Organic & Biomolecular Chemistry

Accepted Manuscript

This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/obc

Replacement of the CysA7-CysB7 disulfide bond with a 1,2,3-triazole linker causes unfolding in insulin glargine

Geoffrey Williams,[a] Kathryn Lee,[b] Xun Li,[b] Garth J. S. Cooper,[b] Margaret A. Brimble*[a,c]

[a] *The School of Chemical Sciences, University of Auckland, 23 Symonds St, Auckland 1010, New Zealand*

[b] *The School of Biological Sciences, University of Auckland, 3 Symonds St, Auckland 1010, New Zealand*

[c] *Maurice Wilkins Centre for Molecular Biodiscovery, University of Auckland, Private Bag 92019, Auckland 1010, New Zealand*

e-mail: m.brimble@auckland.ac.nz

Organic & Biomolecular Chemistry Accepted Manuscript

Abstract

Two analogues of insulin glargine containing a 1,4-disubstituted 1,2,3-triazole group in place of the CysA7-CysB7 disulfide bond were prepared using CuAAC click chemistry to efficiently join the peptide chains. The resulting insulin analogues were analysed by circular dichroism spectroscopy to assess whether this modification compromised the folding pattern of the native form. Investigations, including an *in-vivo* murine study, revealed that these analogues were not biologically active and that the structures were significantly unfolded, an outcome which suggests that maintaining a precise inter-chain distance is critical to the structure of the insulin hormone.

Introduction

Produced in the pancreas, insulin is a protein hormone with a pivotal role in the governing of blood glucose homeostasis and metabolism.[1] Although the hormone was first identified in 1921,[2] its primary structure was not reported until 1955[3] and a further decade elapsed before it was synthesised by chemical means, achieved using solution-phase[4] and by the then newly developed solid-phase techniques.[5] The 3-dimensional structure was solved using X-ray crystallography, firstly in 1969.[6] Insulin is comprised of 51 amino acids divided into two peptide chains, A and B, which are held together by inter-chain disulfide bonds between CysA7 and CysB7 and between CysA20 and CysB19. A third, intra-chain disulfide bond can also be found in the A chain between CysA6 and CysA11 (see Figure 1).

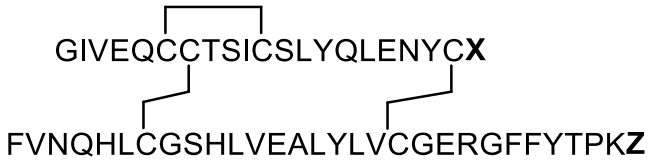


Figure 1. Insulin (**1**): **X**=Asn, **Z**=Thr; Insulin glargine (**2**): **X**=Gly, **Z**=ThrArgArg

Organic & Biomolecular Chemistry Accepted Manuscript

Much effort has been devoted to deciphering insulin's crucial role in vertebrate metabolism.[1] Since the discovery of the insulin receptor - a transmembrane tyrosine kinase - structure-function relationships have been comprehensively explored using site-directed mutagenesis and analysis of a range of naturally occurring mutations.[7] This has led to the identification of a number of key residues of insulin that are strongly involved in receptor binding and which appear highly conserved among vertebrates, namely GlyA1, GlnA5, TyrA19, AsnA21, ValB12, TyrB16, GlyB23, PheB24, PheB25, and TyrB26.[8] These together form a region that long-standing convention has labelled the classical binding surface of insulin, which also overlaps with its dimerization interface.[9] In addition, IleA2 and ValA3, which are not on the surface of the molecule, probably become exposed and interact with the receptor after displacement of the insulin B-chain carboxy terminus during the binding process.[10] Furthermore, residues LeuA13 and LeuB17, which lie outside the classical binding surface region, but are included in the hexamer-forming surface, have also been shown to make contact with the receptor on binding.[11]

Partial crystallographic information of the insulin-insulin receptor complex has provided further insights into key interactions between the hormone and Binding Site 1 of the receptor, consisting of the L1 and αCT domains.[8] These models suggest that the most critical receptor-engaging regions of insulin are the pocket formed by invariant residues ValA3, GlyB8, SerB9 and ValB12 and a hydrophobic indentation formed by residues GlyA1, IleA2, TyrA19, LeuB11, ValB12 and LeuB15. This information serves to more clearly delineate those portions of the insulin molecule that are involved in binding and furthermore appears to confirm that the disulfide bonds of insulin, while they may border the key binding domain, do not appear to be an integral part of it, except in-so-far as they help to maintain its structural integrity. It has been shown[12] that removal of the CysA7-CysB7 disulfide bond from insulin by changing these residues for Ser resulted in a dramatic loss of helical structure and concomitant impaired activity. This bridge between the A and B chains is thus clearly important, but it did raise the

Organic & Biomolecular Chemistry Accepted Manuscript

interesting question of whether the disulfide can be replaced with a non-reducible surrogate and retain the folding integrity of the native molecule. Unlike the other two disulfides of insulin, the CysA7-CysB7 disulfide bond appears to occupy a superficial and sterically uncluttered region of the hormone (Figure 2C), potentially allowing room for alternative bridging species.

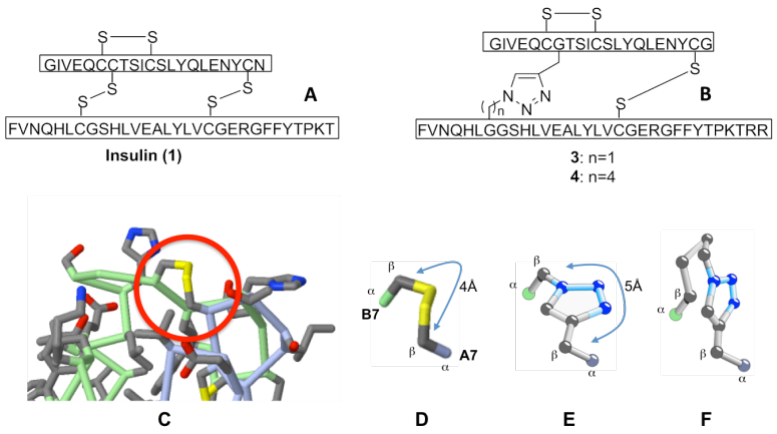


Figure 2. **A**) Insulin **1**; **B**) triazole analogues **3** and **4**; **C**) Crystal structure of insulin (PDB 313Z) indicating the exposed CysA7-CysB7 disulfide bond; **D**) CysA7-CysB7 disulfide bond showing the α and β side-chain atoms; **E** and **F**) stylised representations of potential 'click' triazole mimics of varying lengths.

A similar observation was made for the structurally related Relaxin-Like Factor,[13] and in this case the corresponding disulfide was exchanged for an amide bond to give an analogue which showed some preservation of folding integrity and receptor affinity. However, despite the numerous reported syntheses of insulins – both by global refolding of a single chain in a biomimetic manner[14] and by the stepwise combining of two chains [15, 16] – few examples exist in which the A-chain is crosslinked to the B-chain by disulfide mimics, and these syntheses are characterised by marginal yields.[13, 17] However, one alternative, the CuAAC 'click' reaction[18] which generates the 1,4-disubstituted 1,2,3-triazole ring system, does offer a relatively straightforward and high-yielding means of linking the two chains at the A7 and B7 positions (Figure 2B). Furthermore, the metabolic stability of this this ring system[19] hints at the possibility of an analogue more resistant to in-*vivo* degradation. This notion was recently demonstrated in the case of conotoxin MrIA[20] where the exchange of a disulfide for the

Organic & Biomolecular Chemistry Accepted Manuscript

triazole not only saw retention of its of norepinephrine reuptake inhibitory activity, but also an improvement in stability of the hybrid structure in plasma, presumably resulting from a reduced susceptibility to partial reduction and associated unfolding and proteolysis. In the case of Tachyplesin I[21] a disulfide was similarly exchanged for a triazole, which resulted in a near-perfect mimic of the parent structure and also exhibited comparable antimicrobial activity.

We herein report the synthesis of two insulin analogues in which the CysA7- CysB7 disulfide bond was exchanged for 1,4-disubstituted triazole linkers: analogue **3** with a single carbon atom bridging the triazole ring and the B-chain, and analogue **4** with a correspondingly longer 4-carbon tether between the B-chain and triazole. In the native form of insulin the distance between the β-carbon atoms of the CysA7 and CysB7 residues, is approximately 4 Å (Figure 2D). Thus, a triazole linker ought to span a similar distance, and also allow the same spatial orientation of the analogous Cα-Cβ bonds to minimise distortion of the tertiary structure.[22] The simplest such 1,4-disubstituted triazole mimic (Figure 2E) is planar and rigid and whilst it potentially allows the Cα-Cβ bonds to adopt a similar orientation to those of the native form, the analogous Cβ-Cβ span of about 5 Å suggests it is not an ideal substitute for the disulfide. In fact the 1,5-disubstituted 1,2,3-triazole system has been shown to provide a superior scaffold, but synthesis of this triazole form is difficult and low yielding.[22] Therefore to counter the rigidity issue it was proposed that by extending the 1-position of the triazole with a short carbon chain, the increased flexibility may allow formation of a loop and allow closer mapping onto the Cα-Cβ bonds of the native form (Figure 2F).

The intention was firstly to demonstrate that such analogues could be assembled conveniently and then to establish whether the correct folding of the native form was retained and exhibited some measure of biological activity and altered activity profile as a result of this more stable chemical motif.

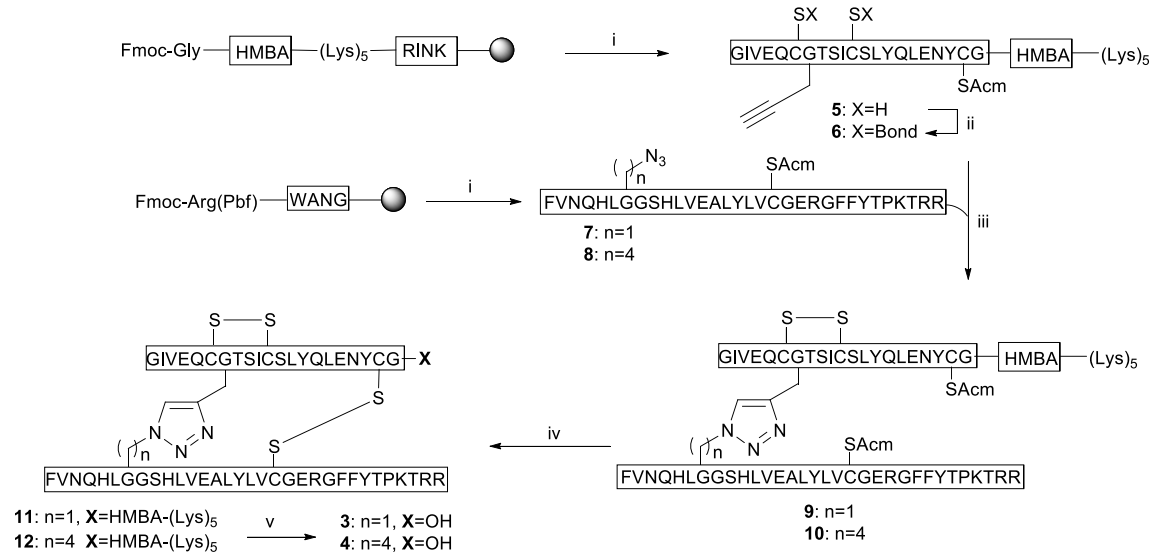
Organic & Biomolecular Chemistry Accepted Manuscript

Results and Discussion

For this study triazole analogues **3** and **4** of insulin glargine **2** rather than insulin itself were investigated for reasons of synthetic tractability;[23] their syntheses, as summarised in Scheme 1, employed a directed stepwise approach to disulfide formation and crosslinking of the A and B chains. The use of CuAAC click chemistry to link peptides quickly and efficiently has become a routine technique in the peptide chemist's arsenal and in the work reported herein it enables a very convenient synthesis of the proposed insulin analogues.

Firstly, the A-chain was synthesised using Fmoc chemistry on ChemMatrix® resin derivatised with a Gly-HMBA-pentalysine[23], incorporating propargyl glycine and Cys(Acm) at positions A7 and A20 respectively, and the conventional Cys(Trt) at positions A6 and A11. Following cleavage from the resin, the free thiols in the resultant peptide **5** at A6 and A11 were converted to an intramolecular disulfide by using a dipyridyl disulfide (DPDS)-mediated oxidation[24] affording peptide **6**. The B-chains were similarly constructed using Fmoc chemistry from Wang resin preloaded with Arg. In one instance CysB7 was substituted with L-β-azidoalanine (n=1) and Cys(Acm) incorporated at position B19 to afford peptide **7** upon cleavage from the resin. In a second example, L-ε-azidonorleucine (n=4) was introduced at position B7 to give peptide **8**.

Organic & Biomolecular Chemistry Accepted Manuscript



Scheme 1. Schematic overview of the synthesis of insulin glargine analogues **3** and **4**. Reagents and conditions: (i) (a) Iterative microwave (CEM) SPPS (5% piperazine then Fmoc-AA, HCTU, NMM, DMF); (b) 94%TFA, 1%TIPS, 2.5% water, 2.5% ethanedithiol; (ii) DPDS (1.1 equiv), 2%MeOH in water; (iii) CuI, TBTA, DMSO then sodium ascorbate, 80 ºC; (iv) $AcOH/I_2$; (v) 0.1 M NaOH.

With the requisite substrates in hand, the intermolecular click reaction to link the chains was then carried out. Firstly, one equivalent of each of **6** and **8** were combined and dissolved in degassed DMSO so that both peptides were at concentrations of 6.7 mM. Four equivalents of aqueous $CuSO_4$ (0.25 M) were then added followed by four equivalents of sodium ascorbate (0.1 M) and the colourless solution heated briefly at 80 ºC. This resulted in a fairly clean conversion to the click product **10,** as indicated by HPLC (Figure 3). Using the same method peptides **6** and **7** were combined to afford click analogue **9**.

Treatment of a dilute solution of **9** or **10** respectively, in chilled aqueous acetic acid with excess iodine effectively removed the two Acm protecting groups from CysA20 and CysB19 with concomitant formation of the inter-chain disulfide bond, giving **11** or **12** respectively, in high yield. Finally, the pentalysine tag was hydrolysed with dilute NaOH to cleanly afford the glargine analogues **3** and **4**.

Organic & Biomolecular Chemistry Accepted Manuscript

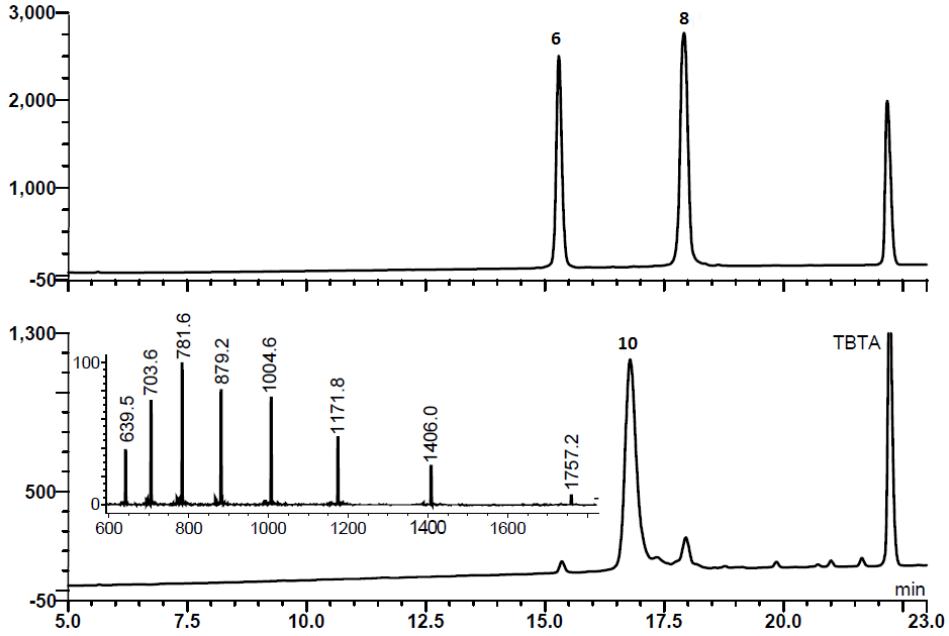


Figure 3. Representative click reaction between the A-chain **6** (Rt 15.3min) and B-chain **8** (Rt 17.9 min) to afford **10** (16.8 min) (Column: Dionex Acclaim C18, 4.6x150mm; gradient 1%-61%B over 25min @ 1mL/min; eluant A = water(0.1%TFA), B=MeCN(0.1%TFA).

The purified analogues **3** and **4** were next analysed by circular dichroism spectroscopy and the spectra compared with that of insulin glargine itself to determine the extent to which the triazole-based tethers disturbed the native structure. The experiments were performed under previously reported conditions[12] to enable facile comparison. Accordingly, solutions of the protein at 0.2 mg/mL concentration in 1 mM HCl were prepared and the spectra recorded at room temperature (Figure 4). Disappearance of the strong positive band at 195-198 nm from the spectrum indicated a marked loss of the native helical structure, with the negative region at 206 nm and shoulder at 222 nm indicative of co-existing, partially folded conformations, probably with a significant alpha helical component.[25] This pattern is identical with that observed for the non-symmetrical destabilisation of the A-chain that occurs in analogues for which both the A7 and B7 Cysteine residues have been exchanged for either Ser[26] or Glu[12] and thus lack the disulfide tether altogether.

Organic & Biomolecular Chemistry Accepted Manuscript

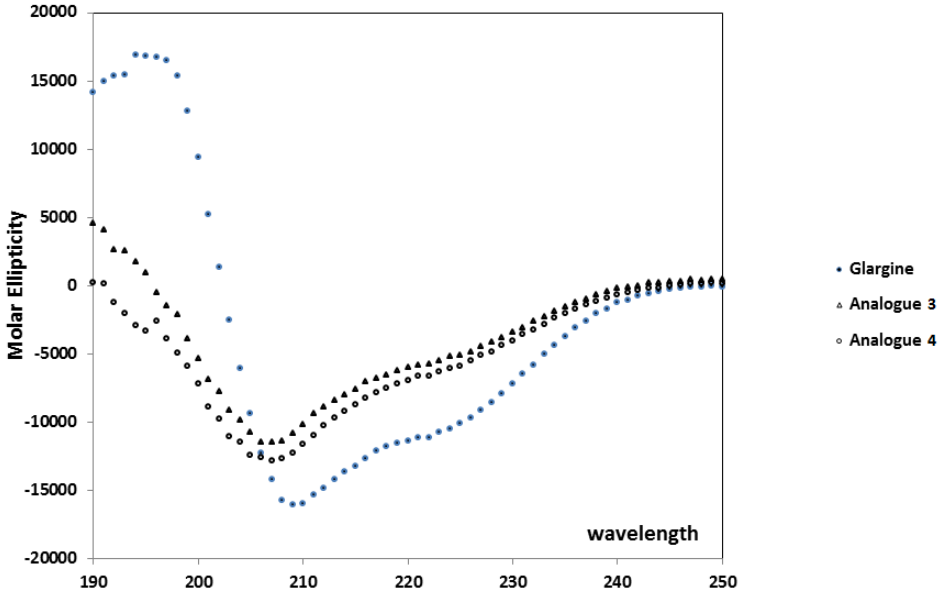


Figure 4. Circular dichroism spectra recorded for insulin glargine and analogues **3** and **4**.

These results were disappointing in that it was hoped that by tethering the A and B chains in close proximity, the propensity of the chains to unfold, as occurs when the disulfide is absent, would be reduced. However, clearly this does not appear to be the case suggesting the bridging distance between the chains must be very tightly controlled to retain overall structural integrity, leaving little margin for manipulation.

Despite this apparently compromised structure, a brief *in-vivo* experiment was performed comparing the blood-glucose lowering effect of Lantus® (insulin glargine) and the triazole mimics **3** and 4 on murine blood glucose levels (Figure 5). As predicted, based on the loss of structure, the triazole-glargine analogues **3** and **4** appeared to exert no significant activity when compared with the control group of animals.

Organic & Biomolecular Chemistry Accepted Manuscript

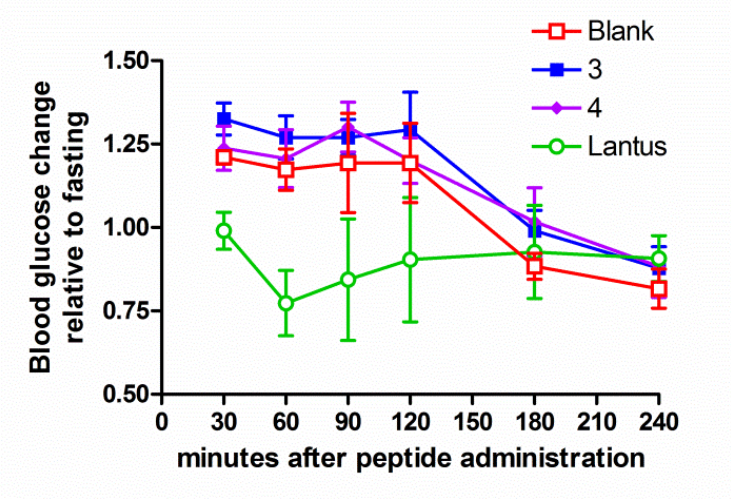


Figure 5. Blood glucose lowering efficacy of purified analogues **3** and **4** as assessed using IPITT. All values are means of blood glucose normalised to fasting (-5min) ± SEM. Repeated measures ANOVA $P<0.0001$; *post hoc* Lantus vs Blank $P<0.01$; 3 and 4 vs Blank P=ns.

Conclusion

In summary, analogues of insulin glargine containing a 1,4-disubstituted 1,2,3-triazole group in place of the CysA7-CysB7 disulfide bond were successfully prepared, illustrating the first instance of using the triazole motif as a potential surrogate for an insulin inter-chain disulfide bond. The resulting compounds were then analysed to assess whether this modification compromised the structural integrity of the native form. It was found that, in addition to being inactive *in-vivo*, these analogues were significantly unfolded, the extent of which effectively mirrored examples in which an A7-B7 tether was completely absent, an outcome that suggests maintaining a precise, rigid inter-chain distance is critical to preserving the structure of the insulin hormone

Acknowledgements

We would like to thank Dr. Jacqueline F. Aitken for her assistance in acquiring the *in-vivo* dataset.

Organic & Biomolecular Chemistry Accepted Manuscript

Supplementary data

Experimental data for the synthesis and biological activity testing of compounds **X** is available free of charge in the online version, at doi:

[1] P. Cohen, *Nat. Rev. Mol. Cell. Biol.* **2006**, *7*, 867-873.

[2] F. G. Banting, C. H. Best, *J. Lab. Clin. Med.* **1922**, *7*, 251-266.

[3] A. P. Ryle, F. Sanger, L. F. Smith, R. Kitai, *Biochem. J.* **1955**, *60*, 541-556.

[4] Y. T. Kung, Y. C. Du, W. T. Huang, C. C. Chen, L. T. Ke, *Sci. Sin.* **1965**, *14*, 1710-1716.

[5] A. Marglin, R. B. Merrifield, *J. Am. Chem. Soc.* **1966**, *88*, 5051-5052.

[6] M. J. Adams, T. L. Blundell, E. J. Dodson, G. G. Dodson, M. Vijayan, E. N. Baker, M. M. Harding, D. C. Hodgkin, B. Rimmer, S. Sheat, *Nature (Lond.)* **1969**, *224*, 491-495.

[7] Y. Kido, J. Nakae, D. Accili, *J. Clin. Endocrinol. Metab.* **2001**, *86*, 972-979.

[8] J. G. Menting, J. Whittaker, M. B. Margetts, L. J. Whittaker, G. K. Kong, B. J. Smith, C. J. Watson, L. Zakova, E. Kletvikova, J. Jiracek, S. J. Chan, D. F. Steiner, G. G. Dodson, A. M. Brzozowski, M. A. Weiss, C. W. Ward, M. C. Lawrence, *Nature* **2013**, *493*, 241-245.

[9] a) E. N. Baker, T. L. Blundell, J. F. Cutfield, S. M. Cutfield, E. J. Dodson, G. G. Dodson, D. M. Hodgkin, R. E. Hubbard, N. W. Isaacs, C. D. Reynolds, et al., *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **1988**, *319*, 369-456; b) T. L. Blundell, G. G. Dodson, D. C. Hodgkin, D. A. Mercola, *Adv. Protein Chem.* **1972**, *26*, 279-402; c) R. A. Pullen, D. G. Lindsay, S. P. Wood, I. J. Tickle, T. L. Blundell, A. Wollmer, G. Krail, D. Brandenburg, H. Zahn, J. Gliemann, S. Gammeltoft, *Nature* **1976**, *259*, 369-373.

[10] a) Q. X. Hua, S. E. Shoelson, M. Kochoyan, M. A. Weiss, *Nature* **1991**, *354*, 238-241; b) S. Ludvigsen, H. B. Olsen, N. C. Kaarsholm, *J. Mol. Biol.* **1998**, *279*, 1-7.

[11] P. De Meyts, *Diabetologia* **1994**, *37 Suppl 2*, S135-S148.

[12] Z. Y. Guo, X. Y. Jia, Y. M. Feng, *Biol. Chem.* **2004**, *385*, 1171-1175.

[13] E. E. Bullesbach, C. Schwabe, *Biochemistry* **2012**, *51*, 4198-4205.

[14] a) M. Avital-Shmilovici, K. Mandal, Z. P. Gates, N. B. Phillips, M. A. Weiss, S. B. Kent, *J. Am. Chem. Soc.* **2013**, *135*, 3173-3185; b) A. P. Tofteng, K. J. Jensen, L. Schaffer, T. Hoeg-Jensen, *Chembiochem* **2008**, *9*, 2989-2996.

[15] F. Liu, E. Y. Luo, D. B. Flora, A. R. Mezo, *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 3983-3987.

[16] F. Liu, E. Y. Luo, D. B. Flora, J. P. Mayer, *Org. Lett.* **2013**, *15*, 960-963.

[17] *US Pat.*, 2 012 022 5811 A1, 2012.

[18] M. J. Roice, I. Johannsen, M. Meldal, *QSAR Comb. Sci.* **2004**, *23*, 662.

[19] J. Hou, X. Liu, J. Shen, G. Zhao, P. G. Wang, *Expert Opin. Drug Discov.* **2012**, *7*, 489-501.

Organic & Biomolecular Chemistry Accepted Manuscript

[20] A. Gori, C. I. Wang, P. J. Harvey, K. J. Rosengren, R. F. Bhola, M. L. Gelmi, R. Longhi, M. J. Christie, R. J. Lewis, P. F. Alewood, A. Brust, *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, 1361-1364.

[21] K. Holland-Nell, M. Meldal, *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 5204-5206.

[22] M. Empting, O. Avrutina, R. Meusinger, S. Fabritz, M. Reinwarth, M. Biesalski, S. Voigt, G. Buntkowsky, H. Kolmar, *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 5207-5211.

[23] M. A. Hossain, A. Belgi, F. Lin, S. Zhang, F. Shabanpoor, L. Chan, C. Belyea, H. T. Truong, A. R. Blair, S. Andrikopoulos, G. W. Tregear, J. D. Wade, *Bioconjug. Chem.* **2009**, *20*, 1390-1396.

[24] K. Maruyama, H. Nagasawa, A. Suzuki, *Peptides* **1999**, *20*, 881-884.

[25] M. A. Fazio, V. X. Oliveira, Jr., P. Bulet, M. T. Miranda, S. Daffre, A. Miranda, *Biopolymers* **2006**, *84*, 205-218.

[26] Q. X. Hua, S. H. Nakagawa, W. Jia, S. Q. Hu, Y. C. Chu, P. G. Katsoyannis, M. A. Weiss, *Biochemistry* **2001**, *40*, 12299-12311.

Organic & Biomolecular Chemistry Accepted Manuscript