incorporated sugar moiety presents, in solution, several tautomeric forms in equilibrium similarly to what free sugar do in solution (Scheme 2.2-3). In fact, the saccaridic unit of reducing sugars is in equilibrium between an open form (thermodynamically un-favored, less than 1%) and several closed forms, notably α-pyranose, β-

74 Horvat, S. and Jakas, A. (2004) *J. Pept. Sci*., **10**, 119.
75 Jakas, A., Katic, A., Bionda, N., and Horvat, S. (2008) *Carbohydr. Res*., ***343***, 2475.

pyranose, α-furanose and β-furanose forms (with the pyranose adducts generally being the most favoured, especially the β form).

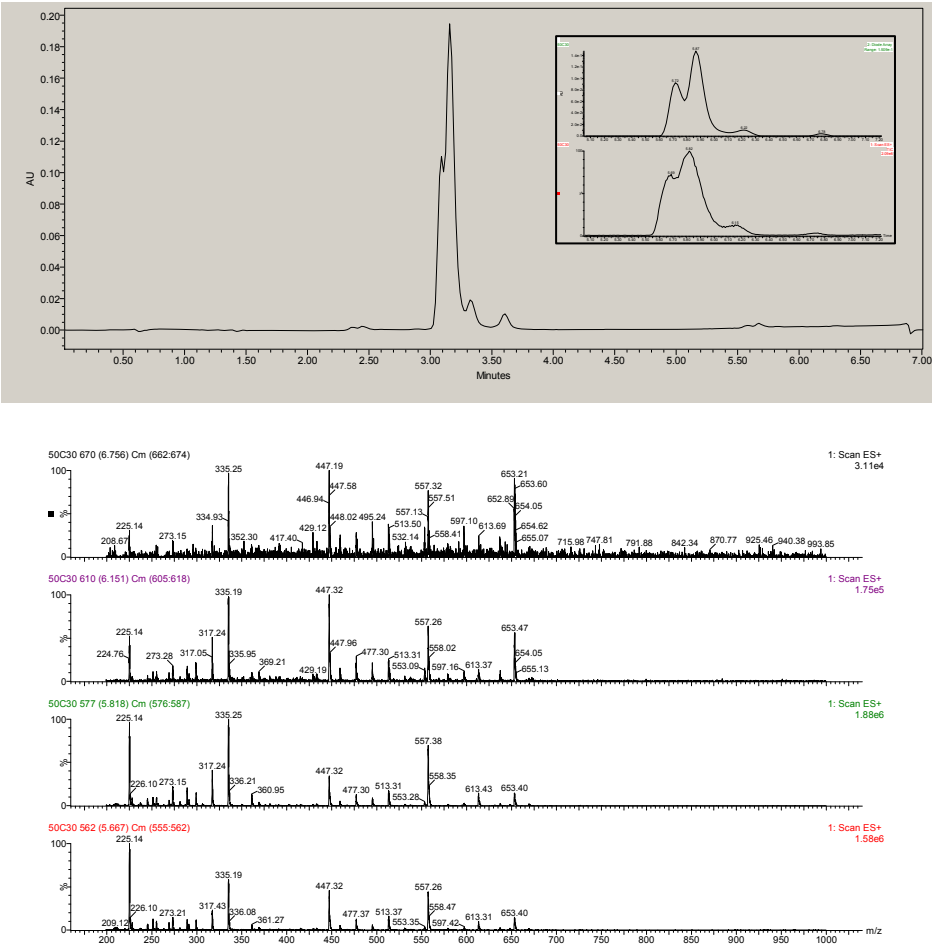
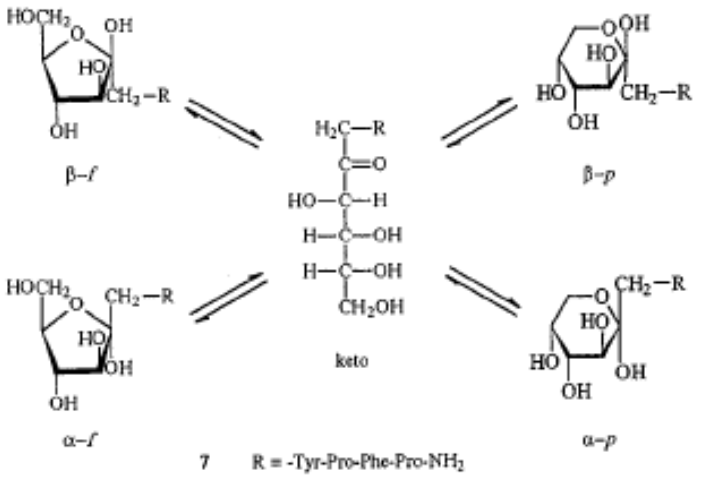


Figure 2.2-2 Above is reported the analytical HPLC chromatogram of product **1** with the zoom of the peak region from LC-MS. Below is reported the mass spectrum of the four peaks.

Interestingly among the four glycated Lysine derivate we have generated, only product 1 displayed such putative tautomerization. Our explanation is that in the case of products **2** and **2a** the equilibrium of isomeric forms is suppressed by the isopropylide protection that doesn't allow the sugar moiety to assume the open form and sub sequentially to tautomerize. However, in the case of



Scheme 2.2-3 In solution tautomeric equilibrium of the glycated adducts.

analog **1a** we have seen none of the tautomerization regardless of the sugar being free from protection. We speculate that the chemical environment influences the tautomerization allowing it to be a thermodynamically driven process or a kinetically driven one. With a simple glycated Lysine we are in the first case, thermodynamic control meaning that the sugar moiety is almost exclusively in the β-pyranose form and the LC chromatogram shows only one peak. With hindered, semi-protected glycated Lysines and with glycated peptides we are in the second case, kinetical control that makes the closed forms over than β-pyranose less un-favored and thus detectable as separate peaks in LC. In this theoric frame, the peaks are not actually the isomeric forms but an effect of the different distribution isomers and each peak area should be proportional to an isomer relative probability of forming.

Finally regarding the purification of Amadori building block, reverse phase flash chromatography (RPFC) proved to be more efficient than direct phase which may be used only with the fully-protected building block (product **2**) with is the less polar one (Figure 2.2–3). Products **1** and **1a**, both with free

hydroxyls, were not separated by RPFC without the presence of acid in the solvents and appear as a single peak However when acid (TFA 0.1-0.05 %) is added the two peaks separates accordingly to the analytical HPLC behavior. Probably without acid the products are in form of corresponding TEA salts while after the TFA addition the free acids are formed and polarity is more relevant in the chromatographic separation.

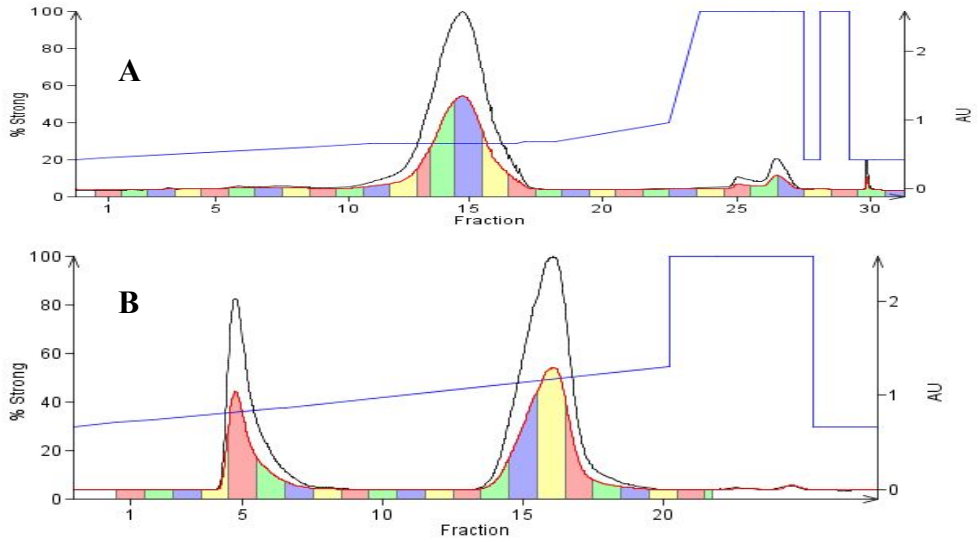


Figure 2.2-3 Above is reported the RP flash chromatogram of a mixture of product **1** and **1a** without acid in the solvents (A) and with acid (B).

2.2.2 *Nα-Fmoc-Lys[Nε-(2,3:4,5-Di-O-Isopropylidene-1-Deoxyfructopyranosyl),Nε-Boc]-OH*

As expected, the fully protected building block, bearing a Boc protection on the ε-amino and two isopropilidene protections on the vicinal hoxydryls of the sugar moiety, proved to be completely stable to the condition of solid phase peptide elongation by Fmoc/tBu strategy.

The sugar moiety is unaffected by the concentrated TFA treatment required for cleavage of the peptide from the resin and side chain deprotection. The isopropilidene protection however are only partially removed by 2-3 hours TFA exposure and the resulting crude peptide (Figure 2.2–4 A) is a mixture of fully deprotected sugar product, mono-protected isopropylidene form and di-protected form. The latter two adducts at progressively longer retention time compare to the unprotected one.

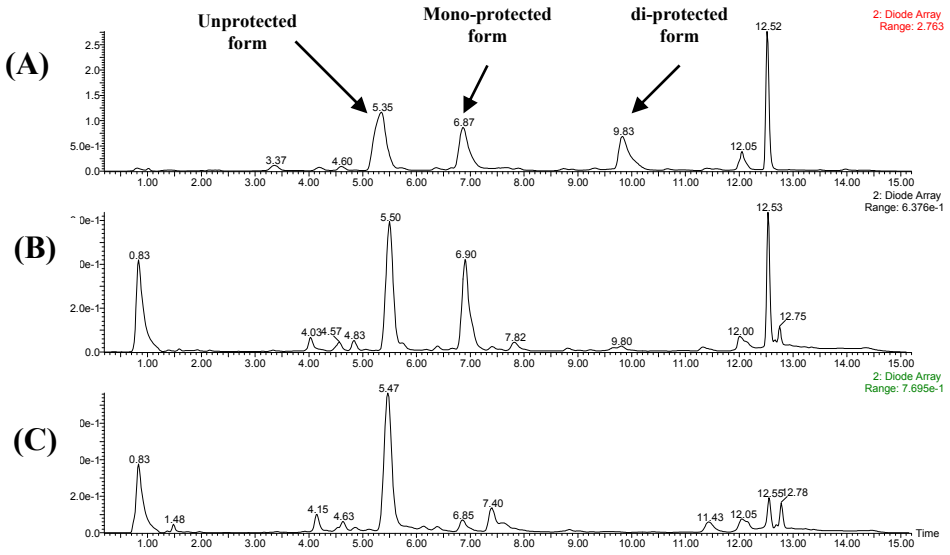


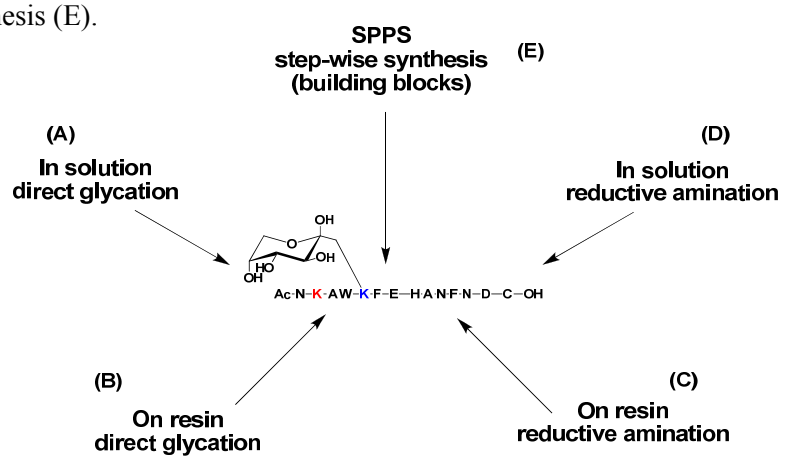
Figure 2.2-4 LC-MS chromatogram of a glycated peptide cleavage after 30 min (A), 2h (B) and second TFA treatment (overall 4h) (C).

An additional TFA treatment is required to completely remove the isopropylidene protection (Figure 2.2–4 B and C). Scavenger must be added to the deprotecting mixture to prevent undesired alkylation due to the presence of carbocations generated in the previous cleavage.

In conclusion the glycated building block represents a highly relabile tool for introducing glycation modifications in specific positions of a peptide sequence

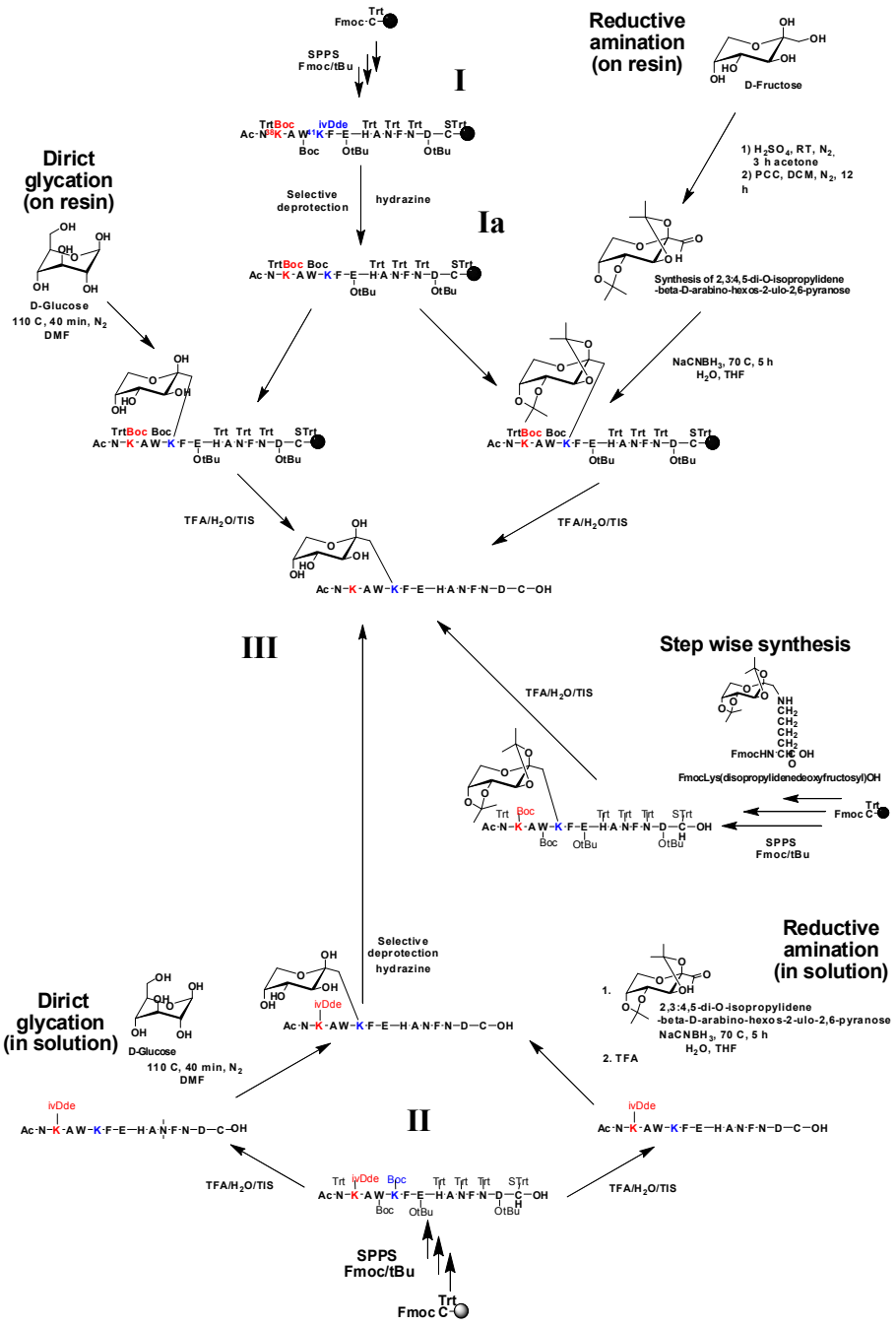
2.2.3 *Systematic approach to the synthesis of a glycated hCD59 peptide antigen*

The ultimate goal of this project was the generation of specific glycated-CD59 antibodies to be used as diagnostic/prognostic tool for diabetes. In order to produce the desired glycated CD59 antigen, to be use for mice and rabbit immunization, five different convergent strategies were carried out (Scheme 2.2–4): (A) in solution direct glycation, (B) on resin direct glycation, (C) on resin reductive amination, (D) in solution reductive amination and stepwise synthesis (E).



Scheme 2.2-4 Convergent strategies for the synthesis of glycated hCD59(37-50)

A first problem that we had to address was the requirement of selectivity for glycation of hCD59(37-50) Lysine41 since the sequence contains the close by Lysine38. Post-synthetic approaches (A,B,C and D) are not selective and require orthogonal protection schemes to achieve a selective glycation on a given position otherwise all the accessible amino groups will react in similar way. On-resin post-synthetic modification (B and C) was achieved by protecting Lysine 41 orthogonally to



Scheme 2.2-5 Synthetic routs of the five convergent strategies attempted.

those protections used for the other side chains in order to be able to selectively deprotect Lys41 leaving al the other protections in place so that

only there the glycation could take place (Scheme 2.2-5). This was achieved using the 1-(4,4'-dimethyl-2,6-dioxocyclohexylidene)-3-methylbutyl (ivDde)[76] protection, cleavable with 2% hydrazine in DMF. The hCD59(37-50) sequence was synthesized by SPPS using Fmoc/tBu strategy obtaining intermediate peptide **I**. After selective deprotection of Lys41 was obtained intermediate peptide **Ia** on which was performed on resin direct glycation and on resin reductive amination.

For the in-solution approaches (A and D) we synthesized the intermediate peptide **II**, a hCD59(37-50) sequence with Lysine38 and Lysine41 respectively protected with ivDde and Boc groups. The product was obtained after concentrated TFA cleavage (with all positions unprotected but Lysine38)

Direct glycation is the oldest among post-synthetic glycation strategies and is normally performed by incubating peptides and proteins in presence of a high excess (50to100 fold) of aldose sugars, glucose usually. In spite of the excess of reactants the glycation is slow and usually the reaction is held for 30 days. We have adopted a recently described method which describes glycation with glucose incubation at 110° in DMF for 30 min[77]. This strategy represents an improvement of the former method because it is faster and the yield of glycation product **III** is good (ca. 70%) but there is a serious problem with undesired byproducts. Such undesired reactions caused probably by oxidation, degradation and polymerization reactions, that gave side products almost impossible to separate from the desired product. The on-resin approach was superior in terms of yield to the in-solution one due to the protection of side chain residues.

The other post-synthetic strategy attempted was reductive amination of the Lysine residue with a specific D-ribose aldehyde derivate, the 2,3:4,5-di-*O*-isopropylidene-*aldehydo*-β-D-*arabino*-hexos-2-ulo-2,6-pyranose, which after

76 Chhabra, S.R., Hothi, B., Evans, D.J., White, P.D., Bycroft, B.W., and Chan, W.C. (1998) *Tetrahedron Lett.*, **39**, 1603.
77 Frolov, A, Singer, D., and Hoffmann, R. (2006) *J. Pept. Sci.*, **12**, 389.

the formation of the corresponding Schiff base can be reduced to the 1-deoxyfructosyl form that is indistinguishable from the derivate form of D-glucose glycation occurring in Nature[78] This strategy proved to be more successful in-solution than on resin where the rate of reaction was slow and remained a substantial amount of un-glycated form. On the contrary in solution, reductive amination was fast and proceeded almost to completion. The draw back of the latter strategy was the necessity of many purification steps which influenced the final yield.

On the other hand, step wise synthesis using our building block **2**, the specific glycated Fmoc-Lysine derivate, Nα-Fmoc-Lys[Nε-1-deoxyfructosyl, Nε-Boc)]-OH79, proved to be completely reliable with no side effects or shortcomings.

78 Brady, R.F. Jr. (1970) *Carbohydr. Res.*, **15**, 35.
79 Carganico, S., Rovero, P., Halperin, J.A., Papini, A.M., and Chorev, M. (2009) *J. Pept. Sci., **15**, 67.*

2.2.4 *Post-synthetic strategies versus building-block approach*

One important feature common to all the preparations was the realization that the glycated, 1-deoxyfructosyl-form, of a the peptide had the retention times in C18, C8 and C4 very close to the corresponding un-glycated form. Under a very shallow, almost isocratic gradient, the glycated and un-glycated forms of our antigen barely starts to separate (Figure 2.2-5).

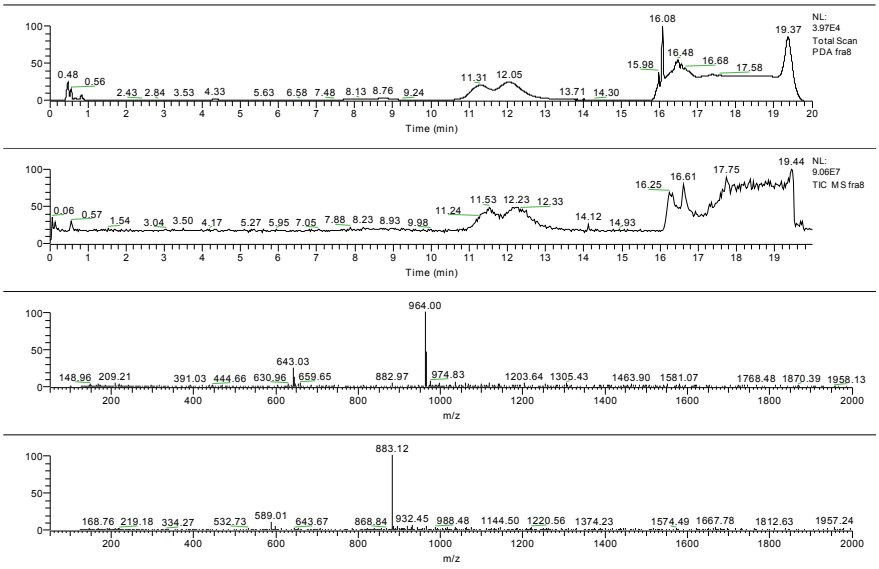


Figure 2.2-5 LC-MS method 15 to 17 % acetonitrile in water in 15 min, chromatogram of a mixture of un-glycated peptide (at 11.31 min, $[M+2H]^{2+}$=883.12) and glycated peptide (at 12.05 min, $[M+2H]^{2+}$=964.00) and glycated peptide.

Further purification by semi-preparative reverse phase HPLC using very shallow and isocratic gradients was unsatisfactory in terms of yield due to the poor separation of the two forms (glycated and un-glycated ones).

In conclusion semi-preparative C18 purification of glycated peptide preparations were found to be strongly dependent on the presence of the un-glycated form of the peptide antigen.

This purification shortcoming undermined the success of the post synthetic strategies (A), (C), and (D), due to the fact that a complete conversion of the starting material into the glycated product was not achieved regardless of the different synthetic conditions attempted. Longer reaction times, higher temperatures or higher excess of reactants lead to formation of di-glycation, that again displayed overlapping retention time with the glycated and un-glycated form suggesting some kind of compensation between the increased polar nature of a glycated form versus the increased steric hindrance. Taking into account the fact that even free Lys and glycated Lysine have very similar retention times, we believe that the introduction of a glycation modification into a peptide sequence generates no significant net change in reverse phase chromatography reaction time.

Method (B) was the only post-synthetic strategy that afforded selectively the desired glycated product however the yield is lowered by the necessity of several purification steps (after peptide synthesis, reductive amination and ivDde deprotection).

Finally, the step-wise pathway using a specific glycated building block proved to be by far the most successful pathway, affording the desired product in high yield, with no trace of the un-glycated form, requiring just one purification step.

2.2.5 *Generation and characterization of monoclonal anti glycated hCD59 antibodies.*

The glycated peptide antigen of hCD59(37-50) was used to immunize rabbits with the aim to generate specific anti-glycated-hCD59 antibodies to be used for identification of the glycated protein in biological samples.

The combined efforts of a specialized private company (Epitomics) and the Laboratory for Translationa Research susseccfully isolate rabbit spleen cells producing a monoclonal antibody that efficiently recognizes glycated hCD59 from urine (Figure 2.2-6). From the comparison of Figure 2.2-6 with Figure 2.2-7 it is possible to conclude that the headlight band corresponds to the same molecular entity and this corresponds to glycated hCD59 because the N-20 antibody only binds to the C-terminus sequence of hCD59. It can be argued that exists the possibility that the protein recognized it is un-glycated hCD59 and not the glycated hCD59. However upon treatment of the samples with $NaCNBH_3$ the monoclonal supernatants are not able any more to recognize the protein (the reducing agent transforms the glycated protein into the correspondent glucitoLysine-derivate). At the same time N-20 still recognized the protein as does a specific anti-glucitoLysine antibody that has been previously developed[80] in the Laboratory for Translational Research.

[80] Sonya Chantel et al. (2009) *JACS,* in press.

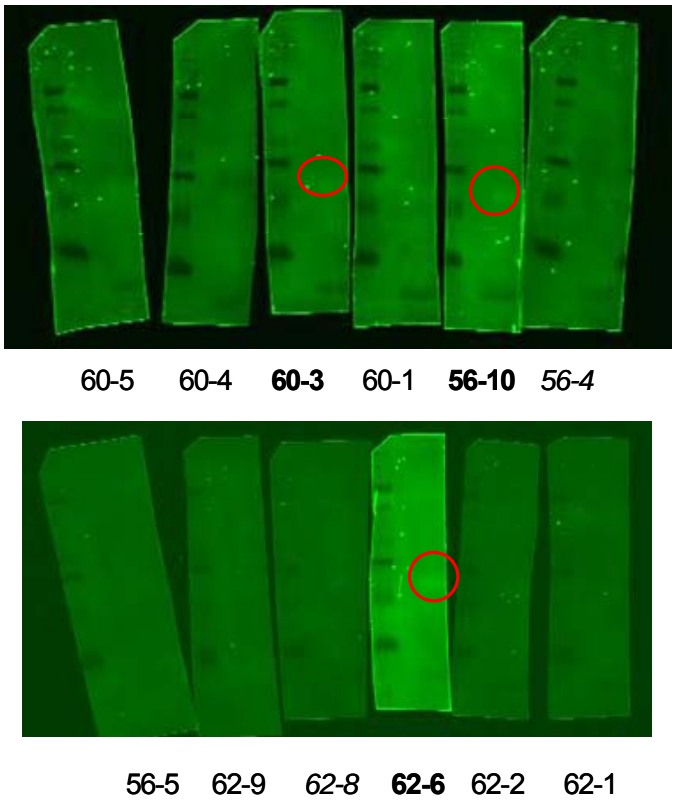


Figure 2.2-6 Western blot to detect β-ME treated human CD59 with different supernatants of rabbit clones (risen with glycated hCD59(37-50) peptide antigen). In red is headlined the glycated CD59 band. Prominent signal is seen for monoclonal 60-3, 56-10, and 62-6 antibodies.

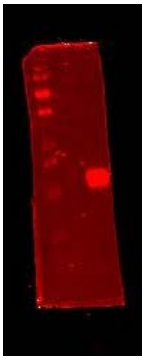


Figure 2.2-7 Blot 60-3 is reprobed with N-20 (anti-C-terminus-hCD59 commercial antibody) which only binds to β-ME treated CD59.

2.3 Synthesis of glycated antigenic probes for auto-antibodies recognition in diabetes

2.3.1 *Antigen-Antibody interactions*

The antigen (Ag) is a substance eliciting an immune response (e.g., the production of specific antibody molecules) when into the tissues of a susceptible animal. It is able to combine antibodies molecules or a receptor for T lymphocytes.[81] Antigens are simple compounds like metabolic ones (*i.e.* sugars, lipids, and hormones), as well as macromolecules (as complex carbohydrates, phospholipids, nucleic acids, and proteins).

Antigens are generally of high molecular weight and commonly are proteins or polysaccharides, despite of antibodies that are proteins and secreted by B cells after triggering an immune response.

Antibodi**s** (Abs, also known as immunoglobulins) are proteins, found in blood or other body fluids of vertebrates, and used by the immune system to identify and neutralize foreign objects, such as bacteria and viruses. Polypeptides, lipids, nucleic acids and many other molecules can be recognized as antigenic compounds by an antibody, while T lymphocytes can only recognize peptides and proteins, as well as hapten proteins and peptides.

The question of T cell recognition of glycopeptides may be important in the immune defense against microorganisms, because many microbial antigens are glycosylated. Bäcklund *et al.* [82] provide evidence that T cell recognition of protein glycans may be crucial also for T cell responses to autoantigens in the course of autoimmune diseases. Glycopeptides with simple sugars have been suitable for studies of the antigen fine specificity of glycopeptide-specific T cells.

---

81 Davies D.R., and Cohen, G.H. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 7.
82 Bäcklund, J., Carlsen, S., Hoger, T., Holm, B., Fugger, L., Kihlberg, J., Burkhardt, H. and Holmdahl, R. (2002) *Proc. Natl. Acad. Sci. USA*, **99***,* 9960.

However, immune responses may also be generated against little molecules, termed haptens, if these are chemically coupled to a larger carrier protein, such as bovine serum albumin, or other synthetic matrices. Moreover, this specific immune response is highly variable and depends also on size, structure and composition of antigens.

Only the small site on an antigen to which a complementary antibody may specifically bind is called an epitope. This is usually one to six monosaccharides or 5–8 amino acid residues on the surface of the antigen. Because antigen molecules are spatially oriented, the epitope recognized by an antibody may be dependent upon the presence of a specific three-dimensional antigenic conformation (a unique site formed by the interaction of two native protein loops or subunits). Moreover, an epitope may correspond to a simple primary sequence region. Such epitopes are described as conformational and linear, respectively. In conformational epitopes, the antibody interacts with amino acids not sequentially linked but spatially closed one to each other, because of the protein folding. The range of possible binding sites is enormous, with each potential binding site having its own structural properties derived by covalent bonds, ionic bonds and hydrophilic and hydrophobic interactions (Figure 2.3–1).

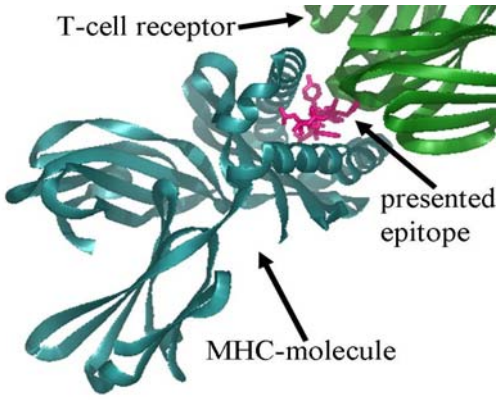


Figure 2.3-1 MHC-I bound epitope is scanned by T-cell receptor.

Proteins can undergo post-translational modifications or proteolysis altering their covalent structure and generating new epitopes. These modified regions are termed neo-antigenic epitopes and can be recognized by other specific antibodies.

The specific association of antigens and antibodies is dependent on hydrogen bonds, hydrophobic interactions, electrostatic forces, and Van der Waals forces. All antigen-antibody binding is reversible, but follows the basic thermodynamic principles of any reversible bimolecular interaction.

2.3.2 *Autoimmunity*

The immune system can occasionally attack self-antigens triggering an autoimmune response, with damages to the tissues on which these antigens are expressed. An autoimmune response can be triggered by several factors. In some cases infections are the starting triggering cause.

Molecular mimicry is the process by which a viral or bacterial infection activates an immune response cross-reacting with self-antigens. Mimicry cannot be simply explained by the structural similarity of two peptides, as elucidated by the original model of autoimmune mimicry,[83] in which a foreign antigen is sufficiently similar to a self-antigen to trigger an autoimmune response (Figure 2.3–2).

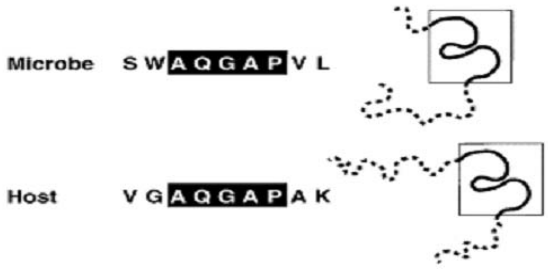


Figure 2.3–2 Molecular mimicry.

83 Oldstone, M.B. A. *Cell* **1987**, *50*, 819-820.

Thus, pathogen epitopes very similar to molecules of the organism may trigger an autoreactive reaction. By this model, autoimmunity can be explained by a response of the immune system towards self-antigens no more tolerated and thus recognized as non-self (Figure 2.3–3).

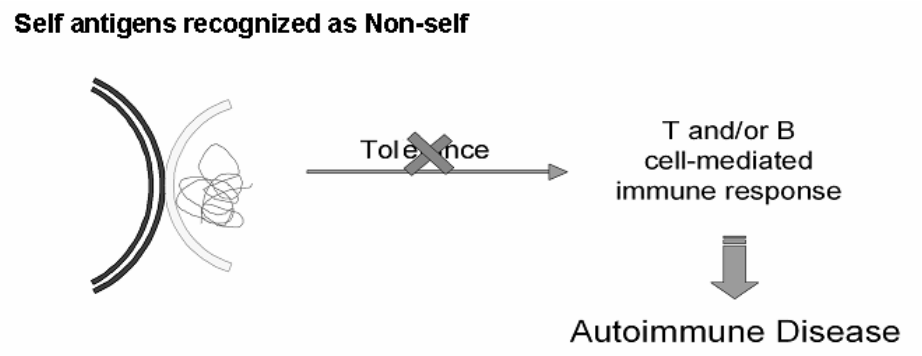


Figure 2.3–3 Autoimmune response.

Autoimmune diseases are generally triggered by T lymphocytes further inducing an autoantibody response. An autoimmune response can cause inflammation and damages to the hit tissue (Figure 2.3–4). Localization of the autoimmune response and damages depends on the disease. In some autoimmune diseases the damages are localized to single organs, while others are systemic.

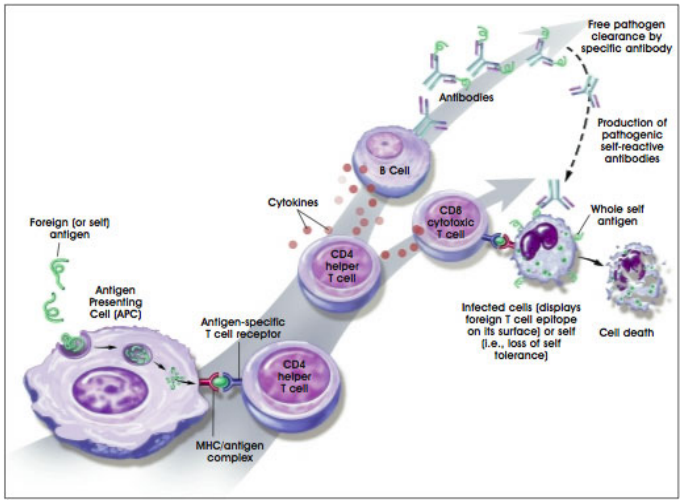


Figure 2.3–4 Mechanism of autoimmune response.

2.3.3 *Chemical Reverse Approach to characterize autoantibodies as autoimmune disease biomarkers*

Autoimmune diseases are due to an "error of evaluation" of the immune system, which no more recognizes self-antigens.[84] It is well-known that autoimmune diseases have a multifactor origin, which includes: genetic predisposition, endogenous, and exogenous elements. Different hypotheses can be made to explain autoimmune processes because of the intricate pathogeneses of such diseases and of the natural complexity of the biochemical mechanisms. Autoimmune diseases are very frequent (they affect at least 5% of the whole population) and have a high social impact, because patients have a long expectation of life during which are subordinated to the follow up of the disease by means of very expensive techniques such as Magnetic Resonance Imaging (MRI) that are not suitable for routine use.

Sera from patients suffering from autoimmune disorders often contain multiple types of autoantibodies. Some autoantibodies can be exclusive of a disease and

84 Rose, N., and Mackay, I. (2006) In *The Autoimmune Diseases*, Elsevier Academic Press, 4th edition, 55.

thus used as biomarkers for diagnosis; others fluctuate with disease exacerbations or remissions and are extremely valuable in the follow-up of patients.[85] Therefore, the antibodies present in patients' serum can be used as disease biomarkers.

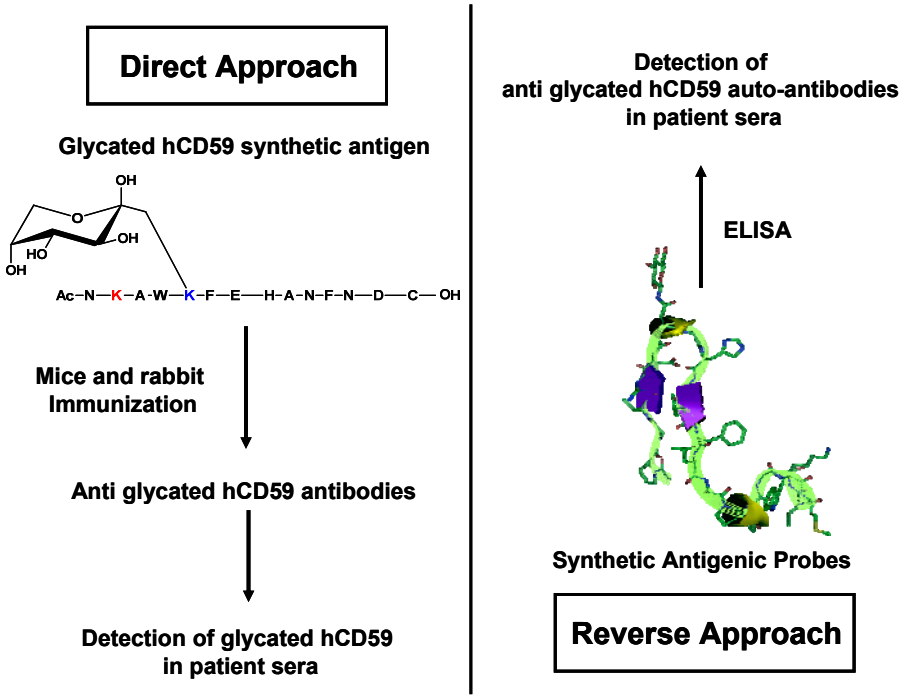
A biomarker is an anatomical, physiological, biochemical parameter, which can be easily determined and used as pointer of normal physiological or pathogenic processes. Molecules present in biological fluids can be identified as biomarkers and used to set up diagnostic/prognostic tools and for monitoring the effects of a therapeutic treatment.

The identification of synthetic peptides as antigenic probes for the characterization of autoantibodies as biomarkers can be achieved by a "Chemical Reverse Approach". This approach is defined "Reverse" because the screening of the antigenic probes is guided by autoantibodies circulating in autoimmune disease patients' biological fluids. The autoantibody recognition drives the selection and optimization of the "Chemical" structure of defined focused peptide libraries. Thus, autoantibodies circulating in patients' biological fluids allow the definition of synthetic post-translationally modified peptides mimicking the neo-antigenic epitopes. Peptides identified by this approach, selectively and specifically recognizing autoantibodies, can be used as antigenic probes in immunoenzymatic assays to detect disease biomarkers.[86]

Summarizing, in the Direct Approach antigenic peptides are used to immunize animals in order to generate antibodies that recognize those specific proteins that the antigens mimic (Scheme 2.3-1). In the Reverse Approach antigenic peptides are used as synthetic probes to recognize autoantibodies in patients' sera. In the first case the native proteins are the biomarkers of the disease, in the second the biomarkers are represented by the autoantibodies.

---

85 Leslie, D., Lipsky, P., and Notkins, A.L. (2001) *J Clin Invest,* **108***,* 1417.
86 Alcaro, M.C., Lolli, F., Migliorini, P., Chelli, M., Rovero, P., and Papini, A.M. (2007) *Chemestry Today*, **25**, 14.



Scheme 2.3-1 The Direct and Reverse Appoarch for the development of diagnostic/prognostic tools of autoimmune disease.

2.3.4 *CSF114: universal scaffold for synthetic antigenic probes*

In previous studies, performed at the Laboratory of Peptide & Protein Chemistry & Biology, it was demonstrated that the glycopeptide [Asn^{31}(Glc)hMOG(30-50)], containing a β-D-glucopyranosyl residue linked to the Asn^{31} residue by an *N*-glucosydic bond, is able to detect autoantibodies in Multiple Sclerosis patients' sera by ELISA experiments.[87]

Conformational studies performed on hMOG(30-50) and its glucosylated analogue established the fundamental role of the glucosyl moiety in autoantibody recognition (Figure 2.3–5). It was observed that the active glucosylated peptide [Asn^{31}(Glc)hMOG(30-50)] and the inactive unglucosylated peptide hMOG(30-50) adopted similar conformations in solution.[88] Therefore, it was hypothesized that the ability of the glucosylated peptide to detect autoantibodies in Multiple Sclerosis was linked to characteristics other than conformation and that the specific autoantibody binding site on MOG glycopeptide was related to the N-linked glucosyl moiety. This result, together with the observation that the N-glucosylated asparagine alone was able to bind Multiple Sclerosis autoantibodies in a solution-phase competitive ELISA experiment, allowed to define that the minimal epitope is the Asn(Glc) moiety.

87 Mazzucco, S., Mata, S., Vergelli, M., Fioresi, R., Nardi, E., Mazzanti, B., Chelli, M., Lolli, F., Ginanneschi, M., Pinto, F., Massacesi, L., and Papini, A.M. (1999) *Bioorg. Med. Chem. Lett.*, **9**, 167.

88 Carotenuto, A.; D'Ursi, A. M.; Nardi, E.; Papini, A. M., and Rovero, P. (2001) *J. Med. Chem.*, **44**, 2378.

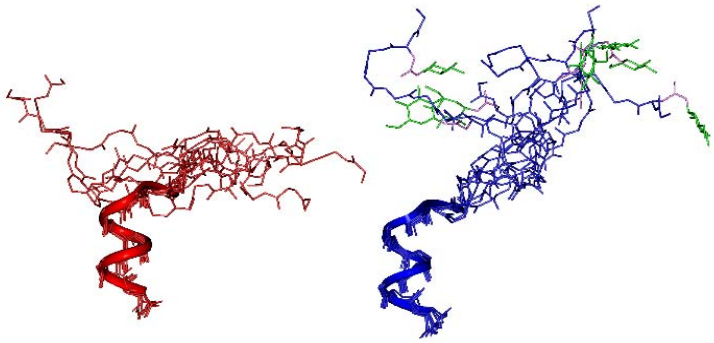


Figure 2.3–5 Conformational analysis of peptides hMOG(30-50) (in red) and [Asn^{31}(Glc)hMOG(30-50)] (in blue).[88]

Hence, the recognition properties of the molecule were optimized through the design and screening of focused libraries of glycopeptides. A specific antigenic probe, CSF114(Glc), was developed to identify a family of autoantibodies, as biomarkers of Multiple Sclerosis correlating with disease activity.

The CSF114(Glc) structure was obtained after a structure-based design confirming that antigen conformation is fundamental for the correct exposition of the minimal epitope Asn(Glc) recognizing specific antibodies. CSF114(Glc) is characterized by a β-hairpin structure in which the Asn(Glc) epitope is on the tip of type I' β-turn.[89] CSF114(Glc) showed a high specificity for Multiple Sclerosis autoantibodies, because no antibody reactivity was detected in other autoimmune diseases or other infective neurological diseases.[90] Therefore, CSF114(Glc) is the first MSAP for detecting specific autoantibodies that can be used as relabile biomarkers for the practical evaluation of the disease activity in a subpopulation of MS patients. CSF114(Glc) was selected by an

89 Carotenuto, A., D'Ursi, A.M., Mulinacci, B., Paolini, I., Lolli, F., Papini, A.M., Novellino, E., and Rovero, P. (2006) *J. Med. Chem*., **49**, 5072.

innovative “Chemical Reverse Approach” that starting from synthetic peptides univocally characterized, can screen autoantibody populations present in sera of patients (Figure 2.3–6).

It was demonstrated that specific probes characterised by β-hairpin structures are able to follow up disease activity in a statistically significant number of Multiple Sclerosis patients. It was hypothesized that CSF114(Glc) β-hairpin structures could be mimetic of aberrantly glucosylated native antigens. Moreover a specific immunoaffinity column based on CSF114(Glc) allowed isolation of disease specific autoantibodies recognising only myelin and olygodentriocytes. [91]

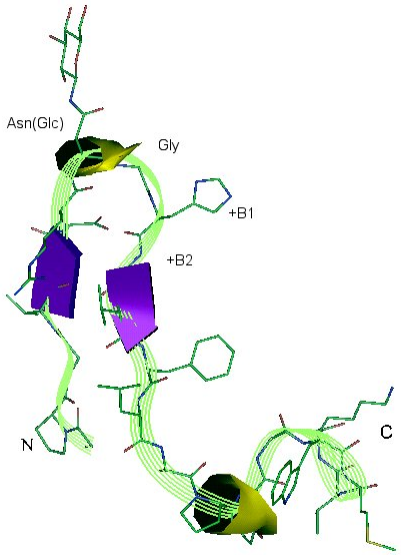


Figure 2.3–6 Calculated structures of CSF114(Glc). Ribbon diagram of the lowest energy conformer of 200 calculated structures of CSF114(Glc) derived from NMR data.

90 Lolli, F., Mazzanti, B., Pazzagli, M., Peroni, E., Alcaro, M.C., Sabatino, G., Lanzillo, R., Brescia, M., Santoro, V., Gasperini, L., Galgani, C., D’Elios, S., Zipoli, M.M.; Sotgiu, V., Pugliatti, S., Rovero, P., Chelli, M. and Papini, A.M. (2005) *J. Neuroimmunol,* **167**, 131.

91 “Glycopeptides, their preparation and use in the diagnosis or therapeutic treatment of Multiple Sclerosis”. *Inventors*: Papini, A.M., Rovero, P., Chelli, M. and Lolli, F. *Applicant*: University of Florence, Italy. PCT International application (2003) WO 03000733 A2. Italian Patent n. 0001327122 (27/04/2005).

Up to now, glucosylation, a post-translational modification not common in eucariotic proteins, has been detected only in bacterial glycoproteins.[92] Putative glucosylation of myelin proteins, by still unknown mechanisms, could transform self-antigens in non-self ones and trigger an autoimmune response. More than one protein could be aberrantly glucosylated creating neo autoantigens.

To further investigate the role of sugars and linkage between sugar and the amino acid, in autoantibody recognition, a collection based on glycoamino acids diversity is fundamental to understand this special molecular mechanism of an antibody mediated Multiple Sclerosis.

92 Wieland, F., Heitzer, R., and Schaefer, W. (1983) *Proc. Natl. Acad. Sci. U.S.A.*, **80**, 5470.

2.3.5 *Generation of a panel of Synthetic Antigenic Probes for the diagnosis of type I and type II diabetes*

Three different peptides bearing posttranslational modifications and corresponding unmodified analogs (without the PTMs, Table 2.3-1) were synthesize to be used as synthetic antigenic probes for the identification of anti-glycated-epitopes autoantibodies in type I and II diabetes patient sera.

The peptide scaffold CSF114 was selected for it is capacity to expose in best way to the solvent a residue in position 7 which is the top of a β-turn. On that specific position was introduced a glycated Lysine using the glycated Fmoc-Lysine derivate **2** obtaining peptide **IV**. On CSF114 position 7 was also introduced a residue of N-glycosylated Asparagine (using a specific developed for this purpose in PeptLab) affording peptide **V**.

Product **VII** was obtained introducing building block **2** into position 41 of hCD59(37-50).

(^{8}Lys(1-deoxyfructosyl) CSF114	TPRVERuGHSVFLAPYGWMVK	**IV**
[(^{8}Asn(Glc)] CSF114	TPRVERvGHSVFLAPYGWMVK	**V**
CSF114	TPRVERNGHSVFLAPYGWMVK	**VI**
[^{41}Lys(1-deoxyfructosyl)]hCD59(37-50)	NKAWuEHANFNDC	**VII**
hCD59(37-50)	NKAWKEHANFNDC	**VIII**

Table 2.3-1 Sequences of the synthesized antigenic probes u= glycated Lysine, v= glycosylated Asparagine.

2.4 Screening of diabetic patient sera by non-competitive ELISA

The following section describes the work originated by the collaboration with Dr. Francesco Dotta of the Diabetic Department of the University of Siena who provided the sera of type I and type II diabetic patients.

2.4.1 *Enzyme-Linked Immunosorbent Assay*

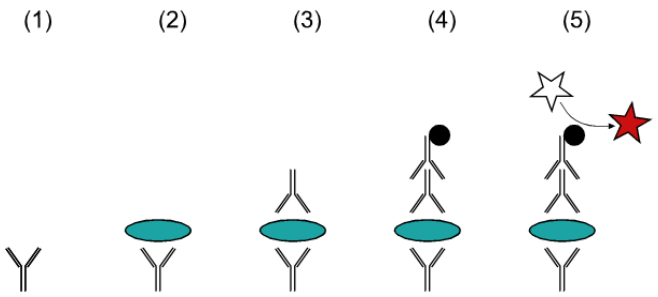
This technique can exploit the capability of antibodies to detect biomolecules with a high sensibility. Immunoassays are based on an antigen/antibody interaction to identify a target compound or a class of compounds. Concentrations of analytes are identified through the use of a sensitive colorimetric reaction. The concentration of the analyte depends on the Lambert-Beer equation and is thus determined by the intensity of color in the sample. The concentration can be determined accurately with a photometer or a spectrophotometer.

Immunoassays take advantage of the ability of antibodies to bind selectively to a target antigen present in a sample matrix and characterized by a specific chemical structure.

One of the most used immunoassay technique is the Enzyme Linked ImmunoSorbent Assay (ELISA) (Scheme 2.4-1), a technique introduced by Engvall and Perlmann,[93] and used for the detection of antigens and antibodies. The sensitivity of this technique is comparable to that of a radioimmunoassay.

The two components involved in ELISA are: a solid phase, to which a specific antigen or antibody is coated, and an enzyme-labeled anti-antibody conjugated to the corresponding antibody or antigen. The enzyme acts on an appropriate substrate, releasing a coloured compound that can be easily detected by a spectrophotometer.

93 Egall, E., and Perlmann, P. (1972) *J. Immunol.,* , **109**, 129.



Scheme 2.4-1 A sandwich *ELISA*. (1) Plate is coated with a capture antibody; (2) sample is added, and any antigen present binds to capture antibody; (3) detecting antibody is added, and binds to antigen; (4) enzyme-linked secondary antibody is added, and binds to detecting antibody; (5) substrate is added, and is converted by enzyme to detectable form.

2.4.1.1 Types of ELISA

Direct ELISA is used to detect an antigen after it has been coated to the solid phase. An antibody, conjugated with a label, is then incubated with the captured antigen. After washing off excess of the conjugate and incubating with a substrate and a chromogen, the presence of an expected color indicates a specific antibody-antigen interaction.

Indirect ELISA is useful for the detection of antibodies using specific antigens. Once again an antigen is adsorbed onto a solid phase. The first, or primary antibody, is incubated with the antigen, then the excess is washed off. The bound antibody is detected after incubation with an enzyme labeled with specific anti-immunoglobulin (secondary antibody) followed by the enzyme substrate.

Competitive ELISA is useful for identification and quantitation of either antigens or antibodies. In antigen determination, the antigen present in the sample competes for sites on the antibody with labeled antigen added to the medium. The color change will be inversely proportional to the amount of antigen in the sample. Competition principle can be exploited in different

ways. In a competitive ELISA, a patient's serum and an antigen-specific conjugate are co-incubated with a captured antigen. The amount of color developed is inversely proportional to the amount of antigen-specific patient Ig present.

Inhibition ELISA works similarly to competitive. One antigen is coated on the plate and the other is added at various concentrations in the presence of the antibody. The antibody binds to the free antigen in solution rather than to the coated antigen. In this case, the free antigen inhibits the antibody binding to the coated antigen. This is particularly useful in determining the identity of specific antigens or antibodies.

2.4.1.2 Immunological assays using CSF114 analogues as antigens

The autoantibody recognition by CSF114 analogues as antigens was evaluated by competitive ELISA on Multiple Sclerosis patients' sera. The inhibition curves (Figure 2.4-1) showed that the glycopeptides Asn7RibCSF114 does not present activity, in fact it was not able to inhibit anti-CSF114(Glc) autoantibodies in Multiple Sclerosis patients. On the other hand, glycopeptide Phe7GlcCSF114 displayed inhibitory activity only at higher concentration. CSF114(Glc) is the glycopeptide with the lower IC_{50} value (concentration required for 50% inhibition).

In conclusion, the glycopeptide containing the Asn(Glc) residue showed the higher affinity to autoantibodies in Multiple Sclerosis patients' sera. These results demonstrate again the crucial importance of the *N*-glycosydic bond between the sugar and the amino acid and the role of the sugar moiety for autoantibody recognition in Multiple Sclerosis patients' sera.

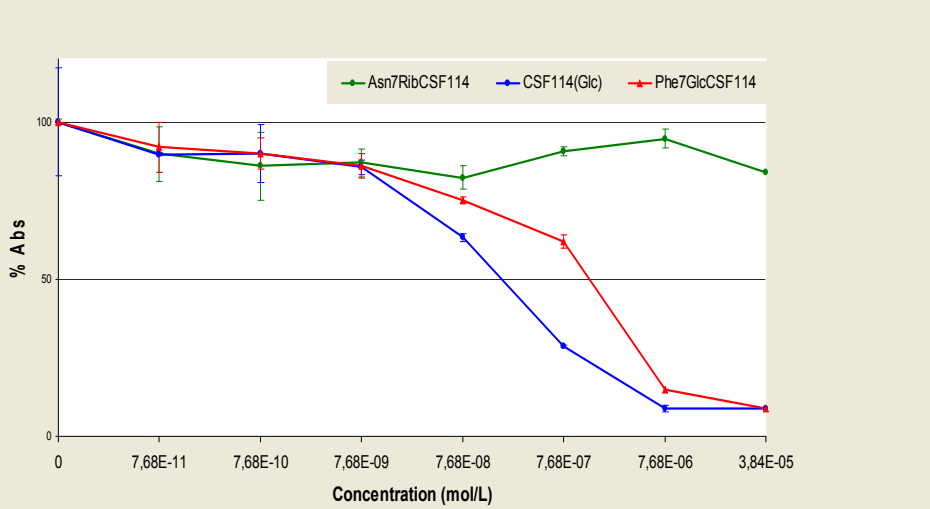


Figure 2.4-1 Inhibition curves of anti-CSF114(Glc) antibodies with the three CSF114-type glycopeptides. Results are expressed as % of a representative Multiple Sclerosis positive serum (y axis). Concentrations of peptides as inhibitors are reported on the x axis.

2.4.2 *ELISA screening of type I diabetes patients' sera*

The interaction between a given peptide and the ELISA plate depends on several factors including plate type, dilution buffer, pH, and temperature. Moreover, each peptide binds to the plate surface in a different way and this is mostly unpredictable. In order to find the best conditions for the assay, each synthetic antigen peptide was tested on a standard non-competitive ELISA versus a reduced number of healthy blood donors and patients' sera. Eight different conditions were used for each product (four different coating solutions each one applied to two different blocking solutions).

After optimization of the coating/blocking conditions the peptides **IV-VIII** (Table 2.4-1) were tested under the optimized conditions.

(^{8}Lys(1-deoxyfructosyl) CSF114	TPRVERuGHSVFLAPYGWMVK	**IV**
[(^{8}Asn(Glc)] CSF114	TPRVERvGHSVFLAPYGWMVK	**V**
CSF114	TPRVERNGHSVFLAPYGWMVK	**VI**
[^{41}Lys(1-deoxyfructosyl)]hCD59(37-50)	NKAWuEHANFNDC	**VII**
hCD59(37-50)	NKAWKEHANFNDC	**VIII**

Table 2.4-1 Sequences of the synthesized antigenic probes. u= glycated Lysine, v= glycosylated Asparagine.

To each set of values we associated a cut-off of + 2 SD (the average of the healthy blood donors' values plus two times the standard deviation of the healthy blood donors' values). The cut-off represents the minimum value of absorbance that can be taken as positive during the assay. The second SD addition minimizes the possibility to have a false positive.

Peptides **IV-VI** didn't show any recognition for IgG but displayed significant signals for IgM, in particular peptide **VI** (bearing the glucosylated Asparagine) Figure 2.4-2 A and B. The data, in spite of being preliminary (not fully

optimize testing conditions and low number of patients' and healthy donors), have been object of a standard diagnostic treatment. In fact, to each value for patient sera has been subtracted the blank sera values average and the healthy donors blank average. In addition a cut-off of ā (healthy signals average) + 2 SD (standard deviation) and not just merely ā was used. Under a less strict treatment, that does not consider the standard deviation, the recognition increases as is shown in Table 2.4-2.

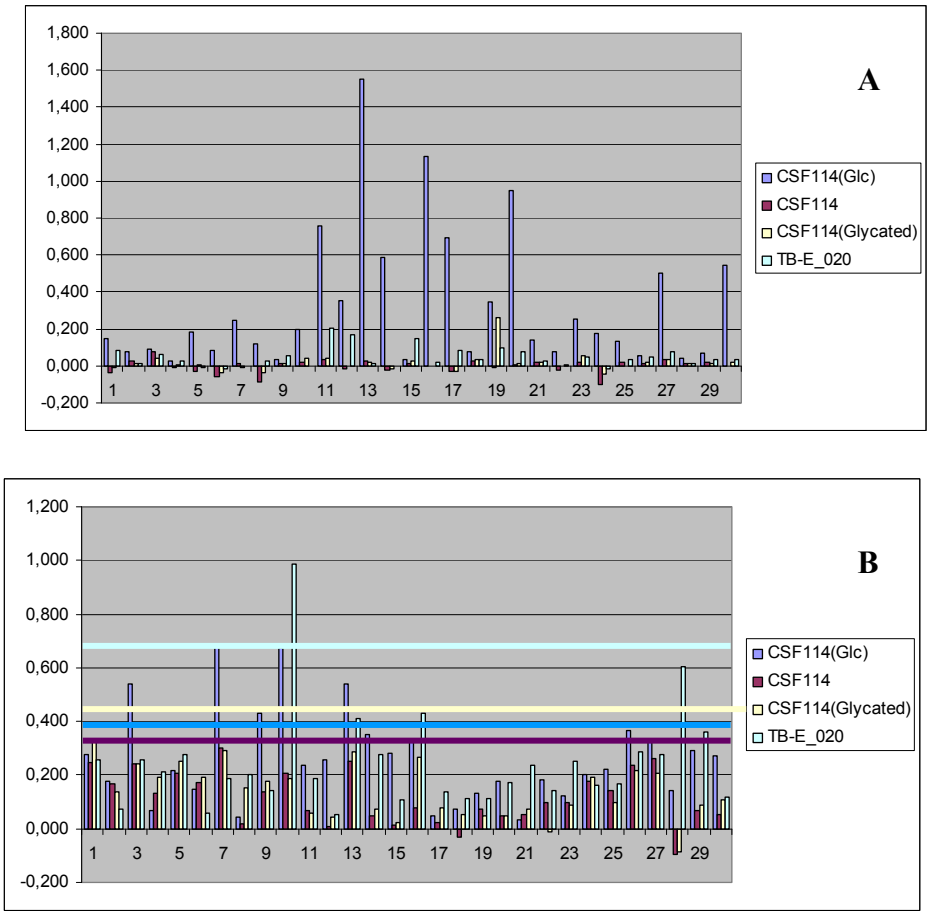


Figure 2.4-2 Non-competitive ELISA assay of peptides IV-VI against type I diabetic patients' sera and cut-off from healthy blood donors values. Results for IgG (**A**) and IgM(**B**). TB_E 20 is an unrelated peptide.

IgG	IV	V	VI	IgM	IV	V	VI
Recognition*	6	0	0	Recognition*	17	0	0
Recognition**	17	0	0	Recognition**	16	10	28

Table 2.4-2 IgM and IgG autoantibodies recognition in patients' sera (%) for peptides IV-VI with strict standard diagnostic treatment* and with preliminary data treatment**.

Glycated hCD59 **VII** displayed a high recognition for both IgG and IgM under a preliminary treatment of the data (Figure 2.4-4, Figure 2.4-4, and Table 2.4-3) and retains a significant recognition (10-15%) even upon addition of twice the standard deviation to the cut off.

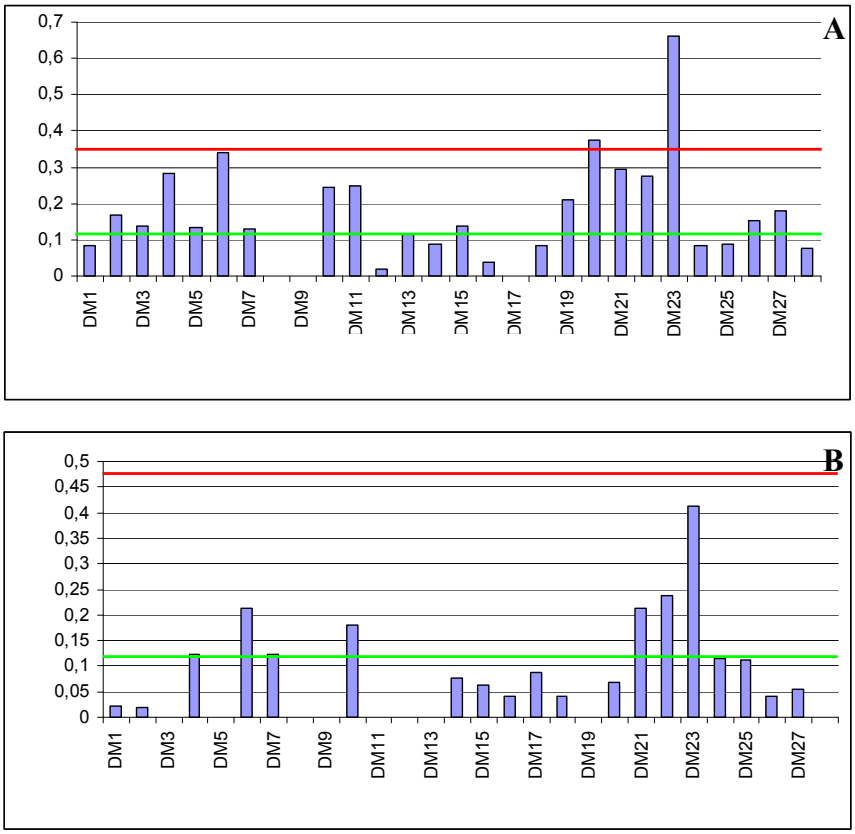


Figure 2.4-3 Non-competitive ELISA test of peptides glycated hCD59 **VII** (**A**) and un-glycated hCD59 **VIII** (**B**) against type I diabetic patients' sera and cut-off from healthy blood donors values. Results for IgG.

On the other hand, unglycated hCD59 **VIII** showed, as expected, less recognition compared to its glycated analog under both data treatments

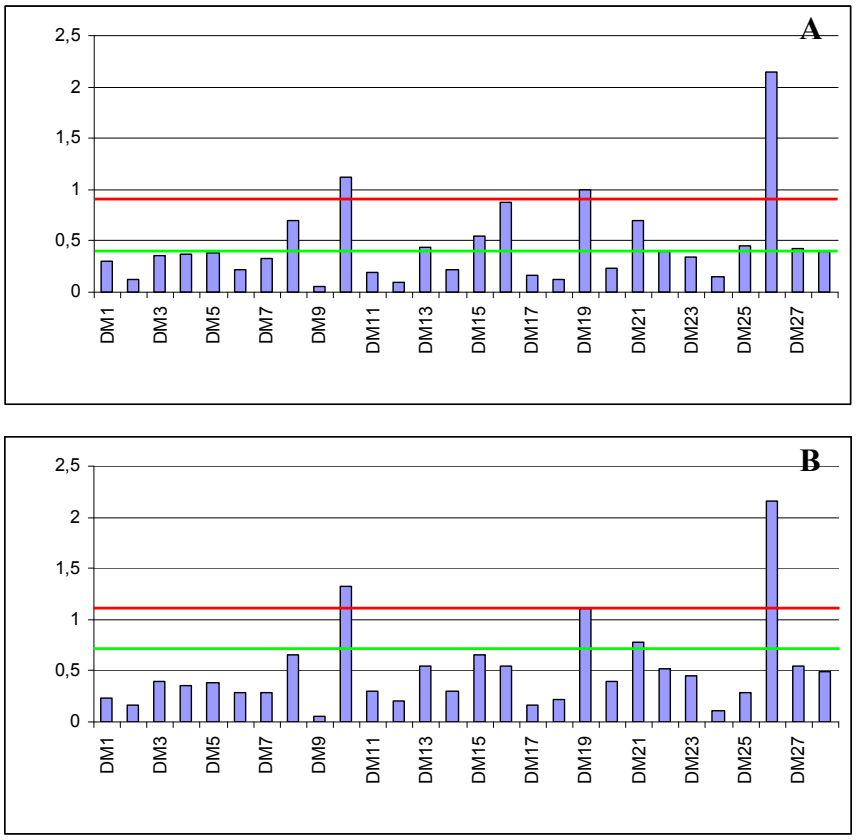


Figure 2.4-4 Non-competitive ELISA test of peptides glycated hCD59 **VII** (**A**) and un-glycated hCD59 **VIII** (**B**)against type I diabetic patients' sera and cut-off from healthy blood donors values. Results for IgM.

IgG	**VII**	**VIII**	**IgM**	**VII**	**VIII**
Recognition*	10	0	Recognition*	15	13
Recognition**	60	33	Recognition**	45	20

Table 2.4-3 IgM and IgG autoantibodies recognition in patients' sera (%) for peptides VII and VIII with strict standard diagnostic treatment* and with preliminary data treatment**.

Figure 2.4–5 and Figure 2.4–6 report the difference between the values of glycated hCD59(37-50) and of unmodified hCD59(37-50) for IgG and IgM in type I diabetic patients' sera (A) and healthy blood donors (B).

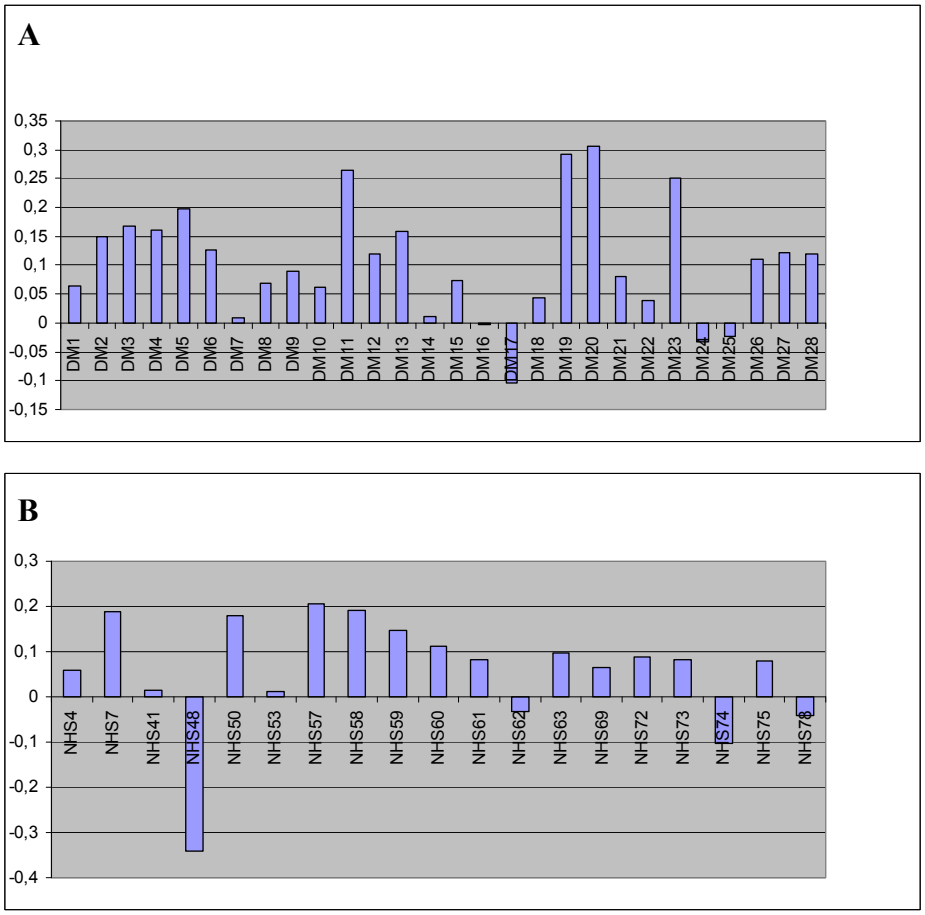


Figure 2.4–5 Difference between the values of glycated hCD59(37-50) and of unmodified hCD59(37-50) for IgG in type I diabetic patients' sera (A) and healthy blood donors (B).

As can be seen from Figure 2.4–6 A the, almost every single patients' serum has IgG antibodies recognizing the glycated peptide with higher

affinity compared to the un-glycated one (the difference is a positive value). This may be explained by the presence of specific anti-glycated antibodies in patients' sera.

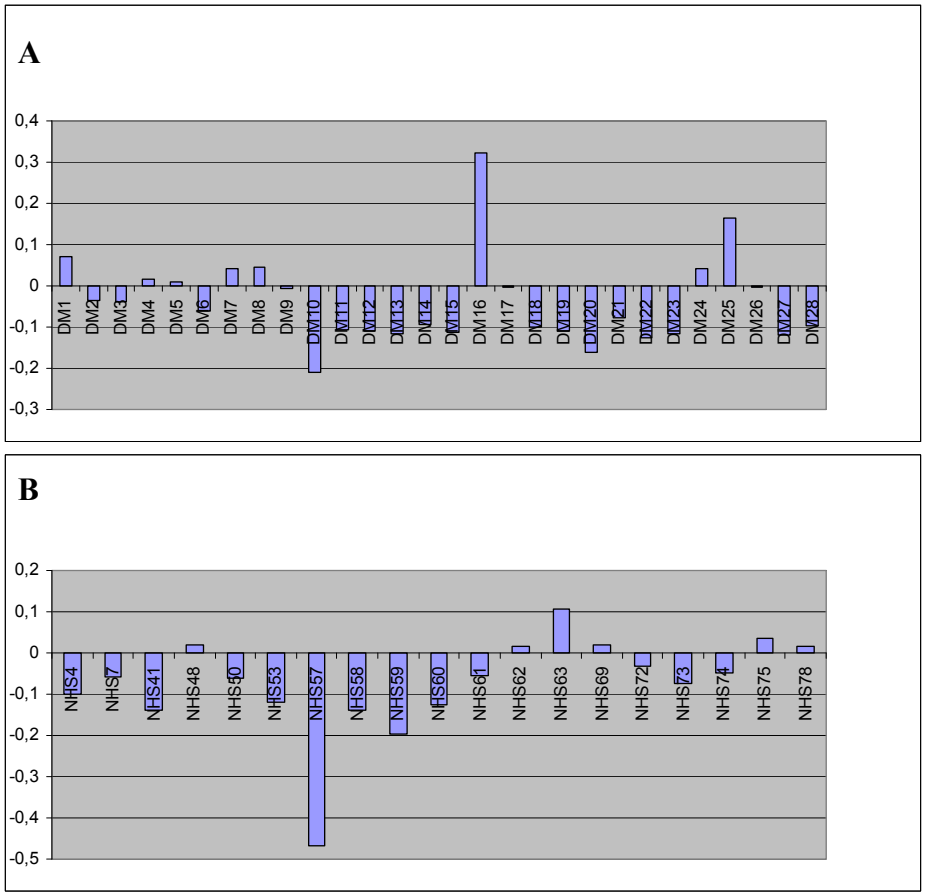
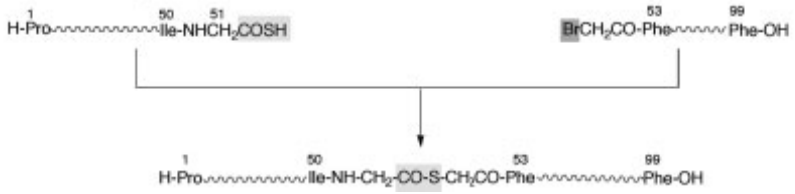


Figure 2.4–6 Difference between the values of glycated hCD59(37-50) and of unmodified hCD59(37-50) for IgM in type I diabetic patients' sera (A) and healthy blood donors (B).

2.5 Protein synthesis by Native Chemical Ligation

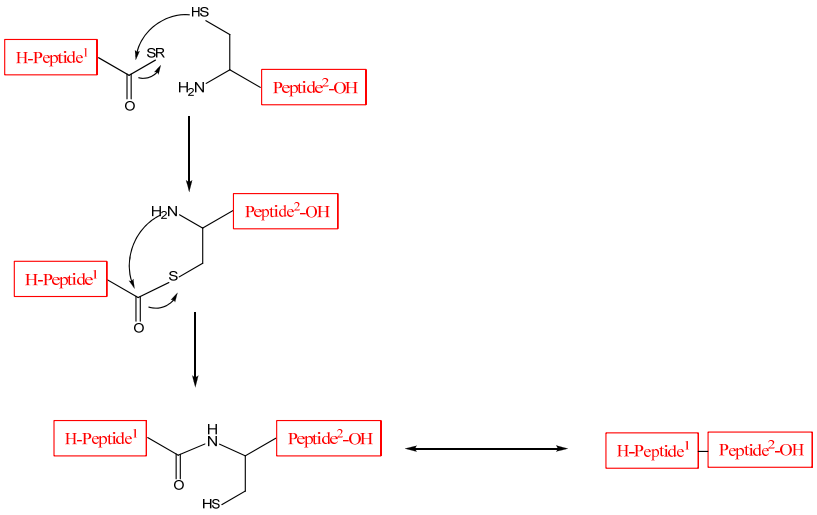
Chemical Ligations (CL) as methods to link together two or more peptide fragments have long been used for the synthesis of small proteins and have been an alternative to the strategy of fragment condensation of fully protected peptides in organic solvents[94]. The CL strategies are based on the presence in each peptide fragment of a unique, mutually reactive functionality, which enables a chemoselective reaction between the two components. The advantage is that this reaction can be performed using partially protected peptides in aqueous or semi-aqueous solvents. One example of CL is the famous synthesis of human immunodeficiency virus-1 protease (HIV-1 PR)[95] whose chemical basis was a nucleophilic substitution reaction between an SH group of a thioacid attached to the C-terminus of a peptide sequence, and an alkyl bromide attached to the N-terminus of the second fragment, leading to the formation of a thioester bond at the ligation site (Scheme 2.5-1). However the drawback of CL methodologies is the generation of an unnatural, meaning not amide, bond at the ligation point.



Scheme 2.5-1 Total synthesis through chemical ligation of the HIV-1 PR analogue.

94 Kimmerlin, T. and Seebach, D. (2005) *J. Peptide Res.*, **65**, 229.
95 Schnolzer, M. and Kent, S.B. (1992) *Science,* **256**, 221.

In recent years, there was an explosion of interest in protein assembly technologies such as native chemical ligation (NCL)[96] and its related expressed protein ligation (EPL)[97]. NCL rely on the principle of chemoselective reaction between two protein fragments, one containing a C-terminal thioester and the other containing a free N-terminal Cysteine residue The reaction is performed in aqueous solution in the presence of excess of thiols with unprotected peptides (a protection is needed only in the presence of 2 equivalents of N-terminal Cysteine residues. The components combine to give a peptide with an amide bond therefore "native" at the point of ligation differently from the bond formed in other types of ligation (which is not an amide one).



Scheme 2.5-2 Mechanism of Native Chemical Ligation.

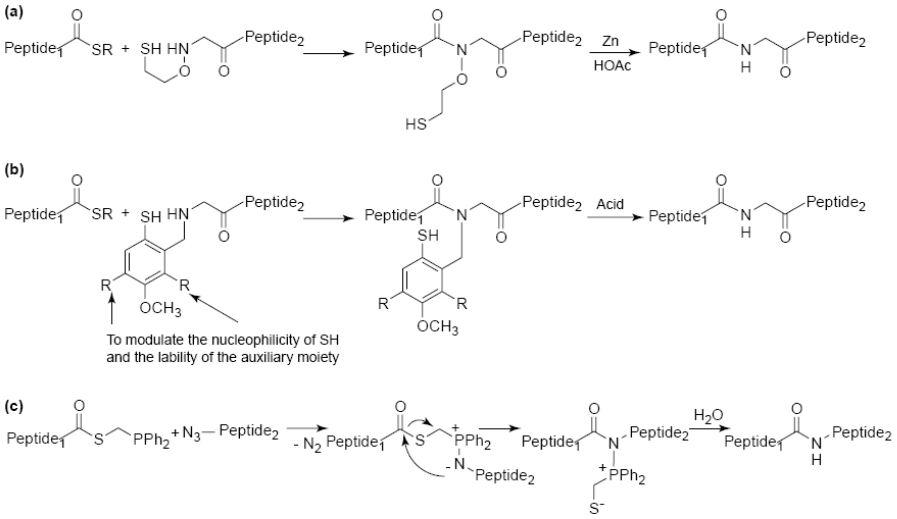
The mechanism involved is a rate-determining transthioesterification between the thiol group of the N-terminal Cysteine and the peptide-2 and the C-terminal thioester group of peptide-1 followed by a rapid S->N acyl shift, which occurs via a five-member-ring transition state (

96 Dawson, P.E., Muir, T.W., Clark-Lewis, I., and Kent, S.B. (1994) *Science*, **266**, 776.
97 T.W., Muir, D., Sondhi, and P.A. Cole, (1998) *Proc. Natl. Acad. Sci. USA*, **95**, 6705.

Scheme 2.5-2)[98]. The ability of NCL strategies to reproduce natural peptide bonds and the fact that it is possible to form it in aqueous solution and in the absence of protecting groups, has placed this powerful technology at the forefront of protein synthesis.

Nonetheless, two limitations remain associated with NCL[99]: the necessity to have an N-terminal free Cysteine and a C-terminal thioester.

Several approaches have been developed to overcome the prerequisite of having an N-terminal Cysteine residue. For example, homoCysteine can be used for NCL and then can be methylated after ligation to render the Met-containing protein[100]. Furthermore, Cys itself can be desulfurized to give an Ala residue[101]. This strategy opens the possibility of using peptides containing thio-derivatives of the common amino acids, which after chemical ligation and desulfurization can give peptides with other proteinogenic amino acids[102].



Scheme 2.5-3 Extended applicability of NCL by using N-ethanethiol (a) and N-2-sulfanylbenzyl (b) auxiliaries or by Staudinger ligation (c)

98 Macmillan, D. (2006) *Angew. Chem. Int. Ed.*, **45**, 7668.
99 Albericio, F. (2004) *Current Opinion in Chemical Biology*, **8**, 211.
100 Tam, J.P. and Yu, Q. (1998) *Biopolymers*, **46**, 319.
101 Yan, L.Z. and Dawson, P.E. (2001) *J Am Chem Soc*, **123**, 526.
102 Clive, D.L., Hisaindee, S., and Coltart, D.M. (2003) *J Org Chem*, **68**, 9247.

Furthermore, N-ethanethiol [Scheme 2.5-3(a)][103] and N-2-sulfanylbenzyl auxiliaries [Scheme 2.5-3(b)][104] have been investigated. These moieties, which contain the thiol required for the ligation, are removed after the protein is formed by treatment with Zn/AcOH and under acid conditions, respectively.

A more general strategy involves Staudinger ligation between a C-terminal phosphinothioester and an N-terminal azide, which gives an amide that does not have residual atoms [Scheme 2.5-3(c)][105]. The initial intermediate is an iminophosphorane and this rearranges to an amidophosphonium salt that after hydrolysis gives the amide.

The second and most serious limitation of NCL is related to the preparation of the thioester peptide, a stage that has mainly been performed using Boc/Bzl chemistry[106]. For larger polypeptide domains and protein domains intein-based bacterial expression systems are used[107]. Unfortunately, the Boc methodology requires the use of HF, which is extremely toxic and not well suited for synthesis of phospho-[108] and glycopeptides[109]. The Fmoc–based methodology, on the other hand, is an attractive alternative as it does not employ HF and hence provides access to the synthesis of phospho- and glycopeptides in good yield. However, the poor stability of the thioester functionality to strong nucleophiles such as piperidine, which are used for the deprotection of the N^{α}-Fmoc group, seriously limits the use of this methodology[110].

In order to overcome this limitation different approaches have been described for C-terminal peptide thioester formation with Fmoc/tBu-SPPS usually involving the use of special resins that release the peptide as a C-terminal

---

103 Canne, L.E., Bark, S.J., and Kent, S.B. (1996) *J Am Chem Soc*, **118**, 5891.
104 Vizzavona, J., Dick, F., and Vorherr, T. (2002) *Bioorg Med Chem Lett*, **12**, 1963.
105 Nilsson, B.L., Hondal, R.J., Soellner, M.B., and Raines, R.T. (2003) *J. Am. Chem. Soc.*, **125**, 5268.
106 Camarero, J.A., Cotton, G.J., Adeva, A., and Muir, T.W. (1998) *J.Pept. Res.*, **51**, 303.
107 Perler, F.B., and Adam, E. (2000) *Curr. Opin. Biotechnol.*,**10**, 377.
108 Huse, M., Holford, M.N., Kuriyan, J., and Muir, T.W. (2000) *J. Am. Chem. Soc.*, **122**, 8337.
109 Macmillan, D. and Bertozzi, C.R. (2004) *Angew. Chem. Int. Ed.Engl.,* **43**, 1355.
110 Camarero, J.A. and Mitchell, A.R. (2005) *Protein and Peptide Letters*, **12**, 723.

thioester under specific cleavage conditions such as the acylsulfonamide[111], the aryl-hydrazine[112], BAL[113], PAM[114] resins.

We have chosen a different approach which consists on the direct thioesterification of the free carboxylic acid of a fully protected peptide[115].

111 Ingenito, R., Bianchi, E., Fattori, D., and Pessi, A. (1999) *J. Am. Chem. Soc.*, **121**, 11369.
112 Camarero, J.A., de Yoreo, J.J., and Mitchell, A.R., (2004) *J. Org. Chem.*, **69**, 4145.
113 Brask, J., Albericio, F., and Jensen, K.J. (2003) *Org. Lett.*, **5**, 233.
114 Swinnen, D. and Hilvert, D. (2000) *Org. Lett.*, **2**, , 789.
115 von Eggelkraut-Gottanka, R., Klose, A., Beck-Sickingera, A.G., and Beyermann, M. (2003) *Tetrahedron Letters,* **44,** 3551.

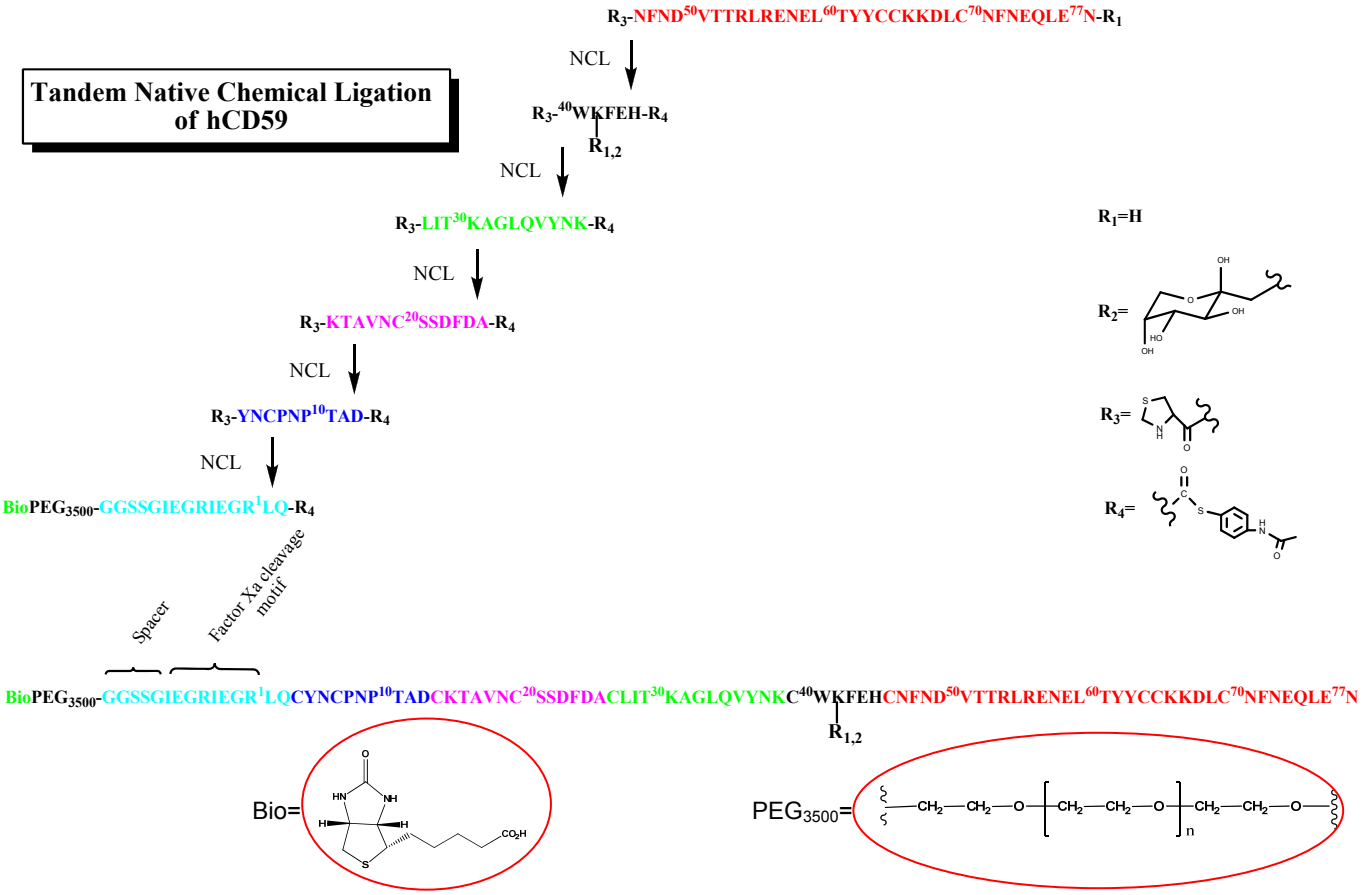
2.5.1 *Total synthesis of glycated hCD59 by Tandem Native Chemical Ligation*

Our working hypothesis is that hyperglycemia-driven glycation of the hCD59 protein destabilizes the Membrane attach complex homeostasis leading to diabetic complication in humans[50]. In order to achieve a diagnostic-prognostic tool for diabetes and diabetes complications, we have raised anti-glycated hCD59 antibodies using synthetic peptide antigens. Those antibodies are able to evidence the presence of glycated hCD59 in diabetic patient's biopsies[116]. Now we plan the generation of more specific monoclonal antibodies, to be risen from the immunization of rabbits using as an antigen the whole glycated hCD59.

Since ex vivo extraction of glycated hCD59 is highly problematic due to its low concentration we have set up a synthetic strategy for the total synthesis of the protein both wild type and bearing the posttranslational modification (glycation).

We proposed the total synthesis of hCD59(1-77) by tandem Native Chemical Ligation which allows the sequential ligation of multiple peptide fragments (Scheme 2.5-4). The sequence has been divided into six peptide units (Table 2.5-1) shaped as C-terminal p-acetamidophenol thioesters and N-terminal thioazolidine (Thz) protected Cysteine except the C-terminal fragment which has been left with a free N-terminal Cysteine and free C-terminal carboxylic acid. The Thz protecting group allows the tandem mode by masking the N-terminal Cysteine of the incoming peptide fragment during the ligation reaction.

[116] Halperin, J. and Chorev, M., unpublished results.



Scheme 2.5-4 Strategy for the totals synthesis of Bio-PEG-hCD59(1-77) wild type (^{41}K=H) and glycated (^{41}K =1-deoxyfructosyl).

and, after deprotection with methoxylamine, leaving it accessible at a later stage for the following coupling.

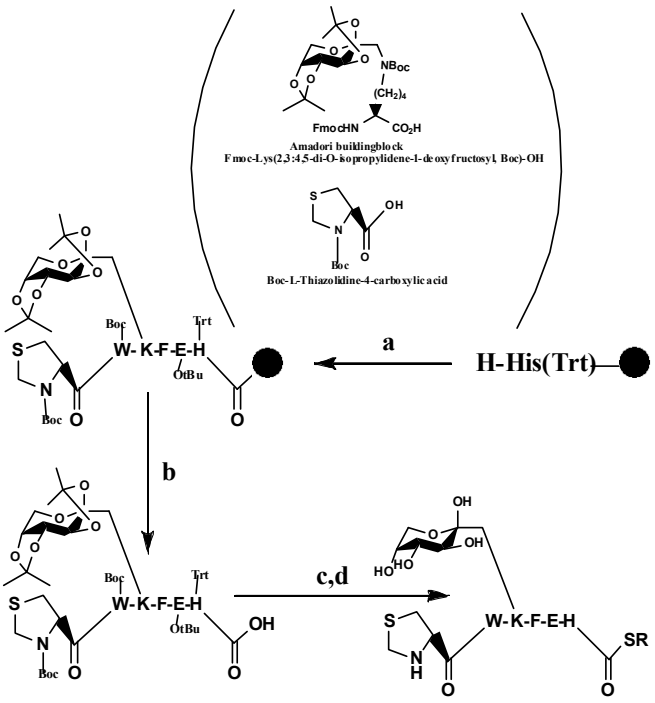
IX	H-^{45}CNFNDVTTRLRENELTYYCCKKDLCNFNEQLE^{77}N-OH
X	Thz-^{39}CWK*FE^{44}H-COSR
XI	Thz-^{39}CWKFE^{44}H-COSR
XII	Thz-^{26}CLITKAGLQVYN^{38}K-COSR
XIII	Thz-^{13}CKTAVNCSSDFD^{25}A-COSR
XIV	Thz-^{3}CYNCPNPTA^{12}D-COSR
XV	BioPEG$_{3500}$-GGSSG-IEGRIEGR^{1}LQ-COSR

Table 2.5-1 Sequence of the six ligation fragments of glycated and unglycated hCD59(1-77). K* is Lysine(1-deoxyfructosyl); Bio is Byotine; PEG_{3500} is a Polyetylene-glycol chain of ca. 3500 of MW; GGSSG is a spacer; IEGRIEGR is the factor X_a cleavage motif.

The direct thioesterification of the C-terminal carboxylic acid with p-acetamidophenol has two advantages with respect to other strategies to obtain a C-terminal thioester. First of all it is relatively stable compared to benzyl and phenyl thioesters commonly used in NCL so that thioester group can survive several weeks if the peptide is maintained lyophilized at -4°C. Moreover the synthesis of the peptide thioester is straightforward, involving a thioesterification with p-acetamidothiophenol[115] of the peptide fully protected but in the C-terminal carboxylic position as shown in Scheme 2.5-5 (the procedure applies for all the peptide fragments except the C-terminal one). The peptide is synthesized on a super-acido labile 2Cl-trityl resin. Cleavage of the peptide from the resin occurs with dilute acetic acid. Under this conditions all the side chain protection are left in place. No special resins, auxiliaries or cleavage protocols are needed and, in addition, p-acetamidothiophenol is a solid were the common thiols use for similar purposes are volatile liquids toxic and with a very bad smell.

The proposed synthesis of hCD59 incorporates at the N-terminal portion a Biotin-polietylene-glycol residue in order to give to the final adduct additional

biochemical properties and a better aqueous solubility. Finally, the Bio-PEG portion will be separated from the protein sequence by a spacer and a protease X_a cleavage motif that would allow selective cleavage of the unnatural tail to obtain the free protein sequence for structural characterization.



Scheme 2.5-5 Synthesis of the second ligation fragment bearing the glycation(a) Fmoc SPPS on 2-Cl-Trt resin; (b) acetic acid/TFE/DCM (1:1:8); (c) 3eq p-acetamidophenol, 3q PyBOP, DIEA; (d) TFA/H2O/TIS (95:2.5:2.5) 3h.

Preliminary attempts demonstrated that the ligation is highly favored and proceeds with excellent yields up to the second coupling stage. The week points of the "Tandem" synthetic strategy are the semi-preparative HPLC purifications cycles needed after each ligation and Thz deprotection steps. To othercome this difficulty we are undertaking molecular size-mediated centrifuge-filtration of the products on the assumption that the big ligate fragment will be retained on filter with a specific pore size while the much

lighter incoming peptide fragments will be filtrated together with the small molecules as thiol catalysts, methoxylamine, detergents, salts etc.

2.6 Monitoring glycation by amino acid analysis

Recent improvements in HPLC and mass spectroscopy technologies allow to detected metabolites in biological fluids even at extremely low concentrations. This can be applied to the diagnosis-prognosis of various despises provided that a solid correlation is established between a specific biomarker and the pathology.

We have decided to apply the concept of glycation to the analysis of metabolites in the urine and sera of type I and type II diabetic patients.

2.6.1 *Amino acids and short peptide sequences in biological fluids*

The presence of free amino acids in human sera has been identified since 1912[117]. The concentration of a given amino acids is the result of a balance between its consumption during protein synthesis, its uptake from the diet and its formation through endogenous protein degradation Some common values of free amino acid concentrations in human sera are reported in Table 2.6-1. It must be remembered, however, that those concentration change importantly depending on a number of physiological conditions such as age, sex, lifestyle etc. Interestingly, they vary according to pathological conditions also. The latter consideration can be used to set up diagnostic prognostic assays working on the assumption that a given disease may generate a specific pattern of free amino acids both natural and post-translationally modified in the sera. In our diabetic study we assume the existence of free glycated amino acids, especially Lysine residues.

[117] Stein, A. (1954), *J.of Biological Chemistry*, **7**, 24

Amino acids	Serum concentration mg/mL
Aspartic acid	0.21
Threonine	1.73
Serine	1.27
Proline	2.05
Glycine	1.65
Alanine	2.42
Valine	2.12
Cysteine	1.05
Methionine	0.24
Isoleucine	0.58
Leucine	1.19
Tyrosine	0.83
Phenylalanine	0.78
Ornitine	0.61
Lysine	1.51
Histidine	1.24
Arginine	1.49

Table 2.6-1 Common values of free amino acid concentration in human serum.

An excessive rise or decrease of the plasma-a.a. concentration often leeds to pathological conditions as is the case of a prolonged protein-poor diet which brings insufficient plasma amino acid content and problems in children developing. The regulatory activity of a.a. has been demonstrated for several cases of gene expression of specific enzymes[118] The details of such mechanisms are however unknown. Off-balanced amino acids homeostasis is held responsible of several pathological conditions Table 2.6-2

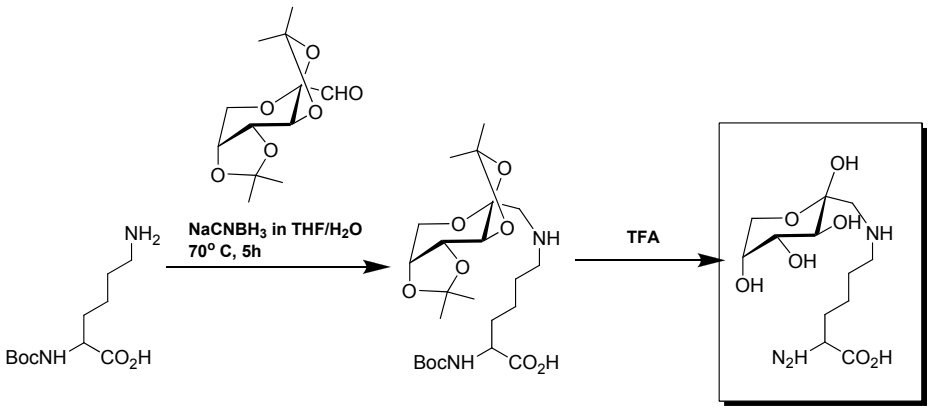
118 Roger, W.J. (1998) In *Biochemical Individuality The basis for genetotrophic concept*, Keats Publishing, New Canaan, Connecticut, , 40.

Disease	Incidence	Therapy
PKU (phenylketonuria)	1:10000	diet
MSUD (maple syrup urine disease)	1:100000	Diet
OmoCysteineuria	1:150000	Vitamin B6 Betain
Ipertyrosinemia	1:1000000	Diet
Citrulinemia	1:100000	glucose,sodium benzoate, phenylacetate
Ipoarginasis	1:100000	Diet

Table 2.6-2 Disease cause by off-balanced amino acids homeostasis.

2.6.2 *Analysis of free glycated amino acids in diabetic patients' sera*

We planned to investigate wether free glycated Lysine could be present in type I and type II diabetic patients' sera and if its presence could be used to set up a diagnostic/prognostic tool for diabetes. At the same time the presence of glycated Lysine could detected among the free amino acids generated from overall hydrolysis of the proteins of diabetic patients' sera, provided that the glycation moiety survives the hydrolysis conditions (HCl 6M at 110°) which is something that should be investigated. In both cases we planned to use as an HPLC analytical standard the free glycated form of Lysine (Scheme 2.6-1) obtained by TFA deprotection of N^{α}-Boc-LysN^{ε}-(2,3;4,5-di-O-isopropylidene-1-deoxyfructosyl)-OH.



Scheme 2.6-1 Synthesis of the free form of glycated Lysine.

2.7 Conclusions and future developments A

Final goal of this project is the development of a diagnostic/prognostic tool for type I and type II diabetes. A convergent approach of multiple different strategies proceeding in parallel has been adopted (section 1).

To this aim, a panel of new Fmoc-Lysine derivates bearing a glycation modification has been developed for the SPPS of glycated peptides and proteins (section 1.1).

A systematic approach to the synthesis of a glycated hCD59(37-50) peptide antigen has been carried out (section 2.2.3).

The glycated antigen has been used to produce specific anti-glycated hCD59 antibodies that efficiently recognize the glycated protein (hCD59) *in vivo* (section 2.2.5).

A panel of glycated antigenic peptide probes has been generated and tested with un-competitive ELISA experiments against type I diabetic patients' sera (section 2.3.5). Preliminary results showed the presence of specific autoantibodies anti-glycated-hCD59 in a subfamily of diabetic patients (section 2.4.2).

The strategy for total synthesis by Native Chemical Ligation of glycated and un-glycated hCD59 protein has been set up (section 2.5.1).

Future developments will be the optimization of the biochemical assays based on the synthetic peptide probes and antibodies generated as well as the completion of the NCL synthesis of hCD59. It will also be undertaken an analytical study of the glycated metabolites of diabetic patients' sera (section 2.6.2); and finally, the synthetic peptide probes will be tested in uncompetitive ELISA assays for type II diabetic patients' sera.

3 PART B: Cyclic clicked peptides as anticancer drugs

The following section describes the synthesis and properties of clicked cyclopeptides derived from the binding site of eIF4E binding protein. Such peptides are proposed as new anticancer drugs in the context of inhibition of translation initiation. The introduction of a triazole-bridge in a peptide sequence by click chemistry is proposed as a way to stabilize the secondary structure of the sequence and hence enhancing the affinity in peptide-protein interactions. At the same time, the presence of the bridge may improve the peptide bio-stability and favoring intracellular drug-delivery.

3.1 Stabilization of peptides in α-helical conformation

Examination of complexes of proteins with other biomolecules reveals that proteins tend to interact with partners via folded subdomains, in which the backbone possesses secondary structure. α-Helices constitute the largest class of protein secondary structures, and play a major role in mediating protein–protein interactions[119]. Significantly, the average length of helical domains in proteins is rather small and spans two to three helical turns (or eight to twelve residues)[120]. These complexes suggest that it may be possible to develop short helices that potentially participate in selective interactions with biomolecules. However, peptides rarely retain their conformation once excised from the protein. Much of their ability to specifically bind their intended targets is lost because they adopt an ensemble of shapes rather than the biologically relevant one. This happens because a short peptide in solution does not give rise to a

119 Guharoy, M., and Chakrabarti, P. (2007) *Bioinformatics*, **23**, 1909.
120 Barlow, D.J. and Thornton, J.M. (1988) *J Mol Biol*, **201**, 601.

number of weak interactions (hydrogen bond and van der Waals interactions) big enough to stabilize a specific conformation. The proteolytic instability of peptides is an additional factor that limits their utility as reagents in molecular biology and drug discovery. Finally, peptides have a low tendency to penetrate biological membranes due to their mostly hydrophylic nature that is in contrast with the highly hydrophobic nature of the central core of the phospholipidic bilayer. Figure 3.1–1 illustrates the different approaches[121] that have been adopted either to stabilize or mimic an α-helix, with the overall aim of endowing peptidic and non-peptidic oligomers with conformational rigidity, proteolytic stability, and the desired array of protein-like functionality. These approaches can be divided into three general categories: helix stabilization, helical foldamers, and helical surface mimetics.

Helix stabilizing methods based on side chain crosslinks and hydrogen-bond surrogates preorganize amino acid residues and initiate helix formation; mini-proteins that display helical domains are also part of this category. Helical foldamers, such as β-peptides and peptoids, are composed of amino acid analogs and are capable of adopting conformations similar to those found in natural proteins. Helical surface mimetics utilize conformationally restricted scaffolds with attached functional groups that resemble the i, i + 4, i + 7 pattern of side chain positioning along the face of an α-helix.

Stabilization of peptides in the helical structure not only reduces their conformational heterogeneity but also substantially increases their resistance to proteases. In fact these enzymes typically bind their substrates in the extended conformation. In addition it has recently been reported that peptides with a stabilized α-helical conformation display a high membrane permeability.

---

121 Henchey, L.K., Jochim, A.L., and Arora, P.S. (2008)*Current Opinion in Chemical Biology*, **12**, 692.

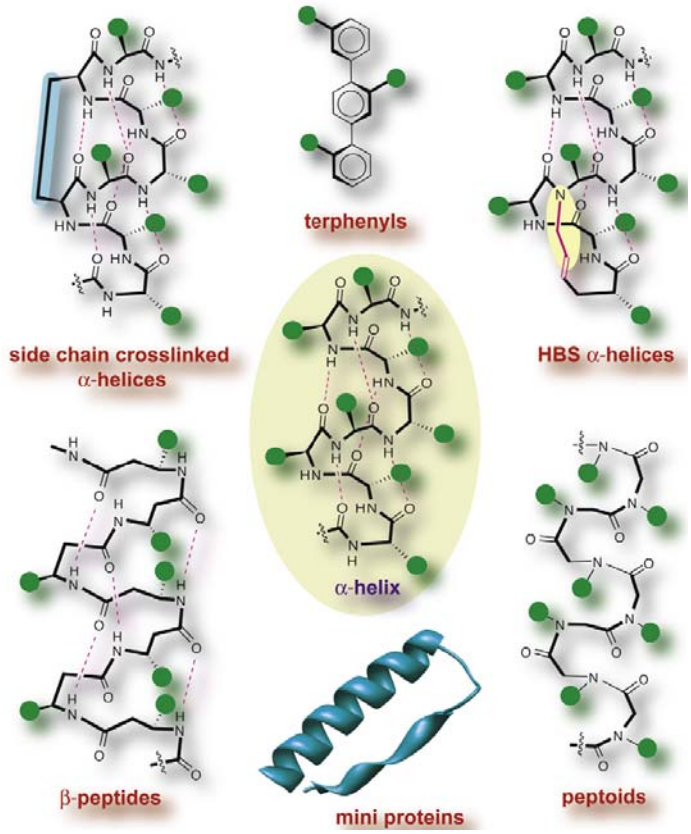


Figure 3.1–1 Stabilized helices and non-natural helix mimetics: Several strategies to stabilize the α-helical conformation in peptides or mimic this domain with non-natural scaffolds have been described. Recent advances include β-peptide helices, terphenyl helix-mimetics, mini-proteins, peptoid helices, side chain crosslinked α-helices, and the hydrogen bond surrogate (HBS) derived α-helices. Green circles represent amino acid side chain functionality.

In fact, Verdine and coworkers[122] have applied the olephin methatesis reaction for the formation of peptide side-chain to side-chain hydrocarbon bridge, Figure 3.1–2 (A). The bridge constraint or "staple" stabilizes the α-helical conformation of a peptide that efficiently promotes apoptosis of cancer cells by triggering intracellular receptors. As shown in Figure 3.1–2 (B) and (C) such peptides almost completely enter the cells.

122 Kim, Y.W.and. Verdine G.L (2009)*Bioorganic and Medicinal Chemistry Letters,* **19**, 2533.

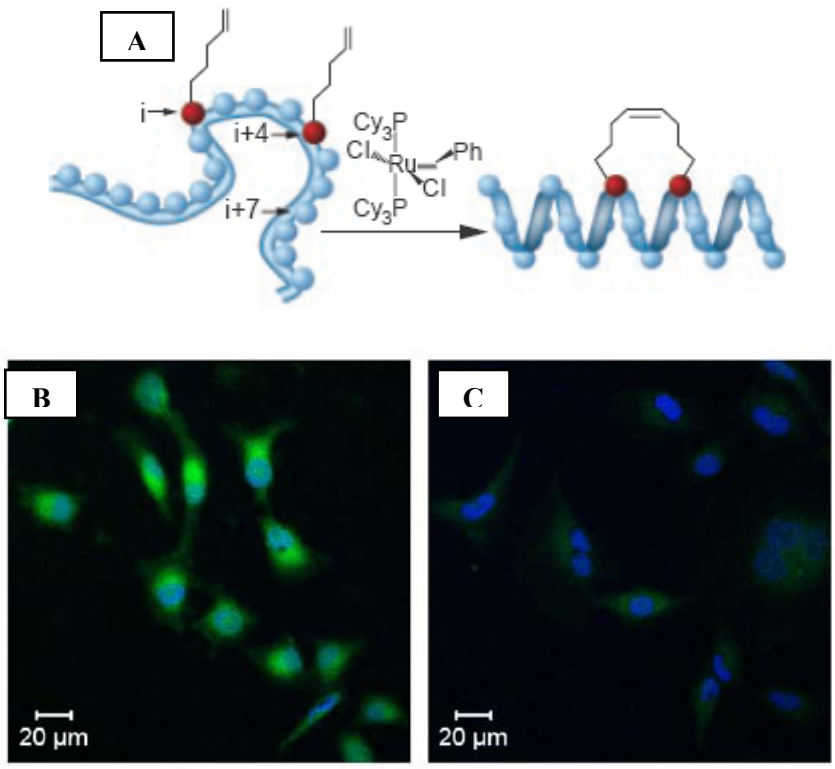


Figure 3.1–2 (A) Stapled peptides with side chain to side chain hydrocarbon bridge (C) and (D) Confocal microscope images of HeLa cells treated with 10 µM of fluoresceinated peptides S,S-SAHBa (C) and R,R-SAHBa (D).

Moving from the considerations described above, we have focused our efforts on the synthesis of peptides with an α-helical conformation stabilized by a side-chain to side-chain triazole bridge generated by an intermolecular click reaction.

3.1.1 *Tridimensional active conformation*

In nature, protein functions including enzyme activities are often regulated through conformational change triggered by ligand binding or post-translational modification at specific sites.[123] For example systems that control protein–protein and peptide–protein interactions, have been designed by introducing mutations into the protein at the interface of the interaction to reduce the binding affinity and by addition of small compounds to restore binding conformation.[124]

The conformational state of receptor protein determines the functional state of a receptor. Peptides are flexible molecules, which tend to adopt a large number of conformations in solution but assume well-defined conformation only when bound to their receptors. Peptides should have active conformation that requested receptor to trigger the biological response. The peptide conformation generated by a given amino acid sequence is fundamental to express its biological activity. Peptides or proteins that do not have the correct conformation do not display their biological role.

Therefore, the binding alters the chemical conformations, i.e. the three dimensional structure. The receptor bound conformation may be poorly populated in solution but this structure may be promoted by incorporating conformational constraints into the peptide. If the conformation stabilized by the constraint closely resembles the structure responsible for bioactivity, this modification can increase potency and selectivity of the resulting peptide. Some flexibility should be retained in the constrained molecule so that the side-chain pharmacophoric groups may adopt orientations analogous to those in the bioactive conformation of the native peptide. The conformation of a peptide can be stabilized by introduction of bridges of various sizes between

123 Pearl, L.H., and Barford, D. (2002) *Curr. Opin. Struct. Biol*., **12**, 761.
124 Guo, Z., Zhou, D., and Schultz, P.G. (2000) *Science*, **288**, 2042.

different parts of the molecule. The bridge can either be local and occurring within a single amino acid residue or being global and linking distant parts of the sequence. In this context most combinatorial libraries have been developed in order to modify R-groups appended to a common core scaffold where R-groups are typically chosen to maximize diversity, but their relative spatial orientations are usually fixed by the geometric constraints of the scaffold. Cyclic peptides offer the possibility of conveniently varying both scaffold geometry and R-group functionality. For example, parameters such as ring size can have a dramatic effect on cyclopeptides conformations, allowing access to structurally diverse species based on simple modifications in their linear sequences. Cyclization affects the degrees of freedom of all residues within the ring and thus a macrolyte should adopt a more defined conformation than the equivalent linear sequence. In fact, the active conformation in cyclic peptides can give superpotent analogues in matched cases.[125] In addition conformational constraints provide the basis for receptor selectivity; often different receptors bind the same flexible substrate in different conformations (Figure 3.1–3).

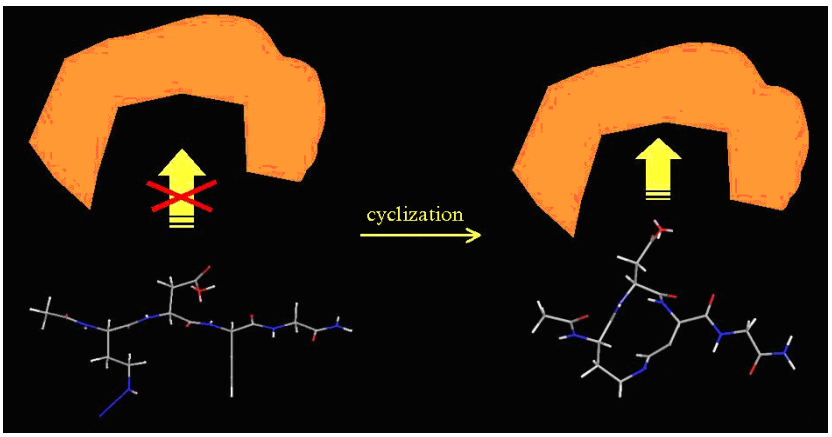


Figure 3.1–3 Differences concerning the interactions ligand-receptor between a linear peptide and cyclic peptide.

125 Kessler H. (1982)*Angew. Chem. Int. Ed. Engl.*, **21**, 512.

It was shown that conformational constraints of peptide structures can induce conformations in an often predictable way.[126] If the bioactive conformation is unknown, the spatial orientation of pharmacophoric groups on a distinct backbone conformation can be systematically screened. The procedure involves a shift of one (or more) D-amino acid around a distinct scaffold. The functional groups of the side-chains and their neighborhoods are retained but their spatial arrangement can be adjusted. If one of these conformations matches the bound conformation (e.g. the conformation of its biological receptor) superactivity can be expected. In addition, the constraints often prevent binding to related receptor subtypes resulting in higher selectivity. This procedure is applied to design a potent lead structure for an anticancer drug.

A peptide mimetic thus embodies the conformational and molecular characteristics thought to be important for biological activity of the native sequence. Mimetics may exhibit enhanced potency and can be more selective for various receptor sub-types than their parent sequence but several generations of variants may need to be prepared before a drug candidate emerges.

126 Matter, H., and Kessler H. **(**1995) *J. Am. Chem. Soc.*, **117**, 3347.

3.1.2 *Types of cyclopeptides*

Characterization of possible relationship between peptide structure and biological activity are often aided by reduction in diversity of the possible conformational states that a linear peptide can adopt by generating cyclic analogues.[127] Intramolecular side-chain-to-side-chain cyclization is an established approach to achieve stabilization of specific conformations, and has been employed to achieve rigidification that results in restricting the conformational freedom. Biologically active cyclic peptides designed and prepared specifically from a linear cyclic peptide approach, have been used to possess several additional attributes including: (i) increased agonist or antagonist potency; (ii) prolonged biological activity and extended pharmacokinetics; (iii) increased stability to enzymatic degradation; and (iv) increased specificity for a particular receptor.[128]

Cyclic peptides are polypeptide chains whose amino and carboxyl *termini*, or two amino acids side chains, are themselves linked together by an amide (CO-NH), disulfide (S-S), carbon-carbon (CH_2-CH_2 or CH=CH), reduced amide (CH_2-NH), methylene thioester (CH_2-S), methylene sulfoxide (CH_2-SO), methylene ether (CH_2-O), thioamide (CS-NH), keto methylene (CO-CH_2), aza (NH-NR-CO) bond or recently 1,4-[1,2,3]triazolyl bridge (Figure 3.1–4).

127 Felix, A.M., Wang, C.T., Heimer, E.P., and Fournier, A. (1988) *Int. J.Peptide Protein Res*., **31**, 231.
128 Hruby, V.J. (1982) *Life Sci*. **31**: 189.

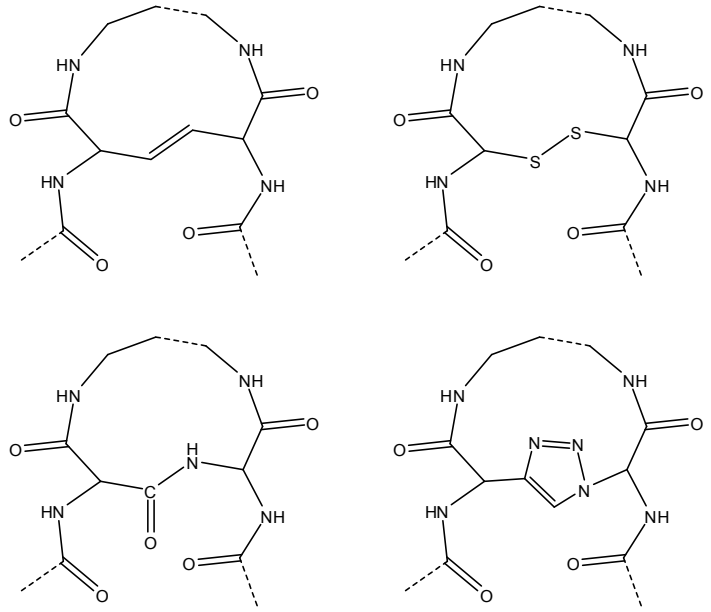


Figure 3.1–4 Common types of cyclopeptides.

In a number of studies, side-chain-to-side-chain cyclization has been performed by lactam bridge formation between *N*- and *C-termini*, side-chain and the *N*- or *C-terminus*, or two side-chains, to generate highly helical and biologically active peptides.[129, 130] Moreover the introduction of the amide isosteres also results in local and global changes dipole moments and in the pattern of intramolecular and peptide-receptor hydrogen-bond formation. Thus, incorporation of amide bond isosteres cannot only improve *in vivo* stability as the mimetic is no longer a substrate for peptidases, but can improve selectivity towards receptor sub-types, changing pharmacological functions and enhancing pharmacokinetic properties.

On the other hand cyclization in the form of a disulfide bridge between two Cysteines, or other thiol-containing amino acids, is the most abundant post-translational modification resulting in side-chain-to-side-chain cyclization. However, under certain redox potentials, the disulfide bridge will behave as a

129 Kapurniotu, A. and Taylor, J.W. (1995) *J. Med. Chem.*, **38**, 836.
130 Chorev, M., Roubini, E., McKee, R.L., Gibbons, S.W., Goldman, M.E., Caufield, M.P., and Rosenblatt, M. (1991) *Biochemistry*, **30**, 5968.

relatively transient modification yielding either the reduced linear form or generate a variety of intermolecular disulfide containing products.

On the other hand, ruthenium complexes have been applied to the ring closing metathesis (RCM) reactions of a number of dienic substrates. The substrate scope includes rings of 6 to 20 members. In addressing macrocyclic peptides, a class of tetrapeptide disulfides inspired the synthesis of the carbon-carbon bond analogs. For example, replacement of Cysteine residues by allylglycines resulted in the acyclic precursors which were subjected to RCM to afford the corresponding macrocycles. In addition, several macrocycles, which were not based upon disulfide bridge-containing species found in nature , were prepared.[131] Moreover the carba replacement (CH_2-CH_2) is non-polar and does not allow the possibility of intramolecular or peptide-receptor hydrogen bonding, while the reduced amide (CH_2-NH) unit is conformationally different from the amide bond because it does not have any double-bond character.

All of these types of cyclization require orthogonal protection of side chains to afford a peptide cyclization. While side-chain to side-chain cyclization of peptide sequences has been successful in many instances, a number of factors are known to significantly influence the efficiency of the cyclization reaction and the yield of the desired cyclic peptide product.

Therefore cyclizations, which do not require complicated orthogonal protection schemes, are of great interest. The recently Cu(I)-catalyzed azide–alkyne 1,3-dipolar Huisgen's cycloaddition [132, 133, 134] as a prototypic "Click chemistry reaction"[135] presents a promising opportunity to develop a new paradigm for intramolecular side-chain-to-side-chain cyclization in peptides.

131 Miller, S.J., Blackwell, H.E., Grubbs, R.H. (1996) *J. Am. Chem. Soc*., **118**, 9606.
132 Huisgen, R. (1984) In *1,3-Dipolar Cycloaddition Chemistry* (Ed: A. Padwa), Wiley, New York, 1.
133 Rostovtsev, V.V., Green, L.G., Fokin, V.V., and Sharpless, K.B. (2002) *Angew. Chem. Int. Ed. Engl.*, **41**, 2596.
134 Tornoe, C.W., Christensen, C., and Meldal, M. (2002) *J. Org. Chem.*, **67**, 3057.
135 Kolb, H.C. and Sharpless, K.B., (2001) *Angew. Chem. Int. Ed. Engl.*, **40**, 2004.

In fact, the 1,4-[1,2,3]triazolyl bridge offers interesting mode to generate a structural constraint.

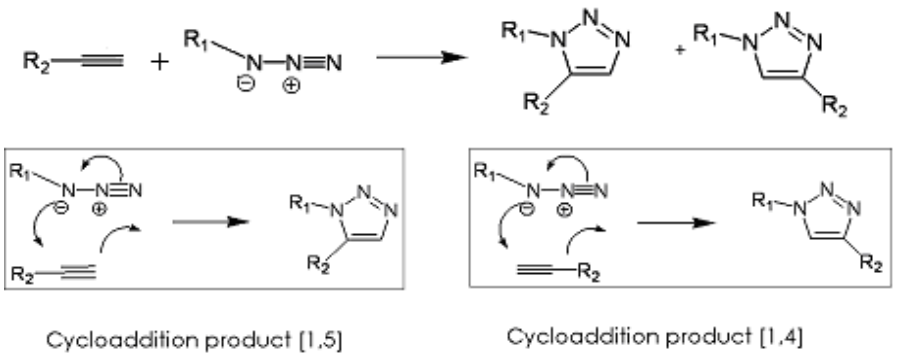
3.2 Click Chemistry

Click chemistry is an entire reaction group in which new substances are generated by joining small units together with heteroatom links (C–X–C).[135] A set of stringent *criteria* that a process should meet to be useful in this context are defined as follows:

- simple reaction conditions (ideally, the process should be insensitive to oxygen and water);
- very high yields;
- only inoffensive by-products that can be removed without using chromatographic methods;
- to be stereospecific (but not necessarily enantioselective);
- high thermodynamic driving force, usually greater than 20 kcal.mol^{-1}.

One of the most interesting reactions that can enter in the "click chemistry" definition are the cycloaddition reactions involving heteroatoms, such as Diels-Alder and, especially, 1,3-dipolar cycloaddition, that provide fast access to a variety of five- and six-membered heterocycles. In particular, Huisgen dipolar cycloaddition of azides and alkynes can ideally meet all prerequisites for an efficient "click chemistry".

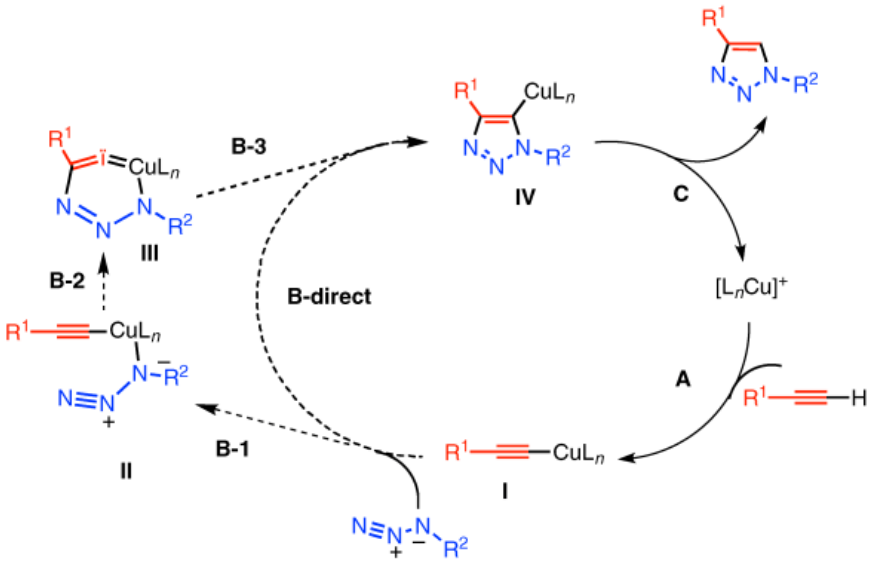
The chemistry of alkynyl and azido group is completely orthogonal to the chemistry of all endogenous functional groups in proteins; an example of their unique reactivity is the irreversible formation of triazole rings (Scheme 3.2-1).



Scheme 3.2-1 [1,2,3]-Triazole formation by [3+2] cycloaddition of an alkyne and an azide.

Usually the cycloaddition between azides and alkynes is carried out in refluxing toluene leading to mixture of the 1,4- and 1,5- regioisomers, but modified proteins and peptides may not survive in these strong conditions. On the other hand, what makes azides unique for click chemistry purposes is their stability toward H_2O and O_2. Meldal and co-workers[134] and Sharpless and co-workers[133] developed a mild and efficient method to generate 1,4-disubstituted [1,2,3]-triazoles by metal-catalyzed reaction using Cu(I) salts as catalyst even in the presence of H_2O.

The catalytic cycle of the Cu(I)-catalized cycloaddiction, proposed by Sharpless and coworkers[133] (Scheme 3.2-2), starts with the formation of a copper(I) acetylide I (as expected, no reaction is observed with internal alkynes) The reaction procedes either through a concerted [2+3] cycloaddition (B-direct) or with a stepwise, annealing sequence (B-1, B-2, and B-3), which proceeds via the a six-membered copper-containing intermediate III.



Scheme 3.2-2 Catalytic cycle of the Cu(I)-catalized cycloaddiction.

3.2.1 *Development of a new strategy based on click chemistry reaction to generate cyclopeptides*

We generated 1,4-[1,2,3]triazolyl-containing cyclopeptides by Huisgen 1,3-dipolar cycloaddition, designing appropriate modified non-coded amino acids bearing on the side-chain alkynyl or azido functions, and correctly protected on the α-functions to be introduced in peptide sequences using Fmoc/tBu solid-phase strategy. This cycloaddition is a simple chemical transformation that improves the resistance towards proteolytic degradation *in vivo*, and allowing the selective formation of covalent adducts in order to obtain modified bioconjugates.[136]

The 1,4-disubstituted [1,2,3]triazolyl serves as a rigid linking unit mimicking a *trans*-amide bond positioning the substituents in positions 1 and 4 at 5.1 Å apart, which is only slightly longer than the distance between two carbons separated by a *trans*-amide bond (3.9 Å). It has a slightly larger dipole moment (~5 Debye), which bisects the ring plane near atoms N3 and C5, and has the capacity of the N2 and N3 electron lone pairs to serve as hydrogen bond acceptors (Figure 3.2–1).

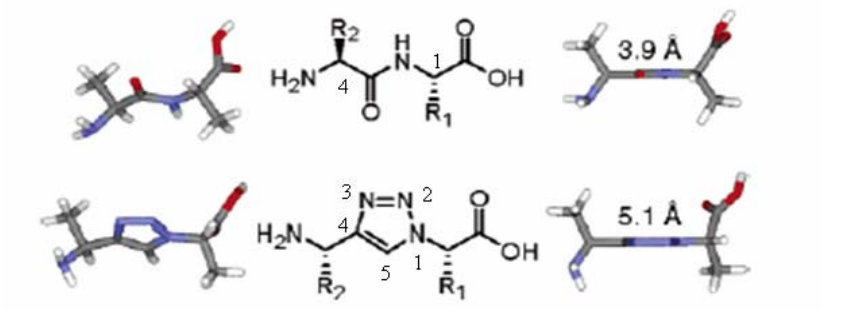


Figure 3.2–1 Topological similarities between amides and 1,2,3-triazoles.

136 Köhn, M., and Breinbauer, R. (2004)*Angew. Chem. Int. Ed.*, , **43**, 3106.

3.3 Inhibition of translation initiation in anticancer therapy, role of eIF4E binding protein

Anticancer therapy by inactivation of Translation Initiation is one of the most important projects developed at the Laboratory for Translational Research of Harvard Medical School. Principles and methodologies of peptide click chemistry have been applied to this subject.

Most cancer therapies target some specific characteristic of cancer cell in order to suppress or hinder their proliferation. Unfortunately few approaches are fully selective and in most cases the strategy of action relays on the concept of damaging cancer cells more than normal cells. This means that most existing treatments target one peculiar metabolic feature that is more important for cancer cells than for normal ones. For example genetic instability, which is higher for cancer cells (and essential for tumor developing) can be enhanced (by radiation or mutagenic compounds) to a point that is unbearable for cancer cells and damaging but still acceptable for normal cells. Similarly it has been demonstrated that cancer cells need to overexpress a set of proteins critical for their proliferation. It is possible to take advantage of this need by inhibiting translation initiation and thus making the protein synthesis of cancer cells not efficient enough for their proliferation.

Protein synthesis is regulated at multiple levels but most of the regulation occurs at the initiation stage of translation that is the process by which the genetic information transcripted into m-RNA is converted into proteins[137] (Figure 3.3–1). The 3' and 5' terminal portions of eukaryotic m-RNA bind a protein complex colle eIF4E which in turn is linked to the ribosomes.

137 Sonenberg, N.and Hinnebusch, A.G. (2009) *Cell* **136**, 731.

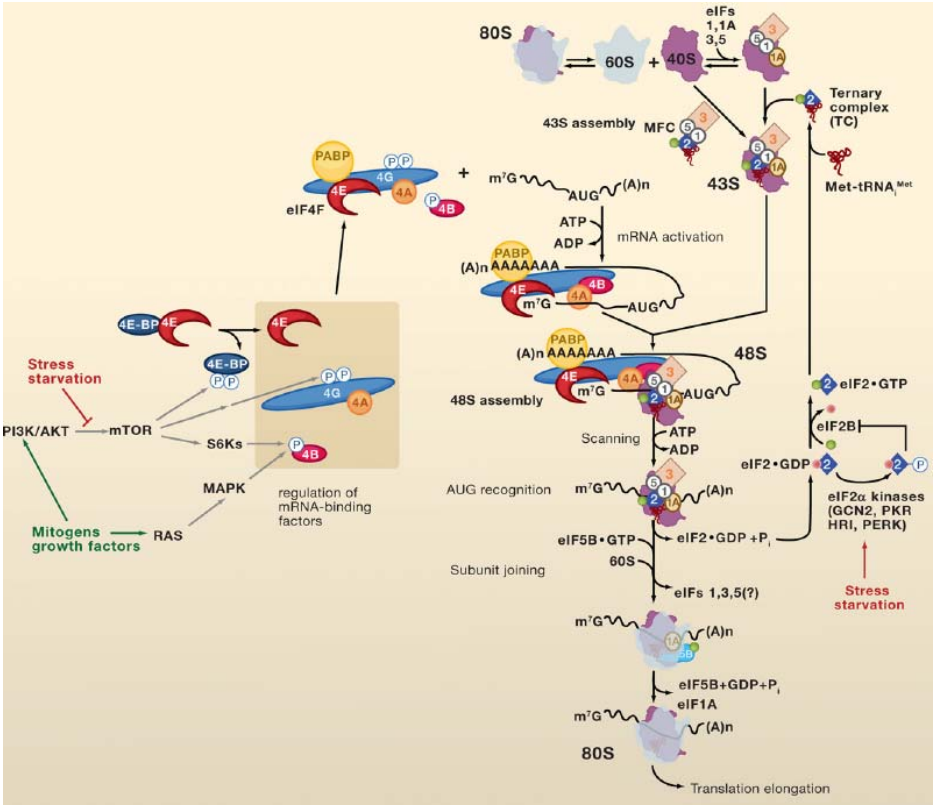


Figure 3.3–1 Regulation of Translation Initiation in eukaryotic cells.

The most important stage of translation initiation is the assembly of the eIF4F complex of proteins. The molecular basis of this process has being studied intensively and one of its key stages is considered to be the binding of the cap-protein eIF4E to the 5'-cap extremity of m-RNA. However, if the regulatory protein 4E-BP (4E binding protein) binds to eIF4E the mRNA is not able any more to interact with eIF4E and thus Translation is suppressed. Figure 3.3–2.

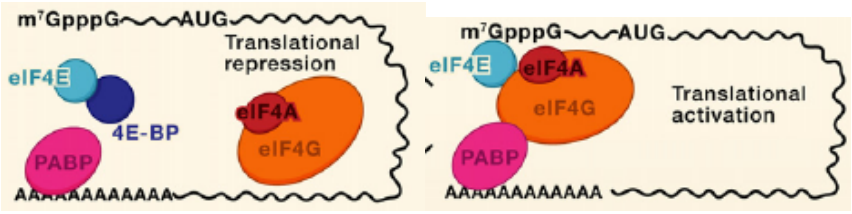


Figure 3.3–2 Repression/activation of Translation Initiation depending on the binding of eIF4E-mRNA to the eIF4A/eIF4G complex and regulated by 4E-BP.

Our goal is the generation of synthetic modified peptides whose sequence reproduces the binding site of the 4E-BP. By inhibiting Translation Initiation such peptides represent potential anticancer agents. We plan moreover to stabilize with side-chain to side-chain triazole bridges the secondary structure of those peptides that are known to assume an α-helical conformation.

3.3.1 *Selection of the model system, eIF4E binding protein peptide*

The development and study of the new intermolecular side-chain-to-side-chain [1,2,3]triazole containing peptide modification was carried out in the context of a molecular model for those structural and biological information still known. In previous work a peptide bearing a fluorescein residue, 4GII-FITC, was synthesized in the Laboratory for Translational Research of the Harvard Medical School:

KKQYDRELLDFQFK(Fluorescinyl)

The 14 amino acid sequence of such peptide reproduces the sequence of the binding site of 4GII (also called 4E-BP) which is the main eIF4E binding protein and acts as an inhibitor (Figure 3.3–2 shows the structure, determined by x-ray analysis, of an eIF4E/eIF4E binding protein peptide complex[149]). The synthetic peptide has proved to possess a high affinity for the protein and remarkable inhibiting effects.

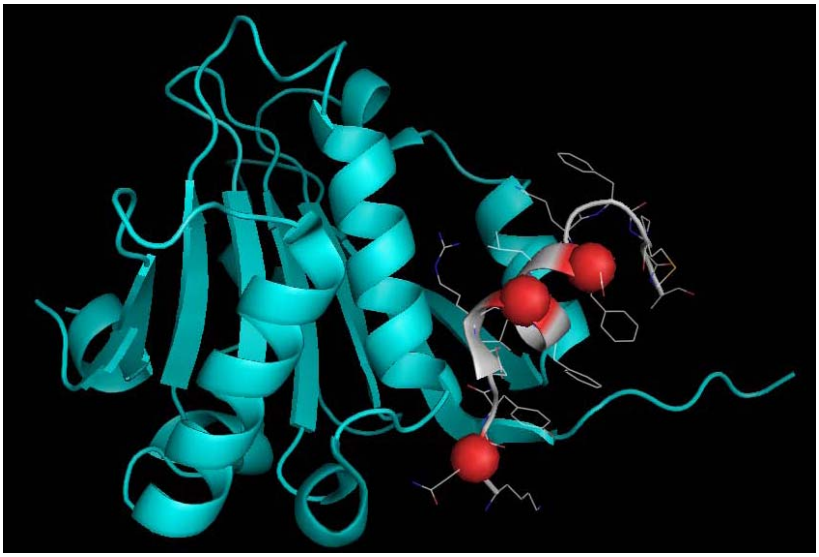
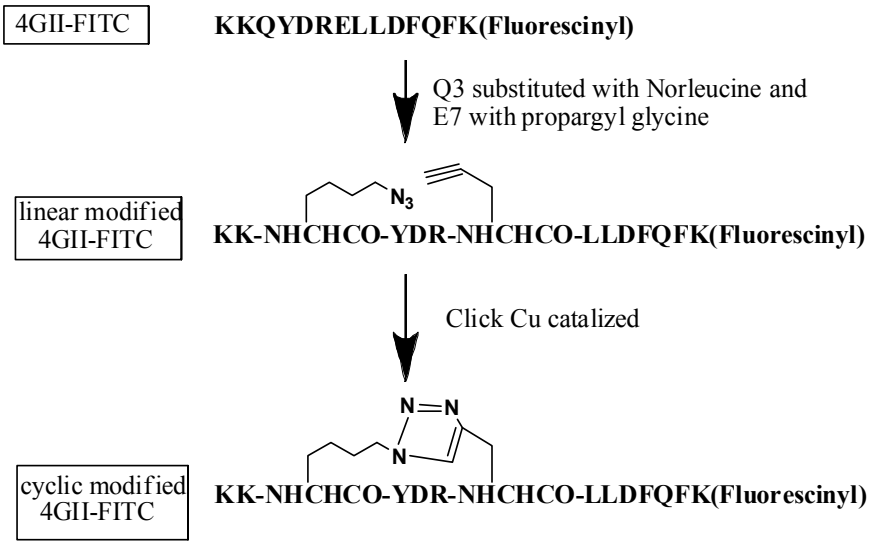


Figure 3.3–3 Structure, determined by x-ray analysis, of a eIF4E/eIF4E binding protein peptide complex.

Our aim was to develop cyclic, clicked eIF4E b.p. peptide analogs modifying the sequence of 4GII-FITC with non-coded alkynyl and azido amino acids specifically synthesized for click chemistry (Scheme 3.3-1).The triazole bridge in i/i+4 positions, stabilizing the α-helical structure of the peptide, should be able to increase its affinity for eIF4E: The fuorescein residue should be useful for Fluorescence Polarization assays for binding affinity measurements.



Scheme 3.3-1 Scheme of **4GII-FITC** modification and cyclization of the linear eIF4E b.p. peptide analogs.

3.4 New collection of amino acids to develop clicked peptides

The 1,4-disubstituted-1,2,3-triazolyl moiety as an amide bond surrogate, specifically assembled through biorthogonal Cu^{I}-catalyzed Huisgen 1,3-dipolar [3 + 2] cycloaddition of an alkynyl to an azido function, generated an unmet need for specifically designed amino-acid-derived building blocks. To this end, there have been several reports in which non-coded α-amino acids modified by ω-azido and ω-alkynyl functionalities were used as biorthogonal reporter peptides and proteins.[134, 138, 139, 140] In addition to intermolecular side-chain-to-side-chain click reaction between α-amino acids modified by ω-azido and ω-alkynyl functions these non-coded amino acids can be used also to conjugate the peptide of interest through an orthogonal reaction with haptens, tags, and labels to generate special reagents.

To date, diazo transfer reaction was used to generate α-azido-acids in solution and α-azido-peptides on solid support.[141] ω-Azido α-amino acids such as β-azidoalanine were prepared from either the salt of α-amino-β-propiolactone[142] or the protected homoserinol. On the other hand only propargylglycine (Pra) is commercially available. It is evident that there is an unmet need for syntheses that will furnish an extensive homologous series of ω-azido and ω-alkynyl-N^{α}-protected amino acid as s tools to generate 1,4-disubstituted-1,2,3-triazolyl containing peptides or for various applications. For example, combinatorial libraries of peptides containing ω-azido- or ω-alkynyl-α-amino acids can be used to generate libraries of higher order by diversifying their structure through click reactions with a variety of respective alkynyl and azido-containing reagents.

---

138 Deiters, A., Cropp, T.A., Mukherji, M., Chin, J.W., Anderson, C., and Schultz, P.G. (2003) *J. Am. Chem. Soc*., **125**, 11782.
139 Dondoni, A., Giovannini, P.P. and Massi A. (2004) *Org Lett.*, **6**, 2929.
140 Lin, H. and Walsh, C. T. (2004) *J Am Chem Soc.*, **126**, 13998.
141 Punna, S., Kuzelka, J., Wang, Q., and Finn, M. G. (2005) *Angew. Chem., Int. Ed.*, **44**, 2215.
142 Arnold, L.D., May, R.G., and Vederas, J.C. (1988) *J. Am. Chem. Soc.*, **110***,* 2237.

Therefore, we developed an efficient and convenient synthetic pathway to generate non-coded N^{α}-Fmoc-ω-azido- and N^{α}-Fmoc-ω-alkynyl-amino acids suitable for Fmoc/*t*-Bu SPPS (Figure 3.4–1). The structural diversity enabled by the small library of building blocks will allow the introduction of 1,4-disubstituted-1,2,3-triazolyl moieties into peptides of interest as modifiers of physicochemical and biological properties.

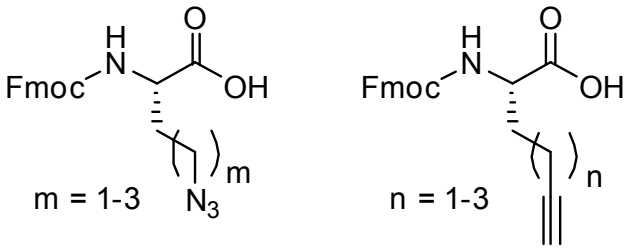
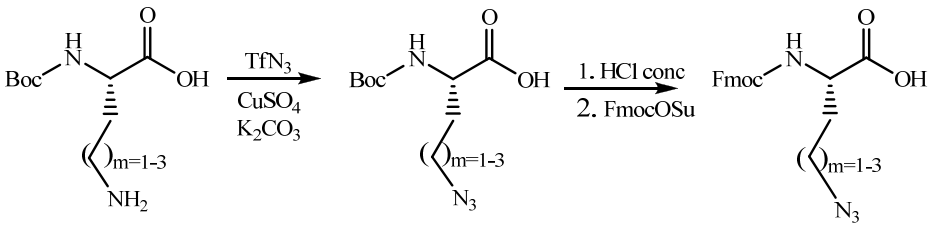


Figure 3.4–1 N^{α}-Fmoc-ω-azido- and N^{α}-Fmoc-ω-alkynyl-α-amino acids for Fmoc/*t*-Bu SPPS.

3.4.1 *N^{α}-Fmoc-ω-azido-α-amino acids*

An efficient and convenient methodology for generating organic azides is Cu^{II}-catalyzed diazo transfer from amines by trifluoromethanesulfonyl azide.[143] Diazo transfer from triflic azide occurs effectively with Cu^{2+}, Ni^{2+} or Zn^{2+} as catalysts. The process is amenable to scale-up, can be carried out using commercially available reagents and does not require anhydrous conditions.

Since the transformation does not involve the C^{α}, it is accomplished with a complete retention of the original configuration of the α-amino acid residues. In this manner N^{α}-Boc-Lysine, N^{α}-Boc-Ornithine, and N^{α}-Boc-AminoButirric acid were converted in good yield to the respective N^{α}-Fmoc-ε-azido-Norleucine (**3**), N^{α}-Fmoc-δ-azido-Norvaline (**4**) and N^{α}-Fmoc-γ-azido-hSerine (**5**) (Scheme 3.4-1).



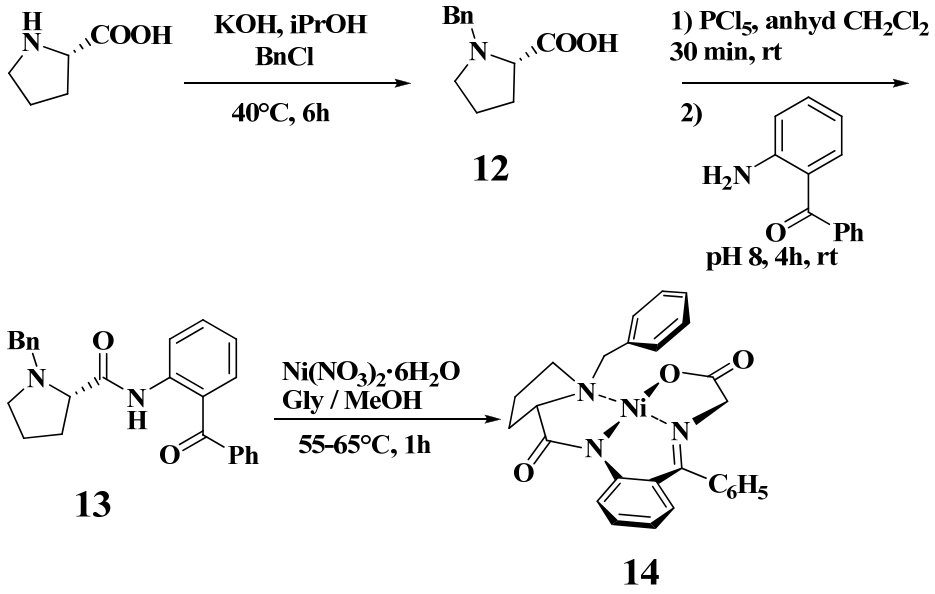
Scheme 3.4-1 General procedure for the synthesis of building blocks **3, 4** and **5**.

This strategy required the use of the N^{α}-Boc-protected amino acids as starting material because of their stability in these reaction conditions.

143 Nyffeler, P.T., Liang, C.H., Koeller, K.M. Wong, C.-H. (2002) *J. Am. Chem. Soc.* **124**, 10773.

3.4.2 *N^{α}-Fmoc-ω-alkynyl-α-amino acids*

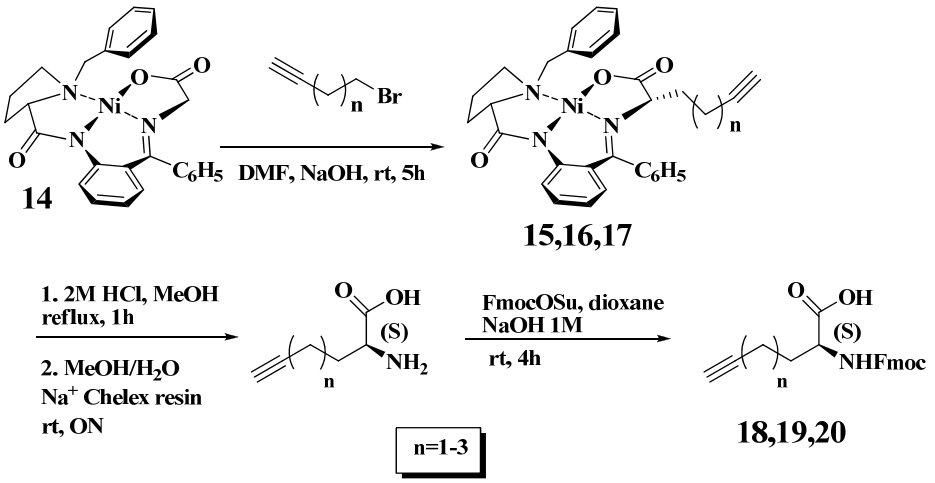
The strategy to synthesize the ω-azido-α-amino acid homologs containing $(CH_2)_n$ (where n=2,3 and 4) on the side chain, employed elaboration of glycine by chiral synthesis. The Ni^{II} complex of the Schiff base [Ni^{II}-(*S*)BPB-Gly] (**14**) derived from glycine and (*S*)-2-(*N*-benzylprolyl)aminobenzophenone (BPB) (**13**) was used as a chirality inducer during the C^{α}-alkylations (Scheme 3.4-2).[144]



Scheme 3.4-2 Synthesis of the chirality inducer [Ni^{II}-(*S*)BPB-Gly].

Alkylations through the *si*-face of the glycine enolate are largely favored to lead to (S)-α-amino acids.

144 Belokon, Y.N., Bulychev, A.G., Vitt, S.V., Struchkov, Y.T., Batsanov, A.S., Timofeeva, T. V., Tsyryapkin, V.A., Ryzhov,M.G., Lysova, L.A., Bakhmutov, V.I. and Belikov, V.M. (1985) *J. Am. Chem. Soc*., **107**, 4252.



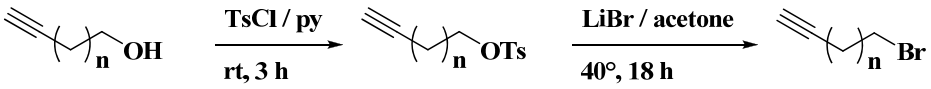
Scheme 3.4-3 Chiral synthesis of L-ω-alkynyl-α-amino acids employing the chirality inducer [Ni^{II}-(*S*)BPB-Gly].

This method is based on a reaction between [Ni^{II}-(*S*)BPB-Gly] and alkynyl bromides (Scheme 3.4-3). Alkylation of glycine by alkynyl bromides in the presence of the chirality inducer [Ni^{II}-(*S*)BPB-Gly], which was carried out according to a procedure described in the literature[144], yielded the corresponding ω-alkynyl-α-amino acids in good diastereoisomeric excess (monitored by UPLC, Table 3.4-1). The separation of the diastereoisomers was successfully achieved by FCC.

Products of alkylation of the Gly-Ni-BPB complex with	*S*,*S*-diastereoisomeric excess (%)	Alkylation yield (%)
6-Bromohex-1-yne (**9**)	88	68
5-Bromopent-1-yne (**10**)	71	58
4-Bromobut-1-yne (**11**)	*63*	41

Table 3.4-1 Diastereoisomeric excess of the alkynylated complexes.

The coupling step between *N*-benzyl-(*S*)-proline (BP) (**12**) and 2-aminobenzophenone was accomplished *in situ* by PCl_5-mediated transformation of the carboxyl function into the acyl chloride, which was used to acylate the 2-aminobenzophenone at pH 8. Alkylation of [Ni^{II}-(*S*)BPB-Gly] by an excess of ω-alkynyl bromide (1.4 equiv.), in the presence of NaOH in anhydrous acetonitrile, proceeded for 5 hours and resulted in a good yield. The ω-alkynyl bromides were prepared in moderate yields from the corresponding alcohols by treating them with TsCl (product **6**, **7**, and **8**) followed by LiBr (product **9**, **10,** and **11**) (Table 3.4-2).



n	OTs (yield %)	Br (yield %)
3	(96)	(31)
2	(93)	(28)
1	(95)	(30)

Table 3.4-2 Synthesis of the ω-alkynyl bromides.

Hydrolysis of the resultant alkylated complex in 2M HCl was complete in one hour. Workup afforded a pale green solid containing the free alkynylated amino acids in the presence of Ni^0 and BPB. The chirality inducer BPB was recovered by washing the crude with acetone and separated from traces of Ni by redissolving the solid in DCM, and washing it with water. The removal of traces of Ni^0 from the free and crude ω-alkynylated α–amino acids, which interfered in the following Fmoc-protection reaction, was accomplished by overnight solid phase extraction of its MeOH/H_2O solution with Na^+ Chelex resin.

Finally, the free ω-alkynylated α-amino acids were N^{α}-protected as Fmoc to yield the building blocks (**18**, **19**, and **20**) in quantities adequate for SPPS by Fmoc/tBu strategy (Table 3.4-3).

Amino acids	Yield (%)	Pure compound
N^{α}-Fmoc-ε-alkynyl-Norleucine (**18**)	32	90 mg
N^{α}-Fmoc-δ-alkynyl-Norvaline (**19**)	39	90 mg
N^{α}-Fmoc-γ-alkynyl-hSerine (**20**)	29	80 mg

Table 3.4-3 Yield and scale of building blocks (**18**, **19**, and **20**).

We hypothesized that the low yields of the final products are due to the difficulties to obtain the pure free amino acids without the presence of salts found after neutralization of the solution, and therefore we used an excess of Fmoc-succinimmide that leads to side products formation decreasing yields of final products.

3.5 Collection of eIF4E binding protein cyclopeptides analogs

Aim of this study is to enhance the affinity of eIF4E binding protein peptide with eIF4E by stabilizing the secondary structure of the peptide. This will be achieved by introduction of a 1,2,3-triazole ring in positions i and i+4 at different levels of the sequence. Under the constraint of such bridge the peptide will assume an α-helical conformation which is the conformation of the binding site of eIF4E binding protein from which the peptide sequence is derived. Moreover, we synthesized a panel of new modified Fmoc-amino acids bearing alkynyl and azide function on the side chains (Figure 3.5–1).

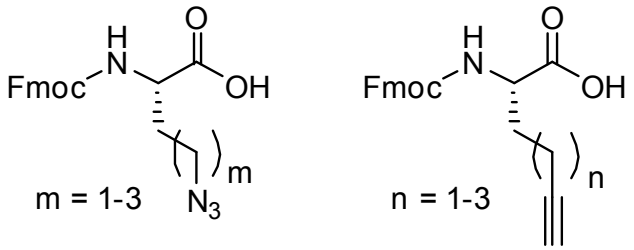


Figure 3.5–1 N^{α}-Fmoc-ω-azido- and N^{α}-Fmoc-ω-alkynyl-α-amino acids for Fmoc/*t*-Bu SPPS.

Regarding the position of the sequence in which introducing the triazole bridge, initially we decided to modified the C-terminal portion (i-i+4=positions 3-7) and we have generated an eIF4E b.p. peptide analog N4+C1 type, product **XVI**, having an azido function on the left and an alkynyl one on the right (Table 3.5-1).

This choice for the triazole position has been shown, with Fluorescence Polarization Assay (see section 3.6.1), to be unfavorable for improving the affinity of the peptide for eIF4E.

	Peptide sequence	CH_2 ring
XVI	Ac-[1]Lys-Lys-Gln-Tyr-Asp-Arg-Glu-Phe-Leu-Leu-Asp-Phe-Gln-Phe-[15]Lys-NH	-
XVII	Ac-Lys-Lys-[2]Nrl(ε-N_3)-Tyr-Asp-Arg-[7]Pra-Phe-Leu-Leu-Asp-Phe-Gln-Phe-Lys-NH	5

Table 3.5-1 Peptide sequences of eIF4E binding protein peptide analogs XV and XVI. Non-coded amino acids [Pra= propargyl-Glycene and Nrl(ε-N_3)= ε-azidoNorleucine] are outlined.

Therefore we chose to shift the triazole bridge position in the middle of the sequence (positions 7-11, Figure 3.5–2)

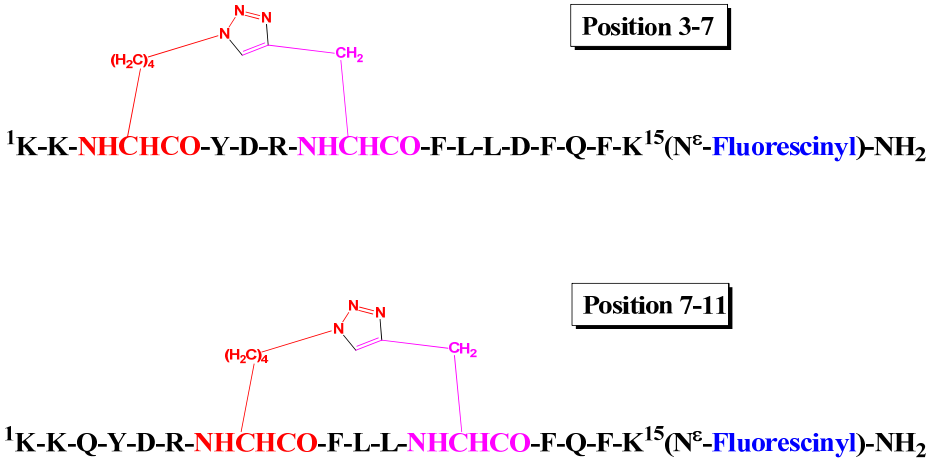


Figure 3.5–2 Triazole bridge positions 3-7 and 7-11 of the eIF4E binding protein peptide analogs.

Moreover two sets of modified amino acids (three alkynes and three azides) in two given positions of a sequence (i-i+4 to stabilize an α-helix) can generate 18 combinations' of triazole bridges with different number of methylenes and different orientations of the triazole. This number comes from the 3^2=9, Nm+Cn, possible combinations having the azide on the left side (position i) with the alkyne on the right (position i+4) (Figure 3.5–3) plus the other 3^2=9 Cn+Nm, possible combinations generated by alkyne on the left side and the azide on the right (Figure 3.5–4).

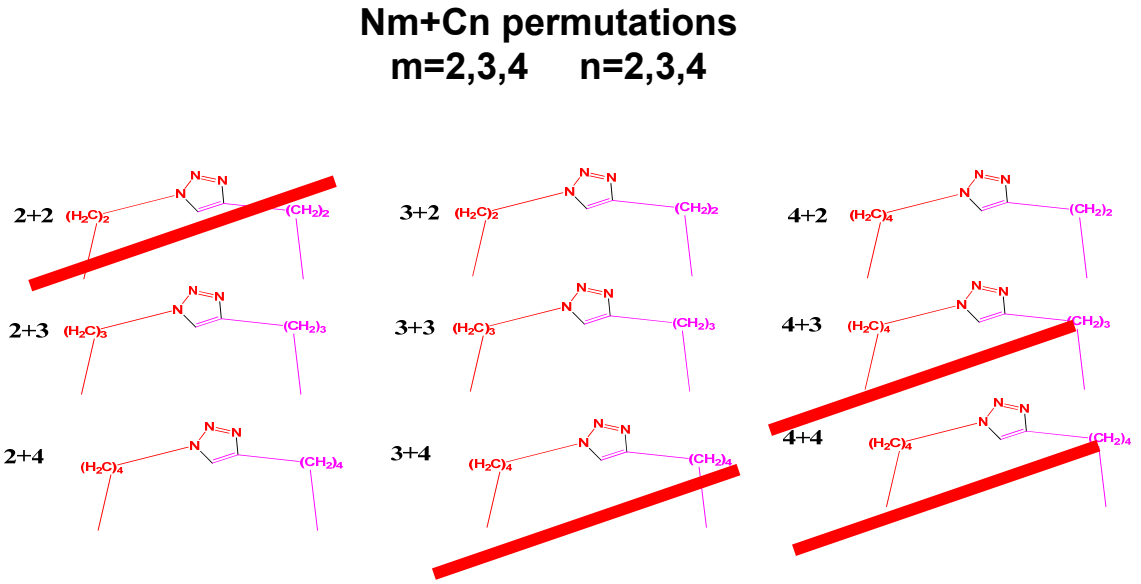


Figure 3.5–3 The 3^2=9 possible combinations of Nm+Cn triazole bridges having the azide on the left side (position i) and the alkyne on the right side (position i+4).

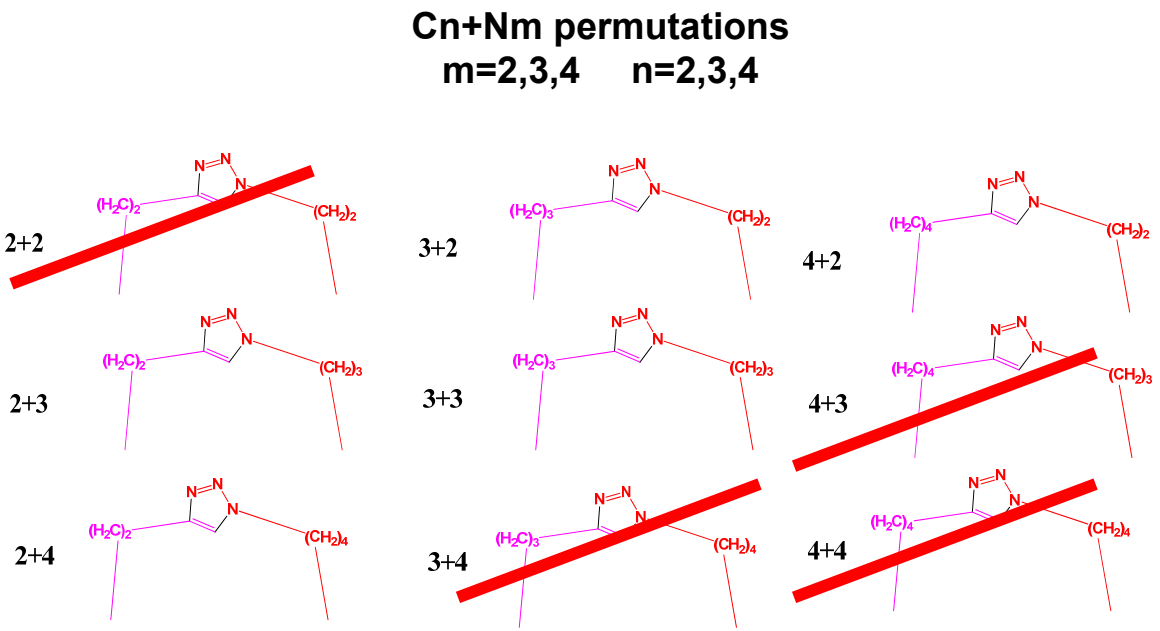


Figure 3.5–4 The 3^2=9 possible combinations of Cn + Nm triazole bridges having the azide on the left side (position i) and the alkyne on the right side (position i+4).

Based on previous experiments on cyclic, clicked peptides[145] we have discarded those combinations generating triazole bridges either too short (4 methylene units) or too long (7-8 methylene units). A number of methylene between 5 and 6 has been shown to be the most favorable for α-helix stabilization. Thus we have planned to focus on 10 out of 18 possible combinations (Table 3.5-2).

N°CH_2	**XVIII**	**XIX**	**XX**	**XXI**
5	N2+C3	C3+N2	N3+C2	C2+N3
N°CH_2	**XXII**	**XXIII**	**XXIV**	**XXV**
6	N2+C4	C4+N2	N4+C2	C2+N4
N°CH_2	**XXV**	**XXVI**		
6	N3+C3	C3+N3		

Table 3.5-2 The selected 10 eIF4E binding protein click peptide with triazole length of 5-6 methylene units. Those products that are under characterization are outlined in green.

Peptide analogs **XVIII**, **XIX**, and **XX** (Table 3.5-3) have been synthesized purified and characterized both with in solution NMR and with Fluorescence Polarization Assay.

	Peptide sequence	**CH_2 ring**
XVIII	Ac-Lys-Lys-Gln-Tyr-Asp-Arg-^{7}hSer(γ-N_3)-Phe-Leu-Leu-^{11}Nvl(δ-yl)-Phe-Gln-Phe-Lys-NH	**5**
XIX	Ac-Lys-Lys-Gln-Tyr-Asp-Arg-^{7}Nvl(δ-yl)-Phe-Leu-Leu-^{11}hSer(γ-N_3)-Phe-Gln-Phe-Lys-NH	**5**
XX	Ac-Lys-Lys-Gln-Tyr-Asp-Arg-^{7}Nvl(δ-N_3)-Phe-Leu-Leu-^{11}hSer(γ-yl)-Phe-Gln-Phe-Lys-NH	**5**

Table 3.5-3 Peptide sequences of eIF4E binding protein peptide analogs **XVIII-XX.** Non-coded amino acids [hSer(γ-N_3)= hSerine(γ-N_3), Nvl(δ-yl)= Norvaline(δ-yl), Nvl(δ-N_3)= Norvaline(δ-N_3), hSer(γ-yl)= hSerine(γ-yl)] are outlined.

[145] Scrima, M., Le Chevalier-Isaad, A., Rovero, P., Papini, A.M., Chorev, M., and D'Ursi A.M. (2009) *Eur JOC* in press.

3.5.1 *Linear peptides*

Peptides **XVI-XX** were generated as described in section 5.6 on an automatic synthesizer starting from Rink amide resin. The first Lysine residue in the C-terminal position was introduced with an ivDde side chain protection for fluorescein coupling. The non-coded amino acids were inserted by manual SPPS in the peptide sequence in the positions i, i+4 as reported in Table 3.5-4.

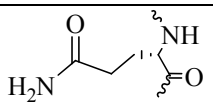
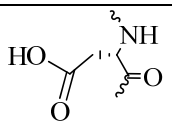
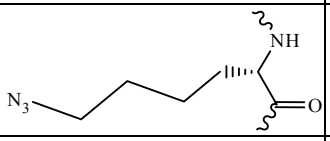
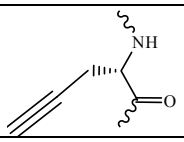
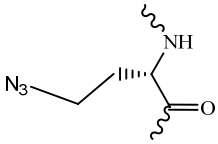
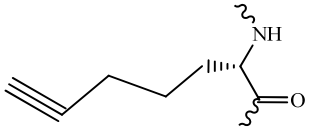
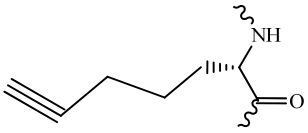
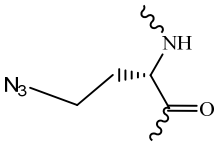
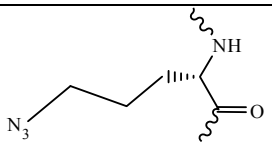
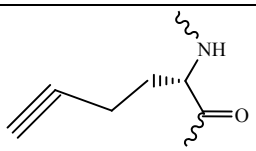
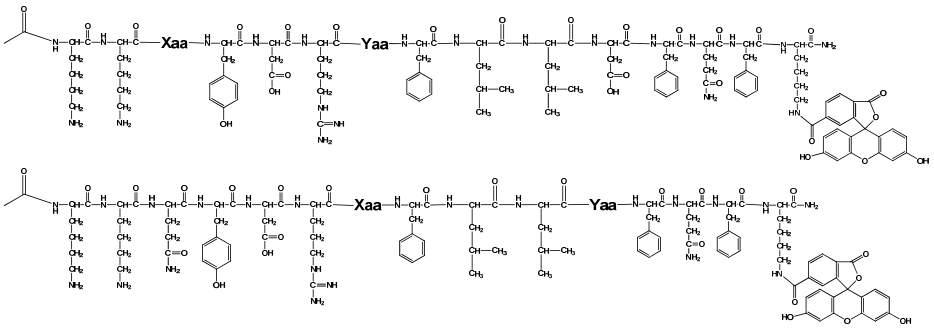
Linear peptide	Xaa	Yaa
XVI		
XVII		
XVIII		
XIX		
XX		

Table 3.5-4 Non-coded amino acids introduced in the peptide sequences in the i, i+4 positions. Peptide **XVI** does not contain any non-coded amino acid.



At the end of the synthesis and after N-terminal α-amino acetylation ivDde group was removed with 2% hydrazine in DMF from the C-terminal Lysine and the ε-amino was coupled to a carboxy-fluorescein residue. The peptides were purified by semi-preparative HPLC and characterized by ESI-MS. Analytical data are reported in Table 3.5-5.

Linear peptide	Rt(min) HPLC	MS^+(ESI) *m/z*	
		found	calculated
XVII	6.3	$[M+H]^+$=2354.92 $[M+H]^{2+}$=1177.96	$[M+H]^+$=2355.62
XVIII	15.4	$[M+H]^+$=2411.28 $[M+H]^{2+}$=1206.14	$[M+H]^+$=2410.7
XIX	15.5	$[M+H]^+$=2410.98 $[M+H]^{2+}$=1205.99	$[M+H]^+$= 2410.7
XX	14.1	$[M+H]^+$=2411.44 $[M+H]^{2+}$=1206.53	$[M+H]^+$= 2410.7

Table 3.5-5 Chemical data of linear peptides **XVIII-XIX**. All HPLC were performed with a gradient 30to70% B in A.

3.5.2 *Cyclopeptides*

3.5.2.1 General features on click chemistry reaction conditions

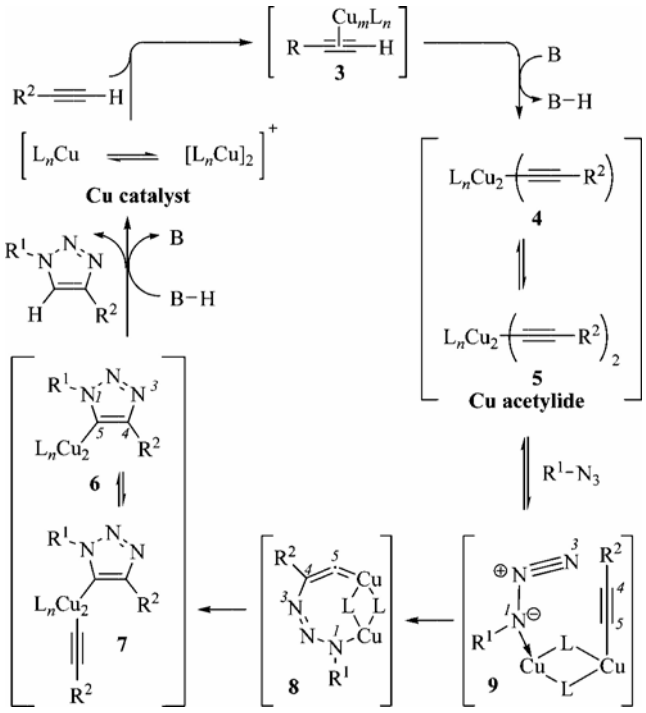
The catalytic cycle of the Cu(I)-catalyzed alkyne-azide "click" cycloaddition (Scheme3.5-1) shows that a source of Cu(I) is fundamental for the regioselectivity. Meldal and co-workers[134] described the use of Cu(I) salts in

the solid phase, while Shrarpless[133] reported solution phase by *in situ* reduction of Cu(II) salts. It is also possible to envisage the formation of Cu(I) from the oxidation of metallic copper.

Alkyne π complexation requires ligand dissociation and is endothermic in acetonitrile. However, in aqueous solution the formation of copper species **4** (Scheme 3.5–1) is exothermic, a result consistent with experimental findings of a rate acceleration in water.

Moreover, the copper coordination lowers the pKa of the alkyne CH, thus making deprotonation in aqueous systems possible without the addition of a base; DIPEA and 2,6-lutidine minimizes the side-product formation.

In this context, the choice of the condition reaction (solvent, base) is crucial for cycloaddition.



Scheme3.5-1 Proposed outline of species involved in the catalytic cycle of the CuI-catalyzed alkyne-azide "click" cycloaddition.[146]

146 Bock, V.D., Hiemstra, H., and van Maarseveen, J.H. (2006)*Eur. J. Org. Chem.*, **14**, 51.

3.5.2.2 Synthesis and characterization of cyclopeptides

Copper catalyzed click cyclization was performed on peptides **XVII-XX** using an excess of copper sulfate (10eq) and ascorbic acid (20eq) in 1:1 tBuOH/H_2O. The reaction was monitored by HPLC and LC-MS. It is important to note that the cyclic, clicked forms have exactly the same mass of the corresponding linear forms so that cannot be distinguished by conventional LC-MS. However the two forms has different retention times, in particular all clicked peptides synthesized so far have a shorter retention time with respect to their corresponding linear forms The cyclic peptides were purified by solid phase extraction (SPE) followed by semi-preparative HPLC purification and characterized by HPLC and ESI-MS (Table 3.5-6).

Peptide	**Linear Rt (min) HPLC**	**Clicked Rt (min) HPLC**	**MS^+(ESI) *m/z***	
			found	**Calculated**
XVIIc	6.3	4.8	$[M+H]^+$=2354.72 $[M+H]^{2+}$=1177.86	$[M+H]^+$=2355.62
XVIIIc	15.4	15.4	$[M+H]^+$=2411.35 $[M+H]^{2+}$=1206.17	$[M+H]^+$=2410.7
XIXc	15.5	15.5	$[M+H]^+$=2410.99 $[M+H]^{2+}$=1206.00	$[M+H]^+$= 2410.7
XXc	14.1	14.01	$[M+H]^+$=2411.21 $[M+H]^{2+}$=1206.10	$[M+H]^+$= 2410.7

Table 3.5-6 Chemical data of clicked peptides **XVIIc-XXc**.

3.5.2.3 Side reactions during peptide click-mediated cyclization

Previous experiments[147] showed that spontaneous cyclization (in absence of metal catalysts) of linear peptides bearing alkyne and azide functions occur to a very little extent as can be seen from Figure 3.5–5.

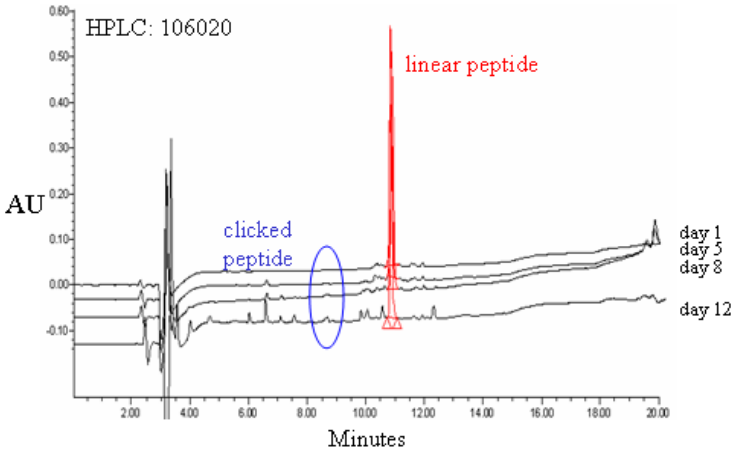


Figure 3.5–5 Example of a typical evolution of the HPLC chromatogram (method : 10% to 60% of B in A for 20 minutes, A = 0.1% TFA in H_2O, B = 0.1% TFA in CH_3CN) obtained in 12 days of solution of a linear peptide developed for click chemistry.

Moreover alkyne-azido functions of modified peptides are stable if kept in dry solid form under low temperature. During peptide click reaction intermolecular cyclization obviously competes with intermolecular oligo-polimerizations between alkyne and azide residues of different molecules. Under a high dilution, of the order of 0.1 μmM, the latter intermolecular reactions are suppressed. However the triazole orientation, the size of the bridge and the position of the bridge itself in the peptide sequences seems to play a role. In our preliminary studies no undesired oligomerization was

[147] Le Chevalier-Isaad A., PhD Thesis 2008.

detected when the triazole was closed to the N-terminal residue of the peptide (analog **XVII**). On the contrary a certain degree of oligomerization was present when the bridge was in the middle of the sequence, as for the case of peptides **XVIII-XX**. The reaction rate itself was slower and complete exhaustion of the linear precursors took longer (3-4 days) compared to **XVII** (12 h). This may imply that a more central position for the bridge could hinder the intermolecular cyclization making the intermolecular reaction more favorable. At the same time we noticed a higher degree of oligomerization for peptide **XVIII** (C3+N2) compared to peptide **XIX** (N2+C3) and peptide **XX** (N3+C2) suggesting that the orientation of the triazole influences the rate of the cyclization. Figure 3.5–6 reports the LCMS chromatogram for click reaction of peptide **XIX**. Together with the linear and cyclic peptides, two more peacks are present being the dimmer and trimer oligomers (the correspondent mass is shown in Figure 3.5–7. More systematic studies will be necessary to elucidate the peptide click cyclization dynamics.

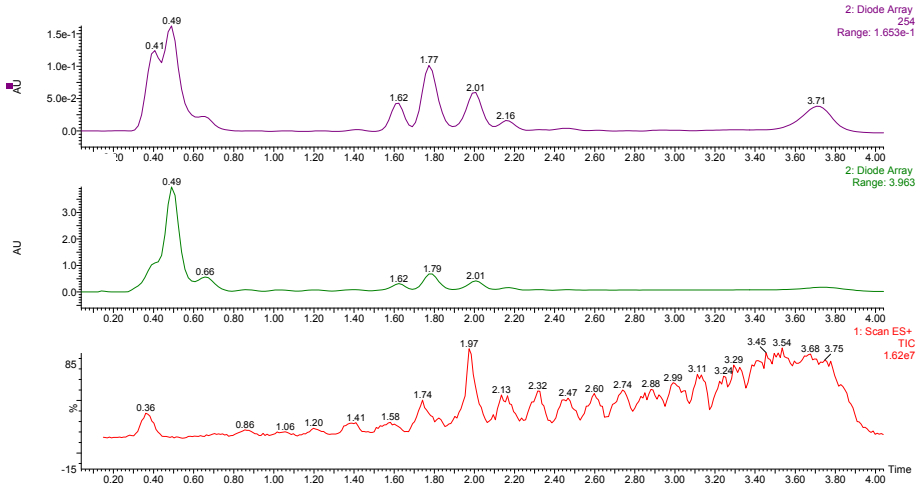


Figure 3.5–6 UPLC-MS of the cyclization reaction of peptide **XIX,** method 20to60% B in A (215nm, 254nm, TIC). The chromatograms show the trimer adduct at 1.62 min, the dimer adduct at 1.77 min, the clicked peptide at 2.02 min, and the linear peptide at 2.16min.

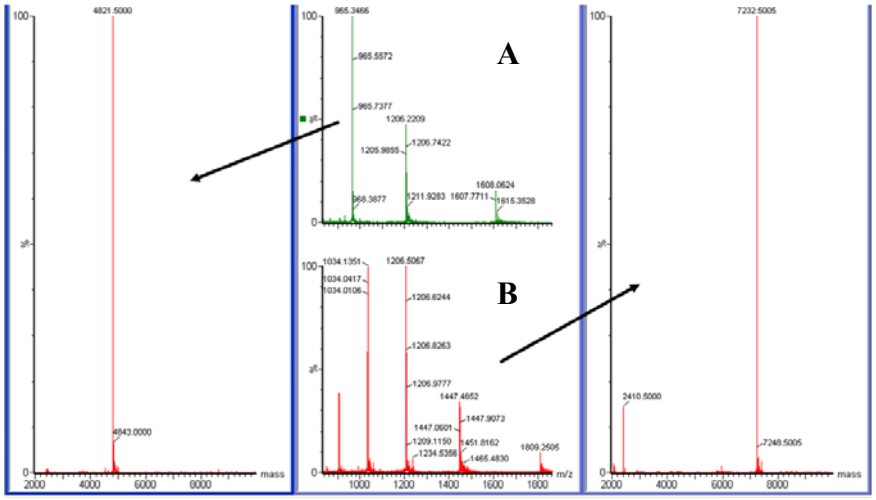


Figure 3.5–7 Mass spectrums of the purified dimmer form (A) and trimer form (B) with corresponding calculated [M+H]+ values.

3.5.2.4 Peptide templated click cyclization

Sharpless and co-worker[148] have demonstrated that the *in situ* click-chemistry approach can be successfully applied to the formation of HIV-1-Pr inhibitors. HIV-1 Protease assembles its own potent inhibitor through formation of the triazole linkage from azide- and alkyne-containing fragments that are themselves poor binders. In fact the protein itself acts as a template for the reaction and greatly increases the rate of formation of the 1,4-triazole product. In this context, to study the ring size influence, we have set up the synthesis of a collection of linear peptides containing alkynyl- and azido-moiety on the side chain backbone for *in situ* template cyclization**.** The assumption is that eIF4E binding protein peptide analogs in presence of eIF4E will fit the protein binding site assuming an α-helical conformation (this is known to happen from x-ray crystal studies of eIF4E/eIF4E b.p. peptides)[149]. Under these conditions the alkyne and azido side chain functions will find themselves in close contact and will cyclise without any need of metal catalysts as has been the case of HIV-1-Pr inhibitors.

Templated click cyclization studies are currently held on linear peptide analogs **XIX** and **XX**.

148 Whiting, M., Muldoon, J., Sharpless, K.B., Elder, J.H. and Fokin, V. V. *Angew. Chem. Int. Ed*. 2006, **45**, 1435.
149 Marcotrigiano, J., Gingras, A.D., Sonenberg, N., Burley, S.K. (1999) *Molecular Cell*, **3**, 707.

3.6 Fluorescence Polarization Assay studies

Fluorescence anisotropy is a method for measuring the binding interaction between two molecules, and can be used to measure the binding constant (or the inverse, the disassociation constant) for the interaction.[150] The basic idea is that a fluorophore excited by polarized light (light whose "waves" only go into one direction) will also emit polarized light. However, if a molecule is moving, it will tend to "scramble" the polarization of the light by radiating at a different direction from the incident light (Figure 3.6–1). The "scrambling" effect is greatest with fluorophores freely tumbling in solution and decreases with decreased rates of tumbling. Protein interactions can be detected when one of the interacting partners is fused to a fluorophore: upon binding of the partner molecule a larger, more stable complex is formed which will tumble more slowly (thus, increasing the polarization of the emitted light and reducing the "scrambling" effect).

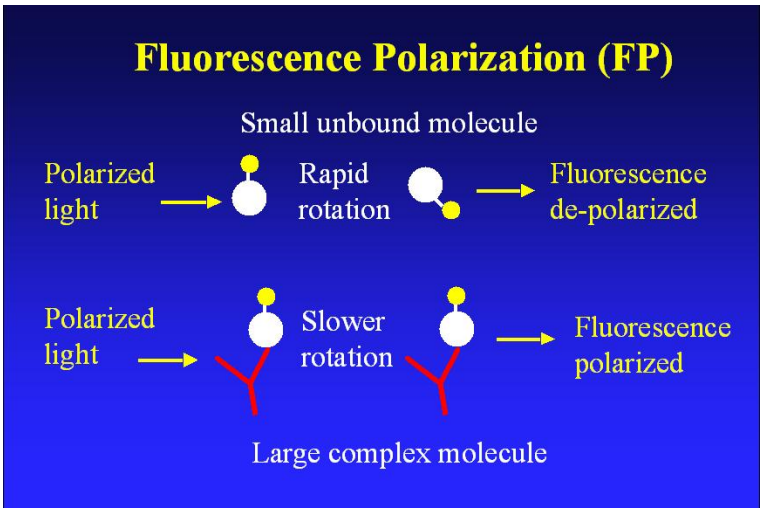


Figure 3.6–1 The principle of Fluorescence Polarization.

150 Jolley, M.E., Stroupe, S.D., Schwenzer, K.S., Holen, J. T., and Kelso, D.M. (1981) *CLIN. CHEM.* **27**, 1575.

This technique works best if a small molecule is fused to a fluorophore and binds to a larger partner[151] (this maximizes the difference in signal between bound and unbound states). If the fluorophore is attached to the larger protein in a binding pair, the difference in polarization between bound and unbound states will be smaller (because the unbound protein will already be fairly stable and tumble slowly to begin with) and the measurement will be less precise.By titrating the amount of one of the proteins, a binding curve can be generated (the amount of polarization observed is proportional to the amount of protein complex formed, which is proportional to the concentration of the binding partners in solution). Mathematical models can be applied to this binding curve to determine the binding constant of the protein interaction[152]. In another application of this technique, it is also possible to measure the folding of a protein, since an unfolded peptide chain will tumble differently than a folded one, giving a difference in polarization

151 A. Fowler, D. Swift, E. Longman, and M. Coldwell (2002) *Analytical Biochemistry,* **308**, 223.
151 Hess Kenny, C., Kriz, R., and Ellestada, G. (2003) *Analytical Biochemistry,* **323**, 224.

3.6.1 *Measurement of eIF4E-affinity of eIF4E binding protein peptide analogs.*

In a first experiment three FP assays have been done on the linear unmodified peptide **XVI**, linear modified peptide **XVII** and cyclic peptide **XVIIc**. We tested the affinity for eIF4E, applying a 10nM concentration of those peptides to all the plate positions and using an increasing concentration of protein. The results showed that linear peptide **XVI** (4GII-FITC), unmodified peptide coming from the sequence of an eIF4E b.p. (4GII), and his modified linear analog (peptide **XVII** with two residues modified with alkynyl and azido groups), has a similar affinity (ca. 150nM) towards eIF4E (Figure 3.6–2). These remarkable findings suggested that the introduction into a peptide sequence of non-coded, alkynyl and azido amino acids doesn't affect its biological properties.

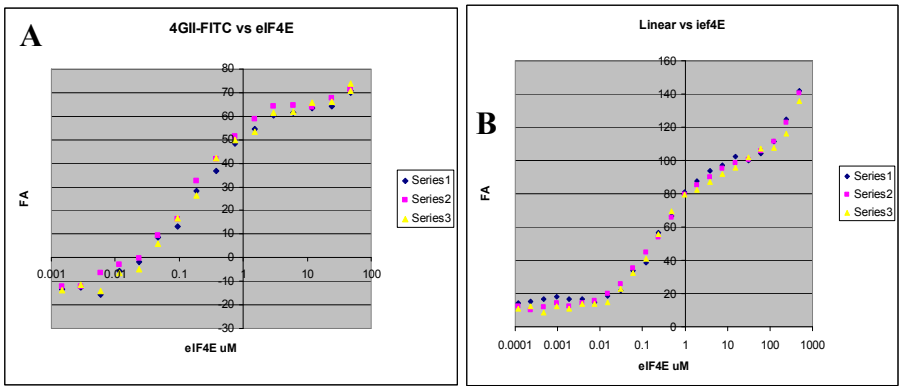


Figure 3.6–2 Fluorescence polarization assay for the measure of binding affinity towards eIF4E of 4GII-FITC (unmodified linear peptide **XVI)** (**A**) and linear modified peptide **XVII** (**B**).

On the other hand, the clicked, cyclic form of peptide **XVII** (peptide **XVIIc**) had a very poor affinity (even with high concentration of protein, 500uM, the curve didn't reach saturation) (Figure 3.6–3). These data imply that the triazole bridge formation completely suppress the ability of the sequence to fit the

proteins binding site. This suggested that the bridge was set in a non convenient position and, as can be seen in Figure 3.6–3, where is shown the structure of the complex eIF4E/eIF4E binding protein peptide, at the bottom of the peptide, because of the introduction of the triazole bridge, a new helix turn was formed where originally there was a coil. Besides, the Tyrosine at position 4, which is the most important residue of the motif, is moved from its correct position when the bridge is formed.

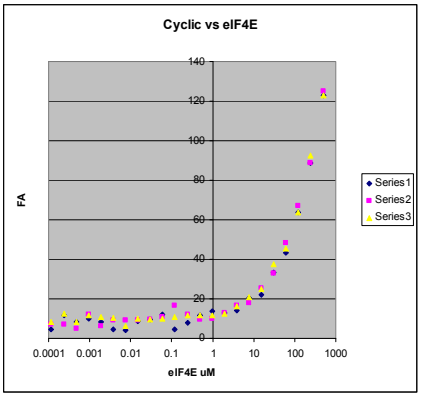


Figure 3.6–3 Fluorescence polarization assay for the measure of binding affinity towards eIF4E of peptide **XVIIc** (clicked, cyclic form of linear modified peptide **XVII**).

Starting from these results we planned the synthesis of new modified analogs of 4GII-FITC with the bridge stabilizing the helix part of the structure between positions 7 and 11.

We synthesized a small collection of cyclic, clicked peptides, generated by click chemistry reaction with different ring sizes (section 3.5).

In Figure 3.6–4 are summarized the Fluorescence Polarization measurements of 4E-BP linear peptide analogs **XVII-XX** and the corresponding cyclic peptides **XVIIc-XXc** against the eIF4E protein.

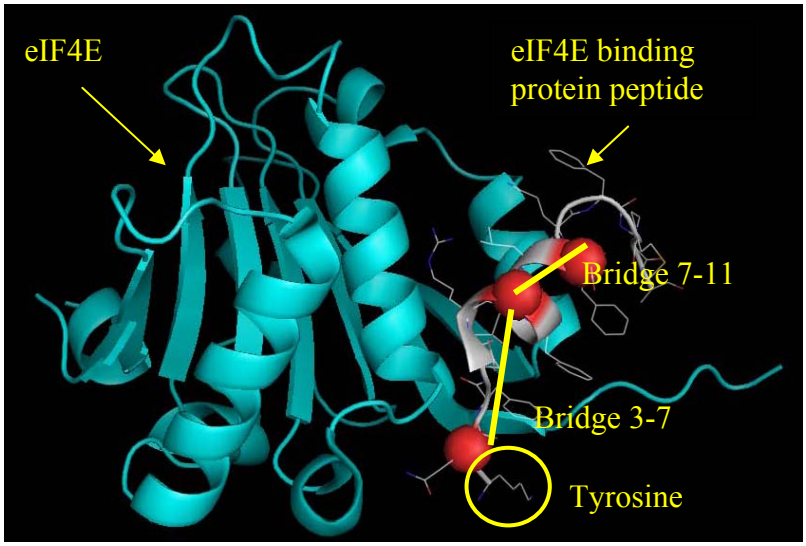


Figure 3.6–3 Structures of the complex eIF4E/eIF4E binging protein (from x-ray analysis). In evidence the position of the triazole bridge 3-7 in peptide XVIIc (between residue 3 and 7) and the tyrosine residue in position 4 that is destabilized by the bridge formation.

The titration curves show that the new cyclic analogs (**XVIIIc**, **XIXc**, and **XXc**) has a far better affinity for eIF4E than the first developed cyclic peptide (**XVIIc**), around 15 nM of k_d versus no affinity. This implies that the decision to shift the triazole bridge position from residues 3-7 to residues 7-11 was the correct one. The fact that the new cyclic peptides bind consistently one order of magnitude less tightly (k_d ca. 150 nM) than the linear counterparts.

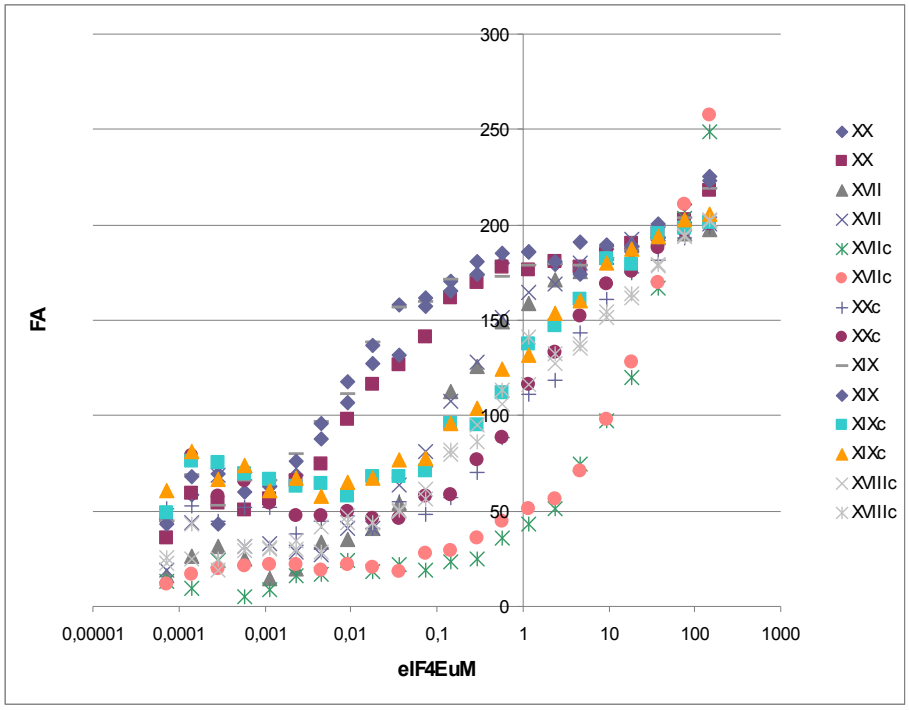


Figure 3.6–4 Fluorescence Polarization mesurments of 4E-BP linear peptide analogs **XVII-XX** and the corresponding cyclic peptides **XVIIc-XXc** against eIF4E.

The latter conclusion means that under the chosen ring size parameters (a 2+3/3+2 triazole bridge) the cyclization does not improve affinity over the linear peptides. However the fact that different cyclic analogs have different affinity towards the protein is a proof that our hypothesis of side-chain to side-chain triazole bridge α-helix stabilization is indeed correct.

As future developments, it remains to find which is the best combination of modified alkyne and azido amino acids to get the best α-helical stabilization.

3.7 Conformational studies

In solution, most of the peptides assume multiple flexible conformations. Determination of the dominant conformers and evaluation of their populations is the aim of peptide conformational studies, in which theoretical and experimental methods play complementary roles.[153] Cyclic peptides typically assume multiple conformations; these conformations are rather flexible, with torsional angles of the backbone (φ, ψ) as well as of the side chain groups (χ_i) fluctuating within large intervals (Figure 3.7–1).

In addition, coupling constants between N*H* and C$^{\alpha}$*H* can give information about the average values of the peptide backbone torsional φ angles.

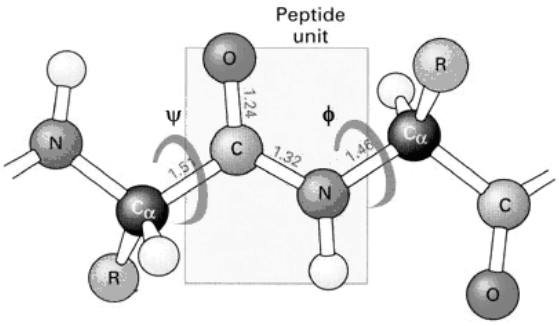


Figure 3.7–1 φ and ψ Dihedral angles.

Moreover, another goal of conformational investigations is to determine the relationship between conformation and activity of biologically important peptides (e.g. SAR: structure-activity relationships studies). Numerous biological results are strongly supported by conformational investigations, clearly indicating that biological peptide activity is determined not only by the

153 Bierzyñski, A. (2001) *Acta Biochimica Polonica*, **48**, 1091..

presence of specific functions binding to a target protein, but also dramatically depends on the conformational properties of the whole peptide structure.

Cyclopeptides **XIXc** (**C3+N2**) and **XXc** (**N3+C2**) (Figure 3.7–2) were analyzed by NMR and CD, to understand the influence of the alkyl ring size containing the triazolyl moiety, the spatial orientation of triazole ring and the spatial arrangement of all the side chains that is fundamental for the interactions with the eIF4E protein.

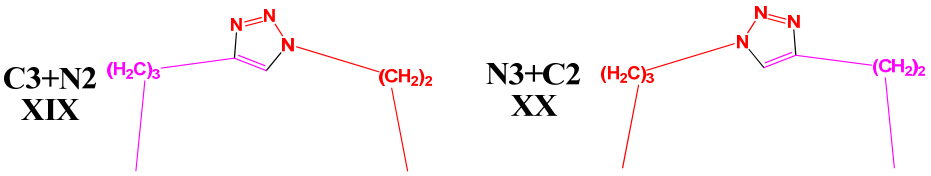


Figure 3.7–2 Triazole bridge orientation in eIF4E b.p. peptide analogs **XIXc** and **XXc** which are isomers having a different triazole orientation.

3.7.1 *NMR and CD structural characterization of cyclopeptides XIXc and XXc*

This part of research was followed in cooperation with Prof. A. M. D'Ursi and M. Scrima of the department of Pharmaceutical Sciences of the University of Salerno.

A preliminary screening of the conformational preferences of **XIXc** and **XXc** [1,2,3]triazolyl-containing peptides as a function of the environment was performed by means of circular dichroism (CD) spectroscopy. CD is a form of spectroscopy based on the differential absorption of left- and right-handed circularly polarized light. It can be used to determine the structure of macromolecules, including the secondary structure of proteins. In fact, CD spectroscopy is widely used to determine the secondary structures of proteins. Within the UV region from 180 up to 240 nm each of the structures α-helices, β-sheets, and the remaining, unordered part of the polypeptide backbone, usually referred as "random coil", contribute in different ways to the peptide spectrum.

CD spectra of both **XIXc** and **XXc** recorded in water (pH=5) presented a negative band at 201 nm. CD spectra (Figure 3.7–3 and Figure 3.7–4) recorded in water and water/HFA (50:50, v/v) mixtures[154] (a secondary structure stabilizing solvent), show for **XXc** the negative bands at 208 and 222 nm and an additional positive band at 192 nm, for **XIXc** presented two minima around 220nm. A single value deconvolution method[154] estimates for both cyclopeptides in water almost 80% of random coil structures, for cyclopeptides in water/HFA high amounts of alfa-helical structures for **XXc** and significant amounts of beta-structures for **XIXc**.

154 Johnson, J.W. (1990) In *Protein secondary structure and circular dichroism: a practical guide*, 205.

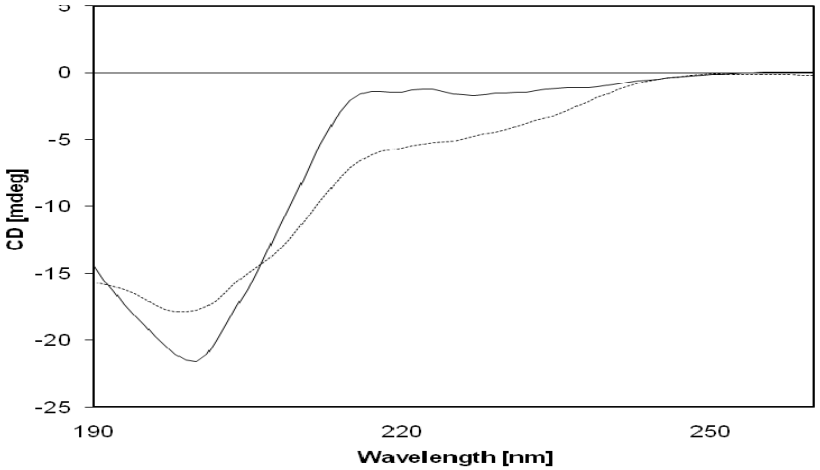


Figure 3.7–3 CD spectra of **XXc** (Continue line) and **XIXc** (dotted line) in H_2O pH 5.

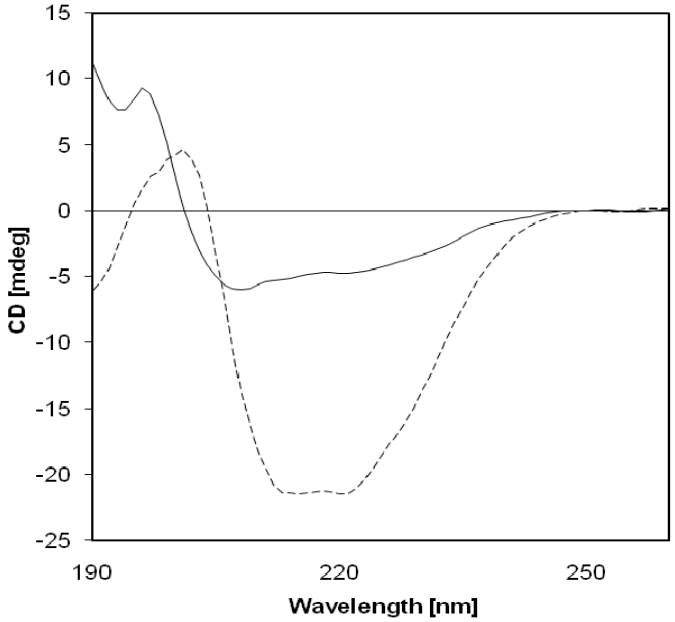


Figure 3.7–4 CD spectra of **XXc** (Continue line) and **XIXc** (dotted line) in H_2O/HFA 50/50 v:v.

These preliminary results require further detailed analyses, nevertheless, they prelminarly suggest a different conformational behavior of **XIXc** and **XXc** which on chemical point of view are different only for the orientation of the triazolyl ring.

A preliminary NMR analysis was undertaken in water solution. NMR spectra of sc152 were acquired in water solution at pH 5.5. To exclude potential aggregation, we recorded the 1D proton spectra of the cyclopeptides at a concentration range spanning 1-0.1 mM. At a peptide concentration of 0.1 mM, the peptides did not display any noticeable effect of aggregation. Chemical shift assignments of the proton spectra of XXc (Table 3.7-1) were achieved via the standard systematic application of DQF-COSY,155 TOCSY,156 and NOESY157 experiments, using the SPARKY158 software package according to the procedure of Wüthrich.159

It is well known that water solutions are considered the most biocompatible media suitable for NMR analysis of biomolecules. However water solutions enhance the flexibility of short peptides and avoid the collection of a sufficient NMR data to build reliable 3D models. To overcome this proble usually mixtures of water and organic solvents are used. The preliminary screening of the conformational analysis in water solution is advisable.

In our case the preliminary NMR investigation of XXc strted with the acquisition of NMR experiments in water solutions. Figure 3.7–5 shows the low-field region of NOESY spectrum of XXc in water solution. In agreement with the expectations for a spectrum in water solution a high number of TOCSY and NOESY correlations is not observable. In any case a careful

155 Piantini, U., Sorensen, O.W., Ernst, R.R., (1982) *J. Am. Chem. Soc.*, **104**, 6800.
156 Bax, A. and Davis, D.G. (1985) *J. Magn. Res.*, **63**, 207.
157 Jeener, J., Meyer, B.H., Bachman, P., and Ernst, R.R. (1979) *J. Chem. Phys.*, **71**, 4546.
158 Goddard, T.D., Kneller, D.G., SPARKY 3 NMR software. University of California, San Francisco (2001).
159 Wüthrich, K. *NMR of Proteins and Nucleic Acids*. John Wiley and Sons: New York, (1986), 44.

analysis of the correlations present, allowed the assignment of a high percent of proton resonances (Table 3.7-1).

The amide and fingerprint regions of NOESY spectra of peptide XXc in water (Figure 3.7–5) showed a significant number of non-trivial dNN(i, i+1) dαN(i, i+3) medium-range NOE correlations. These correlations suggest that in spite of the flexibility of the peptide and the polarity of the aqueous environment, XXc has an intrinsic tendency to assume ordered turn-helical structures.

Cyclopeptide XXc in Water							
Residue	**HN**	**$C^{\alpha}H$**	**$C^{\beta}H$**	**$C^{\gamma}H$**	**$C^{\delta}H$**	**$C^{\varepsilon}H$**	**Others**
Lys1	8.105	4.673	1.879/1.624	1.259	1.532	2.814	Qζ7.759
Lys2	8.167	4.144	1.756/1.532	1.226	1.449	2.782	Qζ7.637
Gln3	8.107	4.112	2.831	2.411			
Tyr4	8.278	4.546	2.740/2.607		7.204	7.123	
Asp5	7.911	4.682	2.569				
Arg6	8.082	4.679	1.412	1.645		2.952	Qζ7.755
Orn7	7.847	4.387	1.653	1.705	3.694		
Phe8	8.395	4.362	2.828		6.490	7.165	Hζ6.884
Leu9	8.468	4.451	1.760	1.490	0.797/0.724		
Leu10	8.454	4.672	1.499	1.396	0.752/0.700		
Orn11	7.751	4.391	1.602	1.982			
Phe12	8.262	4.724	2.946		7.290	7.400	Hζ7.270
Gln13	7.922	4.451	3.399	3.029			
Phe14	8.261	4.675	2.821		7.254	7.277	Hζ 7.269
Lys15	8.399	4.100	1.855/1.721	1.278	1.516	2.836	7.751

Table 3.7-1 Chemical shifts of **XXc** in water solution pH 5.5 (600MHz, 300K).

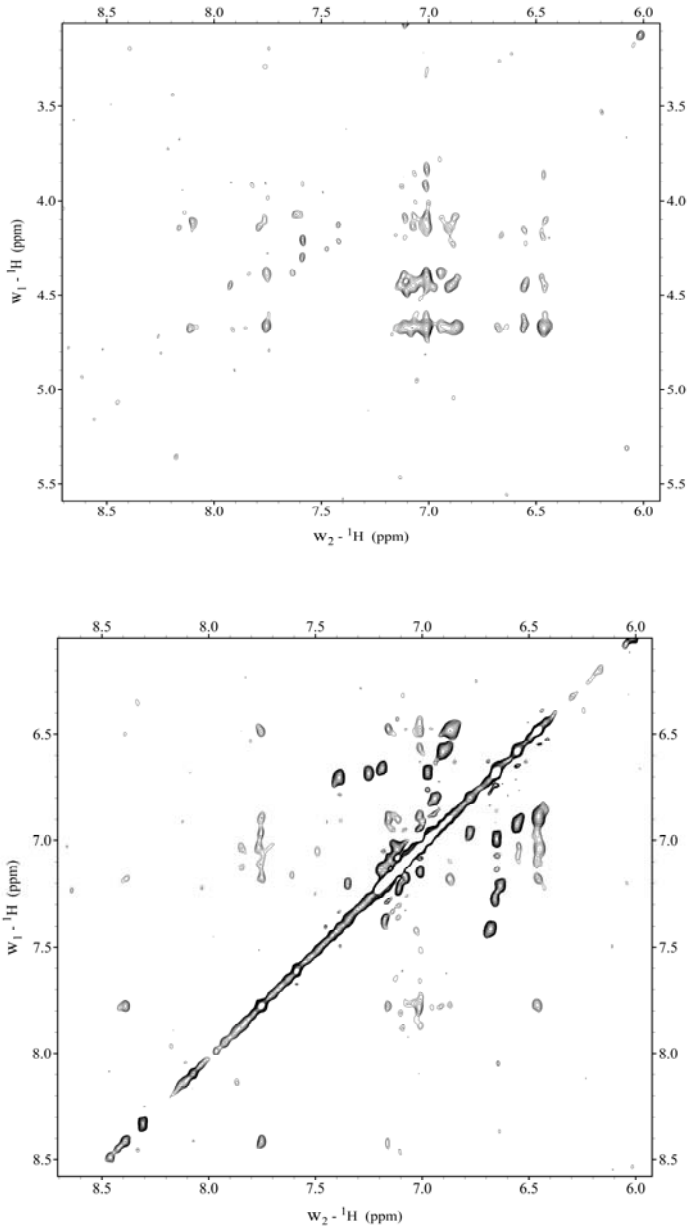


Figure 3.7–5 CH-α (top) and NH-NH (bottom) regions of NOESY spectrum of **XXc** in water pH. 5.5. The spectra were recorded at 600MhH and 300K.

3.8 Conclusions and future developments B

The goal of this project was the development of new eIF4E binding protein peptide inhibitors with stabilized conformation to be used as suppressors of Translation Initiation in a context of cancer therapy (section 3.3).

A collection of N^{α}-Fmoc protected unnatural amino acids bearing on the side chain azide or alkynyl functions was synthesized and introduced by SPPS in the fluoresceinated-4E-BP(621-636) peptide sequence to afford by Cu(I) catalyzed Huisgen reaction, a new collection of cyclopeptides containing the triazolyl moiety (section 3.5). The side-chain-to-side chain cyclization of linear peptides generated *via* click chemistry lead to cyclopeptides containing the triazolyl moiety linked to the α-carbon of the amino acids by alkyl chains of different lengths.

The collection of linear and cyclic peptides has been tested with Fluorescence Polarization Assay to measure the affinity for the eIF4E protein. The position and length of the triazole-bridge appears to play a critical role in enhancing and decreasing the affinity of the binding protein peptide analogs (4E-BP) for the target protein (eIF4E) (section 3.6.1).

A NMR/CD conformational study has been carried out on the cyclopeptides analogs containing triazolyl moiety. From preliminary data the triazole bridge seems, as expected, to stabilize the α-helical structure of the peptides (section 3.7.1).

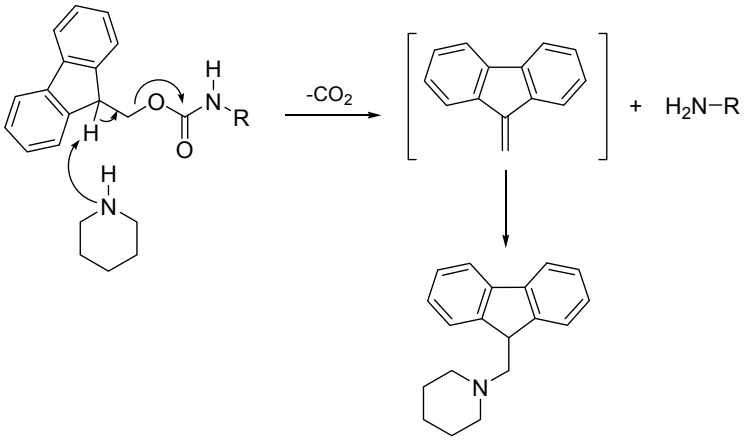
The completion of the synthesis and characterization of the 4E-BP cyclopeptides collection is in progress. The peptides presenting the best affinity with eIF4E will be selected for experiments of templated click cyclization free from metal catalyst and in presence of the protein (eIF4E) (3.5.2.4). The cell membrane permeability of the fluoresceinated peptides will also be assessed with a confocal microscope.

4 PART C: The Fmoc/tBu Solid Phase Peptide Synthesis

The following section contains material coming from the book chapter "Orthogonal protecting group in Fmoc/tBu strategy" edited by the author together with Prof. Anna Maria Papini under press in "Amino, Peptides, and Proteins in Organic Chemistry", Wiley.

4.1 The Fmoc/tBu-strategy

The Fmoc/tBu combination is an orthogonal system that was described in 1978[160], a decade and a half after Boc/Bzl chemistry, and it nowadays, by far, the most widely used strategy for the SPPS. It is based on the Nα-protecting group 9-fluorenylmethoxycarbonyl developed by Carpino and coworkers in 1970[161]. Nα-Deprotection is mediated by a base, usually piperidine (Scheme 4.1-1), so the amino group is available for the next coupling as soon as the protector is removed, and there is no loss of side-chain protectors during chain assembly[162].



Scheme 4.1-1 Piperidine mediated Fmoc deprotection

160 Chang, C.D. and Meienhofer, J. (1978) *International Journal of peptide and Protein research,* 539.
161 Carpino, L.A. and Han, G.Y. (1970) *Journal of the American Chemical Society*, **92**, 5748.
162 Benoiton, N.L. (2006) In *Chemistry of Peptide Synthesis*, Taylor and Frensis, 142.

Final deprotection is by acid of moderate strength, so no special equipment is necessary (contrary of what is required for Boc/Bzl strategy), and the linker need not be especially stable to acid. Fmoc–amino acids cost more than Boc–amino acids, but their use involves one step less (since the acid employed for N^{α}-Boc deprotection leaves the amino group protonated, a neutralization step is required to convert the amino group to a nucleophile), thus reducing consumption of solvent, which is a significant cost savings[163].

163 Atherton, E., Logan, C.J., and Sheppard, R.C. (1981) *Journal of the Chemical Society Perkin Transactions 1,* 538.

4.1.1 *Orthogonal protecting groups in Fmoc/tBu strategy*

4.1.1.1 Arginine

During SPPS the three nitrogen atoms of the guanidine group of Arginine (Figure 4.1–1) being strongly nucleophilic are prone to alkylation and sub sequential Ornitine formation upon base–mediated decomposition[164] and therefore need to be protected. However, in common practice, most protecting groups block only the ω-nitrogen. In addition, free unprotected Arginine residues tend to cyclize upon activation of the α-carboxylic group to form δ-lactams

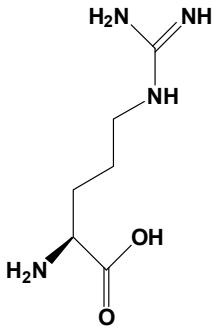
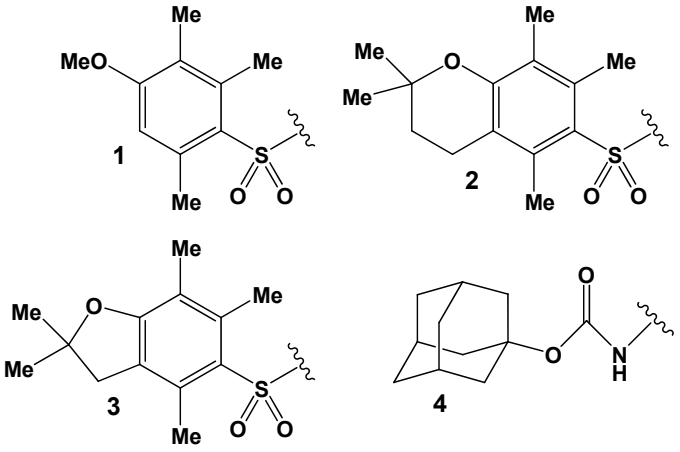


Figure 4.1–1 The guanidino group of Arginine.

In Fmoc/tBu strategy the most commonly used protecting groups of Arginine are the arylsulfonyl-based derived from the tosyl group (Tos)[165] such as 4-methoxy-2,3,6-trimethylbenzenesulfonyl (Mtr) group **1**, now superseded by the two cyclic ether derivates 2,2,5,7,8-pentamethylchroman-6-sulfonyl (Pmc) **2** and 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf) **3** group. Mtr removal requires several hours of TFA treatment and often causes sulfonation of Trp residues[166], which can be avoided using 1M TMSBr in TFA[167].

164 Rink, H., Sieber, P., and Raschdorf, F. (1984) *Tetrahedron Letters*, **25**, 621.
165 Ramage, R., Green, J., and Blake, A.J. (1991) *Tetrahedron,* **47**, 6353.
166 Sieber, P. (1987) *Tetrahedron Letters,* **28**, 1637.
167 Fujii, N., Otaka, A., Sugiyama, N., Hatano, M., and Yajima, H. (1987) *Chemical and Pharmaceutical Bulletin*, **35**, 3880.

Moreover long TFA treatment can cause O-sulfonation of Ser and Thr which can be suppressed adding thiocresol to the cleavage cocktail[168]. Pmc group being much acid sensitive than Mtr can be removed faster and the Trp/Tyr modifications are less pronounced[169] and can be overcome if the Trp indole ring is Boc protected[170]. The Pbf group[171], the dihydrofuran analog of Pmc group, is at present the most widely used Arginine protecting group. It has proved to be more acid labile than Pmc (its removal is 1-2 time faster) and generates less alkylation than the other arylsulfonyl-protecting groups.

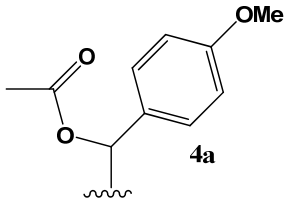


Other reported protecting strategies have not gained popularity due to several different shortcomings. For example, the nitro group (NO) is used to protect the ω-Nitrogen and can be removed with H_2/Pd or with hydrazinium monoformate and magnesium[172] but it is not completely stable under coupling conditions[173]. It generally requires long deprotection times and it is prone to

168 Jaeger, E., Remmer, H.A., Jung, G., Metzger, J., Oberthur, W., Rucknagel, K.P., Schafer, W., Sonnenbichler, J., and Zetl, I. (1993) *Biological Chemistry Hoppe Seyler*, **5**, 349.
169 Green, J., Ogunjobi, O.M., Ramage, R., and Stewart, A.S. (1988) *Tetrahedron Letters,* **29**, 4341.
170 White, P. (1992) in *Peptides, Chemistry, Structure and Biology, Pro. 12th American peptide Symposium*, Smith, J.A. and Rivier, J.E. (Eds.), ESCOM, Leiden, 537.
171 Carpino, L.A., Shroff, H., Triolo, S., Mansour, E.M., Wenschuh, H., and Albericio, F. (1993) *Tetrahedron Letters*, **34**, 7829.
172 Gowda, D.C. (2002) *Tetrahedron Letters*, **43**, 311.
173 Wunsch, E. (1974) In *Houben-Weyls methods der Organischen Chemie*, Muller, E. (Ed.), **15**, Parts 1 and2, Thieme, Stuttgart.

generate partially reduced products[174]. The urethane protections, ω-Boc[175] and δ, ω-bis adamantyloxycarbonyl $(Adoc)_2$ [176] **4** are unable to completely suppress the guanidine group nucleophilicity. On the other hand bis-urethane Arginine derivate with both ω and ω'-Nitrogen Boc protected[177] does not show any side reaction on deprotection but it is highly hindered adduct and the coupling time need to be extended. The trityl group is not normally used for Arginine protection because it gives an adduct with poor solubility in DMF/DCM[178].

A completely different approach to the problem of Arg protection is the use of a suitably protected Ornitine residue that can be converted into Arginine at the end of the synthesis by guanylation with reagents such as 1H-pyrazole-1-carboxamide hydrochloride[179]. An additional level of orthogonality is needed for the Ornitine δ-amino group, which has to be selectively deprotected before the conversion. Protecting groups such as the 1-(4-methoxyphenyl)ethyloxycarbonyl (Mpeoc) [180] **4a**, cleavable under mild acidic conditions, have been specifically developed for this application.



4a

174 Turàn, A., Patthy, A., and Bayusz, S. (1975), *Acta Chimica Academiae Scientiarum*, **85**, 327.

175 Gronvald, F.C., Johansen, N.L., and Lundt, F.G. (1981), In *Peptides 1980*, Brunfeldt K. ed, Scriptor. Copenhagen, 111.

176 Presentini, A. and Antonui, G. (1986), *International Journal of Peptide and Protein Research*, **27**, 123.

177 Verdini, A., Lucietto, P., Fossati, G., and Giordani, C. (1992) *Tetrahedron Letters,* **33**, 6541.

178 Caciagli, V. and Verdini, A.S. (1988) In *Peptide Chemistry 1987,* (Shiba T. and Sakakibara S. eds.), Protein Research Foundation Osaka, 283.

179 Bernatowicz, M.S., Wu, Y., and Matsueda, G.R. (1992), *The Journal of Organic Chemistry*, **57**, 2497.

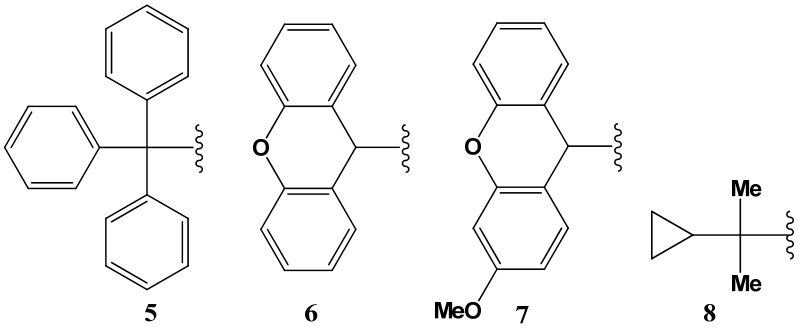
180 Bernatowicz, M.S. and Matsueda, G.R. (1994) In *Peptides, Chemistry, structure and Biology, proc. 13th American Peptide Symposium*, Hodges, R.S. and Smith, J.A. (Eds.) ESCOM, Leiden, 107.

4.1.1.2 Asparagine and Glutamine

Asparagine and Glutamine could in principle be incorporated into peptides without protection but the unprotected derivates display a low solubility in solvents commonly used in peptide synthesis and have low coupling rates. In addition the amide side chain, especially Asparagine's one, can suffer partial dehydratation on activation[181]. Finally, although more common for the Boc/Bzl chemistry, Glutamine in the N-terminal position can undergo weak acid catalized cyclization forming pyroglutamyl residues that cause the truncation of the peptide sequence[182].

Side chain protection prevents all this undesired reactions and, in addition, inhibits hydrogen bond interactions of the amide, which stabilizes secondary structures causing incomplete deprotection and reduced coupling rate.

The most common amide protecting group for Fmoc/tBu SPPS is the triphenylmethyl group (Trt)[183] **5**, which requires care in the choice of the cleavage scavengers since it generates stable carbocations that tend to alkylate Tryptophan. Such alkylation is reduced by using the 9-xanthenyl (Xan) group **6** and its 2-methoxy derivate (2-Moxan)[184] **7** but, on the other hand, it generates less soluble derivates.



5 **6** **7** **8**

181 Gausepohl, H., Kraft, M., and Frank, R. (1989) *International Journal of Peptide and Protein Research,* **34**, 287.
182 Dimarchi, R., Tam, J., and Merrifield, R. (1982) *Journal of Peptide Research*, **19**, 88.
183 Sieber, P. and Riniker, B. (1991) *Tetrahedron Letters*, **32**, 739.
184 Han, Y., Solè, N., Tejibrant, J., and Barany, G. (1996) *Journal of Peptide Research*, **9**, 166.

When Asn is the N-terminal amino acid, trityl deprotection it is slower, due to the vicinity of the free α-amino group[185]. In this case cleavage time needs to be extended or methyltrityl protection can be used instead of Trt.

N-dimethylcyclopropylmethyl (Dcmp) group **8** represents a convenient alternative to Trt with several advantages like rapid removal (even at the N-terminal position), faster coupling rate (due to the minor steric hindrance) and better solubility in DMF[186].

To overcome the problem of the slow and troublesome attachment to most resins the resin can be linked to the amide side chain[187] (Figure 4.1–2) instead to the carboxylic function (that needs to be protected during the synthesis)

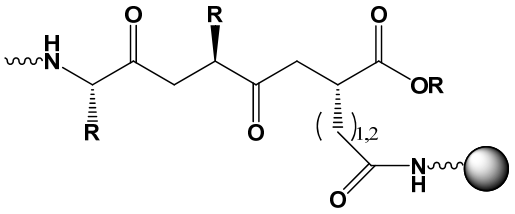
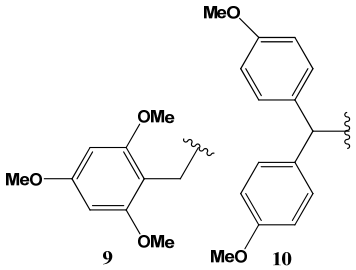


Figure 4.1–2 Side chain anchoring of C-terminal Asn and Gln containing peptides

More used in the past than now, before the Trt protection gained popularity, are 2,4,6-trimethoxybenzyl group (Tmob)[188] **9** and 4,4'-dimethoxybenzhydryl group (Mbh)[189] **10** both less acid-labile and soluble than Trt[183].



185 Friede, M., Denery, S., Neimark, J., Kieffer, S., Gausepohl, H., and Briand, J. (1992) *Journal of Peptide Research*, **5**, 145.

186 Carpino, L.A., Chao, H., Ghassemi, S., and Mansour, E.M. (1995) *The Journal of Organic Chemistry*, **60**, 7718.

187 Greipohl, G., Knolle, J., and Stuber, W. (1990) *International Journal of Peptide and Protein Research*, **35**, 281.

188 Weyand, F., Steglich, W., Bjarnason, J., Ahktar, R., and Chytil, N. (1968) *Chemische Berichte*, **101**, 3623.

189 Konig, W. and Geiger, R. (1970) *Chemische Berichte*, **103**, 2041.

4.1.1.3 Aspartic and Glutamic acid

The carboxylic acid side chains of Asp and Glu need to be protected during peptide synthesis in order to prevent amide bond formation with incoming amino acids and, as a consequence, branching of the peptide[190].

Although protected Asp and Glu residues could still be affected from side reactions, particularly acid or base (particularly in Fmoc/tBu-chemistry) catalized cyclization to form aspartimides and glutarimides respectively[191]. Subsequent hydrolysis of the imide-containing peptides lead to a mixture of the desired peptide along with a product, called β-peptide in which the side chain carboxylic group forms part of the backbone and a β-piperidide adduct[192].

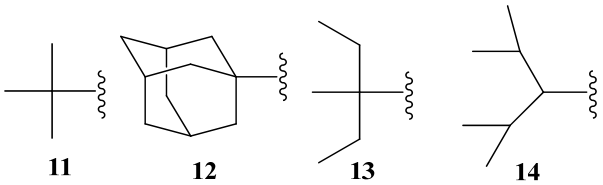
The reaction occurs less often with Glu[193] and is highly sequence dependent. Susceptible sequences are Asp-Xxx with Xxx being Gly, Asn, Ala, and Gln[194].

Since the beginning of Fmoc/tBu SPPS, Asp and Glu have been successfully protected by t-butyl (tBu) **11** group, which is base stable and TFA labile[288].

Equally favourable properties are displayed by 1-adamantyl (1-Ada) protection[195] **12**. Both tBu and 1-Ada minimize piperidine-catalized aspartimide formation although several bulky tBu derivates give better results in this respect like the 3-methylpent-3-yl (Mpe)[196] group **13** and β-2,4-dimethyl-3-penty (Dmp)[197] group **14**.

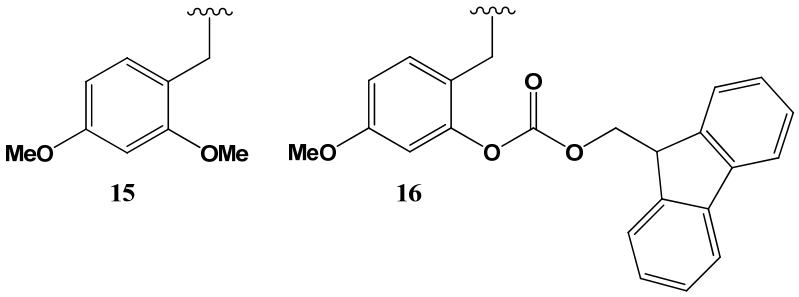
---

190 Natarajan, S. and Bodanszky, M. (1976) *The Journal of Organic Chemistry*, **41**,1269.
191 Tam, J., Riemen, M., and Merrifield, R. (1988) *Journal of Peptide Research*, **1**, 6.
192 Dolling, R., Beyermann, M., Haenel, J., Kernchen, F., Krause, E., Franke, P., Brudel, M., and Bienert, M. (1994) *Journal of the Chemical Society, Chemical Communications*, 853.
193 Kates, S. and Albericio, F. (1994) *Letters in Peptide Science*, **1**, 213.
194 Yang, Y., Sweeney, V., Schneider, K., Thornqvist, S., Chait, B., and Tam, J. (1994) *Tetrahedron Letters*, **35**, 9689.
195 Okada, Y., Igushi, S., and Kawasaki, K., (1987) *Journal of the Chemical Society, Chemical Communications*, 1532.
196 Karlstrom, A. and Unden, A. (1996) *Tetrahedron Letters*, **37**, 4243.
197 Karlstrom, A. and Unden, A. (1995) *Tetrahedron Letters*, **36**, 3909.



11 **12** **13** **14**

Aspartimide formation is also greatly reduced by adding either 1-hydroxybenzotriazole (HOBt) or 2,4-dinitrophenol to the piperidine deprotection solution[198] but can be completely eliminated, especially with susceptible sequences in long peptide sequences, only by employing amide-backbone protection for the introduction of residues preceding Asp. This is achieved using the 2-hydroxy-4-methoxybenzyl (Hmb)[199] **15** or the 2,4-dimethoxybenzyl (Dmb)[200] **16** (only applied to glycine residues) amide protecting groups. Hmb and Dmb, removed contemporary with the final cleavage, prevent undesired side reactions and suppress aggregation during chain extension.

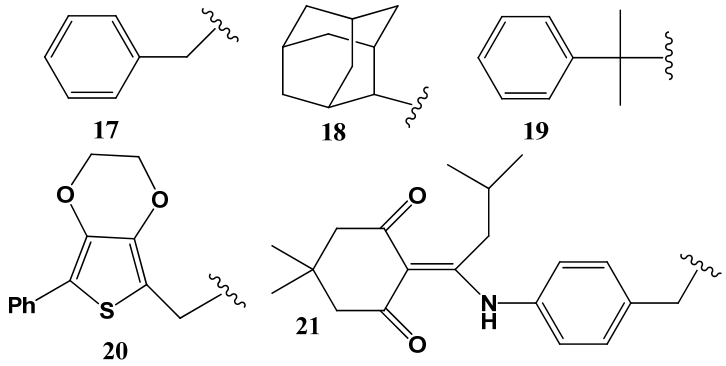


15 **16**

When an additional degree of orthogonality is required, as for example in the case of lactam-bridged peptide synthesis, a number of different Asp and Glu protecting groups exist. For example the benzyl group (Bzl) [201] **17**, the 2-amantyl (2-Ada) [202] **18**, and allyl eters[277] (thought more prone to imide

198 Martinez, J. and Bodansky, M. (1978) *International Journal of Peptide and Protein Research*, **12**, 277.
199 Quibell, M., Owen, D., Packmann, L., and Johnson, T. (1994) *Journal of the Chemical Society, Chemical Communications*, 2343.
200 Zahariev, S., Guarnaccia, C., Pongor, C.I., Quaroni, L., Cemazarc, M. and Pongora, S. (2006) *Tetrahedron Letters,* **47**, 4121.
201 Benoiton, L. (1962) *Canadian Journal of Chemistry*, **40**, 570.
202 Okada, Y. and Igushi, S., (1988) *Journal of the Chemical Society, Perkin Transactions I*, 2129.

formation than tBu) all removed by palladium-catalized transfer to a suitable nucleophile. Alternatively super-acid-labile groups can be used such as the 2-phenyl isopropyl (Pp)[203] group **19**, removable in presence of tBu/Boc with 1% TFA in DCM and the phenyl-3,4-ethylenedioxy-2-thenyl (EDOT-Ph)[204] **20** cleaved by 0.1-0.5% TFA in DCM.



Another orthogonal protecting group is 4-[N-[1-(4,4'-dimethyl-2,6-dioxocyclohexylidene)-3-methylbutyl]aminobenzyl group (Dmab)[205] **21**, removed with 2% hydrazine in DMF. With aspartimide susceptible sequences it is recommended to use Dmab along with additional precautions such as backbone-amide protection[206].

For applications in native chemical ligation it has recently been described a new photo-labile protecting group, {7-[bis(carboxymethyl)amino]coumarin-4-yl}methyl (BCMACM)[207] group **22**, removed with UV irradiation at 405 nm.

203 Kunz, H., Waldmann, H., and Unverzagt, C. (1985) *International Journal of Peptide and Protein Research*, **26**, 493.
204 Isidro-Llobet, A., Alvarez, M., and Albericio, F. (2008) *Tetrahedron Letters*, **49**, 3304.
205 Chan, W., Bycroft, B., Evans, D., and White, P. (1995) *Journal of the Chemical Society, Chemical Communications*, 2209.
206 Nski, J.R., Lewandowska, B., Mucha, P., and Retowski, P. (2008) *Journal of Peptide Science*, **14**, 335.
207 Briand, B., Kotzur, N., Hagen, V., and Beyermann, M. (2008) *Tetrahedron Letters,* **49**, 85.

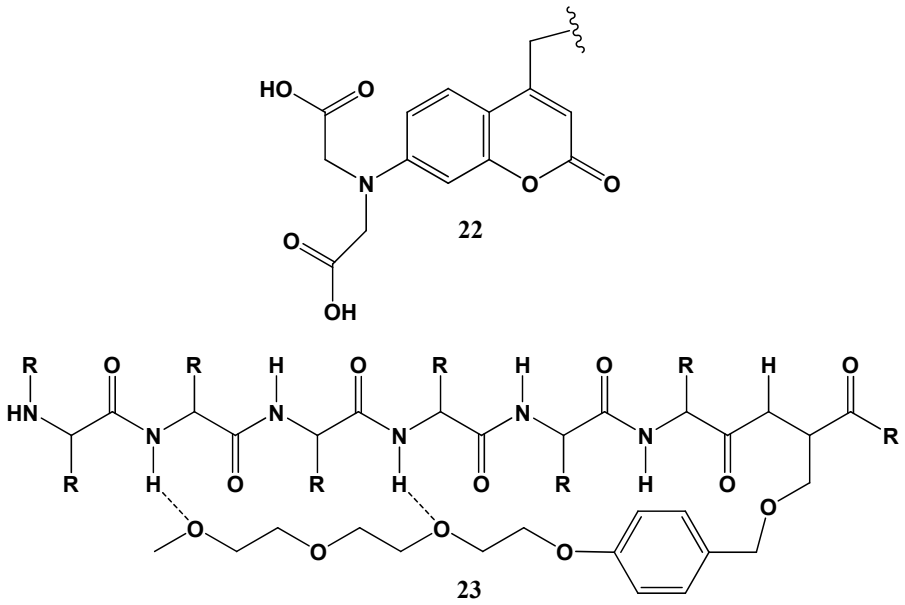


Figure 4.1–3 Internal solvation with the TEGBz group.

A special carboxylic protecting group, the 4-(3,6,9-trioxadecyl)oxybenzyl (TEGBz)[208] **23**, has been developed to suppress aggregation of those 'difficult sequences', in which intermediate resin-bound peptide chains associate into extended β-sheet type structures. TEGBz forms hydrogen bond with the backbone amino groups enabling the so called 'internal solvation' that inhibits aggregation by enhancing backbone linearity (Figure 4.1–3).

208 Kocsis, L., Bruckdorfer, T., and Orosz, G. (2008) *Tetrahedron Letters,* **49**, 7015.

4.1.1.4 Cysteine

In SPPS protection of the Cysteine side chain sulfhdryl group is mandatory otherwise it would easily undergo alkylation and acylation. Free Cysteine residues are also prone to oxidation, even by atmospheric oxygen, to form intra- and inter-molecular disulfide bonds.

The chemistry of Cysteine protecting groups is particularly rich due to the key importance of this amino acid in forming inter- and intra-molecular disulfide bridges and in consideration of the versatile reactivity of the thiol group.

With the Fmoc/tBu strategy trityl (Trt)[209] **24** is the most used Cysteine protection. Since acid S-detritylation is an equilibrium reaction it needs to be driven to completion by capture of the forming carbocation. This can be achieved by adding H_2O, thiols and especially silanes. Triisopropylsilane (TIS) in particular has to be preferred over triethylsilane[210] which can lead to reduction of free Trp residues. In spite of its popularity Cys(Trt) is prone to racemization (up to 10-20%) during peptide coupling by base mediated *in situ* activation[211], especially with the TBTU/DIPEA system. The use of HBTU, PyBOP/HOBt or preactivated reagents such as symmetrical anhydrides, OPfp esters and DIPCDI/HOBt minimize this problem[212]. Enantiomerization occurs also with the attachment of Cys(Trt) to Wang type resins and during chain extension when Cys(Trt) is the C-terminal residue. The use of Cl-trityl resins is recommended[213].

---

209 Fujii, N., Otaka, A., Funakoshi, S., Bessho, K., and Yajima, H., (1987) *Journal of the Chemical Society, Chemical Communications,* 163.

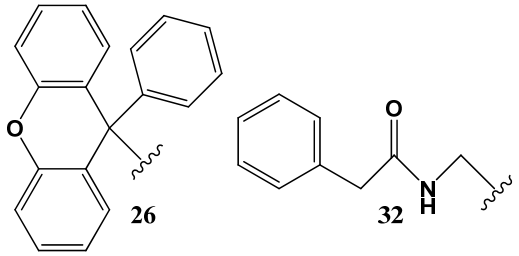
210 Pearson, D.A., Blanchette, M., Blaker, M.L., and Guindon, C.A., (1989) *Tetrahedron Letters*, **30**, 2739.

211 Kaiser, T., Nicholson, G.J., Kohlbau, H.J., and Voelter, W. (1996) *Tetrahedron Letters*, **37**, 1187.

212 Angell, Y.M., Alsina, J., Albericio, F., and Barany, G. (2002) *Journal of Peptide Research,* **5**, 292.

213 Fujiwara, Y. (1994) *Chemical and Pharmaceutical Bulletin*, **42**, 724.

In alternative to Trt, other protecting groups removed with concentrated TFA[214] are 1,4,6-trimethoxybenzyl (Tmob) **25** and 9-phenylxanthen-9-yl (pixyl) **26** whereas monomethoxytrityl (Mmt) **27**, 9H-xanthen-9-yl (xanthyl) **28** and 2-methoxy-9H-xanthen-9-yl (2-Moxan) **7** are more acid-labile[215] and can be selectively cleaved in presence of tBu groups.



26 **32**

The t-buthyl (tBu), S-(1-adamantyl) (1-Ada) **29**, acetamidomethyl (Acm) **30**, trimethylacetamidomethyl (Tacm) **31**, and phenylacetamidomethy (Phacm) **32** groups (last two developed to avoid formation of thiazolidine-2-carboxylic acid) are stable to acid and compatible with both Boc and Fmoc SPPS strategies. Those groups can be removed in several ways (see Table 4.1-1) enabling concomitant disulfide bridge formation even multiple, selective cyclization if used in combination[216]. Phacm has an additional level of orthogonality since it is enzimatically-cleavable by penicillin G acylase[217]

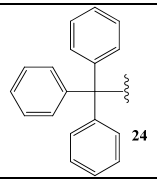
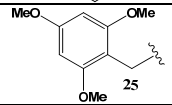
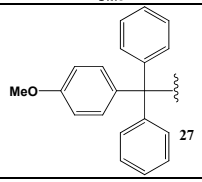
A different protecting approach is to use mixed disulfides such as the S-tert-butylmercapto (StBu) group **33**, which are stable to TFA and are removed with thiols[218] or tributylphoshine[219]. Coupling efficiency is reported to be highly

214 Munson, M.C., Garcia-Echevarria, C., Albericio, and F., Barany, G. (1992) *The Journal of Organic Chemistry,* **57**, 3013.
215 Han, Y. and Barany, G. (1997) *The Journal of Organic Chemistry*, **62**, 3841.
216 Albericio, F. (2000) In *Fmoc Solid Phase Peptide Synthesis* W. C. Chan and White, P.D. (Eds.) Oxford University press, oxford, 77.
217 Greiner, G. and Hermann, P. (1991) In *Peptides (1990)* Giralt E., Andreu D., Eds. ESCOM, Leiden, 277.
218 Weber, U. and Hartter, P. (1970) *Hoppe-Seyler's Zeitschrift fur Physiologische Chemie*, **351**, 1384.
219 Beekman, N.J., Schaaper, W.M., Tesser, G.I., Dalsgaard, K., Kamstrup, S., Langeveld. J.P., Boshuizen. R.S., and Meloen. R.H. (1997) *Journal of Peptide Research*, **50**, 357.

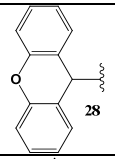
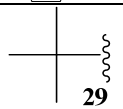
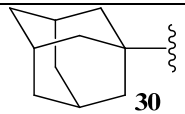
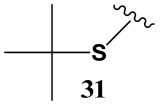
sequence dependent[220]. Allyl-based Cys protecting groups are base labile and therefore cannot be used in Fmoc/tBu-SPPS[221].

Finally, several non-conventional protecting groups have been developed for those Chemical Ligations based on Cysteine. The thiazolidine (Thz)[222] **34** protection has found a special application allowing to mask N-terminal Cysteines during Tandem Native Chemical Ligation (TNCL) reaction. It simultaneously protects the α-amino and the side chain thiol groups of protected N-terminal free Cysteines and is stable to acids and can be removed in aqueous conditions in presence of methoxylamine. The thiosulfonate group (S_2O_3)[223] can be used as well for Cysteine protection during TNCL. It is introduced with sodium tetrathionate ($Na_2S_4O_6$) in a solvent, which can then be removed by treatment with dithiothreitol (DTT).

220 Berangere, D. and Trifilieff, E. (2000) *Journal of Peptide Science*, **6**, 372.
221 Loffet, A. and Zhang, H.X. (1993) *International Journal of Peptide and Protein Research*, **42**, 346.
222 Bang, D. and Kent, S.B.H. (2004) *Angewandte Chemie International Edition*, 43, 2534.
223 Sato, T. and Aimoto, S., (2003) *Tetrahedron Letters,* **44**, 8085.

Protecting group	Structure	Removed by	Stable to
Trityl (Trt)	 24	Dil TFA/scavengers, Ag(I)[224], Hg(II)[225], RSCl, I_2[226], Tl(III) trifuoracetate[241] TFA/DMSO/Anisole[227]	Base, nucleophiles, RSH
1,4,6-trimethoxybenzyl (Tmob)	 25	Dil TFA/DCM/silanes[228] I_2[229], Tl(III) trifuoracetate	Base, nucleophiles
monomethoxytrityl (Mmt)	 27	Dil TFA/DCM/TIS[230], AcOH/TFE/DCM (1:2:7)[231], I_2[232]	Base, nucleophiles, RSH

224 Zervas, L. and Photaki, I. (1962) *Journal of the American Chemical Society*, **84**, 3887.
225 Photaki, I., Taylor-Papadimitriou, J., Sakarellos, C., Mazarakis, P. and Zervas, L. (1970) *Journal of the American Chemical Society*, 2683.
226 Kamber, B. and Rittel, W. (1968) *Helvetica Chimica Acta*, **51**, 2061.
227 Otaka, A., Koide, T., Shide, A. and Nobutaka, F. (1991) *Tetrahedron Letters*, **32**, 1223.
228 Munson, M.C., Garcìa-Echeverrìa, C., Albericio, F. and Barany, G. (1992) *The Journal of Organic Chemistry,* **57**, 3013.
229 Munson, M.C. and Barany, G. (1993) *Journal of the American Chemical Society*, **115**, 10203.
230 Pearson, D.A., Blanchette, M. and Baker, M.L. (1989) Guindon CA., *Tetrahedron Letters,* **30**, 2739.
231 Barlos, K. Gatos, D., Kallitsis, J., Papaphotiu, G., Sotitiu, P., Wenqing, Y., and Scafer, W. (1989) *Tetrahedron Letters*, **30**, 3943.
232 Barlos, K., Gatos, D., Kallitsis, J. Papaphotiu, G., Poulos, C., and Tsegenidis, T. (1991) *International Journal of Peptide and Protein Research*, **38**, 562.

9H-xanthen-9-yl (Xan)	 **28**	Dil TFA/TIS[233] I_2, Tl(III) trifuoracetate[234]	Base, nucleophiles
t-buthyl (tBu)	 **29**	TFMSA[235], TMSBr/TFA/RSH[236], tetrafluoroboric acid[237], Hg(II) acetate[238], Tl(III) trifuoracetate[241], $MeSiCl_3/Ph_2SO/TFA$[239], TFA/DMSO/Anisole	TFA, Ag(I), base, I_2, RSH
1-adamantyl (1-Ada)	 **30**	TfOH/TFA/RSH[240], Hg(II) acetate[238], Tl(III) trifuoracetate[241]	TFA, Ag(I), base, I_2,
S-tert-buthylsulfanyl (S*t*Bu),	 **31**	RSH[242], $NaBH_3$[243], phosphines[244]	TFA, TFMSA, base, RSCl

233 Han, Y. and Barany, G. (1997) *The Journal of Organic Chemistry*, **62**, 3841.
234 Hargittai, B. and Barany, G.J. (1999) *Journal of Peptide Research,* **54**, 468.
235 McCurdy, S. (1989) *Journal of Peptide Research*, **2**, 147.
236 Wang, H., Miao, Z., Lai, L., and Xu, X. (2000) *Synthetic Communications*, **30**, 727.
237 Akaji, K., Yoshida, M., Tatsumi, T., Kimura, T., Fujiwara, Y., and Kiso, Y. (1990) *Journal of the Chemical Society, Chemical Communications*, 288.
238 Atherton, E., Pinori, M., and Sheppard, R. (1985) *Journal of the Chemical Society, Perkin Transactions I*, 2057.
239 Akaji, K., Tatsumi, T., Yioshida, M., Kimura, T., Fujiwara, Y., and Kiso, Y. (1991) *Journal of the Chemical Society, Chemical Communications*, 167.
240 Fujii, N., Otaka, A., Funakoshi, S., Watanabe, T., Akaji, K., and Yajima, H. (1987) *Chemical and Pharmaceutical Bulletin*, **35**, 2339.
241 Yajima, H., Fujii, N., Funakoshi, S., Watanabe, T., Murayama, E., and Otaka, A. (1988) *Tetrahedron*, **44**, 805.
242 Threadgill, M. and Gledhill, A. (1989) *The Journal of Organic Chemistry*, **54**, 2940.
243 Wunsch, E. (1974) In *Houben-Weyl*, **15/1**, 789.
244 Moroder, J., Gemeiner, M., Gohring, W., Jaeger, E., and Wunsch, E. (1981) In *Peptides 1980* Brunfeldt K., Ed. Scriptor Copenhagen, 121.

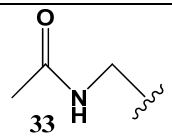
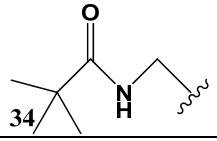
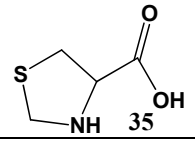
acetamidomethyl (Acm),	 33	Hg(II) acetate[245], Ag(I)/TFA[246], I_2[247], Tl(III) trifuoracetate[241], $AgBF_4$/TFA[248], AgTMS/DMSO-aq. HCl[249]	TFA, TFMSA, base, RSH
trimethylacetamidomethyl (Tacm)	 34	$AgBF_4$/TFA[248], Hg(II) acetate[250], I_2[251], Tl(III) trifuoracetate,	TFA, TFMSA, base RSH
Thiazolidine (Thz)	 35	Methoxylamine/H_2O[252]	TFA, base, RSH,

Table 4.1-1 Deprotection and deprotecti n/oxidation conditions of the most common Cys protecting groups (concomitant disuldife formation in green)

245 Sakakibara, S. (1995) *Biopolymers*, **37**, 17.
246 Fujii, N., Otaka, A., Watanabe, T., Okamachi, A. Tamamura, H., Yiajima, H., Inagaki, Y., Nomizu, M., and Asano, K. (1989) *Journal of the Chemical Society, Chemical Communications*, 283.
247 Kamber, B. (1971) *Helvetica Chimica Acta*, **54**, 927.
248 Yoshida, M., Akaji, K., Tatsumi, T., Fujiwara, Y., Kimura, T., and Kiso, Y. (1990) *Chemical and Pharmaceutical Bulletin*, **38**, 273.
249 Tamamura, H., Otaka, A., Nakamura, J., Okubo, K., Koide, T., Ikeda, K., Ibuka, T., and Fujii, N. (1995) *International Journal of Peptide and Protein Research*, **45**, 312.
250 Xu, Y. and Wilcox, D.E. (1998) *Journal of the American Chemical Society*, **120**, 7375.
251 Kiso, Y., Yoshida, M., Kimura, T., Fujiwara, Y., and Shimokura, M. (1989) *Tetrahedron Letters*, **30**, 1979.
252 Wu, B., Warren, J.D., Chen J., Chen, G., Hua. Z., and Danishefskya, S.J. (2006) *Tetrahedron Letters,* **47**, 5219.

4.1.1.5 Histidine

Under standard SPPS conditions the two imidazole nitrogens of unprotected Histidine react with electrophilies such N,N'-cyclohexylcarbodiimide[253], catalyze acyl-transfer reactions[254], and above all, promote the racemisation of the chiral α-carbon[255].

To prevent both alkylation and racemisation the π nitrogen of the imidazole ring must be protected or rather the protection can be located on the τ position, reducing the nucleophilicity of the π nitrogen by inductive effects. Between the two non-equivalent nitrogen atoms (π and τ), of the imidazole ring a rapid proton exchange takes place and makes the two tautomers inseparable (Figure 4.1–4).

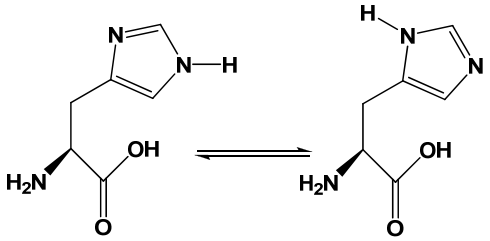


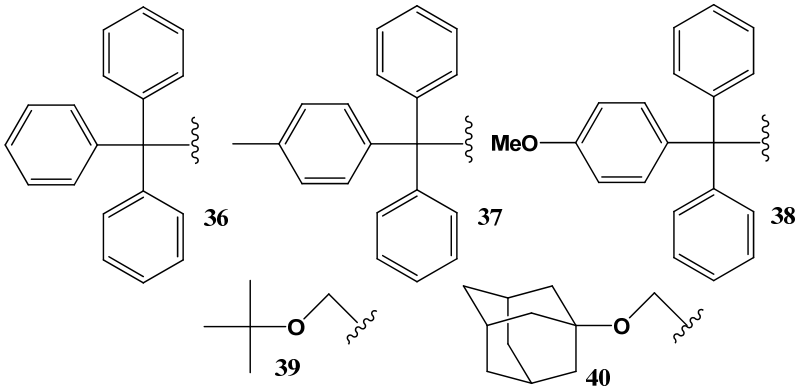
Figure 4.1–4 Tautomeric equilibrium of Histidine side chain imidazole ring

The two positions have almost the same basicity, however their nucleophilicity is significantly different and upon reaction with electrophiles the N-τ product is usually the major one. Thus regiospecific protection of the π position first requires an orthogonal protection of the τ nitrogen and counts for the fact that the synthesis of π products is often troublesome.

253 Rink, H. and Riniker, B. (1974) *Helvetica Chimica Acta*, **57**, 831.
254 Bodansky, M., Fink, M., Klausner, Y., Natarajan, S., and Tatemoto, K. (1977) *The Journal of Organic Chemistry*, **42**, 149.
255 Jones, J. and Ramage, W. (1978) *Journal of the Chemical Society, Chemical Communications*, 472.

The trytyl group (Trt) **36** on the τ-position is the most commonly used protecting group of His for the SPPS with Fmoc strategy[256]. It is stable, commercially available at an affordable price and its mild acidolitic deprotection is fast and smooth. Despite early cautious claims, Trt protection keeps racemisation at a very low rate under normal SPPS conditions and it is regarded to be an exception to the rule that τ-located protection does not completely suppress racemisation[257]. Except in those cases where significant steric hindrance is displayed as for the coupling of His to Pro in which case a small amount of enantiomerization (5%) occurs[258]. However racemisation becomes a serious issue upon esterification of His carboxylic group. For example when hydroxyl-resins are used and His is the first amino acid of a sequence. In this case enatiomerization can be reduced using the Trt group for the α nitrogen protection as well[259]. The best solution to this problem is the use of 2-clorotrityl resin, which can be estherified without racemisation[260]. The protection with the super acid-labile methyltrityl (Mtt) **37** and momomethoxytrityl (Mmt) **38** groups have also been described[261].

256 Barlos, K., Papaioannu, D., and Theodoropoulos, D. (1982) *The Journal of Organic Chemistry,* **47**,1324.
257 Harding, S., Heslop, I., Jones, J., and Wood, M. (1992) In *Peptides 1994, Proceedings of the 23th European peptide Symposium*, Maia H, Ed ESCOM, Leiden, 641.
258 Mergler, F., Dick, F., Sax, B., Shwindling, J., and Vorherr, T.H. (2001) *Journal of Peptide Science*, **7**, 502.
259 Sieber, P. and Riniker, B. (1987) *Tetrahedron Letters,* **28**, 6031.
260 Barlos, K., Chatzi, O., Gatos, D., and Stravropoulos, G. (1991) *International Journal of Peptide and Protein Research,* **37**, 513.
261 Barlos, K., Chatzi, O., Gatos, D., Stravropoulos, G., and Tsegenidis, T. **(**1991) *Tedrahedron Letters*, **32**, 475.



In the case of His rich peptides the tert-butoxymethyl (Bum) **39** protection of the π imidazole position is recommended instead of τ-Trt because it minimizes racemization during peptide coupling reactions[262]. A significant drawback, common to all π-protected derivates, is their difficult synthesis due to the presence of the more reactive τ-position. Bum cleavage with TFA requires slightly longer times compare to tBu, Boc, and Trt. In the case of sequences containing N-terminal Cys, methoxylamine should be added to the cleavage mixture as a scavenger since Bum deprotection generates formaldehyde that can mask Cys as thioazolidine (Thz) giving an adduct with a 12 mass unit difference [263,258].

Another protection of the π imidazole position is the 1-adamantyloxymethyl (1-Adom) **40**, whose derivatives are more soluble in organic solvents than Bum ones and give better synthetic yields[264].

262 Colombo, R., Colombo, F., and Jones, J. (1984) *Journal of the Chemical Society, Chemical Communications,* 292.
263 Gesquiere, J., Najib, J., Diesis, E., Barbry, D., and Tartar, A. (1992) In *Peptides Chemistry and Biology, Proceedings of the 12th American peptide Symposium*, Smith J. and River J Eds. ESCOM, Leiden, 641.
264 Okada, Y., Wang, J., Yamamoto, T., and Mu, Y. (1996),*Chemical and Pharmaceutical Bulletin,* **44**, 871.

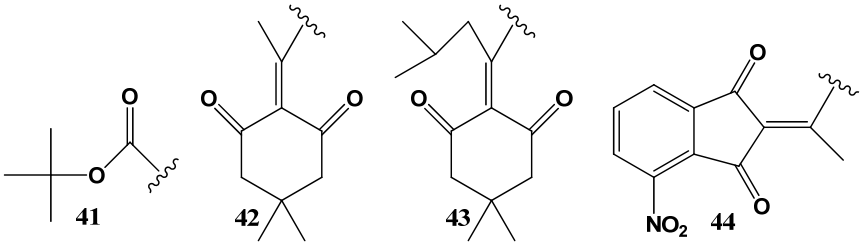
4.1.1.6 Lysine

Lysine side chain protection it is a must in solid phase peptide synthesis, otherwise the ε-amino group would react with acylating agents leading to uncontrolled peptide branching.

The standard protection in Fmoc-strategy is the t-butyloxycarbonyl (Boc) **41**, cleaved with concentrated TFA, which represents a perfect combination together with Fmoc/tBu[265].

Lysine residues are often post-synthetically modified in modern peptide synthesis of bioactive and modified peptides. Usually the goal is to introduce, on a given sequence, post-translational modifications as for example glycations[266] and glycosylations[267] or taking advantage of the ability to form amide bonds between Lysine side chains and molecular devices bearing carboxylic groups[268]. This latter application applies to a broad range of peptide modifications like linking to cromophores, chelating agents, radioactive molecules and many more substituents each conferring a specific property to the modified peptide. The amide bond formation is a straightforward one and highly compatible with on resin peptide synthesis. For all those applications of selective modifications a broad panel of orthogonal Lysine side chain protecting groups was developed.

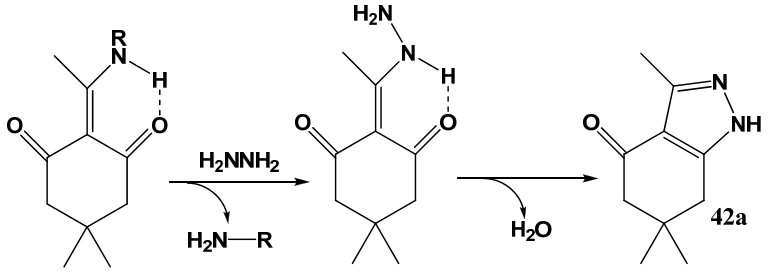
265 Schwyzer, R. and Rittel, W. (1961) *Helvetica Chimica Acta*, **44**, 159.
266 Carganico, S., Rovero, P., Halperin, J.A., Papini, A.M., and Chorev, M. (2009) *Journal of Peptide Science*, **15**, 67.
267 Paolini, I., Nuti, F., Pozo-Carrero, M., Barbetti, F., Kolesinska, B., Kaminski, Z.J., Chelli M., and Papini, A.M. (2007) *Tetrahedron Letters,* **48**, 2901.
268 Grandjean, C., Rommens, C., Gras-Masse, H., and Melnyk, O., (1999) *Tetrahedron Letters*, **40**, 7235.



Particularly popular among orthogonal Lysine side chain protection is 1-(4,4'-dimethyl-2,6-dioxocyclohexylidene)ethyl group (Dde) **42** removed with 2% hydrazine in DMF[269] (Scheme 4.1-2). However Dde has same limitations, it is partially labile to piperidine (once removed it can migrate to other unprotected Lysines), and then hydrazine partially removes the N^{α}-Fmoc[270]. In addition hydrazine can also reduce the Alloc protection preventing its subsequent removal[271]. In order to prevent this, allyl alcohols need to be added as scavengers when Dde is deprotected.

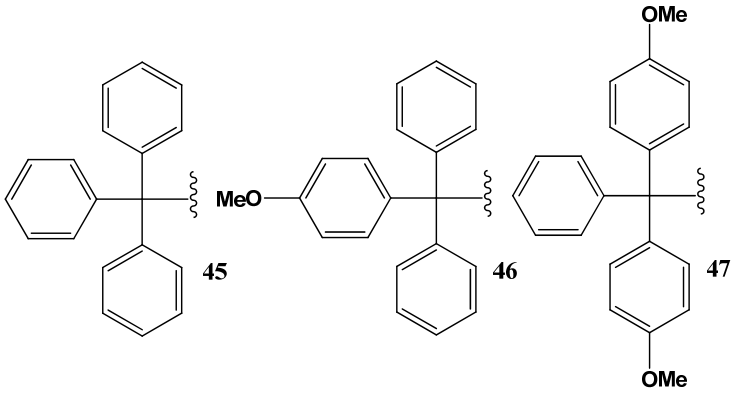
Several hindered Dde variants, in particular 1-(4,4'-dimethyl-2,6-dioxocyclohexylidene)-3-methylbutyl (ivDde)[272] **43**, and the structurally similar 2-acetyl-4-nitroindane-l,3-dione (Nde)[273] **44**, completely overcome those side reactions. Dde, ivDde, and Nde deprotection can be monitored both spectrophotometrically at various UV wavelengths and by mass spectrometry because of the different hydrazine adducts **42a**. Nde removal can be also followed by a change of colour of the resin and solution.

269 Rohwedder, B., Mutti, Y., and Mutter, M., (1998) *Tetrahedron Letters*, **39**, 1175.
270 Augustyns, K., Kraas, W., and Jung, G. (1998) *Journal of Peptide Research,* **51**, 127.
271 Eichler, J., Lucka, W.A., and Houghten, R.A. (1994) *Journal of Peptide Research,* **7**, 300.
272 Chhabra, S.R., Hothi, B., Evans, D.J., White, P.D., Bycroft, B.W., and Chan, W.C. (1998) *Tetrahedron Letters,* **39**, 1603.
273 Kellam, B., Bycroft, B.W., Chan, W.C. and Chhabra, S.R. (1998) *Tetrahedron,* **54**, 6817.



Scheme 4.1-2 Mechanism of hydrazine mediated Lys(Dde) deprotection.

The acid sensitivity of the trityl group, removed with 20% TFA in DCM, is increased by the introduction of electron spending substituents leading to super-acid-labile groups as 4-methyltrityl (Mtt)[274] **45** monomethoxytrityl (Mmt) **46**, and dimethoxytrityl (Dmt)[275] **47**.

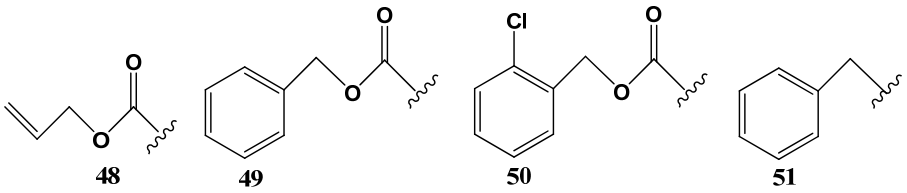


Mtt and Dmt deprotection occurs with 1% TFA in DCM or (DCM/HFIP/TFE/TES 6.5:2:1:0.5) allowing selective removal in presence of other acid-labile protecting groups, like tBu and Boc, that are cleaved by concentrated TFA[276]. Mmt deprotections takes place at even milder conditions, with AcOH/TFE/DCM (1:2:7) and even in the presence of

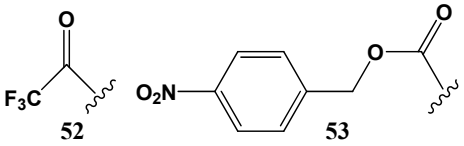
274 Alestras, A., Barlos, K., Gatos, D., Koutsogianni, S., and Mamos, P. (1995) *International Journal of Peptide and Protein Research,* **45**, 488.
275 Matysiak, S., Bidicke, T., Tegge, W., and Frank, R. (1998) *Tetrahedron Letters*, **39**, 1733.
276 Barlos, K., Gatos, D., Chatzi, O., Koutsogianni, S., and Schaefer, W. (1993) In *Peptides 1992, Proceedings of the 22th European Peptide Symposium* C. H. Schneider and A. N.Eberle (Eds), ESCOM, Leiden, 283.

hydrophilic resins such as Tentagel and cellulose that display quenching effect for Mtt and Dmt removal.

The N-ε function in Lysine can bear the allyloxycarbonyl (Alloc) protection **48**, compatible with both Boc/Bzl and Fmoc/tBu strategies and can be removed by palladium catalyst in presence of nucleophiles like NMM[277]. Other groups removed by palladium catalyzed hydrogenolysis, like benzyloxycarbonyl (Z) **49** and 2-chlorobenzyloxycarbonyl (2-Cl-Z)[278] **50** or the benzyl group (Bzl)[279] **51** are also used.



The trifluoroacetyl group (Tfa) **52**, stable to both acid and mild base, removed by strong alkaline aqueous solutions or sodium borohydride, although more common of the Boc-strategy, has found some applications in the Fmoc-chemistry as well[280].



Another semipermanent side-chain protection of Orn and Lys is p-nitrobenzyloxycarbonyl (pNZ) **53** for Fmoc/tBu chemistry that does not result in partial removal of N^{α}-Fmoc that occurs when groups such as Alloc derivates are used for the same application due to the formation of highly

277 Lyttle, M. and Hudson, D. (1992) In *Peptides Chemistry and Biology. Proceedings of the 12th American Peptide Symposium*, Smith J. River J. ESCOM, Leiden, 583.
278 Erickson, B. and Merrifield, R., (1973) *Journal of the American Chemical Society,* **95**, 3757.
279 Huang, Z., Su, X., Du, J., Zhao, Y., and Li, Y., (2006) *Tetrahedron Letters,* **47**, 5997.
280 Stetsenko, D.A. and Gait, M.J. (2001) *Bioconjugate Chemistry*, **12**, 576.

basic free amine[281]. Furthermore, pNZ, removed by 6 $MSnCl_2$ and 1.6 mM HCl/dioxane in DMF, can be used in conjunction with p-nitrobenzyl ester (pNB) to prepare cyclic peptides[282].

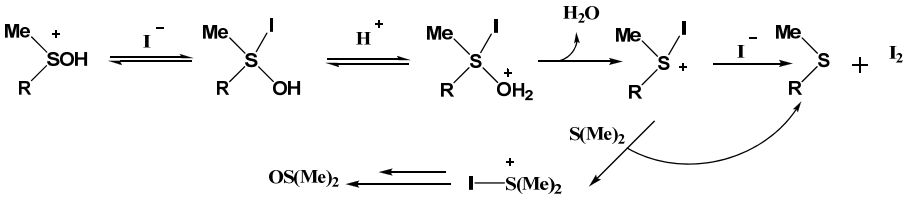
Finally, the Fmoc protection of both α- and ε-nitrogen represents a specialized derivate for simultaneous multiple peptide synthesis, where side chain branching is desired after α-amino deprotection[283] .

281 Farrera-Sinfreu, J., Royo, M., and Albericio, F. (2002) *Tetrahedron Letters*, **43**, 7813.
282 Llobet, A. I. (2005) *Tetrahedron Letters,* **46**, 7733.
283 Tam, J. (1988) *The Proceedings of the National Academy of Science US*, **85**, 5409.

4.1.1.7 Methionine

The thioether function of Methionine can undergo acid catalyzed oxidation forming methionyl sulfoxide and can also be alkylated. In standard Fmoc-based peptide synthesis is common practise to use unprotected Methionine because the use of the proper scavengers like thioanisole[284] in the cleavage mixture greatly reducing both the side reactions.

Sulfoxides are also generated by prolonged air exposure but can be reduced back to the sulphide in several ways such as NH_4I/Me_2S in TFA at low temperatures[285] (Scheme 4.1-3) or NaI/CH_3COCl in DMF[286]



Scheme 4.1-3 Acid NH_4I/Me_2S mediated Met(O) reduction mechanism.

Another approach to the oxidation problem, although more common of the Boc strategy, is introducing Methionine as the sulfoxide derivate Met(O)[287] prepared by treatment with H_2O_2 and then reducing it at a convenient time of the synthesis.

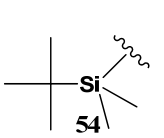
284 Yajima, H., Kanaki, J., Kitajima, M., and Funakoshi, S. (1980) *Chemical and Pharmaceutical Bulletin*, **28**, 1214.
285 Vilaseca, M., Nicolfis, E., Capdevila, F., and Giralt, E. (1998) *Tetrahedron*, **54**, 15273.
286 Norris, K., Halstrom, J., and Brunfeldt, K. (1971) *Acta Chemica Scandinavica*, **25**, 945.
287 Iselin, B. (1961) *Helvetica Chimica Acta*, **44**, 61.

4.1.1.8 Serine and Threonine

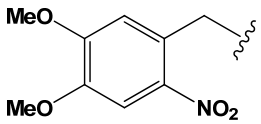
The side reactions affecting unprotecded hydroxyl group of Ser, Thr and Tyr side chains, under the standard synthetic conditions used in peptide chemistry, are less severe compared to amino and carboxylic function. As a consequence a number of synthesis describing unprotected amino acid incorporation has been reported, especially in the past. However unprotected hydroxyl groups can suffer acylation and dehydratation and it is therefore normal practice to protect them.

Ser and Thr have very similar behaviour and characteristics but the minor steric hindrance of the former's hydroxyl makes it more easy to protect but also more reactive toward acylating reagents.

The classical protection of such amino acids is tBu removed under strong acidic conditions[288].Selective deprotection of hydroxyl side chains is often performed during post-synthetic modification of Ser and Thr, especially with phosphorylation and glycosylation. This can be achieved by the trytyl protection, removed orthogonally from tBu and Boc under mild acidic conditions, 1%TFA and 5% TIS in DCM[289] or 20% dichloroacetic acid in DCM[290]. Another acid-lable group is t-butyldimethylsilyl (TBDMS) **54** which can be removed with AcOH/THF/H_2O/ 3:1:1[291].



54



55

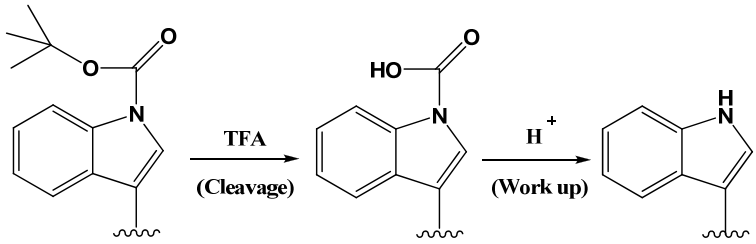
288 Chang, C., Waki, M., Ahmad, M., Meienhofer, J,. Lundell, E., and Huang, J. (1980) *International Journal of Peptide and Protein Research*, **15**, 59.
289 Barlos, K., Gatos, D., and Koutsogianni, S. (1998) *Journal of Peptide Research*, **51**, 194.
290 Coba, M.P., Turyn, D., and Pena, C. (2003) *Journal of Peptide Research*, **61**,17.
291 Fisher P. (1992) *Tetrahedron Letters,* **33**, 7605.

The photocleavable 4,5-dimethoxy-2-nitrobenzyl group[292] **55** has been used for Serine protection. Finally Ser and Thr side chains can be protected as benzyl esthers, removed with TFA.

292 Pirrung, M.C. and Nunn, D.S. (1992) *Bioor. Med. Chem. Lett,* **2**, 1489.

4.1.1.9 Tryptophan

The two main side reactions affecting Tryptophan residues in Fmoc-chemistry are oxidation and alkylation of the indole ring by carbonium ions generated during the TFA cleavage[293]. Alkylation of unprotected Tryptophan could be kept under control by using EDT as a cleavage mixture scavenger[294]. However sulfonation by the by-products of Mtr, Pmc and Pbf protected Arginines cannot be completely eliminated. This problem has been solved by developing t-butoxycarbonyl (Boc) protected Trp[295], which generates after the cleavage step a TFA stable-N-carboxy indole intermediate, capable of reducing the susceptibility of the heterocyclic ring to the electrophilic attack[296] (Scheme 4.1-4). The carbamic acid derivate, associated with a 44 mass unit gain, is not stable in solution and decomposes slowly in water during the routine work-up stages, leaving the indole ring free.



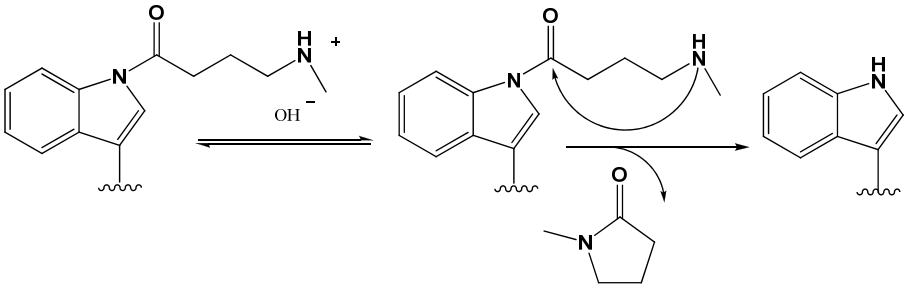
Scheme 4.1-4 Mechanism of Trp(Boc) deprotection.

293 Fields, C. and Fields, G. (1993) *Tetrahedron Letters,* **34**, 6661.
294 Fields, G., Noble, R. (1990) *International Journal of Peptide and Protein Research*, **35**, 161.
295 White, P. (1992) *In Peptides, Chemistry, Structure and Biology, Pro. 12th American peptide Symposium*, J.A. Smith and J.E.Rivier(Eds.), ESCOM, Leiden, 537.
296 Franzen, H., Grehn, L., Ragnarsson, U. (1984) *Journal of the Chemical Society, Chemical Communications*, 1699.

Tryptophan can also suffer partial reduction by triethylsilane used as a scavenger of trityl groups[297] and is affected by the presence of silver salts (used for removal of several Cysteine-protecting groups). The latter side effect can be avoided by adding an excess of free Tryptophan in the deprotection mixture[298]

Also the allyloxycarbonyl (Alloc) protection eliminates the oxidation-alkyklation problems and, although unstable to 20% piperidine, it is stable to DBU that needs therefore to be used for Fmoc removal[299].

Recently a new TFA stable Tryptophan protection has been proposed, 4-(N-methylamino)butanoyl (Nmbu), whose function is to improve the solubility of the peptides in view of HPLC purification[300]. Nmbu is stable to TFA cleavage and on treatment of the purified peptide at pH 9.5 it undergoes an intramolecular cyclization reaction (Scheme 4.1-5) that results in the fully deprotected peptide and N-methylpyrrolidone.



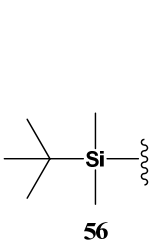
Scheme 4.1-5 Mechanism of Trp(Nmbu) deprotection.

297 Pearson, D., Blanchette, M., Guindon, C. (1989) *Tetrahedron Letters*, **30**, 2739.
298 Najib, J., Letailleur, T., GesQuire, J., Tartar, A. (1996) *Journal of Peptide Science*, **2**, 309.
299 Vorherr, T., Trzeciak, A., Bannwarth, W., (1996) *International Journal of Peptide and Protein Research,* **48**, 553.
300 Wahlström, K. and Undén, A. (2009) *Tetrahedron Letters,* **50**, 2976.

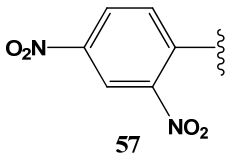
4.1.1.10 Tyrosine

Tyrosine side chain, if left unprotected during the peptide synthesis, is prone to O-acylation because the basic conditions of Fmoc removal generates the phenolate anion which is acylated due to its strong nucleophilicity. Tyrosine is commonly protected with tBu group[301], which generates very little of the 3-alkylated product (see side reaction section).

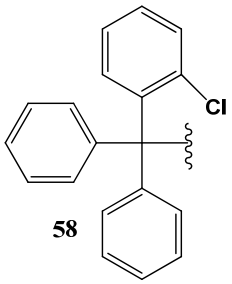
The TBDMS ether protection **56**, although less acid-labile than the corresponding analogs of Serine and Threonine, can nethertheless be removed selectively using tetrabuthylammonium fluoride (TBAF)[291].



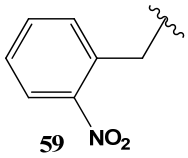
56



57



58



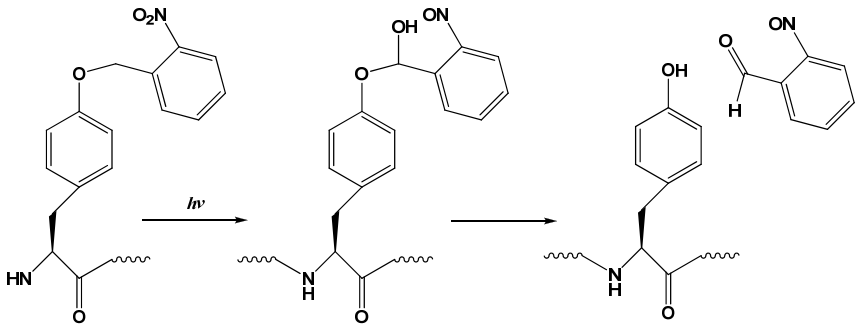
59

Despite some misunderstandings[302], the 2,4-dinitrophenyl (Dnp) group **57**, deprotect with 2-thiophenol/pyridine/DMF 2:1:10, is another suitable choice for Tyrosine selective protection as for example during on resin post-synthetic modifications (such as phosphorilation or glycosylation). However since Dnp is readily cleaved by 20% piperidine or 2 % DBU in DMF[303], Dnp-protected Tyrosine should be employed as the N-terminal residue or could be immediately modified after incorporation. Another possibility is

301 Adamson, J., Blaskowitch, M., Groenvelt, H., Lajoie, G. (1991) *The Journal of Organic Chemistry,* **56**, 3447.
302 Doherty-Kirby, A., Lajoie, G. In *Solid-Phase Synthesis A Practical Guide* (2000), Kates, S.A. and Albericio F., Marcel Dekker Inc, 148.
303 Philosof-Oppenheimer, R., Pecht, I., Fridkin, M. (1995) *International Journal of Peptide and Protein Research*, **45**, 116.

using the more flexible 2-ClTrt group **58** that can be selectively removed any time of the synthesis with 1% TFA / 5% TIS in DMF[289].

Same photo-cleavable Tyrosine protecting groups have been reported, particularly useful for the synthesis of those molecular devices whose activity is controlled by light (caged compounds). One example is the 2-nitrobenzyl group (NB) **59**, removed by UV light[304] (Scheme 4.1-6). Upon irradiation the exited nitro compound abstracts a hydrogen from the benzylic position and the intermediate rapidly rearranges into a nitoso hemiacetal



Scheme 4.1-6 Mechanism of Tyr(NB) deprotection.

304 Tatsu, Y., Shigeri, Y., Sogabe, S., Yumoto, N., and Yoshikawa, S. (1996) *Biohemical and Biophysical Research Communications,* **227**, 688.

5 EXPERIMENTAL PART A

5.1 Materials and methods

The chemicals were purchased from Sigma-Aldrich and used without further purification. TLC were carried out on silica gel precoated plates (Merck; 60 Å F254) and spots located with: (a) UV light (254 and 366 nm), (b) ninhydrin (solution in acetone), (c) Cl_2/toluidine, (d) fluorescamine, (e) I_2, (f) a basic solution of permanganate [$KMnO_4$ (3 g), K_2CO_3 (20 g), and NaOH (0.25 g) in water (300 ml)], (g) 10% H_2SO_4 in EtOH. Flash Column Chromatography (FCC) was performed on Merck silica gel 60 (230-400 mesh) according to Still et al.[305]

^1H and ^{13}C NMR spectra were recorded at 400 and 100 MHz, and 200 and 50 MHz respectively, on a Varian spectrometer in deuterated solutions and are reported in parts per million (ppm), with solvent resonance used as reference. Melting points were determined on a Büchi mod. 510 apparatus. Elemental analyses were performed on a Perkin Elmer 240 C Elemental Analyzer. Infrared spectra were recorded on a Perkin Elmer mod. BX II FT-IR spectrometer. The $[\alpha]_D$ were obtained on Perkin Elmer mod. 343 polarimeter in cell of 1 dm. Products were analyzed and characterized by ACQUITY UPLC (Waters Corporation, Milford, Massachusets) coupled to a single quadrupole ESCI-MS (Micromass ZQ) using a 2.1 x 50 mm 1.7 μm ACQUITY BEH C18 at 30 °C, with a flow rate of 0.45mL/min. The solvent systems used were A (0.1% TFA in H_2O) and B (0.1% TFA in CH_3CN). High performance liquid chromatography -grade acetonitrile (MeCN) was purchased from Carlo Erba (Italy).

---

[305] Still, W. C.; Kahn, M.; Mitra, A. *J. Org. Chem.* **1978**, *43*, 2923–2925.

Protected amino acids were obtained from Novabiochem AG (Laufelfingen, Switzerland). TBTU and HOBt were purchased from Iris-Biotech. Peptide-synthesis grade N,N-dimethylformamide (DMF) was purchased from Scharlau (Barcelona, Spain). TFA, DCM, Piperidine, Ac_2O, and NMM were purchased from Aldrich.

Peptides were analyzed by analytical RP-HPLC (Alliance, model 2695 equipped with a diode array detector, Waters) using a Jupiter C18 (5 μm, 250 × 4.6 mm) column (Phenomenex) at 1 mL/min. The solvent systems used were A (0.1% TFA in H_2O) and B (0.1% TFA in CH_3CN). Peptides were purified by preparative RP-HPLC (model 600, Waters) on a Jupiter C18 or C8 column (10 μm, 25 cm × 10 mm) at 4 mL/min by using the same solvent systems reported above.

5.2 Glycated building blocks for the synthesis of post-translationally modified glycated peptides and proteins

5.2.1 *Nα-Fmoc-Lys[Nε-(Deoxyfructopyranosyl)]-OH*

Nα-Fmoc-Lys-OH (1 g, 2.71 mmol, 1 eq) and D-glucose (1.22 g, 6.77 mmol, 2.5 eq) were suspended in 40 ml of anhydrous DMF under N2 atmosphere[70]. The reactionmixture was thenwarmed at 110 ◦C for about 10 min: Fmoc-Lys-OH dissolved and before the colour of the reaction mixture changed from yellow to brown (index of diglycation), the solution was immediately cooled in an ice bath. DMF was evaporated under vacuum and the crude residue was purified by RP-FC using a gradient 15–40% B in A over 20 min, elution occurred at □25% B. Homogeneous fractions were then evaporated and the water solution lyophilized yielding pure **1a** as a white, highly hygroscopic solid (390 mg, 67%).

LC-ESI-MS (m/z): $[M+H]^+$ calcd for $C_{27}H_{34}N_2O_9$ 531.23; found, 531.38. M.p.= 115°C (caramelization and decomposition were observed). $[\alpha]^{26}_D$ = −20.70 (c = 1.015, H_2O). Elemental analysis calcd for $C_{27}H_{34}N_2O_9$ ·TFA·H_2O: C, 49.86; H, 5.92; N, 4.01; found: C, 49.88; H, 6.26; N, 4.10.

^{1}H NMR (D_2O, 400 MHz): δ 7.2 (broad s, 4H, fluorenyl 4-H, 5-H, 1-H and 8-H), 6.9 (broad s, 4H, fluorenyl 3-H, 6-H, 2-H and 7-H), 4.1-3.9 (m, 2H, CH_2, Fmoc), 3.85 (broad s, CH Fmoc) 3.8 (broad s,CH_α, Lys), 3.8-3.7 (m, CHOHx3, 1-deoxyfructosyl), 3.6 (t, 2H, J=10.0 Hz, 1-deoxyfructosyl), 3.1-2.9 (m,CH_2NH, 1-deoxyfructosyl), 2.8-2.6 (m, $CH_{2\varepsilon}$, Lys), 1.6-1.2 (m, 4H, $CH_{2\beta,\delta}$), 1.1-0.8 (m, $CH_{2\gamma}$, Lys)

^{13}C NMR ($CDCl_3$, 50 MHz) δ 175.8 (C, COOH), 162.7 (q, J=35.4 Hz, C, COOH, TFA), 157.1 (C, CONH, Fmoc), 116.3 (q, J=291.0 Hz, CF_3, TFA), 143.2 (CH_{Ar} x2, Fmoc), 140.6 (CH_{Ar} x2, Fmoc), 127.5 (CH_{Ar} x2, Fmoc), 126.9 (CH_{Ar} x2, Fmoc), 124.7 (CH_{Ar} x2, Fmoc), 119.7(CH_{Ar} x2, Fmoc), 95.3

(C_4, C1 1-deoxyfructosyl), 69.5, 69.3, 68.9 (CHOH x3, 1-deoxyfructosyl), 69.4 (CH_2, 1-deoxyfructosyl), 63.9 (CH_2, Fmoc), 53.9, 52.8 (CH_2NH ideoxyfructosyl, C_α Lys) 48.0 ($CH_2\varepsilon$, Lys), 46.7 (CH, Fmoc), 30.5 ($CH_2\beta$, Lys), 24.8 ($CH_2\gamma$, Lys), 22.2 ($CH_2\delta$, Lys).

5.2.2 *Nα-Fmoc-Lys[Nε-(Deoxyfructopyranosyl),Nε-Boc]-OH*

A solution of Boc_2O (514 mg, 2.35 mmol, 2.5 eq) in 5 mL of MeOH was added, in N_2 atmosphere, under stirring at 0° C, to a solution of **1a** (500 mg 0.94 mmol, 1eq). The reaction was left under stirring at rt for 1.5 h and then the solvent was evaporated and the residue purified by RP-FC with a linear gradient of 30-55% B in A over 20 min. Elution occurred at ~50% B. Homogeneous fractions were then evaporated and the water solution lyophilized yielded pure **1** as a white hygroscopic solid (270 mg, 45%).

LC-ESI-MS (m/z): $[M+Na]^+$ calcd for $C_{32}H_{42}N_2O_{11}$ 653.28; found, 653.26. M.p. = 105°-106°C. $[\alpha]^{26}_D = -16.18$ (c = 0.94, MeOH). Elemental analysis calcd for $C_{32}H_{42}N_2O_{11}\cdot 5H2O$: C, 53.32; H, 7.27; N, 3.88; found: C, 53.03; H, 7.09; N, 4.09.

^{1}H NMR ($CDCl_3$, 400 MHz): δ 7.8-7.6 (m, 2H, fluorenyl 4-H, 5-H), 7.6-7.4 (m, 2H, fluorenyl 1-H and 8-H), 7.4-7.2 (m, 4H, fluorenyl 3-H, 6-H, 2-H and 7-H), 4.5-4.3 (m, 2H, CH_2, Fmoc), 4.3-4.1 (m, 2H, CH Fmoc and CH_α Lys), 4.1-3.8 (m, CHOHx3, 1-deoxyfructosyl), 3.8-3.5 (m, 1-deoxyfructosyl), 3.5-2.9 (m, 4H, CH_2NH 1-deoxyfructosyl and $CH_{2\varepsilon}$ Lys), 2.0-1.4 (m, 4H, $CH_{2\beta,\delta}$), 1.4 (s, CH_3x3, Boc), 1.4-1.0 (m, $CH_{2\gamma}$, Lys).

^{13}C NMR ($CDCl_3$, 50 MHz) δ 174.9 (C, COOH), 158.5 (C, CONH, Boc), 156.0 (C, CONH, Fmoc), 143.5 (CH_{Ar} x2, Fmoc), 141.1 (CH_{Ar} x2, Fmoc), 127.6 (CH_{Ar} x2, Fmoc), 127.0 (CH_{Ar} x2, Fmoc), 125.0 (CH_{Ar} x2, Fmoc), 119.9 (CH_{Ar} x2, Fmoc), 98.7 (C_4, C1 1-deoxyfructosyl), 81.4 (C_4, Boc), 70.6, 70.1, 69.5 (CHOH x3, 1-deoxyfructosyl), 67.2 (CH_2, 1-deoxyfructosyl), 63.3

(CH_2, Fmoc) 54.6, 53.6 (CH_2NH ideoxyfructosyl, C_α Lys) 50.0 ($CH_2\varepsilon$, Lys), 47.2 (CH, Fmoc), 31.9 ($CH_2\beta$, Lys), 28.5 (CH_3x3, Boc), 27.8 ($CH_2\delta$, Lys), 22.2 ($CH_2\gamma$, Lys).

5.2.3 *Nα-Fmoc-Lys[Nε-(2,3:4,5-Di-O-Isopropylidene-1-Deoxyfructopyranosyl)]-OH*

A fresh solution of $NaCNBH_3$ (600 mg, 9.5 mmol, 2.5eq) in 5 mL H_2O/THF (1:1, v/v) was added to a stirred solution of N^{α}-Fmoc-Lys-OH (1.4 g, 3.8 mmol, 1eq) and 2,3:4,5-di-*O*-isopropylidene-*aldehydo*-β-D-*arabino*-hexos-2-ulo-2,6-pyranose[70,71,72] (2.45 g, 9.5 mmol, 2.5eq) in 15 mL H_2O/THF (1:1, v/v) under N_2 at 50° C. After 4h the solvent was removed under reduced pressure and the crude product was purified by RP-FC, eluted with the linear gradient 30-50% B in A over 20 min (eluted at ~37% B). The homogeneous fractions were pooled, evaporated and liophylized yielding the pure **2a** as a white solid (500 mg, 22 %).

LC-ESI-MS (m/z): $[M+H]^+$ calcd for $C_{33}H_{42}N_2O_9$ 611.29; found, 611.17. M.p.= 92°-93°C. $[\alpha]^{26}_D$= −12.04 (c= 0.89, MeOH). Elemental analysis calcd for $C_{33}H_{42}N_2O_9 \cdot TFA \cdot 2H_2O$: C, 55.26; H, 6.23; N, 3.68; found: C, 55.35; H, 6.15; N, 3.39.

^{1}H NMR ($CDCl_3$, 400 MHz): δ 7.74 (d, 2H, *J*= 7.4 Hz, fluorenyl 4-H and 5-H), 7.59 (d, 2H, *J* = 7.4 Hz, fluorenyl 1-H and 8-H), 7.37 (t, 2H, *J*= 7.4 Hz, fluorenyl 3-H and 6-H), 7.28 (t, 2H, *J*= 7.4 Hz, fluorenyl 2-H and 7-H), 5.85 (broad d, NH_α), 4.65-455 (m, CH Fmoc), 4.4-4.3 (3H, CH_2, Fmoc and CH_α Lys), 4.25-4.15 (m, CHOHx3, 1-deoxyfructosyl), 3.8-3.7 (m, CH_2 1-deoxyfructosyl), 3.35-3.30 (m, CH_2NH 1-deoxyfructosyl), 3.30-3.20 (m,$CH_{2\varepsilon}$ Lys), 1.95-1.65 (m, 4H, $CH_{2\beta,\delta}$) 1.5-1.45 (m, $CH_{2\gamma}$, Lys) 1.47, 1.41, 1.33, 1.28 (s, CH_3x4, isopropylidene).

^{13}C NMR ($CDCl_3$, 50 MHz) δ 174.5 (C, COOH), 161.9 (q, J=38.2 Hz, C, COOH, TFA), 156.2 (C, CONH, Fmoc), 116.5 (q, J=290.9 Hz, CF_3, TFA), 143.6 (CH_{Ar} x2, Fmoc), 141.1 (CH_{Ar} x2, Fmoc), 127.6 (CH_{Ar} x2, Fmoc), 127.0 (CH_{Ar} x2, Fmoc), 125.1 (CH_{Ar} x2, Fmoc), 119.8 (CH_{Ar} x2, Fmoc), 110.0 (C_4, isopropylidene), 109.3 (C_4, isopropylidene), 99.56 (C_4, C1 1-deoxyfructosyl), 72.2, 70.3, 69.6 (CHOH x3, 1-deoxyfructosyl), 67.1 (CH_2, 1-deoxyfructosyl), 61.58 (CH_2, Fmoc) 54.3, 53.7 (CH_2NH ideoxyfructosyl, $C_α$ Lys) 48.7 (CH_2ε, Lys), 47.2 (CH, Fmoc), 31.7 (CH_2β, Lys), 26.1, 25.9, 24.6, 24.1 (CH_3x4, isopropylidene), 25.5 (CH_2γ, Lys), 22.0 (CH_2δ, Lys).

5.2.4 *Nα-Fmoc-Lys[Nε-(2,3:4,5-Di-O-Isopropylidene-1-Deoxyfructopyranosyl),Nε-Boc]-OH*

A stirred solution of **2a** (500 mg 0.81 mmol, 1eq) in methanol (5 mL) at 0° C was treated with Boc_2O (445 mg, 2.07 mmol, 2.5 eq). The reaction was left stirring at rt for 1.5 h and then the solvent was evaporated and the residue purified by RP-FC employing a linear gradient of 50-100% B in A over 20 min (the product 2 eluted at ~70% B). Acidification of the concentrate, obtained after evaporation of acetonitrile from the pooled fractions, to pH 3 with acetic acid, resulted in a white precipitate that was filtered off, washed with water and dried under vacuum. The pure **2** was obtained as a white solid (390 mg, 67 %). LC-ESI-MS (m/z): $[M+H]^+$ calcd for $C_{38}H_{50}N_2O_{11}$ 711.34; found, 711.43. M.p.= 97-98° C. $[\alpha]^{26}_D$= −16.97 (c= 1.03, MeOH). Elemental analysis calcd for $C_{38}H_{50}N_2O_{11}\cdot H_2O$: C, 62.62; H, 7.19; N, 3.84; found C, 62.56; H, 7.29; N, 3.84.

^{1}H NMR ($CDCl_3$, 400 MHz): δ 7.75 (d, 2H, *J*= 7.4 Hz, fluorenyl 4-H and 5-H), 7.59 (d, 2H, *J* = 7.4 Hz, fluorenyl 1-H and 8-H), 7.38 (t, 2H, *J*= 7.4 Hz, fluorenyl 3-H and 6-H), 7.30 (t, 2H, *J*= 7.4 Hz, fluorenyl 2-H and 7-H), 5.45 (broad d, $NH_α$), 4.6-4.55 (m, CH Fmoc), 4.50-4.5 (m, $CH_α$ Lys), 4.45-4-30

(m, CHORx3, 1-deoxyfructosyl), 4.25-4.15 (m, CH_2 Fmoc) 3.75-3.60 (m, CH_2 1-deoxyfructosyl), 3.35-3.25 (m, CH_2NH 1-deoxyfructosyl), 3.35-2.25 (m,$CH_{2\varepsilon}$ Lys), 2.00-1.70, 1.55-1.45, 1.35-1.25 (m, 6H, $CH_{2\beta,\delta,\gamma}$, Lys) 1.44 (s, CH_3x3, Boc),1.50, 1.47, 1.34, 1.31 (s, CH_3x4, isopropylidene).

^{13}C NMR ($CDCl_3$, 50 MHz) δ 175.5 (C, COOH), 159.1 (C, CONH, Boc), 156.1 (C, CONH, Fmoc), 143.7 (CH_{Ar} x2, Fmoc), 141.1 (CH_{Ar} x2, Fmoc), 127.6 (CH_{Ar} x2, Fmoc), 127.0 (CH_{Ar} x2, Fmoc), 125.0 (CH_{Ar} x2, Fmoc), 119.8 (CH_{Ar} x2, Fmoc), 108.8 (C_4, isopropylidene), 108.0 (C_4, isopropylidene), 104.5 (C_4, C1 1-deoxyfructosyl), 80.8 (C_4, Boc), 71.4, 70.7, 70.5 (CHOH x3, 1-deoxyfructosyl), 67.2 (CH_2, 1-deoxyfructosyl), 61.3 (CH_2, Fmoc), 53.9, 51.4 (CH_2NH ideoxyfructosyl, C_α Lys) 48.5 ($CH_2\varepsilon$, Lys), 47.3 (CH, Fmoc), 32.5 ($CH_2\beta$, Lys), 28.6 (CH_3x3, Boc), 27.4 ($CH_2\delta$, Lys), 26.4, 26.1, 25.1, 24.1 (CH_3x4, isopropylidene), 22.7 ($CH_2\gamma$, Lys).

5.3 Synthesis of a glycated hCD59 antigen

5.3.1 *Synthesis of hCD59(37-50) peptide analogs*

Intermediate peptide **I** and **II** for post-synthetic modifications and peptide **III** (stepwise synthesis) were synthesized by SPPS according to the methodologies described in chapter 5.6 (Table 5.3-1)

	Sequence and side chain protections
I	Ac-N(Trt)-K(Boc)-A-W(Boc)-K(ivDde)-F-E(OtBu)-H(Trt)-A-N(Trt)-F-N(Trt)-D(OtBu)-C(Trt)-OH
II	Ac-N(Trt)-K(ivDde)-A-W(Boc)-K(Boc)-F-E(OtBu)-H(Trt)-A-N(Trt)-F-N(Trt)-D(OtBu)-C(Trt)-OH
III	Ac-N(Trt)-K(Boc)-A-W(Boc)-K*-F-E(OtBu)-H(Trt)-A-N(Trt)-F-N(Trt)-D(OtBu)-C(Trt)-OH

Table 5.3-1 Peptide analogs for the convergent synthesis of glycated hCD59(37-50) antigen. K* is Fmoc Lysine derivate 2 (section 5.2.4).

5.3.2 *On resin direct glycation*

The peptide **I** was synthesized by SPPS and acetylated according to the procedures described in section 5.6. Residue 41 was then deprotected from ivDde by treatment with hydrazine 2% in DMF for 10min. This removal was repeated 3 times and monitored by LCMS (ivDde deprotection-product mass=221) obtaining intermediate peptide **Ia**. Direct glycation was performed on resin with a 40-fold excess (base on the resin substitution) of Glucose under stirring at 110 C and N2 atmosphere in DMF, stopped after 45 min. The reaction was monitored with microcleavage LCMS analysis, linear gradient of 10-20% B in A over 15 min, Rt=4.66 for peptide **III**
The mass spectrometry shows the presence of un-glycated peptide and of an overall poor conversion of the free epsilon amino into the Amadori form due to the presence of undesired oxidation reactions and di-glycation.

[M+H]+= 1927.2	[M+H]2+= 963.96	glycated product
[M+H]+= 1765	[M+H]2+= 883	un-glycated product
[M+H]+= 1969.5	[M+H]2+= 985	oxidation-product
[M+H]+= 2089	[M+H]2+= 1045	di-glycated-product

The crude peptide was purified by flash chromatography Waters delta prep3000, delta-pak cartridge C18 15 microm, 300 Angstrom 2inc (diam) x 12 inc, 50 mL/min HPLC analysis using a very shallow gradient shown indeed two peaks (glycated and un-glycated with the oxidation product, of mass 985, present in both. Further purification by semi-preparative reverse phase HPLC using very shallow and isocratic gradients was unsatisfactory in terms of yield due to the poor separation of the two forms (glycated and un-glycated).

5.3.3 *On resin reductive amination*

Peptide **I** was treated as in 5.3.2 achieving intermediate peptide **Ia** and the Amadori modification was on Lys 41 by reductive amination treating the peptidyl-resin with a 25-fold molar excess of 2,3:4,5-di-O-isopropylidene-β-D-arabino-hexos-2-ulo-2,6-pyranose and NaCNBH3 2.5 eq in DMF (1mL/g of peptidyl-resin) at 70 C under N2 atmosphere and left under agitation for 4h. The reaction was monitored by LCMS of minicleavages of 2-3 mg of resin using reagent R (TFA:Aniso:Thioanisol:Ethandithiol 90:5:3:2 v/v). Cleavage from the resin and work up according to 5.6. LCMS showed presence of partially protected isopropylidene so the crude peptide was treated with TFA/TIS (95:5) and precipitated in cold ether after 2 h. The crude peptide was purified as in 5.3.2. Purified peptide, LCMS linear gradient of 10-60% B in A over 20 min, Rt= 4.58 min. The peak appeared in the first place to correspond to a pure product by LCMS analysis show it was a mixture of glycated ([M+H]2+= 963.9) and non-glycated peptide ([M+H]2+= 883.1) HPLC analysis using a very shallow gradient shown indeed two peaks. Under the described conditions we were unable to glycated all the free epsilon-amino groups, at least 40% of them remained un-glycated. Under more vigorous conditions as higher T, reaction time and excess of reactants the amount of glycated product increased but di-glycation was seen to a significant extent. Further purification by semi-preparative reverse phase HPLC using very shallow and isocratic gradients was unsatisfactory in terms of yield due to the poor separation of the two forms (glycated and un-glycated). To solve the problem of isolating the glycated product from the non-glycated a purification using a phenyl-boronic column was attempted with good results in terms of separation but with poor yield due to the low loading capacity of those columns.

5.3.4 *In solution direct glycation*

Peptide **II** was synthesized by SPPS, acetylated and cleaved from the resin according to the procedures described in section 5.6 obtaining intermediate peptide **IIa**. To a solution of **IIa**, with position 38 ivDde protected, a 40-fold excess of Glucose was added under stirring at 110 C and N2 atmosphere. Under this temperature conditions the peptide was found to be not stable. At lower temperature the stability increases but the glycation rate is much lower and that leaves the problem to separate the glycated product from the un-glycated one. After purification as in as in 5.3.2 a solution of glycated peptide III was treated with 2% hydrazine in DMF 5min repeated three times to removed ivDde. The Amadori modification was found to be very susceptible to hydrazine. Degradation occurred at >2% or even 2% during liophylization. In conclusion it has been impossible to removed ivDde in presence of the unprotected Amadori modification without completely destroying the sugar moiety.

5.3.5 *In solution reductive amination*

Peptide **II** was treated as in 5.3.45.3.2 achieving intermediate peptide **IIa**. Reductive amination was performed on peptide **IIa** with a 25-fold molar excess of 2,3:4,5-di-O-isopropylidene-β-D-arabino-hexos-2-ulo-2,6-pyranose and NaCNBH3 2.5 eq in DMF (1mL/mg of peptide) at 70 C under N2 atmosphere and left under agitation. The reaction was monitored by LCMS and stopped after 20 min. Purification as in 5.3.2. Purified glycated peptide **IIa** was deprotection from ivDde as in 5.3.4. The resulting solution was purified again as in 5.3.2. The purified peptide was dissolved in 2mL TFA/H2O/TIS (95:2.5:2.5) in order to removed the isopropylidene protections affording peptide **III**.

Starting material (intermediate peptide **IIa**) [M+H]+ =1970 [M+2H]2+=986.4 [M+3H]3+=658; Glycated peptide **III** protected [M+H]+ =2214 [M+2H]2+=1107.5 [M+3H]3+=738.8; Glycated peptide **III** (glycated peptide, OH free) [M+H]+ =1927.4 [M+2H]2+=964.2 [M+3H]3+=643.2.

5.3.6 *Step-wise synthesis by building block approach of Nα-Ac[Lys41(Nε-1-deoxyfructosyl)]hCD59(37-50)-OH (III)*

Peptide **III** was synthesized by SPPS, acetylated and cleaved from the resin according to the procedures described in section 5.6. The Amadori modification in position 41 was stepwise using the glycated Fmoc Lysine derivate, **2** (Nα-Fmoc-Lys[Nε-(2,3:4,5-di-O-isopropylidene-1-deoxyfructosyl,Nε-Boc)] OH) affording the desired with excellent yield

The other three synthesized Amadori s, without Boc protection in the ε-NH_2 **2a** or without isopropylidene protection on the sugar hydroxyls **1** and **1a**, were successfully incorporated into the peptide sequence with Fmoc SPPS strategy. However attempts to couple the subsequent amino acids failed. After the standard TFA cleavage the peptide was still partially isopropylidene protected and displays three main peaks, corresponding to di-isopropylidene protected, mono-isopropylidine protected and glycated product. To completely remove the isopropylidene protection, additional TFA treatment was required (95% TFA, 5% TIS) for 3h.

The crude product was purified as in in 5.3.2.

5.4 Immunological studies

5.4.1 *General procedure for solid-phase non competitive indirect ELISA (SP-ELISA)*

Antibody titers were determined in SP-ELISA.[306] 96-Well activated Polystyrene ELISA plates (Limbro Titertek, ICN Biomedicals, Inc., Aurora, Ohio, USA) were coated with 1 μg/100 μl/well of peptides or glycopeptides in pure carbonate buffer 0.05 M (pH 9.6) and incubated at 4 °C overnight. After 5 washes with saline containing 0.05% Tween 20, non-specific binding sites were blocked by Fetal Calf Serum (FCS), 10% in saline Tween (100 μl/well) at r.t. for 60 minutes. Sera diluted from 1:100 to 1:100.000 were applied at 4 °C for 16 h in saline Tween 10% FCS. After 5 washes, we added 100 μl/well of alkaline phosphatase conjugated anti human IgM or IgG Fab2-specific affinity purified antibodies (Sigma, St. Louis, Missouri, USA) diluted 1:500 in saline Tween/FCS. After an overnight incubation and 5 washes, 100 μl of substrate solution consisting of 2 mg/mL *p*-nitrophenylphosphate (Sigma, St. Louis, Missouri, USA) in 10% diethanolamine buffer was applied. After 30 minutes, the reaction was blocked with 50 μl of 1 M NaOH and the absorbance read in a multichannel ELISA reader (SUNRISE, TECAN, Austria) at 405 nm. ELISA plates, coating conditions, reagent dilutions, buffers, and incubation times were tested in preliminary experiments. Each serum was individually titrated to check for parallelism of antibody absorbances in dilutions. Within-assays and between-assays coefficient of variations were below 10%. Subclass specific anti-IgG conjugates (IgG1, IgG2, IgG3, IgG4, Southern Biotech, Birmingham, AL, USA) were used to detect the IgG-antibodies subclasses in

306 Loomans, E.E., Gribnau, T.C., Bloemers, H.P., Schielen, W.J. (1998) *J. Immunol. Methods,* **221**, 119.

parallel experiments. The antibody levels revealed by SP-ELISA are expressed both as antibody titer (sample dilution which reaches the average plus three standard deviations of blanks) and as absorbance value at a dilution of 1:100 as a ratio of positive controls in the same experiment. Positive samples were analysed twice to evaluate the differences between the two determinations. The references values were set as the mean + 2 SD of the control groups. Within- and between-assays coefficients of variations were below 10%.

5.4.2 *Coating e blocking optimization*

The ELISA plats were coated in with 10 and 20 μg/mL of peptides dilloved in PBS buffer (pH 7.2) and carbonate buffer (pH 9.6). Each plate was then treated with two kind of blocking agents BSA 3% and FBS 5% in PBS buffer. A reduced number of helthy blood donors and patient sera was tested under this conditions as described in section 5.4.1. The best coating/blocking conditions for each peptide were those that lowered the most the Abs values for helthy blood donors compare to the patient sera

5.5 Protein synthesis by Chemical Ligation

5.5.1 *Synthesis of the Bio-PEG-Spacer-hCD59(1-77) fragments for NCL*

The seven fragments for the tandem NCL assembly of Bio-PEG-Spacer-hCD59(1-77) (**XVI**) glycated and wild type (**XVII**) (Table 5.5-1) were synthesized according to the procedures described in chapter 5.6. The C terminal fragment **IX** was obtained using Wang resin and left with the N-terminal Cysteine unprotected.

Fragments from **X** to **XV** were synthesized on 2-Cl-Trityl resin which allowed a diluted acid cleavage of the peptides with all the positions protected but the C-terminal carboxylic function. All those sequences have an N-terminal Cysteine masked as a thiazolidine except product **XV** whose N-terminal residue was coupled to a biotin-PEG_{3500} residue. A glycated Lysine residue was in position 41 of peptide **X** using the glycated Fmoc Lysine **2**. After cleavage the crude was coupled with an excess of p-acetamidothiophenol/HOBt/DIPEA (2:4:4) in acetonitrile to afford the corresponding fully protected C-terminal p-acetamidothioester peptides (Thz-peptide-COSR) in good yields. The crudes were purified by direct phase flash chromatography and then deprotected of the side chain protections with a standard concentrated TFA cleavage affording the corresponding C-ter thioester/N-ter thiazolidine peptides free of side chain protections that were purified by reverse phase flash chromatography.

IX	H-^{45}CNFNDVTTRLRENELTYYCCKKDLCNFNEQLE^{77}N-OH
X	Thz-^{39}CWK*FE^{44}H-COSR
XI	Thz-^{39}CWKFE^{44}H-COSR
XII	Thz-^{26}CLITKAGLQVYN^{38}K-COSR
XIII	Thz-^{13}CKTAVNCSSDFD^{25}A-COSR
XIV	Thz-^{3}CYNCPNPTA^{12}D-COSR
XV	BioPEG$_{3500}$-GGSSGIEGRIEGR^{1}LQ-COSR

Table 5.5-1 Fragments for fragments for the tandem NCL assembly of Bio-PEG-Spacer-hCD59(1-77) (**XVI**) glycated and wild type (**XVII**).

5.6 Solid Phase Peptide Synthesis

5.6.1 *General procedure for in batch and manual SPPS*

Peptides were synthesized on a manual batch synthesizer (PLS 4×4, Advanced ChemTech) using a Teflon reactor (10 mL), following the Fmoc/tBu SPPS procedure. The resin was swelled with DMF (1 mL/100 mg of resin) for 20 min before use.

Peptide synthesis was performed repeating the cycle described as following for each amino acid:

- Swelling: DMF (1 mL/100 mg of resin) for 5 min
- Fmoc-deprotection: resin is washed twice with 20% piperidine in DMF (1 mL/100 mg of resin, one wash for 5 min followed by an other wash for 20 min);
- resin-washings: DMF (3×5 min);
- coupling: scale employing TBTU/HOBt/NMM (2.5eq.:2.5 eq.:3.5 eq.) as the coupling system and 2.5 eq. of the Fmoc protected amino acids, except for Xaa and Yaa (1.5 eq.), in DMF (1 mL/100 mg of

resin) for 40 min. Each coupling was checked by Kaiser test;[307] in our case all tests were negative therefore it was not necessary to repeat the coupling reaction.

– resin-washings: DMF (3×5 min) and DCM (1×5 min).

Kaiser test procedure: to a small amount of peptide-resin placed in a test tube, three drops for each of the following solutions were added: ninhydrin (5 g) in ethanol (100 mL); phenol (80 g) in ethanol (20 mL); KCN (2 mL of 1 mM aqueous solution) in pyridine (98 mL). The tube is heated at 100 °C for 5 min. A positive test (resin beads and solution appears strongly blue-violet) states the presence of at least 5% free amino groups.

5.6.2 *General procedure for peptide acetylation*

The amine functions of the N-terminal fragment of peptides were acetylated in 2 cycles: the first one in 30 minutes followed by the second one in 1.5 hour using Ac_2O/ NMM as reagents. The reaction was monitored by Kaiser test.

5.6.3 *General procedure of deprotection, cleavage and purification of free peptide*

Peptides cleavage from the resin and deprotection of the amino-acids side chains were carried out with TFA/anisole/1,2-ethanedithiol/phenol/H_2O solution (94:2:2:2:2 v/v/v/v/v). The cleavage was maintained for 3 h with vigorous stirring at room temperature. Resins were filtrated and washed with TFA. After partial evaporation under nitrogen flux, filtrates were precipitated

307 Kaiser, E., Colescott, R.L., Bossinger, C.D., Cook, P.I. (1970) *Anal. Biochem.*, ***34***, 595.

with cold diethyl ether, collected by centrifugation, dissolved in H_2O and lyophilized with an Edwards apparatus, model Modulyo.

Peptides were purified by semipreparative RP-HPLC using methods and solvent system as reported. Fractions were checked with UPLC-ESIMS.

5.6.4 *General procedure for solid-phase extraction SPE*

SPE are performed on RP-C18 LiChroprep columns. Main steps are reported here:

Column washings with MeOH (3 column volumes) and CH_3CN (3 column volumes)

Column conditioning with H_2O (3 column volumes)

Dissolving the peptide in H_2O (1 column volume), checking the pH that should be neutral.

Adsorbing peptide solution on silica for 3 times

Eluting with H_2O (3 column volumes)

Eluting with 5%, 10%, 15%, 20% of H_2O/CH_3CN (column volume for each concentration), and 100% of CH_3CN.

Fractions were checked by analytical UPLC-ESIMS, and then lyophilized.

6 XPERIMENTAL PART B

6.1 Synthesis of N^{α}-Fmoc-ω-azido-α-amino acids

6.1.1 *Synthesis of N^{α}-Fmoc-ε-azido-Norleucine-OH*

Tf_2O (1.35 mL, 8.13 mmol) was added dropwise to a vigorously stirred mixture of NaN_3 (2.635 g, 40.5 mmol) in H_2O (6.5 mL) and CH_2Cl_2 (11 mL) at 0 °C. The resulting mixture was allowed to warm to room temperature and stirring was continued for 2 h. The water layer was extracted with CH_2Cl_2 (2 × 4 mL) and the combined organic layers were washed with saturated aqueous Na_2CO_3 (12.5 mL). The resulting solution of TfN_3 in CH_2Cl_2 was then slowly added to a solution of N^{α}-Boc-lysine (1.0 g, 4.06 mmol), K_2CO_3 (0.84 g, 6.08 mmol), and $CuSO_4 \cdot 5H_2O$ (0.01 g, 0.04 mmol) in H_2O (13 mL) and MeOH (27 mL). The mixture was stirred overnight and the reaction was checked by TLC (*i*PrOH—AcOEt—H_2O 6/1/3, revealed with (a) and (b), R_f 0.81). The organic solvents were evaporated under reduced pressure. The water layer was acidified to pH 6 with concd HCl, diluted with 0.25 M of phosphate buffer at pH 6.2 (25 mL), and extracted with CH_2Cl_2 (4 × 50 mL). The organic layers were washed with brine (25 mL), dried over Na_2SO_4, and concentrated under vacuum. The colorless oil was purified using column of RP-18 LiChroprep by solutions of different concentrations of H_2O/CH_3CN to afford 6-azido-Boc-L-norleucine (0.451 g, 41%).

Cleavage of the Boc protecting group of 6-azido-Boc-L-norleucine (1.079 g, 3.96 mmol) was achieved by treatment with an excess of concd HCl (2.5 mL) at room temperature for 6 h. The residue was dissolved in water (5.0 mL) and lyophilized. A solution of 2,5-dioxo-1-pyrrolidinyl 9*H*-fluoren-9-ylmethyl carbonate (Fmoc-OSu; 1.469 g, 4.36 mmol) in dioxane (20 mL) was then added dropwise to a solution of the deprotected amino acid in

dioxane (30 mL). A solution of 1 M NaOH was subsequently slowly added to pH 8–9 and the reaction mixture was stirred at room temperature for 1 h. The reaction was checked by TLC (CH_2Cl_2—MeOH 9/2, UV, R_f 0.58). Water (12 mL) was added and the solution was acidified with 2 M HCl until pH 3. The product was extracted with CH_2Cl_2 (3 × 30 mL), dried with anhyd Na_2SO_4, and the solvent removed under vacuum. The crude material was purified by flash chromatography on silica gel (CH_2Cl_2 to CH_2Cl_2—MeOH 10/1) to obtain the pure 6-azido-Fmoc-L-norleucine **3** (496 mg, 32%) as a yellow oil. RP-UPLC: R_t 1.51 min (50 to 100% of B in 3 min). IR (KBr): 2100 cm^{-1} (N_3). ESI-MS: *m/z* calcd for $C_{21}H_{22}N_4NaO_4$ $[M + Na]^+$: 417.15; found 417.2. $[\alpha]_D$ –2.5 (*c* 1.0, MeOH). ^1H NMR ($CDCl_3$, 400 MHz,): δ 7.74 (d, 2H, $J_{3,4} = J_{5,6}$ = 7.4 Hz, fluorenyl 4-H and 5-H), 7.54 (d, 2H, $J_{1,2} = J_{7,8}$ = 7.4 Hz, fluorenyl 1-H and 8-H), 7.37 (*pseudo* t, 2H, fluorenyl 3-H and 6-H), 7.28 (*pseudo* t, 2H, fluorenyl 2-H and 7-H), 6.19 (broad s, COOH), 5.46 (m, 1H, NH), 4.49–4.33 (m, 3H, CH_2–O and α-H), 4.18 (t, 1H, *J* = 6.4 Hz, fluorenyl 9-H), 3.24–3.21 (m, 2H, ε-H_2), 1.70–1.42 (m, 6H, 3 × CH_2). ^{13}C NMR ($CDCl_3$, 100 MHz): δ 176.97 (COOH), 156.35 (CONH), 143.75, 143.60, and 141.28 (fluorenyl C-4a, C-4b, C-8a, and C-9a), 127.74, 127.06, and 125.01 (fluorenyl C-2 to C-7), 120.00 (fluorenyl C-1 and C-8), 67.14 (CH_2–O), 53.90 (C-α), 51.02 (C-ε), 47.07(fluorenyl C-9), 31.68 (CH_2), 28.31 (CH_2), 22.55 (CH_2). Anal. Calcd for $C_{21}H_{22}N_4O_4$: C, 63.95; H, 5.62; N, 14.20. Found: C, 64.01; H, 5.58; N, 14.23.

6.1.2 *Synthesis of N^{α}-Fmoc-δ-azido-Norvaline-OH*

Cleavage of the Boc protecting group of 5-azido-Boc-L-norvaline (prepared from Boc-L-Orn-OH as in 6.1.1) (1.3 g, 5.01 mmol) was achieved by treatment with an excess of concd HCl (6 mL) at room temperature for 6 h. The residue was dissolved in water and lyophilized. A solution of Fmoc-OSu

(1.88 g, 4.9 mmol) in dioxane (40 mL) was added dropwise to a solution of the deprotected amino acid in dioxane (20 mL). A solution of 1 M NaOH was subsequently slowly added until pH 8 and the reaction mixture stirred at room temperature for 3 h. The reaction was checked by TLC (CH_2Cl_2—MeOH 9/2, UV, R_f 0.58). Water (30 mL) was added and the solution was acidified with 2 M HCl until pH 3. The product was extracted with CH_2Cl_2 (3 × 50 mL), dried over Na_2SO_4, and the solvent removed under vacuum. The crude was purified by flash chromatography on silica gel (CH_2Cl_2 to CH_2Cl_2—MeOH 10/1) to obtain the pure 5-azido-Fmoc-L-norvaline **4** (1.3 g, 69%) as a yellow oil. RP-UPLC: R_t 1.37 min (50 to 100% of B in 3 min). IR (KBr): 2100 (N_3) cm^{-1}. ESI-MS: *m/z* calcd for $C_{20}H_{20}N_4NaO_4$ $[M + Na]^+$ 403.14; found 403.3. $[\alpha]_D$ –2.3 (*c* 1.0, MeOH). ^1H NMR ($CDCl_3$, 400 MHz,): δ 7.76 (d, 2H, $J_{3,4} = J_{5,6}$ = 7.6 Hz, fluorenyl 4-H and 5-H), 7.61 (*pseudo* d, 2H, $J_{1,2} = J_{7,8}$ = 7.6 Hz, fluorenyl 1-H and 8-H), 7.40 (*pseudo* t, 2H, fluorenyl 3-H and 6-H), 7.31 (*pseudo* t, 2H, fluorenyl 2-H and 7-H), 6.16 (broad s, COOH), 5.34 (m, 1H, NH), 4.45–4.40 (m, 3H, CH_2–O and α-H), 4.22 (t, 1H, *J* = 6.6 Hz, fluorenyl 9-H), 3.37–3.30 (m, 2H, δ-H_2), 2.01–1.46 (m, 4H, 2 × CH_2). ^{13}C NMR ($CDCl_3$, 100 MHz): δ 175.72 (COOH), 156.72 (CONH), 143.75, 143.57, and 141.33 (fluorenyl C-4a, C-4b, C-8a, and C-9a), 127.76, 127.08, and 125.00 (fluorenyl C-2 to C-7), 120.02 (fluorenyl C-1 and C-8), 67.12 (CH_2–O), 53.16 (C-α), 50.76 (C-δ), 47.15 (fluorenyl C-9), 29.62 (CH_2), 24.81 (CH_2). Anal. Calcd for $C_{20}H_{20}N_4O_4$: C, 63.15; H, 5.30; N, 14.73. Found: C, 63.09; H, 5.25; N, 14.80.

6.1.3 *Synthesis of N^{α}-Fmoc-γ-azido-hSerine-OH*

Cleavage of the Boc protecting group of Boc-Abu(γ-N_3)-OH (obtained from Boc-Abu(γ-N_3)-OH as in 6.1.1) (0.748 g, 2.9 mmol) was achieved by treatment with an excess of TFA (10 mL) at room temperature for 10 min. The reaction was checked by TLC (AcOEt—*n*-hexane 1/1; revealed with (a), R_f 0.10). TFA was removed by flushing with N_2 and the residue dissolved in water and lyophilized. The methyl ester was hydrolyzed by stirring with 1 M NaOH (5 mL) at room temperature for 6 h. The solution was then treated with concd HCl to pH 7 and lyophilized to afford the free amino acid. A solution of Fmoc-OSu (0.843 g, 2.5 mmol) in dioxane (6 mL) was added dropwise to a solution of the deprotected amino acid in dioxane (10 mL). A solution of 1 M NaOH was subsequently slowly added until pH 8 and the reaction mixture stirred at room temperature for 3 h. The reaction was checked by TLC (CH_2Cl_2—MeOH 9/2, UV, R_f 0.58). Water (7.5 mL) was added and the solution was acidified with concd HCl until pH 3. The product was extracted with CH_2Cl_2 (3 × 20 mL), dried over Na_2SO_4, and the solvent removed under vacuum. The crude was purified by flash chromatography on silica gel (CH_2Cl_2 to CH_2Cl_2—MeOH 10/1) to obtain the pure N^{α}-Fmoc-γ-azido-hSerine-OH **5** (124 mg, 27%) as yellow oil. RP-UPLC: R_t 1.31 min (50 to 100% of B in 3 min). IR: 2100 cm^{-1} (N_3). ESI-MS: *m/z* calcd for $C_{19}H_{18}N_4NaO_4$ $[M + Na]^+$ 389.12; found 389.4. $[\alpha]_D$ –11.5 (*c* 1.0, MeOH). ^1H NMR ($CDCl_3$, 400 MHz,): δ 7.75 (*pseudo* d, 2H, *J* = 7.6 Hz, fluorenyl 4-H and 5-H), 7.54 (*pseudo* d, 2H, *J* = 7.4 Hz, fluorenyl 1-H and 8-H), 7.39 (*pseudo* t, 2H, fluorenyl 3-H and 6-H), 7.31 (*pseudo* t, 2H, fluorenyl 2-H and 7-H), 6.14 (broad s, COOH), 5.63 (m, 1H, NH), 4.53–4.41 (m, 3H, CH_2–O and α-H), 4.21 (t, 1H, *J* = 6.8 Hz, fluorenyl 9-H), 3.42–3.39 (m, 2H, γ-H_2), 2.19–1.96 (m, 6H, 3 × CH_2). ^{13}C NMR ($CDCl_3$, 100 MHz): δ 172.71 (COOH), 156.26 (CONH), 143.53 and 141.29 (fluorenyl C-4a, C-4b, C-8a,

and C-9a), 127.76, 127.08, 125.04, 124.99 (fluorenyl C-2 to C-7), 120.00 (fluorenyl C-1 and C-8), 67.17 (CH_2-O), 51.70 (C-α), 47.68 (C-γ), 47.09 (fluorenyl C-9), 31.21 (CH_2). Anal. Calcd for $C_{19}H_{18}N_4O_4$: C, 62.29; H, 4.95; N, 15.29. Found: C, 62.36; H, 4.99; N, 15.24.

6.2 Synthesis of N^{α}-Fmoc-ω-alkynyl-α-amino acids

6.2.1 *General procedure for the synthesis of p-toluenesulfonate derivatives*

p-Toluenesulfonyl chloride (6.49 mmol) in pyridine (10 mL) was added to a stirred solution of the alcohol derivative (5.9 mmol) at 0 °C. The reaction was stirred for 4 h at rt and checked by TLC [AcOEt—*n*-hexane 5/4; revealed with (a), (e)]. The reaction mixture was neutralized with 2 M HCl, the *p*-toluensulfonate derivative extracted with $CHCl_3$, and the organic layer evaporated under vacuum to afford the product as a pale yellow oil.

5-Hexyn-1-ol 1-(4-methylbenzene sulfonate).(**6**) Yield 96%. ^{1}H-NMR ($CDCl_3$, 200 MHz): δ 7.78 and 7.34 (AA'BB' system, 4H, *J* = 8.4 Hz, MeC_6H_4), 3.57 (t, 2H, *J* = 6.2 Hz, 1-H_2), 2.45 (s, 3H, MeC_6H_4), 2.13–2.09 (m, 2H, 4-H_2), 1.90 (t, 1H, 6-H), 1.93–1.72 (m, 4H, 2-H_2 and 3-H_2).

4-Pentyn-1-ol 1-(4-methylbenzene sulfonate) (**7**) Yield 93%. ^{1}H-NMR ($CDCl_3$, 200MHz): δ 7.77 and 7.35 (AA'BB' system, 4H, *J* = 8.4 Hz, MeC_6H_4), 4.12 (t, 2H, *J* = 5.8 Hz, 1-H_2), 2.43 (s, 3H, MeC_6H_4), 2.32–2.15 (m, 2H, 3-H_2), 1.93 (t, 1H, 5-H), 1.91–1.82 (m, 2H, 2-H_2).

6.2.2 *General procedure for the synthesis of bromo derivatives*

Lithium bormide (6.67 mmol) was added in portions to a stirred solution of the *p*-toluensulfonate derivative (4.45 mmol) in acetone (10 mL) at rt. The reaction mixture was heated at 40 °C for 20 h. The reaction was checked by TLC (AcOEt—hexane 5/4; revealed with (e), R_f 0.75). The solvent was evaporated by flushing with N_2. The residue was treated with *n*-hexane, filtration of the residue and evaporation of DCM under nitrogen flux, afforded the bromo derivative as a yellow oil.

6-Bromohex-1-yne. (**9**) Yield 31%. ^{1}H NMR ($CDCl_3$, 200 MHz): δ 3.41 (t, 2H, J = 7.2 Hz, 6-H_2), 2.21 (dt, 2H, J = 3.0 Hz, 7.4 Hz, 4-H_2), 2.02–1.91 (m, 2H, 3-H_2), 1.96 (t, 1H, J = 3.8 Hz, 1-H), 1.71-1.60 (m, 2H, 5-H_2) .

5-Bromopent-1-yne (**10**) Yield 28%. ^{1}H NMR ($CDCl_3$, 200 MHz): δ 3.44 (t, 2H, J = 6.4 Hz, 5-H_2), 2.75 (dt, 2H, J = 2.6 Hz, 7.0 Hz, 3-H_2), 2.02-1.95 (m, 2H, 4-H_2), 1.86 (t, 1H, J = 3.6 Hz, 1-H).

6.2.3 *Synthesis of the Chiral Inductor BPB*

A solution of 10 g (86,8 mmol) of (S)-proline and 18,5 g (330 mmol) of KOH in 70 mL of iPrOH was prepared with stirring at 40°C. As soon as the solution became transparent, slow addition of freshly distilled BnCl (18,30 g, 130 mmol) was added under stirring at the same temperature for 6h. The reaction mixture was neutralized with concentred aqueous HCl until pH 5-6 (indicator paper), then was added to the reaction mixture $CHCl_3$ (30 mL) with stirring. The mixture was left overnight, then filtered and the precipate was washed with $CHCl_3$. The $CHCl_3$ solutions were combined and evapored, the residue was treated with cold acetone and the precipitate of crude BP

filtered and additionally washed with acetone. Some BP was also recovered from the acetone washings. The crude material was dried in air and then over P_2O_5 in vacuo to give 10,23 g of (*S*)-*N*-benzylproline (BP) (**12**) (57 %).

Product **12** (3.467g, 16.9 mmol) was added at RT under N_2 to a stirred freshly-prepared transparent solution of PCl_5 (7.035g, 33.8 mmol) in anhyd CH_2Cl_2 (55 mL). After 30 min cold petroleum ether was added and the acyl chloride precipitated as an oil. The oil was dissolved in anhyd CH_2Cl_2 (60 mL) under N_2 and 2-aminobenzophenone (3.33 g, 16.9 mmol) was added in one portion, followed by Et_3N to pH 8. The mixture was stirred for 4 h at rt, then washed with a saturated solution of Na_2CO_3 and twice with H_2O. The organic layer was evaporated under reduced pressure. The crude BPB was recrystallized from EtOH. Some product was also recovered from the EtOH washings. The material was dried under vacuum over P_2O_5 to give the chiral inductor (*S*)-2-(*N*-Benzylprolyl)aminobenzophenone (BPB) (**13**) (1.873 g, 29%). ESI-MS: *m/z* calcd for $C_{25}H_{24}N_2O_2$ $[M + H]^+$: 385,18; found 385,2. ^1H NMR ($CDCl_3$, 400 MHz): δ 11.52 (s, 1H, NH), 8.56 (d, 1H, *J*= 8.4 Hz, Bn), 7.79-7.36 (m, 9H, Bn), 7.15 (m, 4H, Bn), δ_A= 3.92, δ_B= 3.59 (syst AB, 2H, J_{AB}= 12.8 Hz, CH_2Bn), 3.32 (dd, 1H, $J_{\alpha,\beta}$= 4.4 Hz, $J_{\alpha,\beta'}$=10.0 Hz, H_α), 3.22 (dd, 1H, $J_{\delta,\delta}$= $J_{\delta,\gamma}$= 6.4 Hz, H_δ), 2.41 (dd, 1H, $J_{\beta,\beta}$= 8.8 Hz, $J_{\beta,\gamma}$= 16 Hz, $H_{\beta'}$), 2.26 (ddd, 1H, $J_{\delta\delta}$= 6.4 Hz, $J_{\delta\gamma}$=12.8 Hz, $J_{\delta\gamma}$= 22 Hz $H_{\delta'}$), 1.96 (ddd, 1H, $J_{\beta,\beta}$= 8.8 Hz, $J_{\beta,\gamma}$= 4.4 Hz, $J_{\beta\gamma}$= 16.4 Hz, H_β), 1.85-1.76 (m, 2H, $H_{,\gamma}$,$H_{,\gamma'}$). ^{13}C NMR ($CDCl_3$, 100 MHz): δ 198.03 (Ph-CO-Ph), 174.64 (COOH), 139.16, 138.54, 138.12, 133.37, 132.55, 132.48, 130.11, 129.12, 128.30, 128.15, 127.05, 125.32, 122.19, 121.52 (18 Ar), 68.25 (C_α), 59.82 (CH_2Bn), 53.85 (C_δ), 30.98 (C_β), 24.14 (C_γ).

6.2.4 *Synthesis of the [Gly-Ni-BPB] complex*

A solution of KOH 1,02 g (18,2 mmol) in 4 mL of MeOH was poured into a variously stirred mixture of BPB 1 g (**13**) (2,6 mmol), $Ni(NO_3)_2.6H_2O$ 1,511g (5,2 mmol), glycine 0,975 g (13 mmol) in 9,1 mL of MeOH under N_2 at 40-50°C. The resulting mixture was stirred at 55-65°C for 1h (a prolonged heating of the reaction mixture might result to a partial racemization of the BPB moiety), and then neutralized with 1mL of AcOH diluted in 39 mL of water. The separated crystalline solid was filtered and washed with water. The crude material was dried in air and then over P_2O_5 in vacuo to give 1,06 g of product **14** (81 %). ESI-MS: *m/z* calcd for $C_{27}H_{25}N_3NiO_3$ $[M + H]^+$: 498,20 ; found 498,3. ^{1}H-NMR ($CDCl_3$, 200MHz): δ 8.27 (d, 1H, *J*= 8.8 Hz, Bn), 8.07 (d, 1H, *J*= 7.2 Hz, Bn), 7.51-7.30 (m, 6H, Bn), 7.24-6.69 (m, 6H, Bn), δ_A= 4.48, δ_B= 3.69 (syst AB, 2H, J_{AB}= 6.2 Hz, CH_2Bn), 3.51-3.31 (m, 2H, $H_{\alpha pro}$), 2.63-2.41 (m, 3H, $H_{\delta pro}$, $H_{\alpha\alpha'Gly}$), 2.13-1.91 (m, 4H, $H_{\beta'pro}$, $H_{\delta'Pro}$, $H_{\beta Pro}$, $H_{,\gamma Pro}$, $H_{,\gamma'Pro}$). ^{13}C-NMR ($CDCl_3$, 50MHz): δ 181.19 (CO_{Gly}), 177.18 (CO-N-Ph), 171.49 (C=N), 134.53, 133.11, 132.16, 131.66, 129.67, 129.53, 129.28, 129.06, 128.87, 126.19, 125.60, 125.12, 124.21, 120.82 (18 Ar), 69.96 ($C_{\alpha Pro}$), 63.23 (CH_2Bn), 61.38 ($C_{\alpha Gly}$), 57.60 ($C_{\delta Pro}$), 30.90 ($C_{\beta Pro}$), 23.89 ($C_{\gamma Pro}$).

6.2.5 *General procedure for the alkylation of the Gly-Ni-BPB complex with bromoalkynes*

To a stirred mixture of Gly-Ni-BPB (**14**) (1.99 g, 4 mmol) in anhyd CH_3CN (17.5 mL) were added, under N_2, finely powdered NaOH (0.4 g, 10 mmol) and bromoalkyne (6.01 mmol). After 5 h, the reaction mixture was treated with 0.1 M HCl (59 ml) and the red product extracted with CH_2Cl_2 (4 × 40 mL), dried over $MgSO_4$, and the solvent removed under vacuum. The crude

was purified by flash chromatography on silica gel (CH_2Cl_2—Me_2CO 2/1) affording the product as a red amorphous solid.

6.2.6 *General Procedure for the hydrolysis of the alkylated complexes and Fmoc protection of the free amino acid.*

A solution of the alkylated complex (1.33 mmol) in MeOH (22.5 mL) was added to warm 2 M HCl (16 mL) and the mixture refluxed for 1 h. After cooling to rt, 1 M NaOH was added until pH 6 and the solvent removed under vacuum. The solid residue was washed with acetone, the dried solid product was dissolved in MeOH—H_2O 15/20 (70 mL) and then gently swirled overnight with Chelex 100 H^+ resin, converted from its Na^+ form. The mixture was filtered and the resin washed with water, the layers of combined filtrates were evaporated under vacuum, and the residue lyophilized.

A solution of FmocOSu (1.51 mmol) in dioxane (15 mL) was added dropwise to the lyophilized product (1.37 mmol) dissolved in dioxane (15 mL) and then 1M NaOH was added until pH 8. The reaction mixture was stirred at RT for 4 h, after which time, water (7.5 mL) was added and the solution acidified with 2 M HCl to pH 3. The product was then extracted with CH_2Cl_2 (3 × 20 mL), dried over Na_2SO_4 and the solvent removed under vacuum. The crude was purified by FCC (CH_2Cl_2 to CH_2Cl_2—MeOH 10/1) to obtain the pure amino acid as a yellow oil.

N^{α}-Fmoc-ε-alkynyl-Norleucine-OH (**18**) Yield 31%. RP-UPLC: R_t 1.49 min (50–100% of B in 3 min). $[\alpha]_D$ –3.1 (*c* 1.0, MeOH). ESI-MS: *m/z* calcd for $C_{23}H_{23}NNaO_4$ $[M + Na]^+$ 400.15; found 400.3. ^1H NMR ($CDCl_3$, 400 MHz): δ 7.75 (*pseudo* d, 2H, *J* = 7.6 Hz, fluorenyl 4-H and 5-H), 7.59 (*pseudo* d, 2H, *J* = 7.6 Hz, fluorenyl 1-H and 8-H), 7.37 (*pseudo* t, 2H, fluorenyl 3-H

and 6-H), 7.28 (*pseudo* t, 2H, fluorenyl 3-H and 6-H), 5.79 (broad s, COOH), 5.48(m, 1H, NH), 4.44–4.38 (m, 3H, CH_2–O and α-H), 4.21 (t, 1H, *J* = 6.8 Hz, fluorenyl 9-H), 2.08–1.99 (m, 3H), 1.94 (t, 1H, *J* = 2.4 Hz, HC≡C), 1.80–1.75 (m, 1H), 1.58–1.42 (m, 4H, 2 × CH_2). ^{13}C NMR ($CDCl_3$, 100 MHz): δ 176.63 (COOH), 156.17 (CONH), 143.83, 143.67 and 141.29 (fluorenyl C-4a, C-4b, C-8a, and C-9a), 127.71, 127.06, 125.04 (fluorenyl C-2 to C-7), 119.98 (fluorenyl C-1 and C-8), 83.97 (HC≡*C*), 68.69 (CH_2-O), 67.06 (H*C*≡C), 53.83 (C-α), 47.15 (fluorenyl C-9), 31.73 and 27.81 (C-β and δ), 24.31 (C-γ), 18.15 (C-ε). Anal. Calcd for $C_{23}H_{23}NO_4$: C, 73.19; H, 6.14; N, 3.71. Found: C, 73.09; H, 6.19; N, 3.81.

N^{α}-Fmoc-δ-alkynyl-Norvaline-OH (**19**) Yield 28%. RP-UPLC: R_t 1.49 min (50–100% of B in 3 min). $[\alpha]_D$ –3.0 (*c* 1.0, MeOH). ESI-MS: *m/z* calcd for $C_{22}H_{21}NNaO_4$ $[M + Na]^+$ 386.14; found 386.2. ^{1}H NMR ($CDCl_3$, 400 MHz): δ 7.73 (d, 2H, *J* = 7.2 Hz, fluorenyl 4-H and 5-H), 7.57 (d, 2H, *J* = 7.4 Hz, fluorenyl 1-H and 8-H), 7.39 (*pseudo* t, 2H, fluorenyl 3-H and 6-H), 7.30 (*pseudo* t, 2H, fluorenyl 2-H and 7-H), 6.60 (broad s, COOH), 5.51 (m, 1H, NH), 4.43–4.35 (m, 3H, CH_2–O and α-H), 4.18 (t, 1H, *J* = 6.6 Hz, fluorenyl 9-H), 2.08–1.99 (m, 3H), 1.94 (t, 1H, *J* = 2.4 Hz, HC≡C), 1.80–1.75 (m, 1H), 1.58–1.42 (m, 2H, CH_2). ^{13}C NMR ($CDCl_3$, 100 MHz): δ 177.06 (COOH), 156.26 (CONH), 143.81, 143.62 and 141.27 (fluorenyl C-4a, C-4b, C-8a, and C-9a), 127.70, 127.06 and 125.05 (fluorenyl C-2 to C-7), 119.96 (fluorenyl C-1 and C-8), 83.49 (HC≡*C*), 69.11 (CH_2-O), 67.06 (H*C*≡C), 53.77 (C-α), 47.11 (fluorenyl C-9), 31.33 (C-β), 24.28 (C-γ), 18.01 (C-δ). Anal. Calcd for $C_{22}H_{21}NO_4$: C, 72.71; H, 5.82; N, 3.85. Found: C, 72.80; H, 5.87; N, 3.80.

.

6.3 General procedure for the synthesis of clicked peptides

Linear eIF4E b.p peptide analogs for click chemistry, **XVI-XX**, were synthesize and purify according to the procedures described in chapter 5.6. The modified amino acids carrying the alkynyl and azido modifications where introduced on the peptide sequences by manual SPPS using a low excess (1.1-1.2) under high concentration of activating agents (4-5 fold) and long coupling time (1.5 h).

For click cyclization the linear pure peptides (3.1 μmol), $CuSO_4.5H_2O$ (43,4 μmol), and ascorbic acid (40.3 umol) where dissolved in 4 mL of H_2O/tBuOH 2:1. The mixture was stirred at room temperature overnight and the solution was concentrated and lyophilized. To remove copper derivatives from the crude and to purify the clicked peptides, a SPE purification was performed followed by semi-preparative HPLC purification to afford pure clicked peptide (97% purity).

The cyclic peptides **XVIc-XXc** were characterization by UPLC-MS. The analytical data are reported in Table 3.5-6 .

6.4 Circular dichroism spectrometry

All CD spectra were recorded on a Jasco J-810 spectropolarimeter using cells of 1 mm path length. The pH of the samples was adjusted to 6.6 with aqueous phosphate buffer. After pH adjustment, samples were lyophilized and dissolved in water, or in water containing 50% (v/v) HFA to obtain a final peptide concentration of 0.02 mM. Spectra were the average of ten scans from 190 to 260 nm, recorded with a band width of 0.5 nm at scan rate of 5 nm/min.

6.5 NMR spectrometry for conformational studies

Samples for NMR were prepared by dissolving lactam- and [1,2,3]triazolylcontaining peptides in 0.5 mL of aqueous phosphate buffer (pH 5.5). NMR spectra were recorded on a Bruker DRX-600 spectrometer. One-dimensional (1D) NMR spectra were recorded in the Fourier mode with quadrature detection. The water signal was suppressed by a low-power selective irradiation in the homogenous mode. DQF-COSY,[155] TOCSY,[156] and NOESY[157] experiments were run in the phase sensitive mode using quadrature detection in ω_1 by time-proportional phase increments of the initial pulse.[308] Data block sizes comprised 2048 addresses in t_2 and 512 equidistant t_1 values. Before Fourier transformation, the time domain data matrices were multiplied by shifted $\sin^2$ functions in both dimensions. A mixing time of 70 ms was used for the TOCSY experiments. NOESY experiments were run at 300 K with mixing times in the range of 100-250 ms. The qualitative and quantitative analyses of DQF-COSY, TOCSY and NOESY spectra were obtained using the SPARKY[158] interactive program package.

[308] Marion, D. and Wuthrich, K. (1983) *Biochem. Biophys. Res. Commun.*, **113**, 967.

7 Supplementary material

Rapid Communication

Journal of PeptideScience

Received: 22 October 2008 Revised: 10 November 2008 Accepted: 11 November 2008 Published online in Wiley Interscience: 17 December 2008

(www.interscience.com) DOI 10.1002/psc.1105

Building blocks for the synthesis of post-translationally modified glycated peptides and proteins

Stefano Carganico,[a,b] Paolo Rovero,[a,c] Jose A. Halperin,[d,e] Anna Maria Papini[a,b] and Michael Chorev[d,e]*

Growing interest in synthetic peptides carrying post-trasiational modifications, in general, and the Amadori modification in particular, raises the need for specific building blocks that can be used in stepwise peptide synthesis. Herein, we report the synthesis of N^{α}-Fmoc-Lys-OH derivatives containing N^{ε}-1-deoxyfructopyranosyl moiety. Copyright © 2008 European Peptide Society and John Wiley & Sons, Ltd.

Supporting information may be found in the online version of this article

Keywords: glycation; Amadori rearrangement; Lys(N^{ε}-1-deoxyfructopyranosyl); Amadori-containing building block; site-specific modification; SPPS

Glycation of proteins through non-enzymatic reactions between glucose or other reducing sugars and reactive amino groups represents one of the more abundant processes involved in post-translational modification of proteins [1]. Spontaneous and reversible condensation of a reducing sugar and a free amino group of a protein forms an aldimine also known as the Schiff base that undergoes a rearrangement into the more stable ketoamine known also as the Amadori product [2]. In the case of glucose, the initially formed Schiff base rearranges into the more stable 1-deoxyfructopyranosyl moiety. Subsequent dehydration, condensation, fragmentation, oxidation, and cyclization reactions lead to the irreversible formation of advanced glycation end products (AGEs). This process leads to inactivation of proteins and is involved in pathologies such as senile cataract [3], arteriosclerosis [4], vascular complications of diabetes [5], dysfunction of skin collagen [6], and neurodegenerative diseases such as Alzheimer's disease [7,8] and Parkinson disease [9].

Growing evidence suggests that glycation occurs preferentially at specific glycation motifs characterized by acidic amino acids, mainly glutamate and lysine residues that catalyze the glycation of nearby lysines [10,11]. Proximity to histidine either in the primary or in the secondary structure was also suggested to promote glycation of adjacent lysines [12,13]. Recent interest to fully characterize the glycation products and to use them as biomarkers and antigens for diagnosis and prognosis of disease, monitoring its progress and evaluation of the efficiency of therapy generated the need for glycated peptides representing the glycation motifs specifically modified by the 1-deoxyfructopyranosyl. Today, syntheses of site-specific Amadori-modified peptides are carried out on partially protected synthetic peptides in which only the lysyl residues designated for glycation are exposed while the rest are protected [14–17]. This approach involves orthogonal protection strategies and suffers from low yields and elaborated purification schemes.

Stepwise assembly of site-specific Amadori-modified peptides requires N^{α}-protected-N^{ε}-glycated-Lys building blocks and represents a fully controlled and effective synthetic strategy. Herein, we report the synthesis, purification, and characterization of N^{α}-Fmoc, N^{α}-Boc, N^{ε}-(1-deoxyfructopyranosyl)lysine building blocks needed for Fmoc-based solid phase synthesis of Amadori-modified peptides.

This study offers a controlled side-specific introduction of N^{ε}-Amadori-modified Lys residue into synthetic peptides during a stepwise assembly either in solution or in solid phase methodologies. This strategy will overcome three major problems associated with the modification of already assembled peptides: (i) lack of site-specificity in the introduction of the modification, (ii) need for elaborate orthogonal protection scheme in an effort to achieve site-specificity, and (iii) extremely low yields and complicated reaction mixtures due to side reactions following the direct thermal glycation. Adapting the conditions for generating Amadori peptides by direct thermal glycation in the presence of excess D-glucose [14] to the direct glycation of N^{α}-Fmoc-lysine led to the synthesis of N^{α}-Fmoc-Lys[N^{ε}-1-deoxyfructopyranosyl]-OH (1a) in 67% yield (Scheme 1, pathway A). Preliminary attempt to use

* Correspondence to: Michael Chorev, Laboratory for Translational Research, Harvard Medical School, One Kendall Square, Building 600, Cambridge, MA 02139, USA. E-mail: michael_chorev@hms.harvard.edu

a Laboratory of Peptide and Protein Chemistry and Biology, Polo Scientifico e Tecnologico, University of Firenze, Sesto Fiorentino I-50019, Italy

b Dipartimento di Chimica Organica, Polo Scientifico e Tecnologico, University of Firenze, Via della Lastruccia 13, Sesto Fiorentino I-50019, Italy

c Dipartimento di Scienze Farmaceutiche, University of Firenze, Via Ugo Schiff 3, Polo Scientifico e Tecnologico, Sesto Fiorentino I-50019, Italy

d Department of Medicine, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115, USA

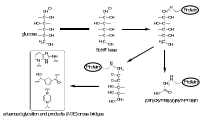
e Laboratory for Translational Research, Harvard Medical School, One Kendall Square, Building 600, Cambridge, MA 02139, USA

Synthetic strategies to Advanced Glycation End products peptides involved in diabetes

S. Carganico[1], M.A. Bonache[1], M.C. Alcaro[1], M. Chelli[1], P. Rovero[1], A.M. Papini[1], M. Chorev[2], A. Lapolla[3], P. Traldi[4],
[1]Laboratory of Peptide & Protein Chemistry & Biology and CNR-ICCOM, Polo Scientifico e Tecnologico, Università di Firenze, Sesto Fiorentino (FI),
[2]Laboratory for Translational Research, Harvard Medical School, Cambridge, Massachusetts, [3]Dipartimento di Scienze Mediche
[4]Laboratorio MS, IBP-CNR, Padova, Italy

Glycation and AGE products

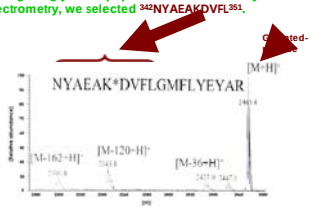
Diabetes complications seem to be related to the Maillard reaction occurring between a free amino group of a protein and glucose or a reducing sugar. Early glycation products undergo further complex reactions to become irreversibly crosslinked heterogeneous derivatives, usually called advanced glycation endproducts (AGEs) [1].



Some AGEs have been structurally characterized, including N-(carboxymethyl)Lysine (Cml), which has been demonstrated to be a major immunological epitope among AGEs [3]. However, these structurally identified AGEs account for only a small percentage of AGE that occur in vivo.

Previous results

Traldi *et al.* reported the analysis of enzymatic digestions of glycated human serum albumin (HSA) by MALDI TOF mass spectrometry [2].
Among the glycated peptides characterized by mass spectrometry, we selected ^{342}NYAEAKDVFL^{351}.



Characterization of AGE-modified synthetic HSA peptide fragments and study of their possible role in Diabetes auto-antibody recognition

Peptide selection

e selected from HSA sequence a peptide fragment known to e glycated at a Lysine residue [1].

^{342}NYAEAKDVF^{351}

Spontaneous glycation of the peptide

vitro incubation with glucose can be performed under ysiological conditions to achieve non enzimatic ontaneous glycation of the peptide and successive GE products formation [6].

Future goals

Autoantibody recognition through ELISA test in diabetes patients' sera using [Cml^{347}]HSA(342-351) as synthetic antigenic probe

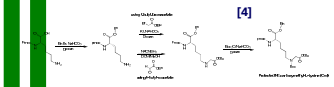
LC-MS monitorig of HSA(342-351) spontaneous glycation

Acknowledgments:
Ente Cassa di Risparmio di Firenze,
FIRB internazionalizzazione 2005 RBIN04TWKN

References:
[1] Y.M. Li, *et al.*, J. Imm. Methods. **1997**, 207, 79.
[2] Traldi *et al.*, J. Mass Spectrom. **2006**, 41(9), 1179.
[3] Reddy S. *et al.*, Biochem. **1995**, 34, 10872.
[4] Hamachi I. *et al.*, Chem. Eur. J. **1999**, 5, 1503.
[5] Gruber P. *et al.*, J. Peptide Res. **2005**, 66, 111.
[6] Wiley J. *et al.*, J. Mass Spectrom. **2003**, 38, 199.

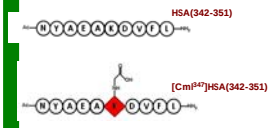
Chemoselective Synthesis of CarboxyMethylLysine (Cml) protected for Fmoc/tBu Solid Phase-Peptide Synthesis (SPPS)

[4]



[5]

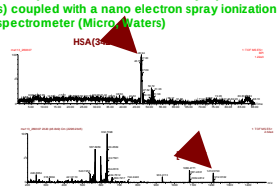
Automatic Fmoc/tBu SPPS of HSA(342-351) and of [Cml^{347}]HSA(342-351)



Multiple automatic synthesizer Apex396 (Aapptec)

LC/ESI Q-TOF MS

HSA(342-351) has been characterized by liquid chromatography (CapLC, Waters) coupled with a nano electron spray ionization Q-TOF mass spectrometer (MicroMass Waters)



CapLC-MicroMS Waters

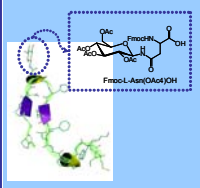
Post translationally modified peptides for an efficient detection of autoantibody biomarkers of autoimmune diseases

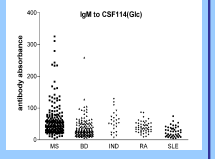
Stefano Carganico[1], Francesca Nuti[1], Elisa Peroni[1], Claudia Alcaro[2], Mario Chelli[1], Paola Migliorini[3], Carlo Selmi[4], Francesco Lolli[5], Michael Chorev[6], Paolo Rovero[7], and Anna Maria Papini[1]

[1]PeptLab c/o Dept of Organic Chemistry, University of Firenze; [2]Toscana Biomarkers S.r.l., Siena; [3]Laboratory of Clinical Immunology, Dept of Internal Medicine, University of Pisa; [4]Div. of Internal Medicine, San Paolo School of Medicine, University of Milano; [5] PeptLab c/o Dept of Neurological Sciences & Azienda Ospedaliera Careggi, University of Firenze; [6]Laboratory for Translational Research, Harvard Medical School, USA; [7] PeptLab c/o Dept of Pharmaceutical Sciences, University of Firenze.

Multiple Sclerosis

The CSF114(Glc)-based not competitive solid phase ELISA enabled the reproducible and effective detection of IgM autoantibodies to this glycopeptide in a significant population (30%) of MS patients. The antibodies detected in the present study were typical of Relapsing-Remitting patients (RR-MS). Therefore, a CSF114(Glc)-based immunoassay on sera has an important prognostic value in monitoring MS disease progression guiding the optimal therapeutic treatment.



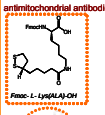


CSF114(Glc) is a mimetic of in vivo auto-antigens triggeringauto-immunoresponse

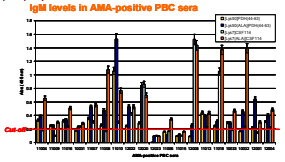
Lolli et al. Proc. Natl. Acad. Sci., U.S.A., (2005), 102, 10273; A.M. Papini Nature Medicine (2005) 11(1), 13; A.M. Papini, et al. Applicant: University of Florence, Italy. PCT International application (2003) WO 03000733 A2.; F. Lolli, et al. The glycopeptide CSF114(Glc) detects serum antibodies in Multiple Sclerosis. J. Neuroimmunol. (2005) 167, 131-137.

PrimaryBiliary Cirrhosis (PBC)

More than 95% of PBC patients have detectable levels of autoantibodies to PDC-E2 [Pyruvate Dehydrogenase Complex (PDC-E2)] and in general these react with a region of the molecule containing a lipoamide lysine residue [Lys(ALA)]. It has been hypothesized that the lipoamide in PDC-E2 serves as a xenobiotic target becoming immunogenic and initiates or perpetuates an antimitochondrial antibodies (AMA) response.



IgM levels in AMA-positive PBC sera



Lys^3(ALA)]CSF114 exposing lipoamide residue on the tip of a β-turn structure is able to detect the best antibody titre in PBC patients' sera

Long S. A., Quan C., Van de Water J., Nantz M. H., Kurth M. J., Barsky D., Colvin M E., Lam K. S., Coppel R. L., Ansari A., and Gershwin M. E., J. Immunol., 2001, 2956-2963.

Our hypothesis on the molecular mechanism of antibody-mediated autoimmune diseases

Changes at the level of protein post-translational modifications (PTMs) may create neoantigens triggering an autoimmune response in which autoantibodies are biomarkers of disease activity.

Post-translationally modified synthetic peptides can be useful tools for mimicking neoantigens responsible of the autoimmune response.

Identification of **disease biomarkers** using a "**Chemical Reverse Approach**" based on **synthetic post-translational modified peptides** for fishing-out **antibodies** as **biomarkers** in patients' biological fluids

CSF114 may be proposed as "*Universal Peptide Scaffold*" because of its β -turn structure, exposing at the best the aberrant PTM specific for antibody recognition in autoimmune diseases

M. C. Alcaro, *et al.*, *Chem.Today*, (2007), *25*, 14; Papini, A.M. *J. Pep Sci* (2009), in press.

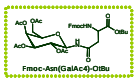
Rheumatoid Arthritis

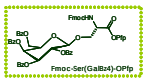
Rheumatoid Arthritis is a common systemic autoimmune disease and it is characterized by inflammation of the synovial membrane of diarthrodial joints. Large number of activated leukocytes infiltrate the synovial membrane leading to progressive destruction of cartilage and bone.

Autoimmune response in RA

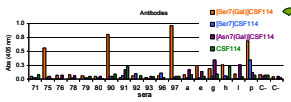
Galactosylated peptides for RA

[Asn^3(Gal)]CSF114	*N*-glycopeptide
[Ser^7(Gal)]CSF114	*O*-glycopeptide





SP-ELISA

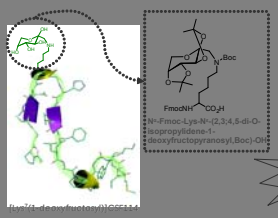


The *O*-glycopeptide [Ser^7(Gal)]CSF114 recognizes antibodies in 30% ca of RA analyzed sera

Papini, A.M. et al. Filing date 16/10/2007. Applicant: Toscana Biomarkers Srl. EPC Application (2007) 07118566.4

Type I diabetes

Glycation of proteins through non-enzymatic reactions between glucose or other reducing sugars and reactive amino groups represents one of the more abundant processes involved in post-translational modification of proteins. Diabetic pastients, because of their characteristic iperglicemia, develop high levels of glycated proteins.



Carganico S, Rovero P, Halperin JA, Papini AM, Chorev M. *J Pept Sci.* (2009) 67-71

OUTLOOK

In conclusion, modified peptides as synthetic probes characterizing families of antibodies in biological fluids are a suitable option for the development of multiple diagnostic/prognostic immunoassays, increasing sensitivity of diagnostics

8 ABBREVIATIONS

Ab: antibody

Ac: acetyl

Ag: antigen

AMBER: assisted model building with energy refinement

ATP: adenosine triphosphate

BD: blood donors

Boc: tert-butoxycarbonyl

BOP : benzotriazole-1-yl-oxy-tris-(dimethylamino)-phosphonium hexafluoro--phosphate

BP: *N*-benzyl-(*S*)-proline

BPB: (*S*)-2-(*N*-benzylprolyl)aminobenzophenone

BSA: bovine serum albumin

cAMP: cyclic adenosine monophosphate

CD: circular dichroism

CDMT : 2-Chloro-4,6-dimethoxy-1,3,5-triazine

CNS: central nervous system

DBU: 1,8-Diazabicycloundec-7-ene

DCM: dichloromethane

DIEA: diethylamine

DMF: *N*,*N*-dimethylformamide

DMTMM-BF_4: 4-(4,6-dimethoxy-1,3,5-triazin-2-yl-)-4-methylmorpholinium tetrafluoroborate

DNA: deoxyribonucleic acid

DQF-COSY: double quantum filtered correlated spectroscopy

EDT: 1,2-ethanedithiole

ELISA: enzyme-linked immunosorbent assay

FCC: flash column chromatography

Fmoc: 9-H-fluoren-9-yl-methoxycarbonyl

Fmoc-OSu: *N*-(9-fluoren-9-yl-methoxycarbonyloxy)-succinimide

GABA: γ-aminobutyric acid

HFA: hexafluoroacetone

HOBt: 1-hydroxybenzotriazole

HPLC: high performance liquid chromatography

IC_{50}: inhibitory concentration of 50

IP: inositol phosphate

IP3: inositol 1,4,5-trisphosphate

MOG: myelin oligodendrocyte glycoprotein

MRI: magnetic resonance imaging

mRNA: messenger ribonucleic acid

MS: multiple sclerosis

MSAP: multiple sclerosis antigenic probe

NMM: *N*-methylmorpholine

NMR: nuclear magnetic resonance

NOE: nuclear Overhauser effect

NOESY: nuclear Overhauser enhancement spectroscopy

ON: overnight

PBS: phosphate buffered saline

Ph: phenyl

PrG: propargylglycine

PTH: parathyroid hormone

PTH1-Rc: human PTH type 1 receptor

PTHrP: parathyroid hormone-related protein

PTM: post-translational modification

RCM: ring-closing metathesis

RP-HPLC: reverse phase-high performance liquid chromatography

Rt: retention time

RT: room temperature

SPE: solid phase extraction

SP-ELISA: solid-phase ELISA

SPPS: solid-phase peptide synthesis

TBTU:2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate

tBu: tert-butyl

TFA: trifluoroacetic acid

TFE: trifluoroethanol

THF: tetrahydrofuran

TLC: thin layer chromatography

TOCSY: total correlated spectroscopy

UPLC: ultra performance liquid chromatogra

Ringraziamenti:

un grazie di cuore a tutti quanti di PeptLab, grazie Anna Maria , grazie Michael e soprattutto grazie Francesca di avermi sopportato!!